Abstract

Novel Protein Regulators of Human Ribosome Biogenesis

Mason Alexander McCool

2023

Ribosome biogenesis is the essential process that all living organisms need to produce the protein synthesizing machinery within cells. It has become apparent in recent years that ribosome biogenesis in higher eukaryotes, like humans, is increasingly complex. When this process becomes dysregulated it can lead to diseases, such as cancer when upregulated and ribosomopathies when inhibited, emphasizing the importance of learning more about the process of making ribosomes. To identify novel regulators of human ribosome biogenesis, the Baserga laboratory previously developed and performed a genome-wide RNAi screen. Here, I build upon those findings to interrogate the function of a handful of these novel human ribosome biogenesis regulators. After discovery that one of the novel regulators, CRK, likely represents an siRNA off-target effect, I helped validate the original screen results by performing a new 5-EU incorporation assay testing for more functional roles in nucleolar ribosomal RNA biogenesis. Next, I examined the molecular functions of high-confidence individual hits in the ribosome biogenesis pathway. I discovered that NOL7 is the likely ortholog of yeast Bud21 and is required for early pre-ribosomal RNA stability. I showed that large subunit biogenesis factors, RSL24D1 and the PeBoW (PES1-BOP1-WDR12) complex, interact with and regulate RNA polymerase I levels. Finally, I studied the cytidine deaminase, APOBEC3A's, function in large ribosomal subunit biogenesis. In doing so, I reveal for the first time the possibility that the pre-ribosomal RNA can be edited by this enzyme during its maturation. These results underscore the nuances of human ribosome biogenesis regulation and push forward our understanding of this essential process with implications in treatment of disease, primarily cancer.

Novel Protein Regulators of Human Ribosome Biogenesis

A Dissertation
Presented to Faculty of the Graduate School
Of
Yale University
In Candidacy for the Degree of
Doctor of Philosophy

By

Mason Alexander McCool

Dissertation Director: Susan J. Baserga, M.D., Ph.D.

May 2023

Table of Contents Page

Introduction

Results

Results

Results

Results

Results

List of Figures and Tables

Figures, Chapter 1

Figures, Chapter 2

Tables, Chapter 2

Figures, Chapter 3

Figures, Chapter 6

Tables, Chapter 6

Chapter 1

Ribosome biogenesis in health and disease

The general process of making ribosomes in humans

Ribosome biogenesis, the process by which ribosomes are made, is an essential and complex series of events across all organisms. It is required for cellular growth, proliferation, and organismal development. In eukaryotes, it occurs across cellular organelles, starting in the nucleolus, through the nucleoplasm, and ending in the cytoplasm with the formation of mature ribosomes poised to translate mRNA into protein. Ribosome biogenesis is the most energy intensive process within a cell; in yeast it accounts for ~75% of transcription [60% pre-ribosomal RNA (pre-rRNA), 15% mRNAs encoding ribosomal proteins] (Warner et al. 2001; Turowski and Tollervey 2015). It requires the coordinated effort of all 3 RNA polymerases, over 200 assembly factors, several small nucleolar ribonucleoproteins, and about 80 ribosomal proteins just to name the most established factors. Due to its tight connection with cellular growth, ribosome biogenesis is tethered to many other cellular processes and responds to a variety of external stimuli. Much of our early understanding of ribosome biogenesis came from studies completed in baker's yeast *Saccharomyces cerevisiae* (Woolford and Baserga 2013). However, the process becomes increasingly complex in higher eukaryotes, specifically in humans. Understanding the molecular underpinnings of human ribosome biogenesis is a current area of extensive research.

In humans, ribosome biogenesis begins in the cell nucleolus, which form at "nucleolar organizing regions" (NORs) present on the p-arms of the 5 acrocentric chromosomes (13,14,15, 21, and 22) (Figure 1.1). There are on average approximately 200-400 tandemly repeated copies of the ribosomal DNA (rDNA) across these 10 locations in a diploid human cell. Within each rDNA repeat there is an intergenic spacer (IGS) region (30 kb) that separates each of the 47S pre-rRNA sequences (13 kb) and harbors regulatory functions. This IGS contains regulatory elements including a promoter, long non-coding RNAs [promoter RNA (pRNA), promoter and pre-rRNA antisense RNA (PAPAS)], and other repetitive sequences (Mayer et al. 2006; Bierhoff et al. 2010; Zhao et al. 2018). Depending on its epigenetic status, the IGS regulates whether an rDNA repeat is active, which only approximately half are in a cell at a given time (McStay and Grummt 2008; Schlesinger et al. 2009). The variability among these repeats and how they dictate activity has become increasingly clear. Thanks to advancements in long read sequencing technology, this past year the repetitive p-arms of the acrocentric chromosomes, including the rDNA repeats, have been sequenced (Nurk et al. 2022). This information will help to build our understanding of active versus silenced NOR regulation in human cells [reviewed in (Hori et al. 2023)].

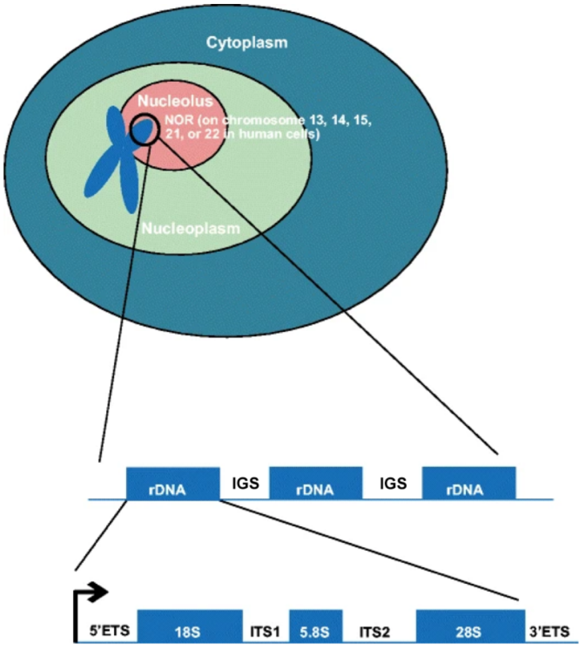


Figure 1.1: Schematic of nucleolar organizing regions (NORs) in human cells. NORs are present on the p-arms of the 5 acrocentric chromosomes that contain the ribosomal (r)DNA repeats (rDNA + intergenic spacers). The transcribed region contains 5’ and 3’ external transcribed spacer (ETS), 2 internal transcribed spacer (ITS), and 3 of the 4 mature ribosomal RNA sequences. Figure adapted from (Farley et al. 2015), not to scale.

On active rDNA repeats, the process of ribosome biogenesis commences with recruitment of RNA polymerase I (RNAP1) to the rDNA. First upstream binding factor (UBF) binds the upstream control element (UCE) then recruits selectivity factor 1 (SL-1) and preinitiation factor 3 (RRN3) carries RNAP1 to the rDNA promoter to form the preinitiation complex (PIC) [reviewed in (Pitts and Laiho 2022)]. rDNA transcription produces the 47S pre-rRNA primary transcript (Figure 1.2) and is the rate-limiting step of the entire ribosome biogenesis pathway. Therefore, it is somewhat unsurprising that it is so tightly regulated (Laferte et al. 2006).

The RNAP1 transcribed polycistronic 47S precursor contains 3 of the 4 mature rRNA sequences (18S, 5.8S, and 28S) (Figure 1.2). The mature rRNA species arise after a series of endo and exonucleolytic processing steps to remove transcribed spacer sequences within its 5' end, internal, and 3' end (5'ETS, ITS1, ITS2, and 3'ETS). There are multiple processing pathways that eventually yield the same mature rRNA products. These rRNAs are joined by the 5S rRNA which is transcribed by RNAP3 outside the nucleolus on chromosome 1 [reviewed in (Ciganda and Williams 2011)].

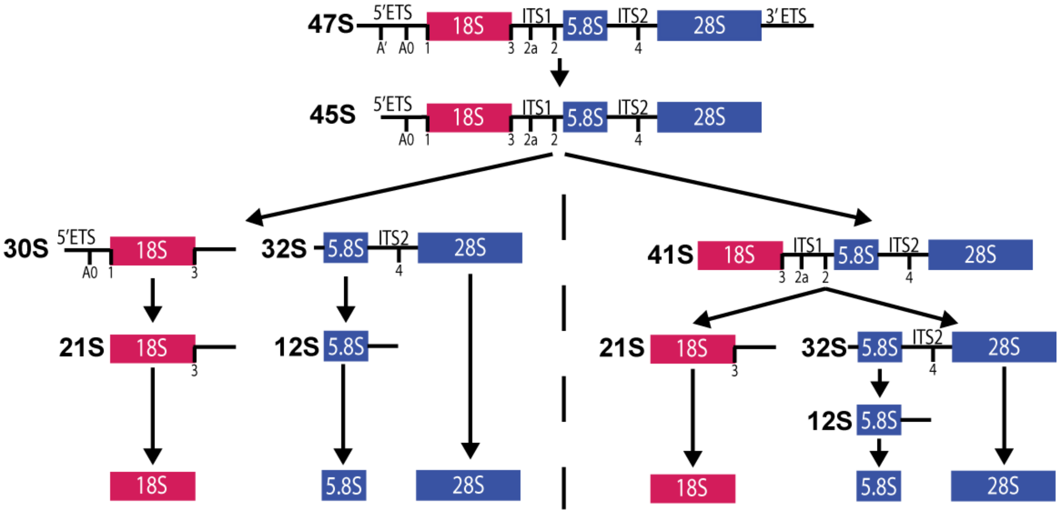


Figure 1.2: Human pre-rRNA processing pathway. The primary 47S pre-rRNA transcript undergoes a series of cleavage steps. The external transcribed spacer (ETS) and internal transcribed spacer (ITS) sequences are removed to yield the mature 18S (red, small subunit) and 5.8S and 28S rRNAs (blue, large subunit).

The rRNA is also heavily modified, primarily by 2'-O-methylation (2'-O-Me) and pseudouridylation (Ψ). 2'-O-Me is catalyzed by C/D box small nucleolar ribonucleoproteins (snoRNPs) [C/D box snoRNAs, fibrillarin (enzyme), NOP56, NOP58, and SNU13] and Ψ is catalyzed by H/ACA box snoRNPs [H/ACA box snoRNA, NHP2, NOP1, GAR1, dyskerin (enzyme)]. The snoRNAs guide, through base pairing, their associated proteins to specific locations on the pre-rRNA for modification. While most modifications are not required, they are present in functional regions of the mature ribosomes [reviewed in (Thorenoor and Slaby 2015; Sloan et al. 2017)].

On the other hand, there are unique snoRNAs that guide cleavage steps (U3, U8, U14, U17, and U22). The eukaryotic U3 snoRNA is required for the 5'ETS cleavage steps to yield the mature small subunit 18S rRNA and the vertebrate specific U8 is required for ITS2 cleavage to yield the mature large subunit 5.8S and 28S rRNAs (Peculis 1997; Langhendries et al. 2016; McFadden and Baserga 2022). Since these snoRNAs guide cleavage steps, they are essential to ribosome production.

Simultaneously during the processing steps of ribosome biogenesis, the rRNAs fold and associate with ribosomal proteins with the help of assembly factors to yield the pre-small 40S subunit (18S rRNA containing) and pre-large 60S subunit (5S, 5.8S, and 28S rRNA containing) [reviewed in (Bassler and Hurt 2019)]. These pre-ribosomal subunits are exported individually from the nucleus into the cytoplasm through the nuclear pore complex with the help of certain subunit specific factors [reviewed in (Nerurkar et al. 2015)]. Upon reaching the cytoplasm, these mature ribosomal subunits can translate mRNA into protein.

Many cellular signaling pathways converge on the nucleolus, specifically RNAP1 transcription, including the multifunctional mTOR and MAPK pathways (Figure 1.3) (Hua et al. 2022). A core component of these signaling pathways is MYC, a transcription factor for a broad number of genes. With many of its targets involved in ribosome production, it has earned the title of “master regulator of ribosome biogenesis” [reviewed in (Carroll et al. 2018; Destefanis et al. 2020; Brown et al. 2022)]. MYC is potentially even directly involved in RNAP1 transcription (Arabi et al. 2005; Grandori et al. 2005). Furthermore, the AKT/mTOR pathway can activate RNAP1 transcription and modulate translational output [reviewed in (Iadevaia et al. 2014)]. Overall, ribosome biogenesis and function are very sensitive to a variety of extracellular signals.

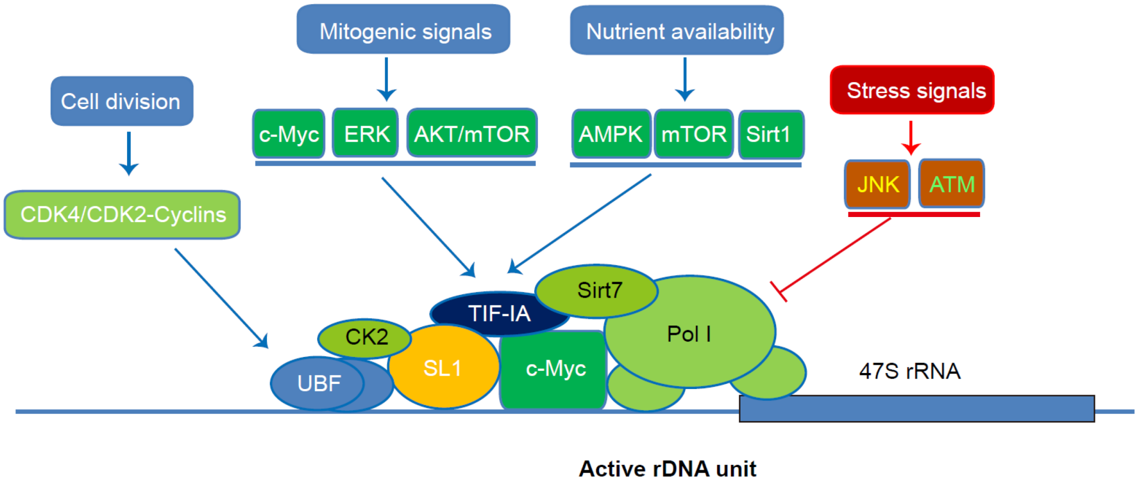


Figure 1.3: RNA polymerase I (Pol I) is regulated by a variety of signaling pathways to promote and inhibit rDNA transcription. These signaling pathways converge on various Pol I machinery and transcription factors. Taken from (Hua et al. 2022).

An additional layer of complexity added by microRNAs and long non-coding RNAs

While many known ribosome biogenesis regulatory factors are proteins or small nucleolar ribonucleoproteins, microRNAs (miRNAs) and long non-coding RNAs (lncRNAs) are emerging as a novel modulatory layer controlling ribosome production. miRNAs and lncRNAs modulate core ribosome biogenesis processes including RNAP1 pre-rRNA transcription, pre-rRNA processing, and ribosome assembly as well as nucleolar structural maintenance and global translation (Bryant et al. 2023). miRNAs and lncRNAs play diverse roles in the regulation of ribosome biogenesis, forming an additional dense layer of control over cellular growth and translational output. My colleague Carson Bryant and I published a review manuscript on these non-coding regulators in 2020 (McCool et al. 2020).

What happens when there are defects in making ribosomes?

Nucleolar Stress Response

While nucleolar function is tightly regulated by signaling pathways, the nucleolus itself acts as a hub for transmitting signals to various cellular processes upon defects or perturbations in ribosome biogenesis. Rubbi and Milner (2003) made key early discoveries on the nucleolar stress response by showing that stressors that disrupt normal nucleolar morphology (an indication of nucleolar stress) all also activate TP53 (p53). They showed disruptions in nucleolar morphology, revealed by the relocalization of nucleolar proteins (i.e. NPM1, Fibrillarin, UBF, etc.) diffusely into the nucleoplasm or into nucleolar "caps" (foci) residing on the outside of the remaining nucleolus structure [reviewed in (Boulon et al. 2010; Yang et al. 2018)]. If ribosome biogenesis factors are not localized properly within the nucleolus, then they cannot perform their correct function in making ribosomes, which is capable of being "sensed".

The most established of these nucleolar stress sensing mechanisms is through p53 as Rubbi and Milner (2003) discovered. The main connection is through the protein components of the 5S RNP that are not being incorporated into functional ribosomes, leading to p53 stabilization. Briefly, this occurs when there are free ribosomal proteins RPL5 (uL18) and RPL11 (uL5) and the 5S rRNA. Together, they bind MDM2 (p53's E3 ubiquitin ligase) to inhibit it from degrading p53 (Figure 1.4) (Farley and Baserga 2016). This p53 stabilization ultimately leads to cell cycle arrest and apoptosis [reviewed in (Lindstrom et al. 2022)]. Along with RPL5 and RPL11, other ribosome biogenesis proteins, namely HEATR3, have been implicated in p53 stabilization through this pathway (Hannan et al. 2022). Additionally, there are other p53 independent stress pathways that can be activated upon nucleolar dysfunction (i.e. p21 and p27) [reviewed in (James et al. 2014)].

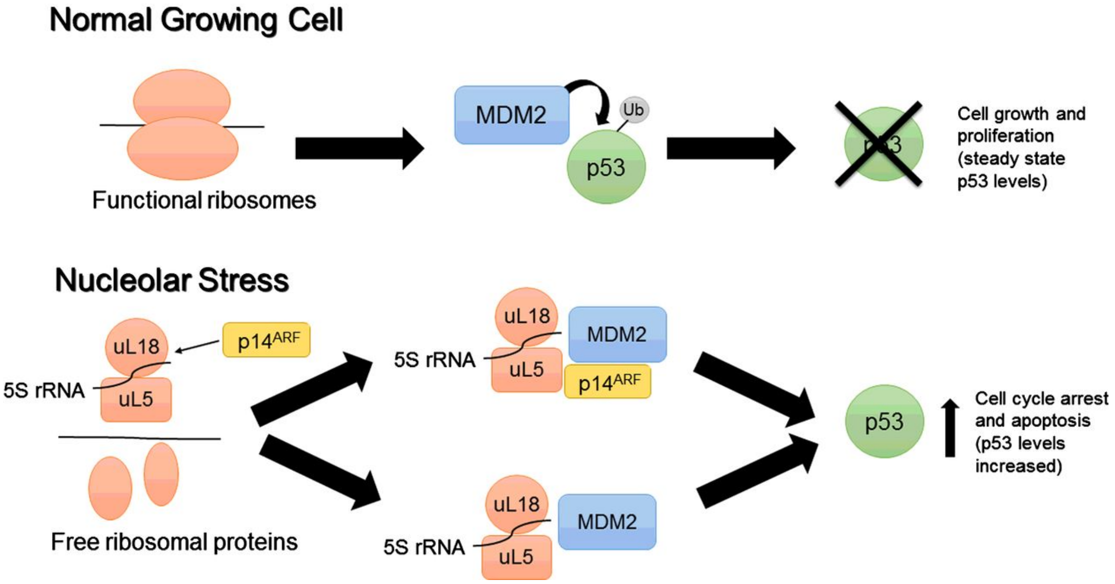


Figure 1.4: The nucleolar stress response leads to p53 (TP53) stabilization through the inhibition of MDM2 function. Under normal conditions MDM2's E3 ubiquitin ligase function targeting p53 is not inhibited and p53 levels remain lower. Dysfunctional ribosome biogenesis leads to the presence of free ribosomal proteins [RPL5 (uL18) and RPL11 (uL5)] and the 5S rRNA that form the 5S RNP complex (along with $p14^{ARF}$ sometimes) that can bind to inhibit MDM2. This leads to p53 stabilization and downstream apoptosis. Taken from (Farley and Baserga 2016).

Ribosomopathies

Nucleolar stress induced apoptosis can have detrimental impacts in ribosomopathies, the human developmental diseases that arise due to aberrant ribosome biogenesis. Ribosomopathies can be caused by mutations in factors required to make ribosomes along all the steps of the ribosome biogenesis pathway and typically manifest with tissue specific defects. Neural crest cell lineage defects have been attributed many times to the increased sensitivity these cells have towards p53 mediated apoptosis compared to other cell lineages [reviewed in (Ross and Zarbalis 2014; Farley-Barnes et al. 2019)]. Further supporting this idea is that, in some cases, the associated tissue specific phenotypes can be rescued in model systems (zebrafish or frogs) by codepletion or knockout of p53. An example of p53 loss resulting in rescue of tissue specific phenotypes includes a study completed in *Xenopus* embryos lacking Nol11 in a collaboration between the Baserga and Khokha laboratories (Griffin et al. 2015).

Nonetheless, complicating our understanding of these diseases further, ribosomopathy neural crest and p53 dependent defects are not always the underlying mechanism of disease pathogenesis.

Another hypothesis for causation of tissue specific defects are “specialized ribosomes”. This term encompasses active ribosomes containing variations in rRNA sequence, rRNA modifications, and associated ribosomal proteins, which can have differential translation capacity for various mRNAs [Reviewed in (Xue and Barna 2012; Genuth and Barna 2018; Guo 2018)]. Thus, it is hypothesized that these differences in translation are important for tissue differentiation during development. In the context of ribosomopathy mutations this has not been specifically explored. However, early results indicate RPL10A (uL1) containing ribosomes are important for early mesoderm development in mice by regulating translation of mRNAs involved in Wnt signaling (Genuth et al. 2022).

In contrast, a different hypothesis, for the tissue-specific nature of ribosomopathies is the “ribosome concentration” hypothesis. This hypothesis proposes that reduced ribosome levels in ribosomopathies will differentially effect mRNA translation and therefore expression levels in a given cell type during development [reviewed in (Mills and Green 2017)]. How ribosomopathies manifest as diverse tissue specific defects, when on a molecular level they affect all cell types, remains an area of ongoing research (Farley-Barnes et al. 2019).

What happens when too many ribosomes are being made?

Connecting the nucleolus to cell proliferation and cancer

For over 100 years, cancer pathologists have associated larger and irregular nucleoli with a worse cancer prognosis (Pianese 1896; Penzo et al. 2019), even without knowing much at all about its cellular function until the 1960’s (Birnstiel et al. 1966; Wallace and Birnstiel 1966). In 1986, Ploton et al. used more quantitative silver staining method (AgNor) to confirm these previously more subjective observations (Ploton et al. 1986). Building upon these numerous observations, a positive correlation between nucleolar size and cell proliferation rate was made by (Derenzini et al. 1998). They argued that it is not the enlarged nucleoli that are associated with cancer itself; but that it is more specifically, that enlarged nucleoli are associated with increased cell proliferation, which is not necessarily a property of all tumors.

Further complicating the connection between ribosome biogenesis and cancer is the observation that some ribosomopathies that cause bone marrow failure [including: Diamond-Blackfan anemia (DBA), Shwachman-Diamond syndrome (SDS), dyskeratosis congenita (DC), cartilage hair hypoplasia (CHH), and a type of myelodysplastic syndrome (del 5q)] predispose the afflicted individual to cancer, termed Dameshek's riddle (Dameshek 1967). It seems contradictory that diseases arising as a result of defective ribosome biogenesis, and reduced numbers of ribosomes, can lead to cancer, which has an increased number of ribosomes. However, over time, new findings argue that there are compensatory mutations that occur to initiate cancer in these individual (Lipton et al. 2021; Reilly and Shimamura 2022). For example in SDS, caused by mutations in the large subunit biogenesis factor SBDS [reviewed in (Warren 2018)], in patients who develop cancer, there have been observations of activation mutations in oncogenes or de-activation mutations in tumor suppressors. Specifically, p53 mutations have been recorded, offering a mechanism for cells to evade the nucleolar stress response, leading to cancer (Lindsley et al. 2017). These phenomena highlight the need for a happy medium level of ribosome production to maintain a healthy state.

Targeting the nucleolus to treat cancer

The strong connection between the nucleolus and cancer has made it a promising target for the development of cancer therapeutics. While most aspects of the ribosome biogenesis pathway have been probed to some extent for potential therapeutic strategies (Figure 1.5) (Catez et al. 2019), rDNA transcription inhibitors have been the most promising thus far. Among these, CX-5461 was proposed as an RNAP1 inhibitor, but more recent studies suggest this is a secondary effect from "topoisomerase poising" whereby DNA intercalating compounds covalently trap topoisomerases on DNA leading to double strand breaks and subsequent DNA damage (Bruno et al. 2020). BMH-21 has been shown to selectively lead to the degradation of RNAP1 through transcription elongation inhibition (Peltonen et al. 2014a; Peltonen et al. 2014b; Fu et al. 2017; Jacobs et al. 2022). The U3 and U8 snoRNPs, whose regulation is a subject of study here, have been proposed to be targets for cancer therapeutics to inhibit pre-rRNA processing (Langhendries et al. 2016). Progress in understanding the players and process of human ribosome biogenesis will allow for development of more precise therapeutic options in the future.

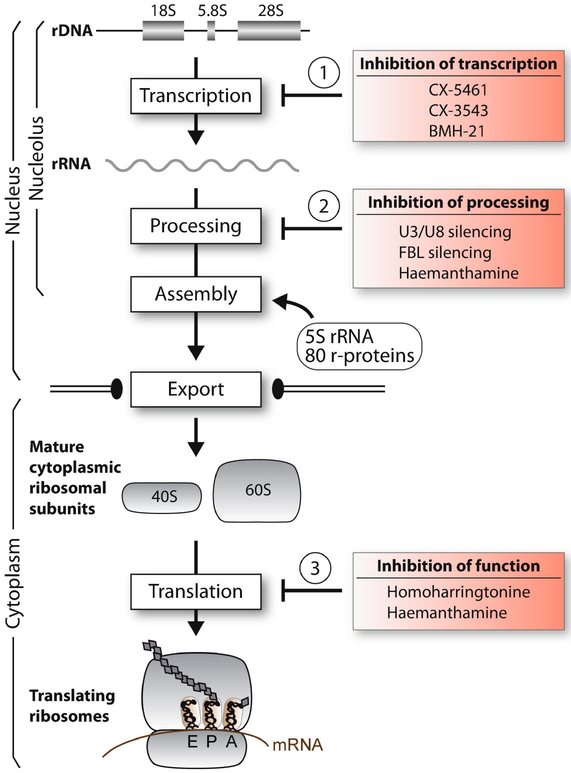


Figure 1.5: Targeting human ribosome biogenesis for cancer therapy. Various stages of the ribosome pathway are being targeted with proposed therapeutic avenues or drugs in various stages of development currently. Taken from (Catez et al. 2019).

Screening platform to identify novel human ribosome biogenesis factors

As touched on above, changes in ribosome biogenesis can lead to alterations in nucleolar morphology and number. The Baserga laboratory took advantage of these changes to identify novel ribosome biogenesis factors in human cells. The first observation from the laboratory was that depletion of NOL11, a protein essential for rDNA transcription and pre-rRNA processing, reduced the average number of nucleoli in MCF10A cells from 2-3 to 1 by immunostaining with antibodies to the nucleolar protein, fibrillarin (Freed et al. 2012). Furthermore, MCF10A cells offered a unique opportunity for identifying significant changes in nucleolar number due to their tighter distribution of nucleoli per nucleus compared to other human cell lines under normal growth conditions (Farley et al. 2015). The Baserga laboratory took this a step further and performed a genome-wide siRNA screen to identify novel ribosome biogenesis factors based on their depletion leading to reductions in nucleolar number in MCF10A cells (Figure 1.6) (Farley-Barnes et al. 2018). This successful screening campaign led to the identification of 139 high-confidence hits (>3-standard deviations from the screen mean) and about 500 additional factors (>2 standard deviations from mean) that are also likely required for ribosome biogenesis.

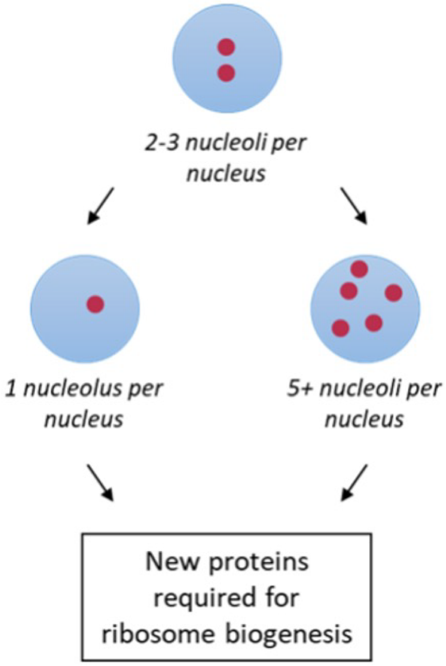


Figure 1.6: The Baserga laboratory's screening platform identifies novel ribosome biogenesis factors by observing changes in nucleolar number. Protein factors were identified whose depletion caused significant changes in nucleolar number, either a decrease to more cells containing 1 nucleolus or increase to more cell containing 5 or more nucleoli. Image created by Lisa Ogawa.

The Baserga laboratory also observed that depletion of certain factors led to an increase in the average number of nucleoli in MCF10A cells from 2-3 to 5+. One hundred thirteen (113) high confidence hits were identified whose depletion led to an increase in nucleolar number (Figure 1.6) (Ogawa et al. 2021). These hits were enriched for proteins involved in the cell cycle and are also required for RNAP1 transcription.

While it is not clear how defects in ribosome biogenesis result in changes in nucleolar number in MCF10A cells, here I will report on several novel factors required for making ribosomes in human cells revealed by these genome-wide siRNA screens.

Novel protein regulators of human ribosome biogenesis (Overview)

Chapter 2: The candidate ribosome biogenesis factor, CRK, was the unintended study of an siRNA off-target effect

Published in 2018, the Baserga laboratory discovered CRK as a novel regulator of human ribosome biogenesis through an unbiased genome-wide siRNA screen in

MCF10A breast epithelial cells (Farley-Barnes et al. 2018). Up until the discovery of its role in ribosome biogenesis, CRK had mainly been studied as a proto-oncoprotein involved in several receptor and non-receptor tyrosine kinase (RTK and NRTK) signaling pathways to control outputs including proliferation and cell migration (Sriram and Birge 2010). My results indicated a critical role for CRK in controlling the early pre-rRNA processing steps of ribosome biogenesis. More specifically, CRK's siRNA depletion caused a decrease in the levels of the essential small nucleolar RNA (snoRNA) U8 and yielded similar pre-rRNA processing defects as those observed upon depletion of the U8 snoRNA. I showed that CRK likely regulates U8 stability by rescuing U8 levels and partial rescue of associated processing defects upon co-depletion of CRK with either of the U8 negative stability regulators, the decapping enzymes DCP2 or NUDT16.

During validation experiments to show that CRK is responsible for the observed phenotypes, I was unable to recapitulate my results using an updated siRNA (siON-TARGET, Dharmacon) technology. Therefore, it is likely that my previous findings are attributable to an siRNA off-target effect of my originally used siGENOME siRNAs that target CRK. While disappointing, this sparked a new direction for my work and for more rigorous validation of the hits identified in our laboratory's screens.

Chapter 3: Efforts to validate RNAi screen hits and elucidate their roles in nucleolar rRNA biogenesis

After the discovery that my results studying CRK were most likely due to an off-target effect, our laboratory took steps to further validate hits within the original genome-wide siRNA screens for nucleolar regulators. First, I ordered a custom library containing 702 siON-TARGET siRNA pools (an improved technology compared to the siGENOME pools used for genome-wide screening) targeting the top hits from both the one nucelolus and 5+ nucleoli screens. Next, Lisa Ogawa performed the laboratory's nucleolar number assay in quadruplicate to produce a shorter list of higher confidence hits. From this, further validation of these hits was completed by Laura Abriola and Ty Brown who performed siRNA deconvolution experiments to test if multiple of the individual siRNAs within the pool could produce same change in nucleolar number phenotype. These results and availability to use this targeted library for more assays has proved valuable for robust identification and follow-up studies of these novel ribosome biogenesis factors.

Studies of the regulation of nucleolar function are critical for ascertaining insights into the basic biological underpinnings of ribosome biogenesis, and for future development of therapeutics to treat cancer and ribosomopathies. Numerous high-throughput primary assays based on morphological alterations of the nucleolus can indirectly identify hits affecting ribosome biogenesis. However, there is a need for a more direct high-throughput assay for nucleolar function to further evaluate hits. Carson Bryant developed a miniaturized, high-throughput 5-EU assay that enables specific calculation of nucleolar rRNA biogenesis inhibition, based on co-staining of the nucleolar protein fibrillarin. The assay utilizes two siRNA controls, a negative non-targeting siRNA control and a positive siRNA control targeting RNA Polymerase 1 (RNAP1; *POLR1A*), and specifically quantifies median 5-EU signal within nucleoli. I validated the 5-EU assay on 68 predominately nucleolar hits from a high-throughput primary screen, showing that 58/68 hits significantly inhibit nucleolar rRNA biogenesis. Furthermore, I utilized this assay to test for roles in nucleolar rRNA biogenesis for 702 top hits from our laboratory's previous genome-wide siRNA nucleolar number screens. This new method establishes direct quantification of nucleolar function in high-throughput, facilitating closer study of ribosome biogenesis in health and disease.

Chapter 4: Nuances of early pre-ribosomal RNA stability regulation revealed through the study of the yeast Bud21 ortholog, Human Nucleolar Protein 7 (NOL7)

The main components of the essential cellular process of eukaryotic ribosome biogenesis are highly conserved from yeast to humans. Among these, the transcription-U3 Associated Proteins (t-UTPs) are a small subunit processome subcomplex that coordinate the first two steps of ribosome biogenesis in rDNA transcription and pre-18S processing. While we have identified the human counterparts of most of the yeast Utps, the homologs of yeast Utp9 and Bud21 (Utp16) have remained elusive. In this study, I find NOL7 is the likely ortholog of Bud21. Previously described as a tumor suppressor through regulation of antiangiogenic transcripts, previous Baserga laboratory members and I now show that NOL7 is required for early pre-rRNA stability and pre-18S processing in human cells. These roles lead to decreased protein synthesis, induction of the nucleolar stress response, and defects in cell cycle progression upon NOL7 depletion. Beyond Bud21's nonessential role in yeast, we establish human NOL7 as an essential UTP that is necessary for both pre-rRNA transcription and processing.

Chapter 5: Large ribosomal subunit maturation factors, RSL24D1 and the PeBoW complex, associate with RNA Polymerase I to regulate it levels and activity

While human ribosome assembly is largely evolutionarily conserved, many of the regulatory details underlying its control and function have not yet been well-defined. The nucleolar protein RSL24D1 was originally identified as a factor important for 60S ribosomal subunit biogenesis. In addition, the PeBoW (BOP1-PES1-WDR12) complex has been well-defined as required for pre-28S rRNA processing and cell proliferation. In this study, Amber Buhagiar and I show that RSL24D1 depletion impairs both pre-ribosomal RNA (pre-rRNA) transcription and mature 28S rRNA production, leading to decreased protein synthesis and p53 stabilization in human cells. Surprisingly, each of the PeBoW complex members is also required for pre-rRNA transcription. I demonstrate that RSL24D1 and WDR12 co-immunoprecipitate with the RNA polymerase I subunit, RPA194, and regulate its steady state levels. These results uncover the dual role of RSL24D1 and the PeBoW complex in multiple steps of ribosome biogenesis and provide evidence implicating large ribosomal subunit biogenesis factors in pre-rRNA transcription control.

Chapter 6: The human cytidine deaminase, APOBEC3A, is required for large ribosomal subunit biogenesis, revealing the potential for pre-rRNA editing

Cancer initiates as a consequence of genomic mutations, and its subsequent progression relies on increased production of ribosomes to maintain high levels of protein synthesis for uncontrolled cell growth. Recently, cytidine deaminases have been uncovered as sources of mutagenesis in cancer. To form more established connections between these two cancer driving processes, I interrogated the cytidine deaminase family of proteins for potential roles in human ribosome biogenesis. I identified and validated APOBEC3A and APOBEC4 as novel ribosome biogenesis factors through our laboratory's established screening platform for the discovery of regulators of nucleolar function in MCF10A cells. I show that APOBEC3A is required for cell cycle progression and global protein synthesis. More specifically, I highlight APOBEC3A's role within the processing and maturation steps that form the large subunit ribosomal RNA (rRNA). Through an innovative nuclear RNA sequencing methodology, I identify candidate APOBEC3A C-to-U editing sites on the pre-rRNA and pre-mRNAs for the first time. My work reveals the exciting possibility that the pre-rRNA can be edited during its maturation. More broadly, I found an additional function of APOBEC3A in cancer pathology, expanding its relevance as a target for cancer therapeutics.

Chapter 2

A cautionary report: the candidate ribosome biogenesis factor, CRK, was the unintended study of an siRNA off-target effect

Introduction

CRK functions as an adapter protein to transmit extracellular signals

CRK was first discovered as an oncogene in avian sarcoma virus necessary for chicken embryo fibroblasts transformation, thus termed Chicken Tumor no. 10 (CT10) Regulator of Kinase (CRK) (Mayer et al. 1988). Its Src homology 2 (SH2) domain recognizes phosphorylated tyrosine residues present on receptor and non-receptor tyrosine kinases (RTKs and NRTKs) (Matsuda et al. 1991; Pawson 1995). CRK transmits these signals through its SH3 domain by recognizing proline rich motifs on downstream targets (Knudsen et al. 1995). CRK has two alternative splice isoforms (CRK-I and CRK-II). The longer CRK-II isoform contains an additional nonfunctional SH3 domain that serves as a site of negative regulation (Feller et al. 1994; Kobashigawa et al. 2007).

CRK is a novel ribosome biogenesis factor

The Baserga laboratory identified 139 novel human ribosome biogenesis factors in a high-throughput phenotypic genome-wide siRNA screen in MCF10A cells published in 2018 (Farley-Barnes et al. 2018). The screen found many known nucleolar regulators as well as previously unidentified and unexpected ones. Interestingly, 101 hits were not localized to the nucleolus based on three databases (Ahmad et al. 2009; Jarboui et al. 2011; Thul et al. 2017), suggesting they are involved in yet to be uncovered signaling pathways regulating ribosome biogenesis. One such hit was the adapter protein CRK. Published follow-up experiments indicated that CRK plays a role in the pre-LSU rRNA processing step of ribosome biogenesis, and thus reduces global translation upon siRNA depletion in MCF10A cells (Farley-Barnes et al. 2018). CRK's role in ribosome biogenesis was validated by siRNA deconvolution in which all four of the individual siRNAs that had been pooled for the initial screen produced the same one nucleolus phenotype (Farley-Barnes et al. 2018). These strong preliminary results gave me confidence that CRK was an interesting novel ribosome biogenesis factor worthy of further investigation.

CRK is required for craniofacial development in both mouse models and in human disease

The *Crk*-null mouse (-/-) exhibits craniofacial defects including improper nasal development and cleft palate (Park et al. 2006). Similarly, Miller-Dieker lissencephaly syndrome is caused by a microdeletion (17p13.3) and is characterized by facial dysmorphia, brain malformation, and slowed growth phenotypes (Reviewed in (Blazejewski et al. 2018)). While there are only a limited number of reported cases of this disease (less than 30), inclusion of *CRK* within this deletion region results in more severe craniofacial abnormalities and growth deficits (Nagamani et al. 2009; Bruno et al. 2010; Ostergaard et al. 2012; Barros Fontes et al. 2017). Recently, a 17p13.3 deletion cohort also exhibited neural crest lineage leukoencephalopathy phenotypes similar to those observed upon U8 small nucleolar RNA (snoRNA) mutations, which I propose CRK regulates (Jenkinson et al. 2016; Emrick et al. 2019) (see below).

While not studied in connection to ribosome biogenesis, mutations in receptor tyrosine kinase (RTK) signaling pathways, in which CRK is involved, result in various neural crest lineage specific defects [reviewed in (Dinsmore and Soriano 2018)]. Based on these examples, there are several levels of connection between CRK, ribosome biogenesis, and craniofacial development. To date, we do not know how these mutations lead to craniofacial dysmorphology. Elucidating the molecular mechanisms of CRK depletion in ribosome biogenesis will further our understanding of the underlying causes of these tissue specific diseases.

Paucity of studies on the regulation of pre-rRNA processing

There are well-established signaling pathways that regulate rDNA transcription [mTOR (Hannan et al. 2003; Mayer and Grummt 2006) and cMYC (Arabi et al. 2005;

Grandori et al. 2005)] and translation output [AKT/mTOR, reviewed in (Iadevaia et al. 2014)]. However, how signaling regulates pre-rRNA processing steps is under studied. Recently, there has been a novel pathway proposed that regulates U8 snoRNA function, an essential pre-rRNA processing factor, via stability regulation by the decapping enzyme DCP2 (Gaviraghi et al. 2018). Defining how CRK fits into this regulation will reveal how snoRNAs and pre-rRNA processing are regulated to modulate overall ribosome function in human cells. Additionally, most of these signaling pathways have previously been studied only in the context of cancer [reviewed in (Gaviraghi et al. 2019)]. Studying CRK's connection to pre-rRNA processing expands the significance to include the context of embryonic development and ribosomopathies.

Control of U8 stability and pre-LSU processing through its 5' cap

Most snoRNAs are spliced from introns and thus do not contain a canonical 7-methyl guanosine (m^7G) cap present on RNA polymerase II transcribed genes; however, a few independently transcribed snoRNAs, such as U8, contain a m^7G cap that is converted into a 2,2,7-methylguanosine (TMG) cap by TMG synthase 1 (TGS1) (Boon et al. 2015). Regulation of capping by decapping enzymes can serve as an additional point of transcript stability regulation. The first protein discovered to regulate U8 levels is the U8 specific decapping enzyme nudix hydrolase 16 (NUDT16) in humans (X29 in *Xenopus tropicalis*) (Tomasevic and Peculis 1999; Ghosh et al. 2004). More recently a second enzyme DCP2 (NUDT20), the mRNA decapping complex enzyme [reviewed in (Li and Kiledjian 2010; Arribas-Layton et al. 2013)], has been implicated in U8 decapping by translocation into the nucleolus via the tumor suppressor PNRC1 in HeLa cells (Gaviraghi et al. 2018). It remains to be further explored how U8 stability is regulated in human cells and its implications in ribosome biogenesis and disease.

Here, I present results pointing towards CRK's regulation of pre-LSU processing through U8 snoRNA stability in human tissue culture cells. More specifically, CRK depletion leads to pre-LSU processing defects, lower U8 levels, and ultimately a reduction in global protein synthesis in MCF10A and HeLa cells. I validated CRK depletion both on the mRNA and protein level and showed that U8 levels could be rescued by codepletion of its negative regulator, DCP2. These experiments were completed with the same siGENOME siRNAs (Dharmacon) and thus it was important to ensure the lack of off-target effects through validation experiments. In doing so, I could not recapitulate these initial results on CRK's role in ribosome biogenesis with an updated siRNA technology siON-TARGET (Dharmacon), while still observing CRK depletion. Thus, these results were most likely the cause of an siRNA mediated off-target effect, raising our concern for other screen hits (Chapter 3). Furthermore, this led to an unsuccessful pursuit of identifying other U8 stability regulators in the LSm protein complex.

Results

CRK depletion with Dharmacon siGENOME siRNAs reveals a likely role in U8 mediated pre-LSU rRNA processing

Validation of CRK depletion with siGENOME siRNAs

When validating an siRNA methodology, it is important to first detect reliable and significant depletion of your target of interest. I tested for efficient siRNA knockdown after 72 h treatment with siGENOME (Dharmacon) siRNAs targeting *CRK* by measuring *CRK* mRNA levels by qRT-PCR and protein levels by western blot of both CRK isoforms relative to si-nontargeting (siNT) treatment in both MCF10A and HeLa cells (Figure 2.1). My results show that both CRK-I and CRK-II levels are significantly decreased upon siGENOME siRNA treatment (Figure 2.1). Interestingly, it is unclear if these siRNAs also depleted CRK's paralog, CRK-L, as measured by qRT-PCR (Figure 2.2). However, it is less than 2-fold, a much lesser extent than *CRK* depletion on the mRNA level.

Furthermore, this indicates there is most likely no compensatory effects of CRK-L to take over for CRK's function in ribosome biogenesis.

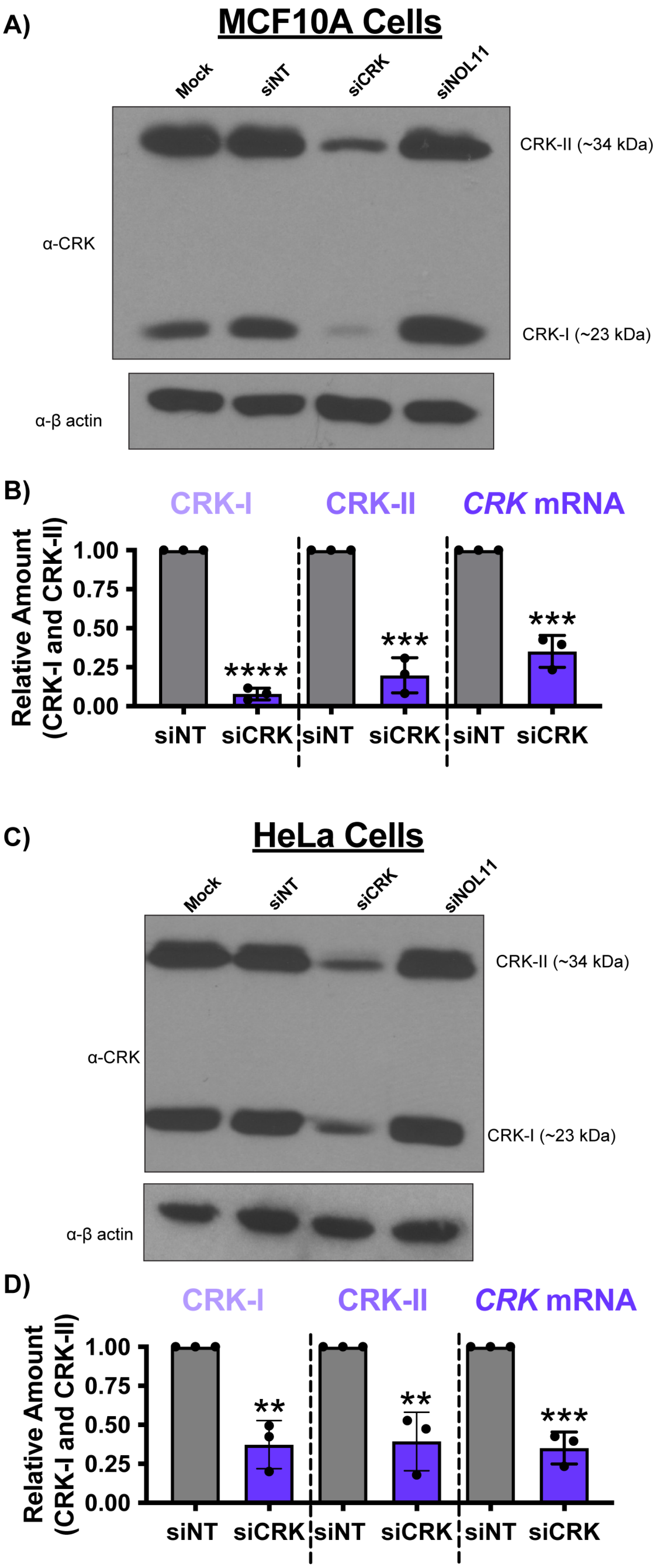

A)
MCF10A Cells
Mock
siNT
siCRK
siNOL11
α-CRK
CRK-II (~34 kDa)
CRK-I (~23 kDa)
α-β actin
B)
CRK-I
CRK-II
CRK mRNA
Relative Amount
(CRK-I and CRK-II)
1.00
0.75
0.50
0.25
0.00
****
***
***
siNT
siCRK
siNT
siCRK
siNT
siCRK
C)
HeLa Cells
Mock
siNT
siCRK
siNOL11
α-CRK
CRK-II (~34 kDa)
CRK-I (~23 kDa)
α-β actin
D)
CRK-I
CRK-II
CRK mRNA
Relative Amount
(CRK-I and CRK-II)
1.00
0.75
0.50
0.25
0.00
**
**
***
siNT
siCRK
siNT
siCRK
siNT
siCRK

Figure 2.1: Confirmation of CRK-I and CRK-II depletion in MCF10A and HeLa cells using siGENOME (Dharmacon) siRNAs. **(A, B)** Western blot and qRT-PCR confirmation of CRK-I and CRK-II knockdown in MCF10A cells. **(A)** Mock, siNT, and siNOL11 siRNAs are shown as negative controls. Representative western blot images using α-CRK and α-β-actin antibodies. **(B)** Quantification of CRK-I, CRK-II, and *CRK* mRNA levels is reported relative to siNT and normalized to the β-actin loading control for western blotting. $2^{-\Delta\Delta Ct}$ values, relative to a siNT negative control and 7SL internal control primer. Graph indicates mean ± SD, n = 3 biological replicates. Data were analyzed by Student's unpaired t-test followed where **** $p \leq 0.0001$, *** $p \leq 0.001$. **(C, D)** Western blot and qRT-PCR confirmation of CRK-I and CRK-II knockdown in HeLa cells. Same as above except in HeLa cells.

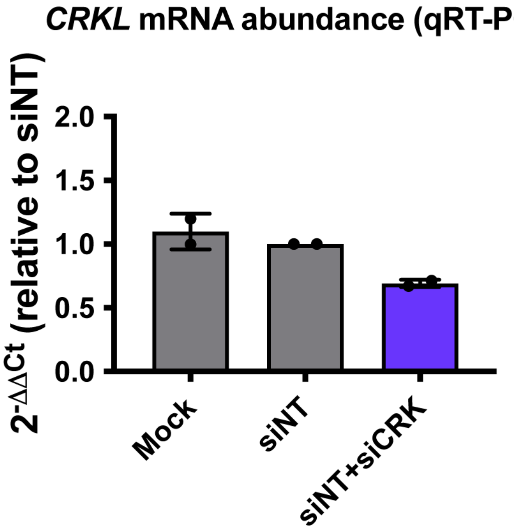


Figure 2.2: siGENOME (Dharmacon) siRNAs targeting *CRK* potentially reduces CRK's paralog, *CRKL*'s mRNA levels. Quantification of *CRKL* mRNA levels is reported as $2^{-\Delta\Delta Ct}$ values, relative to a siNT negative control and 7SL internal control primer. Mock transfection is an additional negative control. Graph indicates mean ± SD, n = 2 biological replicates.

CRK depletion with siGENOME siRNAs produces pre-rRNA processing defects

Previously, Katie Farley-Barnes observed a pre-rRNA processing defect upon CRK siRNA (siGENOME) depletion in MCF10A cells by northern blotting (Farley-Barnes et al. 2018). I performed these same northern blots in both MCF10A and HeLa cells after CRK siRNA (siGENOME) depletion using an ITS1 probe to see if I could recapitulate her results in an additional cell line. As expected, CRK depletion led to a defect in processing with a reduction in the 30S and 21S pre-rRNA precursor, indicating a defect early in the pre-rRNA processing pathway (Figure 2.3). Furthermore, there was no obvious upwards mobility shift of the 30S pre-rRNA precursor (30S+1) which would be indicative of defective 5’ETS processing (Prieto and McStay 2007; Mullineux and Lafontaine 2012). Additionally, there was a modest increase in a 36S pre-rRNA intermediate relative to its upstream precursors, which can accumulate upon defective ITS2 processing, which is mediated by the U8 snoRNA (Peculis and Steitz 1993; Langhendries et al. 2016). Due to a significant defect in pre-rRNA processing after CRK depletion, but no obvious point in the pre-rRNA processing pathway that is inhibited, I decided to test if there were any differences in the amount of mature 18S or 28S rRNAs produced.

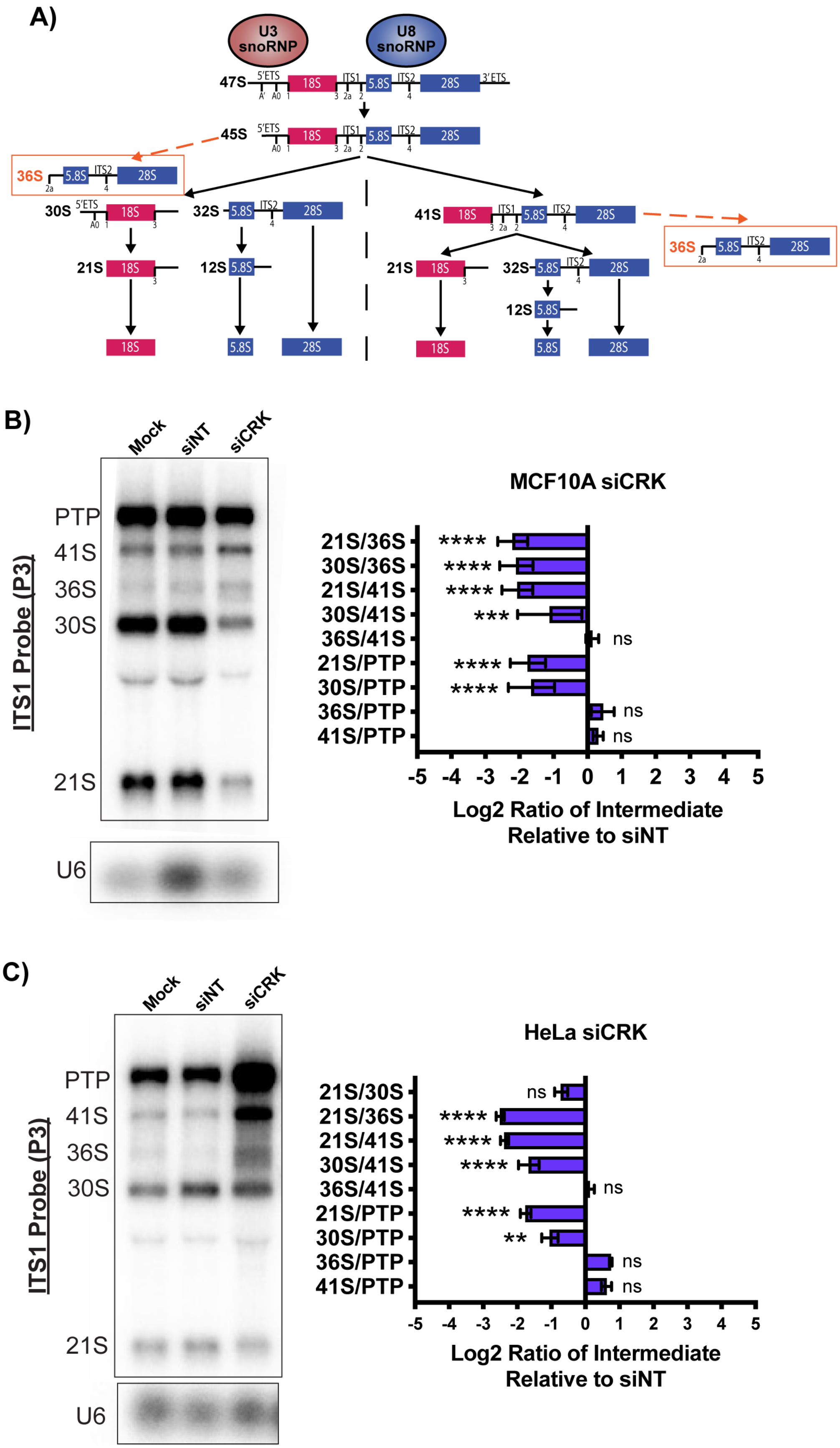
A)
U3 snoRNP
U8 snoRNP
47S
5'ETS
18S
ITS1
5.8S
ITS2
28S
3'ETS
A'
A0
1
3
2a
2
4
45S
36S
30S
32S
41S
21S
12S
B)
Mock
siNT
siCRK
PTP
41S
36S
30S
21S
ITS1 Probe (P3)
U6
MCF10A siCRK
21S/36S
30S/36S
21S/41S
30S/41S
36S/41S
21S/PTP
30S/PTP
36S/PTP
41S/PTP
****
****
****
***
ns
****
****
ns
ns
-5 -4 -3 -2 -1 0 1 2 3 4 5
Log2 Ratio of Intermediate Relative to siNT
C)
Mock
siNT
siCRK
PTP
41S
36S
30S
21S
ITS1 Probe (P3)
U6
HeLa siCRK
21S/30S
21S/36S
21S/41S
30S/41S
36S/41S
21S/PTP
30S/PTP
36S/PTP
41S/PTP
ns
****
****
****
ns
****
**
ns
ns
-5 -4 -3 -2 -1 0 1 2 3 4 5
Log2 Ratio of Intermediate Relative to siNT

Figure 2.3: CRK is required for pre-rRNA (LSU) processing in MCF10A and HeLa cells. **(A)** pre-rRNA processing in human cells. The 47S pre-rRNA precursor [47S+45S = primary transcript plus (PTP)] undergoes a series of modification and processing steps to yield the mature 18S, 5.8S, and 28S rRNAs. The U3 snoRNP mediates 5'ETS processing. The steps that produce the 18S rRNA (small subunit) are indicated in red, and the steps that produce the 5.8S and 28S rRNAs (large subunit) are indicated in blue. Upon inhibition of pre-rRNA cleavage in the ITS2, mediated by the U8 snoRNP, accumulation of an aberrant 36S (cut in ITS1 at site 2a instead of 2) occurs (orange). The U8 snoRNA is essential for processing of the rRNAs that make the large ribosomal subunit. **(B-C)** CRK is required for pre-rRNA processing in MCF10A and HeLa cells. **(B)** (Left) Representative northern blot using an ITS1 probe to measure steady-state levels of pre-rRNAs that lead to the formation of the small ribosomal subunit (SSU) in MCF10A cells. A U6 small-nuclear RNA probe is used as a loading control. Mock transfected and siNT are negative controls. (Right) Quantification of defective pre-rRNA processing results. Ratio analysis of multiple precursors (RAMP) (Wang et al. 2014) quantification of northern blotting using an ITS1 probe. Log_2 fold change of pre-rRNA ratios were quantified relative to siNT negative control. Three biological replicates plotted mean ± SD. Data were analyzed by two-way ANOVA, **** $p \leq 0.0001$, p *** ≤ 0.001. **(C)** Same as in (B) except in HeLa cells.

I measured mature 18S and 28S rRNA levels after siCRK (siGENOME) depletion in MCF10A and HeLa cells by Agilent BioAnalyzer analysis. There was a trending decrease in the 28S/18S rRNA ratio in both cell lines tested (Figure 2.4). The reduced 28S relative levels, along with trending increases in 36S relative levels, indicates CRK might play a preferential role in pre-LSU rRNA processing. However, more replicates of this experiment would be valuable to determine if these small differences are significant. Based on these results, I decided more focused assays were required to elucidate CRK's role in pre-rRNA processing.

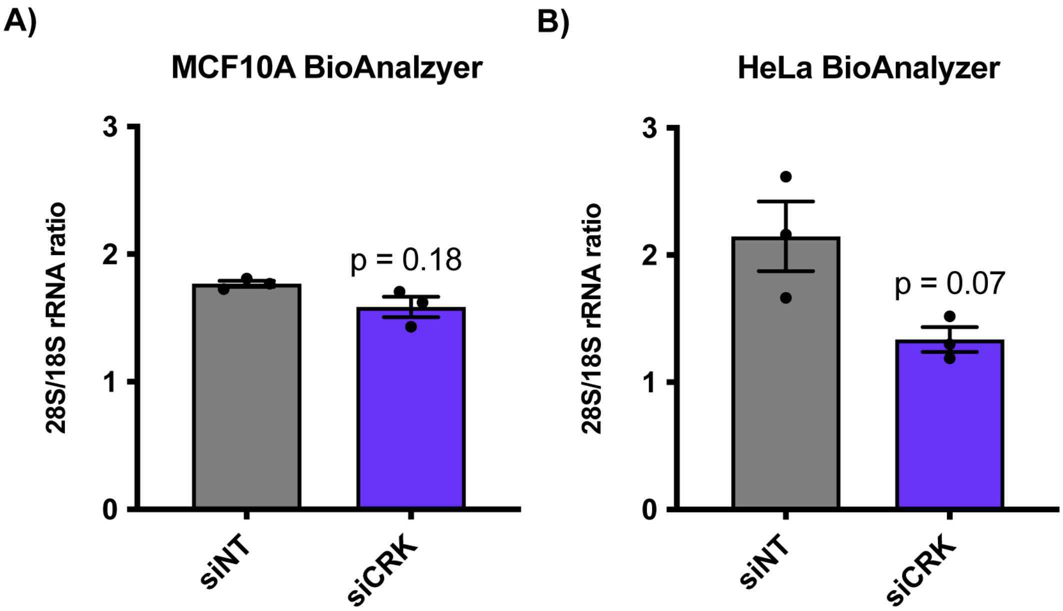


Figure 2.4: CRK is potentially required to maintain normal mature 28S rRNA levels in MCF10A and HeLa cells. **(A)** CRK is likely required for production of a normal 28S/18S mature rRNA ratio in MCF10A cells. Agilent BioAnalyzer analysis of ratio of mature 28S to 18S rRNAs from MCF10A cells depleted of CRK. 3 biological replicates, plotted mean ± SD. Data were analyzed by Student's t-test relative to siNT negative control, p-values indicated. **(B)** Same as in (A) except in HeLa cells.

CRK depletion with siGENOME siRNAs reduces U8 snoRNA levels

U3 (SSU) and U8 (LSU) are two essential snoRNAs required for cleavage steps in the pre-rRNA processing pathway in vertebrates (Peculis 1997; Langhendries et al. 2016). First, I aimed to recapitulate previously observed pre-rRNA processing results upon U3 and U8 depletion in MCF7 cells by northern blotting (Langhendries et al. 2016) in MCF10A and HeLa cells to compare to the CRK depletion pre-rRNA processing defect. I depleted both U3 and U8 for 48 h using antisense oligos (ASOs) (Langhendries et al. 2016) and performed northern blotting using an ITS1 probe as done previously in MCF10A cells (Farley-Barnes et al. 2018; Farley-Barnes et al. 2020). As expected, the ASOs depleted U3 and U8 respectively (Figure 2.5A, B) compared to a scrambled (SCR) ASO, however it seems to be their levels are dependent on each other to some extent. Also, they produced their expected pre-rRNA processing defects (Figure 2.5C, D), where U3 depletion led to presence of an aberrant 30S+1 precursor and U8 depletion led to an aberrant 36S precursor in both MCF10A and HeLa cells.

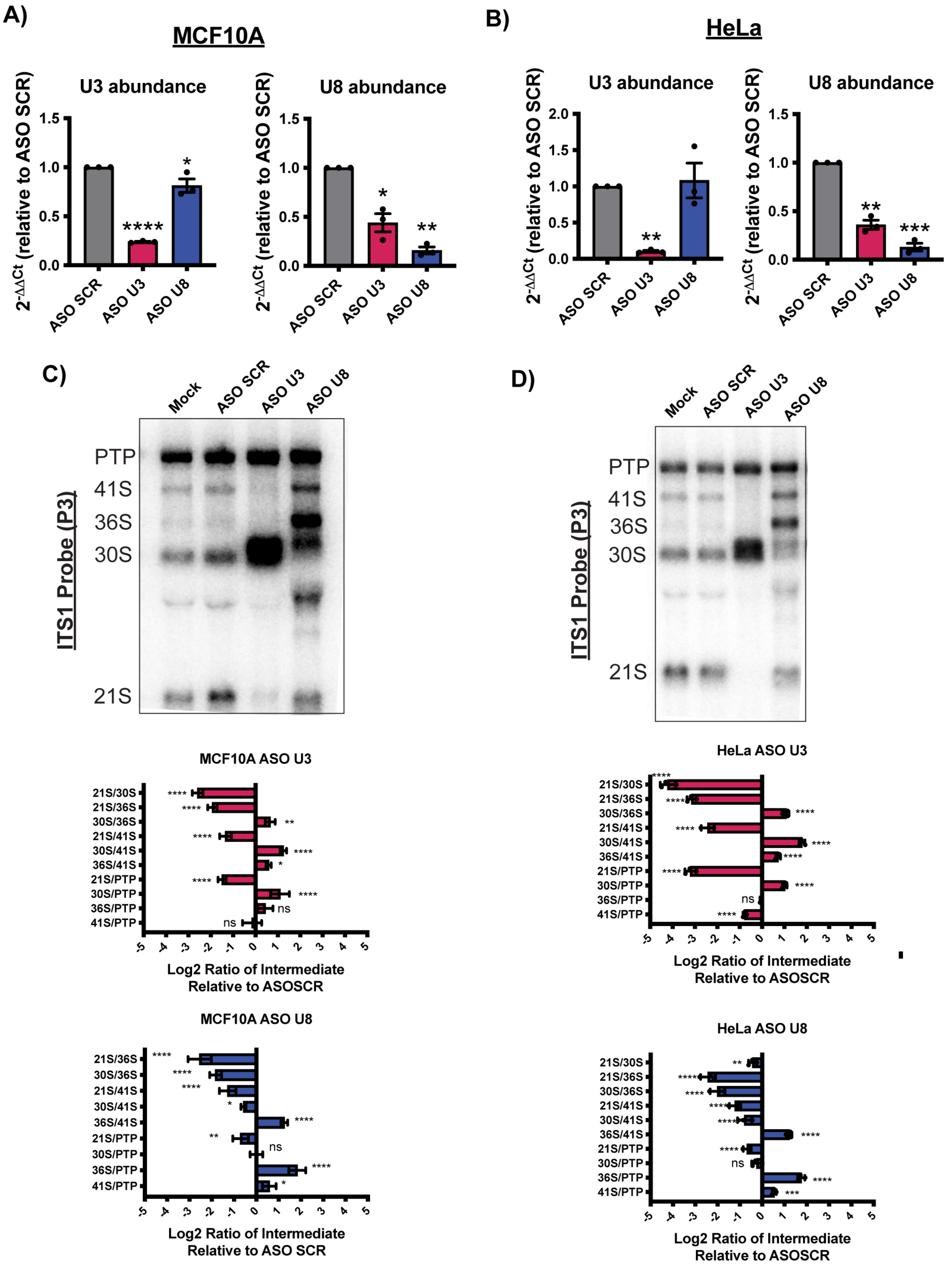


Figure 2.5: Antisense oligos (ASOs) against U3 and U8 reduce their levels and produce their expected pre-rRNA processing defects in MCF10A and HeLa cells. **(A)** ASO U3 and ASO U8 reduce their target transcript levels in MCF10A cells after 48 h. qRT-PCR measuring U3 (red) and U8 (blue) snoRNA transcript levels in MCF10A cells. $2^{-\Delta\Delta Ct}$ were measured relative to 7SL internal control and siNT negative control sample. 3 technical replicates of 3 biological replicates, plotted mean ± SD. Data were analyzed by Student's t-test, **** $p \leq 0.0001$, *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$. **(B)** Same as in A except in HeLa cells. **(C)** (Top) Representative northern blot using an ITS1 probe to measure steady-state levels of pre-rRNAs that lead to the formation of the small ribosomal subunit (SSU) in MCF10A cells. Mock transfected and ASO SCR are negative controls. (Bottom) Quantification of U3 or U8 defective pre-rRNA processing results. Ratio analysis of multiple precursors (RAMP) (Wang et al. 2014) quantification of northern blotting using an ITS1 probe. Log_2 fold change of pre-rRNA ratios were quantified relative to ASO SCR negative control. Three biological replicates plotted mean ± SD. Data were analyzed by two-way ANOVA, **** $p \leq 0.0001$, *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$. **(D)** Same as in (C) except in HeLa cells.

I tested if CRK was required to maintain their function by examining U3 and U8 levels after CRK depletion in MCF10A and HeLa cells to gain insight into the pre-rRNA processing defects I had observed. I measured U3 and U8 levels by qRT-PCR following CRK siRNA (siGENOME) depletion. Aligning with my previous trending results in decreased 28S mature rRNA levels and increased 36S pre-rRNA precursor levels, I observed a significant decrease in U8 levels (~2-fold), but no significant changes in U3 levels in either MCF10A or HeLa cells (Figure 2.6). Since CRK is required to maintain U8 levels specifically and not U3, I explored by which mechanism this could be, by regulating transcription or stability.

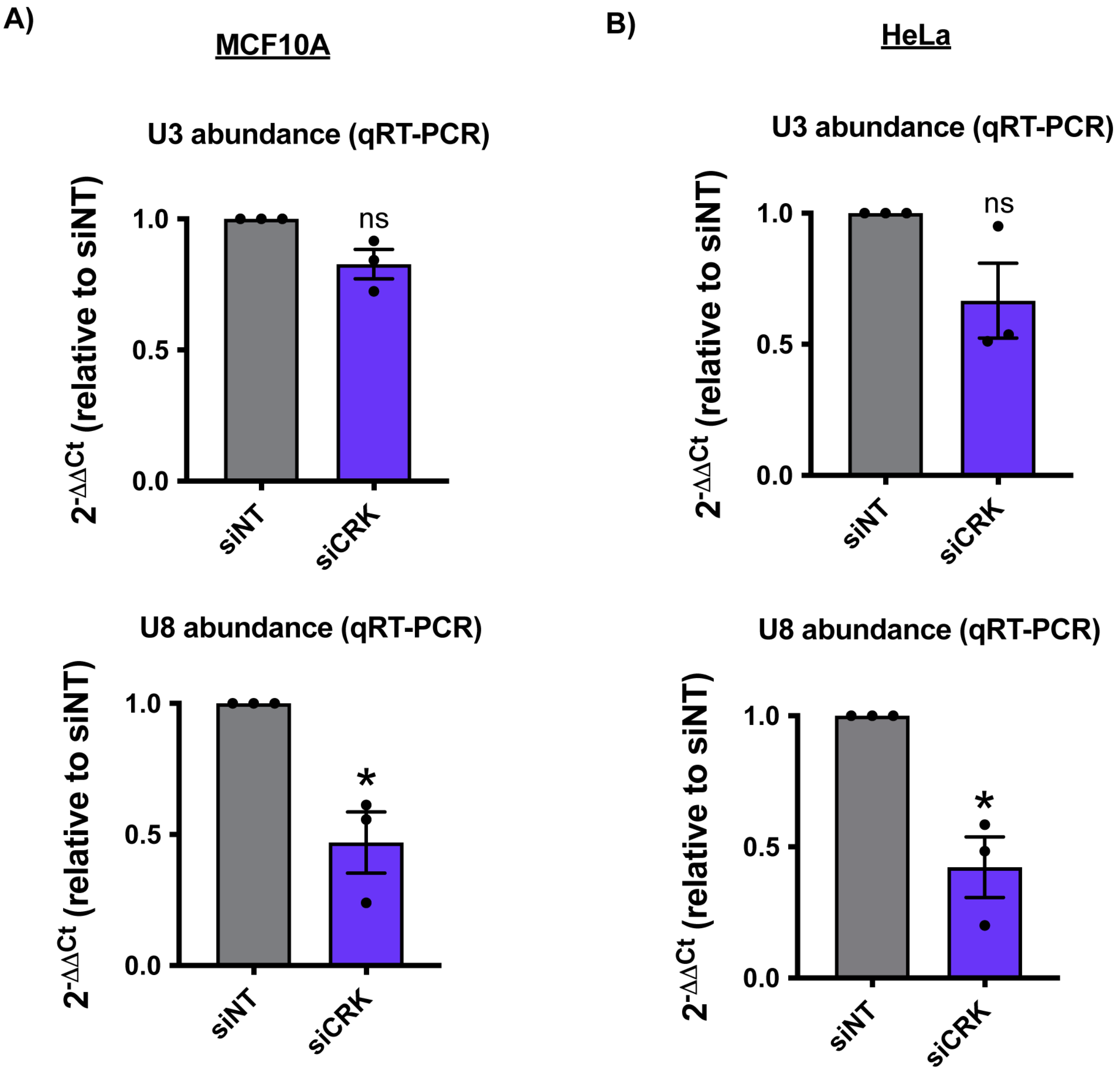


Figure 2.6: CRK siRNA (siGENOME, Dharmacon) depletion in MCF10A and HeLa cells results in reduced U8 snoRNA steady-state levels. **(A)** qRT-PCR measuring U3 (top) and U8 (bottom) snoRNA transcript levels in MCF10A cells. $2^{-\Delta\Delta Ct}$ were measured relative to 7SL internal control and siNT negative control sample. 3 technical replicates of 3 biological replicates, plotted mean ± SD. Data were analyzed by Student's t-test, * $p \leq 0.05$. **(B)** Same as in (A) except in HeLa cells.

Investigating CRK's role in U8 stability and pre-rRNA processing through U8 decapping enzymes DCP2 and NUDT16

I hypothesized if CRK positively regulates U8 stability then co-depletion with a negative regulator would rescue U8 levels. Therefore, I chose to co-deplete CRK with either NUDT16 (a U8 specific decapping enzyme) (Tomasevic and Peculis 1999; Ghosh et al. 2004) or DCP2 (a more promiscuous U8 decapping enzyme) (Gaviraghi et al. 2018) and measure both U8 levels and pre-rRNA processing after these co-depletions (Figure 2.7). Additionally, I would expect that individually depleting DCP2 or NUDT16 alone would increase U8 levels compared to normal levels.

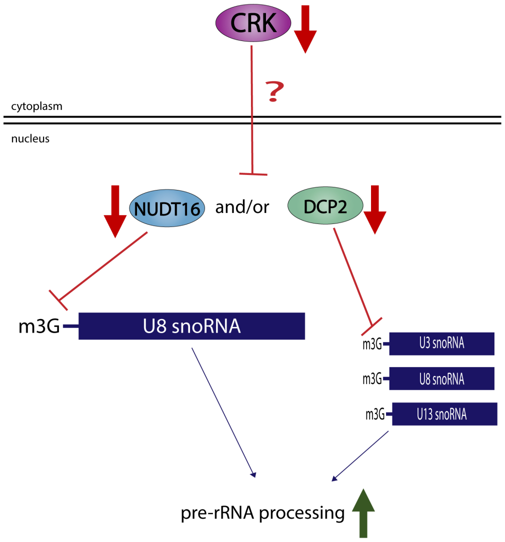


Figure 2.7: Hypothesized pathway by which CRK regulates U8 stability through signaling to control NUDT16 and/or DCP2 decapping function of tri-methyl guanosine (TMG) capped small nucleolar (sno)RNAs. If CRK is codepleted with either NUDT16 or DCP2 then it will be able to restore U8 stability and function.

First, I confirmed robust siRNA (siGENOME) mediated depletion of both *NUDT16* and *DCP2* at the mRNA level by qRT-PCR in HeLa cells (Figure 2.8A). Interestingly, NUDT16 siRNA depletion increases *DCP2* mRNA levels, suggesting a compensatory mechanism by DCP2 when NUDT16 is unable to function. Next, I tested U3 and U8 snoRNA levels in MCF10A cells after CRK, NUDT16, or DCP2 individual and co-depletions by siGENOME siRNAs. As before, CRK siRNA depletion reduced U8 snoRNA levels but did not change U3 snoRNA levels. Furthermore, none of the depletion conditions changed U3 snoRNA levels significantly, but there were trending increases upon any depletion of the decapping enzymes (Figure 2.8B). Unexpectedly, there was an increase in U8 snoRNA levels only after DCP2 siRNA depletion and not NUDT16, however this could be explained by my previous results (Figure 2.8A), where

NUDT16 depletion increases DCP2 levels which could compensate in U8 decapping and subsequent maintenance of U8 snoRNA steady state levels (Figure 2.8B). Lastly, CRK likely regulates U8 snoRNA stability since U8 snoRNA levels can be resuced by co-depletion with either of its decapping enzymes DCP2 or NUDT16. Next, I confirmed that there was also rescue of a U8 snoRNA mediated processing defect that was observed upon CRK depletion.

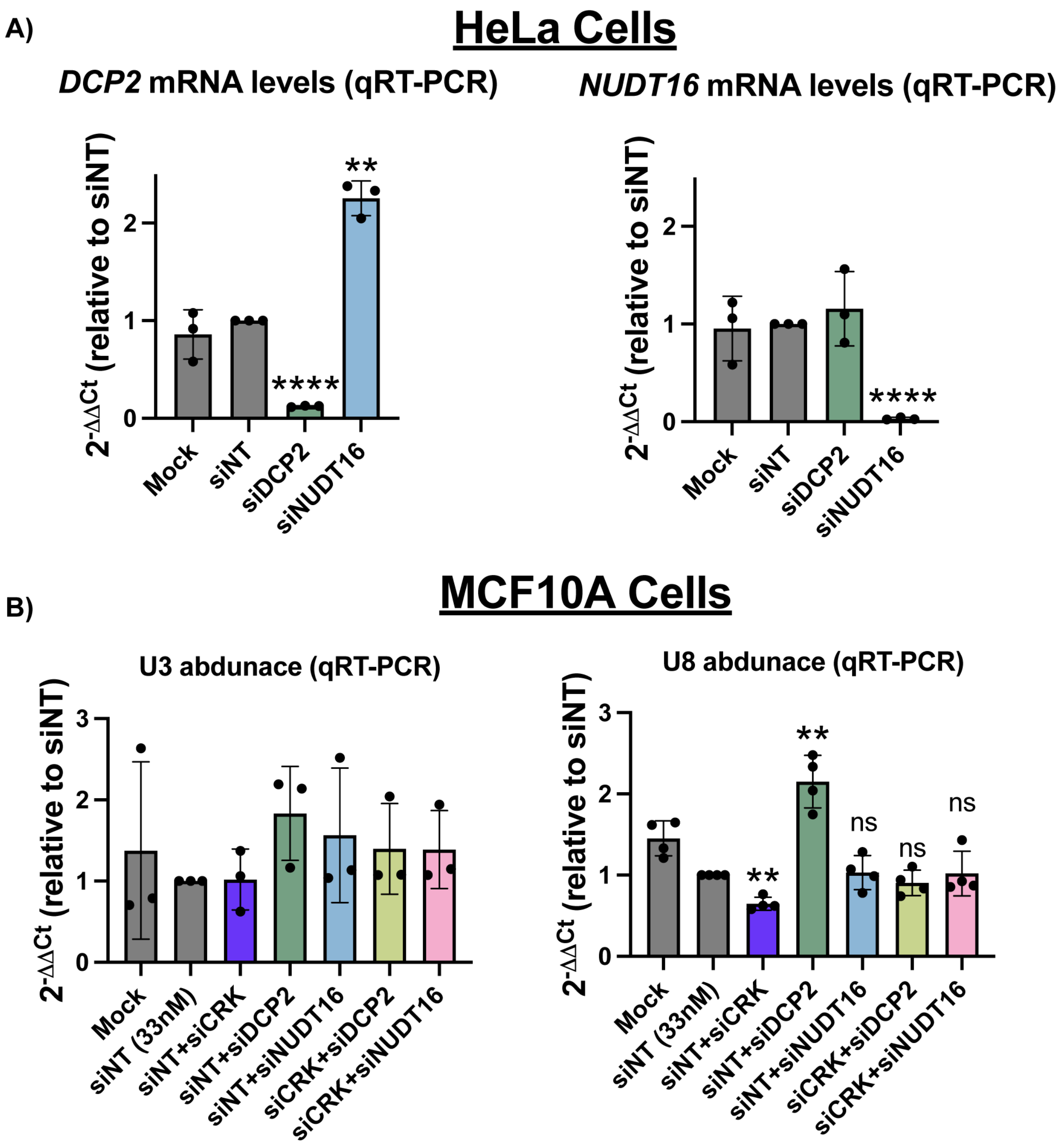


Figure 2.8: DCP2 and NUDT16 siRNA (siGENOME) depletion in HeLa cells results in reduces their target transcript steady-state levels, while NUDT16 siRNA depletion increases DCP2 levels. Co-depletion of CRK with either DCP2 or NUDT16 rescues U8 snoRNA steady-state levels, while U3 snoRNA levels remain unchanged, DCP2 depletion alone can increase U8 snoRNA levels. **(A)** qRT-PCR measuring *DCP2* (left) and *NUDT16* (right) mRNA transcript levels in HeLa cells. $2^{-\Delta\Delta Ct}$ were measured relative to 7SL internal control and siNT negative control sample. 3 technical replicates of 3 biological replicates, plotted mean ± SD. Data were analyzed by Student's t-test, **** $p \leq 0.0001$, *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$. **(B)** Same as in (A) except measuring U3 (left) and U8 (right) snoRNA levels in different depletion conditions in MCF10A cells. siRNA depletion was conducted with a total siRNA concentration of 33 nM (i.e. 16.5 nM for each individual siRNA). 3 technical replicates of 3 biological replicates was completed for U3, while 3 technical replicate of 4 biological replicates was completed for U8.

I performed pre-rRNA processing northern blots in MCF10A cells upon these same co-depletion treatments to see if in addition to U8 snoRNA levels, the CRK processing defect could be rescued as well. Co-depletion of CRK with either DCP2 or NUDT16 was able to partially rescue the CRK depletion pre-rRNA processing defect (Figure 2.9). More specifically, the 36S precursor levels were decreased and the 30S and 21S intermediates were increased closer to their normal levels in siNT treated cells. This corroborates my results showing U8 snoRNA levels being rescued and that CRK plays a role in maintaining U8 snoRNA stability by an undiscovered mechanism. However, given these striking results, I expect that CRK is able to signal to either DCP2 and/or NUDT16 to ultimately control pre-rRNA processing in human cells. Only partial pre-rRNA processing rescue through co-depletion is likely due to our incomplete understanding of the redundant roles of DCP2 and NUDT16 in U8 TMG cap regulation. Further supporting CRK's hypothesized connection to U8 stability, CRK was shown in a high-throughput interactome study to interact with two mRNA decapping complex members, enhancer of mRNA decapping 4 and dead-box helicase 6 (EDC4 and DDX6) in K652 cells (Brehme et al. 2009). These interactions offer a possible mode of action for CRK's regulation of U8 decapping through DCP2.

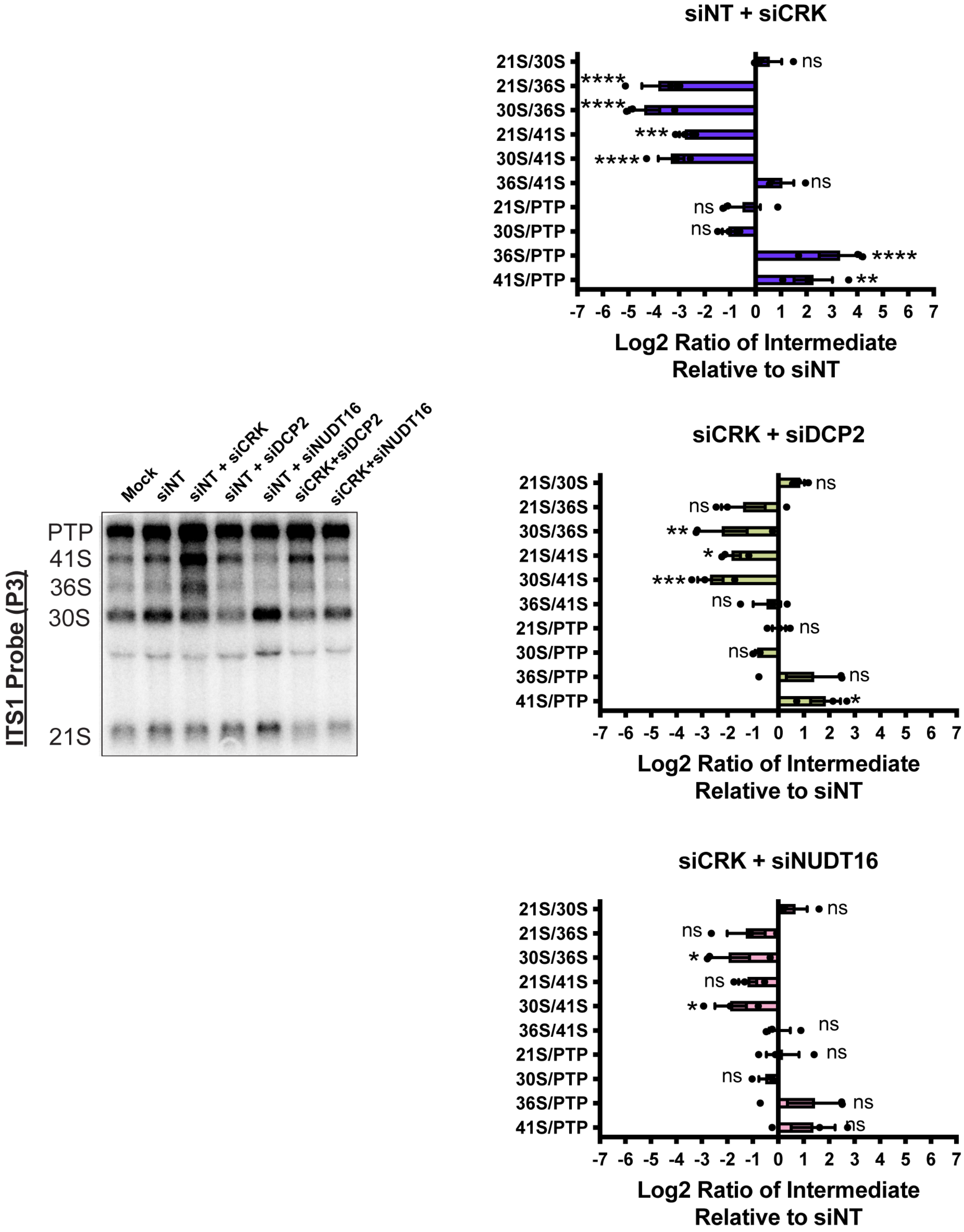


Figure 2.9: CRK mediated pre-rRNA processing defects can be partially rescued by co-depletion with either DCP2 or NUDT16 in MCF10A. (Left) Representative northern blot using an ITS1 probe to measure steady-state levels of pre-rRNAs that lead to the formation of the small ribosomal subunit (SSU) in MCF10A cells. Mock transfected and

siNT are negative controls. (Right) Quantification of pre-rRNA processing results in CRK (top), CRK+DCP2 (middle), CRK+NUDT16 (bottom) siRNA depletion conditions. Ratio analysis of multiple precursors (RAMP) (Wang et al. 2014) quantification of northern blotting using an ITS1 probe. Log_2 fold change of pre-rRNA ratios were quantified relative to siNT negative control. Three biological replicates plotted mean ± SD. Data were analyzed by two-way ANOVA, **** $p \leq 0.0001$, *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$.

CRK depletion with Dharmacon siON-TARGET siRNAs reveals the previous siGENOME results were likely due to an siRNA off-target effect

To confirm these exciting results revealing a potential mechanism for CRK's requirement in pre-rRNA processing, I aimed to recapitulate them using an updated siRNA technology from Dharmacon (siON-TARGET) which are designed to reduce siRNA off-target effects (Jackson et al. 2006) and have different siRNA target sequences compared to the siGENOME siRNAs I used previously (Table 2.1). Furthermore, this alternative method of depletion would be more likely to have different off-target effects in addition to fewer. First, I confirmed CRK depletion by siON-TARGET siRNAs on the mRNA level by qRT-PCR in MCF10A cells (Figure 2.10A). The siON-TARGET siRNAs led to an even greater reduction in CRK mRNA levels than siGENOME siRNAs (Figure 2.1 and 2.10A). Next, I tested for changes in U8 snoRNA levels after CRK siRNA (siON-TARGET) depletion in MCF10A cells. Unexpectedly, I did not observe any changes in U8 snoRNA levels as I had previously after siGENOME depletion (Figure 2.10A, 2.6). I also did not observe any changes in mature 28S/18S rRNA ratio as quantified by BioAnalyzer analysis after CRK siRNA (siON-TARGET) depletion in MCF10A cells (Figure 2.10B). Next, to assess whether CRK still regulates pre-rRNA processing in the same way, yet not through U8 stability, I performed pre-rRNA processing northern blots as previously. Agreeing with no changes in U8 snoRNA levels, after CRK siRNA (si-ONTARGET) depletion I did not observe any noticeable defect in pre-rRNA processing in MCF10A cells using an ITS1 probe (Figure 2.10C). Finally, I wanted to observe if CRK still played a role in making ribosomes by performing a puromycin incorporation assay to measure global protein synthesis (Schmidt et al. 2009). Aligning with all the results using siON-TARGET siRNAs to deplete CRK thus far, I did not observe a significant decrease in global protein synthesis in MCF10A cells (Figure 2.10D). Since I was able to detect significant CRK depletion but none of the previously identified defects in pre-rRNA processing (Figure 2.3), U8 snoRNA steady-state levels (Figure 2.6), or global protein synthesis (Farley-Barnes et al. 2018), it is likely that the results I observed were due to an siRNA off-target effect from the siGENOME siRNAs.

Table 2.1: CRK targeting siRNA sequences.

Members of siRNA pool	siGENOME (Dharmacon) sequences (5' -> 3')	siON-TARGET (Dharmacon) sequences (5' -> 3')
#1	GGAGACAUCUUGAGAAUCC	GGACAGCGAAGGCAAGAGA
#2	UCCCUUACGUCGAGAAGUA	GAAUAGGAGAUCAAGAGUU
#3	GGACAGCGAAGGCAAGAGA	GGUGAGCUGGUAAAGGUUA
#4	GGGACUAUGUGCUCAGCGU	GGACAAGCCUGAAGAGCAG

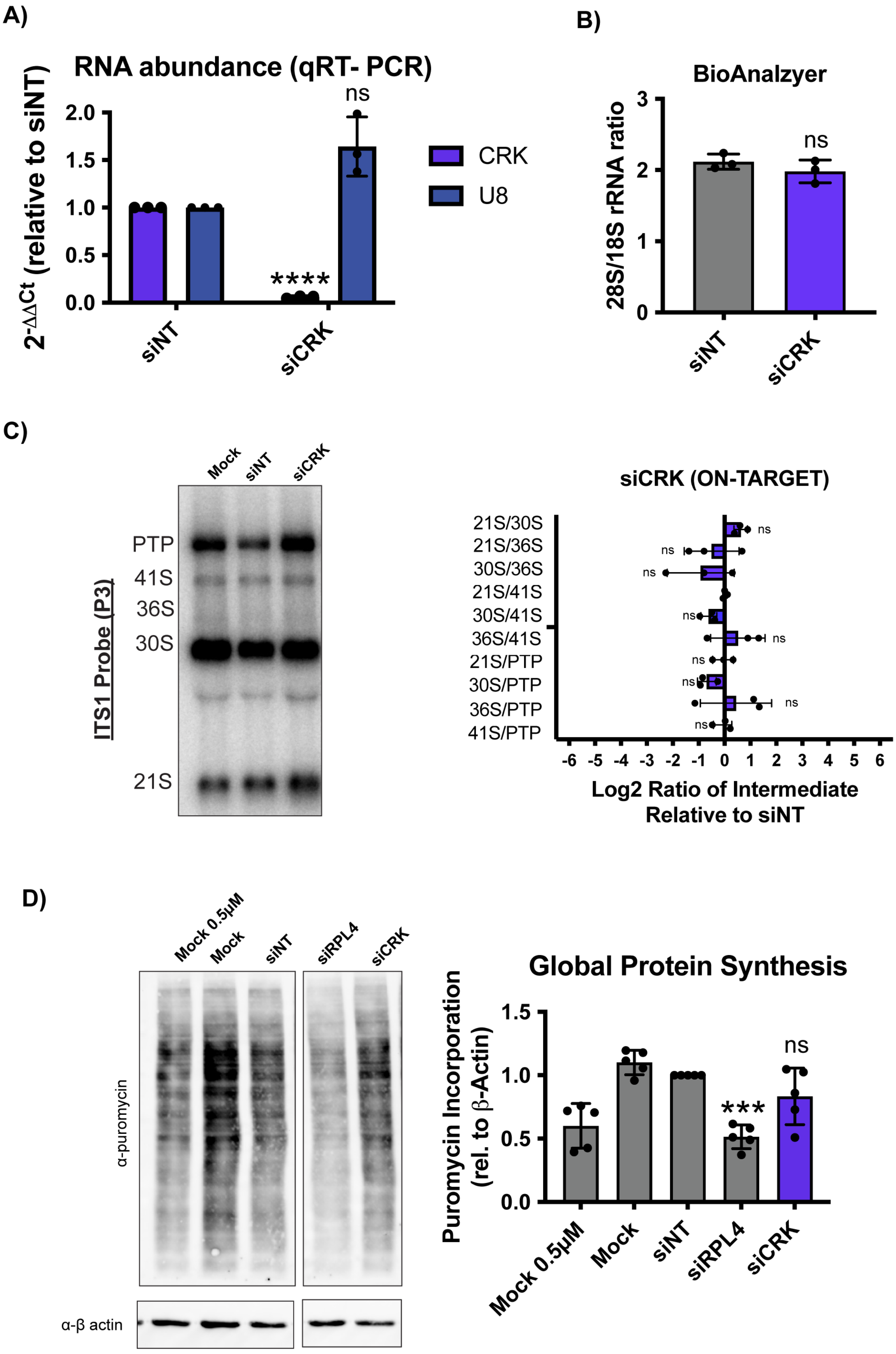

A)
RNA abundance (qRT- PCR)
$2^{-\Delta\Delta Ct}$ (relative to siNT)
2.0
1.5
1.0
0.5
0.0
ns
****
CRK
U8
siNT
siCRK
B)
BioAnalzyer
28S/18S rRNA ratio
3
2
1
0
ns
siNT
siCRK
C)
Mock
siNT
siCRK
ITS1 Probe (P3)
PTP
41S
36S
30S
21S
siCRK (ON-TARGET)
21S/30S
21S/36S
30S/36S
21S/41S
30S/41S
36S/41S
21S/PTP
30S/PTP
36S/PTP
41S/PTP
ns
-6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6
Log2 Ratio of Intermediate Relative to siNT
D)
Mock 0.5µM
Mock
siNT
siRPL4
siCRK
α-puromycin
α-β actin
Global Protein Synthesis
Puromycin Incorporation (rel. to β-Actin)
1.5
1.0
0.5
0.0
***
ns
Mock 0.5µM
Mock
siNT
siRPL4
siCRK

Figure 2.10: Depletion of CRK using siON-TARGET (Dharmacon) siRNAs does not produce the previously observed ribosome biogenesis defects using siGENOME siRNAs. **(A)** CRK siON-TARGET depletion reduces *CRK* mRNA levels but does not change U8 snoRNA steady-state levels in MCF10A Cells. qRT-PCR measuring *CRK* (left) and U8 (right) transcript levels in MCF10A cells. $2^{-\Delta\Delta Ct}$ were measured relative to 7SL internal control and siNT negative control sample. 3 technical replicates of 3 biological replicates, plotted mean ± SD. Data were analyzed by Student's t-test, **** $p \leq 0.0001$. **(B)** CRK siON-TARGET depletion does not alter normal mature 28S/18S rRNA ratios in MCF10A. Agilent BioAnalyzer analysis of ratio of mature 28S to 18S rRNAs from MCF10A cells depleted of CRK using siON-TARGET siRNAs. Three biological replicates, plotted mean ± SD. The data were analyzed by Student's t-test relative to siNT negative control. **(C)** CRK siON-TARGET depletion does not alter pre-rRNA processing in MCF10A cells. (Left) Representative northern blot using an ITS1 probe to measure steady-state levels of pre-rRNAs that lead to the formation of the small ribosomal subunit (SSU) in MCF10A cells. Mock transfected and siNT are negative controls. (Right) Quantification of pre-rRNA processing. Ratio analysis of multiple precursors (RAMP) (Wang et al. 2014) quantification of northern blotting using an ITS1 probe. Log_2 fold change of pre-rRNA ratios were quantified relative to siNT negative control. Three biological replicates plotted mean ± SD. Data were analyzed by two-way ANOVA. **(D)** CRK siON-TARGET depletion does not change global protein synthesis in MCF10A cells. After 72 h siRNA depletion, 1μM puromycin was added for 1 h to measure translation. (Left) Representative western blot using an α-puromycin antibody. Mock and siNT are negative controls, siRPL4 is a positive control, and Mock 0.5μM is a control to indicate robust quantification. α-β-actin is shown as a loading control. (Right) Quantification of puromycin signal normalized to β-actin signal and relative to siNT negative control. 5 biological replicates plotted mean ± SD. Data were analyzed by one-way ANOVA with Dunnett's multiple comparisons test, *** $p \leq 0.001$.

Testing other candidate U8 stability regulators, the LSm protein complex

After testing DCP2 and NUDT16's regulation of U8 levels and ability to rescue CRK siGENOME depletion reduction in U8 levels, I was inspired to further interrogate other U8 stability regulators. Among these are the LSm (like Sm) complex of proteins which form a heptameric ring (cytoplasmic LSm 1-7, or nuclear LSm 2-8) to bind various RNA species and regulate stability, splicing, and processing [reviewed in (Beggs 2005)]. Furthermore, they have been shown to bind the U8 snoRNA in *Xenopus* through a conserved sequence (Tomasevic and Peculis 2002) and regulate pre-rRNA processing in yeast (Kufel et al. 2003). Mutations within the LSm binding site on U8 are associated with cerebral microangiopathy leukoencephalopathy with calcifications and cysts (LCC), emphasizing the importance of U8's interaction with the LSm complex (Jenkinson et al. 2016). However, the molecular mechanisms regarding LSm regulation of U8 snoRNA function in humans remain understudied to date. Therefore, I aimed to fill this gap in knowledge. I hypothesized that depletion of LSm complex members would reduce U8 stability and its steady-state levels.

To test for changes in U8 levels after LSm complex member depletion, I performed qRT-PCR in MCF10A cells after siRNA depletion of LSm3 and the nuclear specific LSm protein, LSm8. I leveraged U6 snRNA as a positive control, since the nuclear LSm complex has been shown to be important for its stability previously through interaction with its 3' end (Mayes et al. 1999; Zhou et al. 2014). Surprisingly, I did not observe any changes in U8 snoRNA or U6 small nuclear RNA (snRNA) steady-state levels after LSm3 or LSm8 siRNA depletion in MCF10A cells (Figure 2.11A). Contrary to the expected decreases in U6 and U8 abundance, I observed a trending increase after LSm depletion, which cannot be explained by the current understanding of LSm function. I was able to successfully confirm depletion of both LSm3 and LSm8 upon siRNA depletion as well (Figure 2.11B). The cause of these unexpected results remains to be explored, whether they are due to technical limitations, compensations, or a yet to be discovered mechanism of LSm regulation of U6 snRNA and U8 snoRNA.

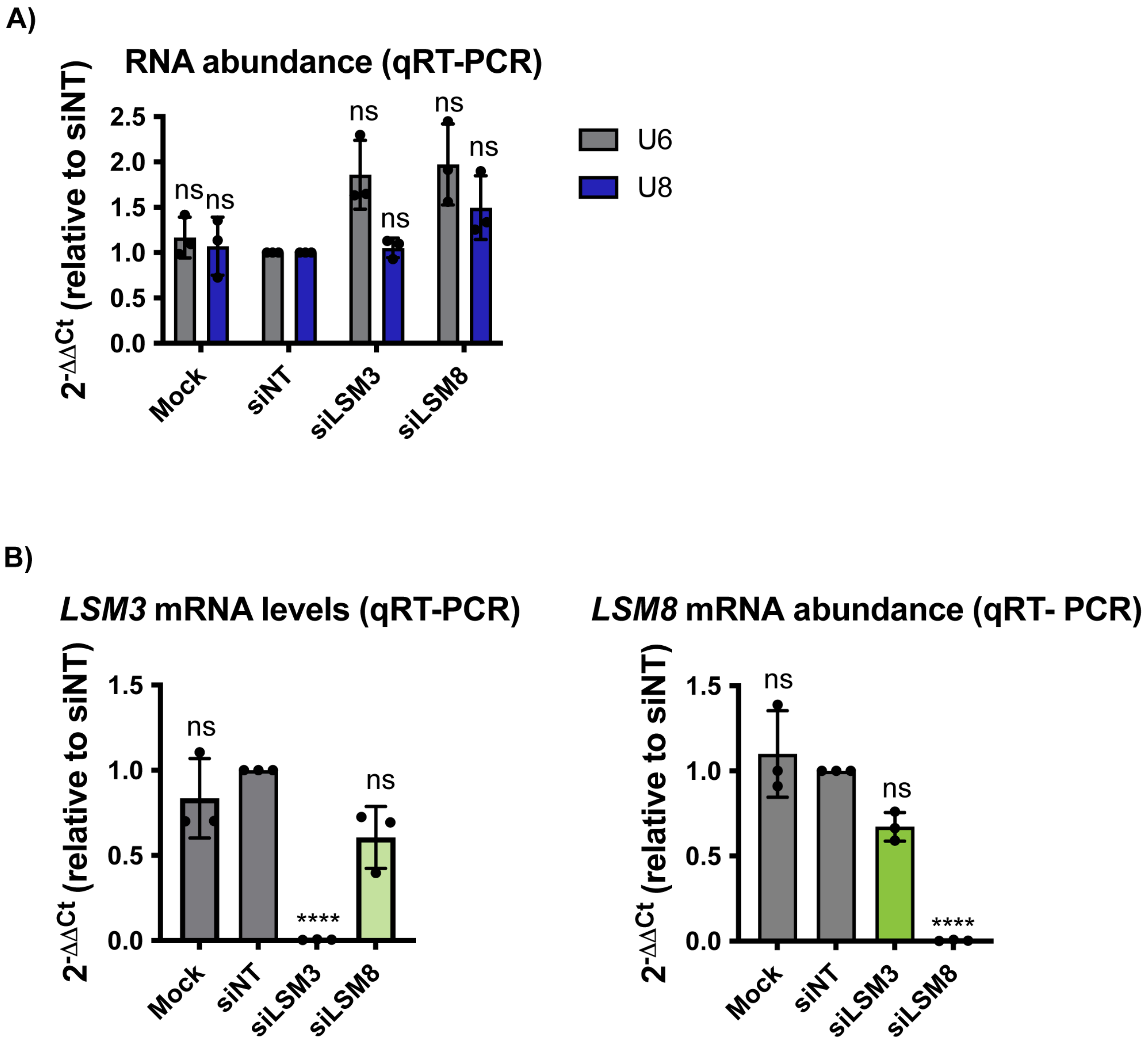


Figure 2.11: LSm3 and LSm8 siRNA (siON-TARGET, Dharmacon) depletion does not alter U6 or U8 levels in MCF10A cells. **(A)** qRT-PCR measuring U6 (grey) and U8 (blue) steady-state transcript levels in MCF10A cells. $2^{-\Delta\Delta Ct}$ were measured relative to 7SL internal control and siNT negative control sample. Mock transfection is an additional negative control. 3 technical replicates of 3 biological replicates, plotted mean ± SD. Data were analyzed by Student's t-test. **(B)** siON-TARGET siRNAs efficiently deplete *LSm3* and *LSm8* transcript levels in MCF10A cells. Same as in (A) except *LSm3* (left) and *LSm8* (right) were measured.

Discussion

Off-target effects are a concern in any RNAi based experiment and require rigorous confirmation to ensure they are not producing the observed phenotypes in experiments. First, I was able to confirm depletion of the target of interest (CRK) and associated phenotypes and Katie Farley-Barnes successfully completed deconvolution of the siRNA pool of 4, confirming all 4 individual siRNAs produce the one-nucleolus phenotype. However, the next step I took with an alternative siRNA chemistry did not yield the same results or any discernable defects in pre-rRNA processing, U8 snoRNA levels, or global protein synthesis. Thus, I attributed these previous results to be likely caused by an siRNA off-target effect. Even if multiple depletion methods are successful, it is important to perform rescue experiments where the expression of an siRNA-resistant version of your target is expressed in cells and the observable defects are restored to normal. Ultimately, thanks to rigorous work, I was able to show that CRK was not actually required for U8 snoRNA stability.

While the DCP2 holoenzyme, decapping complex, and other interacting partners are known for DCP2 (Brehme et al. 2009), NUDT16 RNA targets, localization, and interacting partners remain a mystery in human cells. NUDT16 was considered U8 specific in Xenopus (Tomasevic and Peculis 1999; Ghosh et al. 2004), but it is unclear if NUDT16 has the same specificity in human cells since studies to date have used exogenously expressed tagged versions of NUDT16 or correlative assays in cells to identify localization and putative targets (Lu et al. 2011; Anadon et al. 2017; Zhang et al. 2020). Due to this lack of clarity, I worked to make a useful tool for NUDT16 study by obtaining a human NUDT16 specific antibody. I cloned, expressed, and purified recombinant full length human NUDT16 protein in *E. coli* alongside a high school student Charlie Seymour. I then sent the purified protein to Pocono Rabbit Farm for antibody sera production. Unfortunately, after a preliminary test indicated a NUDT16 antibody capable of detecting endogenous NUDT16 by western blot of MCF10A and HeLa cell lysate, more rigorous testing indicated that this antibody was not specific. Therefore, more work will be necessary in the future to develop the proper tools needed to study NUDT16 function in human cells.

There are several other factors that could regulate U8 snoRNA stability beyond the ones tested here. For example, U8 is a C/D box snoRNA that forms together with protein components, the C/D box snoRNP (Fibrillarin, NOP56, NOP58, SNU13, and the TMG cap is added by TSG1 (Figure 2.12A). Furthermore, contrary to the pan-eukaryotic U3 snoRNA, U8 is vertebrate specific (Peculis and Steitz 1993; Langhendries et al. 2016), increasing the likelihood for yet to be discovered regulators. Mutations across the transcript are associated with disease (Jenkinson et al. 2016), highlighting the importance of the entire transcript (Figure 2.12B). Understanding U8's function in human cells, especially any pathways that control its function (albeit not through CRK signaling), will be vital for treatment of U8 associated disease, both ribosomopathies and cancer.

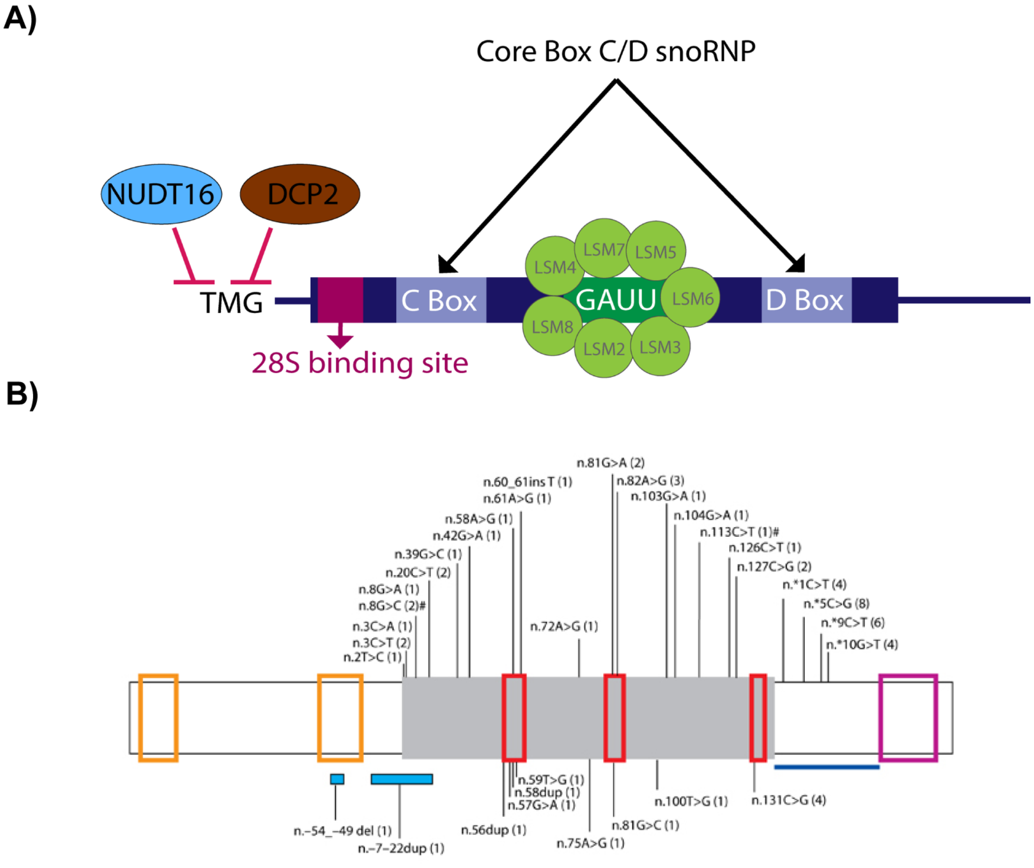


Figure from Jenkinson et al 2016

Figure 2.12: U8 has many functional and regulatory sequences and interacting partners. Several mutations along U8 are associated with the disease cerebral microangiopathy leukoencephalopathy with calcifications and cysts (LCC). **(A)** Schematic of U8 with important sequences and interacting partners indicated. **(B)** Schematic of U8 showing

disease associated mutations, taken from (Jenkinson et al. 2016). Red boxes indicate C box, LSm protein binding site, and D box sequences from left to right.

Materials and Methods

Cell culture

MCF10A cells (ATCC, CRL-10317) were subcultured in Dulbecco's modified Eagles' medium/Nutrient mixture F-12 (Gibco, 1130-032) containing horse serum (Gibco, 16050), 10 µg/mL insulin (Sigma, I1882), 0.5 µg/mL hydrocortisone (Sigma H0135), 100 ng/mL cholera toxin (Sigma, C8052), and 20 ng/mL epidermal growth factor (Peprotech, AF-100-15). HeLa cells (ATCC, CCL-2) cells were grown in DMEM (Gibco, 41965-062) with 10% fetal bovine serum (FBS, Gibco, 10438026). Cell lines were maintained at 37°C, in a humidified atmosphere with 5% CO_2.

RNAi

All siRNAs were purchased from Horizon Discovery Biosciences (previously Dharmacon). siGENOME or siON-TARGET siRNAs were used as indicated. siRNA transfection was performed using Lipofectamine RNAiMAX Transfection Reagent (Invitrogen, 13778150) per manufacturer's instructions with a final siRNA concentration of 33 nM for all assays (co-depletions had a combined siRNA concentration of 33 nM).

Protein harvesting and western blotting

Total protein was harvested from cells by scraping followed by PBS rinse. Cells were lysed using AZ lysis buffer (50 mM Tris pH 7.5, 250 mM NaCl, 1% Igepal, 0.1% SDS, 5 mM EDTA pH 8.0) with protease inhibitors (cOmplete^{TM} Protease Inhibitor Cocktail, Roche, 11697498001) for 15 minutes at 4°C by vortexing. Lysed cells were spun at 21000 RCF for 15 minutes at 4°C, supernatant was harvested and protein concentration was determined by the Bradford assay (Bio-Rad). Either 50 µg or 25 µg (puromycin blots only) of total protein was separated by SDS-PAGE and transferred to a PVDF membrane (Bio-Rad, 1620177) for blotting. The following primary antibodies were used: α-puromycin (Kerafast, EQ0001), α-β-actin (Sigma Aldrich, A1978), and α-CRK (Sigma Aldrich, MABC172). Secondary antibodies were α-mouse HRP conjugated (GE Healthcare NXA931). Images were acquired by exposing to film and analyzed using ImageJ software.

qRT-PCR

Total cellular RNA was extracted using TRIzol (Life Technologies, 5596018) per manufacturer's instructions. Prior to cDNA synthesis, all $A_{260/230}$ values were above 1.7 by NanoDrop (ThermoFisher, ND2000CLAPTOP). cDNA was made from 1µg total RNA using iScript gDNA clear cDNA synthesis kit (Bio-Rad, 1725035) with random primers. qPCR was performed with iTaq Universal SYBR Green Supermix (Bio-Rad, 1725121). The following amplification parameters were used: initial denaturation 95 ºC for 30 s, 40 cycles of 95 ºC for 15 s and 60 ºC for 30 s. Subsequent melt curve analysis was performed to ensure a single product, 95 ºC for 15 s, then gradual (0.3 ºC/15 s) increase from 60 ºC to 94.8 ºC. Gene specific primers were used. Amplification of 7SL RNA was used as an internal control and relative RNA levels were determined using comparative C_T method ($\Delta\Delta C_T$). Three technical replicates were performed for each biological replicate.

Northern blotting

Total cellular RNA was extracted using TRIzol (Life Technologies, 5596018) per manufacturer's instructions. For each sample, 3 µg of total RNA was resolved on a denaturing 1% agarose/1.25% formaldehyde gel using Tri/Tri buffer (Mansour and Pestov 2013) and transferred to a Hybond-XL membrane (GE Healthcare, RPN 303S). UV-crosslinked membranes were stained with methylene blue (0.025% w/v) and imaged. Blots were hybridized to ^{32}P radiolabeled DNA oligonucleotide probe (P3 (ITS1) 5'-AAGGGGTCTTTAAACCTCCGCGCCGGAACGCGCTAGGTAC-3') or U6 (5'-GAATTTGCGTGTCATCCTTGCGCAGGGGCCATGCTAA-3') (Hwang et al. 2007) and detected by a phosphorimager (Bio-Rad Personal Molecular Imager) and images were quantified using the associated software. Ratio-analysis of multiple precursors (RAMP) was performed (Wang et al. 2014).

BioAnalyzer

For each sample, 100 ng/µL total RNA in nuclease-free water was submitted for Agilent BioAnalyzer analysis, performed by the Yale Center for Genome Analysis.

Puromycin incorporation assay

Global protein synthesis was assessed as in (Schmidt et al. 2009). Following 72 h siRNA depletion, 1 μM puromycin (or 0.5 μM puromycin for Mock 0.5 control) was added to the media for 1 h to label nascent polypeptides. Then I proceeded with western blotting with α-puromycin antibodies as above.

Chapter 3

Efforts to validate RNAi screen hits and elucidate their roles in nucleolar rRNA biogenesis

Introduction

As I discovered myself, one concern using siRNAs is the presence of many potential off-target effects. These arise due to the inevitable partial complementarity (6 bases or more seed) of siRNAs to many transcripts in addition to possessing full complementarity (19 base pair match) towards a specific transcript of interest. While full complementarity allows for efficient Ago2 cleavage and degradation of the target of interest, partial complementarity in the 3'UTR of mRNAs can lead to siRNAs functioning as microRNAs (miRNAs) to inhibit translation or enhance decay of several transcripts (Figure 3.1) (Jackson et al. 2003; Birmingham et al. 2006). On some occasions, these manifest themselves in observable experimental phenotypes or introduce increased experimental noise (Fedorov et al. 2006). In RNAi based screening methods this leads to false positives and negatives (Sigoillot et al. 2012), highlighting the importance of follow-up and validation experiments when using this technology.

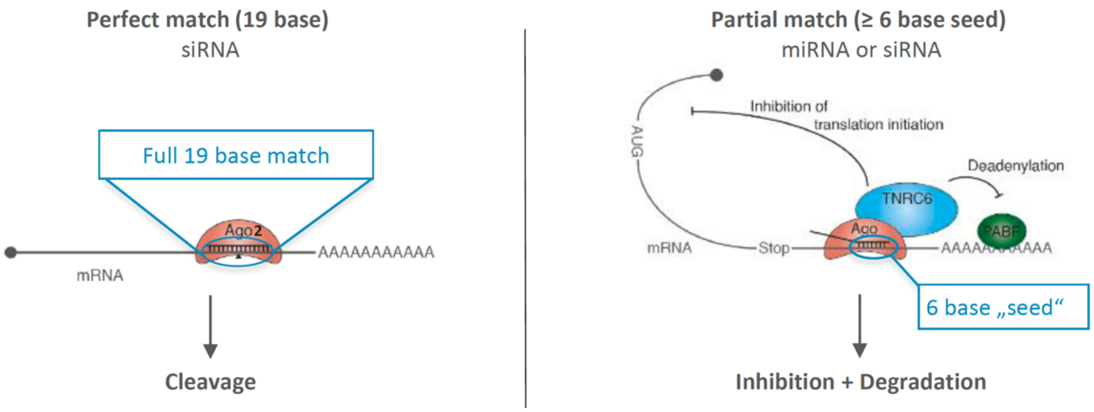


Figure 3.1: siRNAs can efficiently deplete their target transcript of interest through full complementarity (left) and also be "micro-RNA-like" to deplete or inhibit translation of many transcripts through partial binding of their 3'UTRs (right). Figure taken from (siTOOLs BioTech, Technote 1, https://www.sitoolsbiotech.com/technotes.php)

The Baserga laboratory previously performed a genome-wide RNAi-based screen to identify novel regulators of nucleolar function that upon depletion lead to either decreases or increases in normal nucleolar number in MCF10A cells (Farley-Barnes et al. 2018; Ogawa et al. 2021). The screen identifying factors whose depletion led to decreases in nucleolar number (Farley-Barnes et al. 2018) identified 191 top hits and the screen identifying factors whose depletion led to increases in nucleolar number (Ogawa et al. 2021) identified 186 top hits that were greater than 3 standard deviations from their respective screen mean percent effects. These were narrowed down to 139 and 113 hits respectively after filtering hits for expression and cell viability. Fortunately, siRNA pool deconvolution (to observe multiple of the individual siRNAs comprising the pool that was screened with produced the one-nucleolus or 5+ nucleoli phenotypes) and rigorous follow-up experiments on subsets of hits from both screens revealed diverse phenotypes, where depletion of screen hits inhibited various parts of the entire ribosome biogenesis pathway. This provides convincing evidence that the screen hits as a whole were not driven by an enrichment of a small amount of off-target seed sequences targeting only one or a few nucleolar regulating transcripts.

Even with this confidence and impressive screen statistics from these screens, it is clear more validation work must be completed when following up individual screen hits. For example, there was only one replicate completed for these original screens to identify factors that when depleted produced a change in nucleolar number (Farley-Barnes et al. 2018; Ogawa et al. 2021), increasing likelihood of false positives and false negatives. Furthermore, I was able to identify a probable off-target effect through my study of CRK (Chapter 2). Thus, I aimed to help the laboratory produce a higher confidence hit list for individual follow-up studies by building upon these initial screening results.

Given the importance of nucleolar function in human health and disease, the creation of more robust tools for measuring rRNA biogenesis within the nucleolus is essential for understanding the basic biological mechanisms through which ribosome biogenesis (RB in this chapter) can be regulated, as well as for developing next-generation small molecule or biologic therapeutics. In the past decade, a number of studies using high-throughput screening (HTS) have elucidated novel mechanisms through which human RB is regulated, including the above screens for nucleolar number by our laboratory (Tafforeau et al. 2013; Badertscher et al. 2015; Farley-Barnes et al. 2018; Ogawa et al. 2021). Several candidate therapeutics targeting the nucleolus have also been discovered with HTS chemical library or natural product campaigns (Drygin et al. 2011; Peltonen et al. 2014b; Scull et al. 2019; Ferreira et al. 2020; Kirsch et al. 2020). While several HTS modalities for monitoring nucleolar form and morphology have been described (Farley-Barnes et al. 2018; He et al. 2018; Stamatopoulou et al. 2018), none of these platforms directly measure nucleolar rRNA biogenesis, or the synthesis and accumulation of nascent pre-rRNA within the nucleolus. To date, the lack of a direct high-throughput assay for nucleolar rRNA biogenesis constrains our ability to select for and validate the most promising candidate regulators of RB.

To monitor nucleolar function in a high-throughput manner, Carson Bryant sought to adapt a 5-ethynyl uridine (5-EU) assay for nucleolar rRNA biogenesis to an accessible, miniaturized format. The 5-EU assay has been successfully used to quantify changes in nucleolar transcriptional activity by several other groups in a variety of systems including human tissue culture cells (Bai et al. 2013; Lafita-Navarro et al. 2016; Cheng et al. 2017; Hayashi et al. 2017; Calo et al. 2018; Hayashi et al. 2018; Rossetti et al. 2018; Stamatopoulou et al. 2018; Dong et al. 2021). A key limitation in almost all of these studies is that total cellular or total nuclear 5-EU is quantified, rather than solely nucleolar 5-EU. Because only nucleolar signal corresponds to biogenesis of the primary pre-rRNA, quantifying total 5-EU leads to increased background from nascent transcription by RNA polymerases other than RNAP1.

Here, I report a targeted rescreen of the top hits from the nucleolar number screen using siON-TARGET siRNA technology from Dharmacon. Compared to the siGENOME library with which theses original screens were completed, the siON-TARGET siRNAs are designed to produce fewer off-target effects (Jackson et al. 2006). Thus, this technology allowed us to maintain the same screening platform but to improve upon our method of depletion. First, I helped curate a list of hits to rescreen which was completed by Lisa Ogawa. Next, I leveraged a miniaturized high-throughput 5-EU incorporation assay developed by Carson Bryant to test these rescreened hits for a role in nucleolar rRNA biogenesis (a more direct readout of nucleolar function than the original screens). From these results, we arrived at a more stringent hit list with an associated information-rich dataset of novel human ribosome biogenesis factors deserving of follow-up mechanistic studies. Additionally, the nucleolar rRNA biogenesis assay results showed that it actually reported a combination measurement of pre-rRNA transcription, processing, stability, and assembly. Our results prompt an expansion of the field's conceptualization of nucleolar 5-EU incorporation experiments in general, which, at measurable time points, report not only on RNAP1 transcription, but more broadly on nucleolar rRNA biogenesis. Overall, the targeted nucleolar number re-screen and nucleolar rRNA biogenesis assay provided us with a high confidence list of novel ribosome biogenesis factors and additional mechanistic insight into which step(s) of ribosome biogenesis they take part in. Some of the results presented here I published in a co-first author manuscript with Carson Bryant in *Open Biology* (Bryant et al. 2022) or are parts of other projects leading to published or in-progress manuscripts (Chapters 4-6).

Results

Curating a list of hits for targeted rescreening

The original screen was completed on a genome-wide siGENOME siRNA library from Dharmacon targeting a total of 18,107 protein coding genes. From the hit list, I collaborated with and consulted with the entire Baserga laboratory and chose to use siON-TARGET siRNAs to validate only the top hits. More specifically, I chose the top one-nucleolus hits that were at least 2 standard deviations from the screen mean (N = 768) to be more inclusive than the published stringent 3 standard deviation cutoff (Farley-Barnes et al. 2018). These 2 standard deviation hits still produced a one-nucleolus percent effect of at least 91%, almost reaching the positive control siUTP4 (100% one-nucleolus percent effect). For the 5+ nucleoli hits, I adhered to the 3 standard deviation from the screen mean (N=186) list since the top hits produced a more mild quantifiable phenotype (i.e. the 5+ percent effect was only at least 25% in the 3 standard deviation hit list). Next, I filtered these hits for those that were expressed in at least 1 of 4 MCF10A RNA-seq datasets as calculated by Carson Bryant using a (zTPM > -3) cutoff to call expression [BioProjects: PRJNA290557 (GEO GSM1829628), PRJNA384982 (GEO: GSM2593351, GSM2593352, GSM2593353), PRJNA530983 (GEO: GSM3711368, GSM3711369), PRJNA647393 (GEO: GSM4667014, GSM4667015, GSDM4667016)]; a cutoff for a gene being active previously established by (Hart et al. 2013). Two previous hits were excluded due to Dharmacon having siON-TARGET siRNAs targeting them (OSTCP1 and ZNF322P1). I also included other genes of interest in this library including: (YARS, XPOT, NOL7, DCP2, and NUDT16) other genes tested in Chapter 5 were added manually to the plates] for a total of 702 siON-TARGET siRNA pools (Figure 3.2) for rescreening.

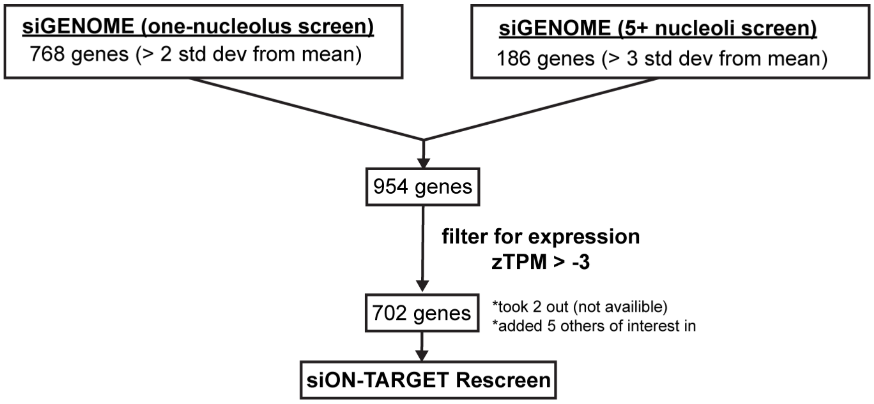


Figure 3.2: Choosing siON-TARGET siRNAs to use for re-screening the hits from the Baserga laboratory's original screen for nucleolar regulators. (Farley-Barnes et al. 2018) had 768 hits (> 2 standard deviations from mean) and (Ogawa et al. 2021) had 186 hits (> 3 standard deviations from the mean). Hits were filtered for expression, availability of reagents, and other genes of interest were added for a total of 702 targets to be re-screened.

Targeted rescreen and siRNA deconvolution of the original one-nucleolus hits

Here, I will focus on briefly recapping the results from re-screening the previous hits that were found to reduce nucleolar number upon their depletion. Lisa Ogawa took this library and performed a targeted rescreen in quadruplicate using the same exact protocol as established previously (Farley-Barnes et al. 2018; Ogawa et al. 2021). For this rescreen screen a cutoff of > 3 standard deviations from the negative (siNT) control was used. Of the 587 original one-nucleolus hits, 147 passed this cutoff in at least 2 of the 4 replicates. As expected, CRK was not a hit in this rescreen, only producing an average one-nucleolus percent effect of -3.16%. While Lisa Ogawa obtained a similar number of hits as was in the original number screen, we have more confidence in this list of proteins playing a role in ribosome biogenesis.

To reduce the likelihood that a hit is a result of an siRNA off-target effect, Laura Abriola from the YCMD and Ty Brown performed siRNA deconvolution on the 147 one-nucleolus hits as well as the 5+ rescreen hits identified. This dataset was valuable when I followed up on individual hits (Chapters 4-6).

Nucleolar rRNA biogenesis assay as an additional screening high-throughput method

Development of a 5-EU assay to quantify nucleolar rRNA biogenesis

To achieve specific quantification of nucleolar rRNA biogenesis, Carson Bryant introduced a 5-EU labeling step into our previously established screening platform for counting nucleolar number (Farley-Barnes et al. 2018), which utilizes CellProfiler (McQuin et al. 2018) to segment nuclei and nucleoli in images of cells immunofluorescently stained for DNA and the nucleolar protein, fibrillarin (FBL) (Figure 3.3). In his new protocol, MCF10A breast epithelial cells are reverse-transfected with siRNA duplexes for 72 h. For one hour following the transfection period, the cells are treated with 1 mM 5-EU, which is incorporated into nascent transcripts. Since the bulk of cellular transcription occurs in the nucleolus, most of the 5-EU label is incorporated into nucleolar nascent pre-rRNA (Figure 3.3). The cells are fixed and immunofluorescently stained for DNA and FBL, after which nascent RNA is visualized *in situ* by performing a bio-orthogonal click reaction to covalently label the 5-EU alkyne moiety with an azide fluorophore (AF488 azide) (Figure 3.3). The cells are then imaged and analyzed with CellProfiler to specifically quantify nucleolar rRNA biogenesis across all control and unknown wells.

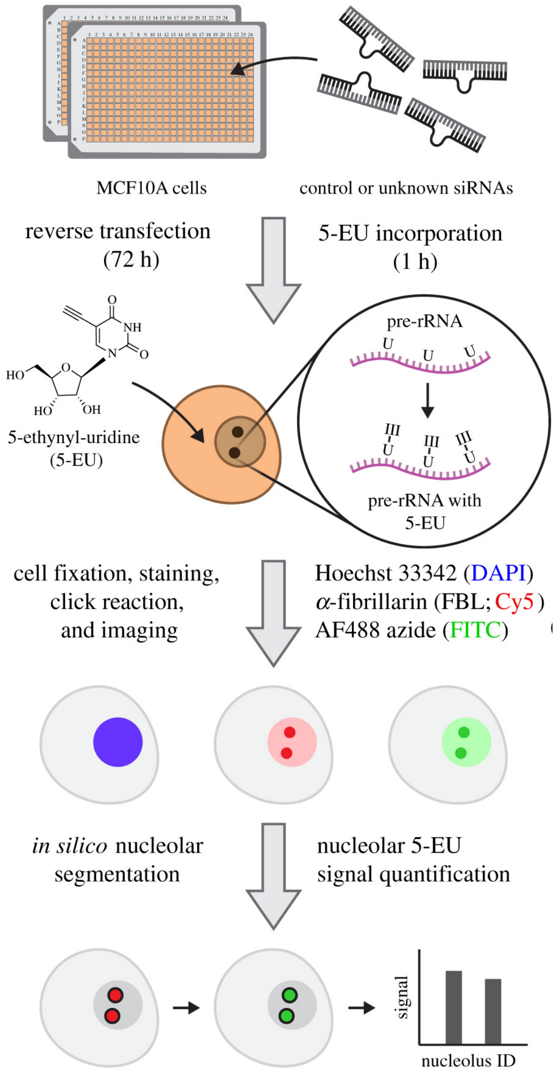


Figure 3.3: Schematic of the 5-EU assay protocol. MCF10A cells are reverse-transfected in 384-well imaging plates with control or unknown siRNAs for 72 h. Following target depletion, 5-EU is incorporated into nascent RNA transcripts for 1 h, with the majority of label incorporated into nascent pre-ribosomal RNA (pre-rRNA). Treated cells are fixed and stained for DNA (Hoechst 33342, DAPI channel) and the nucleolar protein FBL (Cy5 channel). 5-EU in nascent transcripts are conjugated to an azide fluorophore (AF488 azide, FITC channel) via a copper-catalyzed click reaction. After fluorescent imaging, cell nuclei and nucleoli are segmented *in silico* with CellProfiler, and nucleolar-specific 5-EU signal is quantified for each nucleolus object identified. Figure generated by Carson Bryant, taken from (Bryant et al. 2022)

Carson optimized the 5-EU assay to use a non-targeting siRNA as a negative control (siNT), and an siRNA targeting *POLR1A*, the largest subunit of RNAP1 also known as RPA194, as a positive control (siPOLR1A) (Figure 3.4). RNAP1 inhibition by POLR1A depletion strongly reduces the nucleolar 5-EU signal to a degree consistent with acute treatment with BMH-21, a potent small molecule inhibitor of RNAP1 (Peltonen et al. 2014a; Peltonen et al. 2014b) (Figure 3.4, compare siNT to siNT + BMH and siPOLR1A). However, it is clear that residual nucleoplasmic 5-EU signal remains even after RNAP1 inhibition, (Figure 3.4) (siNT + BMH and siPOLR1A), emphasizing the importance of only quantifying 5-EU staining within the nucleolus via FBL co-staining.

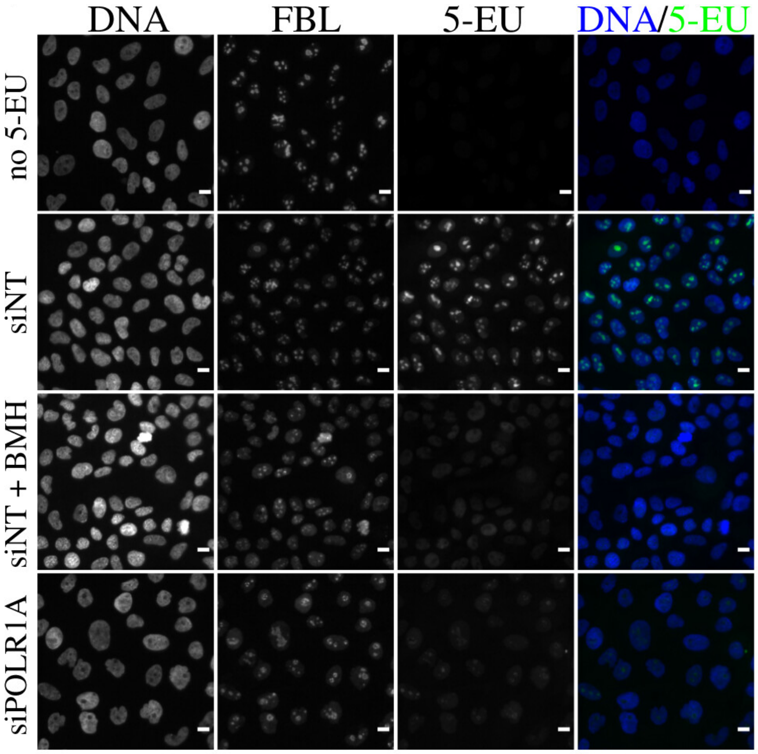


Figure 3.4: RNAP1 inhibition specifically inhibits nucleolar 5-EU incorporation. No 5-EU, experiment without 1 h 5-EU incorporation. Treatment with a non-targeting siRNA (siNT) leads to a high 5-EU signal within the nucleolus and moderate nucleoplasmic background signal. Acute treatment with BMH-21 (siNT + BMH) or siRNA-mediated depletion of POLR1A (siPOLR1A) decreases nucleolar 5-EU signal, although nucleoplasmic background remains. DNA (Hoechst staining), FBL (staining), 5-EU (5-EU staining) and DNA/5-EU (combined Hoechst and EU staining). Scale bars, 10 μm. Figure generated by Carson Bryant, taken from (Bryant et al. 2022)

To achieve nucleolar 5-EU quantification during analysis, images of DNA and FBL staining (Figure 3.5, panels 1 and 2) were first used to segment nuclei and nucleoli by CellProfiler (Figure 3.5, panels 3 and 4), respectively. Then, the median 5-EU signal within each nucleolus was measured (Figure 3.5, panel 5), enabling aggregate quantification analysis per treatment condition across every nucleolus within each well

(Figure 3.5, panel 6). Final calculation of mean signals, percent inhibitions (by normalization to the negative and positive controls), and screening statistics including signal-to-background (S/B) and Z' factor.

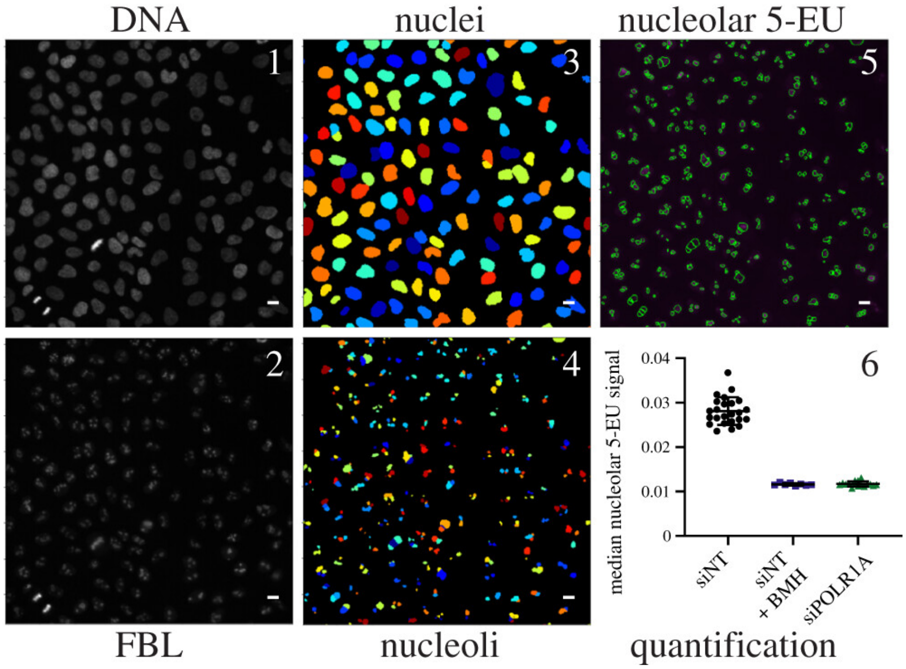


Figure 3.5: Schematic of CellProfiler segmentation and nucleolar 5-EU quantification. Panels 1 and 2, raw images of DNA and FBL staining. Panels 3 and 4, nuclei or nucleoli segmented by CellProfiler from DNA or FBL staining, respectively. Rainbow coloring identifies object number. Panel 5, overlay of segmented nucleoli (green) on top of 5-EU staining (magenta). Panel 6, quantification of median nucleolar 5-EU signal for nucleoli in cells treated with siNT, siNT and BMH-21, or siPOLR1A. *n* = 24, 8 or 16 wells, respectively. Scale bars, 10 µm. Figure generated by Carson Bryant, taken from (Bryant et al. 2022)

In the final phase of optimization, I studied how siRNA knockdown of known RB factors affected nuclear and nucleolar 5-EU signal. Carson Bryant and I chose to deplete NOL11, a small subunit processome factor critical for pre-rRNA transcription (Freed et al. 2012), or POLR1A, the largest subunit of the RNAP1 complex, as positive controls. I verified robust knockdown of *NOL11* or *POLR1A* mRNA transcripts using RT-qPCR (Figure 3.6A) Compared to treatment with siNT, depletion of NOL11 or POLR1A decreased maximum nuclear signal and median nucleolar signal by roughly 50% in each case (Figure 3.6B), corresponding to control S/B values of 1.9-2.0 for each control (Figure 3.6C). The median signal was chosen to be the measurement of nucleolar rRNA biogenesis because of its better Z’ factor (separation of control means) for both siNOL11 and siPOLR1A.

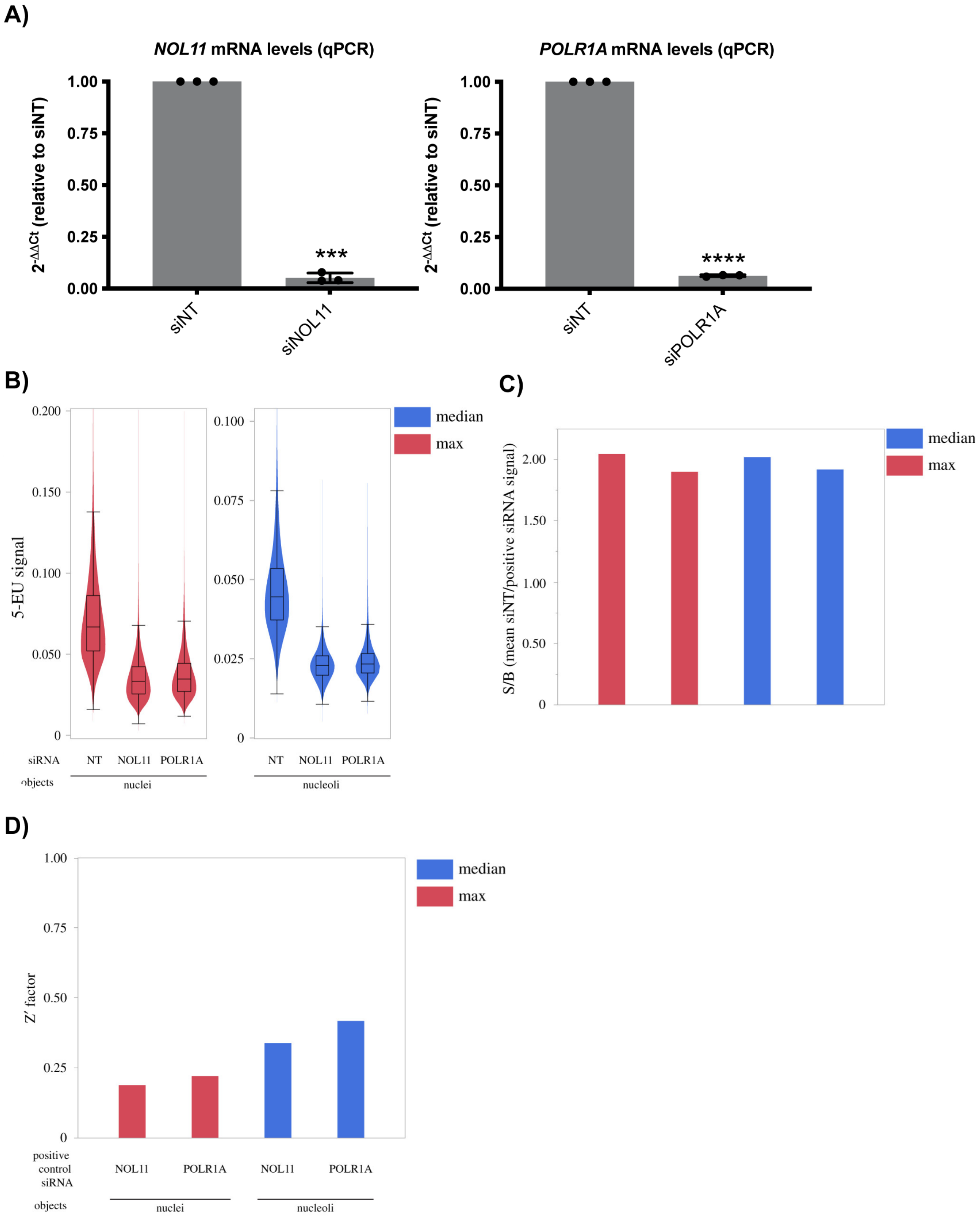


Figure 3.6: Optimization of the miniaturized 5-EU assay for nucleolar rRNA biogenesis. **(A)** Positive controls, siRNA depletion of NOL11 and POLR1A are confirmed by on mRNA level. qRT-PCR measuring *NOL11* (left) and *POLR1A* (right) steady-state transcript levels in MCF10A cells. $2^{-\Delta\Delta Ct}$ were measured relative to 7SL internal control and siNT negative control sample. 3 technical replicates of 3 biological replicates, plotted mean ± SD. Data were analyzed by Student's t-test. **(B)** Maximum nuclear 5-EU signal (red) or median nucleolar 5-EU signal (blue) for cells treated with siNT, siNOL11 or siPOLR1A. $n \geq 130\,000$ cells per siRNA. **(C-D)** Control S/B and Z' factor values for cells from (B) treated with siNOL11 or siPOLR1A as the positive control. Control S/B is calculated as the ratio of mean siNT-treated 5-EU signal divided by mean siNOL11- or siPOLR1A-treated 5-EU signal. Maximum nuclear 5-EU signal (red), median nucleolar 5-EU signal (blue). I generated (A) and Carson Bryant generated (B-D), all taken from (Bryant et al. 2022).

Validation of the 5-EU assay to quantify nucleolar rRNA biogenesis on a subset of known ribosome biogenesis factors

After Carson's optimization, I tested the high-throughput 5-EU assay by siRNA depletion of a subset of 68 previously studied RB factors, including RPs and assembly factors for both ribosomal subunits, as well as core RNAP1 machinery and drivers of transcription such as MYC (Figure 3.7A, Supplementary Table 1). This subset was found within the re-screen library of 702 siRNAs and had previously established roles in RB. I depleted each RB factor over 72 h using siRNA pools in accordance with our protocol, performing the assay in biological triplicate to ensure reproducibility. Carson Bryant then performed the quantification of these images as described above. Strikingly, we found that depletion of 58/68 ribosome biogenesis factors led to a significant (≥ 50%) inhibition of nucleolar 5-EU signal after standardization to the controls (Figure 3A). Images of the assay controls illustrate typical signal levels observed for the negative control siNT, set at 0% inhibition, and the positive control siPOLR1A, set at 100% inhibition (Figure 3.7B, siNT and siPOLR1A). Furthermore, images from the RB factors tested demonstrate the sensitivity of the assay to RNAP1 inhibition, from extreme effects above 100% inhibition (*e.g.* siMYC) to more moderate inhibitory effects (*e.g.* siTRMT112) (Figure 3.7B). Full results from the assay validation are presented in Figure 3.7C and Appendix I.

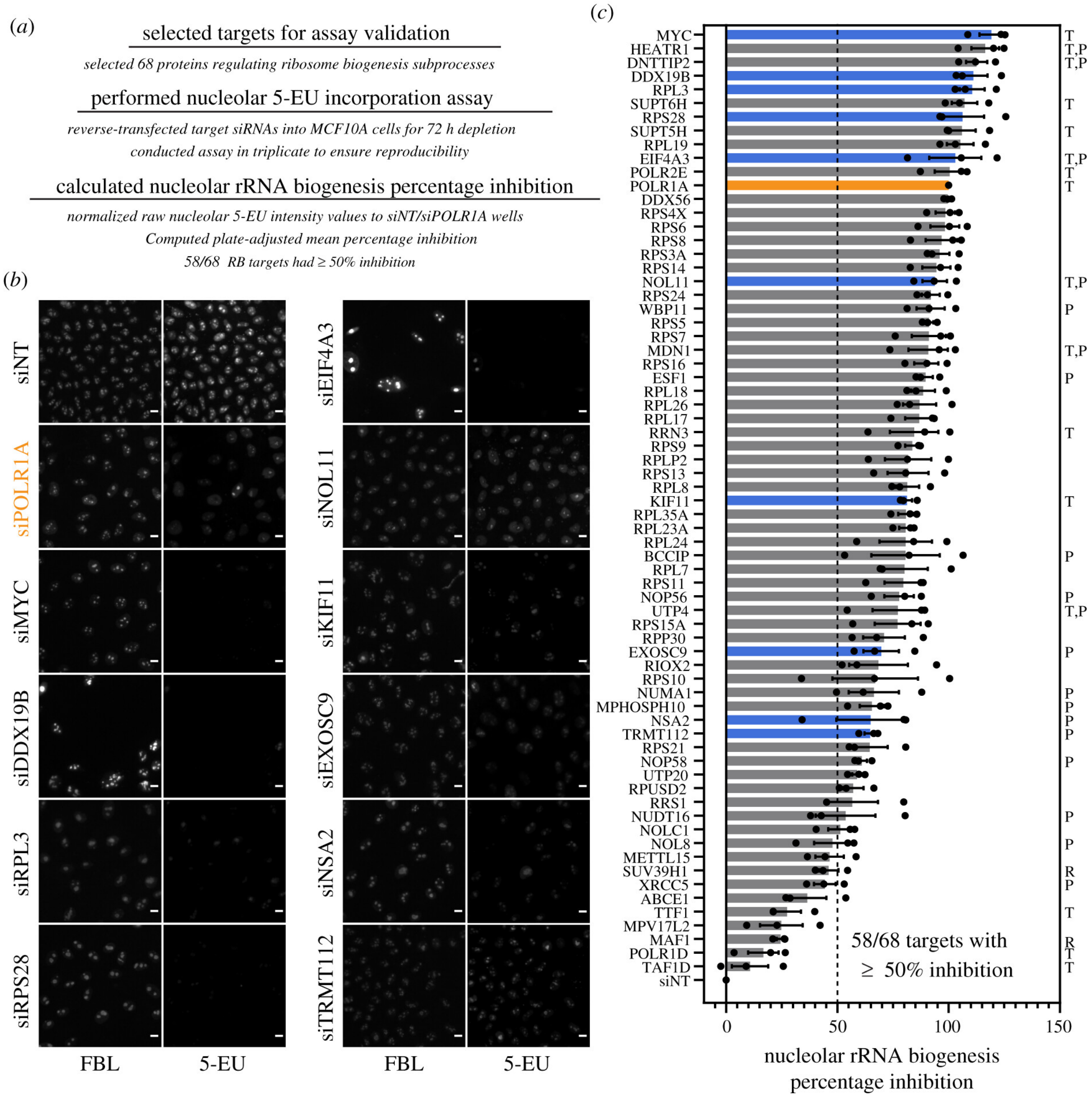


Figure 3.7: Validation of the 5-EU assay for nucleolar rRNA biogenesis on 68 known RB factors. **(A)** Outline of assay validation experiments. Sixty-eight proteins known to regulate RB subprocesses, including RNAP1 transcription and pre-rRNA processing, modification or stability, were selected for assay validation. The 5-EU assay was performed on cells depleted of these factors in biological triplicate, as described. **(B)** Representative images of FBL staining and 5-EU visualization for cells treated with siNT (negative control), siPOLR1A (positive control, orange), or a subset of siRNAs targeting known RB factors. **(C)** Nucleolar rRNA biogenesis percentage inhibition values for cells depleted of each known RB factor. Black dots, individual percentage inhibition values for one biological replicate. Solid bars, mean percentage inhibition ($n = 3$). Orange bar, POLR1A positive control (percentage inhibition = 100%). Blue bars, RB factors illustrated in (B). Letters to right indicate factors involved in RNAP1 transcription (T), pre-rRNA processing (P), or transcription repression (R). Figure generated by Carson Bryant, taken from (Bryant et al. 2022).

Strikingly, we observed 11 targets that resulted in stronger nucleolar rRNA biogenesis inhibition than the positive control, POLR1A; consistent with a mean percent inhibition greater than 100%, 7/11 of these targets are implicated in control of pre-rRNA transcription, including MYC (Grandori et al. 2005), HEATR1/UTP10 (Gallagher et al. 2004; Prieto and McStay 2007; Turi et al. 2018), DNTTIP2/TdIF2 (Koiwai et al. 2011), SUPT6H (Engel et al. 2015), SUPT5H (Farley-Barnes et al. 2018), EIF4A3/DDX48 (Zhang et al. 2011), and POLR2E (Goodfellow and Zomerdijk 2013). Overall, 12/58 factors with a significant inhibition of nucleolar rRNA biogenesis have been implicated in transcription, also including the RNAP1 initiation factor RRN3 (Bodem et al. 2000; Moorefield et al. 2000; Miller et al. 2001), two other t-UTPs, NOL11/UTP8 (Freed et al. 2012) and UTP4 (Freed and Baserga 2010), and the proteins MDN1, a pre-60S assembly factor, and KIF11, a mitotic kinesin essential for RB (Ogawa et al. 2021).

Pre-rRNA processing and modification factors comprised a sizeable subset of factors with significant nucleolar rRNA biogenesis mean percent inhibition. In total, 19/58 factors that inhibited nucleolar rRNA biogenesis were critical for processing, including the t-UTPs HEATR1/UTP10, NOL11/UTP8, and UTP4 (Gallagher et al. 2004; Prieto and McStay 2007; Freed and Baserga 2010; Freed et al. 2012; Turi et al. 2018), the C/D box snoRNP scaffolds NOP56 and NOP58 (Watkins and Bohnsack 2012), as well as other processing factors including DNTTIP2 (Tafforeau et al. 2013), WBP11 (Carbon and Mungall 2021), MDN1 (Raman et al. 2016), ESF1 (Tafforeau et al. 2013; Chen et al. 2018), BCCIP (Ye et al. 2020), RPP30 (Stolc and Altman 1997; Tafforeau et al. 2013), EXOSC9 (Muller et al. 2020), NUMA1 (Farley-Barnes et al. 2018), MPHOSPH10 (Westendorf et al. 1998), TRMT112 (Zorbas et al. 2015), UTP20 (Wang et al. 2007), and NUDT16 (Ghosh et al. 2004). In addition, nucleolar rRNA biogenesis was moderately inhibited for the pre-rRNA modification factors TRMT112 (Zorbas et al. 2015), RPUSD2 (Carbon and Mungall 2021), and NOLC1 (Yang et al. 2000; Werner et al. 2015). Notably, factors involved in transcription had a higher mean percent inhibition than factors involved in processing (83.1% inhibition v. 74.9% inhibition, n = 15 v. n = 22); factors involved in both transcription and processing had a mean percent inhibition of 99.0% (n = 6).

We also noted significant percent inhibition averages for 28 RPs from both the 40S and 60S subunits. Almost all RPs are essential for pre-rRNA biogenesis in the yeast *Saccharomyces cerevisiae* (Ferreira-Cerca et al. 2005; Poll et al. 2009) and in human cells (O'Donohue et al. 2010; Nicolas et al. 2016), compatible with a concomitant observed decrease in nucleolar rRNA biogenesis following their depletion.

Furthermore, of the 10 factors that had a mean percent inhibition value under 50%, five factors were either inhibitors of pre-rRNA transcription, including SUV39H1 (Murayama et al. 2008) and MAF1 (Upadhya et al. 2002; Bonhoure et al. 2020), mitochondrial ribosome biogenesis factors, including METTL15 (Van Haute et al. 2019; Chen et al. 2020) and MPV17L2 (Dalla Rosa et al. 2014), or ribosome recycling factors involved in translation, namely ABCE1 (Zhu et al. 2020).

Surprisingly, the other five RB factors with a mean percent inhibition less than 50% are well-appreciated for playing roles in pre-rRNA transcription, including POLR1D (Russell and Zomerdijk 2006), TAF1D (Gorski et al. 2007), and TTF1 (Evers and Grummt 1995), and in pre-rRNA processing, including NOL8 (Sekiguchi et al. 2006) and XRCC5/Ku86, which also aids TTF1 during RNAP1 termination (Wallisch et al. 2002; Shao et al. 2020). It is possible that these factors were not significantly depleted following transfection, or that, within our timeframe, the 5-EU assay cannot detect a significant change in nucleolar RNA levels as a result of non-concordant changes in both pre-rRNA transcription and stability.

The nucleolar rRNA biogenesis assay on the novel 702 hits

In addition to testing 68 known ribosome biogenesis factors, I also tested the rest of the hit library (N=702) which included many novel ribosome biogenesis factors whose roles have yet to be elucidated. Overall, 401/702 (57%) hits produced a nucleolar rRNA biogenesis percent inhibition greater than our empirical cutoff of 50%. Since the nucleolar rRNA biogenesis assay is a module added to the laboratory's original nucleolar number screening assay, I was also able to gain additional this additional data on top of Lisa Ogawa's data obtained with this same library (see above). Of the one-nucleolus hit subset (N=562), there was a modest positive correlation between one-nucleolus percent effect (Lisa Ogawa's results) vs. nucleolar rRNA biogenesis percent inhibition (Figure 3.8). The modest correlation emphasizes the diversity of hits identified with various 5-EU assay functional readouts (i.e. the one-nucleolus rescreen identified factors involved in various parts of the ribosome biogenesis pathway, not just those with high nucleolar rRNA biogenesis percent inhibition values). I will touch on examples of informative results from this assay in (Chapters 4-6) where I performed individual follow-up analyses on NOL7, RSL24D1, PeBoW (PES1-BOP1-WDR12) and APOBEC3A.

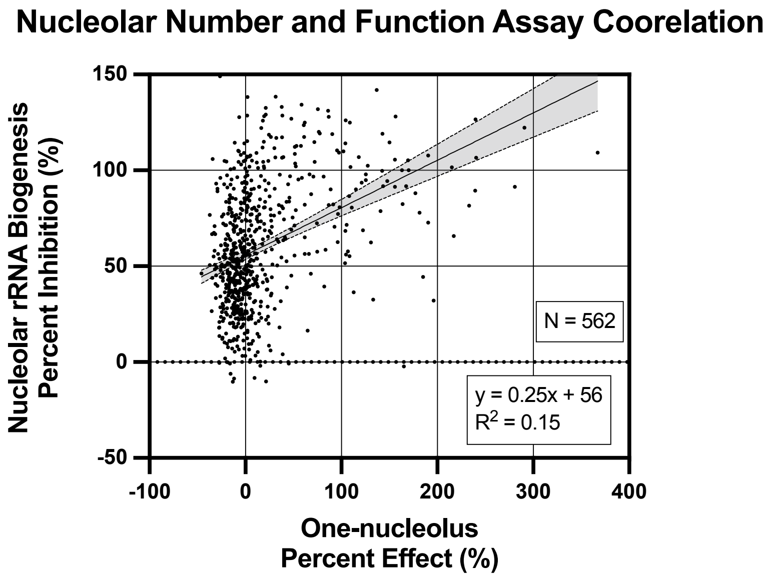


Figure 3.8: The one-nucleolus assay has an extremely weak positive correlation with the nucleolar rRNA biogenesis assay. One-nucleolus rescreened hits (N=562) were plotted for their average (N=3) one-nucleolus percent effect (x-axis) versus their average (N=3) nucleolar rRNA percent inhibition (y-axis) that was measured after siON-TARGET siRNA depletion. Both values were calculated from the same experiments. Simple linear regression was performed, 95% confidence interval (grey shading).

Discussion

siRNA off-target effects can lead to false positives and negatives within high-throughput RNAi screens. Since repeating the genome-wide screen with a new siON-TARGET library was not feasible (Farley-Barnes et al. 2018; Ogawa et al. 2021), Lisa Ogawa and I honed in on the top hits from both screens (N=702) and performed a targeted rescreen on these alone. By rescreening in quadruplicate and utilizing siON-TARGET (Dharmacon, Horizon) technology to reduce off-target effects, our targeted rescreen produced a more condensed list of higher confidence hits for individual follow-up analysis.

More precise, accessible methods for the study of nucleolar function will be critical for illuminating novel ribosome biogenesis regulators and next-generation therapeutics for human disease states including cancer, aging, and rare ribosomopathies. Here, Carson Bryant developed an HT-ready, image-based assay that selectively measures nucleolar rRNA biogenesis in MCF10A breast epithelial cells. Building upon previous HTS techniques, he combined FBL staining of nucleoli and 5-EU incorporation into nascent RNA to measure only the 5-EU signal corresponding to nucleoli. He optimized the parameters of this assay using both small molecule inhibition (BMH-21) and acute siRNA depletion of essential RNAP1 transcription machinery (POLR1A and NOL11). His detailed assay framework can be applied to studies of novel RNAP1 drug inhibitors and cellular regulators of nucleolar rRNA biogenesis, with the potential for adaptation to a variety of cell types. His assay will increase the dimensionality and efficiency of future HTS campaigns focused on the nucleolus, accelerating the discovery of novel modulators of nucleolar function.

After Carson Bryant optimized the 5-EU assay for a miniaturized format, I validated its utility on 68 known RB factors including core RNAP1 components, small (pre-40S) or large (pre-60S) ribosomal subunit-specific processing and assembly factors, pre-rRNA modification factors, and RPs. As expected, all RB factors had a percent inhibition value greater than 0%. While a wide range of percent inhibition values were observed, 58/68 factors (85.5%) had a mean percent inhibition of at least 50%, signaling that the 5-EU assay robustly reports depletion conditions that interrupt nucleolar rRNA biogenesis.

Although our nucleolar 5-EU assay accurately reported the interruption of nucleolar rRNA biogenesis for the vast majority of RB factors studied, we note the following considerations and caveats regarding our method and results. First, nucleolar rRNA biogenesis can be affected by changes in one or more RB subprocesses including pre-rRNA transcription, processing, modification, and binding by RPs, which all occur co-geographically within the nucleolus. Since kinetic studies have defined the rates of human pre-rRNA transcription (Jackson et al. 2000) and initial pre-rRNA processing steps (Popov et al. 2013) to be on the order of minutes, 5-EU label will be distributed across a population of partially-processed or folded nucleolar pre-rRNA intermediates at the end of the assay’s 1 h labeling period. Therefore, nucleolar 5-EU incorporation over the course of 1 h cannot report solely on RNAP1 transcriptional activity, and additional mechanistic assays may be necessary to precisely define how an experimental treatment alters RB following the observation of a 5-EU defect. We highlight the importance of our discovery of the expanded ability of the 5-EU assay to report on defects in multiple RB steps in addition to RNAP1 transcription, which to our knowledge has not been previously considered. Second, a treatment, like 72 h siRNA-mediated depletion of cultured human cells as we have done here, may also have opposing, compensatory effects on multiple RB subprocesses, leading to an artificially low percent inhibition and a false negative result. More broadly, as with any HTS study using RNAi-mediated target depletion, off-target effects or inefficient on-target depletion could lead to false positive or false negative results, respectively (Echeverri et al. 2006; Ou et al. 2012). Third, we have empirically defined a percent inhibition significance cutoff of 50% inhibition because it minimizes the number of incorrectly classified RB factors. However, it is still unclear if there is a more stringent percent inhibition cutoff that would correspond strictly to RB factors regulating RNAP1 transcriptional activity, or cutoffs for other RB subprocesses. Future studies may elucidate the relationship between the roles of a given RB factor and the nucleolar rRNA biogenesis percent inhibition value observed upon its depletion.

The miniaturized 5-EU assay enables direct quantification of nucleolar rRNA biogenesis in high-throughput, providing clearer insight into how targets modulate RB and improving upon previous HTS techniques for studying nucleolar function. The 5-EU assay is also compatible with our previously published assay for nucleolar number (Farley-Barnes et al. 2018; Ogawa et al. 2021), and is likely to be compatible with other high-content assays for ribosome biogenesis that monitor nucleolar architecture by co-staining for nucleolar proteins (He et al. 2018; Stamatopoulou et al. 2018). By extending the dimensionality and specificity of current state-of-the-art assays which indirectly track nucleolar function, the 5-EU assay will permit researchers to focus on the most promising screen candidates earlier, thereby increasing the efficiency of RB-directed screening campaigns. Carson Bryant and I anticipate that the miniaturized 5-EU assay will expedite the identification and definition of novel regulators of RB in basic or translational studies of nucleolar function. Moreover, this assay has provided results for other nucleolar number rescreen hits being followed up currently or potentially in the future by myself (Chapters 4-6) and other laboratory members.

Material and Methods

Cell lines and culture conditions

Human MCF10A breast epithelial cells (ATCC CRL-10317) were cultured in DMEM/F-12 (Gibco 11330032) with 5% horse serum (Gibco 16050122), 10 µg/mL insulin (MilliporeSigma I1882), 0.5 µg/mL hydrocortisone (MilliporeSigma H0135), 20 ng/mL epidermal growth factor (Peprotech AF-100-15), and 100 ng/mL cholera toxin (MilliporeSigma C8052). Cells were incubated at 37 °C in a humidified atmosphere with 5% CO2.

RNAi depletion by reverse-transfection

RNAi depletion was conducted in MCF10A cells as previously reported (Farley-Barnes et al. 2018; Ogawa et al. 2021). MCF10A cells were reverse-transfected into an arrayed 384-well plate library containing small interfering RNA (siRNA) constructs (Horizon Discovery). Assay-ready plates containing 10 µL of 100 nM ON-TARGET siRNAs resuspended in 1X siRNA buffer (Horizon Discovery B-002000-UB-100) were prepared from master library 384-well plates (Horizon Discovery, 0.1 nmol scale) and stored at -80 C. Plates were thawed at room temperature for 30 min and briefly centrifuged at 300 RPM. siRNA controls were freshly diluted in 1X siRNA buffer to 100 nM from a 50 µM frozen stock, and 10 µL of 100 nM control siRNAs were manually pipetted into the assay-ready plates. To each well, 10 µL of a 1:100 (v/v) RNAiMAX:OptiMEM solution was added (Invitrogen 13778-150, Gibco 31985070), after which the plates were briefly centrifuged at 300 RPM and incubated at room temperature for 30 min. MCF10A cells at 70%-80% confluency were trypsinized for 15 min with 0.05% trypsin (Gibco 25300054), resuspended in culture media, counted with a hemacytometer, and diluted in culture medium to a density of 100,000 cells/mL. Thirty µL of cells were dispensed into assay plates using a Multidrop Combi Reagent Dispenser (Thermo Scientific), to achieve a seeding density of 3000 cells/well, a final volume of 50 µL, and a final siRNA concentration of 20 nM. Seeded assay plates were briefly centrifuged at 300 RPM and incubated at 37 °C for 72 h. Ribosome biogenesis factors were screened in triplicate.

Analysis of mRNA knockdown by RT-qPCR

MCF10A cells were seeded at 1×10^5 cells per well in 6-well plates and incubated at 37 °C for 24 h. Cells were reverse transfected with 20 nM siRNA controls using lipofectamine RNAiMAX per manufacturer's instructions for 72 h. RNA was harvested using TRIzol reagent (Life Technologies 15596018) per manufacturer's instructions. RNA used for cDNA synthesis had a minimum A_{260}/A_{230} ratio of 1.7. cDNA was synthesized from 1 µg total input RNA using iScript^{TM} gDNA Clear cDNA Synthesis Kit (BioRad 1725035). qPCR was performed using SYBR Green reagent (BioRad 1725121) and gene-specific primers. Cycling parameters were as follows: initial denaturation 95 ºC for 30 s, 40 cycles 95 ºC for 15 s and 60ºC for 30 s, melt curve analysis 60 ºC to 94.8 ºC in 0.3 ºC increment. Data analysis was completed using the comparative C_T method ($\Delta\Delta C_T$) using *ACTB* mRNA as an internal control.

BMH-21 treatment and 5-ethynyl uridine incorporation

BMH-21 (MilliporeSigma SML1183) was resuspended in DMSO to a working concentration of 50 µM (50X) and stored at -20 C. 5-ethynyl uridine (5-EU, ClickChemistryTools 1261-100) was resuspended in ddH_2O from powder to a working concentration of 50 mM (50X) and stored at -20 C. For BMH-21 treatment, reverse-transfected assay plates were treated 15 min before the end of the 72 h RNAi depletion period. One µL of either DMSO vehicle or of 50 µM BMH-21 was manually added directly to 50 µL medium in the appropriate wells of the assay plates, which were then briefly centrifuged at 300 RPM and incubated for 15 min before 5-EU incorporation and for the remaining 1 h 5-EU treatment period. For 5-EU incorporation into nascent RNA, reverse-transfected assay plates were treated for 1 h after the end of the 72 h RNAi depletion period. One µL of 50 mM 5-EU was manually added directly to 50 µL medium in each well of the assay plates, which were then briefly centrifuged at 300 RPM and incubated for 1 h.

Immunofluorescent staining and click fluorophore labeling

After 5-EU incorporation, cells were gently washed with 30 µL of PBS and fixed with 1% (v/v) paraformaldehyde (Electron Microscopy Sciences 15710-S) diluted in PBS at room temperature for 20 min. Cells were washed twice with 20 µL wash buffer consisting of PBS with 0.05% (v/v) TWEEN 20 (MilliporeSigma P1379), then permeabilized with 20 µL of 0.5% (v/v) Triton X-100 in PBS for 5 min. Cells were washed twice with 20 µL wash buffer and incubated with 20 µL of blocking buffer consisting of 10% (v/v) FBS (MilliporeSigma F0926) diluted in PBS for 1 h at room temperature. FBL primary antibody solution was prepared by diluting supernatant from the 72B9 hybridoma line (Reimer et al. 1987) at 1:500 or 1:250 (v/v) in blocking buffer. After blocking, cells were incubated with 20 µL FBL primary antibody solution for 2 h at room temperature. Cells were washed twice with 20 µL wash buffer and incubated with 20 µL secondary antibody solution, consisting of 1:1000 (v/v) goat anti-mouse AlexaFluor 647 (Invitrogen A-21236) and 3 µg/mL Hoechst 33342 dye in blocking buffer, for 1 h in the dark at room temperature. Immediately before the end of the secondary antibody incubation period, the click reaction cocktail was prepared in PBS by combining 5 µM AFDye 488 azide (ClickChemistryTools 1275-5), 0.5 mg/mL $CuSO_4$ (Acros Organics 197730010), and 20 mg/mL freshly-resuspended sodium ascorbate (Alfa Aesar A15613). Cells were washed twice with 20 µL wash buffer, then treated with 20 µL of click reaction cocktail for 30 min in the dark at room temperature. Cells were washed twice with 20 µL wash buffer, and soaked in 20 µL PBS containing 3 µg/mL Hoechst

33342 dye for 30 min in the dark at room temperature to dissociate excess AFDye 488 azide. Cells were washed twice with 20 µL wash buffer and 40 µL of PBS was added to each well before high-content imaging.

High-content imaging

Stained assay plates were imaged with a GE Healthcare IN Cell Analyzer 2200. Fields of view were acquired at 20X magnification with 2x2 pixel binning (665.63 μm^2) at 16-bit depth using Cy5, DAPI, and FITC channels for FBL, Hoechst, and 5-EU staining, respectively. Laser autofocus was used to automatically determine imaging Z-height. For publication, images were cropped, merged, and labeled with scale bars using ImageJ 1.53i (Schneider et al. 2012).

CellProfiler pipeline and data analysis

Image analysis was conducted with a custom pipeline for CellProfiler 3.1.9 (Carpenter et al. 2006; McQuin et al. 2018). Briefly, nuclei and nucleoli objects were segmented from DAPI and Cy5 channels, respectively, using global two-class Otsu thresholding. Child nucleoli objects were linked to parent nuclei objects using the RelateObjects module. For both object classes, area was measured from DAPI or Cy5 images, and 5-EU intensity was measured from FITC images. Object-level normalized 5-EU intensity metrics including maximum, mean, median, and standard deviation were calculated by CellProfiler. Raw CellProfiler output CSV files including plate metadata were imported into and analyzed with JMP Pro 15.2.0 (SAS Institute). Per-well averages were computed for each 5-EU metric. For each plate, aggregate control well data were used to calculate signal-to-background (S/B) and Z' factor screening statistics. Nucleolar rRNA biogenesis percent inhibition values were calculated for each well as follows:

$$\text{Nucleolar rRNA biogenesis percent inhibition} = \left(1 - \frac{x_i - \bar{x}_{POLR1A}}{\bar{x}_{NT} - \bar{x}_{POLR1A}}\right) * 100\%$$

where x is the average 5-EU metric value over all objects in a well, x_i is the well metric value for a non-control well, $\bar{x}_{NT}$ and $\bar{x}_{POLR1A}$ are averages of all NT or POLR1A control well metric values respectively. Plate-adjusted percent inhibition values were calculated for non-control wells by subtracting the plate's median NT percent inhibition value from each non-control well percent inhibition (Zhang 2011). Nucleolar/nuclear area ratios were calculated for each nucleus by summing the area of all child nucleoli for a given nucleus, then dividing by the area of the specified nucleus. Nucleoli without a valid parent nucleus (parent ID 0) were discarded. Per-well averages were then computed. Optimization data were graphed in JMP. Triplicate data from the ribosome biogenesis factor screen were averaged in JMP and graphed with GraphPad Prism 8 (GraphPad Software).

Chapter 4

Nuances of early pre-ribosomal RNA stability regulation revealed through the study of the yeast Bud21 ortholog, Human Nucleolar Protein 7 (NOL7)

Introduction

Ribosome biogenesis is an essential and conserved process in all living organisms. In eukaryotes these steps begin in the nucleolus with the transcription of the ribosomal DNA (rDNA) by RNA polymerase I (RNAP1) to synthesize the polycistronic pre-rRNA precursor. This pre-rRNA then undergoes a series of modification, processing, and maturation steps while associating with ribosomal proteins to yield the mature small 40S subunit (18S rRNA) and large 60S subunit (28S, 5.8S, and RNAP3-transcribed 5S rRNAs) (Aubert et al. 2018; Bohnsack and Bohnsack 2019). Much work has been completed in baker's yeast, *Saccharomyces cerevisiae* (*S. cerevisiae*), to understand the fundamentals of eukaryotic ribosome biogenesis (Woolford and Baserga 2013). More recent work has highlighted the increased complexities of this process in human cells with new factors, functions, and connections to other cellular processes being uncovered.

The small subunit (SSU) processome is the earliest stable intermediate pre-ribosome complex that forms co-transcriptionally and is marked by the presence of the U3 small nucleolar ribonucleoprotein (U3 snoRNP). Within the SSU processome, there are multiple subcomplexes that are distinguished by their order of assembly (Krogan et al. 2004; Perez-Fernandez et al. 2007; Chaker-Margot et al. 2015). The first subcomplex to form around the pre-rRNA, the transcription U3 associated proteins (t-Utps) or UtpA subcomplex, is required for both transcription and 5'-external transcribed spacer sequence (5'ETS) processing of the primary transcript pre-rRNA. The UtpA subcomplex was first discovered and studied in yeast as a group of essential proteins (Dragon et al. 2002; Gallagher et al. 2004; Krogan et al. 2004). Other Utps, including the UtpB and UtpC subcomplexes, associate with the small subunit processome later and thus are only required for pre-18S processing (Phipps et al. 2011). Recent work has focused on highlighting the role of the SSU processome's components through cryo-EM structures of both yeast (Barandun et al. 2017) and humans (Singh et al. 2021).

To date, most of the human orthologs of the Utps have been identified except for the t-Utp, Utp9, and a non-essential early associating Utp, Bud21 [called Utp16 in (Dragon et al. 2002)] (Figure 1A, top). As a t-Utp, Utp9 coordinates both pre-rRNA transcription and pre-rRNA processing (Gallagher et al. 2004). In part, a direct association with Utp8 is required for its role in pre-18S processing and for an additional role in protein synthesis through direct regulation of mature tRNA nuclear re-export (Eswara et al. 2009; Huang et al. 2010). Bud21 was first discovered as a non-essential protein required for bud site selection in *S. cerevisiae*. Bud21 is among the set of proteins to associate with the 5'ETS but has not been shown to be a part of any of the SSU subcomplexes in yeast (Chaker-Margot et al. 2015). Bud21 is also known to be a Ty1 retrotransposon host factor (Risler et al. 2012). Deletion of Bud21 renders increased hypoxia tolerance (Shah et al. 2011) and improved xylose carbon source utilization (Usher et al. 2011), while Bud21 expression is upregulated upon acetic acid challenge (Cheng et al. 2021). Overall, not much is known about the molecular functions of Utp9 and Bud21 in ribosome biogenesis, even in the well-studied model organism *S. cerevisiae*.

Here, I describe human NOL7 as a likely yeast Bud21 ortholog and establish its function in human cells. NOL7 is required to maintain a normal number of nucleoli in MCF10A cells, an established predictive indicator of proteins that are involved in ribosome biogenesis (Farley-Barnes et al. 2018; Ogawa et al. 2021). More precisely, NOL7 is required for pre-rRNA transcription, early pre-rRNA stability, and pre-SSU rRNA processing, but not for rDNA promoter activity in human cells. Its depletion leads to decreased mature 18S levels and reduced global protein synthesis, and subsequent induction of the nucleolar stress response. My results, as well as others completed in the laboratory before me, present a new role for NOL7 in human ribosome biogenesis that extends beyond Bud21's defined role in yeast to emphasize the increasing complexities of early ribosome biogenesis in human cells.

Initial work on this project was performed by a variety of Baserga Laboratory members. Sam Sondalle made the first observation that NOL7 was a potential t-UTP by examining high-throughput interactome datasets. Sam Sondalle, Lisa Ogawa, and Katie Farley-Barnes produced stable NOL7 si-resistant HeLa cell lines and performed pre-rRNA processing northern blotting experiments. Carson Bryant performed the rDNA promoter activity assay and analyzed high-throghput screening images. Hannah Huang performed NOL7 siRNA knockdown confirmation by qRT-PCR and western blotted for changes in RNAP1 machinery after NOL7 siRNA depletion. I performed the rest of the experimentation, data analysis, and manuscript preparation. This project was a vast collaborative effort. It is currently published as a pre-print on BioRxiv (McCool et al. 2022a) and is under revision at *RNA Biology.*

Results

NOL7 is the likely yeast Bud21 ortholog

I hypothesized that NOL7 might be the human ortholog to either Utp9 or Bud21 (Utp16) due to its nucleolar localization (Zhou et al. 2010) and because of its interactions with components of the SSU processome. I obtained a list of NOL7-associated proteins from high-throughput interactome datasets (Oughtred et al. 2021) and cryoEM structural analysis (Singh et al. 2021) and performed Gene Ontology overrepresentation tests for both biological processes and cellular components using PANTHER ($\text{Log}_2 > 1$, $p < 0.05$) (Mi et al. 2017). NOL7-associated factors' most enriched categories included positive regulation of rRNA processing (GO:2000234) and transcription by RNA polymerase I (GO:0045943) processes, and the t-UTP complex (GO:0034455) and small-subunit processome (GO:0032040) components (Figure 4.1A). NOL7 protein-protein associations are enriched for factors present in the early SSU processome complex, indicative of a t-UTP subcomplex member.

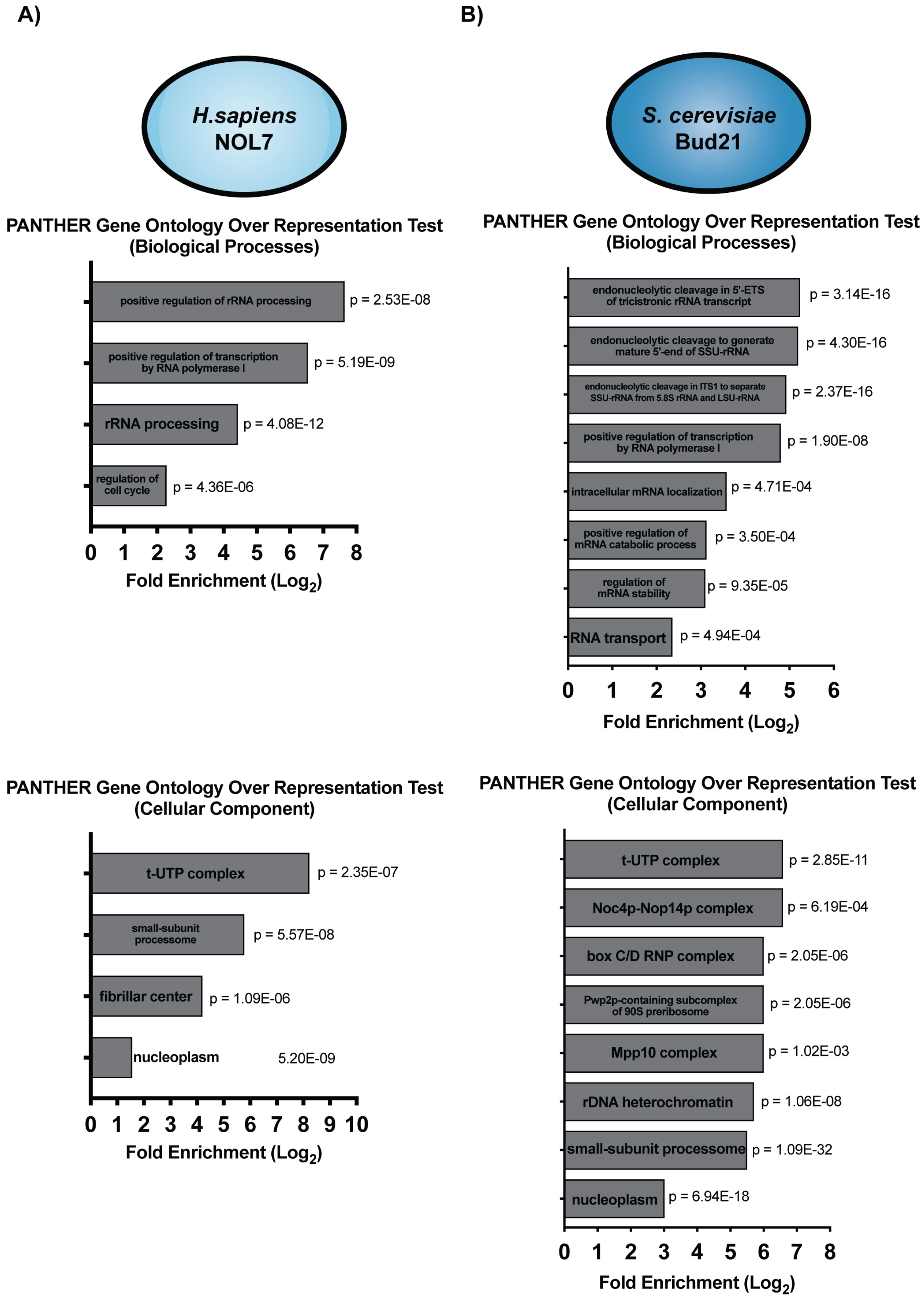
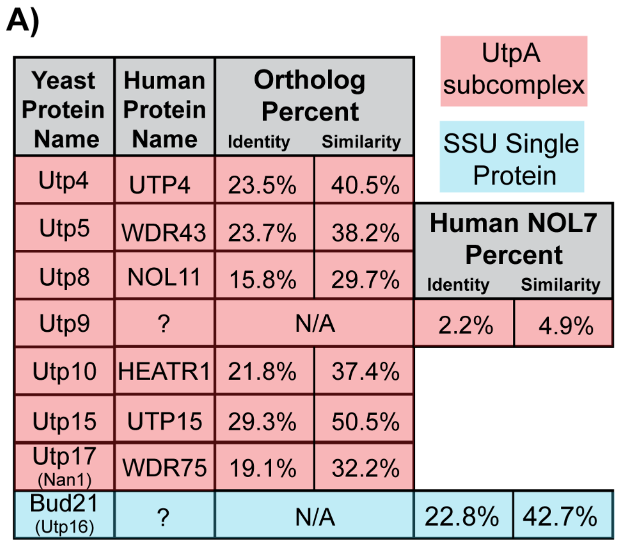

A)
H.sapiens
NOL7
PANTHER Gene Ontology Over Representation Test
(Biological Processes)
positive regulation of rRNA processing
p = 2.53E-08
positive regulation of transcription by RNA polymerase I
p = 5.19E-09
rRNA processing
p = 4.08E-12
regulation of cell cycle
p = 4.36E-06
0 1 2 3 4 5 6 7 8
Fold Enrichment (Log2)
PANTHER Gene Ontology Over Representation Test
(Cellular Component)
t-UTP complex
p = 2.35E-07
small-subunit processome
p = 5.57E-08
fibrillar center
p = 1.09E-06
nucleoplasm
5.20E-09
0 1 2 3 4 5 6 7 8 9 10
Fold Enrichment (Log2)
B)
S. cerevisiae
Bud21
PANTHER Gene Ontology Over Representation Test
(Biological Processes)
endonucleolytic cleavage in 5'-ETS of tricistronic rRNA transcript
p = 3.14E-16
endonucleolytic cleavage to generate mature 5'-end of SSU-rRNA
p = 4.30E-16
endonucleolytic cleavage in ITS1 to separate SSU-rRNA from 5.8S rRNA and LSU-rRNA
p = 2.37E-16
positive regulation of transcription by RNA polymerase I
p = 1.90E-08
intracellular mRNA localization
p = 4.71E-04
positive regulation of mRNA catabolic process
p = 3.50E-04
regulation of mRNA stability
p = 9.35E-05
RNA transport
p = 4.94E-04
0 1 2 3 4 5 6
Fold Enrichment (Log2)
PANTHER Gene Ontology Over Representation Test
(Cellular Component)
t-UTP complex
p = 2.85E-11
Noc4p-Nop14p complex
p = 6.19E-04
box C/D RNP complex
p = 2.05E-06
Pwp2p-containing subcomplex of 90S preribosome
p = 2.05E-06
Mpp10 complex
p = 1.02E-03
rDNA heterochromatin
p = 1.06E-08
small-subunit processome
p = 1.09E-32
nucleoplasm
p = 6.94E-18
0 1 2 3 4 5 6 7 8
Fold Enrichment (Log2)

Figure 4.1: Human NOL7 and yeast Bud21 associate with similar proteins. **(A)** A list of NOL7 interacting proteins was compiled from TheBioGrid (Oughtred et al. 2021) and direct contacts from human pre-A0 cleavage small subunit processome cryo-EM structure pdb: (Singh et al. 2021). A gene ontology overrepresentation test for Biological Processes (top) and Cellular Component (bottom) was completed using PANTHER 17.0. Data were analyzed by Fisher's Exact test, fold enrichment > 2, FDR < 0.05, p < 0.05. **(B)** Same analysis as in (A) except for Bud21 interacting proteins from TheBioGrid and direct contacts from Saccharomyces cerevisiae small subunit processome cryo-EM structure pdb: 5WLC (Barandun et al. 2017) were used.

To address whether Utp9 or Bud21 is the more likely NOL7 yeast ortholog, I performed protein sequence alignments. Interestingly, human NOL7 best aligned in both percent identity and similarity with *S. cerevisiae* Utp16 (22.8% identity, 42.7% similarity) compared to Utp9 (2.2% identity, 4.9% similarity) (Figure 4.2). The level of conservation between NOL7 and Bud21 is similar to that of other UTPs (Figure 4.2), such as human UTP4 with yeast Utp4 (Freed and Baserga 2010) and NOL11 with the single-celled eukaryote species *Capsaspora owczarzaki* Utp8 (Freed et al. 2012). I completed a multiple sequence alignment using CLUSTAL Omega (Sievers et al. 2011) to visualize the conservation of vertebrate NOL7 with Bud21 across various species (Figure 4.3). I observed conservation throughout evolution of the Bud21 sequence. Unexpectedly, I was unable to identify NOL7 orthologs in either *Drosophila melanogaster* [(Hu et al. 2011) DIOPT Version 8.5 2021, https://www.flyrnai.org/cgi-bin/DRSC_orthologs.pl] or in *Caenorhabditis elegans* [(Kim et al. 2018) OrthoList2, http://ortholist.shaye-lab.org/]. Furthermore, the overrepresented Gene Ontology categories of interacting proteins were strikingly similar between Bud21 and NOL7 (Figure 4.1). Based on conservation of both sequence and interaction partners, I concluded that NOL7 is most likely the human ortholog of yeast Bud21.

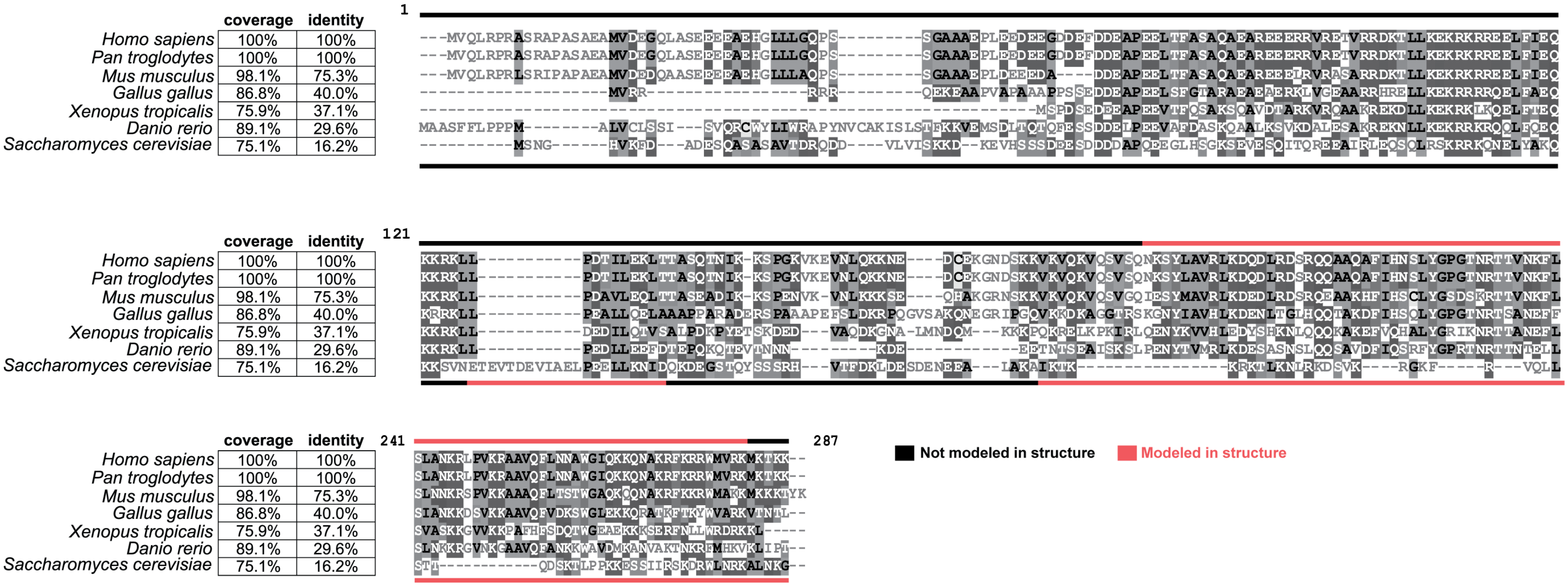


🟥 UtpA subcomplex

🟦 SSU Single Protein

Yeast Protein Name	Human Protein Name	Ortholog Percent		Human NOL7 Percent	
		Identity	Similarity	Identity	Similarity
Utp4	UTP4	23.5%	40.5%		
Utp5	WDR43	23.7%	38.2%		
Utp8	NOL11	15.8%	29.7%		
Utp9	?	N/A		2.2%	4.9%
Utp10	HEATR1	21.8%	37.4%		
Utp15	UTP15	29.3%	50.5%		
Utp17 (Nan1)	WDR75	19.1%	32.2%		
Bud21 (Utp16)	?	N/A		22.8%	42.7%

Figure 4.2: NOL7 is the likely Bud21 (Utp16) ortholog. **(A)** Human NOL7 contains the highest sequence similarity to yeast Bud21 (Utp16). Table of yeast and human t-UTP proteins and their orthologs including the yeast t-Utp interacting protein Bud21. Percent similarity and identity of NOL7 with two yeast Utp members that do not currently have a known human ortholog is indicated.



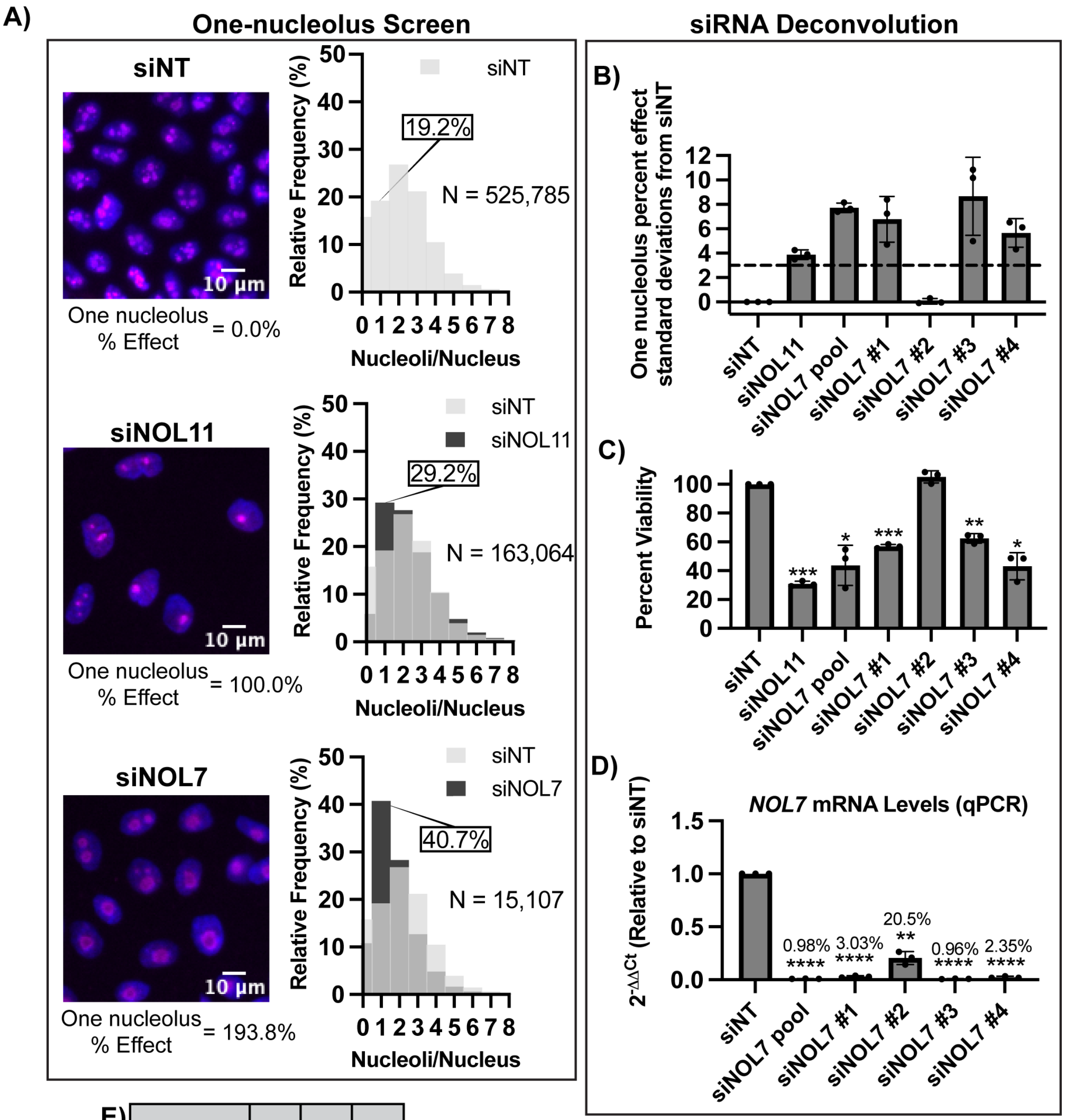
	coverage	identity
Homo sapiens	100%	100%
Pan troglodytes	100%	100%
Mus musculus	98.1%	75.3%
Gallus gallus	86.8%	40.0%
Xenopus tropicalis	75.9%	37.1%
Danio rerio	89.1%	29.6%
Saccharomyces cerevisiae	75.1%	16.2%

Figure 4.3: The Bud21 protein sequence is conserved from Saccharomyces cerevisiae to vertebrate NOL7. Multiple protein sequence alignment of NOL7 in different vertebrate species with S. cerevisiae Bud21 using Clustal Omega (Sievers et al. 2011). Percent coverage and identity values shown in table on left. Consensus disorder (not in modeled in cryo-EM structure, black) and U3 snoRNA associated (modeled in cryo-EM structure, red) indicated with a line either below for Bud21 within yeast (Barandun et al. 2017) or above for human (Singh et al. 2021) small subunit processome structures, respectively.

NOL7 regulates nucleolar function in human cells

NOL7's functional role in human ribosome biogenesis has not yet been firmly established. As a first pass, I took advantage of our laboratory's prior siRNA screening methodology to identify novel regulators of nucleolar function (Farley-Barnes et al. 2018; Ogawa et al. 2021). Briefly, MCF10A cells normally harbor 2-3 nucleoli per cell nucleus. However, upon depletion of ribosome biogenesis factors, this number decreases to 1 or increases to 5 or more nucleoli on average. Since siRNAs against *NOL7* were not present in the original genome-wide siRNA library that the Baserga laboratory had screened and published (Farley-Barnes et al. 2018), I specifically depleted NOL7 using siRNAs (si-ONTARGET pool, Horizon Discovery) in high throughput, applying our workflow pipeline. I observed a significant increase in MCF10A cells with 1 nucleolus (40.7% one-nucleolus cells, percent effect = 193.8%) after NOL7 depletion compared to the negative control non-targeting siRNA (siNT) (19.2% one-nucleolus cells, percent effect = 0%). Notably, NOL7 depletion produced a greater proportion of one-nucleolus harboring cells than depletion of the positive control, the t-UTP, NOL11 (29.2% one nucleolus cells, percent effect = 100%) (Figure 4.4A).

To validate NOL7 as a likely hit, Laura Abriola deconvoluted the si-ONTARGET pool targeting *NOL7* to test the ability of each individual siRNA to produce one-nucleolus containing cells and to mitigate possible off-target effects. Upon analysis of her data, we observed a reduction in nucleolar number using 3 of the 4 individual siRNAs that targeted NOL7, using > 3 standard deviations from siNT negative control as a stringent cutoff (Figure 4.4B). As expected with inhibition of ribosome biogenesis, this one-nucleolus percent effect produced by individual siNOL7 treatments was inversely correlated with cell viability as measured by number of Höechst-stained nuclei remaining after treatment (Figure 4.4C). Moreover, the same 3 siRNAs that produced the one nucleolus phenotype also led to the greatest reduction (> 95%) in *NOL7* mRNA levels as measured by qRT-PCR (Figure 4.4D). Because the other siRNA (siNOL7 #2) did not produce the one-nucleolus phenotype or decrease cell viability, it is possible that the remaining ~20% of normal *NOL7* mRNA levels is enough to maintain normal cellular function (Figure 4.4E). These deconvolution experiments indicate *NOL7* depletion of >95% is required for the reduced nucleolar number and cell viability.



E)

siRNA	One-nucleolus Phenotype?	Viability Decrease?	NOL7 Depletion?
siNOL7 pool N = 15,107	✓	✓	99.0%
siNOL7 #1 N = 14,820	✓	✓	97.0%
siNOL7 #2 N = 27,374	✗	✗	79.5%
siNOL7 #3 N = 16,268	✓	✓	99.0%
siNOL7 #4 N = 11,234	✓	✓	97.7%

Figure 4.4: NOL7 regulates nucleolar function in MCF10A cells. **(A)** NOL7 depletion by an siRNA pool reduces nucleolar number in MCF10A cells. (Left) Representative merged images of nucleoli stained with fibrillarin (magenta) and nuclei stained with Hoechst (blue). siNT was used as a negative control (2-3 nucleoli per nucleus), siNOL11 was used as a positive control (1 nucleolus per nucleus). (Right) Histograms of relative frequency of nucleoli per nucleus from quantification of images. Light grey indicates siNT negative control, dark grey indicates either siNOL11 positive control or siNOL7 treated cells. Overlap with siNT indicated by an intermediate grey color. **(B-D)** Deconvolution of the 4 individual NOL7 siRNAs that constitute the NOL7 siRNA pool in MCF10A cells. **(B)** Testing of the 4 individual NOL7 siRNAs for the one nucleolus percent effect. siNOL7 pool and 3 out of 4 individual siRNAs targeting NOL7 produce an increase in the one-nucleolus percent effect in MCF10A cells. 3 biological replicates, plotted mean ± SD. One-nucleolus percent effect is set relative to standard deviations from the negative control, siNT. siNOL11 is a positive control. A 3 standard deviation cutoff (dashed line) was used to consider an siRNA pool treatment to be a hit and for an individual siRNA treatment to pass the deconvolution criteria of producing the one-nucleolus phenotype. **(C)** Testing of the 4 individual NOL7 siRNAs for cell viability. siNOL7 pool and 3 out of 4 individual siRNAs targeting NOL7 reduce MCF10A cell viability. 3 biological replicates, plotted mean ± SD. Cell viability where siNT, negative control, is set to 100%. siNOL11 is a positive control. Data were analyzed by one-way ANOVA with Dunnett's multiple comparisons test, ** $p \leq 0.01$, *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$. **(D)** Testing of the 4 individual NOL7 siRNAs for NOL7 mRNA levels. siNOL7 pool and 3 out of 4 individual siRNAs targeting NOL7 reduce NOL7 mRNA transcript levels greater than 95% in MCF10A cells. qRT-PCR measuring the primary NOL7 mRNA transcript levels. $2^{-\Delta\Delta Ct}$ were measured relative to 7SL internal control and siNT negative control sample. 3 technical replicates of 3 biological replicates, plotted mean ± SD. Data were analyzed by one-way ANOVA with Dunnett's multiple comparisons test, ** $p \leq 0.01$, **** $p \leq 0.0001$. **(E)** Summary table of siNOL7 pool deconvolution in (B), (C) and (D). Reduction in NOL7 mRNA levels greater than 95% leads to reduction in nucleolar number and cell viability. N = number of cells analyzed sum of 3 independent replicates. One-nucleolus percent effect passes 3 standard deviations from negative control siNT cutoff from panel B. Viability decreased considered if significance was reached from panel C. The percentage of NOL7 mRNA level depletion is calculated from panel D.

To further demonstrate that the results obtained are due to depletion of NOL7, and not due to an off-target effect of siRNAs, previous Baserga laboratory members and I both performed rescue experiments. Previous lab members made a stable HeLa cell line expressing an siRNA-resistant and N-terminally HA-tagged version of NOL7 (resistant to siNOL7 #3 of the siON-TARGET siRNA pool), and I observed that NOL7 mRNA and protein levels depleted by an siRNA can be rescued by overexpressing the si-resistant NOL7. Rescue was detected at both the mRNA level by qRT-PCR and at the protein level by western blotting using an HA antibody (Figure 4.5). HeLa cells were used for these experiments below to highlight the expected universal role of NOL7 in making ribosomes. Additionally, MCF10A is a non-cancerous and a “near-normal” cell line (Soule et al. 1990), thus using HeLa cells allows for a more direct evaluation of whether NOL7 is necessary to drive ribosome biogenesis within the context of cancer.

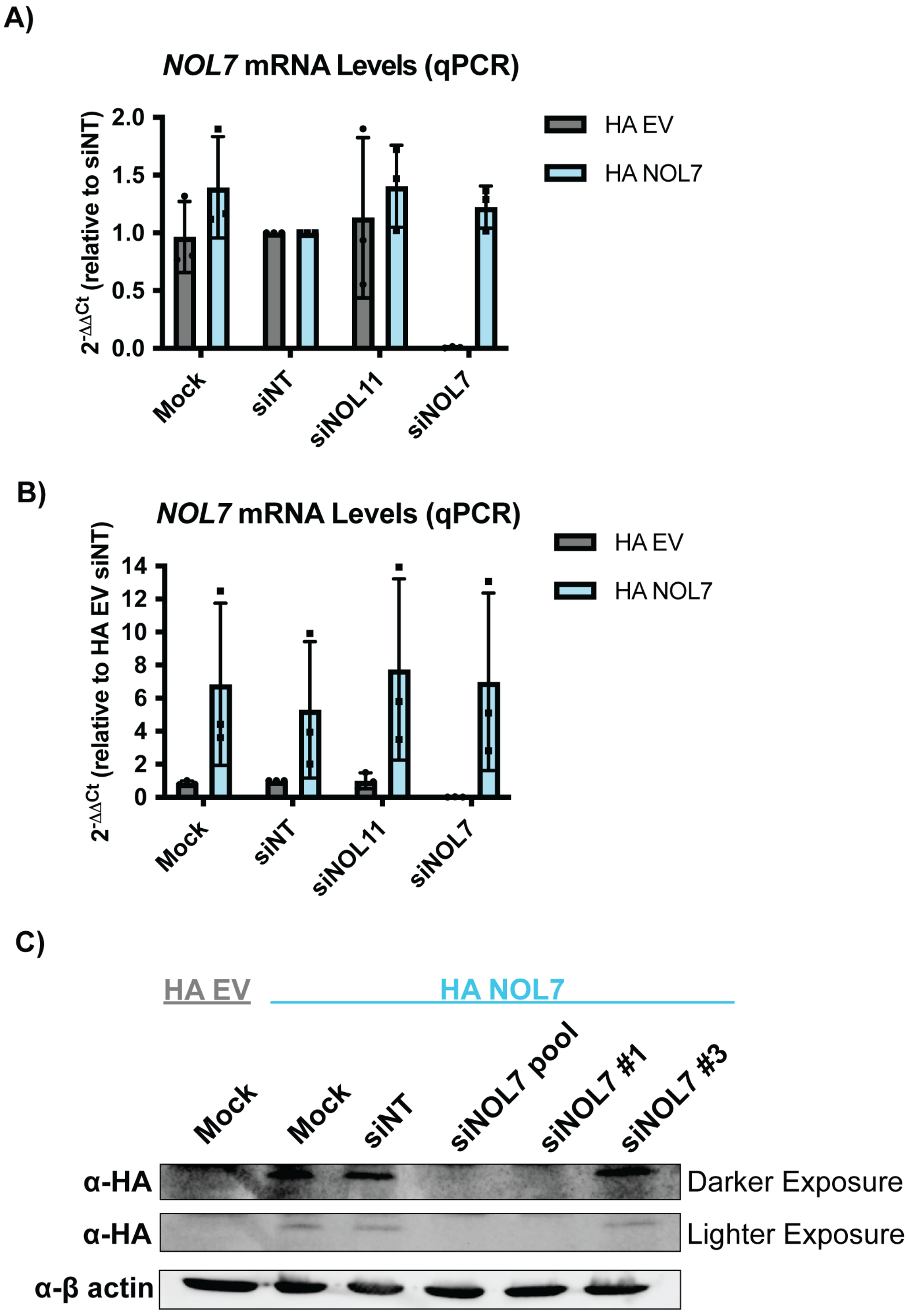


Figure 4.5: NOL7 mRNA and protein levels are rescued upon introduction of an HA-tagged and siRNA resistant version of NOL7 in HeLa cells. **A)** *NOL7* mRNA levels are rescued. qRT-PCR measuring *NOL7* mRNA levels in HeLa cells either expressing empty vector (HA EV) or siNOL7 resistant HA-tagged NOL7 (HA NOL7). $2^{-\Delta\Delta Ct}$ were measured relative to 7SL internal control and siNT negative control sample. siNOL7 treatment is with an individual siNOL7 #3. 3 technical replicates of 3 biological replicates, plotted mean ± SD. **B)** NOL7 mRNA levels are overexpressed in siNOL7 resistant HA tagged NOL7-transduced HeLa cells. Same as in **A)** except all samples were measured relative to siNT HA EV negative control. **C)** NOL7 protein levels are rescued only when upon individual siNOL7 #3 treatment. Western blot using an HA antibody in both HA EV and HA NOL7 transduced HeLa cells. HA EV is a negative control, while Mock and siNT HA NOL7 are positive controls. Darker exposure (top), lighter exposure (bottom), β-actin was used as a loading control.

Because NOL7 has been described as a tumor suppressor in previous studies (Hasina et al. 2006; Zhou et al. 2010; Mankame and Lingen 2012; Doci et al. 2013), I analyzed *NOL7*'s mRNA expression levels in breast and cervical cancer (based on cell lines utilized in this study) in Genotype-Tissue Expression (GTEx) unmatched normal and The Cancer Genome Atlas (TCGA) matched normal and tumor samples (Goldman et al. 2020). At odds with these previous studies that NOL7 is a tumor suppressor, but consistent with its hypothesized role in making ribosomes, *NOL7* mRNA expression is significantly increased in both breast and cervical cancer tissue compared to normal (Figure 4.6). This highlights the relevance of both cell lines used in this study, MCF10A (breast epithelial) and HeLa (cervical cancer) cells. Moreover, I expect NOL7's role in making ribosomes to be conserved to not only these cell lines but in all human cell types. Concordant with these results, a recent study found that NOL7 expression increases and helps drive melanoma proliferation and metastasis (Li et al. 2021). These initial results point towards NOL7 having an important functional role in making ribosomes, and therefore in cell proliferation.

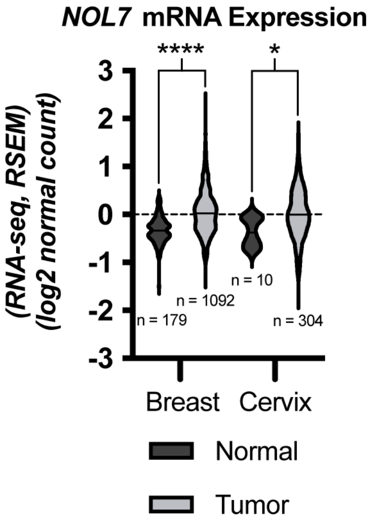


Figure 4.6: Violin plots showing increased expression at the mRNA level for *NOL7* in breast and cervical cancer. Data are from Genotype-Tissue Expression (GTEx) unmatched normal and The Cancer Genome Atlas (TCGA) matched normal and tumor RNA-seq by Expectation-Maximization (RSEM) LOG2 fold expression levels for *NOL7* subtracted from the mean. *NOL7* expression in normal and tumor breast tissue (left), normal and tumor cervical tissue (right). *NOL7* expression in all normal and tumor tissues. Dashed line (set at 0) indicates mean of entire dataset for both normal and tumor expression, black lines indicate mean of individual normal or tumor expression dataset. Data were analyzed by Student's t-test, **** $p \leq 0.0001$, * $p \leq 0.05$.

NOL7 is a component of the SSU processome and necessary for early pre-rRNA stability

Recent cryo-EM structures have offered insight into the location and potential function of Bud21 in yeast (Barandun et al. 2017) and NOL7 in humans (Singh et al. 2021) within the SSU processome (pre-A0 5'ETS cleavage). I summarized the interactions of Bud21 and NOL7 from these structures, showing that both proteins make contacts with the 5' portion of the 5'ETS, the t-UTPs, UTP4 and UTP15, and the U3 snoRNP methyltransferase Nop1/fibrillarin (FBL; Figure 4.7). I hypothesized that these direct interactions would give NOL7 the ability to function in pre-rRNA transcription. Although Bud21 has not been found to be required for rDNA transcription in yeast, it is possible that it was overlooked, since Bud21 is non-essential in yeast (Dragon et al. 2002).

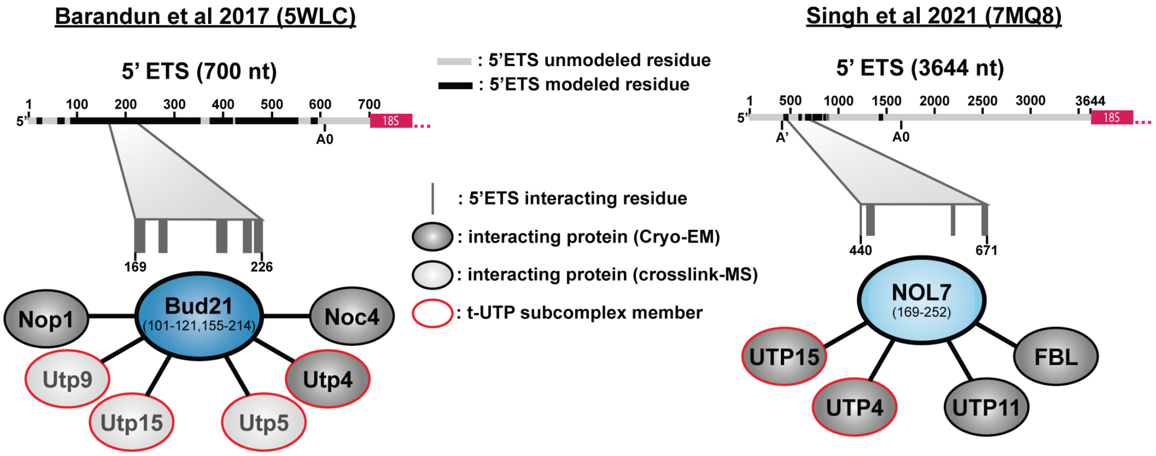
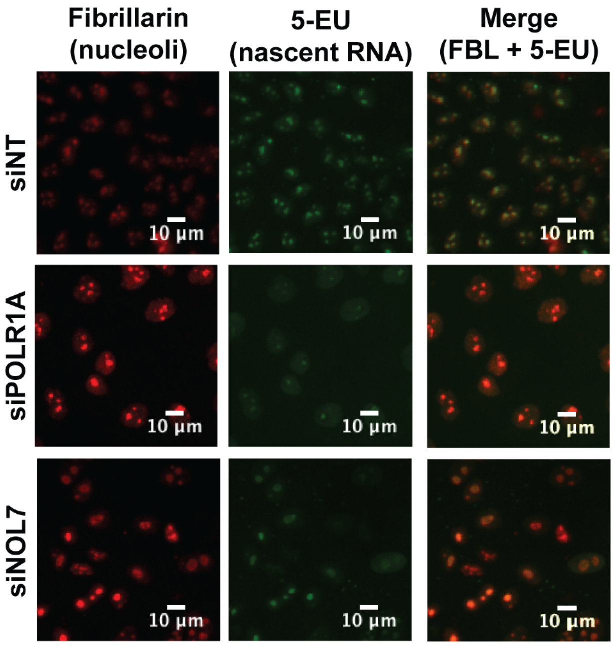


Figure 4.7: Structure interaction summary of Bud21 (Utp16) and NOL7 within the small subunit processome of (Left) *S. cerevisiae* (Barandun et al. 2017) (pdb: 5WLC) and (Right) humans (Singh et al. 2021) (pdb: 7MQ8).

To gain insight into NOL7's potential role in pre-rRNA transcription, I employed an imaging assay developed in our lab to measure nascent nucleolar rRNA biogenesis by 5-ethynyl uridine (5-EU) incorporation and biocompatible click chemistry in MCF10A cells (Bryant et al. 2022). By specifically observing the 5-EU signal residing within the nucleolus, I can measure the amount of nucleolar (pre-)rRNA produced over a one-hour time period as a readout of pre-rRNA transcription and stability. Cells depleted of NOL7 had an inhibition of nucleolar rRNA biogenesis of 76.6%, which is almost to the same extent as the positive control, depletion of the largest subunit of RNAP1, POLR1A (100.0% inhibition; Figure 4.8, Appendix I). The level of inhibition observed is consistent with depletion of previously tested factors that are involved in pre-rRNA transcription and processing, including siNOL11, which exhibited an inhibition of nucleolar rRNA biogenesis of 93.8% in our previous study (Bryant et al. 2022).



siRNA	Nucleolar rRNA Percent Inhibition
siNT	0.0 +/- 17.9%
siPOLR1A	100.0 +/- 5.0%
siNOL7	76.6 +/- 3.1%

Figure 4.8: NOL7 siRNA depletion decreases nucleolar rRNA biogenesis. (Left) MCF10A cells were depleted with siRNAs for 72 h, then fixed cells were stained for nucleoli (fibrillarin, FBL, red) and click chemistry was performed to conjugate AF488 azide to labeled nascent RNA (5-EU, green). siPOLR1A is a positive control and siNT is a negative control. (Right) Table of quantification of the percent inhibition of nucleolar rRNA biogenesis after siNOL7 treatment in MCF10A cells. siPOLR1A positive control treatment set at 100%, siNT negative control set at 0%. 3 biological replicates, mean ± SD.

To more precisely define NOL7's role in pre-rRNA transcription regulation, I measured steady state levels of the primary RNAP1 transcript by qRT-PCR in MCF10A cells depleted of NOL7 (Figure 2D). RNAP1 transcribes the polycistronic 47S pre-rRNA precursor. Cleavage at site A' in the 5'ETS yields the 45S pre-rRNA precursor (Sloan et al. 2014) that is detectable by qRT-PCR (Figure 4.9A) (Woolnough et al. 2016). NOL7 depletion resulted in a significant decrease in 45S pre-rRNA levels compared to siNT, similar to the level of depletion of the positive control, NOL11 (Figure 4.9B). Again, I observed that the decrease in 45S pre-rRNA after NOL7 depletion can be rescued by the expression of an siRNA-resistant version of NOL7, but not an empty vector control in HeLa cells (Figure 4.9C). Taken together, both the nucleolar rRNA biogenesis assay and the 45S qRT-PCR results point toward NOL7 playing a role in the transcription of the pre-rRNA and its early stability prior to 5’ETS processing, suggesting it is required for optimal RNAP1 transcription in human cells.

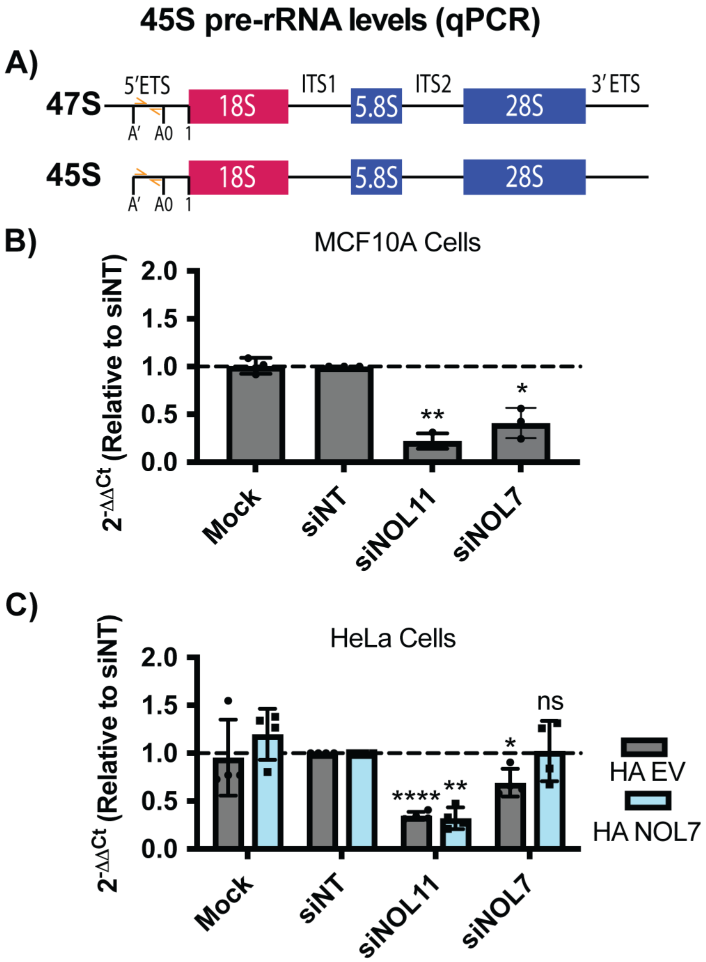


Figure 4.9: qRT-PCR to measure 45S pre-rRNA levels after NOL7 depletion. **(A)** Schematic of 47S primary and 45S pre-rRNA transcripts with qRT-PCR primer locations indicated (orange). **(B)** NOL7 siRNA depletion reduces 45S pre-rRNA levels in MCF10A cells. qRT-PCR measuring the primary 45S pre-rRNA transcript levels. $2^{-\Delta\Delta Ct}$ were measured relative to 7SL internal control and siNT negative control sample. siNOL11 is a positive control. 3 technical replicates of 3 biological replicates, plotted mean ± SD. Data were analyzed by one-way ANOVA with Dunnett’s multiple comparisons test, * $p \leq 0.05$, ** $p \leq 0.01$. **(C)** The reduction of 45S pre-rRNA levels after NOL7 siRNA depletion can be rescued by introduction of an si-resistant version of NOL7. qRT-PCR measuring the primary 45S pre-rRNA transcript in HeLa cells either expressing empty vector (HA EV) or siNOL7 resistant HA-tagged NOL7 (HA NOL7). $2^{-\Delta\Delta Ct}$ were measured relative to 7SL internal control and siNT negative control sample. siNOL11 is a positive control. 3 technical replicates of 4 biological replicates, plotted mean ± SD. Data were analyzed by one-way ANOVA with Dunnett’s multiple comparisons test, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, ns = not significant.

Because NOL7 depletion reduced the steady state levels of early pre-rRNA precursors, we tested if this was a result of reduced rDNA promoter activity in siNOL7 treated MCF10A cells. Carson Bryant used a dual-luciferase rDNA promoter (-410 to +327) activity assay (Ghoshal et al. 2004) to assess pre-rRNA transcription after NOL7 depletion in MCF10A cells. In contrast to the results in 2C and 2D, NOL7 depletion did not reduce rDNA promoter activity compared siNT, while the positive control siNOL11 did (Figure 4.10A). Yale undergraduate, Hannah Huang, also measured protein levels of RPA194 and UBTF by western blot to see if NOL7 depletion reduced their abundance and thus pre-rRNA transcription. However, NOL7 depletion did not change either RPA194 (called POLR1A in Chapter 2), UBTF, or FBL protein levels (Figure 4.10B). NOL7 is thus required to maintain early pre-rRNA precursor levels, but not through regulation of rDNA promoter (-410 to +327) activity in a reporter system.

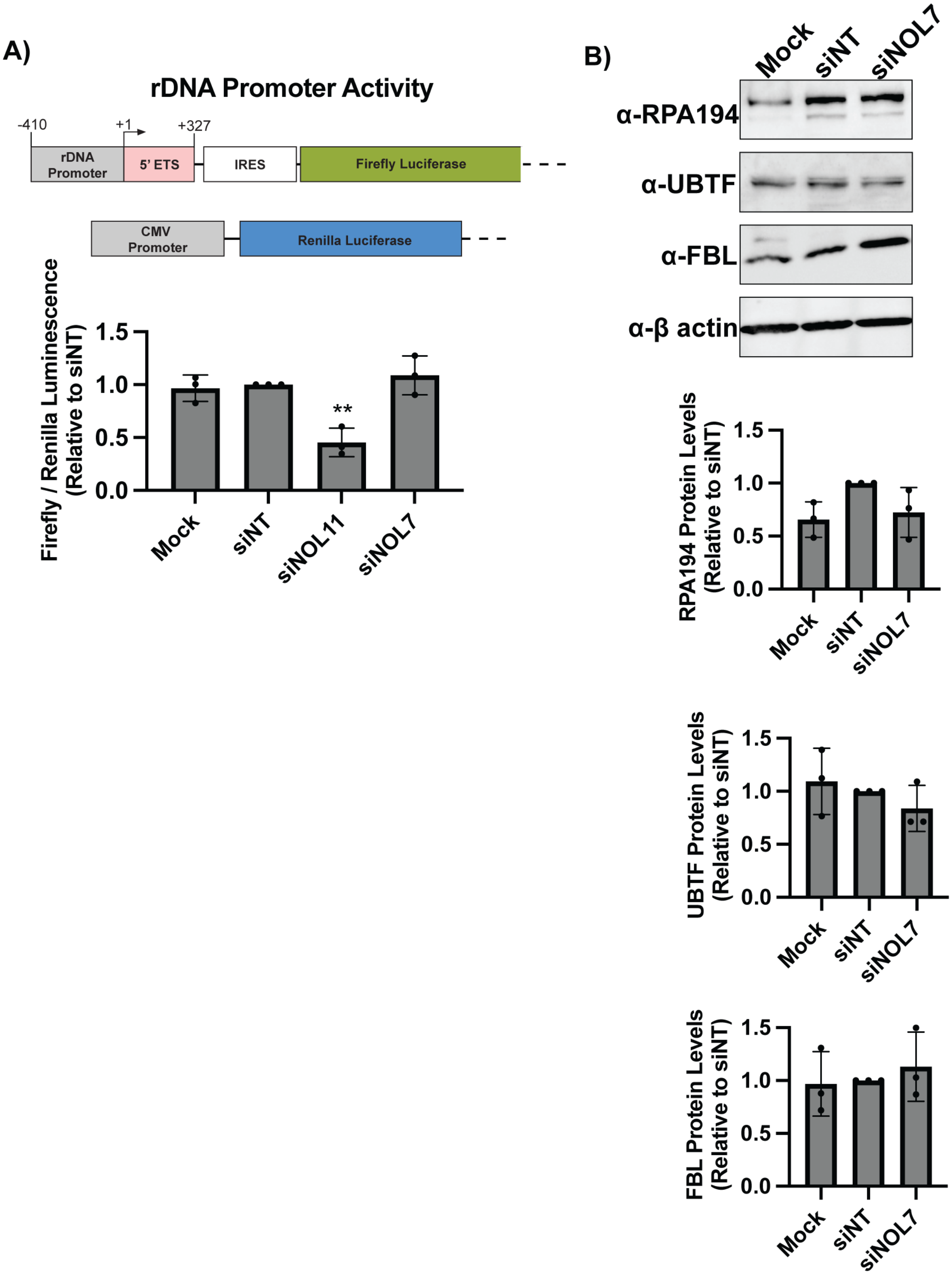


Figure 4.10: NOL7 siRNA depletion does not change rDNA promoter activity or transcription machinery abundance. **(A)** NOL7 siRNA depletion does not change rDNA promoter activity. (Top) Schematic of reporter plasmids used: Firefly (pHrD-IRES-Luc, rDNA promoter reporter) and *Renilla* (CMV, transfection control) luciferase plasmids (Ghoshal et al. 2004). (Bottom) Quantification of rDNA promoter activity. Firefly

luminescence measured relative to *Renilla* luminescence and siNT negative control. siNOL11 is a positive control. 3 technical replicates of 3 biological replicates, plotted mean ± SD. Data were analyzed by one-way ANOVA with Dunnett's multiple comparisons test, ** $p \leq 0.01$. **(B)** NOL7 depletion does not change the protein levels of the RNA Polymerase I transcription machinery (RPA194 and UBTF) or the methyltransferase U3 box snoRNP Fibrillarin (FBL) in MCF10A cells. (Top) Representative western blot for RPA194, UBTF, and FBL where Mock and siNT are negative controls. β-actin was used as a loading control. (Bottom) Quantification of western blots from normalized to β-actin signal. 3 biological replicates, plotted mean ± SD. Data were analyzed by one-way ANOVA with Dunnett's multiple comparisons test.

NOL7 is required for U3 snoRNP mediated 5'ETS pre-rRNA processing to produce the small subunit rRNA

t-UTP's are unique among SSU processome factors as they are required for both the transcription and processing of the pre-18S rRNA (Gallagher et al. 2004; Prieto and McStay 2007; Sloan et al. 2014). Previous Baserga laboratory members and I tested NOL7's role in pre-rRNA processing, specifically of the 5'ETS, which is mediated by the U3 snoRNP (Figure 3A) (Hughes and Ares 1991; Langhendries et al. 2016). Due to NOL7's interaction with the C/D box snoRNP component, FBL, I hypothesized that depletion of NOL7 could impact the stability of the U3 snoRNA and thus its function in 5'ETS processing. I measured U3 and U8 snoRNA levels by qRT-PCR in MCF10A cells following NOL7 siRNA depletion. U8 is a metazoan-specific C/D box snoRNA involved in large subunit processing (Peculis and Steitz 1993; Langhendries et al. 2016). I observed a significant decrease in U3 snoRNA levels upon NOL7 depletion, but not upon depletion of NOL11. The small changes in U8 snoRNA levels upon NOL7 depletion were not significant (Figure 4.11). These decreases in U3 snoRNA levels were not associated with any changes in FBL protein levels as measured by western blotting (Figure 4.10). These results suggest a role for NOL7 in maintaining the stability of the U3 snoRNA.

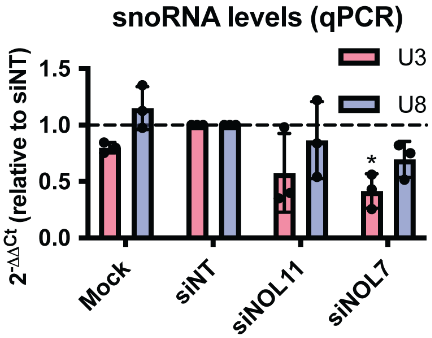


Figure 4.11: NOL7 siRNA depletion in MCF10A cells results in reduced U3 snoRNA steady-state levels. qRT-PCR measuring U3 and U8 snoRNA transcript levels in MCF10A cells. $2^{-\Delta\Delta Ct}$ were measured relative to 7SL internal control and siNT negative control sample. 3 technical replicates of 3 biological replicates, plotted mean ± SD, dotted line at siNT value y = 1. Data were analyzed by one-way ANOVA with Dunnett's multiple comparisons test, * $p \leq 0.05$.

NOL7 was previously identified in a screen for human proteins required for pre-rRNA processing, where its depletion led to a pre-18S processing defect (Tafforeau et al. 2013). To further validate those findings, Lisa Ogawa performed northern blots to assess pre-18S rRNA processing in HeLa cells and I quantified these results using ratio analysis of multiple precursors (RAMP) (Wang et al. 2014) to detect defects within the processing pathway. First, Lisa Ogawa probed ITS1 upstream of cleavage site 2 (probe P3) to measure changes in pre-rRNA intermediates leading to the mature 18S rRNA (Figure 4.12A). Upon depletion of NOL7 and the positive control NOL11, she observed a buildup of 30S pre-rRNA intermediates and a subsequent decrease in 21S pre-rRNA precursors, indicating an inhibition in 5'ETS processing based on my quantification (Figure 4.12B).

Interestingly, in siNOL7 and siNOL11 treated lanes in the northern blots (Figure 4.12B) we noticed an upwards shift in the mobility of the 30S pre-rRNA. It is known that upon depletion of factors required for processing of the 5'ETS, there is a build-up of an aberrant 30S+1 intermediate, with an extended 5' end compared to the 30S intermediate (Figure 4.12A) (Prieto and McStay 2007; Mullineux and Lafontaine 2012). To more specifically detect this defect, Lisa Ogawa used a 5’ETS probe upstream of the A’ cleavage site to measure the levels of the 47S and 30S+1 pre-rRNAs. As expected, she observed a significant increase in the 30S+1 intermediate compared to the upstream 47S primary transcript when both NOL7 and the positive control NOL11 were depleted (Figure 4.12C). This 5’ETS processing defect could be rescued in HeLa cells expressing an siRNA-resistant version of NOL7, but not by an empty vector control (Figure 4.12D). I conclude from these results that NOL7 is required for 5’ETS pre-rRNA processing in human tissue culture cells.

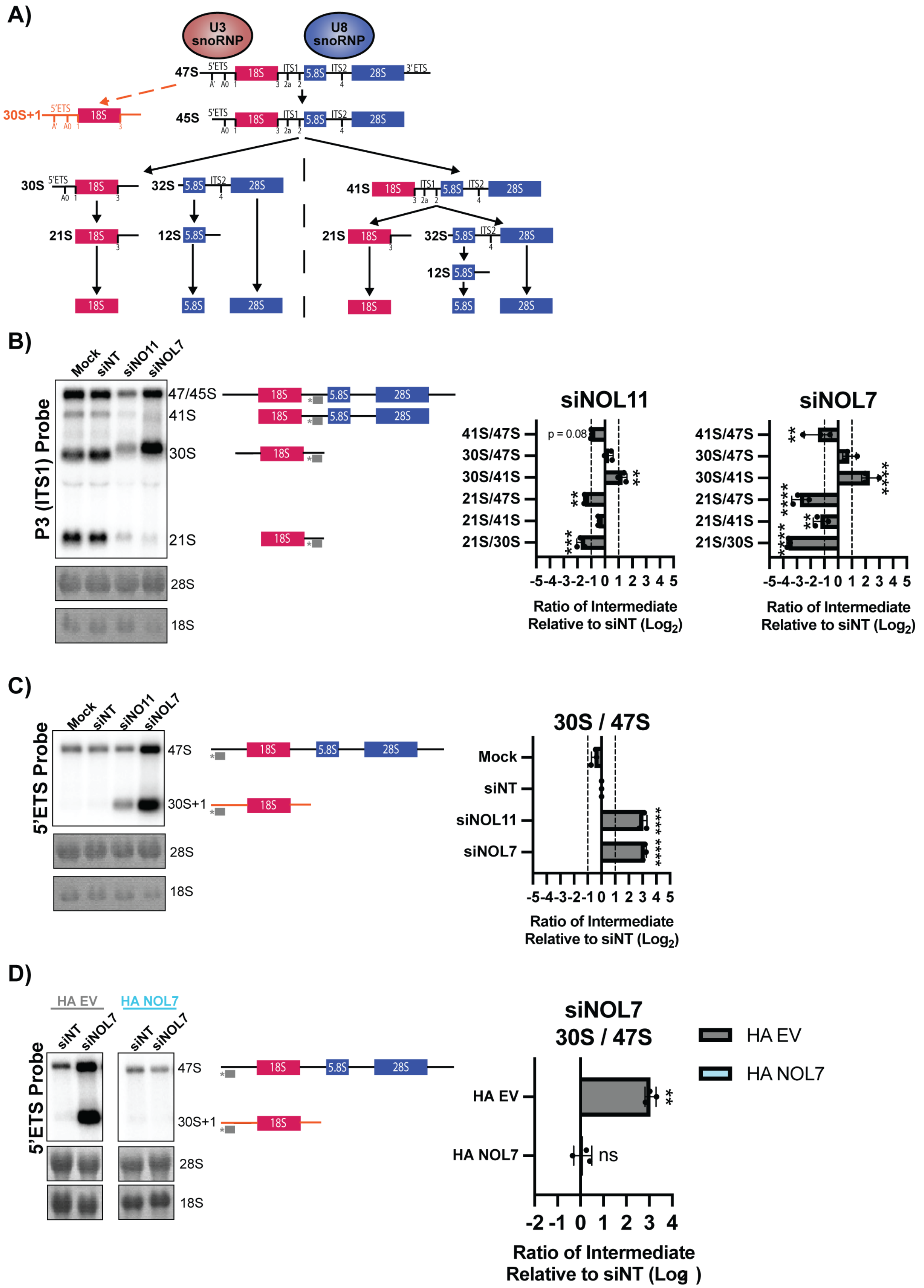


Figure 4.12: NOL7 is required for U3 snoRNA-mediated small subunit pre-rRNA processing in human tissue culture cells. **(A)** pre-rRNA processing in human cells. The 47S pre-rRNA precursor undergoes a series of modification and processing steps to

yield the mature 18S, 5.8S, and 28S rRNAs. Upon inhibition of pre-rRNA cleavage in the 5'ETS, mediated by the U3 snoRNP, accumulation of an aberrant 5'ETS extended 30S+1 precursor occurs (orange). The steps that produce the 18S rRNA (small subunit) are indicated in red, and the steps that produce the 5.8S and 28S rRNAs (large subunit) are indicated in blue. The U8 snoRNA is essential for processing of the rRNAs that make the large ribosomal subunit. **(B)** NOL7 is required for pre-18S rRNA processing in MCF10A cells. (Left) Representative northern blot using a P3 ITS1 probe (indicated in grey) measuring steady state levels of pre-rRNA precursors leading to the 18S rRNA. Methylene blue staining of 28S and 18S was used to show even loading. siNT is a negative control and siNOL11 is a positive control. (Right) Quantification of northern blots using ratio analysis of multiple precursors (RAMP) (Wang et al. 2014) relative to siNT negative control. 3 biological replicates, plotted mean ± SD. Data were analyzed by two-way ANOVA, ** $p \leq 0.01$, *** $p \leq 0.001$, **** ≤ 0.0001. **(C)** NOL7 is required for 5'ETS cleavage in MCF10A cells. (Left) Representative northern blot using a 5'ETS probe (indicated in grey) measuring steady state levels of 47S and 30S+1 pre-rRNA precursors. Methylene blue staining of 28S and 18S was used to show even loading. siNT is a negative control and siNOL11 is a positive control. (Right) Quantification of northern blots using ratio analysis of multiple precursors (RAMP) (Wang et al. 2014) relative to siNT negative control. 3 biological replicates, plotted Log_2 (mean ± SD). Data were analyzed by one-way ANOVA with Dunnett's multiple comparisons test, **** $p \leq 0.0001$. **(D)** Accumulation of the 30S+1 pre-rRNA after siNOL7 depletion can be rescued by introduction of an si-resistant version of NOL7 in Hela cells. (Left) Representative northern blot using a 5'ETS probe measuring steady state levels of 47S and 30S+1 pre-rRNA precursors upon NOL7 depletion and si-resistant rescue with either empty vector (HA EV) or HA-tagged NOL7 (HA NOL7) in HeLa cells. Methylene blue staining of 28S and 18S show even loading. siNT is a negative control. (Right) Quantification of northern blots using ratio analysis of multiple precursors (RAMP) (Wang et al. 2014) relative to siNT negative control. 3 biological replicates, plotted Log_2 (mean ± SD). Data were analyzed by Student's t-test, ** $p \leq 0.05$, ns = not significant.

Since NOL7 is required for pre-18S processing, I tested whether NOL7 depletion would lead to a downstream reduction in the levels of mature 18S rRNA by Agilent BioAnalyzer analysis. There was a significant increase in the 28S/18S rRNA ratio upon NOL7 depletion in MCF10A cells (Figure 4.13A). More specifically, this increased ratio was due to only decreases in 18S levels rather than increases in 28S levels, indicating NOL7's role in the maturation of the 18S rRNA (Figure 4.13B, C). The increase in 28S/18S ratio that was a result of decreases in 18S levels could also be rescued with the introduction of an si-resistant version on NOL7 but not an empty vector control in HeLa cells (Figure 4.13D-F). NOL7 plays a role in the 5'ETS pre-rRNA cleavage that leads to the 18S rRNA, possibly mediated through its interaction with the U3 snoRNP to maintain its function and stability.

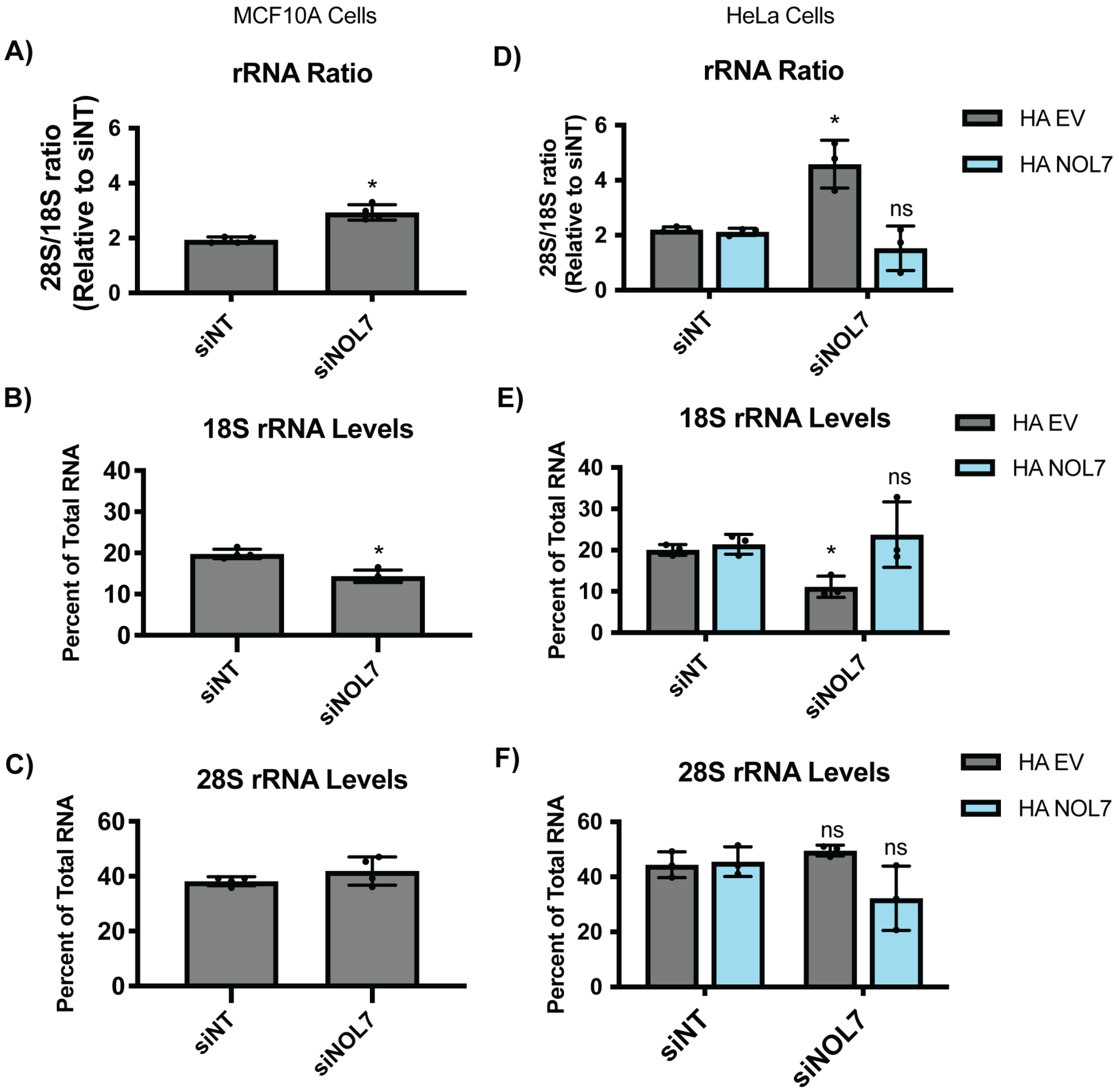


Figure 4.13: NOL7 is required to maintain mature 18S rRNA levels in human tissue culture cells. **(A)** NOL7 is required for production of the a normal 28S/18S mature rRNA ratio in MCF10A cells. Agilent BioAnalyzer analysis of ratio of mature 28S to 18S rRNAs from MCF10A cells depleted of NOL7. 3 biological replicates, plotted mean ± SD. Data were analyzed by Student's t-test relative to siNT negative control, * $p \leq 0.05$. **(B-C)** NOL7 is small subunit (18S) specific in its role in rRNA maturation in MCF10A cells. **(B)** NOL7 is required for the production of the 18S rRNA. Agilent BioAnalyzer analysis of percent of overall RNA levels for 18S rRNA in MCF10A cells depleted of NOL7. 3 biological replicates, plotted mean ± SD. Data were analyzed by Student's t-test relative to siNT negative control, * $p \leq 0.05$. **(C)** NOL7 is not required for the production of the 28S rRNA. Panel as above (B) except that the large subunit (28S) rRNA was measured. **(D)** siNOL7 increase in 28S/18S mature rRNA ratio can be rescued by introduction of an si-resistant version of NOL7 in Hela cells. Agilent BioAnalyzer analysis of ratio of mature 28S to 18S rRNAs upon NOL7 depletion and si-resistant rescue with either empty vector (HA EV) or HA-tagged NOL7 (HA NOL7) in HeLa cells. 3 biological replicates, plotted mean ± SD. Data were analyzed by Student's t-test relative to siNT negative control, * $p \leq 0.05$, ns = not significant. **(E-F)** siNOL7 small subunit (18S) specific role in rRNA

maturation can be rescued by introduction of an si-resistant version of NOL7 in HeLa cells. **(E)** siNOL7 decreases in 18S rRNA levels can be rescued by introduction of an si-resistant version of NOL7. Agilent BioAnalyzer analysis of percent of overall RNA levels upon NOL7 depletion and si-resistant rescue with either empty vector (HA EV) or HA-tagged NOL7 (HA NOL7) in HeLa cells. 3 biological replicates, plotted mean ± SD. Data were analyzed by Student's t-test relative to siNT negative control, * $p \leq 0.05$, ns = not significant. **(F)** siNOL7 and si-resistant version of NOL7 rescue do not change 28S levels. Panel as above in (E) except that large subunit (28S) rRNA was measured.

NOL7 is required for normal levels of protein synthesis and its depletion leads to induction of the nucleolar stress response

I hypothesized that the role for NOL7 in pre-rRNA transcription, early pre-ribosome stability, and pre-18S processing that previous Baserga laboratory members and I observed would lead to downstream effects on ribosome function. Therefore, I performed a puromycin incorporation assay to measure changes in global protein synthesis in MCF10A cells depleted of NOL7 (Schmidt et al. 2009; Farley-Barnes et al. 2018). After siRNA knockdown, cells were treated with 1 μM puromycin, which is incorporated into the nascent polypeptide chain over a time period of 1 h. As expected, depletion of NOL7 led to a significant decrease in global protein synthesis to a similar extent as that of the positive control siNOL11 (Figure 4.14).

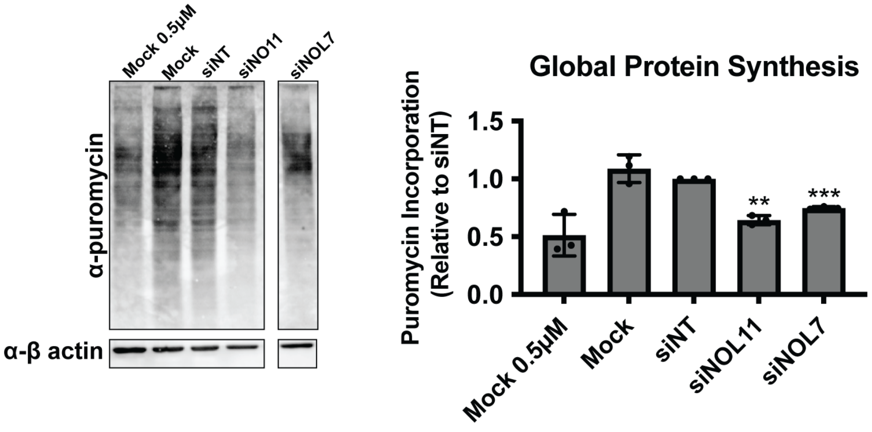


Figure 4.14: NOL7 siRNA depletion reduces global protein synthesis. (Left) MCF10A cells were depleted with siRNAs for 72 h, followed by a 1 h treatment of 1μM puromycin. Total protein was analyzed by western blot. Representative western blot shown where Mock and siNT are negative controls, siNOL11 is a positive control. Mock 0.5μM was included to confirm robustness of the signal quantification. β-actin was used as a loading

control. (Right) Quantification of the decrease in global protein synthesis after NOL7 siRNA depletion in MCF10A cells. Quantification of western blots from (A) normalized to β-actin signal and relative to siNT negative control. 3 biological replicates, plotted mean ± SD. Data were analyzed by one-way ANOVA with Dunnett's multiple comparisons test, ** $p \leq 0.01$.

Because previous Baserga laboratory members and I defined a role of NOL7 in ribosome biogenesis and in maintaining cell viability, I reasoned that depletion of NOL7 would lead to the nucleolar stress response. As a result of impaired ribosome biogenesis, cells can undergo the nucleolar stress response (Rubbi and Milner 2003; Boulon et al. 2010). This leads to changes in nucleolar morphology, induction of TP53 levels, downstream induction of *CDKN1A* (p21) expression, and ultimately cell cycle arrest and apoptosis. To check for the induction of the TP53 mediated nucleolar stress response, I measured TP53 levels in MCF10A cells by western blot after NOL7 depletion. As expected, I observed a significant increase of TP53 levels in NOL7 depleted cells (Figure 4.15A). Additionally, since TP53 is a transcription factor for *CDKN1A*, which plays an important role in inducing cell cycle arrest (el-Deiry et al. 1994), I measured *CDKN1A* mRNA levels by qRT-PCR. After depletion of NOL7 or our positive control NOL11, *CDKN1A* mRNA levels were also strikingly increased (Figure 4.15B). These results are consistent with the observed increase in *CDKN1A* and another cyclin-dependent kinase inhibitor, *CDKN1B* (p27), in melanoma cell lines after NOL7 knockdown (Li et al. 2021).

I expected that these increases in TP53 and *CDKN1A* levels would result in subsequent cell cycle defects after NOL7 depletion in MCF10A cells. Carson Bryant quantified DNA content of both siNT and siNOL7 treated cells by Hoechst staining from images used in the 5-EU nucleolar rRNA biogenesis assay as described in (Chan et al. 2013) which I subsequently analyzed and graphed. Consistent with previous results in a different human cell line (Li et al. 2021), I observed an increase in S-phase cells and an accompanying decrease in sub-G1 cells in NOL7 depleted cells compared to siNT

(Figure 4.15C). Taken together, NOL7 is required for maintaining proper ribosome function and cell cycle progression.

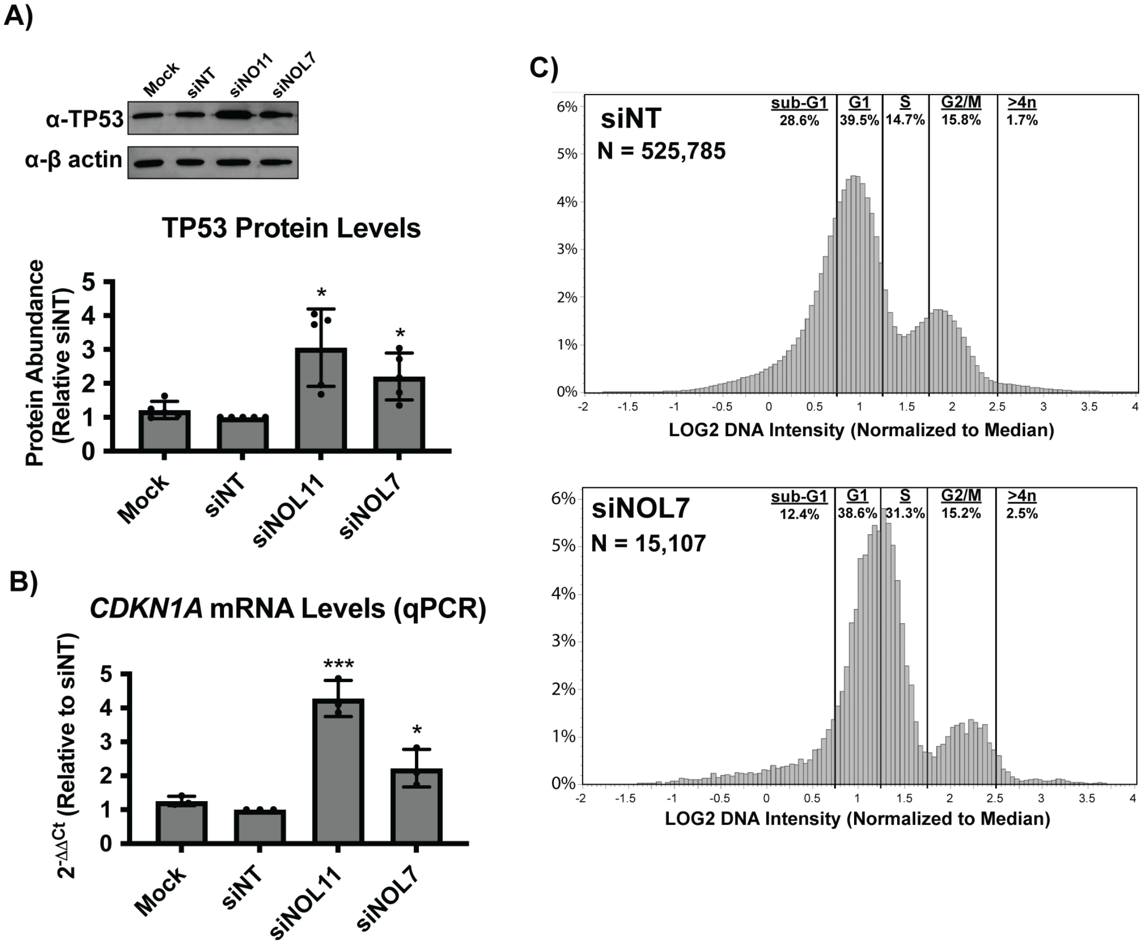


Figure 4.15: NOL7 siRNA depletion causes the nucleolar stress response and cell cycle defects in MCF10A cells. **(A)** NOL7 siRNA depletion leads to increased TP53 protein levels in MCF10A cells. (Top) Representative western blot for TP53 shown where Mock and siNT are negative controls, siNOL11 is a positive control. β-actin is a loading control. (Bottom) Quantification of the increase in TP53 levels after NOL7 siRNA depletion in MCF10A cells. Quantification of western blots from (C) normalized to β-actin signal and relative to siNT negative control. 5 biological replicates, plotted mean ± SD. Data were analyzed by one-way ANOVA with Dunnett's multiple comparisons test, * $p \leq 0.05$. **(B)** NOL7 siRNA depletion significantly increases *CDKN1A* (p21) mRNA levels in MCF10A cells. qRT-PCR measuring *CDKN1A* mRNA levels. $2^{-\Delta\Delta Ct}$ were measured relative to 7SL internal control and siNT negative control sample. siNOL11 is a positive control. 3 technical replicates of 3 biological replicates, plotted mean ± SD. Data were analyzed by one-way ANOVA with Dunnett's multiple comparisons test, *** $p \leq 0.001$, * $p \leq 0.05$. **(C)** NOL7 siRNA depletion inhibits cell cycle progression in MCF10A cells. DNA intensity was measured by Hoechst staining. The Log2 integrated DNA intensities are plotted as a histogram normalized to siNT where G1 peak = 1.0 and G2 peak = 2.0. Cell cycle phases were the following normalized Log2 integrated intensities: sub G-1 < 0.75, G1 = 0.75-1.25, S = 1.25 – 1.75, G2/M = 1.75-2/5, >4n > 2.5. N=the number of cells analyzed.

Discussion

I have shown here that the nucleolar protein, NOL7, plays a critical role in early ribosome biogenesis, specifically in pre-rRNA transcription and pre-18S processing, in human cells. siRNA depletion of NOL7 siRNA leads to a reduction in 47S/45S pre-rRNA levels, inhibition of 5'ETS processing, and reduced mature 18S levels. Each of these endpoints were rescued by the introduction of an si-resistant version of NOL7. These defects lead to a downstream reduction in protein synthesis and to the induction of the nucleolar stress response. These results and interactions with other t-UTPs strongly indicate NOL7's role to be consistent with that of a t-UTP in human ribosome biogenesis.

Some of the roles for NOL7 in human ribosome biogenesis found here were expected based on its orthologous yeast protein, Bud21, and some were not. Previously, the Baserga Laboratory had discovered Bud21 as a member of the SSU processome in yeast, but Bud21 was not previously tested for a role in pre-rRNA transcription in this organism (Dragon et al. 2002; Gallagher et al. 2004). Bud21 is not essential in yeast and its depletion leads to cold-sensitivity (Dragon et al. 2002), a known hallmark of non-essential proteins involved in making ribosomes (Guthrie et al. 1969). In contrast, NOL7 is essential in most human cancer cell lines as found in the DepMap project (https://depmap.org/) by both shRNA and CRISPR genetic perturbation screens across hundreds of cell lines (Tsherniak et al. 2017; McFarland et al. 2018). NOL7's newfound role in ribosome biogenesis is consistent with it being regarded as an essential protein across cancer cell lines. This relevance of NOL7's function in cancer is further emphasized by cancer cells' increased reliance on ribosome biogenesis, leading to new therapeutics being developed to target ribosome biogenesis, specifically RNAP1 (i.e. BMH-21) (Pelletier et al. 2018; Harold et al. 2021). It is possible, given the differences in 5'ETS expansion sequences and in interacting proteins, that NOL7 has gained a more impactful function in pre-rRNA transcription and early pre-rRNA stability compared to yeast Bud21 that underlies its essential nature in humans.

While most of my NOL7 findings are consistent with what is known about the function of other t-UTPs, there are some distinct differences that suggest NOL7's role is not completely aligned with that of all the other t-UTPs. NOL7 is necessary for early pre-rRNA stability and 5'ETS processing that leads to the production of the SSU. More specifically, I discovered that NOL7 is required to maintain 45S pre-rRNA levels and for nucleolar rRNA biogenesis. Each of these assays measures not only pre-rRNA transcription but also pre-rRNA stability. In contrast, in a reporter gene assay for rDNA promoter activity, NOL7 depletion had no effect. Perhaps sequences downstream of the rDNA promoter reporter gene sequences (-410 to +327) are required for NOL7's function (Ghoshal et al. 2004). Consistent with this, only a fraction of the entire 5'ETS (post A' cleavage) is modeled in the human SSU processome structure (Singh et al. 2021), none of which is included in our rDNA reporter gene. Additionally, NOL7 is important for maintaining steady-state levels of the U3 snoRNA, with which it interacts through FBL based on the SSU processome structure (Singh et al. 2021). Although NOL11 depletion did not recapitulate the same results, it remains to be elucidated how other t-UTPs and SSU processome factors alter the stability of early pre-ribosomes and the U3 snoRNA during transcription and the first steps of pre-rRNA processing in human cells.

Our discovery of NOL7's essential role in making ribosomes is at odds with several previous studies on NOL7's role as a tumor suppressor. These studies focused on NOL7's role in post-transcriptional regulation of mRNAs, more precisely mRNA stabilization through direct interactions with mRNAs and their processing machinery within the nucleus (Hasina et al. 2006; Zhou et al. 2010; Mankame and Lingen 2012; Doci et al. 2013). These interactions lead to a tumor suppressor effect due to preferential stabilization of antiangiogenic transcripts, of which TSP-1 has been validated (Doci et al.

2013). It is worth noting that within our PANTHER overrepresentation analysis of Bud21 interacting proteins, but not for the shorter list of known NOL7 interactors, I discovered an enrichment of GO biological processes consistent with this other NOL7 function including: intracellular mRNA localization (GO:0008298), positive regulation of mRNA catabolic process (GO:0061014), regulation of mRNA stability (GO:0043488), and RNA transport (GO:0050658) (Figure S1B). Similarly, there is precedent for other ribosome biogenesis factors having functions in cellular processes outside the nucleolus, including nucleolin and FANCI, in which their roles are regulated by their interacting partners and subcellular localization (Berger et al. 2015; Sondalle et al. 2019). It is unsurprising that the ubiquitous process of making ribosomes has cross-talk mechanisms to coordinate itself with other cellular processes. Our data and Li et al. (2021) strongly suggest NOL7's role to be important for tumor growth and progression, but previous data indicate a role opposing this. It remains to be further explored to what extent and in which contexts NOL7 functions outside of ribosome biogenesis, particularly through modulation of mRNA subsets.

Carson Bryant and I discovered that depletion of NOL7 reduces cell viability and cell cycle progression, most likely as a result of the nucleolar stress response. These results agree with (Li et al. 2021) in which NOL7 depletion increases *CDKN1A* levels and leads to cell cycle S-phase arrest in melanoma cell lines. While the nucleolar stress response occurs as a result of impaired ribosome biogenesis, it is also plausible that NOL7 could act in a more direct manner to modulate the cell cycle. Regulation of cell cycle (GO:0051726) was one of the top overrepresented biological processes categories among NOL7 interacting proteins (Figure S1A). While our data and analysis emphasize NOL7's role in ribosome biogenesis, it remains to be studied if NOL7's interaction with these proteins involved in the cell cycle has a more direct impact on cell cycle regulation and overall cell viability.

NOL7 is another example of the increased complexities of ribosome biogenesis in humans as compared to yeast. We have demonstrated this through NOL7's role in pre-rRNA transcription, processing, and early pre-rRNA and U3 snoRNA stabilities. It is also exemplified by NOL7's other known function outside the nucleolus in regulation of nuclear mRNA stability (Hasina et al. 2006; Zhou et al. 2010; Mankame and Lingen 2012; Doci et al. 2013), adding even more layers of complexity. These results and future studies will certainly help uncover the finer details of human ribosome biogenesis, specifically pre-rRNA transcription and SSU processome function, that will lead to better understanding of therapeutic design options in diseases such as cancer and ribosomopathies.

Materials and Methods

Publicly available expression datasets

Genotype-Tissue Expression (GTEx) unmatched normal and The Cancer Genome Atlas (TCGA) matched normal and tumor expression datasets were obtained through the Xena platform (https://xena.ucsc.edu/) (Goldman et al. 2020). RNA-seq by Expectation-Maximization (RSEM) LOG2 fold expression levels for *NOL7* were subtracted from the mean of the overall normal and tumor tissues combined for graphical visualization.

Sequence alignment and structure analyses

Human NOL7 and *S. cerevisiae* Utp9 and Bud21 protein sequences were used for pairwise alignment using EMBOSS Needle (Needleman and Wunsch 1970) to report percent identity and similarity. Vertebrate protein sequences for NOL7 were used for a multiple sequence alignment along with *S. cerevisiae* Bud21 using Clustal Omega (Sievers et al. 2011).

S. cerevisiae (PDB ID: 5WLC) (Barandun et al. 2017) and human (PDB ID: 7MQ8) (Singh et al. 2021) small subunit processome cryo-EM structures were used to analyze Bud21 and NOL7 protein-protein and 5’ETS protein-RNA interactions respectively. Interface residues buried fraction (>0) was considered an interaction using Biojava-structure-6.0.4 for a given position.

Enrichment analyses

A list of NOL7 interacting proteins was obtained from TheBioGrid (Oughtred et al. 2021), hu.Map 2.0 (Drew et al. 2021), and a cryo-EM structure of the small subunit processome human pre-A0 cleavage (Singh et al. 2021). A list of Bud21 interacting proteins was obtained from TheBioGrid (Oughtred et al. 2021) and a cryo-EM structure of the *S. cerevisiae* small subunit processome (Barandun et al. 2017). These lists were used for PANTHER 17.0 gene ontology overrepresentation analysis (Fisher’s Exact test, reported top main categories where fold enrichment > 2, FDR < 0.05, $p < 0.05$).

Plasmids

Full-length cDNA for human NOL7 was subcloned into Gateway entry vector pDONR221 (Life Technologies) per manufacturer’s instructions. Site-directed mutagenesis to generate synonymous mutations to render NOL7 resistant to siNOL7 #3 (Table S3) was performed using the Change-IT^{TM} Multiple Mutation Directed Mutagenesis Kit (Affymetrix, 78480) per manufacturer’s instructions. NOL7 siRNA-resistant cDNA was subcloned into the pLX301 vector (Addgene, 25895) containing an N-terminal HA epitope tag with an LR reaction (Life Technologies) per manufacturer’s instructions.

Cell culture, media, and lentiviral transduction

MCF10A cells (ATCC, CRL-10317) were subcultured in Dulbecco’s modified Eagles’ medium/Nutrient mixture F-12 (Gibco, 1130-032) containing horse serum (Gibco, 16050), 10 µg/mL insulin (Sigma, I1882), 0.5 µg/mL hydrocortisone (Sigma H0135), 100 ng/mL cholera toxin (Sigma, C8052), and 20 ng/mL epidermal growth factor (Peprotech, AF-100-15). HeLa cells (ATCC, CCL-2) cells were grown in DMEM (Gibco, 41965-062) with 10% fetal bovine serum (FBS, Gibco, 10438026). Cell lines were maintained at 37°C, in a humidified atmosphere with 5% CO_2.

For the high-throughput nucleolar number and 5-EU assays, 3,000 cells/well were reverse transfected in 384-well plates on day 0. For RNA or protein isolation, 100,000 cells/well were seeded into 6-well plates on day 0. For the dual-luciferase assay, 100,000 cells/well were seeded into 12-well plates on day 0.

Lentiviral production and transduction were performed as in (Sondalle et al. 2019). Lentiviruses were packaged in the HEK293FT cell line by co-transfecting pVSV-G, psPAX2, and modified pLX301 vectors (Yang et al. 2011) containing either a siRNA-resistant version of HA-NOL7 CDS or empty vector in a 1:9:10 ratio (1 μg:9 μg:10 μg for a 10-cm plate) with Lipofectamine 2000 (Life Technologies, 11668019) per manufacturer's instructions. HeLa cells were transduced with NOL7 or empty vector containing lentivirus under puromycin antibiotic selection (2 μg/mL).

RNAi

All siRNAs were purchased from Horizon Discovery Biosciences (Table S3). For the 5-EU nucleolar rRNA biogenesis assay of NOL7 ON-TARGETplus pools were used except for the NOL11 positive control where the siGENOME SMARTpool siRNAs were used. For screen validation by deconvolution, the individual ON-TARGET set of four siRNAs that comprised the pool was used. The ON-TARGETplus pools were used in the remaining functional analysis except for northern blotting experiments where siGENOME SMARTpool siRNAs were utilized. Pools of siRNAs were used in all experiments except where indicated that individual siRNAs were tested. siRNA transfection was performed using Lipofectamine RNAiMAX Transfection Reagent (Invitrogen, 13778150) per manufacturer's instructions with a final siRNA concentration of 20 nM for 384-well plate high-throughput assays and 33 nM for all other assays. For the high-throughput screens, cells were reverse transfected in 384-well plates on day 0 where siRNA controls were added to 16 wells each and siNOL7 was added to 1 well for each replicate. For other assays, cells were transfected 24 h after plating.

Nucleolar number assay

Lisa Ogawa followed the (Farley-Barnes et al. 2018) protocol for counting nucleolar number in MCF10A cells in high-throughput. For siRNA pool treatments, including siONT targeting NOL7, a hit was called based on a stringent cutoff if it produced a mean one-nucleolus percent effect greater than or equal to +3 standard deviations (SD) above the mean percent effect for the entire genome-wide screen (122% effect). For siONT deconvolution, an individual siRNA targeting NOL7 was considered validated if it produced a mean one-nucleolus percent effect greater than or equal to +3 SD above the siNT mean, using the siNT SD within each replicate.

5-EU incorporation assay

I monitored nucleolar rRNA biogenesis using a high-throughput protocol Carson Bryant developed in the laboratory (Bryant et al. 2022). Briefly, following 72 h siRNA depletion, MCF10A cells were treated with 1mM 5-ethynyl uridine (5-EU) for 1 h to label nascent RNA. Then cells were fixed, stained for nucleoli [72B9 anti-fibrillarin antibody (Reimer et al. 1987), AlexaFlour 647 goat anti-mouse IGG secondary] and nuclei (Hoëchst), and I performed click chemistry to attach an AF488 azide for 5-EU visualization. Images were acquired with IN Cell 2200 imaging system (GE Healthcare) and a Carson Bryant used a CellProfiler pipeline for analysis (Appendix I).

Cell Cycle Analysis

Carson Bryant and I performed cell cycle analysis as in (Chan et al. 2013; Ogawa et al. 2021). Briefly, we analyzed high-throughput screening images from the 5-EU incorporation assay and measured integrated intensity of Hoechst staining per nucleus. The Log2 integrated intensities of each cell were plotted as a histogram where siNT G1 peak was set at 1.0 and siNT G2 peak set at 2.0. siNOL7 treated cell intensities were normalized to siNT. Cell cycle phases were defined as the following normalized Log2 integrated intensities: sub-G1 < 0.75, G1 = 0.75 – 1.25, S = 1.25 – 1.75, G2/M = 1.75 – 2.5, >4n > 2.5.

Puromycin incorporation Assay

Global protein synthesis was assessed as in (Schmidt et al. 2009). Following 72 h siRNA depletion, 1 μM puromycin (or 0.5 μM puromycin for Mock 0.5 control) was added to the media for 1 h to label nascent polypeptides. Then I proceeded with western blotting as below.

Western blotting

Total protein was harvested from cells by scraping followed by PBS rinse. Cells were lysed using AZ lysis buffer (50 mM Tris pH 7.5, 250 mM NaCl, 1% Igepal, 0.1% SDS, 5 mM EDTA pH 8.0) with protease inhibitors (cOmplete^{TM} Protease Inhibitor Cocktail, Roche, 11697498001) for 15 minutes at 4°C by vortexing. Lysed cells were spun at 21000 RCF for 15 minutes at 4°C, supernatant was harvested and protein concentration was determined by the Bradford assay (Bio-Rad). Either 50 µg or 25 µg (puromycin blots only) of total protein was separated by SDS-PAGE and transferred to a PVDF membrane (Bio-Rad, 1620177) for blotting. The following primary antibodies were used: α-p53 (Santa Cruz, sc-126), α-puromycin (Kerafast, EQ0001), α-HA (clone 12CA5), and α-β-actin (Sigma Aldrich, A1978), α-POLR1A (Santa Cruz sc-48385), α-FBL (Abcam ab226178), α-UBTF (Santa Cruz sc-13125). For detection of primary antibodies, secondary α-mouse-HRP conjugated antibody (GE Healthcare, Life Sciences NA931) was used. Images were acquired using Bio-Rad Chemidoc (12003153) and analyzed using ImageJ software.

qRT-PCR analysis

Total cellular RNA was extracted using TRIzol (Life Technologies, 5596018) per manufacturer's instructions. Prior to cDNA synthesis, all $A_{260/230}$ values were above 1.7

by NanoDrop (ThermoFisher, ND2000CLAPTOP). cDNA was made from 1µg total RNA using iScript gDNA clear cDNA synthesis kit (Bio-Rad, 1725035) with random primers. qPCR was performed with iTaq Universal SYBR Green Supermix (Bio-Rad, 1725121). The following amplification parameters were used: initial denaturation 95 ºC for 30 s, 40 cycles of 95 ºC for 15 s and 60 ºC for 30 s. Subsequent melt curve analysis was performed to ensure a single product, 95 ºC for 15 s, then gradual (0.3 ºC/15 s) increase from 60 ºC to 94.8 ºC. Gene specific primers were used (Table S4). Amplification of 7SL RNA was used as an internal control and relative RNA levels were determined using comparative C_T method ($\Delta\Delta C_T$). Three technical replicates were performed for each biological replicate.

Dual-luciferase reporter assay

After 48 h of siRNA depletion, MCF10A cells were co-transfected with 1000 ng of pHrD-IRES-Luc plasmid (Ghoshal et al. 2004) and 0.1 ng of a CMV-Renilla constitutive internal control plasmid using Lipofectamine 3000 (Thermo Fisher Scientific, L3000015) per manufacturer's instructions. Twenty-four h post plasmid transfection (72 h of siRNA depletion), cells were harvested and luminescence was measured by a Dual-luciferase Reporter Assay System (Promega, E1910) per manufacturer's instructions using a GloMax Discover Microplate Reader (Promega, GM3000). The ratio of pHrD-IRES-Luc / Renilla activity was calculated to control for plasmid transfection efficiency.

Northern blotting

Total cellular RNA was extracted using TRIzol (Life Technologies, 5596018) per manufacturer's instructions. For each sample, 3 µg of total RNA was resolved on a denaturing 1% agarose/1.25% formaldehyde gel using Tri/Tri buffer (Mansour and Pestov 2013) and transferred to a Hybond-XL membrane (GE Healthcare, RPN 303S). UV-crosslinked membranes were stained with methylene blue (0.025% w/v) and imaged. Blots were hybridized to ^{32}P radiolabeled DNA oligonucleotide probes (5'ETS 5'-

CCTCTCCAGCGACAGGTCGCCAGAGGACAGCGTGTCAGC-3’) or (P3 (ITS1) 5’-AAGGGGTCTTTAAACCTCCGCGCCGGAACGCGCTAGGTAC-3’) and detected by phosphorimager (Amersham™ Typhoon™, 29187194). Images were analyzed using ImageJ, ratio-analysis of multiple precursors (RAMP) was performed (Wang et al. 2014).

BioAnalyzer

For each sample, 100 ng/µL total RNA in nuclease-free water was submitted for Agilent BioAnalyzer analysis, performed by the Yale Center for Genome Analysis.

Statistical analyses

Statistical analyses were performed using GraphPad Prism 9.3.1 (GraphPad Software). Tests are described in the associated figure legends.

Chapter 5

Large ribosomal subunit maturation factors, RSL24D1 and the PeBoW complex, associate with RNA Polymerase I to regulate its levels and activity

Introduction

As rDNA transcription is the rate-limiting step of ribosome biosynthesis which underlies cell growth and proliferation, this process is under tight regulation to ensure cellular homeostasis (Laferte et al. 2006; Chedin et al. 2007; Kopp et al. 2007). Pre-rRNA processing and ribosome assembly require the hierarchical association and dissociation of many assembly factors. Some of these factors are required for the formation of both ribosomal subunits, while others are required for the synthesis of one of the two subunits (Bassler and Hurt 2019). Moreover, a group of factors known collectively as the transcription U3 small nucleolar RNA Associated Proteins (t-UTPs), a subcomplex of the SSU processome, are required for the early pre-rRNA transcription steps in eukaryotic cells in addition to their functional role in the processing of the 18S rRNA (Gallagher et al. 2004; Krogan et al. 2004; Prieto and McStay 2007). The t-UTPs associate with the SSU processome to coordinate this action (Krogan et al. 2004). While SSU processome components exhibit dual-functional roles in transcription and processing (Phipps et al. 2011), it remains largely unknown whether there are nucleolar ribosome assembly factors required for LSU biogenesis that play a similar role in RNAP1 transcription.

Among LSU factors, it is well established that the PeBoW (PES1-BOP1-WDR12) complex is required for LSU biogenesis in eukaryotes. It is a conserved complex from yeast (Nop7-Erb1-Ytm1) that is required for ITS2 processing (Holzel et al. 2005). As an extension of this, PeBoW has been shown to be required for cell-cycle progression and cellular proliferation by stimulating ribosome biogenesis (Strezoska et al. 2002; Lapik et al. 2004; Holzel et al. 2005; Grimm et al. 2006). Overexpression of PeBoW proteins has been associated with several cancers (Killian et al. 2006; Fan et al. 2018; Yin et al. 2018; Li et al. 2020; Mi et al. 2021; Vellky et al. 2021; Li et al. 2022). Nonetheless, PeBoW's function in other critical steps of ribosome biogenesis, such as RNAP1 transcription, has not been studied to date.

RSL24D1 is an evolutionarily conserved protein previously studied in yeast with more recent work beginning to elucidate its role in human cells. The *S. cerevisiae* Rlp24 protein (ortholog of mammalian RSL24D1) is required for 27SB pre-rRNA processing to 5.8S and 25S rRNAs via ITS2 cleavage to produce the 60S ribosomal subunit (Saveanu et al. 2003). Rlp24 associates with both early and late stage pre-60S particles in the nucleolus and cytoplasm respectively and is replaced by the homologous ribosomal protein eL24 via the AAA-ATPase Drg1 (Kappel et al. 2012). In humans, RSL24D1 expression in tumor educated platelets has been negatively correlated with early-stage cancer progression and its upregulation has been associated with familial hypercholesterolemia (Li et al. 2015; Ge et al. 2021). Our laboratory's previous RNAi screening campaign identified RSL24D1 as a factor important for the synthesis of ribosomes in human breast epithelial MCF10A cells (Farley-Barnes et al. 2018). More recently, RSL24D1 was also a hit in a screen for factors involved in 60S subunit assembly in HeLa cells (Dorner et al. 2022). However, the molecular mechanisms of how RSL24D1 participates in ribosome assembly in metazoan cells have not yet been fully understood.

This project was undertaken originally by Amber Buhagiar whom I owe deeply for her work originally discovering that RSL24D1 and PeBoW regulate RNAP1 activity. I do not touch on all of her results regarding RSL24D1 and PeBoW here, but all our work is featured in a co-first authored manuscript that was published late last year in *RNA* (McCool et al. 2022b). Amber performed the initial assays establishing RSL24D1's role in humans and results showing RSL24D1 and PeBoW are required for rDNA promoter activity. I then helped confirm these results through the nucleolar rRNA biogenesis assay (Chapter 3) and show interaction between RSL24D1 and WDR12 with RNAP1 which are reported here. Additionally, Lisa Ogawa performed the re-screening, Laura Abriola performed siRNA deconvolution experiments, and Carson Bryant analyzed the high-throughput screening data.

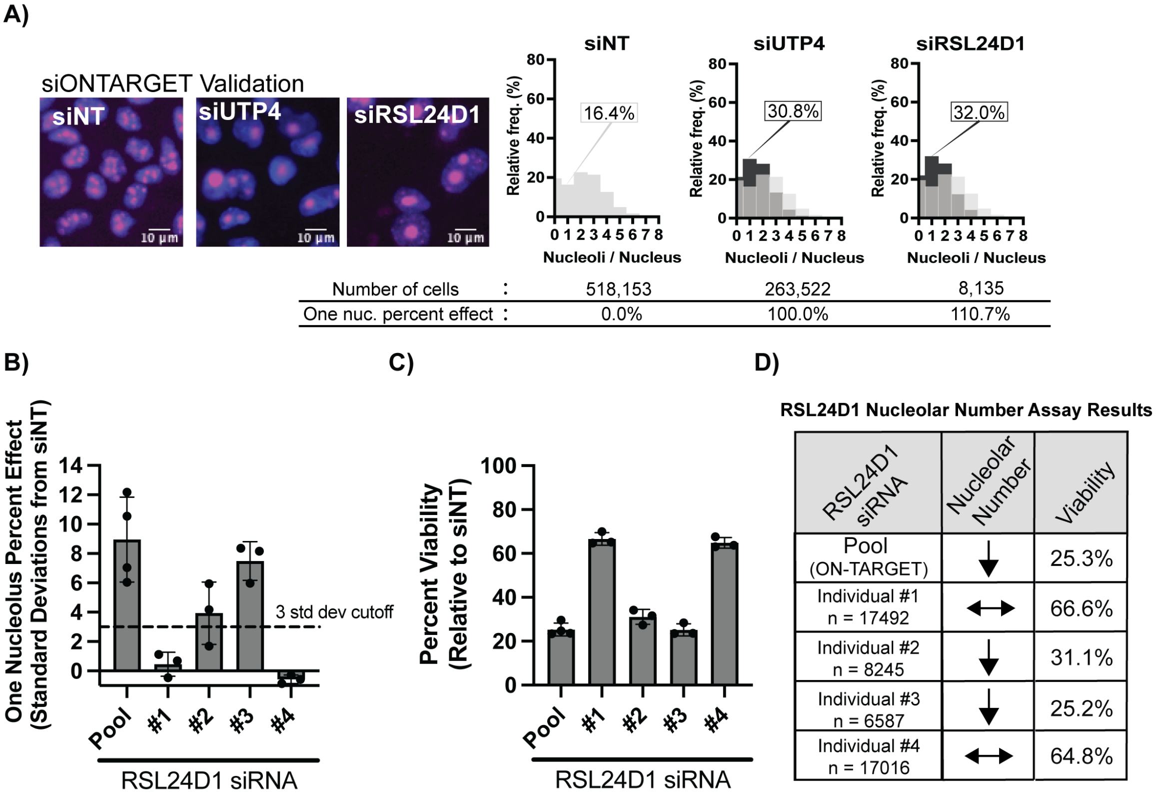
Results

RSL24D1 is required to maintain nucleolar function, indicating yeast Rlp24's general function in conserved in human MCF10A cells

Previously, the Baserga Laboratory performed a genome-wide, phenotypic RNAi screen to uncover novel protein regulators of nucleolar number using a library of siGENOME siRNAs (Horizon Discovery) (Farley-Barnes et al. 2018; Ogawa et al. 2021). RSL24D1 depletion caused a striking decrease in nucleolar number (139.1% effect) relative to the siGFP negative control (0% effect) and siUTP4 positive control (100% effect). This surpassed the extremely stringent screen cutoff used of greater than 3 standard deviations from the mean percent effect for the entire screening population (> 122% effect) (Farley-Barnes et al. 2018). Indeed, siRSL24D1 treatment produced a larger percentage of one-nucleolus harboring MCF10A cells (62.3%) than the positive control siUTP4 (52.4%). These results, consistent with another RNAi screening effort to identify 60S assembly factors (Dorner et al. 2022), strengthened our hypothesis that RSL24D1's role in making ribosomes is conserved from yeast to humans.

To minimize the inclusion of off-target effect induced hits, Lisa Ogawa rescreened RSL24D1 using ON-TARGETplus siRNA reagents (siONT; Horizon Discovery), which are designed to reduce siRNA off-target effects (Jackson et al. 2006). She found that RSL24D1 depletion using a siONT pool reproduces the one-nucleolus phenotype (110.7% effect, 32.0% one-nucleolus cells) relative to the non-targeting (siNT) negative control (0% effect, 16.4% one-nucleolus cells) and the siUTP4 positive control (100% effect, 30.8% one nucleolus-cells) (Figure 5.1). Since this was a targeted validation of our previous siGENOME screen results, I used a cut-off threshold of producing the one-nucleolus percent effect greater than 3 standard deviations from the negative control siNT (> 39.1% effect) for each of the 4 independent screen repetitions.

Laura Abriola at the YCMD performed deconvolution of the RSL24D1 siRNA pool to further validate the role of RSL24D1 in maintaining normal nucleolar number. The individual siRNAs from the siONT pool of 4 siRNAs were evaluated for their effects on nucleolar number and cell viability in RSL24D1-depleted cells. She found that 2 of the 4 individual siRNAs significantly reduced the number of nucleoli from 2-3 to only 1, passing the cut-off criteria of producing an average one-nucleolus percent effect greater than 3 standard deviations from the negative control siNT for 3 independent experiments (Figure 5.1). Additionally, the siONT treatments that produced the one-nucleolus phenotype correlated with a greater reduction in cell viability (Figure 5.1). A correlation between the presence of one nucleolus upon RSL24D1 depletion and cell viability is expected if RSL24D1 is required for ribosome biogenesis, an essential cellular process. Thus, Amber Buhagiar hypothesized that the reduction in nucleolar number upon RSL24D1 depletion was the result of defective ribosome biogenesis.



	siNT	siUTP4	siRSL24D1
Number of cells :	518,153	263,522	8,135
One nuc. percent effect :	0.0%	100.0%	110.7%

RSL24D1 Nucleolar Number Assay Results

RSL24D1 siRNA	Nucleolar Number	Viability
Pool (ON-TARGET)	↓	25.3%
Individual #1 n = 17492	↔	66.6%
Individual #2 n = 8245	↓	31.1%
Individual #3 n = 6587	↓	25.2%
Individual #4 n = 17016	↔	64.8%

Figure 5.1: RSL24D1 depletion using siON-TARGET siRNAs decreases nucleolar number and cell viability. **(A)** RSL24D1 siON-TARGET siRNA depletion reduces nucleolar number in MCF10A cells. (Left) Representative images of nuclei stained with Hoechst (blue) and nucleoli stained with an anti-fibrillarin antibody (pink). siNT was used as a negative control (2-3 nucleoli/nucleus) and siUTP4 was used as a positive control (1 nucleolus/nucleus). (Right) Histograms of the relative frequency of nucleoli per nucleus are shown for the controls and siRSL24D1. There is a decrease in nuclei with 2-3 nucleoli for positive controls and RSL24D1-depleted cells. Light gray bars show the nucleolar number distribution for the siNT negative control, and black bars show the nucleolar number distribution for the indicated positive control or siRSL24D1. Total number of cells analyzed, percentage of cells harboring one nucleolus, and one nucleolus (one nuc.) percent effect for each treatment are indicated. **(B)** siRNA pool and two of four individual RSL24D1 siRNAs (#2 and #3) decrease nucleolar number in MCF10A cells. Oligonucleotide deconvolution of the ON-TARGETplus siRNA pool was used to confirm the activity of individual siRNAs targeting RSL24D1 in the assay for nucleolar number. The graph indicates mean ± standard deviation (SD), the n = 3-4 separate experiments (indicated), and the dotted line is the cutoff of 3 SD from non-targeting (siNT) negative control. **(C)** siRNA pool and 2 of 4 individual RSL24D1 siRNAs (#2 and #3) reduce MCF10A cell viability over 50%. Quantitation of number of cells based on Hoechst staining, negative control (siNT) set at 100%. Graph indicates mean ± standard deviation (SD), n = 3-4 separate experiments. **(D)** Nucleolar number assay and cell viability results correlate for siRNA pool and individual siRNAs targeting RSL24D1. n = total number of cells analyzed for the 3-4 separate experiments completed from (A-C). Down arrow indicates a greater than 3 standard deviations from non-targeting (siNT) negative control one nucleolus percent effect, horizontal arrows indicate less than 3 standard deviations from siNT. Mean percent viability reported from (C).

Supporting RSL24D1's specific role in ribosome biogenesis, it was identified in a previous targeted RNAi depletion screen as a factor potentially involved in LSU processing and assembly in HeLa cells and 60S biogenesis in mouse embryonic stem cells (Durand et al. 2021; Dorner et al. 2022). Moreover, the RSL24D1 yeast ortholog, Rlp24, has been shown to be important for large subunit pre-rRNA processing (Saveanu et al. 2003). Therefore, Amber Buhagiar performed northern blots to measure pre-rRNA processing. Then, Amber Buhagiar and I assessed mature rRNA levels by BioAnalyzer to show that RSL24D1 is required for pre-28S processing and mature 28S levels in MCF10A cells (McCool et al. 2022b). She also confirmed that RSL24D1 is required for global protein synthesis through a puromycin incorporation assay in MCF10A cells. These results indicate that RSL24D1's role in pre-60S maturation is conserved to human cells from yeast.

However, Amber Buhagiar and I wanted to explore further to see if RSL24D1 possesses any additional roles within the ribosome biogenesis pathway. Within this work, we included the study of a predicted RSL24D1 interacting complex of PeBoW (BOP1, PES1, WDR12) by Hu.MAP2.0 (Drew et al. 2020), which is also required for pre-60S biogenesis (Holzel et al. 2005). Here, I report my studies of a new role for RSL24D1 and PeBoW in regulating RNAP1 stability and function.

Confirmation of RSL24D1 and PeBoW siRNA depletion

As an important step to further confirm our observations with siRNAs targeting *RSL24D1*, Amber Buhagiar measured *RSL24D1* mRNA levels, and I measured RSL24D1 protein levels after siONT pool treatment. Both mRNA levels, as measured by qRT-PCR, and protein levels, as measured by western blotting, were greatly reduced, indicating robust knockdown using this method (Figure 5.2A, B). Furthermore, I confirmed PeBoW complex member and Shwachman-Bodian-Diamond Syndrome protein (SBDS) [a negative control in the following experiments since it is an LSU biogenesis factor that acts in downstream maturation steps after nuclear export (Warren 2018))] siRNA depletions using qRT-PCR (Figure 5.2C).

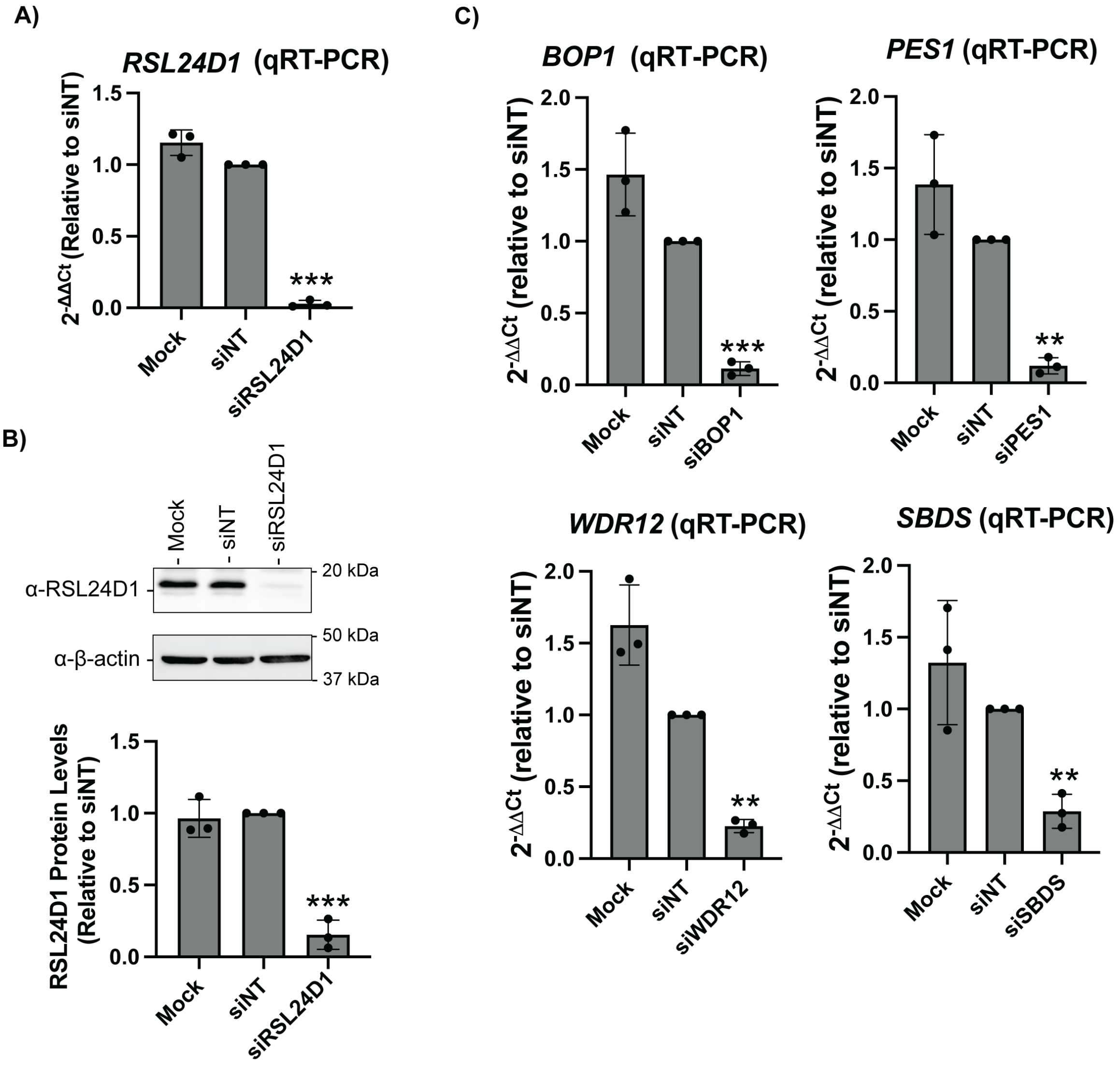


Figure 5.2: Confirmation of siRNA depletion of RSL24D1, PeBoW (BOP1, PES1, WDR12) and SBDS in MCF10A cells. **(A)** qRT-PCR confirmation of *RSL24D1* knockdown in MCF10A cells. $2^{-\Delta\Delta Ct}$ values, relative to a non-targeting (siNT) negative control and 7SL internal control primer, show knockdown of *RSL24D1* by qRT-PCR using the indicated siRNAs. Graph indicates mean ± standard deviation (SD), n = 3 biological replicates. Data were analyzed by Student's t-test followed by multiple testing p-value correction (two-stage linear step-up procedure of Benjamini, Krieger, and Yekutieli) where *** $p \leq 0.001$. **(B)** Western blot confirmation of RSL24D1 knockdown in MCF10A cells. Mock and non-targeting (siNT) siRNAs are shown as negative controls. (Top) Representative western blot images using α-RSL24D1 and α-β-actin antibodies. (Bottom) Quantitation of RSL24D1 levels is reported relative to siNT and normalized to the β-actin loading control. Graph indicates mean ± SD, n = 3 biological replicates. Data were analyzed by one-way ANOVA with Dunnett's multiple comparisons test where ****

p ≤ 0.0001; ns = not significant. **(C)** qRT-PCR confirmation of *BOP1*, *PES1, WDR12*, *SBDS* knockdown in MCF10A cells. $2^{-\Delta\Delta Ct}$ values, relative to a siNT control and 7SL internal control primer, show knockdown of targeted transcripts by qRT-PCR using the indicated siRNAs. Graph indicates mean ± SD, n = 3 biological replicates. Data were analyzed by Student's t-test followed by multiple testing p-value correction (two-stage linear step-up procedure of Benjamini, Krieger, and Yekutieli) where *** p ≤ 0.001, ** p ≤ 0.01.

A connection of RSL24D1 and PeBoW to cancer

Changes in nucleolar number and morphology have been connected with cancer pathology and prognosis for over one-hundred years (Derenzini et al. 2017). Based on our result that RSL24D1 is necessary for maintenance of nucleolar number, I examined *RSL24D1*'s mRNA expression levels in Genotype-Tissue Expression (GTEx) unmatched normal and The Cancer Genome Atlas (TCGA) matched normal and tumor samples (Goldman et al. 2020). Although *RSL24D1* expression is negatively correlated with cancer progression in tumor educated platelets (Ge et al. 2021), this is not what I observed for *RSL24D1*'s expression in tumor compared to normal tissue. *RSL24D1* was significantly more highly expressed in all tumor samples, including breast tumor, compared to normal tissue (Figure 5.3). *RSL24D1*'s increased expression in breast cancer specifically highlights the relevance of its expected role in ribosome biogenesis in MCF10A cells.

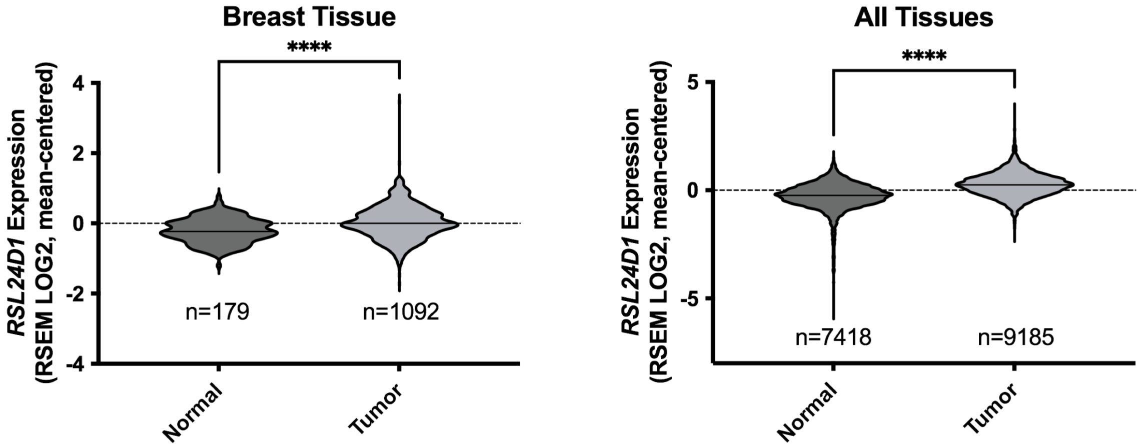


Figure 5.3: *RSL24D1* has increased expression levels in all cancer, including breast cancer. Violin plots of Genotype-Tissue Expression (GTEx) unmatched normal and The

Cancer Genome Atlas (TCGA) matched normal and tumor RNA-seq by Expectation-Maximization (RSEM) LOG2 fold expression levels for *RSL24D1* subtracted from the mean. (Left) *RSL24D1* expression in normal and tumor breast tissue. (Right) *RSL24D1* expression in all normal and tumor tissues. Dashed line (set at 0) indicates mean of entire dataset for both normal and tumor expression, black lines indicate mean of individual normal or tumor expression dataset. Data were analyzed by Student's t-test using GraphPad Prism where **** $p \leq 0.0001$.

RSL24D1 and PeBoW regulate RNA polymerase I levels through their association with RPA194

Previously, Sorino et al. 2020 performed immunoprecipitation of RPA194 (called POLR1A in Chapter 2), the largest subunit of the RNAP1 holoenzyme, followed by mass spectrometry. They identified RSL24D1 and WDR12 of the PeBoW complex as members of the RPA194 interactome (Sorino et al. 2020). These LSU maturation factors were among many confirmed RPA194 interacting proteins detected in their analysis, including NOL11 and UBF (Panov et al. 2006; Freed et al. 2012). To validate whether RSL24D1 and WDR12 interact or are in complex with RPA194, I performed co-immunoprecipitation experiments followed by western blotting. Intriguingly, antibodies against RPA194, but not unconjugated beads, co-immunoprecipitated endogenous RSL24D1 and WDR12 from MCF10A whole cell extracts (Figure 5.4A). However, endogenous PES1 was not co-immunoprecipitated. Treatment of extracts with RNase A did not abrogate these interactions (Figure 5.4A). The reciprocal co-immunoprecipitation with WDR12 also revealed that WDR12 interacts with RPA194 and that this interaction is not abrogated with RNase A treatment (Figure 5.4B). In contrast, WDR12 did not co-immunoprecipitate either RSL24D1 or PES1, a member of the PeBoW complex. While I do not understand the lack of co-immunoprecipitation of WDR12 with its complex partner, PES1, this suggests that there may be a separate pool of WDR12 outside of the PeBoW complex that associates with RNAP1. Furthermore, I have not been able to confirm that RSL24D1 co-immunoprecipitates with PeBoW as predicted by hu.MAP 2.0

(Drew et al. 2020). The RPA194 and WDR12 antibodies used for co-immunoprecipitations and the PES1 antibody used for blotting were tested for specificity by western blotting after siRNA depletion. I showed both antibodies are specific based on an observed decrease in the expected band after siRNA treatment (Figure 5.4C).

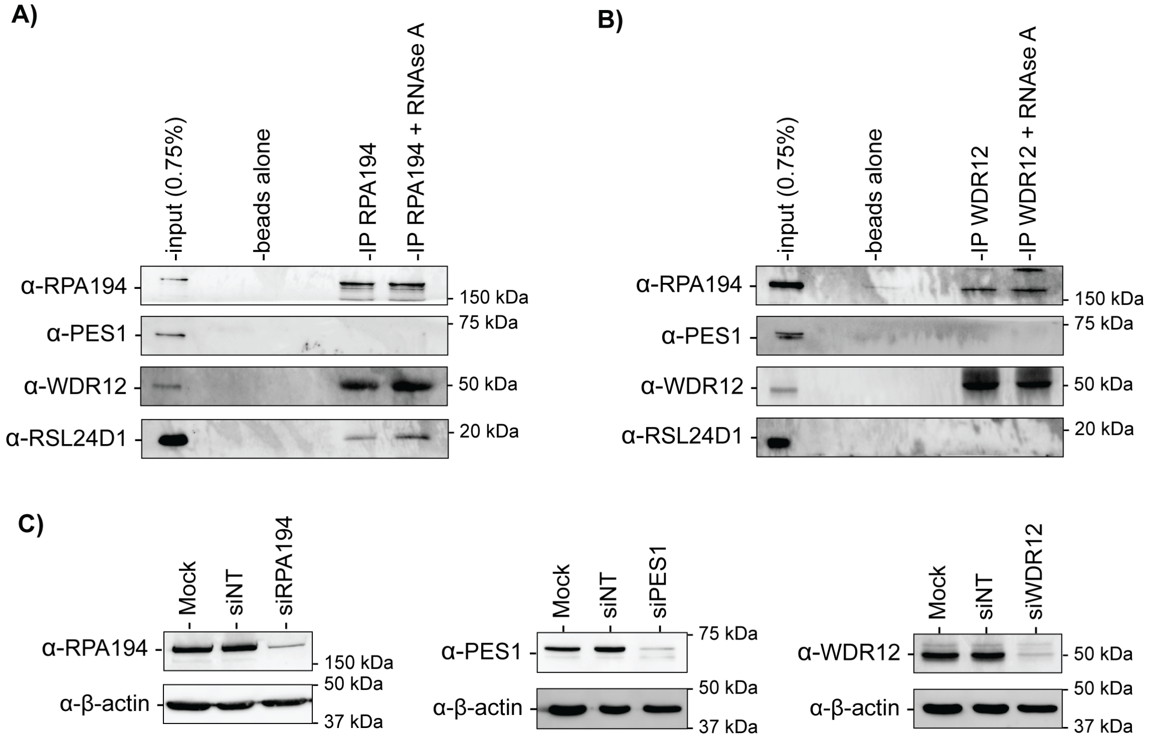


Figure 5.4: RSL24D1 and WDR12 associate with RPA194. **(A)** MCF10A whole cell extracts were immunoprecipitated with α-RPA194 antibody with/without RNase A treatment. Input corresponds to 0.75% of the whole cell extract (WCE) used for immunoprecipitation. Unconjugated protein A beads were used as a negative control. RPA194, PES1, WDR12, and RSL24D1 were detected by western blotting with specific antibodies. Representative western blot images shown are 1 of 2 biological replicates. **(B)** Panel as above, except MCF10A whole cell extracts were immunoprecipitated with α-WDR12 antibodies. **(C)** α-RPA194, α-WDR12 antibodies for immunoprecipitation and α-PES1 antibody for blotting are specific. Western blot of MCF10A cell lysate using α-RPA194 (left), α-PES1 (middle) and α-WDR12 (right) antibodies. α-β-actin was used as an internal loading control. Mock and non-targeting (siNT) were used as negative controls, siRPA194, siPES1, and siWDR12 were used as positive controls to confirm a reduction in levels of their respective expected bands.

This finding led us to ask whether RSL24D1 and PeBoW could also influence ribosome biogenesis at the level of pre-rRNA transcription in addition to their regulatory role in pre-LSU processing steps. First, due to the observed association of RSL24D1 and WDR12 with RPA194, I tested for changes in RPA194 protein levels after RSL24D1 and PeBoW complex member depletion. Mock (no siRNA), siNT, and siSBDS were used as negative controls, while siRPA194 was used as a positive control. siRNA depletion of RSL24D1, WDR12, and the two other PeBoW members, BOP1 and PES1, led to significant decreases in RPA194 levels in MCF10A cells (Figure 5.5). However, depletion of the Shwachman-Bodian-Diamond Syndrome protein (SBDS), an additional negative control, did not have any effects on RPA194 levels. This indicates that it is most likely not a feedback mechanism of impaired later steps of ribosome biogenesis that results in reductions in RPA194 levels.

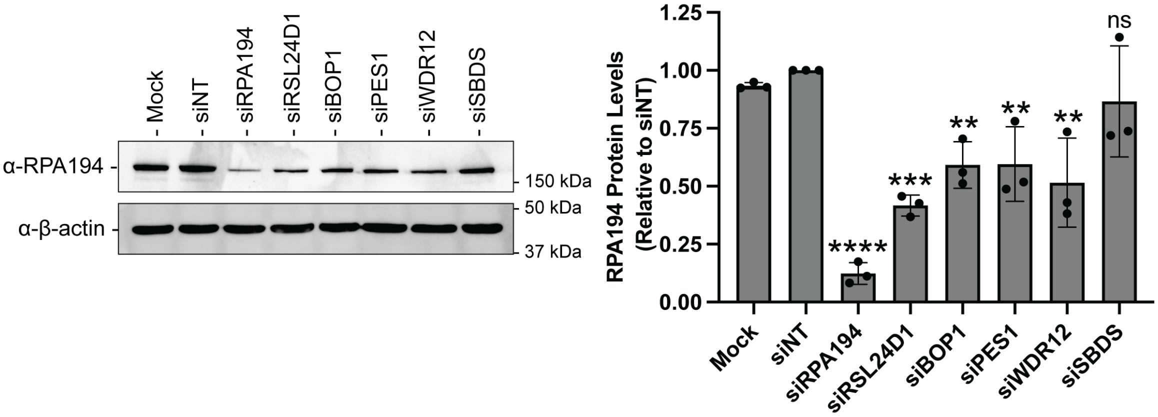
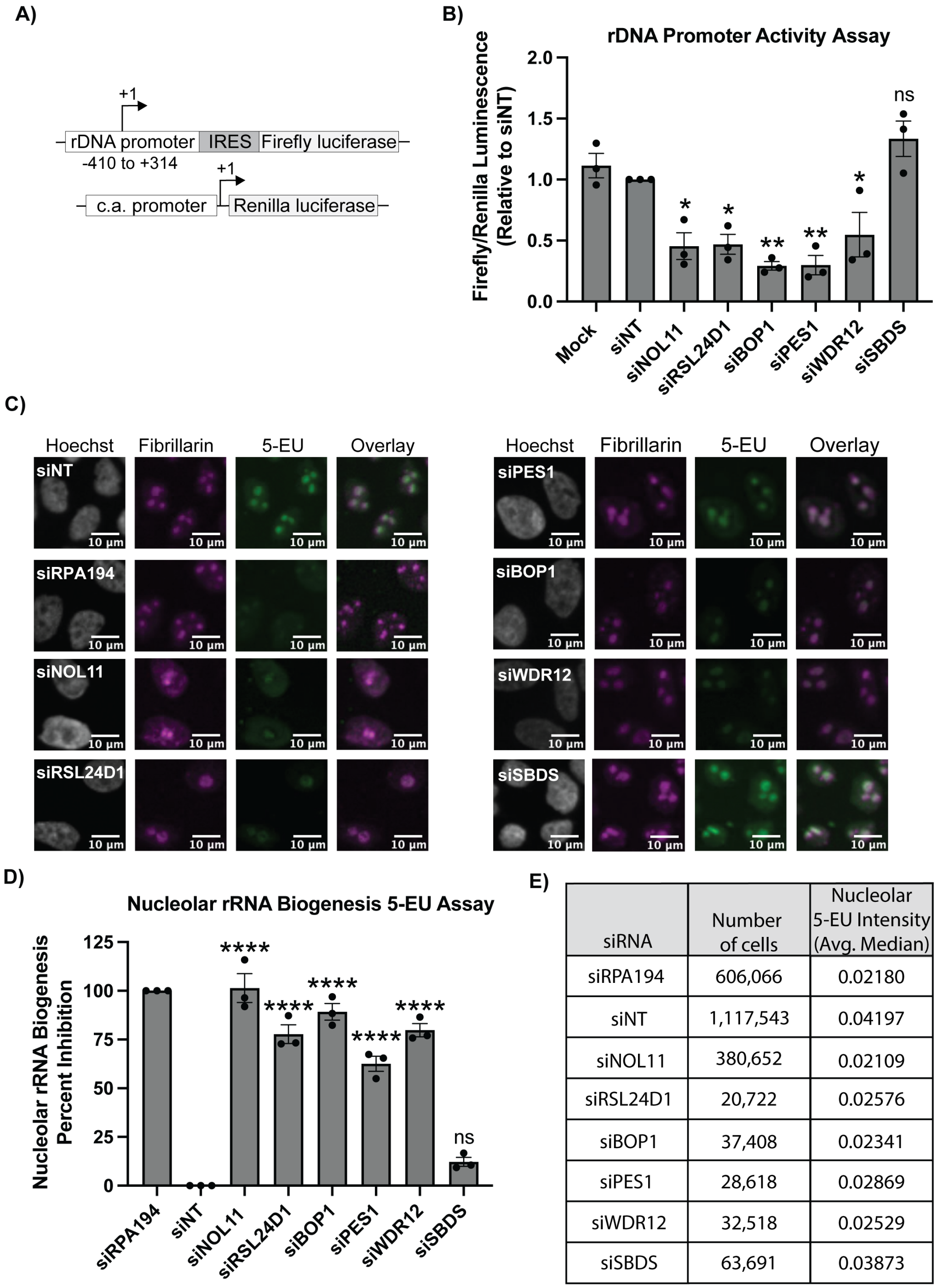


Figure 5.5: Depletion of RSL24D1 and PeBoW reduce RPA194 protein levels in MCF10A cells. Cells were treated with the indicated siRNAs for 72 h and total protein from whole cell extracts was harvested. Mock, non-targeting (siNT), and SBDS siRNAs are shown as negative controls. (Left) Representative western blot images using α-RSL24D1 and α-β-actin antibodies. (Right) Quantitation of RPA194 levels is reported relative to siNT and normalized to the β-actin loading control. Graph indicates mean ± SD, n = 3 biological replicates. Data were analyzed by one-way ANOVA with Dunnett's multiple comparisons test where **** $p \leq 0.0001$; ** $p \leq 0.01$; ns = not significant.

RSL24D1 and the PeBoW complex are required for RNAP1 activity

Since RPA194 is essential for RNAP1 transcription of the pre-rRNA, Amber Buhagiar predicted that RSL24D1 and PeBoW would also be critical for RNAP1 transcription. Amber Buhagiar assessed pre-rRNA transcription by RNAP1 using a dual-luciferase reporter assay. In this system, the pHrD-IRES-Luc plasmid, which contains the firefly luciferase gene under the control of the human rDNA promoter, is co-transfected with a control plasmid containing the Renilla luciferase gene under the control of a constitutively active promoter (Figure 5.6A) (Ghoshal et al. 2004; Farley-Barnes et al. 2018). Amber Buhagiar assayed transcription in RSL24D1-, PES1-, BOP1- or WDR12-siRNA depleted MCF10A cells where mock (no siRNA), siNT and siSBDS were used as negative controls, while siNOL11 was used as a positive control. Remarkably, depletion of RSL24D1 or individual PeBoW members revealed reduced rDNA promoter activity and thus decreased pre-rRNA transcription (Figure 5.6B). I further corroborated this result by *in cellula* pulse labeling with the uridine analog 5-ethynyl uridine (5-EU) and the analysis of its incorporation into nascent nucleolar rRNA through immunofluorescent staining as in (Bryant et al. 2022). Strikingly, cells depleted of RSL24D1 or PeBoW components exhibited a strong reduction in nucleolar rRNA biogenesis relative to siNT and siSBDS negative control cells (Figure 5.6C-E). This is consistent with previous results following depletion of ribosome biogenesis factors required for pre-rRNA processing and transcription (Bryant et al. 2022). RSL24D1 and PeBoW siRNA depletion had a reduction in nucleolar rRNA biogenesis almost to the same extent as siRPA194 and siNOL11 positive controls. Taken together, these results indicate that RSL24D1 and PeBoW modulate rDNA transcription, linking LSU processing factors to RNAP1 transcription regulation.



siRNA	Number of cells	Nucleolar 5-EU Intensity (Avg. Median)
siRPA194	606,066	0.02180
siNT	1,117,543	0.04197
siNOL11	380,652	0.02109
siRSL24D1	20,722	0.02576
siBOP1	37,408	0.02341
siPES1	28,618	0.02869
siWDR12	32,518	0.02529
siSBDS	63,691	0.03873

Figure 5.6: RSL24D1 and PeBoW complex members are required for optimal rDNA transcription. **(A)** Schematic of Firefly (pHrD-IRES-Luc) and Renilla luciferase plasmids for the rDNA promoter activity reporter system (Ghoshal et al. 2004); c.a. promoter: constitutively active promoter. **(B)** Depletion of RSL24D1 or PeBoW complex members leads to reduced rDNA promoter activity. MCF10A cells treated with the indicated siRNAs were transfected with reporter plasmids expressing Firefly and Renilla luciferase. Luminescence was quantified as a ratio in which Firefly gene expression, controlled by the human rDNA promoter, was normalized to Renilla gene expression, controlled by a constitutive promoter, and reported relative to non-targeting siRNA

(siNT). Mock, non-targeting (siNT, and siSBDS were used as negative controls and siNOL11 was used as a positive control. Quantitation of results were reported relative to siNT. Graph indicates mean ± SEM, n = 3 biological replicates. Data were analyzed by one-way ANOVA with Dunnett's multiple comparisons test where ** $p \leq 0.01$, and * $p \leq 0.05$; ns = not significant. **(C)** Nascent nucleolar RNA levels are reduced after siRNA depletion of RSL24D1 and individual PeBoW complex members by 5-EU visualization of nucleolar rRNA biogenesis. Seventy-two hours after siRNA transfection, MCF10A cells were supplemented with 1 mM 5-EU for 1 h before fixation. Representative images shown of fixed cells stained for DNA (Hoechst) and the nucleolar protein fibrillarin (FBL), and click chemistry was performed to conjugate AF488 azide to labeled nascent RNA (5-EU). **(D)** Quantitation of the results in (C). Nucleolar rRNA biogenesis was quantified in MCF10A cells treated with the indicated siRNAs, where strongly reduced 5-EU signal corresponds to RNAP1 inhibition (Bryant et al. 2022). The negative control non-targeting siRNA (siNT) (n = 16 wells per replicate) was set at 0% inhibition and the positive control siRPA194 (n = 16 wells per replicate) was set at 100% inhibition. siNOL11 and siSBDS were additional positive and negative controls, respectively (n = 1 well per replicate). Quantitation of results were reported relative to siNT. Graph indicates mean ± SD, n = 3 biological replicates. Data were analyzed by one-way ANOVA with Dunnett's multiple comparisons test where **** $p \leq 0.0001$; ns = not significant. **(E)** Summary data table of the quantitation in (D). Total number of cells analyzed (sum of 3 independent repetitions) for each siRNA treatment. Average of the median nucleolar 5-EU intensity raw value from 3 independent repetitions for each siRNA treatment. The images were acquired by me, analyzed by Carson, and made into a figure by Amber.

RSL24D1 depletion induces the nucleolar stress response

Nucleolar stress denotes a key cellular response to drugs and environmental insults including impaired ribosome biogenesis due to RNAP1 transcription repression (Rubbi and Milner 2003; Lindstrom et al. 2018). When disrupted, the nucleolus can signal stress by activating the p53 pathway to initiate cell cycle arrest and apoptosis. Previously, PeBoW-complex members have been shown to induce the nucleolar stress response pathway when depleted (Holzel et al. 2005; Grimm et al. 2006). Likewise, I tested whether nucleolar perturbations described above upon RSL24D1 depletion were linked to concomitant TP53 stabilization by western blotting for p53 levels in MCF10A cells depleted of RSL24D1. Indeed, knockdown of RSL24D1 induced p53 accumulation relative to the mock (no siRNA) and siNT negative controls (Figure 5.7A). TP53 stabilization following RSL24D1 knockdown was also observed in a recent genome-wide high-throughput screen conducted in A549 cells (Hannan et al. 2022). I further blotted for TP53 transcriptional target gene CDKN1A (p21) levels to orthogonally validate p53 activation and found these levels to be significantly increased relative to the mock (no siRNA) and siNT negative controls (Figure 5.7B). Additionally, increased *CDKN1A* (*p21)* transcript levels were observed by qRT-PCR in MCF10A cells (Figure 5.7C). These findings show that RSL24D1 depletion increases TP53 and CDKN1A levels, likely as a result of nucleolar stress response induction upon impaired ribosome biogenesis, as has been observed previously upon inhibition of pre-60S biogenesis (Sun et al. 2010; Fumagalli et al. 2012; Daftuar et al. 2013) or RNAP1 activity (Bywater et al. 2012; Peltonen et al. 2014a; Fu et al. 2017).

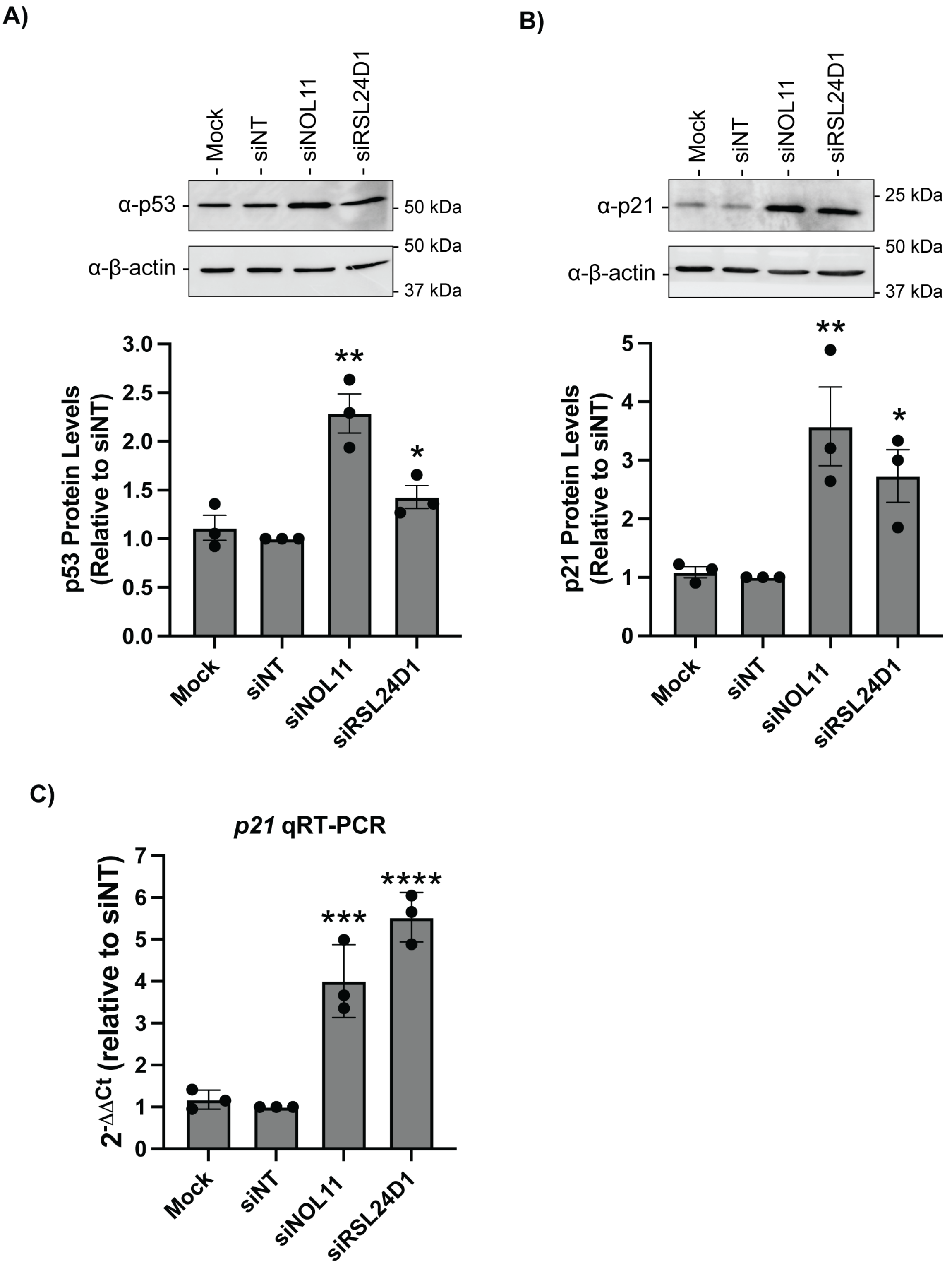


Figure 5.7: RSL24D1 depletion induces the nucleolar stress response. **(A)** RSL24D1 depletion increases p53 protein levels. MCF10A cells were treated with the indicated siRNAs for 72 h and total protein from whole cell extracts was harvested. Mock and non-targeting siRNA (siNT) were used as negative controls and siNOL11 was used as a positive control. (Top): Representative western blot images using α-p53 antibody, and α-

β-actin antibody was used as a loading control. (Bottom): Quantitation of p53 levels is reported relative to siNT and normalized to the β-actin loading control. Graph indicates mean ± SD, n = 3 biological replicates. Data were analyzed by Student's t-test followed by multiple testing p-value correction (two-stage linear step-up procedure of Benjamini, Krieger, and Yekutieli) where * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$. **(B)** RSL24D1 depletion increases p21 protein levels in MCF10A cells. Panel as above, except western blotting was performed using the α-p21 antibody. **(C)** RSL24D1 depletion increases *p21* mRNA levels in MCF10A cells. qRT-PCR was performed and quantified using the $2^{-\Delta\Delta Ct}$ method, relative to a siNT control and 7SL internal control primer. Non-targeting siRNA (siNT) and siNOL11 were used as negative and positive controls respectively. Graph indicates mean ± SD, n = 3 biological replicates. Data were analyzed by Student's t-test followed by multiple testing p-value correction (two-stage linear step-up procedure of Benjamini, Krieger, and Yekutieli) where *** $p \leq 0.001$, **** $p \leq 0.0001$.

Discussion

Here, I report the functional analysis of the putative ribosome biogenesis assembly factor RSL24D1. This protein was previously identified to be important for 60S biogenesis (Durand et al. 2021; Dorner et al. 2022), and consistent with these reports, Amber Buhagiar demonstrated that RNAi mediated depletion of RSL24D1 specifically inhibits the late pre-rRNA processing steps required for 60S subunit formation. Moreover, she showed repression of this pathway by RSL24D1 depletion leads to a global reduction in mRNA translation. Lisa Ogawa showed that disruption of these processes after knockdown of RSL24D1 is associated with a reduction in nucleolar number (Farley-Barnes et al. 2018) and I showed a concomitant activation of the nucleolar stress response pathway. Remarkably, Amber Buhagiar and I uncover an unexpected role for RSL24D1 and the PeBoW complex for the efficient production of pre-rRNA by RNAP1, the first step of ribosome biogenesis. We demonstrate that RSL24D1 and PeBoW are positive regulators of rDNA promoter activity and RPA194 stability. I provide a likely mechanism through my confirmation of RSL24D1 and WDR12 association with RPA194, the largest subunit of the RNAP1 enzyme complex.

In the present study, Amber Buhagiar and I provide strong evidence that LSU-associating assembly factors can also serve as dual-function factors in pre-rRNA transcription and processing. Several assembly factors are known to associate with the nascent pre-rRNA transcript to facilitate RNAP1 recruitment and initiation. A subgroup of these factors, the t-UTPs, assemble synchronously to promote rDNA transcription and stimulate the formation of the SSU processome for the processing of the pre-40S subunit (Gallagher et al. 2004; Krogan et al. 2004; Prieto and McStay 2007). The results presented here are supported by previous studies that have identified nucleolar proteins that regulate both pre-rRNA transcription and pre-LSU processing, including splicing factor HTATSF1 and DNA repair protein FANCI (Corsini et al. 2018; Sondalle et al. 2019). Amber Buhagiar and I extend this framework to include the LSU factors RSL24D1 and PeBoW complex members (PES1-BOP1-WDR12) whose depletion leads to decreased RPA194 levels, down-regulated rDNA promoter activity, strongly decreased nucleolar rRNA biogenesis consistent with roles in pre-rRNA processing and transcription, and impaired pre-LSU processing.

My results point towards both direct and potentially indirect mechanisms of RNAP1 regulation by RSL24D1 and PeBoW. Consistent with the interactome results in Sorino et al. 2020, I observed an association of WDR12 and RSL24D1 with RPA194 by co-immunoprecipitation. Although a reduction in rDNA transcription and RPA194 levels extended to depletion of other PeBoW components, I did not find an association between PES1 and RPA194. This lack of association agrees with previously published PES1 and BOP1 interactomes, where RPA194 was not detected in either (Kellner et al. 2015). Interestingly, PES1 has been identified to be an interacting partner of Upstream Binding Transcription Factor (UBTF) which is required for the recruitment of the pre-initiation complex along the rDNA promoter (Huttlin et al. 2017). Furthermore, all three PeBoW members interact with SIRT7, a positive regulator of RNAP1 (Ford et al. 2006; Tsai et al. 2012).

Another potential explanation for PeBoW members all causing a reduction in rDNA transcription when individually depleted may be due to the interdependence of their steady state protein levels. Previous work has shown that depletion of individual PeBoW members results in a reduction in protein levels of the other complex members in human cells (Rohrmoser et al. 2007; Mi et al. 2021). Further studies investigating the crosstalk between these protein interactions and their influence on RNAP1 transcription activation will be able to untangle the precise mechanism through which RSL24D1, PeBoW complex members, and maybe even other LSU maturation factors, regulate this process.

Surprisingly, I was unable to co-immunoprecipitate PES1 with WDR12. This result was unexpected, as the presence of both proteins is well-established within the PeBoW complex (Holzel et al. 2005; Grimm et al. 2006; Rohrmoser et al. 2007). One potential explanation is that a large pool of endogenous WDR12 outside of the PeBoW complex led to our inability to detect the PES1-WDR12 association by co-immunoprecipitation. Although I validated WDR12 and RSL24D1's association with RPA194, further investigation will be critical to provide insight into the exact physical associations between RSL24D1, PeBoW, and RNAP1 in human cells.

While RSL24D1 and PeBoW gene mutations are not yet known to be implicated in the molecular pathogenesis of ribosomopathies, the diseases of making ribosomes, several were shown to be linked to mutations in genes encoding LSU biogenesis factors. For example, mutations in *SBDS* are associated with Schwachman-Diamond syndrome (Boocock et al. 2003) and defects in *RBM28* cause alopecia, neurological defects, and endocrinopathy (ANE) syndrome (Nousbeck et al. 2008; McCann et al. 2016; Bryant et al. 2021). RSL24D1 was recently reported to be required for murine embryonic stem cell proliferation (Durand et al. 2021) and BOP1 in *Xenopus* anterior development (Gartner et al. 2022). Previously, it has been shown that PES1 is required for zebrafish and mouse embryonic development (Allende et al. 1996; Lerch-Gaggl et al. 2002) and more specifically *Xenopus* neural crest cell migration (Gessert et al. 2007). The cranial cartilage defects observed in *Xenopus* upon loss of BOP1 (Gartner et al. 2022) and PES1 (Gessert et al. 2007) are hallmarks of ribosomopathies (Farley-Barnes et al. 2019). It will be of interest to further dissect the roles of RSL24D1 and PeBoW in ribosome biogenesis as they relate to embryonic development in vertebrates. The prevalence of disorders arising from mutations in genes encoding assembly factors of the large subunit highlight the importance of faithful subunit processing for the steady production of ribosomes to maintain cellular homeostasis.

Materials and Methods

Publicly Available Expression Datasets

Genotype-Tissue Expression (GTEx) unmatched normal and The Cancer Genome Atlas (TCGA) matched normal and tumor expression datasets were obtained through the Xena platform (https://xena.ucsc.edu/) (Goldman et al. 2020). RNA-seq by Expectation-Maximization (RSEM) LOG2 fold expression levels for *RSL24D1* were subtracted from the mean of the overall normal and tumor tissues combined for graphical visualization.

Cell culture and media

MCF10A cells (ATCC CRL-10317) were subcultured in Dulbecco's modified Eagles' medium/Nutrient mixture F-12 (Gibco 1130-032) containing horse serum (Gibco 16050), 10 µg/mL insulin (Sigma I1882), 0.5 µg/mL hydrocortisone (Sigma H0135), 100 ng/mL cholera toxin (Sigma C8052), and 20 ng/mL epidermal growth factor (Peprotech AF-100-15). Cells were maintained at 37 °C, in a humidified atmosphere with 5% CO_2. For the high-throughput nucleolar number and 5-EU screens, 3,000 cells/well were reverse transfected into 384-well plates on day 0. For the dual-luciferase reporter assay, 75,000 cells/well were seeded into 12-well plates on day 0. For RNA or protein isolation, 100,000 cells/well were seeded into 6-well plates on day 0.

RNAi

All siRNAs were purchased from Horizon Discovery Biosciences (Data S1). siGENOME SMARTpool siRNAs were used in the original screen (Farley-Barnes et al. 2018). For the 5-EU secondary screen validation of RSL24D1, PeBoW, and SBDS, ON-TARGETplus pools were used except for the NOL11 positive control which used the siGENOME SMARTpool siRNAs. For screen validation by deconvolution, the individual siONT set of four siRNAs that comprised the pool was used. The ON-TARGETplus pools were used in the remaining functional analysis. siRNA transfection was performed using Lipofectamine RNAiMAX Transfection Reagent (Invitrogen 13778150) as described in (Ogawa et al. 2021). For the high-throughput screens, cells were reverse transfected into 384-well plates on day 0. siNT, siUTP4, siRPA194, and siNOL11 controls were added to 16 wells per plate, while our experimental samples were added to 1 well each for each independent screen repetition. For other assays, cells were transfected 24 h after plating. For siONT deconvolution, an individual siRNA targeting RSL24D1 was considered validated if it produced a mean one-nucleolus percent effect greater than or equal to +3 SD above the siNT mean, using the siNT SD.

5-EU incorporation assay for nucleolar rRNA biogenesis

Following 72 h of siRNA depletion, MCF10A cells were treated with 1 mM 5-ethynyl uridine (5-EU; Click Chemistry Tools 1261) for 1 h to label nascent RNA as in (Bryant et al. 2022). Briefly, cells were washed with PBS, fixed with 1% paraformaldehyde (Electron Microscopy Sciences 15710-S) in PBS for 20 minutes, and permeabilized with 0.5% (vol/vol) Triton X-100 in PBS for 5 minutes. Cells were blocked with 10% (vol/vol) FBS (MilliporeSigma F0926) in PBS for 1 h at room temperature. Nucleoli were stained with 72B9 primary anti-fibrillarin antibody (Reimer et al. 1987), 1:250 for 2 h at room temperature, followed by secondary AlexaFluor 647 goat anti-mouse immunoglobulin G (1:1,000, Invitrogen A-21235) with Hoechst 33342 dye (1:3,000) for nuclei detection for 1 h at room temperature. 5-EU incorporation was visualized by performing the following click reaction in PBS: $CuSO_4{:}5H_2O$ (0.5 mg/mL resuspended in water, Acros Organics 197730010), ascorbic acid (20 mg/mL freshly resuspended in water, Alfa Aesar A15613), and AF488 azide (0.5 µM resuspended in DMSO, Click Chemistry Tools 1275) for 30 min at room temperature. Cells were then soaked with Hoechst 33342 (1:3,000) for 30 min to dissociate excess azide dye. Cell images were acquired using IN Cell 2200 imaging system (GE Healthcare) and analysis was performed using a custom CellProfiler pipeline (Bryant et al. 2022).

Dual-luciferase reporter assay for pre-rRNA transcription

Following 48 h of siRNA depletion, MCF10A cells were co-transfected with 1000 ng of pHrD-IRES-Luc (Ghoshal et al. 2004) and 0.1 ng of a Renilla internal control plasmid using Lipofectamine 3000 (Thermo Fisher Scientific L3000015). Twenty-four h after plasmid transfection, cells here harvested and luminescence was measured using the Dual-luciferase Reporter Assay System (Promega E1910) following the manufacturer’s instructions with a GloMax 20/20 luminometer (Promega). The ratio of pHrD-IRES-luciferase/Renilla activity was calculated to control for transfection efficiency and normalized to the non-targeting control.

qRT-PCR analysis

After 72 h of siRNA-mediated depletion, RNA was extracted using TRIzol (Life Technologies 5596018) as per the manufacturer’s instructions. cDNA preparation was performed using the iScript gDNA Clear cDNA synthesis Kit (Bio-Rad 1725035), and qPCR was performed using iTaq Universal SYBR Green Supermix (Bio-Rad 1725121). All $A_{260/230}$ values were above 1.7 prior to cDNA synthesis. Amplification of the 7SL RNA was used as an internal control, and analysis was completed using the comparative C_T

method (ΔΔC_T). Bio-Rad PrimePCR Assay gene-specific primers were used to test *RSL24D1* mRNA levels (Bio-Rad 10025636; RSL24D1, qHsaCID0021318) and the primers for *BOP1, PES1, SBDS, WDR12,* 7SL are available in (Data S2). Melt curves were performed for each sample to verify the amplification of a single product. Three biological replicates, each with three technical replicates, were measured.

Western blots

Following 72 h of siRNA depletion, cells were collected and total protein was harvested and prepared as in (Farley-Barnes et al. 2018). A Bradford assay (BioRad 5000006) was used to quantify amount of total protein. 25-50 μg of total protein was separated by SDS-PAGE and transferred to a PVDF membrane (Bio-Rad 1620177). Membranes were blocked for 1 h in 5% milk in 1X PBST (or 5% BSA in 1X PBST for western blots of immunoprecipitations) and incubated overnight with the specified primary antibodies at 4 °C. Proteins were detected with the following antibodies: α-RSL24D1 (dilution 1:1,000; Proteintech 25190-1-AP), α-RPA194 (dilution 1:1,000; Santa Cruz sc-48385), α-WDR12 (1:1000; Bethyl Laboratories A302-650A), α-PES1 (1:1000; Bethyl Laboratories A300-902A), α-p53 (dilution 1:5,000; Santa Cruz sc-126), α-p21 (dilution 1:400, Santa Cruz sc-6246), α-puromycin (dilution 1:10,000; Kerafast EQ0001), and α-β-actin (dilution 1:30,000; Sigma Aldrich A1978). The western blots were developed with enhanced chemiluminescence reagents (Thermo Scientific 34096). Images were acquired by digital imaging using the Bio-Rad ChemiDoc Imaging System. Images were quantified with ImageJ software.

Co-immunoprecipitation

Protein A agarose beads (Cell Signaling Technologies 9863S) were washed and incubated in NET150 buffer (20 mM Tris HCl pH. 7.5, 150 mM NaCl, 0.05% NP-40) with 20 μg of the indicated antibodies overnight at 4 °C with nutation. Harvested MCF10A cells were washed with PBS and incubated on ice for 10 min in NET150 buffer (with the addition of 1X protease inhibitors and 4 mM NEM). Total cell extracts were obtained by sonication and cleared by centrifugation at 16,000 g for 10 min at 4 °C after lysis. Supernatants were incubated with either antibody-bound or unconjugated Protein A beads for 2 h at 4 °C with nutation. RNase A treated extracts were treated with 20 µg/mL RNase A (AMRESCO E866). After beads were washed five times with NET150, immunocomplexes were eluted in 2X Laemmli buffer and resolved on a 12% acrylamide gel. Western blotting was performed as described above.

Statistical analyses

All statistical analyses were performed in GraphPad Prism 8.2.1 (Graphpad Software) using the tests described in the figure legends.

Chapter 6

The human cytidine deaminase, APOBEC3A, is required for large ribosomal subunit biogenesis, revealing the potential for pre-rRNA editing

Introduction

An upregulation of ribosome biogenesis and protein synthesis is tightly associated with cancer [Reviewed in (Orsolic et al. 2016; Pelletier et al. 2018)]. The nucleolus is pleiotropic itself (Panov et al. 2021) and so are various ribosome biogenesis factors with dual-functions in other cellular processes (Berger et al. 2015; Ogawa and Baserga 2017; Sondalle et al. 2019). A main connection to the nucleolus is cell cycle progression through the TP53 mediated nucleolar stress response (Rubbi and Milner 2003; Boulon et al. 2010; Hannan et al. 2022). This stress response can result in nucleolar structure alterations, which are attributes of cancer cells; however, cancer cells exhibit increased number and size of nucleoli, correlating with greater ribosome production (Pianese 1896; Derenzini et al. 1998). The reliance that proliferating cancer cells have on increased number of ribosomes is exemplified by the fact there are multiple cancer therapeutics in various stages of development that target the ribosome biogenesis pathway directly (Catez et al. 2019; Zisi et al. 2022).

While cell growth in cancer relies on increased ribosome biogenesis, carcinogenesis is driven by somatic genomic DNA damage and mutation (Hanahan and Weinberg 2011). One somewhat newly implicated factor contributing to genomic instability is the aberrant function of the cytidine deaminase family of proteins (Siriwardena et al. 2016; Grillo et al. 2022). Cytidine deaminases convert cytosine to thymine on single-stranded DNA (ssDNA) or to uracil on RNA with established biological roles in the immune response through hypermutation of the antibody variable region [activation induced cytidine deaminase (AICDA)] (Muramatsu et al. 2000; Feng et al. 2021), mutation of infecting viral genomes, and endogenous retroviral elements (apolipoprotein B mRNA editing catalytic polypeptide-like 3 subfamily members (APOBEC3s) (Harris and Dudley 2015; Willems and Gillet 2015). The first studied APOBEC, APOBEC1, (Greeve et al. 1991; Teng et al. 1993) edits *apoB* mRNA and other RNAs, (Rosenberg et al. 2011) while still maintaining activity towards ssDNA (Harris et al. 2002). On the other hand, APOBEC2 and APOBEC4 have no observed catalytic activity (Harris et al. 2002; Marino et al. 2016). These useful editing functions can have unintended and deleterious consequences when the host genome is targeted. APOBEC3A and APOBEC3B expression is detected in many cancer types and is associated with attributing their own classified Catalogue of Somatic Mutations in Cancer (COSMIC) mutational signatures (single base substation signatures (SBS) 2 and 13) being present in over 50% of human cancer types (Roberts et al. 2013; Alexandrov et al. 2020; Mertz et al. 2022). To date, APOBEC3A is emerging as the more prominent cytidine deaminase in editing the genome in cancer (Cortez et al. 2019; Petljak et al. 2022).

On top of genomic mutations, RNA editing is arising as another contributing element to the molecular pathogenesis of cancer (Baysal et al. 2017). Largely studied in an immune cell context, APOBEC3A (Niavarani et al. 2015; Sharma et al. 2015; Sharma et al. 2017; Alqassim et al. 2021) and APOBEC3G (Sharma et al. 2016; Sharma et al. 2019) possess mRNA editing activity. Similar to its increased expression resulting in off-target genomic DNA editing, APOBEC3A has been shown to edit mRNAs when expressed in cancer (Jalili et al. 2020) and this is possibly true for other APOBEC3 family members as well (Asaoka et al. 2019). While not completely understood, APOBEC3A is a multifaceted DNA/RNA editing enzyme that can provide a defense against viruses, but its non-discriminatory editing also contributes to cancer disease progression through aberrant editing of the genome and potentially of mRNAs.

In this chapter, I provide evidence that APOBEC3A regulates ribosome biogenesis in human cells, another pathway that drives cancer. Previously, our laboratory identified APOBEC3A and APOBEC4 in a genome-wide siRNA screen as regulators of nucleolar function (Farley-Barnes et al. 2018). Lisa Ogawa, myself, and the

Yale Center for Molecular Discovery (Laura Abriola and Yulia Surovtseva) validated APOBEC3A and APOBEC4 as two novel ribosome biogenesis factors (Chapter 3). Furthermore, I establish that APOBEC3A is required for cell cycle progression and for maintaining normal levels of protein production in MCF10A and HeLa human cell lines. More precisely, I show that APOBEC3A is required for proper LSU maturation and pre-LSU rRNA processing. I investigated how APOBEC3A regulates ribosome biogenesis and the cell cycle through a novel nuclear RNA-sequencing experiment, identifying candidate sites of APOBEC3A C to U editing on the pre-rRNA and pre-mRNAs for the first time. I show a potential direct function of APOBEC3A in ribosome biogenesis by identification of candidate editing sites on the pre-LSU rRNA. I also observed a potential indirect function by identification of APOBEC3A candidate editing sites on pre-mRNAs encoding nucleolar proteins and cell cycle regulators. My results point towards APOBEC3A being a modulator of ribosome synthesis, linking its expression to cancer cell proliferation through a novel mechanism.

I have submitted a manuscript with the results within this chapter where I am first author which is currently available as a pre-print on bioRxiv. The experiments in this chapter were performed by myself with the exception that Carson Bryant analyzed the high-throughput screening images for the nucleolar rRNA biogenesis assay and Laura Abriola at the YCMD performing the high-throughput deconvolution experiments.

Results

Investigating The Cancer Genome Atlas (TCGA) and previous high-throughput screens provide strong evidence that APOBEC3A regulates human ribosome biogenesis

Due to the various connections between cytidine deaminases and cancer, I analyzed all 11 known human cytidine deaminases' mRNA expression levels across cancer types using Genotype-Tissue Expression (GTEx) unmatched normal and The

Cancer Genome Atlas (TCGA) matched normal and tumor samples (Goldman et al. 2020). I observed that 9 out of 11 cytidine deaminases have significantly higher expression in tumor versus normal tissue (Figure 6.1). Since this does not provide any evidence for a functional role within the context of tumorigenesis or cancer progression, I sought to examine any correlations between cytidine deaminase expression and patient survival across cancer types using the associated patient survival data (Goldman et al. 2020). After stratifying expression into either high or low based on the mean expression level, 8 out of the 11 cytidine deaminases examined had high expression levels associated with a significant reduction in survival probability (Figure 6.2). Based on these two analyses, 7 out of the 11 cytidine deaminases are more highly expressed in cancer versus normal tissue and their higher expression is associated with decreased survival probability (Figure 6.2). I aimed to build upon these correlations with analysis of cytidine deaminase function in cancer by interrogating this protein family for yet to be discovered cellular roles, specifically ribosome biogenesis related roles.

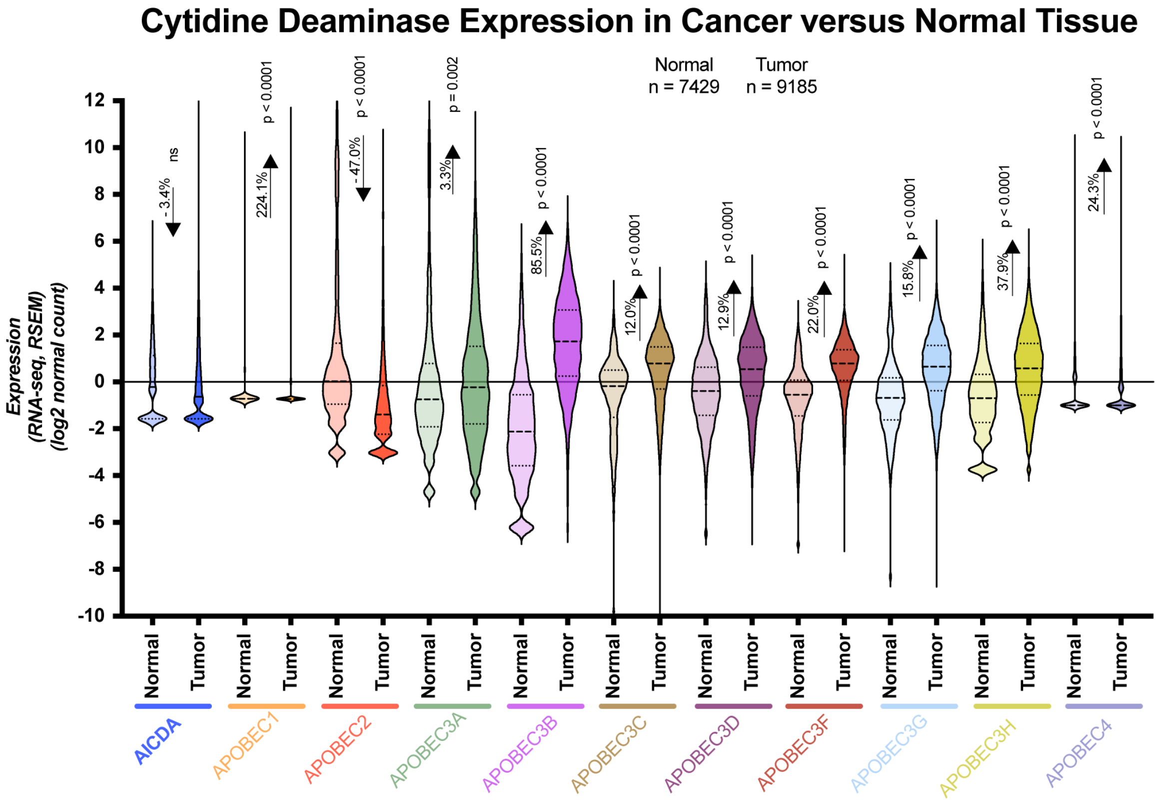


Figure 6.1: Nine out of the 11 human cytidine deaminases are more highly expressed in tumor versus normal tissue. Violin plots from Genotype-Tissue Expression (GTEx) unmatched normal and The Cancer Genome Atlas (TCGA) matched normal and tumor RNA-seq by Expectation-Maximization (RSEM)(Goldman et al. 2020) Log_2 fold expression levels for human cytidine deaminases subtracted from the mean. Mean of tumor (N = 9185) and normal (N = 7429) cytidine deaminase expression set at 0 (horizontal line). For each indicated transcript, median (heavy dashed line), quartiles (light dashed line), normal data (light shading), tumor data (darker shading). Percent change in expression between normal versus tumor indicated for each cytidine deaminase transcript. Data were analyzed by Student's t-test, significant p-values are reported on the graph.

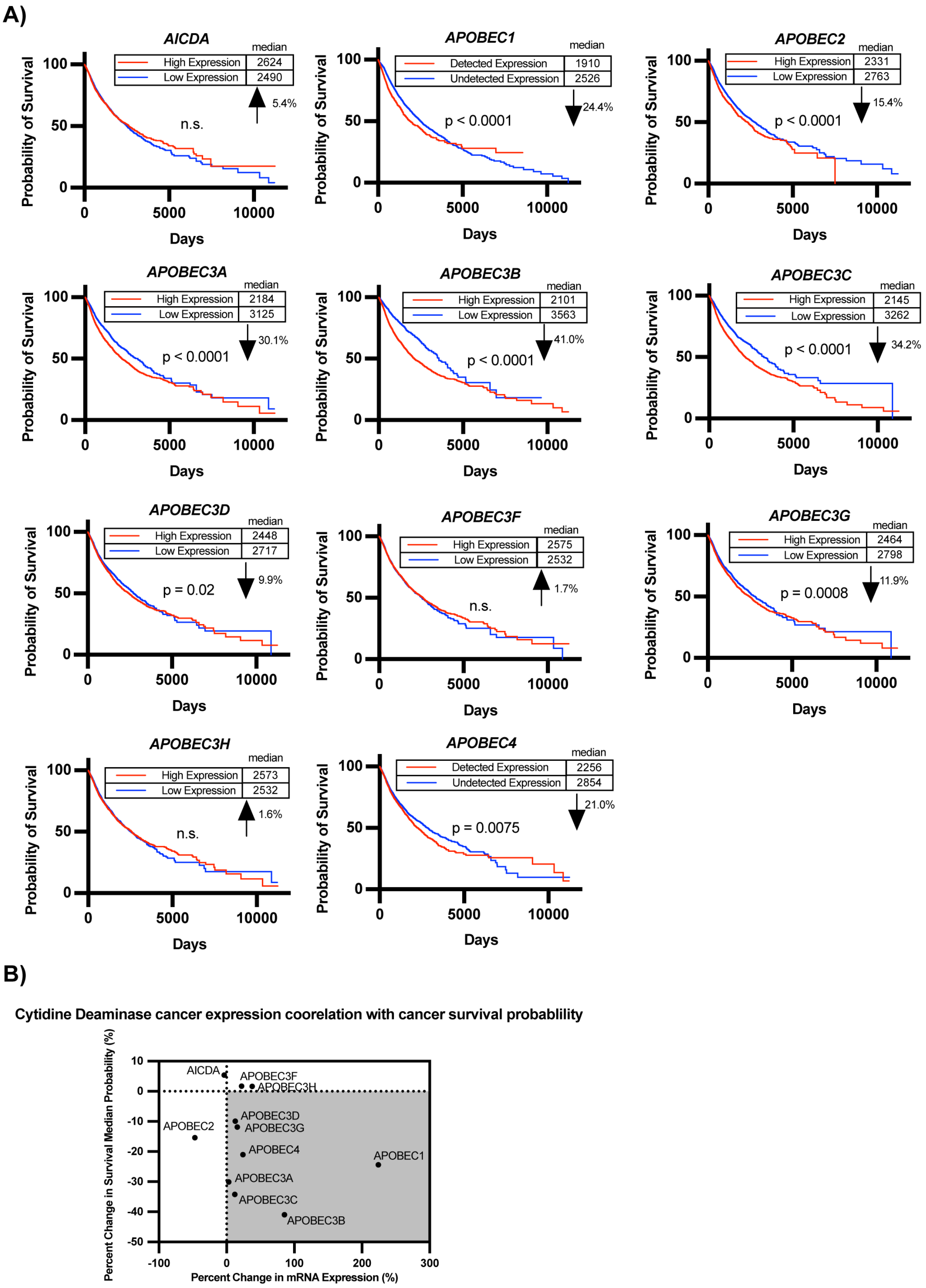


Figure 6.2: Seven out of the 11 human cytidine deaminases are more highly expressed in tumor versus normal tissue and their higher expression is correlated with a decrease

in survival. **(A)** Higher expression of 8 out of the 11 human cytidine deaminases correlates with decreased survival probability. Survival data was obtained from Genotype-Tissue Expression (GTEx) unmatched normal and The Cancer Genome Atlas (TCGA) matched normal and tumor samples using the Xena platform.(Goldman et al. 2020) Samples were stratified by either high or low cytidine deaminase expression and survival curves were generated. For cytidine deaminases where the median expression was 0, data was stratified by either detected or undetected levels of expression. Median survival (days) and the percent difference between low and high expression groups were reported. Data were analyzed by Kaplan-Meier survival analysis, significant p-values reported on respective graphs. **(B)** Seven out of the 11 human cytidine deaminases exhibit higher expression in tumor versus normal tissue and higher expression is associated with lower survival probability. Percent change in mRNA expression from (S1) (x-axis) and percent change in median survival from (A) (y-axis) was plotted for each cytidine deaminase. An increase in expression and a decrease in survival quadrant is indicated by the grey background.

Next, I leveraged our laboratory's previously published high-throughput siRNA screen for novel regulators of nucleolar function (Chapter 1, 3) (Farley-Barnes et al. 2018). In the previously published screen from our laboratory, out of the 10 cytidine deaminases analyzed, only APOBEC3A and APOBEC4 were hits. Their siRNA depletion produced a one-nucleolus percent effect to an even greater extent than the positive controls siUTP4 and siNOL11, while the other cytidine deaminases did not meet this threshold (Figure 6.3) (Farley-Barnes et al. 2018). As expected, this result correlated with a decrease in cell viability, both of which were remarkably lower than the total screen median of 83.7% compared to non-targeting siRNA (siNT) negative control (100%) (Figure 6.3) (Farley-Barnes et al. 2018). These results indicate that APOBEC3A and APOBEC4 are potentially involved in making ribosomes, while it is less likely that other cytidine deaminase family members possess a function in ribosome biogenesis.

A)

Cytidine Deaminase	Activity / Localization
APOBEC1	Pan-cellular
APOBEC2	? (No activity)
APOBEC3A	Pan-cellular
APOBEC3B	Nuclear
APOBEC3C	Pan-cellular
APOBEC3D	Cytoplasmic
APOBEC3F*	Cytoplasmic
APOBEC3G	Cytoplasmic
APOBEC3H	Pan-cellular
APOBEC4	? (No activity)
AICDA	Pan-cellular

*not in screening library

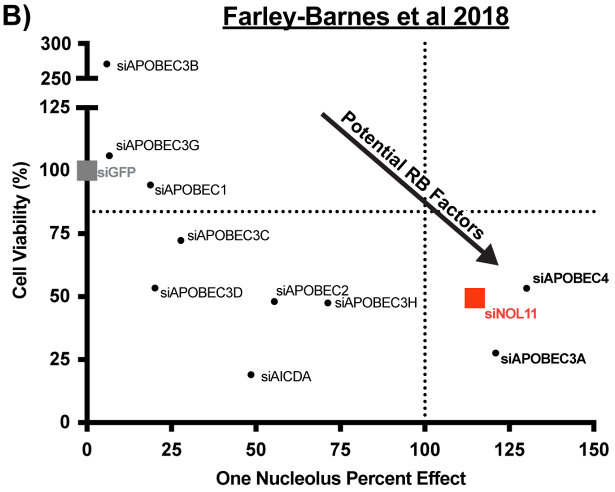


Figure 6.3: The cytidine deaminases, APOBEC3A and APOBEC4, are novel ribosome biogenesis factors in MCF10A cells. **(A)** The cytidine deaminase family of proteins have diverse activities and sub-cellular localizations. This table lists all 11 known human cytidine deaminases with their sub-cellular localizations or if they do not have known catalytic activity. This information regarding cytidine deaminase function was curated from the following sources: (Smith et al. 2012; Lackey et al. 2013; Revathidevi et al. 2021). **(B)** The cytidine deaminases APOBEC3A and APOBEC4 were hits in our laboratory's previous genome-wide siRNA screen for ribosome biogenesis regulators in MCF10A cells (Farley-Barnes et al. 2018). Graph of data on cell viability (y-axis) and one-nucleolus harboring cells (one-nucleolus percent effect, x-axis) from 72 h siRNA treatment targeting 10 of 11 known human cytidine deaminases. siGFP (grey) was a negative control (100% viability, 0% one-nucleolus percent effect) and siUTP4 (100% one-nucleolus percent effect, vertical dotted line) and siNOL11 (red) were positive controls. The entire screen cell viability median was 83.7% (horizontal dotted line). siAPOBEC3A and siAPOBEC4 treatments (bold, lower right quadrant) passed the thresholds for both one-nucleolus percent effect and reduced cell viability.

Furthermore, I took advantage of other available published ribosome biogenesis screening data to provide more insight on APOBEC3A, APOBEC4, and other cytidine deaminases functions (Wild et al. 2010; Tafforeau et al. 2013; Badertscher et al. 2015; Ogawa et al. 2021; Dorner et al. 2022). However, while some possess potential roles in 40S biogenesis (Badertscher et al. 2015), no cytidine deaminases met the threshold required to be deemed a hit in any of these other ribosome biogenesis screens (Table 6.1). Thus, I did not include any other cytidine deaminases in my downstream analyses.

Table 6.1: Previous ribosome biogenesis related screen results of human cytidine deaminase and indicated ranking within screen. Screen cutoffs at bottom.

Cytidine Deaminase	Farley-Barnes et al 2018 (One-nucleolus screen, MCF10A cells)	Ogawa et al 2021 (5+ nucleoli screen, MCF10A cells)	Wild et al 2010 (Targeted 40S and 60S subunit screen, HeLa cells)	Tafforeau et al 2013 (Targeted pre-rRNA processing, HeLa cells)	Badertscher et al 2015 (40S subunit, HeLa cells)	Doerner et al 2022 (60S subunit, HeLa cells)	Hannan et al 2022 (p53 stabilization, A549 cells)
AICDA	No (4,366 / 18,107)	No (15,662 / 18,107)	Not tested	Not tested	No (10,135 / 19,601)	No (17,611 / 20,482)	No (3,834 / 18,120)
APOBEC1	No (11,018 / 18,107)	No (17,498 / 18,107)	Not tested	Not tested	No (15,428 / 19,601)	No (17,312 / 20,482)	No (13,283 / 18,120)
APOBEC2	No (3,421 / 18,107)*	No (9,463 / 18,107)	Not tested	Not tested	No (3,330 / 19,601)	No (3,300 / 20,482)	No (10,701 / 18,120)
APOBEC3A	Yes (208 / 18,107)	No (3,377 / 18,107)	Not tested	Not tested	No (19,150 / 19,601)	No (6,207 / 20,482)	Yes (784 / 18,120)
APOBEC3B	No (14,557 / 18,107)	No (5,080 / 18,107)	Not tested	Not tested	No (2,562 / 19,601)*	No (15,629 / 20,482)	No (6,158 / 18,120)
APOBEC3C	No (8,579 / 18,107)	No (4,290 / 18,107)	Not tested	Not tested	No (14,732 / 19,601)	No (5,563 / 20,482)	No (1,026 / 18,120)
APOBEC3D	No (10,627 / 18,107)	No (952 / 18,107)*	Not tested	Not tested	No (3,537 / 19,601)	No (9,368 / 20,482)	No (17,561 / 18,120)
APOBEC3F	Not tested	Not tested	Not tested	Not tested	No (2,089 / 19,601)*	No (7,597 / 20,482)	Not tested
APOBEC3G	No (14,370 / 18,107)	No (5,620 / 18,107)	Not tested	Not tested	No (1,484 / 19,601)*	No (4,813 / 20,482)	No (5,008 / 18,120)
APOBEC3H	No (1,848 / 18,107)*	No (17,261 / 18,107)	Not tested	Not tested	No (10,207 / 19,601)	No (2,835 / 20,482)	No (17,916 / 18,120)
APOBEC4	Yes (127 / 18,107)	No (15,428 / 18,107)	Not tested	Not tested	No (14,248 / 19,601	No (10,447 / 20,482)	No (4,476 / 18,120)

Hit Cutoffs	Farley-Barnes et al 2018 (One-nucleolus screen, MCF10A cells)	Ogawa et al 2021 (5+ nucleoli screen, MCF10A cells)	Wild et al 2010 (Targeted 40S and 60S subunit screen, HeLa cells)	Tafforeau et al 2013 (Targeted pre-rRNA processing, HeLa cells)	Badertscher et al 2015 (40S subunit, HeLa cells)	Doerner et al 2022 (60S subunit, HeLa cells)	Hannan et al 2022 (p53 stabilization, A549 cells)
Published Hit	> 122% effect	> 25% effect	N/A	N/A	multiple scoring parameters and validation, top 302	multiple scoring parameters and validation, top 310	> 2-fold p53 increase
Potential (arbitrary)	* > 50% effect	* top 10%	N/A	N/A	* $p < 0.05$	* $p < 0.05$	* > 1.33 fold p53 increase

Validation of APOBEC3A as a ribosome biogenesis factor

To validate the screening results, I repeated the established nucleolar number assay in triplicate with an updated siRNA technology (si-ONTARGET, Horizon Discovery) which reduces off-target effects (Jackson et al. 2006) (Chapter 3). I observed a significant increase in cells harboring 1 nucleolus after APOBEC3A depletion (46.3% one-nucleolus cells, 224.8% effect), but only a modest increase in one-nucleolus harboring cells after APOBEC4 depletion (23.9% one-nucleolus cells, 33.6% effect) compared to the negative control siNT (19.2% one-nucleolus cells, 0% effect) and positive control siNOL11 (29.2% one-nucleolus cells, 100% effect) treatments (Figure 6.4). Furthermore, Laura Abriola at the YCMD deconvoluted the si-ONTARGET pool by testing each of the 4 individual siRNA's ability to produce the one-nucleolus phenotype. I used a stringent >3 standard deviation from the negative control (siNT) as a cutoff for considering the one-nucleolus phenotype to be produced for both the pool and individual siRNA treatments. For siAPOBEC3A treatment, both the pool and 2 out of 4 individual siRNAs passed the cutoff for reducing nucleolar number. While the average of the siAPOBEC4 pool and 2 out of 4 individual siRNAs surpassed the 3 standard deviation cutoff, only 1 of the 3 siAPOBEC4 pool replicates surpassed this cutoff (Figure 6.5, top). Again, decreases in nucleolar number correlated with concomitant decreases in cell viability (Figure 6.5, bottom). Due to the stronger and more consistent results with siAPOBEC3A compared to siAPOBEC4, I decided to focus solely on APOBEC3A as the more promising novel ribosome biogenesis factor.

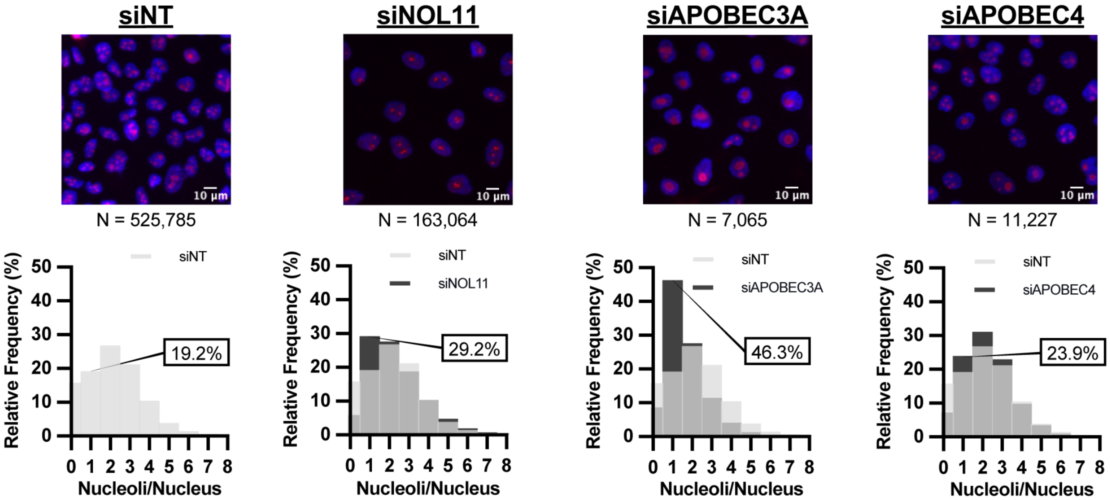


Figure 6.4: APOBEC3A and APOBEC4 depletion by siON-TARGET pools reduces nucleolar number in MCF10A cells. (Top) Representative merged images of nuclei stained with Hoechst (blue) and nucleoli stained with α-fibrillarin (magenta), N = number of nuclei (cells) analyzed. Non-targeting siRNA treatment (siNT) was used as a negative control (2-3 nucleoli per nucleus) and siNOL11 was used as a positive control (reduction in nucleolar number). (Bottom) Histograms of relative frequency of nucleoli per nucleus. siNT (light grey), siNOL11, siAPOBEC3A, and siAPOBEC4 (dark grey), overlap between siNT and other siRNA treatments (intermediate grey). Percentage of one-nucleolus harboring cells is indicated for each siRNA treatment.

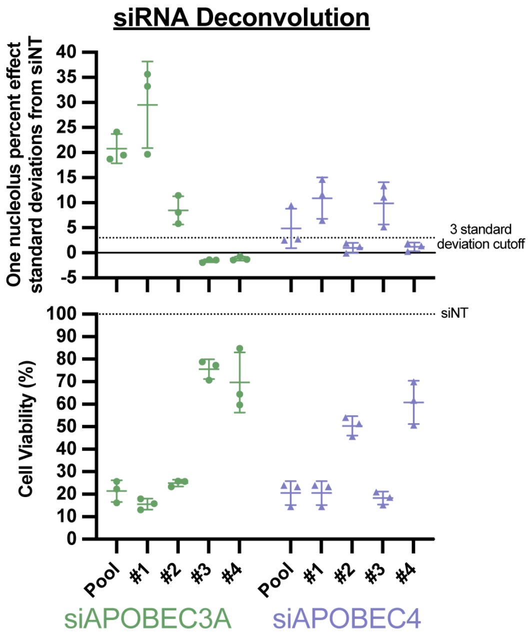


Figure 6.5: siRNA deconvolution for siAPOBEC3A and siAPOBEC4 siON-TARGET pools. (Top) siAPOBEC3A and siAPOBEC4 pool and 2 out of 4 individual siRNA treatments significantly reduce nucleolar number in MCF10A cells. One-nucleolus percent effect set relative to standard deviations from negative control, siNT. 3-standard deviation from negative control cutoff (horizontal dashed-line). 3 biological replicates plotted mean ± SD. (Bottom) The same pool and individual siRNA treatments in (Top) that reduce nucleolar number reduce cell viability. Percent cell viability set relative to siNT treatment (100%, horizontal dashed line). 3 biological replicates plotted mean ± SD.

Additionally, I validated APOBEC3A depletion by western blotting using an antibody that detects both APOBEC3A and 3B. I observed a significant reduction in APOBEC3A levels, but not APOBEC3B levels after treatments with the APOBEC3A specific siRNA pool and the two individual siRNAs that passed deconvolution (siRNAs #1 and #2) (Figure 6.6A). Interestingly, siNT treatment alone produced a trending increase in APOBEC3A protein levels indicating that APOBEC3A expression might be induced upon siRNA treatments. APOBEC3A mRNA levels are typically very low in most cell types, and in some cases undetectable by qRT-PCR (Middlebrooks et al. 2016; Udquim et al. 2020). Consequently, I was unable to reliably detect *APOBEC3A* mRNA by qRT-PCR in MCF10A cells but did successfully observe depletion of *APOBEC3A* mRNA levels by digital droplet PCR (ddPCR) after siAPOBEC3A pool treatment (Figure

6.6B). Based on these preliminary results, I considered APOBEC3A to be a strong candidate ribosome biogenesis factor worthy of further investigation.

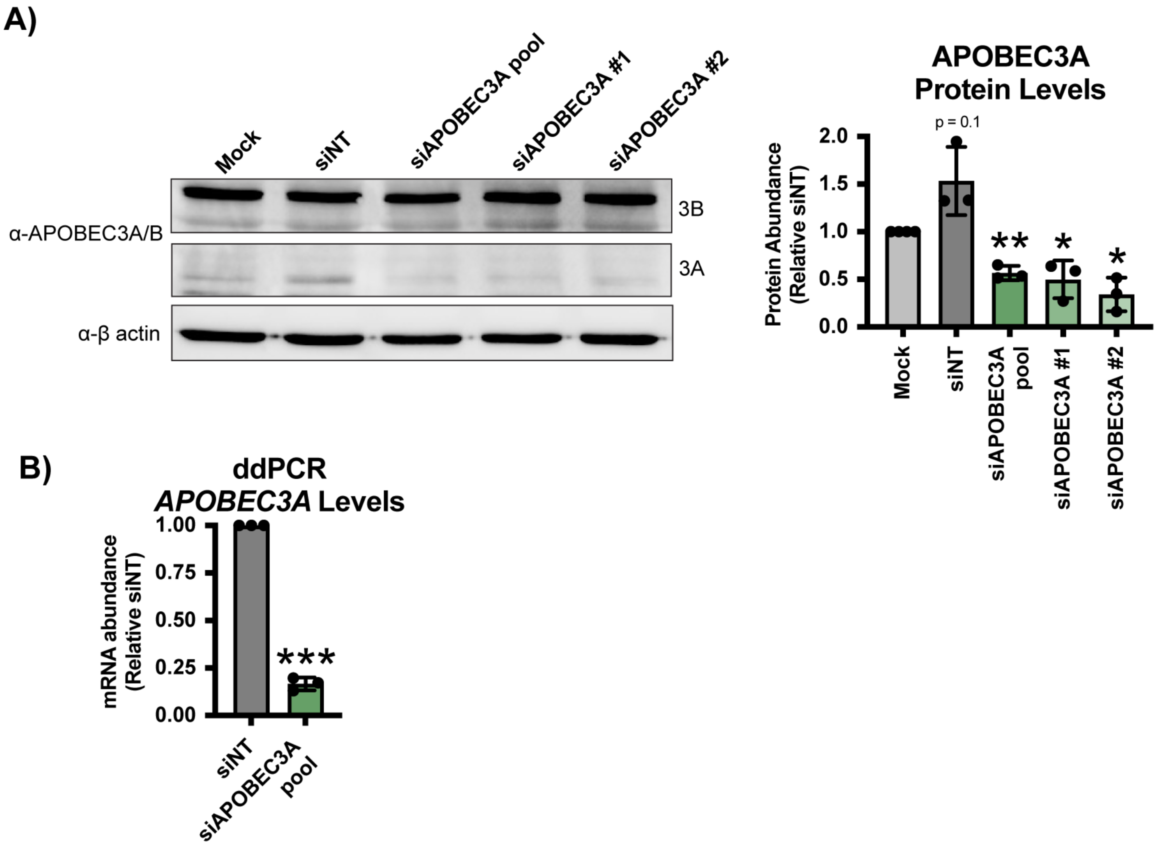


Figure 6.6: APOBEC3A targeting siRNAs efficiently deplete APOBEC3A on protein and mRNA level. **(A)** Non-targeting (siNT) siRNA treatment increases APOBEC3A protein levels, while APOBEC3A siRNAs decrease APOBEC3A protein levels in MCF10A cells. (Left) Representative western blot using an α-APOBEC3A/B antibody. α-β-actin is as a loading control. (Right) Quantification of APOBEC3A protein levels normalized to β-actin signal and relative to Mock. Three or 4 biological replicates plotted mean ± SD. Data were analyzed by Student's t-test * $p \leq 0.05$, ** $p \leq 0.01$. **(B)** Digital droplet PCR (ddPCR) confirming *APOBEC3A* mRNA depletion after siAPOBEC3A pool treatment. ddPCR measuring *APOBEC3A* mRNA levels normalized to *ACTB* internal control and relative to siNT negative control. 3 technical replicates of 3 biological replicates, plotted mean ± SD. Data were analyzed by Student's t-test *** $p \leq 0.001$.

APOBEC3A is required for cell cycle progression and protein synthesis

Because APOBEC3A's acute depletion for 72 h led to reduced cell viability, I predicted that it would also result in the dysregulation of cell cycle progression. Ribosome biogenesis and the cell cycle are interconnected, (Visintin and Amon 2000; Boisvert et al. 2007) with dysregulation of both having consequences in cancer pathogenesis (Derenzini et al. 2017). Using images collected from the screening assay, Carson Bryant and I quantified DNA content from Hoechst staining of siNT and siAPOBEC3A treated MCF10A cells as in (Chan et al. 2013; Ogawa et al. 2021). APOBEC3A siRNA depleted cells displayed a sharp increase in S-phase cells and an accompanying decrease in sub-G1 and G1 phase cells compared to siNT treated cells (Figure 6.7A).

To connect APOBEC3A's potential role in ribosome biogenesis with the cell cycle, I tested for the induction of the nucleolar stress response (Rubbi and Milner 2003; Boulon et al. 2010; Hannan et al. 2022) after APOBEC3A siRNA depletion. The nucleolar stress response results in a change in nucleolar morphology, increased levels of TP53 and its transcriptional target CDKN1A (el-Deiry et al. 1994) and ultimately cell cycle arrest and apoptosis. I tested for an induced nucleolar stress response by measuring TP53 and CDKN1A protein levels by western blotting after siAPOBEC3A treatment in MCF10A cells, which have wild-type p53 (Soule et al. 1990; Weiss et al. 2010). I observed both a significant increase in TP53 and CDKN1A levels in siAPOBEC3A treated cells compared to siNT treated cells (Figure 6.7B). These results agree with a genome-wide RNAi screen in A549 cells to identify proteins whose depletion leads to TP53 stabilization, where APOBEC3A was a top hit (Table 5.1) (Hannan et al. 2022). Taken together, these results indicate that acute depletion of

APOBEC3A induces the nucleolar stress response which leads to inhibition of cell cycle progression and to reduced cell viability.

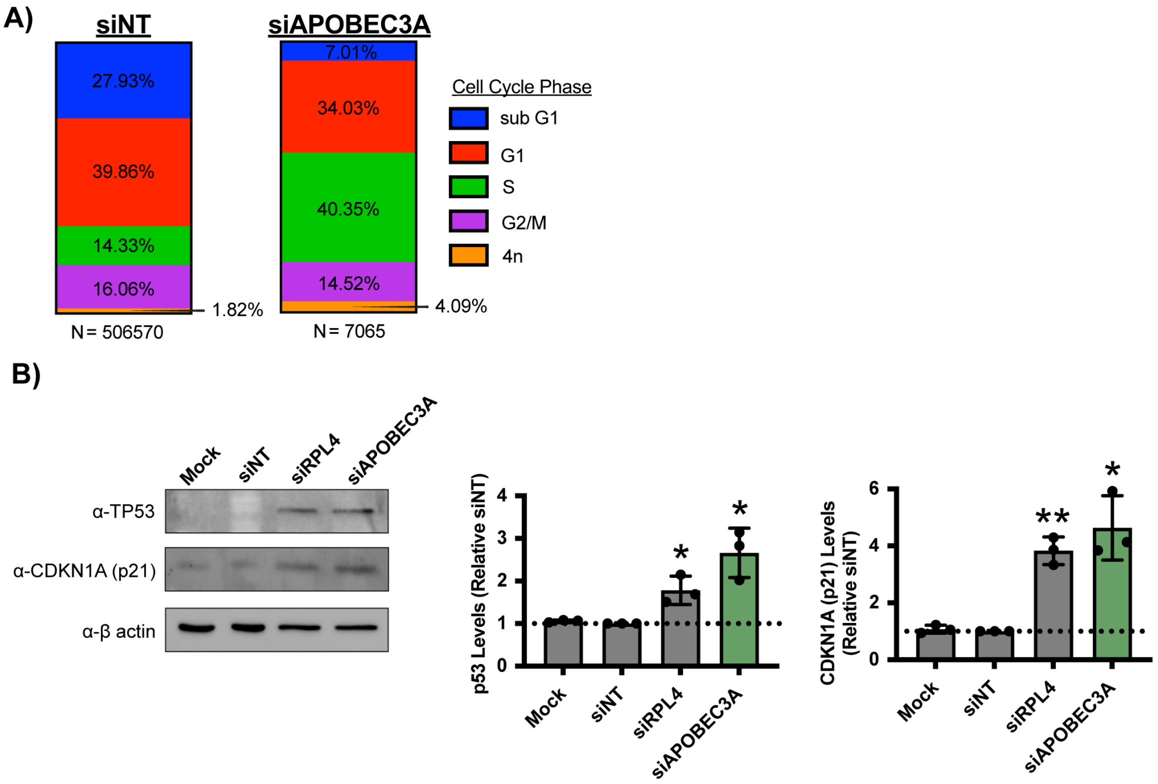


Figure 6.7: APOBEC3A siRNA depletion inhibits cell cycle progression and induces the nucleolar stress response. **(A)** siAPOBEC3A depletion (siRNA pool) leads to changes in cell cycle distribution in MCF10A cells. DNA intensity was measured by Hoechst staining (from images in Figure 1C) and the Log_2 DNA intensities were normalized to siNT negative control where its G1 peak = 1.0 and G2 peak = 2.0. Cell cycle phases were determined by the following normalized Log_2 integrated intensities. Sub G1 < 0.75, G1 = 0.75-1.26, S = 1.25-1.75, G2/M = 1.75-2.5, >4n > 2.5. Data was plotted as relative proportion of cells within each phase indicated by color. N = number of cells. **(B)** siAPOBEC3A depletion (pool) induces the nucleolar stress response in MCF10A cells. (Left) Representative western blot using α-TP53 and α-CDKN1A (p21) antibodies. Mock and siNT are negative controls. α-β-actin is as a loading control. Right) Quantification of TP53 and CDKN1A (p21) protein levels normalized to α-β-actin signal and relative to siNT negative control. 3 biological replicates plotted mean ± SD. Data were analyzed by Student's t-test ** p ≤ 0.01, * p ≤ 0.05.

Since the nucleolar stress response is caused by a reduction in mature ribosome levels (Rubbi and Milner 2003; Boulon et al. 2010), I predicted that APOBEC3A siRNA depletion would result in a reduced pool of functional ribosomes, and thus decreased protein synthesis. I used a puromycin incorporation assay (Schmidt et al. 2009) to measure changes in global protein synthesis upon siAPOBEC3A treatment of MCF10A cells. After siRNA depletion, 1 μM puromycin was added, which is incorporated into nascent polypeptides, for 1 h. I measured puromycin incorporation over this time period by western blotting with α-puromycin antibodies and observed a significant decrease in global protein synthesis after APOBEC3A depletion. The reduction in protein synthesis upon APOBEC3A depletion was similar in extent to depletion of the positive control, an LSU ribosomal protein, RPL4 (L4) (Figure 6.8A). Furthermore, this result was recapitulated using both individual siRNAs that target APOBEC3A (#1 and #2), which both resulted in a significant reduction in protein synthesis when transfected into MCF10A cells (Figure 6.8B).

Ribosome biogenesis is required across cell types. Therefore, I expected that APOBEC3A's function in making ribosomes would be conserved to other cell types as well. I confirmed APOBEC3A's overall function in making ribosomes in HeLa cells. Because MCF10A cells are not a cancer cell line, Hela cells provide us with a more direct test of APOBEC3A's role in a cancer context. Moreover, APOBEC3A has been shown to be a probable driver of cervical cancer pathogenesis [Reviewed in (Warren et al. 2017; Revathidevi et al. 2021)]. As I observed in MCF10A cells, I was able to confirm both decreased APOBEC3A protein levels and global protein synthesis after siAPOBEC3A treatment compared to siNT in HeLa cells (Figure 6.8C, D). Taken together, my results show that APOBEC3A is required for cell cycle progression, most likely due to in part to its depletion leading to the nucleolar stress response and decreased translation.

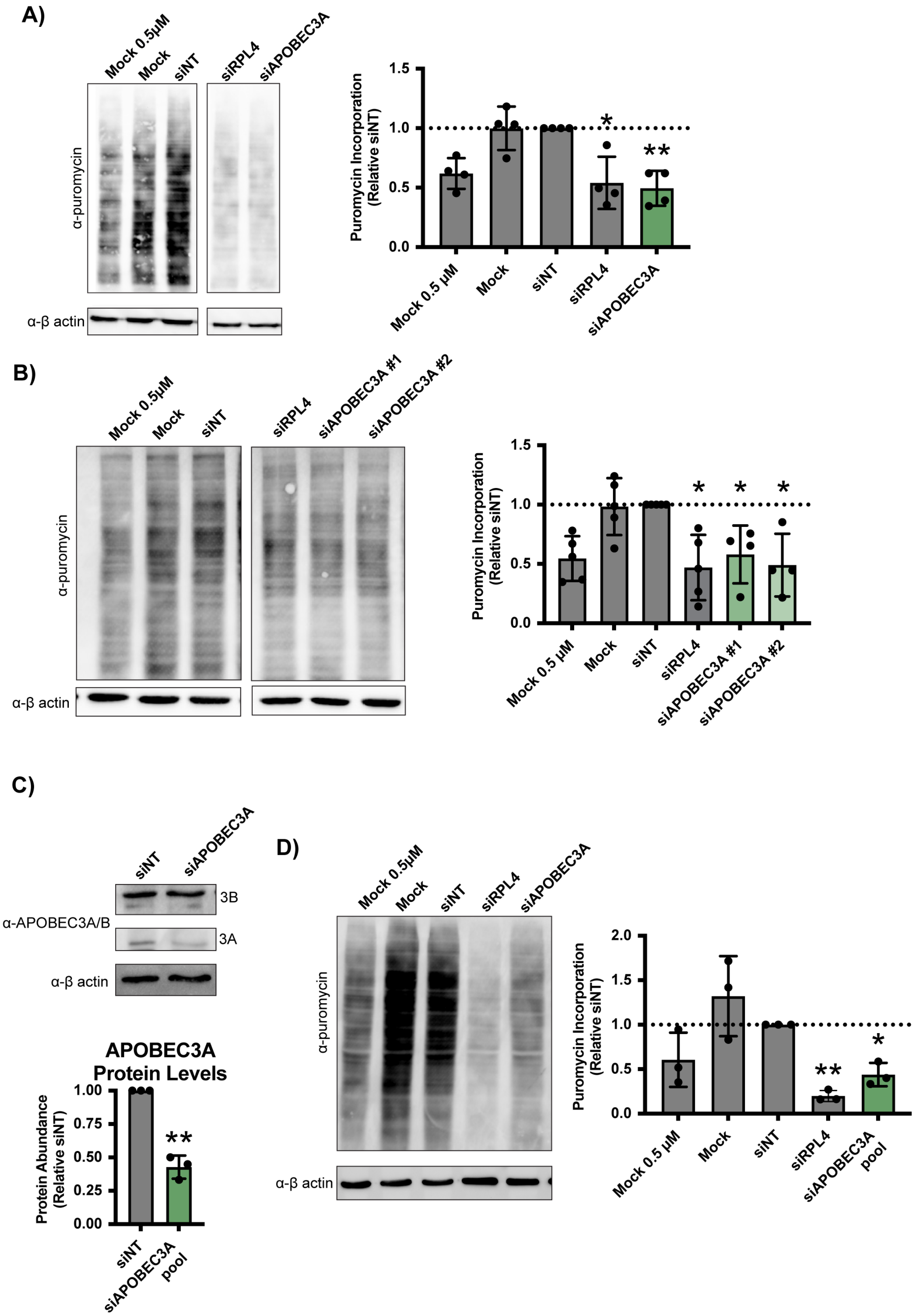

A)
Mock 0.5μM
Mock
siNT
siRPL4
siAPOBEC3A
α-puromycin
α-β actin
Puromycin Incorporation (Relative siNT)
1.5
1.0
0.5
0.0
*
**
Mock 0.5 μM
Mock
siNT
siRPL4
siAPOBEC3A
B)
Mock 0.5μM
Mock
siNT
siRPL4
siAPOBEC3A #1
siAPOBEC3A #2
α-puromycin
α-β actin
Puromycin Incorporation (Relative siNT)
1.5
1.0
0.5
0.0
*
*
*
Mock 0.5 μM
Mock
siNT
siRPL4
siAPOBEC3A #1
siAPOBEC3A #2
C)
siNT
siAPOBEC3A
3B
α-APOBEC3A/B
3A
α-β actin
APOBEC3A Protein Levels
Protein Abundance (Relative siNT)
1.00
0.75
0.50
0.25
0.00
**
siNT
siAPOBEC3A pool
D)
Mock 0.5μM
Mock
siNT
siRPL4
siAPOBEC3A
α-puromycin
α-β actin
Puromycin Incorporation (Relative siNT)
2.0
1.5
1.0
0.5
0.0
**
*
Mock 0.5 μM
Mock
siNT
siRPL4
siAPOBEC3A pool

Figure 6.8: siAPOBEC3A depletion (pool) reduces global protein synthesis in MCF10A and HeLa cells. **(A)** siAPOBEC3A depletion (pool) reduces global protein synthesis in MCF10A cells. After 72 h siRNA depletion, 1μM puromycin was added for 1 h to measure translation. (Left) Representative western blot using an α-puromycin antibody. Mock and siNT are negative controls, siRPL4 is a positive control, and Mock 0.5μM is a control to indicate robust quantification. α-β-actin is as a loading control. (Right) Quantification of puromycin signal normalized to β-actin signal and relative to siNT negative control. 4 biological replicates plotted mean ± SD. Data were analyzed by one-way ANOVA with Dunnett's multiple comparisons test, ** $p \leq 0.01$, * $p \leq 0.05$. **(B)** Depletion with individual siAPOBEC3A #1 and #2 reduces global protein synthesis in MCF10A cells. After 72 h siRNA depletion, 1μM puromycin was added for 1 h to measure translation. (Left) Representative western blot using an α-puromycin antibody. Mock and siNT are negative controls, siRPL4 is a positive control, and Mock 0.5μM is a control to indicate robust quantification. α-β-actin is as a loading control. (Right) Quantification of puromycin signal normalized to β-actin signal and relative to siNT negative control. Four biological replicates were plotted mean ± SD. Data were analyzed by one-way ANOVA with Dunnett's multiple comparisons test, * $p \leq 0.05$. **(C)** Confirming APOBEC3A protein depletion after siAPOBEC3A pool treatment in HeLa cells. (Top) Representative western blot using an α-APOBEC3A/B antibody. siNT is a negative control. α-β-actin is as a loading control. (Bottom) Quantification of APOBEC3A protein levels normalized to β-actin signal and relative to siNT negative control. Three biological replicates plotted mean ± SD. Data were analyzed by Student's t-test ** $p \leq 0.01$. **(D)** siAPOBEC3A pool treatment reduces global protein synthesis in HeLa cells. After 72 h siRNA depletion, 1μM puromycin was added for 1 h to measure translation. (Left) Representative western blot using an α-puromycin antibody. Mock and siNT are negative controls, siRPL4 is a positive control, and Mock 0.5μM is a control to indicate robust quantification. α-β-actin is as a loading control. (Right) Quantification of puromycin signal normalized to β-actin signal and relative to siNT negative control. Three biological replicates plotted mean ± SD. Data were analyzed by one-way ANOVA with Dunnett's multiple comparisons test, ** $p \leq 0.01$, * $p \leq 0.05$.

APOBEC3A is required for nucleolar rRNA biogenesis by regulating steps downstream of rRNA transcription

To gain insight into APOBEC3A's role in making ribosomes, I utilized an assay previously developed in our laboratory by Carson Bryant (Chapter 3) to measure nascent nucleolar rRNA biogenesis (Bryant et al. 2022). Briefly, after siRNA knockdown, MCF10A cells were treated with 5-ethynyl uridine (5-EU) for 1 h followed by biocompatible click chemistry to visualize the 5-EU signal. By co-staining with the nucleolar marker fibrillarin, I can then measure the 5-EU signal residing within the nucleolus. This readout combines both rRNA transcription and the stability of nucleolar pre-rRNA. I found that siAPOBEC3A treated cells displayed a modest decrease in nucleolar 5-EU signal, corresponding to a 50.8% nucleolar rRNA biogenesis percent inhibition relative to the negative control siNT (0% inhibition) and positive control siPOLR1A (the largest subunit of RNAP1, 100% inhibition) (Figure 6.9A, B, Appendix I). Based on our previously published results testing established ribosome biogenesis factors using this assay (Bryant et al. 2022), depletion of factors required for pre-rRNA transcription or both pre-rRNA transcription and processing result in higher nucleolar rRNA biogenesis percent inhibition values of greater than ~80%. Factors only required in pre-rRNA transcription or maturation result in lower percent inhibition values ~50-80%. Thus, it is likely that APOBEC3A is not required for pre-rRNA transcription but plays a role somewhere downstream in the ribosome biogenesis pathway.

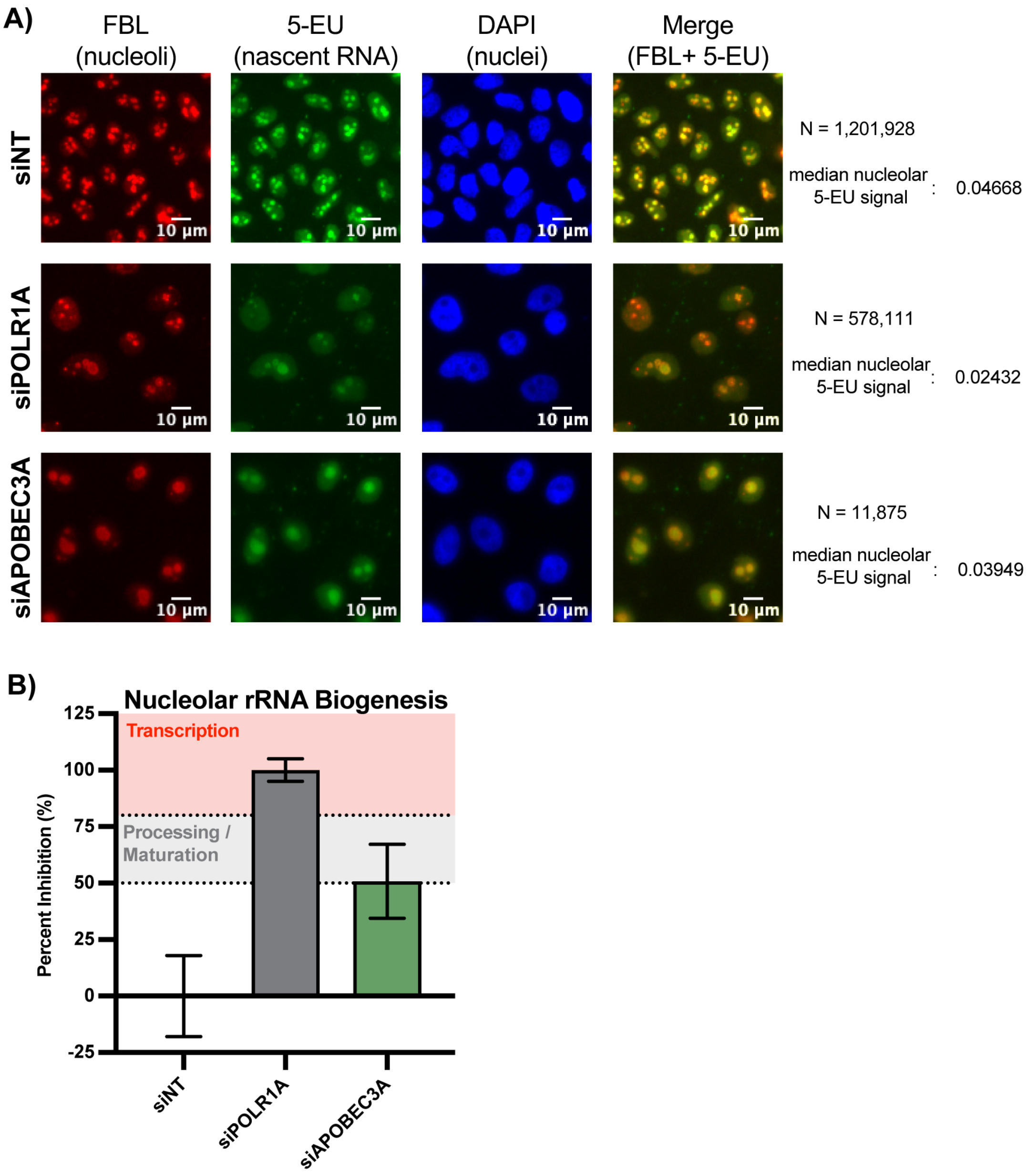


Figure 6.9: siAPOBEC3A depletion (pool) modestly reduces nucleolar rRNA biogenesis in MCF10A cells. **(A)** After 72 h siRNA depletion, 1mM 5-EU was added for 1 h to measure nucleolar rRNA biogenesis. Representative images of nucleoli stained with α-fibrillarin (FBL, red), 5-ethynl uridine (5-EU) visualized by click-chemistry attached AF488 azide (green), nuclei stained with DAPI (blue), and FBL and 5-EU merged (yellow). siNT is a negative control and siPOLR1A (large subunit of RNA polymerase I) is a positive control. N = number of nuclei (cells) analyzed in all 3 biological replicates and median nucleolar 5-EU signal (nucleolar rRNA biogenesis) reported. **(B)** Quantification of nucleolar rRNA biogenesis percent inhibition. siNT negative control is set to 0% and siPOLR1A positive control is set to 100%. Based on results in Bryant et al 2022(Bryant et al. 2022), factors required for pre-rRNA transcription have a percent inhibition > ~80% (red background) and factors only required for pre-rRNA processing /

maturation have a percent inhibition ~50-80% (grey background). Three biological replicates are plotted mean ± SD.

I utilized multiple assays that are more direct readouts of pre-rRNA transcription to confirm that APOBEC3A is not required for the first step in making ribosomes. In one approach, I used qRT-PCR to measure steady state levels of the primary pre-rRNA transcript in MCF10A cells depleted of APOBEC3A. While RNAP1 transcribes the 47S pre-rRNA precursor, immediate cleavage at site A' yields the more readably detectable 45S pre-rRNA precursor, so this assay is measuring both 47S and 45S levels (Figure 6.10A) (Sloan et al. 2014; Woolnough et al. 2016). siAPOBEC3A treatment did not significantly alter 47S/45S pre-rRNA levels compared to the negative control siNT treatment; however, the positive control siPOLR1A treatment did significantly reduce these levels (Figure 6.10B).

As another readout of pre-rRNA transcription, I used a dual-luciferase reporter (Ghoshal et al. 2004) to test for changes in rDNA promoter activity upon APOBEC3A depletion. After siRNA depletion, the firefly luciferase reporter, under control of the rDNA promoter (-410 to +327), is measured relative to a constitutive *Renilla* luciferase control 24 h post plasmid transfection (Figure 6.10C) (Ghoshal et al. 2004). siAPOBEC3A treatment did not significantly reduce rDNA promoter activity compared to siNT treatment, while the positive control, siPOLR1A treatment, did significantly reduce rDNA promoter activity (Figure 6.10D). Collectively, APOBEC3A is likely required for ribosome biogenesis during step(s) downstream of pre-rRNA transcription.

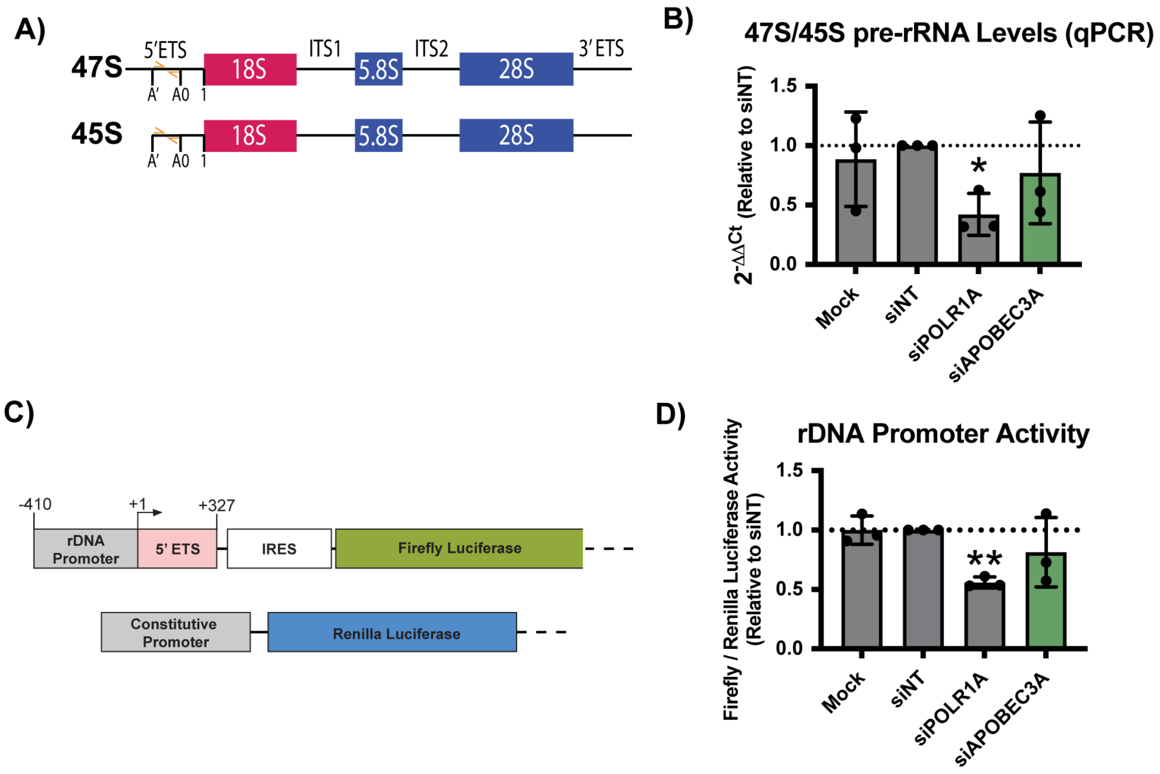
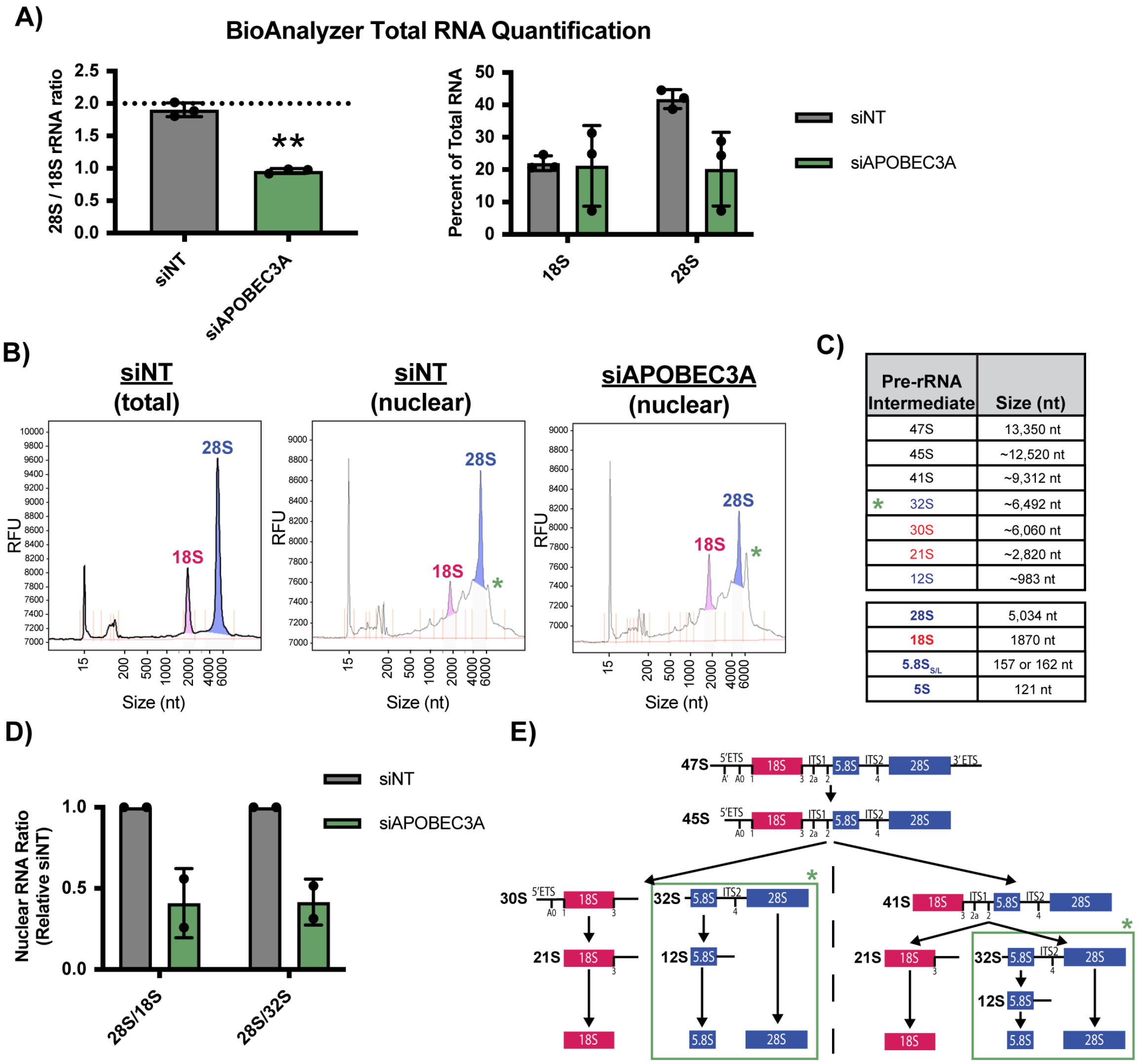


Figure 6.10: APOBEC3A is not required for rDNA transcription in MCF10A cells. **(A)** Schematic of 47S/45S pre-rRNA transcript measured by qRT-PCR in (D) using the indicated primers (orange). **(B)** siAPOBEC3A depletion (pool) does not significantly reduce 47S/45S pre-rRNA transcript levels. qRT-PCR was performed to measure primary 47/45S pre-rRNA transcript levels. Mock and siNT are negative controls and siPOLR1A is a positive control. $2^{-\Delta\Delta CT}$ measured relative to 7SL internal control and siNT negative control. Three technical replicates of 3 biological replicates plotted mean ± SD. Data were analyzed by Student's t-test, * $p \leq 0.05$. **(C)** Schematic of luciferase reporter plasmids to measure rDNA promoter activity (Ghoshal et al. 2004). (Top) Firefly pHrD-IRES-Luc rDNA promoter reporter plasmid. (Bottom) *Renilla* constitutive promoter, transfection control. **(D)** siAPOBEC3A depletion (pool) does not significantly reduce rDNA promoter activity in MCF10A cells. After 48 h of siRNA depletion, plasmids in (E) were transfected for 24 h for a total of 72 h siAPOBEC3A depletion. Firefly luminescence was measured relative to *Renilla* luminescence and the siNT negative control. siPOLR1A (the large subunit of RNA polymerase I) is a positive control. Three biological replicates plotted mean ± SD. Data were analyzed by Student's t-test, ** $p \leq 0.01$.

APOBEC3A is required for pre-LSU rRNA processing and maturation within the nucleus

To reveal if one ribosomal subunit was preferentially affected by APOBEC3A siRNA depletion, I quantified the levels of mature 18S (SSU) and 28S (LSU) rRNAs on an Agilent Bioanalyzer. I used total RNA to quantify both a 28S/18S rRNA ratio and overall rRNA levels in siAPOBEC3A treated MCF10A cells. MCF10A cells normally have a 28S/18S ratio of ~2 in siNT treated cells, which was significantly decreased 2-fold after APOBEC3A siRNA depletion (Figure 6.11A, left). This decreased ratio can be attributed to reduced 28S rRNA levels specifically and not to changes in 18S rRNA levels (Figure 6.11A, right). The observed reduction in 28S rRNA levels points to APOBEC3A playing a role in LSU pre-rRNA processing and/or maturation.

I also harvested nuclear RNA from MCF10A cells to test if siAPOBEC3A mediated 28S rRNA maturation defects occur before or after ribosome subunit export out of the nucleus. BioAnalyzer analysis of total RNA picks up almost exclusively mature rRNAs, which are in the cytoplasm (Figure 6.11B, left). As I had anticipated, the nuclear BioAnalyzer traces have a much greater proportion of large sized RNAs detected around the 18S and 28S mature rRNA peaks compared to total RNA analysis. Presumably, many of these correspond to pre-rRNA intermediate species (Figure 6.11B, C). I quantified these results by taking the 28S/18S rRNA ratio which, similar to the results with total RNA (Figure 4A), was reduced again ~2-fold in APOBEC3A depleted MCF10A cells compared to siNT treated cells (Figure 6.11D, left). Therefore, siAPOBEC3A treatment of MCF10A cells results in an LSU specific maturation defect within the nucleus. Aside from the expected changes in 28S abundance, one of the most striking differences between the siNT and siAPOBEC3A nuclear RNA BioAnalyzer traces was an increase in abundance of an RNA species around ~6,500 nucleotides in size. This corresponds to the size of the 32S pre-rRNA intermediate, a precursor to the mature 28S and 5.8S LSU rRNAs (Figure 6.11E) (Henras et al. 2015). I quantified the 28S/32S ratio and observed a ~2-fold decrease in APOBEC3A depleted cells compared to the siNT negative control (Figure 6.11D, right). This decreased ratio indicates a pre-rRNA processing defect where there is an aberrant build-up of the 32S pre-rRNA precursor and a subsequent decrease in mature 28S rRNA levels (Figure 6.11E), which I confirmed using northern blotting.



Pre-rRNA Intermediate	Size (nt)
47S	13,350 nt
45S	~12,520 nt
41S	~9,312 nt
* 32S	~6,492 nt
30S	~6,060 nt
21S	~2,820 nt
12S	~983 nt
28S	5,034 nt
18S	1870 nt
$5.8S_{S/L}$	157 or 162 nt
5S	121 nt

Figure 6.11: APOBEC3A is required for pre-LSU (28S rRNA) maturation prior to nuclear export. **(A)** siAPOBEC3A depletion (pool) reduces mature 28S rRNA levels in MCF10A cells. Agilent BioAnalyzer total RNA quantification. (Left) 28S/18S mature rRNA ratio with siNT as a negative control. Expected normal 28S/18S rRNA ratio of ~2.0 (dashed horizontal line). (Right) Percent of total RNA levels for 18S and 28S rRNA with siNT as a negative control. Three biological replicates plotted mean ± SD. Data were analyzed by Student's t-test, ** $p \leq 0.01$. **(B)** Nuclear RNA is enriched for the pre-rRNAs and siAPOBEC3A depletion (pool) increases the abundance of a ~6.5kb RNA species (32S pre-rRNA) in the nucleus of MCF10A cells. Representative Agilent BioAnalyzer traces from total cell RNA (left) and nuclear fraction RNA (middle, right). Green star indicates an apparent nuclear RNA species (~6.5 kb) with increased abundance upon siAPOBEC3A treatment compared to the siNT negative control. **(C)** Table of human pre-rRNA intermediate and mature rRNA species and their approximate length in nucleotides (nt). Green star indicates the 32S pre-rRNA is increased after siAPOBEC3A depletion. **(D)** siAPOBEC3A depletion (pool) reduces nuclear 28S/18S and 28S/32S pre-

rRNA ratios. Quantification of nuclear RNA Agilent BioAnalyzer traces. pre-rRNA ratios are reported relative to the siNT negative control. Two biological replicates are plotted mean ± SD. **(E)** Diagram of pre-rRNA processing in humans. The external transcribed spacer (ETS) and internal transcribed spacer (ITS) sequences are removed to yield the mature 18S, 5.8S, and 28S rRNAs. The green box indicates the expected defect in processing after APOBEC3A depletion, where ITS2 processing is reduced leading to an increase in 32S levels and a decrease in 12S levels.

I performed pre-rRNA northern blotting to substantiate my claim that pre-28S processing defects were occurring after APOBEC3A siRNA depletion in MCF10A cells. I used an internal transcribed spacer sequence 2 (ITS2) probe to measure steady-state levels of the pre-rRNA precursors leading to the maturation of the 5.8S and 28S rRNAs (Figure 6.11E). I observed a noticeable increase in the 32S pre-rRNA intermediate and corresponding decrease in the downstream 12S intermediate in siAPOBEC3A compared to siNT treatment (Figure 6.12A). These results were quantified using ratio analysis of multiple precursors (RAMP) (Wang et al. 2014), which confirmed a significant increase in the 32S precursor and a decrease in the 12S precursor (product of 32S) compared to all of their upstream precursors in the LSU processing pathway (Figure 6.11E, 5.12B). Agreeing with the BioAnalyzer results, methylene blue staining also indicated a decrease in mature 28S/18S rRNA ratio after APOBEC3A siRNA depletion (Figure 6.12C). I also analyzed processing using both individual siRNAs that target APOBEC3A (#1 and #2) and observed the same defect with siRNA #1 and a more modest trending increase in relative 32S levels with siRNA #2 (Figure 6.12D, E). More specifically with siAPOBEC3A #2, there was an increase in 12S levels relative to the 41S and 47S precursors, which still indicates a defect in pre-LSU processing. Only siRNA #1 and not siRNA #2 APOBEC3A depletion produced a decrease in the 28S/18S rRNA ratio by methylene blue staining (Figure 6.12F). Taken together, APOBEC3A is required for pre-LSU rRNA processing during ITS2 cleavage prior to nuclear export in MCF10A cells.

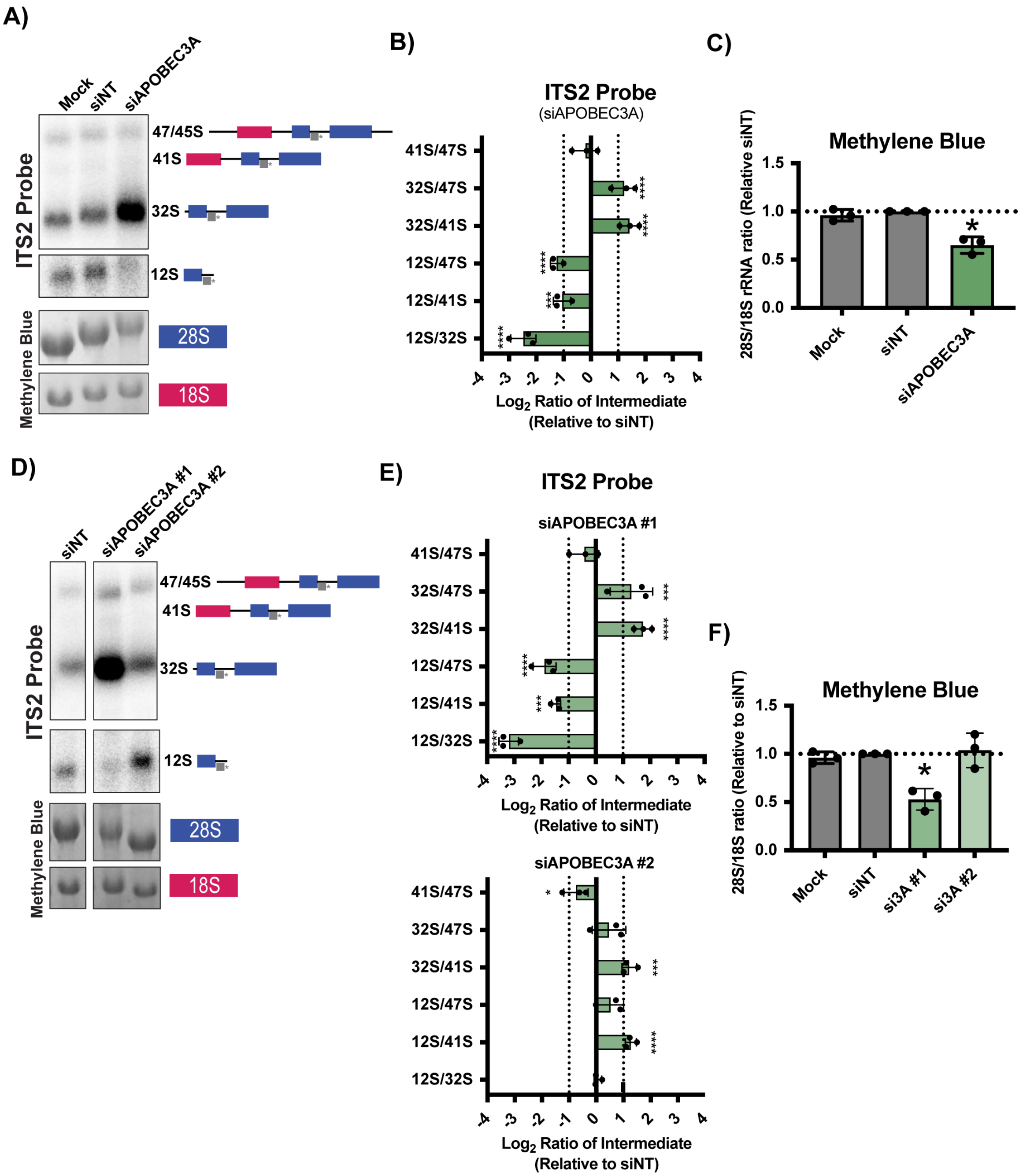


Figure 6.12: siAPOBEC3A depletion (pool) inhibits pre-LSU rRNA processing in ITS2 to reduce LSU maturation. **(A)** (Top) Processing of the 32S to the 12S pre-rRNA is reduced after siAPOBEC3A depletion. Representative northern blot using an ITS2 probe to measure steady-state levels of pre-rRNAs that lead to the formation of the mature 5.8S and 28S rRNAs (LSU). The detected pre-rRNAs are indicated along with the ITS2 probe (grey). (Bottom) Methylene blue staining of the mature 28S and 18S rRNAs as a loading control. Mock transfected and siNT are negative controls. (**B**) Quantification of defective pre-rRNA processing results in (F). Ratio analysis of multiple precursors (RAMP) (Wang et al. 2014) quantification of northern blotting using an ITS2 probe in (A). Log$_2$ fold change of pre-rRNA ratios were quantified relative to siNT negative control. Three biological replicates plotted mean ± SD. Data were analyzed by two-way ANOVA, **** p

≤ 0.0001, *** ≤ 0.001. **(C)** Quantification of mature rRNA levels in (A). Methylene blue staining quantification of mature 28S to 18S rRNA ratio relative to siNT negative control. Three biological replicates plotted mean ± SD. . Data were analyzed by Student's t-test, * $p \leq 0.05$. **(D-F)** Individual siAPOBEC3A #1 and #2 treatment inhibits pre-LSU processing of internal transcribed spacer 2 (ITS2) in MCF10A cells. **(D)** (Top) Representative northern blot using an ITS2 probe measuring steady-state levels of pre-rRNA precursors leading to the formation of the mature 5.8S and 28S rRNAs. Precursors detected indicated along with the ITS2 probe (grey). (Bottom) Methylene blue staining of the mature 28S and 18S rRNAs as a loading control. siNT is a negative control. **(E)** Ratio analysis of multiple precursors (RAMP)(Wang et al. 2014) quantification of northern blotting using an ITS2 probe. Log_2 fold change of pre-rRNA intermediate ratios were quantified relative to siNT negative control. 3 biological replicates plotted mean ± SD. Data were analyzed by two-way ANOVA, **** $p \leq 0.0001$, *** ≤ 0.001. **(F)** Quantification of mature rRNA levels in (D). Methylene blue staining quantification of mature 28S to 18S rRNA ratio relative to siNT negative control. Three biological replicates plotted mean ± SD. Data were analyzed by Student's t-test, * $p \leq 0.05$.

APOBEC3A has candidate editing target sites in the pre-LSU rRNA

The pre-rRNA/rRNA is by far the most abundant RNA in the cell and APOBEC3A has already established roles in editing much less abundant nucleic acids in a cell (Roberts et al. 2013; Niavarani et al. 2015; Sharma et al. 2015; Sharma et al. 2017; Alexandrov et al. 2020; Alqassim et al. 2021; Mertz et al. 2022). I hypothesized that APOBEC3A could also bind or edit the pre-rRNA, leading to a requirement for it in ribosome biogenesis. APOBEC3A has a UC dinucleotide RNA sequence motif preference, where the edited cytidine is preceded by a uracil (Kouno et al. 2017). I scanned the primary 47S pre-rRNA transcript (NR_145144.1) and 5S rRNA [E00204 (Szymanski et al. 2016)] consensus sequences for UC sequence motifs to identify candidate portions of the pre-rRNA that APOBEC3A could bind and/or modify. Based on examining UC sequence motif occurrence at 100 nucleotide intervals across the transcript and its enrichment over its expected occurrence at random (~6.25 per 100 nucleotides), I observed that this motif was more likely to occur within the ITS sequences compared to the mature rRNA sequences on the 47S pre-rRNA (Figure 6.13A). More specifically, there were sites of UC motif enrichment just 5' of cut sites 1 in the 5'ETS, cut site 2a in ITS1, and cut site 4 in ITS2, and an overall decrease below what would be predicted across the 18S rRNA. There were 7 UC sequence motifs identified on the 5S rRNA (Figure 6.13B). I kept these locations of UC motif enrichment in mind during our search for APOBEC3A binding and/or editing sites along the pre-rRNA, especially considering their proximity to important cleavage sites that produce the pre-LSU.

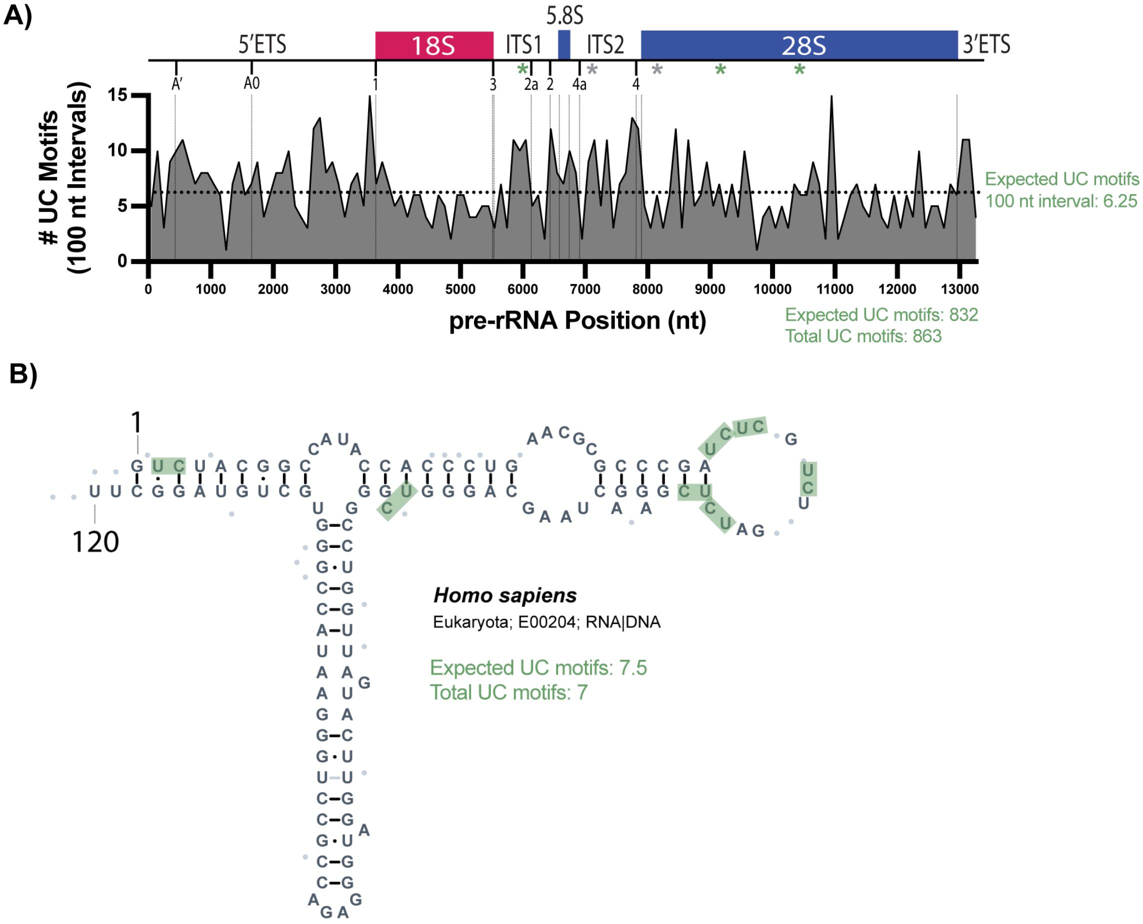
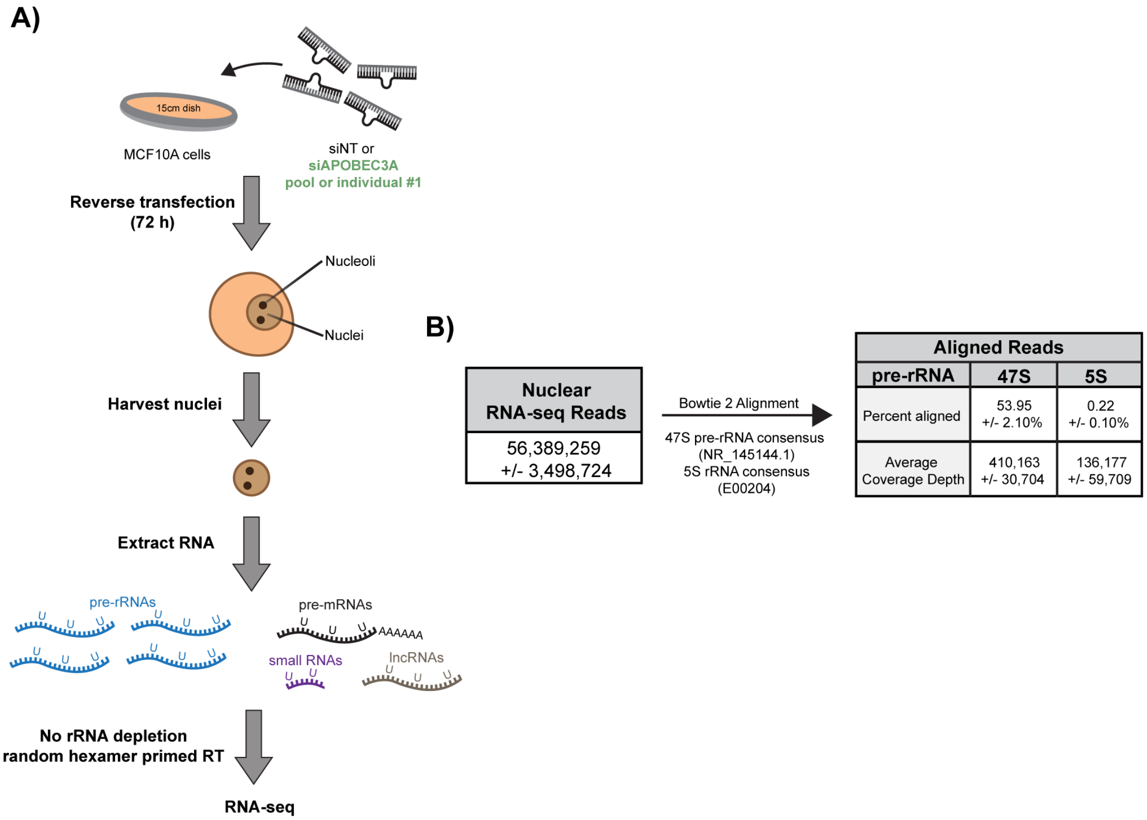


Figure 6.13: APOBEC3A's UC dinucleotide sequence motif is more frequently found in the external and internal transcribed spacer (ETS and ITS) sequences near pre-rRNA cleavage sites and less frequently found in the 18S rRNA coding sequence. **(A)** The occurrence of UC sequence motifs was calculated for every 100 nucleotide (nt) interval across the pre-rRNA and graphed. The occurrence of UC sequence motif at random over 100 nt is predicted to be 6.25, indicated by the dotted horizontal line. Pre-rRNA cleavage sites are indicated with vertical solid lines. **(B)** The human 5S rRNA contains 7 UC dinucleotide sequence motifs. Secondary structure of the human 5S rRNA (E00204) obtained from 5SrRNAdb.(Szymanski et al. 2016) UC sequence motifs highlighted in green. Expected number of UC sequence motifs reported based on occurrence at random of 6.25 per 100 nucleotides.

To identify candidate APOBEC3A target sites on the pre-rRNA, I performed a nuclear RNA-seq experiment followed by variant analysis comparing 3 samples (siNT and siAPOBEC3A (pool and siRNA #1)) treated MCF10A cells in duplicate. After siRNA treatment, I harvested nuclei, extracted RNA, and the Yale Center for Genomic Analysis performed a library prep without rRNA depletion and using random hexamer RT primers to enrich for pre-rRNA reads (Figure 6.14A). I obtained on average about 56 million reads and after alignment to both a 47S pre-rRNA (NR_145144.1) and 5S rRNA (E00204 (Szymanski et al. 2016)) consensus sequences, over half of our reads aligned to these transcripts of interest, indicating successful pre-rRNA enrichment (Figure 6.14B).

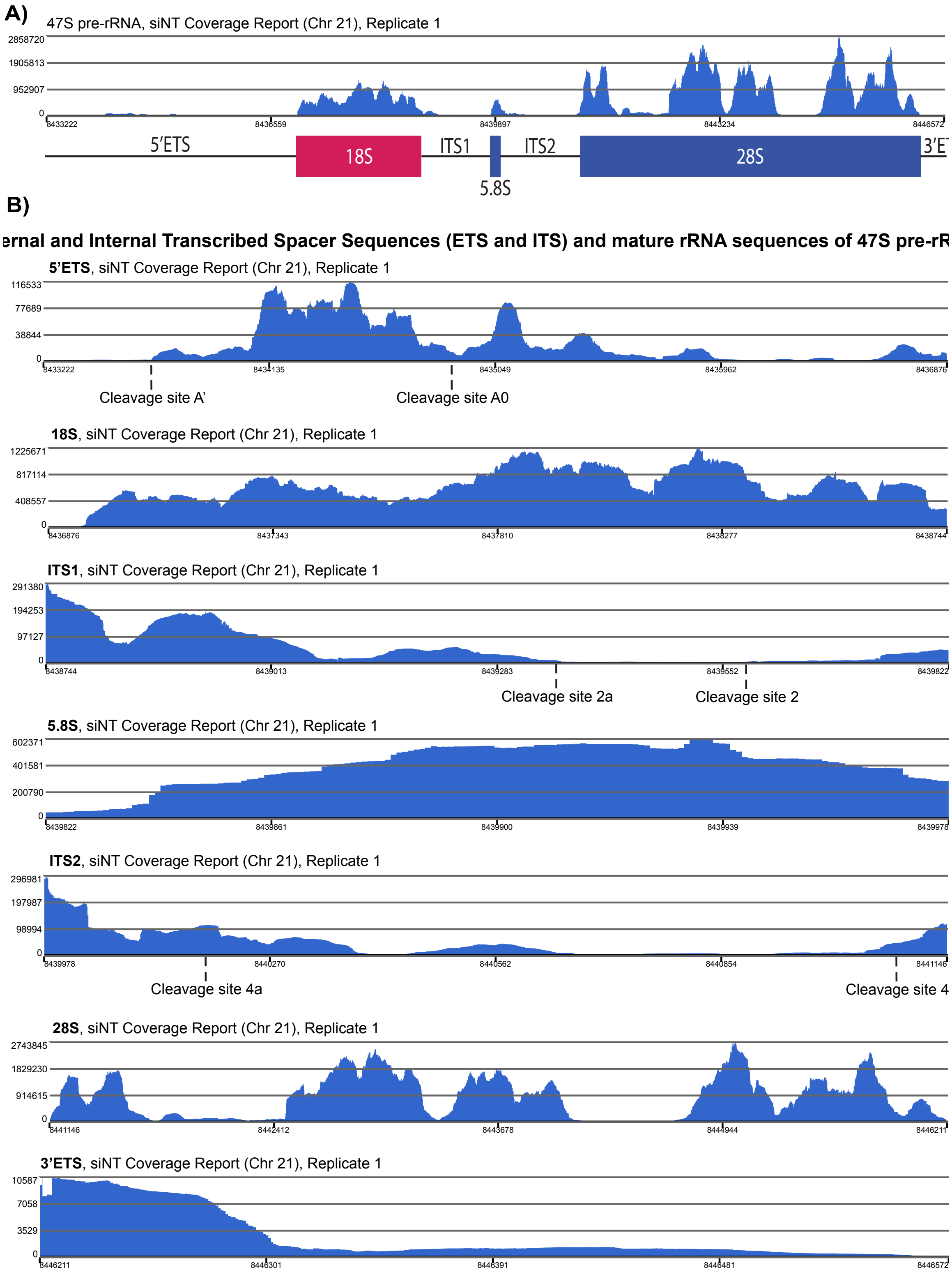


Aligned Reads		
pre-rRNA	47S	5S
Percent aligned	53.95 +/- 2.10%	0.22 +/- 0.10%
Average Coverage Depth	410,163 +/- 30,704	136,177 +/- 59,709

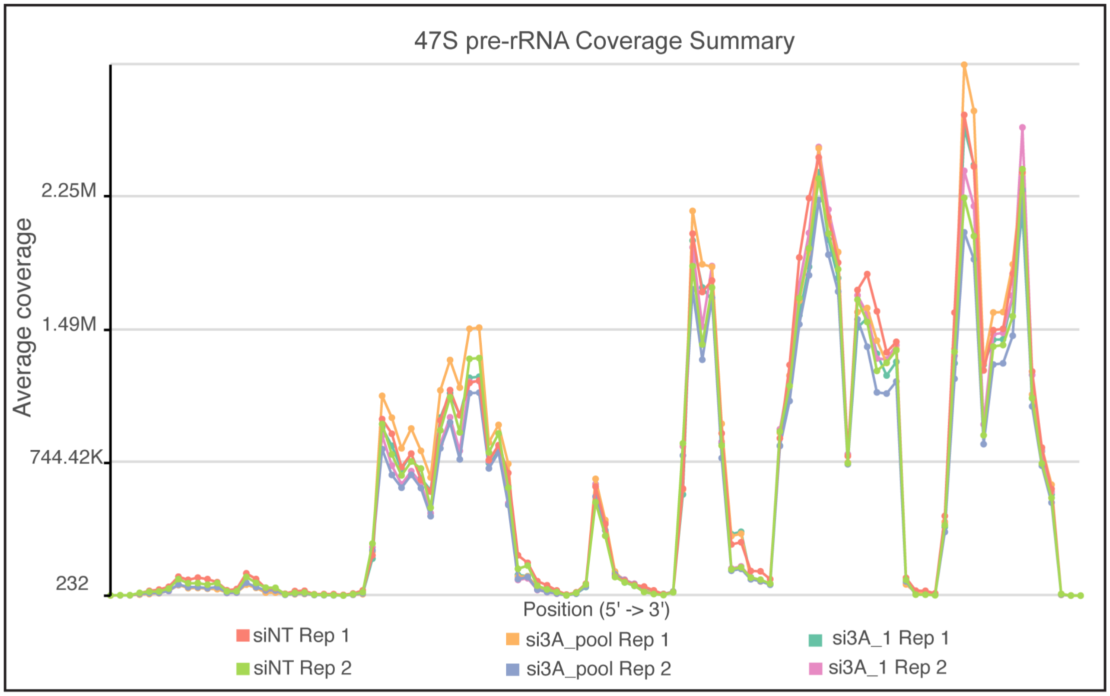
Figure 6.14: Nuclear RNA sequencing as a novel experiment to enrich for pre-rRNA reads. **(A)** Schematic of the nuclear RNA-seq experiment. MCF10A cells were treated with siNT negative control, siAPOBEC3A pool, or siAPOBEC3A individual #1 for 72 h. Nuclear RNA was extracted and submitted for sequencing with no rRNA depletion step and primed with random hexamers for reverse transcription (RT). **(B)** Over half of the nuclear RNA-seq reads aligned to the 47S pre-rRNA and to the 5S rRNA. Average reads, percent alignment, and coverage depth of all nuclear RNA-sequencing runs using Bowtie 2 alignment to either the 47S pre-rRNA (NR_145144.1) or 5S rRNA

(E00204(Szymanski et al. 2016)). Average of 6 biological replicates (2x siNT, 2x siAPOBEC3A pool, 2x siAPOBEC3A #1) ± SD.

Upon closer inspection, my read alignments were enriched for the mature rRNAs, but still had notable coverage across both the external and internal transcribed spacer sequences (Figure 6.15A, B). This coverage was consistent across siRNA treatments within replicates (Figure 6.15C, D).

A)
47S pre-rRNA, siNT Coverage Report (Chr 21), Replicate 1
2858720
1905813
952907
0
8433222
8436559
8439897
8443234
8446572
5'ETS
18S
ITS1
5.8S
ITS2
28S
3'E
B)
ernal and Internal Transcribed Spacer Sequences (ETS and ITS) and mature rRNA sequences of 47S pre-rR
5'ETS, siNT Coverage Report (Chr 21), Replicate 1
116533
77689
38844
0
8433222
8434135
8435049
8435962
8436876
Cleavage site A'
Cleavage site A0
18S, siNT Coverage Report (Chr 21), Replicate 1
1225671
817114
408557
0
8436876
8437343
8437810
8438277
8438744
ITS1, siNT Coverage Report (Chr 21), Replicate 1
291380
194253
97127
0
8438744
8439013
8439283
8439552
8439822
Cleavage site 2a
Cleavage site 2
5.8S, siNT Coverage Report (Chr 21), Replicate 1
602371
401581
200790
0
8439822
8439861
8439900
8439939
8439978
ITS2, siNT Coverage Report (Chr 21), Replicate 1
296981
197987
98994
0
8439978
8440270
8440562
8440854
8441146
Cleavage site 4a
Cleavage site 4
28S, siNT Coverage Report (Chr 21), Replicate 1
2743845
1829230
914615
0
8441146
8442412
8443678
8444944
8446211
3'ETS, siNT Coverage Report (Chr 21), Replicate 1
10587
7058
3529
0
8446211
8446301
8446391
8446481
8446572

C)


47S pre-rRNA Coverage Summary
Average coverage
2.25M
1.49M
744.42K
232
Position (5' -> 3')
siNT Rep 1
siNT Rep 2
si3A_pool Rep 1
si3A_pool Rep 2
si3A_1 Rep 1
si3A_1 Rep 2

D)

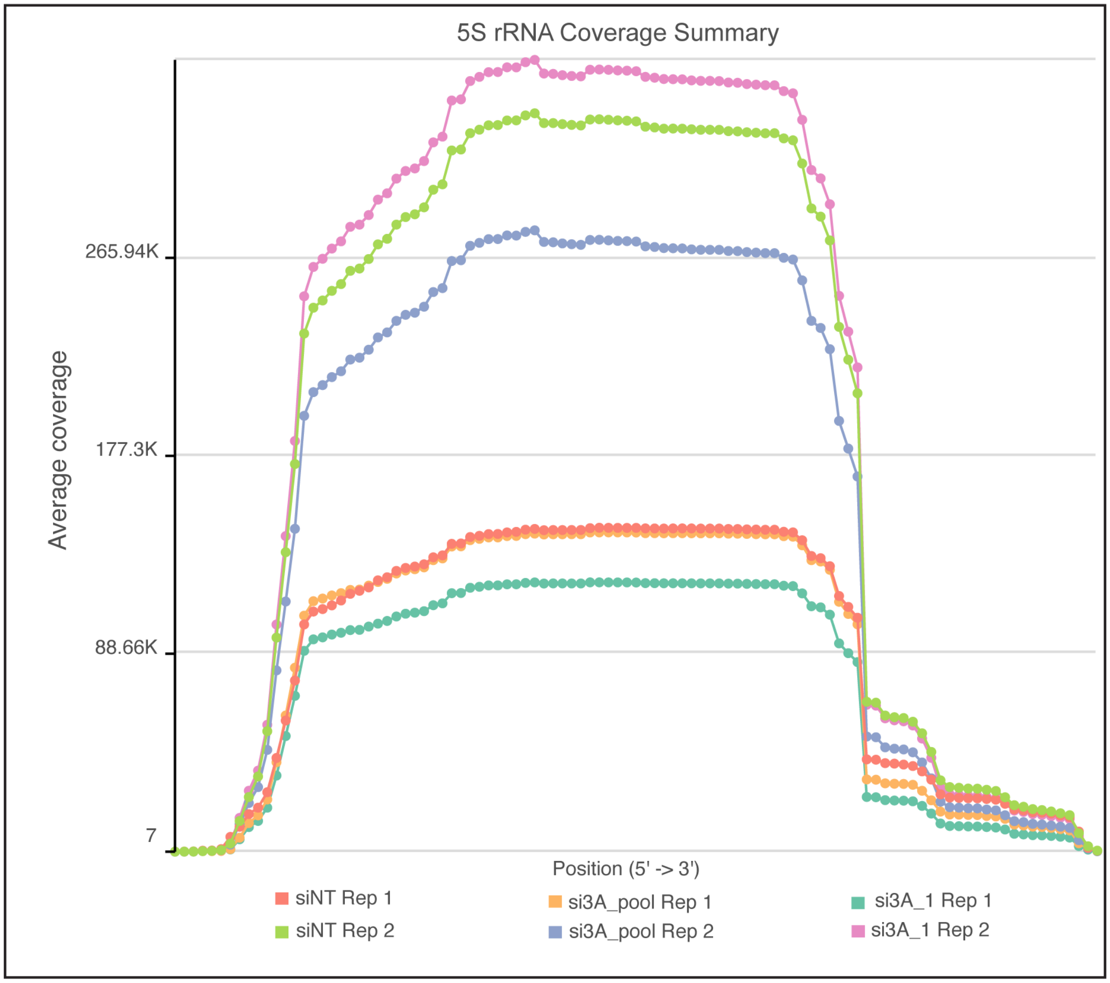
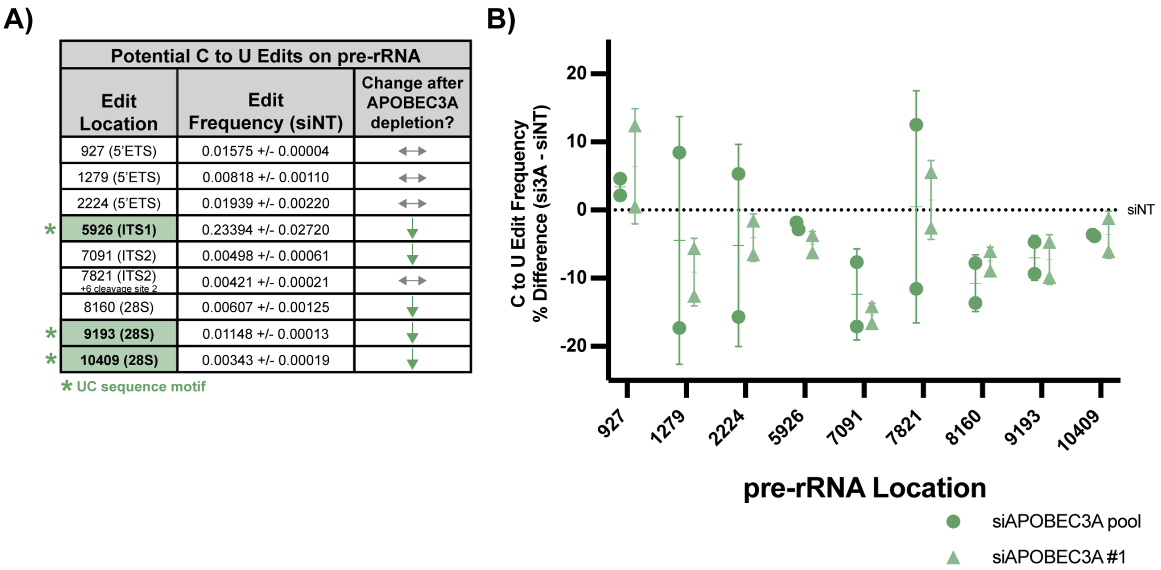

5S rRNA Coverage Summary
Average coverage
265.94K
177.3K
88.66K
7
Position (5' -> 3')
siNT Rep 1
siNT Rep 2
si3A_pool Rep 1
si3A_pool Rep 2
si3A_1 Rep 1
si3A_1 Rep 2

Figure 6.15: (A, B) Nuclear RNA-sequencing reads spanned the entire 47S pre-rRNA with enrichment in the mature rRNA sequences compared to the external and internal transcribed spacer pre-rRNA sequences. **(A)** Coverage report across the 47S pre-rRNA sequence (hg38, chromosome 21) after STAR (Version 2.7.8a) alignment to hg38. Y-axis is average read depth; x-axis is genomic coordinates on chromosome 21 corresponding to the location of the 47S pre-rRNA (RNA45SN1). Data is for nuclear RNA-seq of siNT negative control treated MCF10A cells replicate #1. Image was generated using Partek® Flow® chromosome viewer. **(B)** Same as in (A) except zoomed into the different portions of the 47S pre-rRNA transcript including the 5' and 3' external transcribed spacers (ETSs), internal transcribed spacers (ITSs) 1 and 2, and the mature (18S, 5.8S, and 28S) rRNAs. Cleavage sites indicated below the x-axis. **(C, D)** Nuclear RNA-sequencing read coverage of the 47S pre-rRNA and 5S rRNA was consistent across siRNA treatments within replicates. **(C)** Coverage summary of MCF10A cell nuclear RNA-seq reads aligned to a 47S pre-rRNA consensus sequence (NR_145144.1) for all replicates indicated by color. Y-axis is average coverage depth; x-axis is the entire 47S pre-rRNA sequence. Image was generated using Partek® Flow® chromosome viewer. **(D)** Same as in (C) except for the 5S rRNA (E00204(Szymanski et al. 2016)) sequence.

Using these datasets, I performed variant analysis using LoFreq (Wilm et al. 2012) to identify C to U edit sites on the pre-rRNA that decreased after APOBEC3A depletion. This is similar to what was done previously to assess rDNA variation across individuals and tissues (Parks et al. 2018). I identified 9 sites on the 47S pre-rRNA (none on the 5S rRNA), where a C to U edit was detected in both replicates in siNT treated samples with varying frequencies (0.34% - 23.39%). Out of the 9 identified edit sites, 5 had a decreased edit frequency in all the APOBEC3A depleted samples (2 replicates each of siAPOBEC3A pool and siAPOBEC3A #1), with 3 of those occurring in a UC sequence motif context (Figure 6.16A, Appendix II). Interestingly, all 5 of these sites occurred downstream of the 18S rRNA in locations corresponding to LSU processing intermediates or in the mature 28S rRNA. It is important to note that most of these edits were detected at a modest frequency and a correspondingly modest, but reproducible, reduction in editing frequency was observed upon APOBEC3A depletion (Figure 6.16B, Appendix II). Overall, this RNA-seq experiment revealed a handful of candidate APOBEC3A target locations on the 47S pre-rRNA that are sensitive to the presence of the APOBEC3A, and that APOBEC3A may edit.



A)

Potential C to U Edits on pre-rRNA		
Edit Location	Edit Frequency (siNT)	Change after APOBEC3A depletion?
927 (5'ETS)	0.01575 +/- 0.00004	↔
1279 (5'ETS)	0.00818 +/- 0.00110	↔
2224 (5'ETS)	0.01939 +/- 0.00220	↔
* **5926 (ITS1)**	0.23394 +/- 0.02720	↓
7091 (ITS2)	0.00498 +/- 0.00061	↓
7821 (ITS2) +6 cleavage site 2	0.00421 +/- 0.00021	↔
8160 (28S)	0.00607 +/- 0.00125	↓
* **9193 (28S)**	0.01148 +/- 0.00013	↓
* **10409 (28S)**	0.00343 +/- 0.00019	↓

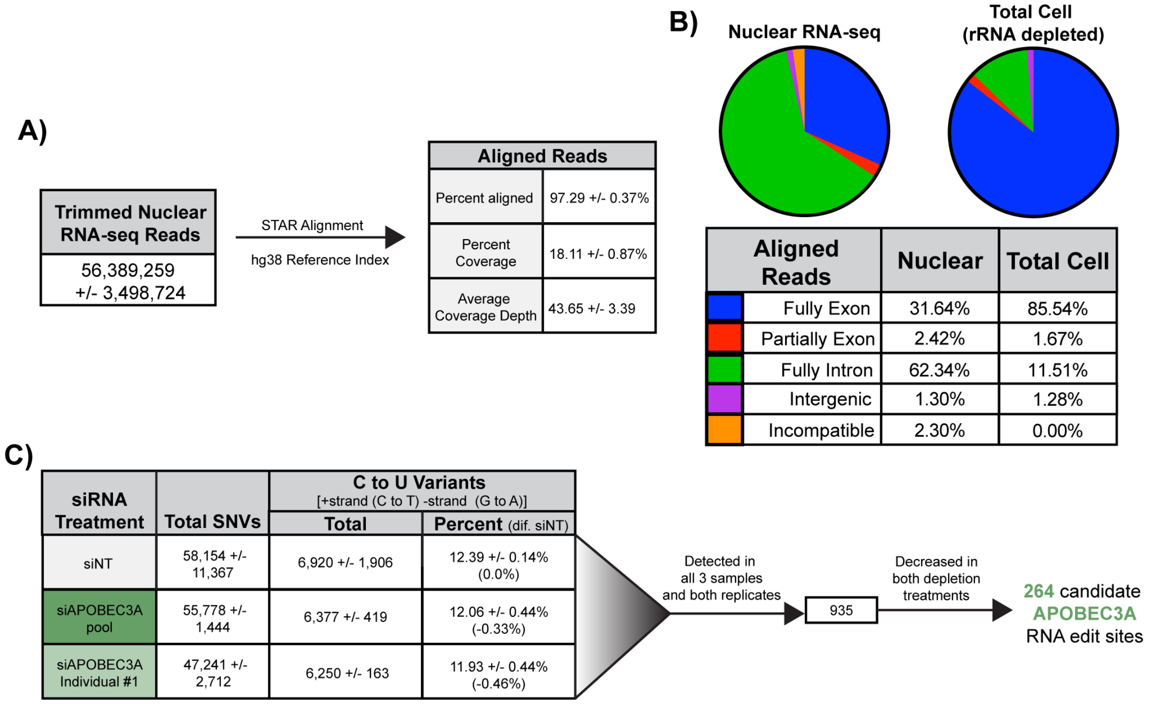
* UC sequence motif

Figure 6.16: Nuclear RNA-seq revealed 9 C to U edits on the pre-rRNA, 5 of which are candidate APOBEC3A edit sites. **(A)** LoFreq variant analysis of the nuclear RNA-seq datasets identified 5 candidate C to U APOBEC3A edit sites in the pre-rRNA (green down arrows). Three of these occur at a UC sequence motif (green highlighted squares). Table of 9 C to U edit locations on the pre-rRNA detected by LoFreq (see Methods) in siNT negative control. Editing frequencies are the mean ± SD of 2 biological replicates. A change in the edit frequency after siAPOBEC3A depletion was called if there was a decrease in editing after siAPOBEC3A pool and siAPOBEC3A #1 depletion in both replicates for each (down green arrow), If not, it was indicated as a non-APOBEC3A edited site (horizontal grey arrows). **(B)** siAPOBEC3A depletion leads to modest but reproducible decreases in editing frequency at 5 candidate C to U variant sites on the pre-rRNA. Percent differences between siAPOBEC3A edit frequency and siNT negative control edit frequency (dotted horizontal line) at C to U edit sites in (E) are graphed. Two biological replicates of siAPOBEC3A pool (circle data points) and siAPOBEC3A #1 (triangle datapoints) are plotted as the mean (horizontal line) ± SD. Locations on the pre-rRNA in nucleotides are indicated on x-axis.

APOBEC3A has candidate editing target sites on nuclear pre-mRNAs that encode nucleolar and cell cycle regulator proteins

Although over half of the nuclear RNA-seq reads aligned to the pre-rRNA, there was still a large portion of other nuclear RNA species that APOBEC3A could regulate. This time I aligned the reads to the hg38 human reference and noticed a prominent increase in the amount of reads that aligned to intronic sequences compared to a previously total cell RNA-seq experiment in MCF10A cells completed in our laboratory (GEO accession GSE154764) (Ogawa et al. 2021) (Figure 6.17A, B). Using the same LoFreq (Wilm et al. 2012) variant analysis, I identified transcriptome-wide C to U variants

(C to U on + strand or G to A on - strand) in siNT and APOBEC3A depleted samples. As expected, there was a modest overall decrease in C to U variants detected in the APOBEC3A depleted samples compared to siNT (Figure 6.17C).



Aligned Reads	
Percent aligned	97.29 +/- 0.37%
Percent Coverage	18.11 +/- 0.87%
Average Coverage Depth	43.65 +/- 3.39

	Aligned Reads	Nuclear	Total Cell
🟦	Fully Exon	31.64%	85.54%
🟥	Partially Exon	2.42%	1.67%
🟩	Fully Intron	62.34%	11.51%
🟪	Intergenic	1.30%	1.28%
🟧	Incompatible	2.30%	0.00%

siRNA Treatment	Total SNVs	C to U Variants [+strand (C to T) -strand (G to A)]	
		Total	Percent (dif. siNT)
siNT	58,154 +/- 11,367	6,920 +/- 1,906	12.39 +/- 0.14% (0.0%)
siAPOBEC3A pool	55,778 +/- 1,444	6,377 +/- 419	12.06 +/- 0.44% (-0.33%)
siAPOBEC3A Individual #1	47,241 +/- 2,712	6,250 +/- 163	11.93 +/- 0.44% (-0.46%)

Figure 6.17: APOBEC3A has candidate edit target sites on 264 nuclear RNA transcripts. **(A, B)** MCF10A nuclear RNA-seq reads aligned to mostly intronic regions of pre-mRNAs. **(A)** Average reads, percent alignment, percent coverage percentage, and coverage depth of all nuclear RNA-sequencing runs using STAR alignment to the hg38 reference index. Average of 6 biological replicates (2x siNT negative control, 2x siAPOBEC3A pool, 2x siAPOBEC3A #1) ± SD. **(B)** Nuclear RNA-seq reads largely align to introns of pre-mRNAs**.** Exon and intron distribution of MCF10A nuclear RNA-seq read alignments (this experiment) vs total RNA-seq (rRNA depleted) read alignment (GEO accession GSE154764)(Ogawa et al. 2021). Percentage of total for each location type is shown in a pie chart and in a table based on color. **(C)** siAPOBEC3A depletion leads to overall reduced C to U edits. There are 264 candidate APOBEC3A C to U edit sites that were revealed by LoFreq variant analysis of nuclear RNAs. The total number of all single nucleotide variants (SNVs) reported by LoFreq was reported as an average of 2 biological replicates ± SD. Total and percent of total C to U variants were reported as an average of 2 biological replicates ± SD (percent difference from siNT negative control). C to U variants were filtered for those that were detected in both replicates of all 3 samples (i.e. all 6 biological replicates) and those that decreased on average for both siAPOBEC3A pool and siAPOBEC3A #1 depletions.

I took the top candidate APOBEC3A editing sites identified for further analysis. A total of 264 sites were chosen that were detected in all 3 samples, in both replicates, and had a decreased editing frequency in both APOBEC3A depletion conditions (Figure 6.17C). Consistent with the read alignment, these top candidate edit sites were detected across all portions of transcripts, but were enriched within intronic regions (Figure 6.18A). As noted, one key factor regulating APOBEC3A's enzymatic function is the presence of a UC sequence motif. Therefore, it was surprising that the 264 sites did not have an enrichment in preceding U presence before the edited C. However, when only looking at the top candidates that had a percent difference in editing frequency between siNT and siAPOBEC3A treatments of greater than 66% (Table 6.2), there was a slight enrichment of a preceding U present (Figure 6.18B). While unexpected, it is important to note that these are the first APOBEC3A candidate sites identified after nuclear enrichment, to our knowledge, due to our unique experimental design.

Examining the APOBEC3A candidate sites in pre-mRNAs, I investigated if there were any trends or consistencies in the types of transcripts which they occur. I took advantage of three nucleolar proteome datasets (Leung et al. 2006; Ahmad et al. 2009; Jarboui et al. 2011; Thul et al. 2017) to test if there was an enrichment of transcripts encoding nucleolar proteins over the expected at random 17.7% within our dataset. Of the 264 sites identified, they occur in 213 transcripts, 184 of which are protein coding (Appendix III). Interestingly, there is almost a 2-fold enrichment in these transcripts encoding nucleolar proteins and this increase is maintained in transcripts that contain sites that have a >33% edit change, >66% edit change, and occur within the UC sequence motif context (Figure 6.18C). Finally, I utilized the STRING database (Szklarczyk et al. 2021) to perform gene ontology overrepresentation analysis of the transcripts that contained our top 264 APOBEC3A candidate sites. The top overrepresented biological process categories include RNA splicing and metabolism (GO:0043484, GO:0016071, GO:0016070) and positive regulation of cell migration, negative regulation of apoptosis and cell death (GO:0030335, GO:0043066, GO:0060548) (Figure 6.18D). While these categories were not enriched for established nucleolar regulators, RNA regulation and apoptosis are overlapping and tightly linked to the ribosome biogenesis pathway (Rubbi and Milner 2003; Boulon et al. 2010; Aubert et al. 2018; Hannan et al. 2022). APOBEC3A potentially edits several pre-mRNA transcripts encoding proteins present in the nucleolus and with roles regulating RNA and the cell cycle.

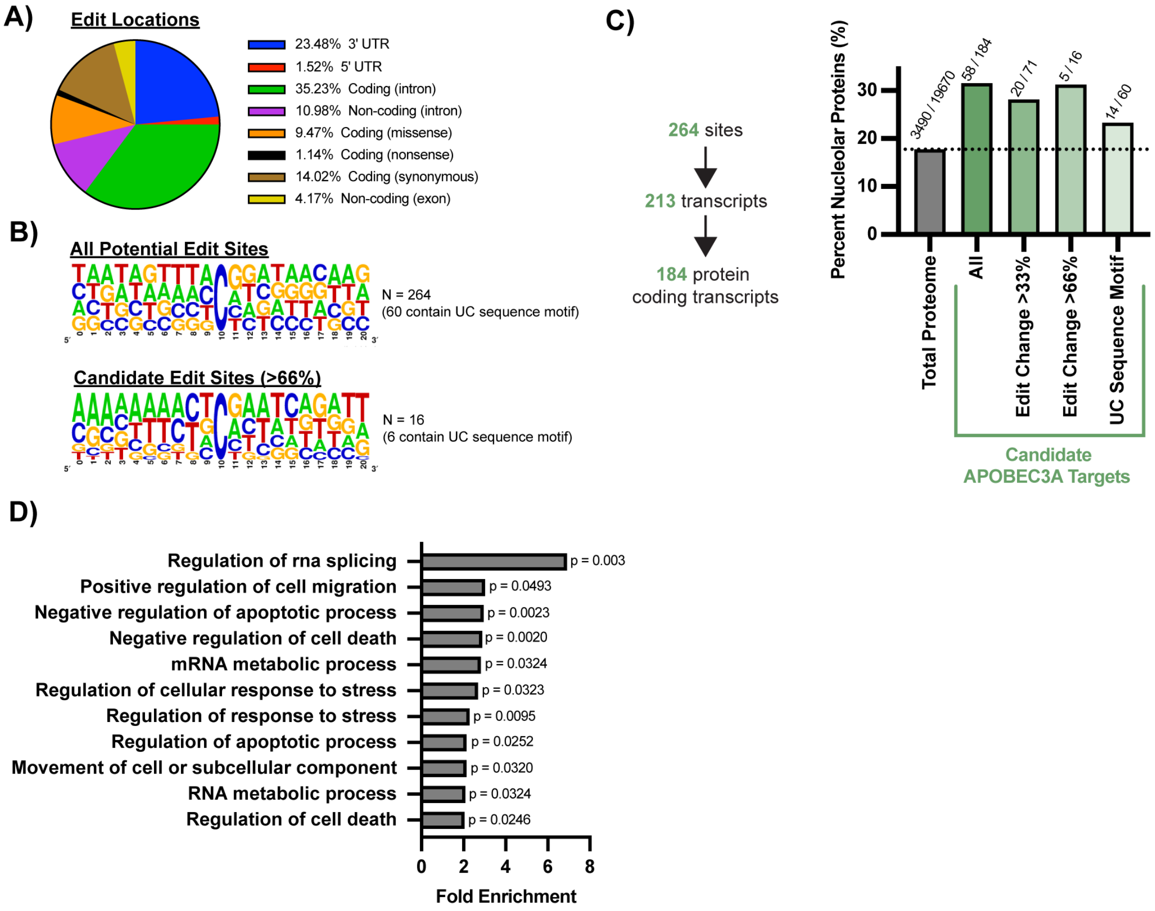


Figure 6.18: APOBEC3A has candidate edit target sites on pre-mRNA transcripts that encode nucleolar proteins and RNA splicing and cell cycle regulators. **(A)** APOBEC3A candidate edit sites on nuclear RNAs are found mostly within intronic regions. The distribution of the 264 candidate APOBEC3A edit sites in (Figure 6.17C) location types are reported as a percentage of total in a pie chart. Locations which include 5' or 3' untranslated regions (UTRs), transcript types (coding or non-coding), and edit type in parentheses (intron, missense, nonsense, synonymous, non-coding exon) are indicated by color. **(B)** The APOBEC3A UC sequence motif is enriched in the top candidate target sites. (Top) sequence logos for all 264 candidate edit sites in (Figure 6.17C) including 10

nucleotides upstream and downstream of the edited C. (Bottom) Same as (Top) except indicating only the top 16 edit sites (exhibiting a greater than 66% decrease in siNT editing frequency compared to the average of the siAPOBEC3A pool and siAPOBEC3A #1 treatments) were used for logo creation. **(C)** The APOBEC3A candidate edit sites are enriched on pre-mRNAs encoding nucleolar proteins. The 264 candidate edit sites were located on 213 transcripts of which 184 are protein coding. An estimate of human nucleolar proteins were determined based on the presence of a protein in at least 1 of 3 proteomic datasets (N = 3490) (Leung et al. 2006; Ahmad et al. 2009; Jarboui et al. 2011; Thul et al. 2017). The total number of human proteins was reported based on Thul et al. 2017 (N = 19,670).(Thul et al. 2017) Proteins that were encoded by transcripts with APOBEC3A candidate targets and subgroups within were tested for enrichment of nucleolar protein coding transcripts by comparison to an estimate of percentage of the proteome that is nucleolar (17.7%, dotted horizontal line) and graphed. **(D)** The APOBEC3A candidate edit sites are enriched on pre-mRNAs involved in RNA metabolism, positive regulators of cell migration and negative regulators of cell apoptosis/death. All 213 transcripts that contain APOBEC3A candidate edit sites were analyzed for overrepresentation of biological function gene ontology categories using the STRING database (Szklarczyk et al. 2021). Categories were reported where fold-enrichment > 2.0 and $p < 0.05$. p-values indicated on graph.

Table 6.2: Top candidate APOBEC3A C to U edit sites on nuclear RNAs.

Genomic Loci	Gene Name	Type	Nucleolar/ UC motif?	siNT Edit Freq.	si3A Edit Freq.	siNT / si3A Edit Ratio	
chr2:240465606	GPC1	missense (R/C)	No / No	0.670	0.318	2.179	
chr6:41965816	CCND3	intronic	**Yes** / **Yes**	0.624	0.298	2.151	*
chr6:111374684	REV3L	missense (T/I)	No / No	0.617	0.347	2.057	
chr1:26700833	ARID1A	intronic	No / No	0.529	0.338	2.028	
chr2:44320435	PREPL	3' UTR	No / **Yes**	0.896	0.534	1.991	
chr10:97680760	AVPI1	intronic	No / No	0.476	0.259	1.988	
chr7:130891504	MIR29A	intronic	No / **Yes**	0.571	0.308	1.975	
chr6:150990351	MTHFD1L	intronic	**Yes** / No	0.588	0.390	1.828	
chr2:151799138	ARL5A	3' UTR	No / No	0.688	0.421	1.820	
chr9:94167287	RP11-2B6.3	intronic	No / No	0.542	0.336	1.801	
chr19:54460050	LENG8	3' UTR	**Yes** / No	0.443	0.290	1.783	
chr13:109194076	MYO16	intronic	No / **Yes**	0.474	0.269	1.780	
chr11:13461680	BTBD10	intronic	**Yes** / **Yes**	0.528	0.366	1.773	*
chr2:69749292	ANXA4	intronic	No / **Yes**	0.550	0.321	1.738	
chr1:47377105	CMPK1	3' UTR	**Yes** / No	0.668	0.453	1.719	
chr10:97671727	PI4K2A	intronic	No / No	0.512	0.385	1.712	

Discussion

Following up on candidate hits from a high throughput phenotypic screen for nucleolar function, I defined a role for APOBEC3A in ribosome biogenesis in the nucleus, revealing for the first time the intriguing possibility that the pre-rRNA is edited in human cells. I discovered that the human cytidine deaminases, APOBEC3A and APOBEC4, are candidate ribosome biogenesis factors through validation of our previous genome-wide siRNA screen (Farley-Barnes et al. 2018) for changes in nucleolar number in MCF10A cells. Biochemical assays on APOBEC3A reveal that its depletion leads to defects in cell cycle progression, likely through the nucleolar stress response, and decreased protein synthesis, likely due to inhibitions in pre-LSU rRNA processing in the nucleus/nucleolus. Furthermore, a novel nuclear RNA-seq experiment indicated select putative APOBEC3A C to U editing sites on the pre-LSU rRNA and pre-mRNAs encoding nucleolar localized proteins and RNA splicing and cell cycle regulators. These results strongly suggest APOBEC3A functions within the nucleus to modulate ribosome biogenesis and cell cycle progression, two interconnected processes, through its editing activity (Figure 6.19).

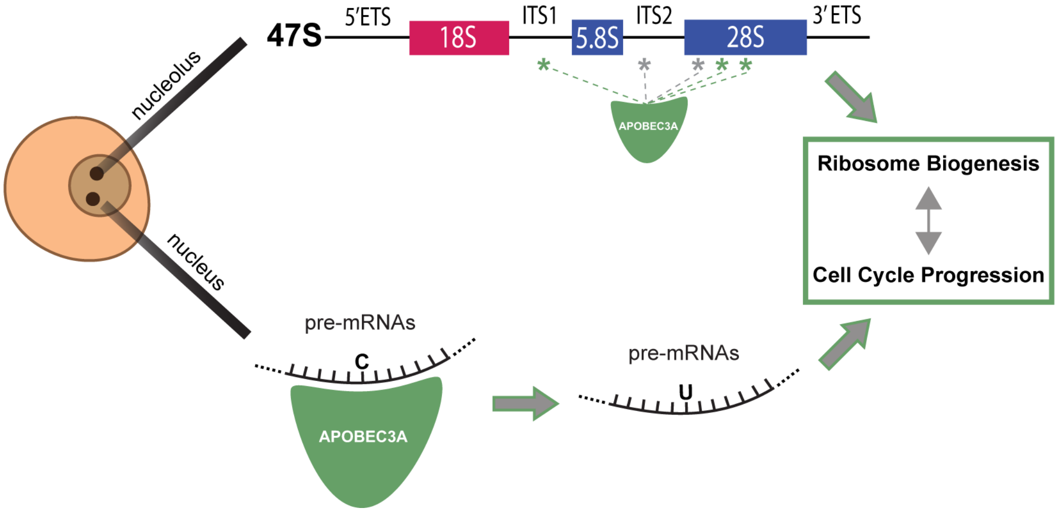


Figure 6.19: Model where APOBEC3A regulates human ribosome biogenesis by editing the pre-rRNA and/or pre-mRNAs that encode nucleolar or cell cycle regulating proteins. Schematic detailing APOBEC3A's potential functions in the regulation of ribosome biogenesis in the nucleus (pre-mRNA editing/binding) or nucleolus (pre-rRNA editing/binding). These function(s) would impact both ribosome biogenesis and the cell cycle, which are connected, either indirectly or directly.

There is good correlation between APOBEC3A's role in pre-LSU rRNA processing and its predicted edit sites. I observed a defect in ITS2 cleavage in the pre-rRNA leading to the formation of the mature LSU 28S and 5.8S rRNAs by northern blotting. Similarly, 4 out of the 5 identified APOBEC3A C to U editing sites on the pre-rRNA are within the pre-LSU (5.8S + ITS2 + 28S rRNA). Thus, the pre-rRNA processing steps affected by APOBEC3A depletion correlate with the concentration of putative APOBEC3A editing sites in the pre-rRNA, pointing towards a potential direct function of APOBEC3A. I also identified candidate edit sites on pre-mRNAs with nucleolar functions, including on ITS2 processing. These APOBEC3A target sites in the nucleus were enriched for pre-mRNAs encoding proteins localized to the nucleolus and for proteins involved in RNA splicing (GO:0043484). Most strikingly within this subgroup of RNA splicing, *DDX17*, encodes an RNA helicase also required for ITS2 pre-rRNA cleavage (Jalal et al. 2007). Thus, APOBEC3A may regulate pre-LSU processing both directly (rRNA) and indirectly (pre-mRNAs encoding nucleolar proteins) based on our identified APOBEC3A candidate target sites.

To our knowledge, this is the first time that a cytidine deaminase's function has been investigated on a nuclear-only subset of RNAs, offering a unique dataset for analysis of the pre-rRNA. Furthermore, since rRNA was not depleted, over 50% of the reads aligned to the 47S pre-rRNA or 5S rRNA transcripts with average read coverages over 400,000 and 100,000 respectively. This deep coverage across the pre-rRNA provided the sensitivity to detect very low frequency (<0.05%) edits and any changes in edit frequency between negative control siNT and APOBEC3A siRNA depletion conditions. While previous studies have observed human rDNA/rRNA variants across rDNA repeats, mature rRNAs, tissues, or individuals (Parks et al. 2018; Nurk et al. 2022; Rothschild et al. 2023) our study allows for identification of variants precisely produced by one enzyme, APOBEC3A.

Additionally, nuclear RNA-seq allowed me to better analyze intronic regions of pre-mRNAs for candidate APOBEC3A C to U edit sites. Nuclear RNA-seq reads aligned to intronic regions almost 5-fold more in comparison to a previous total RNA-seq experiment from MCF10A cells (GEO accession GSE154764) (Ogawa et al. 2021). This predisposed our APOBEC3A editing analysis towards intronic regions which have not yet been explored, as they have been mostly focused on exon regions through targeted or total RNA sequencing (Niavarani et al. 2015; Sharma et al. 2015; Sharma et al. 2017; Alqassim et al. 2021). Unexpectedly, I did not observe robust enrichment of the UC dinucleotide preferred by APOBEC3A for nuclear pre-mRNA (intron enriched) editing. More recent work has highlighted, in addition to the UC motif, the presence of a hairpin secondary structure in APOBEC3A edit sites on DNA (Buisson et al. 2019) and RNA (Barka et al. 2022). This may also contribute to APOBEC3A's preference in the nuclear RNA sites I identified. It remains to be explored how APOBEC3A regulates these pre-mRNAs and if splicing, stability, and/or translation efficiency is altered.

Is APOBEC3A editing RNA or DNA? It is possible that it is APOBEC3A's ssDNA editing function that is leading to downstream changes that I observe as RNA edits. APOBEC3A prefers ssDNA over RNA *in vitro* (Barka et al. 2022). Although it is lowly expressed in MCF10A cells, addition of siRNAs could induce its expression [as suggested by my results in MCF10A cells (Figure 6.6)] causing edits in DNA over the short 72 h time course experiments that are reduced after treatment with siAPOBEC3A. While my results have not answered whether the edits are in RNA or DNA, our observed pre-rRNA processing and maturation defects with no changes in pre-rRNA transcription would point towards APOBEC3A functioning at the RNA level. Possibly, it is merely APOBEC3A's RNA binding ability that regulates pre-LSU maturation, which would partially explain the low edit frequencies and differences I observed. Lastly, increasing the likelihood that our reproducible results are from APOBEC3A's function on RNA, in cultured cancer cell lines, APOBEC3A associated genomic mutation rates are variable and episodic in nature over lengthier time periods than our experiments (Petljak et al. 2019; Petljak et al. 2022). Nonetheless, targeted deep DNA sequencing of the identified edit sites would uncover whether APOBEC3A editing is occurring at the DNA or RNA level.

APOBEC3A’s requirement in ribosome biogenesis strengthens its connection to cancer pathogenesis beyond its current established role in producing mutations across the genome. Previous work has established several examples of proteins working in both DNA repair and ribosome production [reviewed in (Ogawa and Baserga 2017; Lindstrom et al. 2018)]. My work on APOBEC3A highlights another example that links the integrity of the genome to ribosomes biogenesis. Currently, cytidine deaminase inhibitors are being developed [Reviewed in (Grillo et al. 2022)], including against APOBEC3A (Kurup et al. 2022; Serrano et al. 2022), with a focus on their application in cancer therapy (Olson et al. 2018). APOBEC3A’s newfound function in ribosome biogenesis emphasizes it as a top candidate cytidine deaminase target for cancer therapeutics.

Materials and Methods

GTEx and TCGA expression and survival dataset analysis

Genotype-Tissue Expression (GTEx) unmatched normal and The Cancer Genome Atlas (TCGA) matched normal and tumor expression datasets were obtained through the Xena platform (https://xena.ucsc.edu/) (Goldman et al. 2020). RNA-seq by Expectation-Maximization (RSEM) Log2 fold expression levels for respective cytidine deaminase transcripts were subtracted from the mean of the overall normal and tumor tissues combined and graphed. Survival data, also within the same dataset, was obtained through the Xena platform. Samples were stratified by either high or low cytidine deaminase expression and survival curves were generated. For cytidine deaminases where the median expression was 0, data was stratified by either detected or undetected levels of expression.

Cell Culture

MCF10A cells (ATCC, CRL-10317) were subcultured in Dulbecco's modified Eagles' medium/nutrient mixture F-12 (Gibco, 1130-032) containing horse serum (Gibco, 16050), 10 µg/mL insulin (Sigma, I1882), 0.5 µg/mL hydrocortisone (Sigma H0135), 100 ng/mL cholera toxin (Sigma, C8052), and 20 ng/mL epidermal growth factor (Peprotech, AF-100-15). HeLa cells (ATCC, CCL-2) cells were grown in DMEM (Gibco, 41965-062) with 10% fetal bovine serum (FBS, Gibco, 10438026). Cell lines were maintained at 37°C, in a humidified atmosphere with 5% CO_2.

For the high-throughput nucleolar number and 5-EU assays, 3,000 cells/well were reverse transfected in 384-well plates on day 0. For RNA or protein isolation, 100,000 cells/well were seeded into 6-well plates on day 0. For the dual-luciferase assay, 100,000 cells/well were seeded into 12-well plates on day 0. For nuclei isolation, 1,560,000 cells were seeded into 15cm plates on day 0.

RNAi

Horizon Discovery Biosciences siGENOME SMARTpool siRNAs and ON-TARGETplus pool and individual were used for all assays as indicated above. siRNA transfections were completed using Lipofectamine RNAiMAX Transfection Reagent (Invitrogen, 13778150) per manufacturer's instructions with a final siRNA concentration of 20nM for 384-well plate high-throughput assays (nucleolar number and 5-EU) and 33nM for all other assays. For the high-throughput assays, cells were reverse transfected on day 0 and siRNA controls were added to 16 wells and other siRNAs were added to 1 well for each replicate. For all other assays cells were transfected 24 h after plating. All siRNA depletions were performed for 72 h total.

Nucleolar number assay

Carson Bryant and I counted nucleolar number in MCF10A cells in high-throughput as done previously in our laboratory (Farley-Barnes et al. 2018; Ogawa et al. 2021; McCool et al. 2022b), except that siNOL11 was used as a positive control. In short, after 72 h siRNA depletion, cells were fixed, stained for nucleoli [72B9 α-fibrillarin antibody (Reimer et al. 1987), Alexa Fluor 647 goat anti-mouse IGG secondary] and nuclei (Hoechst). Images were acquired with IN Cell 2200 imaging system (GE Healthcare) and a custom CellProfiler pipeline was used for analysis. For both siRNA pool and individual siRNA deconvolution treatments, a hit was called based on a cutoff of a mean one-nucleolus percent effect greater than or equal to +3 standard deviations (SD) above the non-targeting siRNA (siNT) negative control of a given replicate.

5-EU incorporation (nucleolar rRNA biogenesis) assay

I measured nucleolar 5-EU incorporation (rRNA biogenesis) in high-throughput as done previously in our laboratory (Bryant et al. 2022; McCool et al. 2022b). In short, the nucleolar number assay was followed by the addition of a 1 mM 5-ethynl uridine (5-EU) treatment at the end of a 72 h siRNA depletion. An additional click chemistry step was performed to attach an AF488 azide for 5-EU visualization. Carson Bryant used a custom CellProfiler pipeline was also used for analysis to measure median nucleolar 5-EU signal within each nucleolus.

Cell cycle analysis

Carson Bryant and I performed cell cycle analysis as done previously (Chan et al. 2013), including in work performed in our laboratory (Ogawa et al. 2021). In short, integrated Hoechst staining intensity was measured per nucleus from the high-throughput images for the 5-EU incorporation assay. The Log2 integrated intensities of each nucleus was calculated. The siNT treatment G1 peak was set at 1.0 and G2 peak at 2.0 for normalization. Cell cycle phases were defined as the following normalized Log2 integrated Hoechst staining intensities: sub-G1 < 0.75, G1 = 0.75 – 1.25, S = 1.25 - 1.75, G2/M = 1.75 – 2.5, >4n > 2.5.

Protein harvesting and western blotting

Cells were harvested by scraping followed by a PBS rinse. Cells were resuspended in AZ lysis buffer (50 mM Tris pH 7.5, 250 mM NaCl, 1% Igepal, 0.1% SDS, 5 mM EDTA pH 8.0) supplemented with protease inhibitors (cOmplete™ Protease Inhibitor Cocktail, Roche, 11697498001) and lysed by vortexing at 4ºC for 15 minutes. Cell debris was pelleted by spinning at 21000 RCF at 4ºC for 15 minutes and protein containing supernatant was taken. Protein concentration was determined by Bradford assay (Bio-Rad, 5000006). An equal amount of total protein (20 - 50 μg per sample) was separated by SDS-PAGE and transferred to a PVDF membrane (Bio-Rad, 1620177) for blotting.

The following primary antibodies were used: α-APOBEC3A/B (Invitrogen, PA5-104035), α-p21-HRP (Santa Cruz, sc-6246), α-p53-HRP (Santa Cruz, sc-126), α-puromycin (Kerafast, EQ0001), and α-β-actin (Sigma Aldrich, A1978). For detection of non-HRP conjugated primary antibodies, either secondary α-mouse (GE Healthcare, Life Sciences NA931) or α-rabbit (GE Healthcare, Life Sciences NA934V) HRP conjugated antibodies were used. Images were acquired using Bio-Rad Chemidoc (Bio-Rad, 12003153) and analyzed using ImageJ software.

Puromycin incorporation (global protein synthesis) assay

Global protein synthesis was accessed as done previously (Schmidt et al. 2009), including in our laboratory (Farley-Barnes et al. 2018; Sondalle et al. 2019; Ogawa et al. 2021; McCool et al. 2022b). In short, 1 μM (or 0.5 μM for Mock 0.5 control) puromycin was added for 1 h to label nascent polypeptides over that time period at the end of a 72 h siRNA depletion. Protein harvesting and western blotting was carried out as described above.

RNA harvesting

RNA (either from total cell or nuclei purified fractions) was extracted using TRIzol reagent (Life Technologies, 559018) per manufacturer’s instructions. RNA was washed and stored in 75% ethanol as a pellet at -20 ºC or -80 ºC before use in downstream analysis.

Digital droplet PCR

The $A_{260/230}$ of all total cell RNA samples dissolved in nuclease-free water were determined to be above 1.7 by Nanodrop (ThermoFisher, ND2000CLAPTOP) before proceeding to cDNA synthesis. Five µg of each RNA sample was treated with DNase I to digest any genomic DNA. cDNA was synthesized using iScript Advanced cDNA synthesis kit (Bio-Rad, 1725037) per manufacturer's instructions. ddPCR Supermix for Probes (No dUTP) (Bio-Rad, 1863023) was used per manufacturer's instructions with gene-specific primers/probes (Table S7). Droplets were produced using Automated Droplet Generator (Bio-Rad, 1864101). The following amplification parameters were used for PCR: 95 ºC for 10 minutes for enzyme activation, 40 cycles of 94 ºC for 30 s and 55 ºC for 1 minute, then 98 ºC for 10 minutes for enzyme deactivation. Data were acquired using QX200 Droplet Reader (Bio-Rad, 1864003). Number of positive droplets was used as a readout of mRNA expression.

qRT-PCR

qRT-PCR was performed as carried out previously in our laboratory (McCool et al. 2022b). The $A_{260/230}$ of all RNA samples dissolved in nuclease-free water were determined to be above 1.7. cDNA was synthesized from 1 µg of total RNA using iScript gDNA clear cDNA synthesis kit (Bio-Rad, 1725035) with random primers. iTaq Universal SYBR Green Supermix (Bio-Rad, 1725121) was used to perform qPCR with gene-specific primers (Table S7). The following amplification parameters were used: initial denaturation 95 ºC for 30 s, 40 cycles of 95 ºC for 15 s and 60 ºC for 30 s. Subsequent melt curve analysis was performed to ensure a single product, 95 ºC for 15 s, then gradual (0.3 ºC/15 s) increase from 60 ºC to 94.8 ºC. Amplification of the 7SL transcript was used as an internal control and relative RNA levels were determined by using comparative C_T method ($\Delta\Delta C_T$). Three technical replicates were completed for each biological replicate and averaged.

Dual-luciferase reporter assay

rDNA promoter activity was accessed as done previously (Ghoshal et al. 2004), including in our laboratory (Ogawa et al. 2021; McCool et al. 2022b). In short, 48 h after siRNA transfection, MCF10A cells were co-transfected with both 1000 ng of pHrD-IRES-Luc plasmid(Ghoshal et al. 2004) and 0.1 ng of a constitutively expressed *Renilla* containing internal transfection control plasmid (Freed et al. 2012) using Lipofectamine 3000 (Thermo Fisher Scientific, L3000015) per manufacturer's instructions. Twenty-four hours later (72 h after siRNA transfection), cells were harvested by scraping and luminescence was measured by a Dual-luciferase Reporter Assay System (Promega, E1910) per manufacturer's instructions using a GloMax 20/20 luminometer (Promega).

BioAnalyzer RNA quantification

Either 1 µg of total cell or nuclear fraction RNA was submitted at a concentration of 100 ng/µL in nuclease-free water to the Yale Center for Genome Analysis for Agilent BioAnalyzer analysis. Nuclear RNA samples were also used for downstream sequencing analysis (see below).

Northern blotting

Northern blotting was performed as carried out previously in our laboratory (Farley-Barnes et al. 2018; Ogawa et al. 2021; McCool et al. 2022b). In short, 3 µg of total cellular RNA was resolved on a denaturing 1% agarose / 1.25% formaldehyde gel using Tri/Tri buffer (Mansour and Pestov 2013). Separated RNA was transferred to a Hybond-XL membrane (GE Healthcare, RPN303S) and UV-crosslinked. Membranes were stained with methylene blue (0.025% w/v) and imaged using Bio-Rad Chemidoc (Bio-Rad, 12003153). Blots were hybridized to a ^{32}P radiolabeled DNA oligonucleotide probe (ITS2, 5' – AAGGGGTCTTTAAACCTCCGCGCCGGAACGCGCTAGGTAC – 3') and detected using a phosphorimager (Amersham™ Typhoon™, 29187194). Images were analyzed using ImageJ software and ratio-analysis of multiple precursors (RAMP)(Wang et al. 2014) was used for quantification.

Nuclei harvesting

All steps were performed at 4ºC. Cells were harvested by scraping following a PBS rinse. Washed cell pellets were resuspended in Buffer A (10 mM HEPES pH 8.0, 10 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM dithiothreitol, 1X protease inhibitors (cOmplete™ Protease Inhibitor Cocktail, Roche, 11697498001), 4 mM NEM) and swelled for 10 minutes. Swollen cells were dounced using a 7mL dounce (Wheaton, 3575420) for 20 strokes. Dounced cells were centrifuged at 220 RCF for 5 minutes. The pellet containing nuclei was then used for downstream RNA analysis.

Nuclear RNA-seq and variant analysis

Library preparation and RNA-seq was completed by the Yale Center for Genome Analysis. Illumina RNA-seq library prep was completed without rRNA depletion or polyA enrichment step(s). The cDNA library was made by random primed first strand synthesis. Paired-end (2 x 150 bp) reads were collected using the NovaSeq 6000 S4 XP system (Illumina).

RNA-seq analysis was completed using Partek® Flow® software available through the Harvey Cushing / John Hay Whitney Medical Library. For pre-rRNA analysis raw nuclear RNA-seq reads were aligned to either a [47S pre-rRNA consensus sequence (NR_145144.1) or 5S rRNA consensus sequence (E00204, 5SrRNADB(Szymanski et al. 2016)) as reference indexes] using BowTie2 (Version 2.2.5), with default parameters. For transcriptome-wide analysis, raw nuclear RNA-seq reads were trimmed of Illumina Universal adapter sequences. Trimmed reads were aligned to the hg38 reference index using STAR (Version 2.7.8a), using default parameters.

Single nucleotide variants (SNVs) were detected using LoFreq(Wilm et al. 2012) (Version 2.1.3a) against the respective aligned reference above, default parameters, $p < 0.05$, read depth > 10. Data are available on NCBI: accession PRJNA935922.

Comparisons were made to a previously obtained total MCF10A cell RNA-seq dataset completed in our laboratory (GEO accession GSE154764)(Ogawa et al. 2021).

Consensus motif logo visualization

Consensus sequence motif logos were generated using WebLogo (Version 2.8.2).(Crooks et al. 2004) Input sequences included the genomic DNA sequence of 10 nucleotides upstream and downstream of candidate edited cytidines of nuclear RNAs.

Enrichment analyses

A metadataset of nucleolar proteins was created based on 3 previous proteomic nucleolar datasets.(Leung et al. 2006; Ahmad et al. 2009; Jarboui et al. 2011; Thul et al. 2017). Adding all the proteins together that were detected in at least one of these datasets results in an estimate of 3490 nucleolar proteins out of 19670 total in the entire human proteome (17.7%)(Thul et al. 2017) which was used as a baseline for enrichment analysis.

STRING Database (Version 11.5)(Szklarczyk et al. 2021) was used for biological processes gene ontology enrichment analysis. Categories were reported where fold enrichment > 2 and $p < 0.05$.

Statistical analyses

Statistical analyses were performed using GraphPad Prism 9.3.1 (GraphPad Software). Tests are described in the associated figure legends.

Chapter 7

Perspectives and Future Directions

Introduction

Through my dissertation work I have studied some of the complexities and nuances of ribosome biogenesis in human cells. By studying novel factors identified through the Baserga laboratory's high-throughput screens for nucleolar function (Farley-Barnes et al. 2018; Ogawa et al. 2021; Bryant et al. 2022), I was able to learn about proteins that might not have been implicated in ribosome biogenesis otherwise, highlighting the value of these screens. It was through collaborative efforts that I produced a higher-confidence hit list through rescreening and with utilizing Carson Bryant's nucleolar rRNA biogenesis assay. I expanded upon this work through my study of how NOL7, RSL24D1, PeBoW, and APOBEC3A function in human ribosome biogenesis. I was able to obtain some more answers on how ribosome biogenesis is regulated in human cells, but in doing so, opened many areas of future exploration.

Difficulties avoiding of siRNA off-target effects in ribosome biogenesis studies

Off-target effects are inevitable when utilizing RNAi as a knockdown method, however the proper steps can be taken to ensure you are not studying one of these off-target effects. Since ribosome biogenesis is such an essential process, knockout (CRISPR) experiments are not useful. Based on my experience, I have developed a strategy for utilizing siRNAs as a depletion method in the future.

After observing effective depletion of your target of interest upon RNAi treatment by western blotting and qRT-PCR, I would suggest pursuing other depletion methods that may have different off-target effects. First, as presented here, deconvolution of siRNA pools are effective ways to lessen the likelihood of off-target effect observations. Since it is unlikely that multiple individual siRNAs would produce the same off-target effects, recapitulation of phenotypes with individual siRNAs that have the same target is a good first step towards off-target avoidance. Next, although not presented here, I attempted to use other depletion methods, such as siTOOLs siRNA pools [low concentrations of 30 siRNAs all targeting the same transcript (Hannus et al. 2014)] and Vivo Morpholinos to block translation of a target transcript (Morcos et al. 2008) with limited success. There are also more effective and precise methods of depletion, such as inducible protein degradation tags [i.e. dTAG system (Nabet et al. 2018)], which can be time intensive to set up but pay off with more convincing experimental results.

The "gold-standard" of confirming siRNA-based results is to perform rescue experiments, which Baserga laboratory members, including myself, were able to complete in our studies of NOL7 (Chapter 4). Rescue experiments also provide opportunity for even more insightful future studies I have not yet pursued. By introducing an siRNA-resistant version of NOL7, previous Baserga laboratory members and I were able to observe rescue of some of the phenotypes observed, indicating that NOL7 is responsible for those phenotypes and not an off-target effect. While not completed here for APOBEC3A, a successful rescue experiment would ensure that I am not examining off-target effects. Furthermore, rescue experiments can be carried out with disease mutants or individual protein domains to test if they alter biological function. In my studies, this would be especially advantageous to studying APOBEC3A, which has catalytic RNA/DNA editing activity that can be abrogated through a E72A mutation, to determine if its editing activity is necessary for its role in making ribosomes. In general, there are many mutations associated with ribosomopathies and/or cancer that would be advantageous to study through rescue experiments interrogating the ribosome biogenesis pathway in human tissue culture cells.

Continued leveraging and development of high-throughput screening methods to study human ribosome biogenesis

Recent high-throughput screens have been invaluable to the advancement of our understanding of human ribosome biogenesis. Several laboratories, including our own, have performed screens to identify factors involved in and to better define ribosome biogenesis regulation in humans (Wild et al. 2010; Tafforeau et al. 2013; Badertscher et al. 2015; Ogawa et al. 2021; Dorner et al. 2022; Hannan et al. 2022). These screens have laid the groundwork for follow up studies on individual factors in our laboratory (Farley-Barnes et al. 2020; McCool et al. 2022b) (NOL7: Chapter 4, RSL24D1: Chapter 5, APOBEC3A: Chapter 6). Through their utility, they have spurred the development of more vigilant validation steps in our screening campaigns (Chapter 2) and more high-throughput functional assays (Chapter 2) (Bryant et al. 2022). It is clear these studies have been vital to the field, but how do we continue to leverage the already available datasets and produce even more insightful datasets as well?

Just within the hits identified in the Baserga laboratory's screen and re-screening campaigns, there are years' worth of individual follow-up mechanistic studies awaiting to better understand how these novel ribosome biogenesis factors function in human cells. For instance, APOBEC4's function (Chapter 6) has yet to be defined in any capacity in human cells, let alone in ribosome production. Many of the novel factors identified are not known to be nucleolar. Only 107/702 (15.2%) of the rescreened hits were nucleolar based on being present in at least one of three nucleolar proteomic datasets (Leung et al. 2006; Ahmad et al. 2009; Jarboui et al. 2011; Thul et al. 2017). This indicates many possibilities to define new pathways that control ribosome biogenesis from outside the nucleolus, specifically ones converging on pre-rRNA processing (Chapter 2). Excitingly, current members of the Baserga laboratory are investigating some of these non-nucleolar hits involved in other cellular processes, such as within the nuclear pore complex, mitosis, and mitochondrial function.

Beyond merely identifying novel ribosome biogenesis regulators, the most recent high-throughput ribosome biogenesis screens have begun to take the necessary next step towards offering more mechanistic insight. For example, Carson Bryant developed a module to our laboratory's screening method to measure the amount of nucleolar rRNA made over a 1 h time period by 5-EU labeling (nucleolar rRNA biogenesis) (Bryant et al. 2022) and the Kutay Laboratory used another imaging-based assay to measure 60S subunit biogenesis specifically by tracking a GFP tagged RPL29 (Dorner et al. 2022). Our established imaging-based pipeline and siRNA and small molecule libraries leave a myriad of possibilities for more mechanistic high-throughput screening inspired by this recent work. Here, I list ideas for future high-throughput studies that take a step towards mechanistic insight regarding ribosome biogenesis:

1) Recently, the Hannan lab completed an imaging-based screen to test which depleted proteins lead to increases in nuclear TP53 levels, ultimately highlighting that the nucleolar surveillance pathway is essential for TP53 accumulation (Hannan et al. 2022). A factor downstream in the TP53 apoptosis pathway and a candidate for future screening is CDKN1A (p21) (el-Deiry et al. 1994). Furthermore, CDKN1A has been shown to be driven through a p53 independent pathways, including through RPL3 (Russo et al. 2013; James et al. 2014). I propose a high-throughput imaging-based screen to test for CDKN1A induction in both p53 WT and p53 -/- cells to better define p53 independent pathway(s) of nucleolar stress by antibody staining. First, I would test for protein regulators through an RNAi screen and then small molecule inducers of this pathway could also be tested with an eye towards treatment of p53 mutant cancers.

2) During my studies of RSL24D1 and the PeBoW complex (Chapter 5 and (McCool et al. 2022b), I discovered that siRNA depletion of those proteins resulted in a reduction in POLR1A/RPA194 protein levels (Figure 5.5) which is similar to the potent RNAP1 inhibitor BMH-21 (Peltonen et al. 2014a; Peltonen et al. 2014b; Fu et al. 2017; Jacobs et al. 2022). This was an unexpected observation for these LSU biogenesis factors and I expect that there are other non-canonical RNAP1 regulators yet to be discovered. Thus, I propose performing a high-throughput imaging-based screen to detect changes in POLR1A/RPA194 levels and localization in cells. Our additional

nucleolar segmentation module will quantify both total and nucleolar RPA194, allowing for identification of factors or small molecules that dysregulate RPA194 levels and normal localization upon their respective depletion or treatment. This screen would identify more RNAP1 regulators and other small molecules that have similar mechanisms to that of BMH-21.

3) The adaptability of the Baserga laboratory's screening platform provides vast potential. I mentioned two specific ideas above but staining for any protein (provided a reasonable antibody) can be implemented to measure its nucleolar, nuclear, and/or total abundance in high-throughput. 5-EU labeling can be used to measure changes in nucleolar rRNA biogenesis on shorter timescales than the 72 h timepoints used to measure changes in nucleolar number, making it more adaptable to small molecule studies. Furthermore, addition of other modules such as RNA-FISH of the pre-rRNA 5'ETS to measure rDNA transcription more directly or expression of fluorescently tagged ribosomal or ribosome biogenesis related proteins can add another layer of information. This method has already been developed by Antony et al. (2022), termed "47S-FISH-FLOW", where they used fluorescent probes at the extreme 5' end of the 47S pre-rRNA transcript to measure nascent 47S pre-rRNA (Antony et al. 2022).

Xenopus **as a model system for ribosomopathies (NOL7 and other ribosome biogenesis factors)**

Xenopus tropicalis (and *Xenopus laevis)* is an optimal model organism to elucidate the molecular pathology of ribosomopathies in the Baserga laboratory for multiple reasons. First, the Baserga and Khokha Labs have successfully collaborated utilizing *X. troplicalis* to explore the role of ribosome biogenesis factors in craniofacial and overall development previously (Griffin et al. 2015; Griffin et al. 2018; Farley-Barnes et al. 2020). Secondly, *X. tropicalis* as a model organism offers advantages over mice due to their short generation time, extra-utero development, and overall low cost of maintenance. Additionally, mice may not faithfully model some human ribosomopathies, as seen in Diamond Blackfan Anemia (Matsson et al. 2006). Furthermore, *X. tropicalis* is more closely related to humans than zebrafish (Grainger 2012; Barriga et al. 2015).

I propose studying Nol7's role in *Xenopus* development, with the hypothesis that its deletion or depletion will produce similar phenotypes as that of loss of Nol11 (also a Utp) as done previously (Griffin et al. 2015). Preliminary data have already been collected to this end in by John Griffin, Grant Wheeler, and Santosh Maharana at the University of East Anglia in the United Kingdom. In the future it is possible that our laboratory will perform biochemical assays on their frog embryos without Nol7. Studying frog embryonic development will be useful in the future study of novel ribosome biogenesis factors that could be implicated in ribosomopathies.

Validation of APOBEC3A editing and "specialized ribosomes"

Follow-up experiments to validate the candidate APOBEC3A edit sites I identified on both pre-mRNA and pre-rRNA are essential to determine if APOBEC3A is catalyzing these edits in cells. One way of validating would be to detect the presence of APOBEC3A at these locations on RNA (or DNA). Due to the difficulties in making APOBEC3A specific antibodies and its low level of expression in most cell lines, it would be necessary to produce a tagged construct to express for RNA immunoprecipitation experiments. Detection of APOBEC3A at these sites would increase confidence that even if APOBEC3A is not editing to the full extent, it is still interacting with nuclear RNAs at these locations. Another way of confirming these C to U editing sites would be in an overexpression system of APOBEC3A, where you would expect an increase in editing frequency compared to a negative control. Once editing sites are confirmed, more in depth experiments probing the consequences of individual edit locations can be pursued.

Another important caveat is that most of the differences in C to U editing levels were modest after APOBEC3A depletion (especially on the pre-rRNA), therefore it could be that it is merely APOBEC3A binding that is important for its function in these contexts. To answer this, rescue experiments would be critical where APOBEC3A is depleted and an siRNA resistant version of APOBEC3A wild-type or catalytically dead E72A mutant is introduced. If both the wild-type and E72A mutant rescue the observed ribosome biogenesis defects, then it is likely that only APOBEC3A binding is required. If only the wild-type rescues these defects, then it is likely that APOBEC3A catalysis is required, however at what specific locations would remain unclear.

Broadly, more extensive experiments regarding APOBEC3A editing, specifically pre-rRNA editing, would open a new avenue into the field of “specialized ribosomes”. This term encapsulates the observation of active ribosomes containing variations in rRNA sequence, RNA modifications, and associated ribosomal proteins, which can produce differential translation capacity for various mRNAs [Reviewed in (Xue and Barna 2012; Genuth and Barna 2018; Guo 2018)]. Furthermore, these have been shown to be tissue-specific and thus proposed to control tissue differentiation and development [Reviewed in (Norris et al. 2021)] (Genuth et al. 2022) (a connection to the tissue specific phenotypes of ribosomopathies) or potentially increase translation of oncogenic transcripts to drive cancer. APOBEC3A would fit into this context by having the ability to produce a variable pool of pre-ribosomes/ribosomes with different C to U editing sites to impact ribosome formation and/or function.

Connecting to the bigger picture: human disease and therapeutic strategies

In general, all the work presented here is an attempt to obtain a better understanding of how ribosomes are made in humans. While previous work established a general eukaryotic ribosome biogenesis pathway in yeast (Woolford and Baserga 2013), newer research has shown how the regulation of this pathway becomes increasingly complex in humans (Aubert et al. 2018; Bohnsack and Bohnsack 2019). Therefore, the Baserga laboratory's focus currently has been on obtaining a better understanding of these nuances by identifying novel ribosome biogenesis factors and their molecular functions. Based on the aforementioned high-throughput screens and connection between the nucleolus and many cellular processes, there is still much work to be completed towards understanding ribosome production in human cells.

The most promising of therapeutics directly targeting ribosome biogenesis, concentrate on the rate-limiting step of pre-rRNA transcription (Laferte et al. 2006; Chedin et al. 2007; Kopp et al. 2007). BMH-21, a RNAP1 targeting small molecular inhibitor, acts by reducing RPA194 steady state levels (Peltonen et al. 2014b). BMH-21's specific mechanism is through inhibiting RNAP1 transcription elongation, leading to decreased RNAP1 gene occupancy and ultimately reduced RPA194 stability (Jacobs et al. 2022). My work revealed that RSL24D1 and PeBoW depletion leads to similar decreases in RPA194 levels (Chapter 5). While it is unclear if this is a result or causative of impaired rDNA transcription, it highlights the ability to disrupt both LSU biogenesis and rDNA transcription simultaneously. Therefore, I propose investigation into identifying both protein factor regulators and small molecule inhibitors of multiple steps of ribosome biogenesis that possess the potential for more comprehensive control over the entire process.

Instead of inhibiting ribosome biogenesis, which in turn shuts down other cellular processes, like the cell cycle, or vice versa through crosstalk mechanisms (Ogawa and Baserga 2017; Ogawa et al. 2021), why not stop multiple cancer driving forces in concert? It is becoming increasingly clear that several proteins possess roles not only in ribosome biogenesis but in other essential cellular processes as well (Berger et al. 2015; Sondalle et al. 2019; McCool et al. 2022a). These dual functions make them promising targets for the inhibition of multiple cancer driving processes at once. In my studies, APOBEC3A emerges as a promising candidate target to this end. Development of APOBEC3A specific inhibitors with an eye towards cancer treatments has already commenced (Kurup et al. 2022; Serrano et al. 2022). With APOBEC3A's newfound role in ribosome biogenesis (Chapter 6), it becomes an even more promising therapeutic strategy to shut down its role in both producing genomic mutations and ribosome biogenesis at the same time. APOBEC3A is just one example. Keeping in mind that the cell uses the nucleolus as a regulatory hub (Lindstrom et al. 2018), it seems appropriate to leverage this fact in the development of specific cancer therapeutics against these multifunctional proteins.

Work towards a more complete understanding of making ribosomes will provide us with more than just knowledge of what is happening in a healthy cell. Since ribosome biogenesis is such an energy intensive and essential process, it is unsurprisingly dysregulated in disease. An increase in making ribosomes is tightly connected to cancer pathogenesis while defects are associated with ribosomopathies. Thus, increased knowledge of how this pathway works on a molecular level will undoubtedly allow for the development of better and more targeted therapeutics to combat a number of these diverse diseases in the future.

References

Ahmad Y, Boisvert FM, Gregor P, Cobley A, Lamond AI. 2009. NOPdb: Nucleolar Proteome Database--2008 update. *Nucleic Acids Res* **37**: D181-184. doi:10.1093/nar/gkn804.

Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, Boot A, Covington KR, Gordenin DA, Bergstrom EN et al. 2020. The repertoire of mutational signatures in human cancer. *Nature* **578**: 94-101. doi:10.1038/s41586-020-1943-3.

Allende ML, Amsterdam A, Becker T, Kawakami K, Gaiano N, Hopkins N. 1996. Insertional mutagenesis in zebrafish identifies two novel genes, pescadillo and dead eye, essential for embryonic development. *Genes Dev* **10**: 3141-3155. doi:10.1101/gad.10.24.3141.

Alqassim EY, Sharma S, Khan A, Emmons TR, Cortes Gomez E, Alahmari A, Singel KL, Mark J, Davidson BA, Robert McGray AJ et al. 2021. RNA editing enzyme APOBEC3A promotes pro-inflammatory M1 macrophage polarization. *Commun Biol* **4**: 102. doi:10.1038/s42003-020-01620-x.

Anadon C, van Tetering G, Ferreira HJ, Moutinho C, Martinez-Cardus A, Villanueva A, Soler M, Heyn H, Moran S, Castro de Moura M et al. 2017. Epigenetic loss of the RNA decapping enzyme NUDT16 mediates C-MYC activation in T-cell acute lymphoblastic leukemia. *Leukemia* **31**: 1622-1625. doi:10.1038/leu.2017.99.

Antony C, George SS, Blum J, Somers P, Thorsheim CL, Wu-Corts DJ, Ai Y, Gao L, Lv K, Tremblay MG et al. 2022. Control of ribosomal RNA synthesis by hematopoietic transcription factors. *Mol Cell* **82**: 3826-3839 e3829. doi:10.1016/j.molcel.2022.08.027.

Arabi A, Wu S, Ridderstrale K, Bierhoff H, Shiue C, Fatyol K, Fahlen S, Hydbring P, Soderberg O, Grummt I et al. 2005. c-Myc associates with ribosomal DNA and activates RNA polymerase I transcription. *Nat Cell Biol* **7**: 303-310. doi:10.1038/ncb1225.

Arribas-Layton M, Wu D, Lykke-Andersen J, Song H. 2013. Structural and functional control of the eukaryotic mRNA decapping machinery. *Biochim Biophys Acta* **1829**: 580-589. doi:10.1016/j.bbagrm.2012.12.006.

Asaoka M, Ishikawa T, Takabe K, Patnaik SK. 2019. APOBEC3-Mediated RNA Editing in Breast Cancer is Associated with Heightened Immune Activity and Improved Survival. *Int J Mol Sci* **20**. doi:10.3390/ijms20225621.

Aubert M, O'Donohue MF, Lebaron S, Gleizes PE. 2018. Pre-Ribosomal RNA Processing in Human Cells: From Mechanisms to Congenital Diseases. *Biomolecules* **8**. doi:10.3390/biom8040123.

Badertscher L, Wild T, Montellese C, Alexander LT, Bammert L, Sarazova M, Stebler M, Csucs G, Mayer TU, Zamboni N et al. 2015. Genome-wide RNAi Screening Identifies Protein Modules Required for 40S Subunit Synthesis in Human Cells. *Cell Rep* **13**: 2879-2891. doi:10.1016/j.celrep.2015.11.061.

Bai B, Moore HM, Laiho M. 2013. CRM1 and its ribosome export adaptor NMD3 localize to the nucleolus and affect rRNA synthesis. *Nucleus* **4**: 315-325. doi:10.4161/nucl.25342.

Barandun J, Chaker-Margot M, Hunziker M, Molloy KR, Chait BT, Klinge S. 2017. The complete structure of the small-subunit processome. *Nat Struct Mol Biol* **24**: 944-953. doi:10.1038/nsmb.3472.

Barka A, Berrios KN, Bailer P, Schutsky EK, Wang T, Kohli RM. 2022. The Base-Editing Enzyme APOBEC3A Catalyzes Cytosine Deamination in RNA with Low

Proficiency and High Selectivity. *ACS Chem Biol* **17**: 629-636. doi:10.1021/acschembio.1c00919.

Barriga EH, Trainor PA, Bronner M, Mayor R. 2015. Animal models for studying neural crest development: is the mouse different? *Development* **142**: 1555-1560. doi:10.1242/dev.121590.

Barros Fontes MI, Dos Santos AP, Rossi Torres F, Lopes-Cendes I, Cendes F, Appenzeller S, Kawasaki de Araujo T, Lopes Monlleo I, Gil-da-Silva-Lopes VL. 2017. 17p13.3 Microdeletion: Insights on Genotype-Phenotype Correlation. *Mol Syndromol* **8**: 36-41. doi:10.1159/000452753.

Bassler J, Hurt E. 2019. Eukaryotic Ribosome Assembly. *Annu Rev Biochem* **88**: 281-306. doi:10.1146/annurev-biochem-013118-110817.

Baysal BE, Sharma S, Hashemikhabir S, Janga SC. 2017. RNA Editing in Pathogenesis of Cancer. *Cancer Res* **77**: 3733-3739. doi:10.1158/0008-5472.CAN-17-0520.

Beggs JD. 2005. Lsm proteins and RNA processing. *Biochem Soc Trans* **33**: 433-438. doi:10.1042/BST0330433.

Berger CM, Gaume X, Bouvet P. 2015. The roles of nucleolin subcellular localization in cancer. *Biochimie* **113**: 78-85. doi:10.1016/j.biochi.2015.03.023.

Bierhoff H, Schmitz K, Maass F, Ye J, Grummt I. 2010. Noncoding transcripts in sense and antisense orientation regulate the epigenetic state of ribosomal RNA genes. *Cold Spring Harb Symp Quant Biol* **75**: 357-364. doi:10.1101/sqb.2010.75.060.

Birmingham A, Anderson EM, Reynolds A, Ilsley-Tyree D, Leake D, Fedorov Y, Baskerville S, Maksimova E, Robinson K, Karpilow J et al. 2006. 3' UTR seed matches, but not overall identity, are associated with RNAi off-targets. *Nat Methods* **3**: 199-204. doi:10.1038/nmeth854.

Birnstiel ML, Wallace H, Sirlin JL, Fischberg M. 1966. Localization of the ribosomal DNA complements in the nucleolar organizer region of Xenopus laevis. *Natl Cancer Inst Monogr* **23**: 431-447.

Blazejewski SM, Bennison SA, Smith TH, Toyo-Oka K. 2018. Neurodevelopmental Genetic Diseases Associated With Microdeletions and Microduplications of Chromosome 17p13.3. *Front Genet* **9**: 80. doi:10.3389/fgene.2018.00080.

Bodem J, Dobreva G, Hoffmann-Rohrer U, Iben S, Zentgraf H, Delius H, Vingron M, Grummt I. 2000. TIF-IA, the factor mediating growth-dependent control of ribosomal RNA synthesis, is the mammalian homolog of yeast Rrn3p. *EMBO Rep* **1**: 171-175. doi:10.1093/embo-reports/kvd032.

Bohnsack KE, Bohnsack MT. 2019. Uncovering the assembly pathway of human ribosomes and its emerging links to disease. *EMBO J* **38**: e100278. doi:10.15252/embj.2018100278.

Boisvert FM, van Koningsbruggen S, Navascues J, Lamond AI. 2007. The multifunctional nucleolus. *Nat Rev Mol Cell Biol* **8**: 574-585. doi:10.1038/nrm2184.

Bonhoure N, Praz V, Moir RD, Willemin G, Mange F, Moret C, Willis IM, Hernandez N. 2020. MAF1 is a chronic repressor of RNA polymerase III transcription in the mouse. *Sci Rep* **10**: 11956. doi:10.1038/s41598-020-68665-0.

Boocock GR, Morrison JA, Popovic M, Richards N, Ellis L, Durie PR, Rommens JM. 2003. Mutations in SBDS are associated with Shwachman-Diamond syndrome. *Nat Genet* **33**: 97-101. doi:10.1038/ng1062.

Boon KL, Pearson MD, Kos M. 2015. Self-association of Trimethylguanosine Synthase Tgs1 is required for efficient snRNA/snoRNA trimethylation and pre-rRNA processing. *Sci Rep* **5**: 11282. doi:10.1038/srep11282.

Boulon S, Westman BJ, Hutten S, Boisvert FM, Lamond AI. 2010. The nucleolus under stress. *Mol Cell* **40**: 216-227. doi:10.1016/j.molcel.2010.09.024.

Brehme M, Hantschel O, Colinge J, Kaupe I, Planyavsky M, Kocher T, Mechtler K, Bennett KL, Superti-Furga G. 2009. Charting the molecular network of the drug target Bcr-Abl. *Proc Natl Acad Sci U S A* **106**: 7414-7419. doi:10.1073/pnas.0900653106.

Brown IN, Lafita-Navarro MC, Conacci-Sorrell M. 2022. Regulation of Nucleolar Activity by MYC. *Cells* **11**. doi:10.3390/cells11030574.

Bruno DL, Anderlid BM, Lindstrand A, van Ravenswaaij-Arts C, Ganesamoorthy D, Lundin J, Martin CL, Douglas J, Nowak C, Adam MP et al. 2010. Further molecular and clinical delineation of co-locating 17p13.3 microdeletions and microduplications that show distinctive phenotypes. *J Med Genet* **47**: 299-311. doi:10.1136/jmg.2009.069906.

Bruno PM, Lu M, Dennis KA, Inam H, Moore CJ, Sheehe J, Elledge SJ, Hemann MT, Pritchard JR. 2020. The primary mechanism of cytotoxicity of the chemotherapeutic agent CX-5461 is topoisomerase II poisoning. *Proc Natl Acad Sci U S A* **117**: 4053-4060. doi:10.1073/pnas.1921649117.

Bryant CJ, Lorea CF, de Almeida HL, Jr., Weinert L, Vedolin L, Pinto EVF, Baserga SJ. 2021. Biallelic splicing variants in the nucleolar 60S assembly factor RBM28 cause the ribosomopathy ANE syndrome. *Proc Natl Acad Sci U S A* **118**. doi:10.1073/pnas.2017777118.

Bryant CJ, McCool MA, Abriola L, Surovtseva YV, Baserga SJ. 2022. A high-throughput assay for directly monitoring nucleolar rRNA biogenesis. *Open Biol* **12**: 210305. doi:10.1098/rsob.210305.

Bryant CJ, McCool MA, Rosado-Gonzalez GT, Abriola L, Surovtseva YV, Baserga SJ. 2023. Discovery of novel microRNA regulators of ribosome biogenesis. *bioRxiv*. doi:10.1101/2023.02.17.526327.

Buisson R, Langenbucher A, Bowen D, Kwan EE, Benes CH, Zou L, Lawrence MS. 2019. Passenger hotspot mutations in cancer driven by APOBEC3A and mesoscale genomic features. *Science* **364**. doi:10.1126/science.aaw2872.

Bywater MJ, Poortinga G, Sanij E, Hein N, Peck A, Cullinane C, Wall M, Cluse L, Drygin D, Anderes K et al. 2012. Inhibition of RNA polymerase I as a therapeutic strategy to promote cancer-specific activation of p53. *Cancer Cell* **22**: 51-65. doi:10.1016/j.ccr.2012.05.019.

Calo E, Gu B, Bowen ME, Aryan F, Zalc A, Liang J, Flynn RA, Swigut T, Chang HY, Attardi LD et al. 2018. Tissue-selective effects of nucleolar stress and rDNA damage in developmental disorders. *Nature* **554**: 112-117. doi:10.1038/nature25449.

Carbon S, Mungall C. 2021. Gene Ontology Data Archive (Version 2021-07-02).

Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, Guertin DA, Chang JH, Lindquist RA, Moffat J et al. 2006. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol* **7**: R100. doi:10.1186/gb-2006-7-10-r100.

Carroll PA, Freie BW, Mathsyaraja H, Eisenman RN. 2018. The MYC transcription factor network: balancing metabolism, proliferation and oncogenesis. *Front Med* **12**: 412-425. doi:10.1007/s11684-018-0650-z.

Catez F, Dalla Venezia N, Marcel V, Zorbas C, Lafontaine DLJ, Diaz JJ. 2019. Ribosome biogenesis: An emerging druggable pathway for cancer therapeutics. *Biochem Pharmacol* **159**: 74-81. doi:10.1016/j.bcp.2018.11.014.

Chaker-Margot M, Hunziker M, Barandun J, Dill BD, Klinge S. 2015. Stage-specific assembly events of the 6-MDa small-subunit processome initiate eukaryotic ribosome biogenesis. *Nat Struct Mol Biol* **22**: 920-923. doi:10.1038/nsmb.3111.

Chan GK, Kleinheinz TL, Peterson D, Moffat JG. 2013. A simple high-content cell cycle assay reveals frequent discrepancies between cell number and ATP and MTS proliferation assays. *PLoS One* **8**: e63583. doi:10.1371/journal.pone.0063583.

Chedin S, Laferte A, Hoang T, Lafontaine DL, Riva M, Carles C. 2007. Is ribosome synthesis controlled by pol I transcription? *Cell Cycle* **6**: 11-15. doi:10.4161/cc.6.1.3649.

Chen H, Shi Z, Guo J, Chang KJ, Chen Q, Yao CH, Haigis MC, Shi Y. 2020. The human mitochondrial 12S rRNA m(4)C methyltransferase METTL15 is required for mitochondrial function. *J Biol Chem* **295**: 8505-8513. doi:10.1074/jbc.RA119.012127.

Chen JY, Tan X, Wang ZH, Liu YZ, Zhou JF, Rong XZ, Lu L, Li Y. 2018. The ribosome biogenesis protein Esf1 is essential for pharyngeal cartilage formation in zebrafish. *FEBS J* **285**: 3464-3484. doi:10.1111/febs.14622.

Cheng F, Belting M, Fransson LA, Mani K. 2017. Nucleolin is a nuclear target of heparan sulfate derived from glypican-1. *Exp Cell Res* **354**: 31-39. doi:10.1016/j.yexcr.2017.03.021.

Cheng Y, Zhu H, Du Z, Guo X, Zhou C, Wang Z, He X. 2021. Eukaryotic translation factor eIF5A contributes to acetic acid tolerance in Saccharomyces cerevisiae via transcriptional factor Ume6p. *Biotechnol Biofuels* **14**: 38. doi:10.1186/s13068-021-01885-2.

Ciganda M, Williams N. 2011. Eukaryotic 5S rRNA biogenesis. *Wiley Interdiscip Rev RNA* **2**: 523-533. doi:10.1002/wrna.74.

Corsini NS, Peer AM, Moeseneder P, Roiuk M, Burkard TR, Theussl HC, Moll I, Knoblich JA. 2018. Coordinated Control of mRNA and rRNA Processing Controls Embryonic Stem Cell Pluripotency and Differentiation. *Cell Stem Cell* **22**: 543-558 e512. doi:10.1016/j.stem.2018.03.002.

Cortez LM, Brown AL, Dennis MA, Collins CD, Brown AJ, Mitchell D, Mertz TM, Roberts SA. 2019. APOBEC3A is a prominent cytidine deaminase in breast cancer. *PLoS Genet* **15**: e1008545. doi:10.1371/journal.pgen.1008545.

Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Res* **14**: 1188-1190. doi:10.1101/gr.849004.

Daftuar L, Zhu Y, Jacq X, Prives C. 2013. Ribosomal proteins RPL37, RPS15 and RPS20 regulate the Mdm2-p53-MdmX network. *PLoS One* **8**: e68667. doi:10.1371/journal.pone.0068667.

Dalla Rosa I, Durigon R, Pearce SF, Rorbach J, Hirst EM, Vidoni S, Reyes A, Brea-Calvo G, Minczuk M, Woellhaf MW et al. 2014. MPV17L2 is required for ribosome assembly in mitochondria. *Nucleic Acids Res* **42**: 8500-8515. doi:10.1093/nar/gku513.

Dameshek W. 1967. Riddle: what do aplastic anemia, paroxysmal nocturnal hemoglobinuria (PNH) and "hypoplastic" leukemia have in common? *Blood* **30**: 251-254.

Derenzini M, Montanaro L, Trere D. 2017. Ribosome biogenesis and cancer. *Acta Histochem* **119**: 190-197. doi:10.1016/j.acthis.2017.01.009.

Derenzini M, Trere D, Pession A, Montanaro L, Sirri V, Ochs RL. 1998. Nucleolar function and size in cancer cells. *Am J Pathol* **152**: 1291-1297.

Destefanis F, Manara V, Bellosta P. 2020. Myc as a Regulator of Ribosome Biogenesis and Cell Competition: A Link to Cancer. *Int J Mol Sci* **21**. doi:10.3390/ijms21114037.

Dinsmore CJ, Soriano P. 2018. MAPK and PI3K signaling: At the crossroads of neural crest development. *Dev Biol* **444 Suppl 1**: S79-S97. doi:10.1016/j.ydbio.2018.02.003.

Doci CL, Zhou G, Lingen MW. 2013. The novel tumor suppressor NOL7 post-transcriptionally regulates thrombospondin-1 expression. *Oncogene* **32**: 4377-4386. doi:10.1038/onc.2012.464.

Dong C, An L, Yu CH, Huen MSY. 2021. A DYRK1B-dependent pathway suppresses rDNA transcription in response to DNA damage. *Nucleic Acids Res* **49**: 1485-1496. doi:10.1093/nar/gkaa1290.

Dorner K, Badertscher L, Horvath B, Hollandi R, Molnar C, Fuhrer T, Meier R, Sarazova M, van den Heuvel J, Zamboni N et al. 2022. Genome-wide RNAi screen identifies novel players in human 60S subunit biogenesis including key enzymes of polyamine metabolism. *Nucleic Acids Res* **50**: 2872-2888. doi:10.1093/nar/gkac072.

Dragon F, Gallagher JE, Compagnone-Post PA, Mitchell BM, Porwancher KA, Wehner KA, Wormsley S, Settlage RE, Shabanowitz J, Osheim Y et al. 2002. A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature* **417**: 967-970. doi:10.1038/nature00769.

Drew K, Wallingford JB, Marcotte EM. 2020. hu.MAP 2.0: Integration of over 15,000 proteomic experiments builds a global compendium of human multiprotein assemblies. *bioRxiv*: 2020.2009.2015.298216. doi:10.1101/2020.09.15.298216.

Drew K, Wallingford JB, Marcotte EM. 2021. hu.MAP 2.0: integration of over 15,000 proteomic experiments builds a global compendium of human multiprotein assemblies. *Mol Syst Biol* **17**: e10016. doi:10.15252/msb.202010016.

Drygin D, Lin A, Bliesath J, Ho CB, O'Brien SE, Proffitt C, Omori M, Haddach M, Schwaebe MK, Siddiqui-Jain A et al. 2011. Targeting RNA polymerase I with an oral small molecule CX-5461 inhibits ribosomal RNA synthesis and solid tumor growth. *Cancer Res* **71**: 1418-1430. doi:10.1158/0008-5472.CAN-10-1728.

Durand S, Bruelle M, Bourdelais F, Bennychen B, Blin-Gonthier J, Isaac C, Huyghe A, Seyve A, Vanbelle C, Meyronet D et al. 2021. RSL24D1 sustains steady-state ribosome biogenesis and pluripotency translational programs in embryonic stem cells. *bioRxiv*: 2021.2005.2027.443845. doi:10.1101/2021.05.27.443845.

Echeverri CJ, Beachy PA, Baum B, Boutros M, Buchholz F, Chanda SK, Downward J, Ellenberg J, Fraser AG, Hacohen N et al. 2006. Minimizing the risk of reporting false positives in large-scale RNAi screens. *Nat Methods* **3**: 777-779. doi:10.1038/nmeth1006-777.

el-Deiry WS, Harper JW, O'Connor PM, Velculescu VE, Canman CE, Jackman J, Pietenpol JA, Burrell M, Hill DE, Wang Y et al. 1994. WAF1/CIP1 is induced in p53-mediated G1 arrest and apoptosis. *Cancer Res* **54**: 1169-1174.

Emrick LT, Rosenfeld JA, Lalani SR, Jain M, Desai NK, Larson A, Kripps K, Vanderver A, Taft RJ, Bluske K et al. 2019. Microdeletions excluding YWHAE and PAFAH1B1 cause a unique leukoencephalopathy: further delineation of the 17p13.3 microdeletion spectrum. *Genet Med* **21**: 1652-1656. doi:10.1038/s41436-018-0358-0.

Engel KL, French SL, Viktorovskaya OV, Beyer AL, Schneider DA. 2015. Spt6 Is Essential for rRNA Synthesis by RNA Polymerase I. *Mol Cell Biol* **35**: 2321-2331. doi:10.1128/MCB.01499-14.

Eswara MB, McGuire AT, Pierce JB, Mangroo D. 2009. Utp9p facilitates Msn5p-mediated nuclear reexport of retrograded tRNAs in Saccharomyces cerevisiae. *Mol Biol Cell* **20**: 5007-5025. doi:10.1091/mbc.E09-06-0490.

Evers R, Grummt I. 1995. Molecular coevolution of mammalian ribosomal gene terminator sequences and the transcription termination factor TTF-I. *Proc Natl Acad Sci U S A* **92**: 5827-5831. doi:10.1073/pnas.92.13.5827.

Fan P, Wang B, Meng Z, Zhao J, Jin X. 2018. PES1 is transcriptionally regulated by BRD4 and promotes cell proliferation and glycolysis in hepatocellular carcinoma. *Int J Biochem Cell Biol* **104**: 1-8. doi:10.1016/j.biocel.2018.08.014.

Farley KI, Baserga SJ. 2016. Probing the mechanisms underlying human diseases in making ribosomes. *Biochem Soc Trans* **44**: 1035-1044. doi:10.1042/BST20160064.

Farley KI, Surovtseva Y, Merkel J, Baserga SJ. 2015. Determinants of mammalian nucleolar architecture. *Chromosoma* **124**: 323-331. doi:10.1007/s00412-015-0507-z.

Farley-Barnes KI, Deniz E, Overton MM, Khokha MK, Baserga SJ. 2020. Paired Box 9 (PAX9), the RNA polymerase II transcription factor, regulates human ribosome biogenesis and craniofacial development. *PLoS Genet* **16**: e1008967. doi:10.1371/journal.pgen.1008967.

Farley-Barnes KI, McCann KL, Ogawa LM, Merkel J, Surovtseva YV, Baserga SJ. 2018. Diverse Regulators of Human Ribosome Biogenesis Discovered by Changes in Nucleolar Number. *Cell Rep* **22**: 1923-1934. doi:10.1016/j.celrep.2018.01.056.

Farley-Barnes KI, Ogawa LM, Baserga SJ. 2019. Ribosomopathies: Old Concepts, New Controversies. *Trends Genet* **35**: 754-767. doi:10.1016/j.tig.2019.07.004.

Fedorov Y, Anderson EM, Birmingham A, Reynolds A, Karpilow J, Robinson K, Leake D, Marshall WS, Khvorova A. 2006. Off-target effects by siRNA can induce toxic phenotype. *RNA* **12**: 1188-1196. doi:10.1261/rna.28106.

Feller SM, Knudsen B, Hanafusa H. 1994. c-Abl kinase regulates the protein binding activity of c-Crk. *EMBO J* **13**: 2341-2351.

Feng Y, Seija N, Di Noia JM, Martin A. 2021. AID in Antibody Diversification: There and Back Again: (Trends in Immunology 41, 586-600; 2020). *Trends Immunol* **42**: 89. doi:10.1016/j.it.2020.10.011.

Ferreira R, Schneekloth JS, Jr., Panov KI, Hannan KM, Hannan RD. 2020. Targeting the RNA Polymerase I Transcription for Cancer Therapy Comes of Age. *Cells* **9**. doi:10.3390/cells9020266.

Ferreira-Cerca S, Poll G, Gleizes PE, Tschochner H, Milkereit P. 2005. Roles of eukaryotic ribosomal proteins in maturation and transport of pre-18S rRNA and ribosome function. *Mol Cell* **20**: 263-275. doi:10.1016/j.molcel.2005.09.005.

Ford E, Voit R, Liszt G, Magin C, Grummt I, Guarente L. 2006. Mammalian Sir2 homolog SIRT7 is an activator of RNA polymerase I transcription. *Genes Dev* **20**: 1075-1080. doi:10.1101/gad.1399706.

Freed EF, Baserga SJ. 2010. The C-terminus of Utp4, mutated in childhood cirrhosis, is essential for ribosome biogenesis. *Nucleic Acids Res* **38**: 4798-4806. doi:10.1093/nar/gkq185.

Freed EF, Prieto JL, McCann KL, McStay B, Baserga SJ. 2012. NOL11, implicated in the pathogenesis of North American Indian childhood cirrhosis, is required for pre-rRNA transcription and processing. *PLoS Genet* **8**: e1002892. doi:10.1371/journal.pgen.1002892.

Fu X, Xu L, Qi L, Tian H, Yi D, Yu Y, Liu S, Li S, Xu Y, Wang C. 2017. BMH-21 inhibits viability and induces apoptosis by p53-dependent nucleolar stress responses in SKOV3 ovarian cancer cells. *Oncol Rep* **38**: 859-865. doi:10.3892/or.2017.5750.

Fumagalli S, Ivanenkov VV, Teng T, Thomas G. 2012. Suprainduction of p53 by disruption of 40S and 60S ribosome biogenesis leads to the activation of a novel G2/M checkpoint. *Genes Dev* **26**: 1028-1040. doi:10.1101/gad.189951.112.

Gallagher JE, Dunbar DA, Granneman S, Mitchell BM, Osheim Y, Beyer AL, Baserga SJ. 2004. RNA polymerase I transcription and pre-rRNA processing are linked by specific SSU processome components. *Genes Dev* **18**: 2506-2517. doi:10.1101/gad.1226604.

Gartner C, Messmer A, Dietmann P, Kuhl M, Kuhl SJ. 2022. Functions of block of proliferation 1 during anterior development in Xenopus laevis. *PLoS One* **17**: e0273507. doi:10.1371/journal.pone.0273507.

Gaviraghi M, Vivori C, Pareja Sanchez Y, Invernizzi F, Cattaneo A, Santoliquido BM, Frenquelli M, Segalla S, Bachi A, Doglioni C et al. 2018. Tumor suppressor PNRC1 blocks rRNA maturation by recruiting the decapping complex to the nucleolus. *EMBO J* **37**. doi:10.15252/embj.201899179.

Gaviraghi M, Vivori C, Tonon G. 2019. How Cancer Exploits Ribosomal RNA Biogenesis: A Journey beyond the Boundaries of rRNA Transcription. *Cells* **8**. doi:10.3390/cells8091098.

Ge X, Yuan L, Cheng B, Dai K. 2021. Identification of seven tumor-educated platelets RNAs for cancer diagnosis. *J Clin Lab Anal* **35**: e23791. doi:10.1002/jcla.23791.

Genuth NR, Barna M. 2018. Heterogeneity and specialized functions of translation machinery: from genes to organisms. *Nat Rev Genet* **19**: 431-452. doi:10.1038/s41576-018-0008-z.

Genuth NR, Shi Z, Kunimoto K, Hung V, Xu AF, Kerr CH, Tiu GC, Oses-Prieto JA, Salomon-Shulman REA, Axelrod JD et al. 2022. A stem cell roadmap of ribosome heterogeneity reveals a function for RPL10A in mesoderm production. *Nat Commun* **13**: 5491. doi:10.1038/s41467-022-33263-3.

Gessert S, Maurus D, Rossner A, Kuhl M. 2007. Pescadillo is required for Xenopus laevis eye development and neural crest migration. *Dev Biol* **310**: 99-112. doi:10.1016/j.ydbio.2007.07.037.

Ghosh T, Peterson B, Tomasevic N, Peculis BA. 2004. Xenopus U8 snoRNA binding protein is a conserved nuclear decapping enzyme. *Mol Cell* **13**: 817-828. doi:10.1016/s1097-2765(04)00127-3.

Ghoshal K, Majumder S, Datta J, Motiwala T, Bai S, Sharma SM, Frankel W, Jacob ST. 2004. Role of human ribosomal RNA (rRNA) promoter methylation and of methyl-CpG-binding protein MBD2 in the suppression of rRNA gene expression. *J Biol Chem* **279**: 6783-6793. doi:10.1074/jbc.M309393200.

Goldman MJ, Craft B, Hastie M, Repecka K, McDade F, Kamath A, Banerjee A, Luo Y, Rogers D, Brooks AN et al. 2020. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat Biotechnol* **38**: 675-678. doi:10.1038/s41587-020-0546-8.

Goodfellow SJ, Zomerdijk JC. 2013. Basic mechanisms in RNA polymerase I transcription of the ribosomal RNA genes. *Subcell Biochem* **61**: 211-236. doi:10.1007/978-94-007-4525-4_10.

Gorski JJ, Pathak S, Panov K, Kasciukovic T, Panova T, Russell J, Zomerdijk JC. 2007. A novel TBP-associated factor of SL1 functions in RNA polymerase I transcription. *EMBO J* **26**: 1560-1568. doi:10.1038/sj.emboj.7601601.

Grainger RM. 2012. Xenopus tropicalis as a model organism for genetics and genomics: past, present, and future. *Methods Mol Biol* **917**: 3-15. doi:10.1007/978-1-61779-992-1_1.

Grandori C, Gomez-Roman N, Felton-Edkins ZA, Ngouenet C, Galloway DA, Eisenman RN, White RJ. 2005. c-Myc binds to human ribosomal DNA and stimulates transcription of rRNA genes by RNA polymerase I. *Nat Cell Biol* **7**: 311-318. doi:10.1038/ncb1224.

Greeve J, Navaratnam N, Scott J. 1991. Characterization of the apolipoprotein B mRNA editing enzyme: no similarity to the proposed mechanism of RNA editing in kinetoplastid protozoa. *Nucleic Acids Res* **19**: 3569-3576. doi:10.1093/nar/19.13.3569.

Griffin JN, Sondalle SB, Del Viso F, Baserga SJ, Khokha MK. 2015. The ribosome biogenesis factor Nol11 is required for optimal rDNA transcription and craniofacial development in Xenopus. *PLoS Genet* **11**: e1005018. doi:10.1371/journal.pgen.1005018.

Griffin JN, Sondalle SB, Robson A, Mis EK, Griffin G, Kulkarni SS, Deniz E, Baserga SJ, Khokha MK. 2018. RPSA, a candidate gene for isolated congenital asplenia, is required for pre-rRNA processing and spleen formation in Xenopus. *Development* **145**. doi:10.1242/dev.166181.

Grillo MJ, Jones KFM, Carpenter MA, Harris RS, Harki DA. 2022. The current toolbox for APOBEC drug discovery. *Trends Pharmacol Sci* **43**: 362-377. doi:10.1016/j.tips.2022.02.007.

Grimm T, Holzel M, Rohrmoser M, Harasim T, Malamoussi A, Gruber-Eber A, Kremmer E, Eick D. 2006. Dominant-negative Pes1 mutants inhibit ribosomal RNA processing and cell proliferation via incorporation into the PeBoW-complex. *Nucleic Acids Res* **34**: 3030-3043. doi:10.1093/nar/gkl378.

Guo H. 2018. Specialized ribosomes and the control of translation. *Biochem Soc Trans* **46**: 855-869. doi:10.1042/BST20160426.

Guthrie C, Nashimoto H, Nomura M. 1969. Structure and function of E. coli ribosomes. 8. Cold-sensitive mutants defective in ribosome assembly. *Proc Natl Acad Sci U S A* **63**: 384-391. doi:10.1073/pnas.63.2.384.

Hanahan D, Weinberg RA. 2011. Hallmarks of cancer: the next generation. *Cell* **144**: 646-674. doi:10.1016/j.cell.2011.02.013.

Hannan KM, Brandenburger Y, Jenkins A, Sharkey K, Cavanaugh A, Rothblum L, Moss T, Poortinga G, McArthur GA, Pearson RB et al. 2003. 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. doi:10.1128/mcb.23.23.8862-8877.2003.

Hannan KM, Soo P, Wong MS, Lee JK, Hein N, Poh P, Wysoke KD, Williams TD, Montellese C, Smith LK et al. 2022. Nuclear stabilization of p53 requires a functional nucleolar surveillance pathway. *Cell Rep* **41**: 111571. doi:10.1016/j.celrep.2022.111571.

Hannus M, Beitzinger M, Engelmann JC, Weickert MT, Spang R, Hannus S, Meister G. 2014. siPools: highly complex but accurately defined siRNA pools eliminate off-target effects. *Nucleic Acids Res* **42**: 8049-8061. doi:10.1093/nar/gku480.

Harold CM, Buhagiar AF, Cheng Y, Baserga SJ. 2021. Ribosomal RNA Transcription Regulation in Breast Cancer. *Genes (Basel)* **12**. doi:10.3390/genes12040502.

Harris RS, Dudley JP. 2015. APOBECs and virus restriction. *Virology* **479-480**: 131-145. doi:10.1016/j.virol.2015.03.012.

Harris RS, Petersen-Mahrt SK, Neuberger MS. 2002. RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol Cell* **10**: 1247-1253. doi:10.1016/s1097-2765(02)00742-6.

Hart T, Komori HK, LaMere S, Podshivalova K, Salomon DR. 2013. Finding the active genes in deep RNA-seq gene expression studies. *BMC Genomics* **14**: 778. doi:10.1186/1471-2164-14-778.

Hasina R, Pontier AL, Fekete MJ, Martin LE, Qi XM, Brigaudeau C, Pramanik R, Cline EI, Coignet LJ, Lingen MW. 2006. NOL7 is a nucleolar candidate tumor suppressor gene in cervical cancer that modulates the angiogenic phenotype. *Oncogene* **25**: 588-598. doi:10.1038/sj.onc.1209070.

Hayashi Y, Fujimura A, Kato K, Udagawa R, Hirota T, Kimura K. 2018. Nucleolar integrity during interphase supports faithful Cdk1 activation and mitotic entry. *Sci Adv* **4**: eaap7777. doi:10.1126/sciadv.aap7777.

Hayashi Y, Kato K, Kimura K. 2017. The hierarchical structure of the perichromosomal layer comprises Ki67, ribosomal RNAs, and nucleolar proteins. *Biochem Biophys Res Commun* **493**: 1043-1049. doi:10.1016/j.bbrc.2017.09.092.

He JS, Soo P, Evers M, Parsons KM, Hein N, Hannan KM, Hannan RD, George AJ. 2018. High-Content Imaging Approaches to Quantitate Stress-Induced Changes in Nucleolar Morphology. *Assay Drug Dev Technol* **16**: 320-332. doi:10.1089/adt.2018.861.

Henras AK, Plisson-Chastang C, O'Donohue MF, Chakraborty A, Gleizes PE. 2015. An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip Rev RNA* **6**: 225-242. doi:10.1002/wrna.1269.

Holzel M, Rohrmoser M, Schlee M, Grimm T, Harasim T, Malamoussi A, Gruber-Eber A, Kremmer E, Hiddemann W, Bornkamm GW et al. 2005. Mammalian WDR12 is a novel member of the Pes1-Bop1 complex and is required for ribosome biogenesis and cell proliferation. *J Cell Biol* **170**: 367-378. doi:10.1083/jcb.200501141.

Hori Y, Engel C, Kobayashi T. 2023. Regulation of ribosomal RNA gene copy number, transcription and nucleolus organization in eukaryotes. *Nat Rev Mol Cell Biol*. doi:10.1038/s41580-022-00573-9.

Hu Y, Flockhart I, Vinayagam A, Bergwitz C, Berger B, Perrimon N, Mohr SE. 2011. An integrative approach to ortholog prediction for disease-focused and other functional studies. *BMC Bioinformatics* **12**: 357. doi:10.1186/1471-2105-12-357.

Hua L, Yan D, Wan C, Hu B. 2022. Nucleolus and Nucleolar Stress: From Cell Fate Decision to Disease Development. *Cells* **11**. doi:10.3390/cells11193017.

Huang YC, Tseng SF, Tsai HJ, Lenzmeier BA, Teng SC. 2010. Direct interaction between Utp8p and Utp9p contributes to rRNA processing in budding yeast. *Biochem Biophys Res Commun* **393**: 297-302. doi:10.1016/j.bbrc.2010.02.003.

Hughes JM, Ares M, Jr. 1991. Depletion of U3 small nucleolar RNA inhibits cleavage in the 5' external transcribed spacer of yeast pre-ribosomal RNA and impairs formation of 18S ribosomal RNA. *EMBO J* **10**: 4231-4239. doi:10.1002/j.1460-2075.1991.tb05001.x.

Huttlin EL, Bruckner RJ, Paulo JA, Cannon JR, Ting L, Baltier K, Colby G, Gebreab F, Gygi MP, Parzen H et al. 2017. Architecture of the human interactome defines protein communities and disease networks. *Nature* **545**: 505-509. doi:10.1038/nature22366.

Hwang HW, Wentzel EA, Mendell JT. 2007. A hexanucleotide element directs microRNA nuclear import. *Science* **315**: 97-100. doi:10.1126/science.1136235.

Iadevaia V, Liu R, Proud CG. 2014. mTORC1 signaling controls multiple steps in ribosome biogenesis. *Semin Cell Dev Biol* **36**: 113-120. doi:10.1016/j.semcdb.2014.08.004.

Jackson AL, Bartz SR, Schelter J, Kobayashi SV, Burchard J, Mao M, Li B, Cavet G, Linsley PS. 2003. Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol* **21**: 635-637. doi:10.1038/nbt831.

Jackson AL, Burchard J, Leake D, Reynolds A, Schelter J, Guo J, Johnson JM, Lim L, Karpilow J, Nichols K et al. 2006. Position-specific chemical modification of siRNAs reduces "off-target" transcript silencing. *RNA* **12**: 1197-1205. doi:10.1261/rna.30706.

Jackson DA, Pombo A, Iborra F. 2000. The balance sheet for transcription: an analysis of nuclear RNA metabolism in mammalian cells. *Faseb J* **14**: 242-254. doi:10.1096/fasebj.14.2.242.

Jacobs RQ, Huffines AK, Laiho M, Schneider DA. 2022. The small-molecule BMH-21 directly inhibits transcription elongation and DNA occupancy of RNA polymerase I in vivo and in vitro. *J Biol Chem* **298**: 101450. doi:10.1016/j.jbc.2021.101450.

Jalal C, Uhlmann-Schiffler H, Stahl H. 2007. Redundant role of DEAD box proteins p68 (Ddx5) and p72/p82 (Ddx17) in ribosome biogenesis and cell proliferation. *Nucleic Acids Res* **35**: 3590-3601. doi:10.1093/nar/gkm058.

Jalili P, Bowen D, Langenbucher A, Park S, Aguirre K, Corcoran RB, Fleischman AG, Lawrence MS, Zou L, Buisson R. 2020. Quantification of ongoing APOBEC3A activity in tumor cells by monitoring RNA editing at hotspots. *Nat Commun* **11**: 2971. doi:10.1038/s41467-020-16802-8.

James A, Wang Y, Raje H, Rosby R, DiMario P. 2014. Nucleolar stress with and without p53. *Nucleus* **5**: 402-426. doi:10.4161/nucl.32235.

Jarboui MA, Wynne K, Elia G, Hall WW, Gautier VW. 2011. Proteomic profiling of the human T-cell nucleolus. *Mol Immunol* **49**: 441-452. doi:10.1016/j.molimm.2011.09.005.

Jenkinson EM, Rodero MP, Kasher PR, Uggenti C, Oojageer A, Goosey LC, Rose Y, Kershaw CJ, Urquhart JE, Williams SG et al. 2016. Mutations in SNORD118 cause the cerebral microangiopathy leukoencephalopathy with calcifications and cysts. *Nat Genet* **48**: 1185-1192. doi:10.1038/ng.3661.

Kappel L, Loibl M, Zisser G, Klein I, Fruhmann G, Gruber C, Unterweger S, Rechberger G, Pertschy B, Bergler H. 2012. Rlp24 activates the AAA-ATPase Drg1 to initiate cytoplasmic pre-60S maturation. *J Cell Biol* **199**: 771-782. doi:10.1083/jcb.201205021.

Kellner M, Rohrmoser M, Forne I, Voss K, Burger K, Muhl B, Gruber-Eber A, Kremmer E, Imhof A, Eick D. 2015. DEAD-box helicase DDX27 regulates 3' end formation of ribosomal 47S RNA and stably associates with the PeBoW-complex. *Exp Cell Res* **334**: 146-159. doi:10.1016/j.yexcr.2015.03.017.

Killian A, Sarafan-Vasseur N, Sesboue R, Le Pessot F, Blanchard F, Lamy A, Laurent M, Flaman JM, Frebourg T. 2006. Contribution of the BOP1 gene, located on 8q24, to colorectal tumorigenesis. *Genes Chromosomes Cancer* **45**: 874-881. doi:10.1002/gcc.20351.

Kim W, Underwood RS, Greenwald I, Shaye DD. 2018. OrthoList 2: A New Comparative Genomic Analysis of Human and Caenorhabditis elegans Genes. *Genetics* **210**: 445-461. doi:10.1534/genetics.118.301307.

Kirsch VC, Orgler C, Braig S, Jeremias I, Auerbach D, Muller R, Vollmar AM, Sieber SA. 2020. The Cytotoxic Natural Product Vioprolide A Targets Nucleolar Protein 14, Which Is Essential for Ribosome Biogenesis. *Angew Chem Int Ed Engl* **59**: 1595-1600. doi:10.1002/anie.201911158.

Knudsen BS, Zheng J, Feller SM, Mayer JP, Burrell SK, Cowburn D, Hanafusa H. 1995. Affinity and specificity requirements for the first Src homology 3 domain of the Crk proteins. *EMBO J* **14**: 2191-2198.

Kobashigawa Y, Sakai M, Naito M, Yokochi M, Kumeta H, Makino Y, Ogura K, Tanaka S, Inagaki F. 2007. Structural basis for the transforming activity of human cancer-related signaling adaptor protein CRK. *Nat Struct Mol Biol* **14**: 503-510. doi:10.1038/nsmb1241.

Koiwai K, Noma S, Takahashi Y, Hayano T, Maezawa S, Kouda K, Matsumoto T, Suzuki M, Furuichi M, Koiwai O. 2011. TdIF2 is a nucleolar protein that promotes rRNA gene promoter activity. *Genes Cells* **16**: 748-764. doi:10.1111/j.1365-2443.2011.01524.x.

Kopp K, Gasiorowski JZ, Chen D, Gilmore R, Norton JT, Wang C, Leary DJ, Chan EK, Dean DA, Huang S. 2007. Pol I transcription and pre-rRNA processing are coordinated in a transcription-dependent manner in mammalian cells. *Mol Biol Cell* **18**: 394-403. doi:10.1091/mbc.e06-03-0249.

Kouno T, Silvas TV, Hilbert BJ, Shandilya SMD, Bohn MF, Kelch BA, Royer WE, Somasundaran M, Kurt Yilmaz N, Matsuo H et al. 2017. Crystal structure of APOBEC3A bound to single-stranded DNA reveals structural basis for cytidine deamination and specificity. *Nat Commun* **8**: 15024. doi:10.1038/ncomms15024.

Krogan NJ, Peng WT, Cagney G, Robinson MD, Haw R, Zhong G, Guo X, Zhang X, Canadien V, Richards DP et al. 2004. High-definition macromolecular composition of yeast RNA-processing complexes. *Mol Cell* **13**: 225-239. doi:10.1016/s1097-2765(04)00003-6.

Kufel J, Allmang C, Petfalski E, Beggs J, Tollervey D. 2003. Lsm Proteins are required for normal processing and stability of ribosomal RNAs. *J Biol Chem* **278**: 2147-2156. doi:10.1074/jbc.M208856200.

Kurup HM, Kvach MV, Harjes S, Barzak FM, Jameson GB, Harjes E, Filichev VV. 2022. Design, Synthesis, and Evaluation of a Cross-Linked Oligonucleotide as the First Nanomolar Inhibitor of APOBEC3A. *Biochemistry* **61**: 2568-2578. doi:10.1021/acs.biochem.2c00449.

Lackey L, Law EK, Brown WL, Harris RS. 2013. Subcellular localization of the APOBEC3 proteins during mitosis and implications for genomic DNA deamination. *Cell Cycle* **12**: 762-772. doi:10.4161/cc.23713.

Laferte A, Favry E, Sentenac A, Riva M, Carles C, Chedin S. 2006. The transcriptional activity of RNA polymerase I is a key determinant for the level of all ribosome components. *Genes Dev* **20**: 2030-2040. doi:10.1101/gad.386106.

Lafita-Navarro MC, Blanco R, Mata-Garrido J, Liano-Pons J, Tapia O, Garcia-Gutierrez L, Garcia-Alegria E, Berciano MT, Lafarga M, Leon J. 2016. MXD1 localizes in the nucleolus, binds UBF and impairs rRNA synthesis. *Oncotarget* **7**: 69536-69548. doi:10.18632/oncotarget.11766.

Langhendries JL, Nicolas E, Doumont G, Goldman S, Lafontaine DL. 2016. The human box C/D snoRNAs U3 and U8 are required for pre-rRNA processing and tumorigenesis. *Oncotarget* **7**: 59519-59534. doi:10.18632/oncotarget.11148.

Lapik YR, Fernandes CJ, Lau LF, Pestov DG. 2004. Physical and functional interaction between Pes1 and Bop1 in mammalian ribosome biogenesis. *Mol Cell* **15**: 17-29. doi:10.1016/j.molcel.2004.05.020.

Lerch-Gaggl A, Haque J, Li J, Ning G, Traktman P, Duncan SA. 2002. Pescadillo is essential for nucleolar assembly, ribosome biogenesis, and mammalian cell proliferation. *J Biol Chem* **277**: 45347-45355. doi:10.1074/jbc.M208338200.

Leung AK, Trinkle-Mulcahy L, Lam YW, Andersen JS, Mann M, Lamond AI. 2006. NOPdb: Nucleolar Proteome Database. *Nucleic Acids Res* **34**: D218-220. doi:10.1093/nar/gkj004.

Li G, Wu XJ, Kong XQ, Wang L, Jin X. 2015. Cytochrome c oxidase subunit VIIb as a potential target in familial hypercholesterolemia by bioinformatical analysis. *Eur Rev Med Pharmacol Sci* **19**: 4139-4145.

Li JL, Chen C, Chen W, Zhao LF, Xu XK, Li Y, Yuan HY, Lin JR, Pan JP, Jin BL et al. 2020. Integrative genomic analyses identify WDR12 as a novel oncogene involved in glioblastoma. *J Cell Physiol* **235**: 7344-7355. doi:10.1002/jcp.29635.

Li Y, Kiledjian M. 2010. Regulation of mRNA decapping. *Wiley Interdiscip Rev RNA* **1**: 253-265. doi:10.1002/wrna.15.

Li Y, Zhong C, Wang J, Chen F, Shen W, Li B, Zheng N, Lu Y, Katanaev VL, Jia L. 2021. NOL7 facilitates melanoma progression and metastasis. *Signal Transduct Target Ther* **6**: 352. doi:10.1038/s41392-021-00676-3.

Li YZ, Zhang C, Pei JP, Zhang WC, Zhang CD, Dai DQ. 2022. The functional role of Pescadillo ribosomal biogenesis factor 1 in cancer. *J Cancer* **13**: 268-277. doi:10.7150/jca.58982.

Lindsley RC, Saber W, Mar BG, Redd R, Wang T, Haagenson MD, Grauman PV, Hu ZH, Spellman SR, Lee SJ et al. 2017. Prognostic Mutations in Myelodysplastic Syndrome after Stem-Cell Transplantation. *N Engl J Med* **376**: 536-547. doi:10.1056/NEJMoa1611604.

Lindstrom MS, Bartek J, Maya-Mendoza A. 2022. p53 at the crossroad of DNA replication and ribosome biogenesis stress pathways. *Cell Death Differ* **29**: 972-982. doi:10.1038/s41418-022-00999-w.

Lindstrom MS, Jurada D, Bursac S, Orsolic I, Bartek J, Volarevic S. 2018. Nucleolus as an emerging hub in maintenance of genome stability and cancer pathogenesis. *Oncogene* **37**: 2351-2366. doi:10.1038/s41388-017-0121-z.

Lipton JM, Molmenti CLS, Desai P, Lipton A, Ellis SR, Vlachos A. 2021. Early Onset Colorectal Cancer: An Emerging Cancer Risk in Patients with Diamond Blackfan Anemia. *Genes (Basel)* **13**. doi:10.3390/genes13010056.

Lu G, Zhang J, Li Y, Li Z, Zhang N, Xu X, Wang T, Guan Z, Gao GF, Yan J. 2011. hNUDT16: a universal decapping enzyme for small nucleolar RNA and cytoplasmic mRNA. *Protein Cell* **2**: 64-73. doi:10.1007/s13238-011-1009-2.

Mankame TP, Lingen MW. 2012. The RB tumor suppressor positively regulates transcription of the anti-angiogenic protein NOL7. *Neoplasia* **14**: 1213-1222. doi:10.1593/neo.121422.

Mansour FH, Pestov DG. 2013. Separation of long RNA by agarose-formaldehyde gel electrophoresis. *Anal Biochem* **441**: 18-20. doi:10.1016/j.ab.2013.06.008.

Marino D, Perkovic M, Hain A, Jaguva Vasudevan AA, Hofmann H, Hanschmann KM, Muhlebach MD, Schumann GG, Konig R, Cichutek K et al. 2016. APOBEC4 Enhances the Replication of HIV-1. *PLoS One* **11**: e0155422. doi:10.1371/journal.pone.0155422.

Matsson H, Davey EJ, Frojmark AS, Miyake K, Utsugisawa T, Flygare J, Zahou E, Byman I, Landin B, Ronquist G et al. 2006. Erythropoiesis in the Rps19 disrupted mouse: Analysis of erythropoietin response and biochemical markers for Diamond-Blackfan anemia. *Blood Cells Mol Dis* **36**: 259-264. doi:10.1016/j.bcmd.2005.12.002.

Matsuda M, Mayer BJ, Hanafusa H. 1991. Identification of domains of the v-crk oncogene product sufficient for association with phosphotyrosine-containing proteins. *Mol Cell Biol* **11**: 1607-1613. doi:10.1128/mcb.11.3.1607.

Mayer BJ, Hamaguchi M, Hanafusa H. 1988. A novel viral oncogene with structural similarity to phospholipase C. *Nature* **332**: 272-275. doi:10.1038/332272a0.

Mayer C, Grummt I. 2006. Ribosome biogenesis and cell growth: mTOR coordinates transcription by all three classes of nuclear RNA polymerases. *Oncogene* **25**: 6384-6391. doi:10.1038/sj.onc.1209883.

Mayer C, Schmitz KM, Li J, Grummt I, Santoro R. 2006. Intergenic transcripts regulate the epigenetic state of rRNA genes. *Mol Cell* **22**: 351-361. doi:10.1016/j.molcel.2006.03.028.

Mayes AE, Verdone L, Legrain P, Beggs JD. 1999. Characterization of Sm-like proteins in yeast and their association with U6 snRNA. *EMBO J* **18**: 4321-4331. doi:10.1093/emboj/18.15.4321.

McCann KL, Teramoto T, Zhang J, Tanaka Hall TM, Baserga SJ. 2016. The molecular basis for ANE syndrome revealed by the large ribosomal subunit processome interactome. *Elife* **5**. doi:10.7554/eLife.16381.

McCool MA, Bryant CJ, Baserga SJ. 2020. MicroRNAs and long non-coding RNAs as novel regulators of ribosome biogenesis. *Biochem Soc Trans* **48**: 595-612. doi:10.1042/BST20190854.

McCool MA, Bryant CJ, Huang H, Ogawa LM, Farley-Barnes KI, Sondalle SB, Abriola L, Surovtseva YV, Baserga SJ. 2022a. Human Nucleolar Protein 7 (NOL7) is required for pre-rRNA transcription and pre-18S rRNA processing. *bioRxiv (Preprint)*.

McCool MA, Buhagiar AF, Bryant CJ, Ogawa LM, Abriola L, Surovtseva YV, Baserga SJ. 2022b. Human pre-60S assembly factors link rRNA transcription to pre-rRNA processing. *RNA*. doi:10.1261/rna.079149.122.

McFadden EJ, Baserga SJ. 2022. U8 variants on the brain: a small nucleolar RNA and human disease. *RNA Biol* **19**: 412-418. doi:10.1080/15476286.2022.2048563.

McFarland JM, Ho ZV, Kugener G, Dempster JM, Montgomery PG, Bryan JG, Krill-Burger JM, Green TM, Vazquez F, Boehm JS et al. 2018. Improved estimation of cancer dependencies from large-scale RNAi screens using model-based normalization and data integration. *Nat Commun* **9**: 4610. doi:10.1038/s41467-018-06916-5.

McQuin C, Goodman A, Chernyshev V, Kamentsky L, Cimini BA, Karhohs KW, Doan M, Ding L, Rafelski SM, Thirstrup D et al. 2018. CellProfiler 3.0: Next-generation image processing for biology. *PLoS Biol* **16**: e2005970. doi:10.1371/journal.pbio.2005970.

McStay B, Grummt I. 2008. The epigenetics of rRNA genes: from molecular to chromosome biology. *Annu Rev Cell Dev Biol* **24**: 131-157. doi:10.1146/annurev.cellbio.24.110707.175259.

Mertz TM, Collins CD, Dennis M, Coxon M, Roberts SA. 2022. APOBEC-Induced Mutagenesis in Cancer. *Annu Rev Genet* **56**: 229-252. doi:10.1146/annurev-genet-072920-035840.

Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, Thomas PD. 2017. PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res* **45**: D183-D189. doi:10.1093/nar/gkw1138.

Mi L, Qi Q, Ran H, Chen L, Li D, Xiao D, Wu J, Cai Y, Zhang S, Li Y et al. 2021. Suppression of Ribosome Biogenesis by Targeting WD Repeat Domain 12 (WDR12) Inhibits Glioma Stem-Like Cell Growth. *Front Oncol* **11**: 751792. doi:10.3389/fonc.2021.751792.

Middlebrooks CD, Banday AR, Matsuda K, Udquim KI, Onabajo OO, Paquin A, Figueroa JD, Zhu B, Koutros S, Kubo M et al. 2016. Association of germline variants in the APOBEC3 region with cancer risk and enrichment with APOBEC-signature mutations in tumors. *Nat Genet* **48**: 1330-1338. doi:10.1038/ng.3670.

Miller G, Panov KI, Friedrich JK, Trinkle-Mulcahy L, Lamond AI, Zomerdijk JC. 2001. hRRN3 is essential in the SL1-mediated recruitment of RNA Polymerase I to rRNA gene promoters. *EMBO J* **20**: 1373-1382. doi:10.1093/emboj/20.6.1373.

Mills EW, Green R. 2017. Ribosomopathies: There's strength in numbers. *Science* **358**. doi:10.1126/science.aan2755.

Moorefield B, Greene EA, Reeder RH. 2000. RNA polymerase I transcription factor Rrn3 is functionally conserved between yeast and human. *Proc Natl Acad Sci U S A* **97**: 4724-4729. doi:10.1073/pnas.080063997.

Morcos PA, Li Y, Jiang S. 2008. Vivo-Morpholinos: a non-peptide transporter delivers Morpholinos into a wide array of mouse tissues. *Biotechniques* **45**: 613-614, 616, 618 passim. doi:10.2144/000113005.

Muller JS, Burns DT, Griffin H, Wells GR, Zendah RA, Munro B, Schneider C, Horvath R. 2020. RNA exosome mutations in pontocerebellar hypoplasia alter ribosome biogenesis and p53 levels. *Life Sci Alliance* **3**. doi:10.26508/lsa.202000678.

Mullineux ST, Lafontaine DL. 2012. Mapping the cleavage sites on mammalian pre-rRNAs: where do we stand? *Biochimie* **94**: 1521-1532. doi:10.1016/j.biochi.2012.02.001.

Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. 2000. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102**: 553-563. doi:10.1016/s0092-8674(00)00078-7.

Murayama A, Ohmori K, Fujimura A, Minami H, Yasuzawa-Tanaka K, Kuroda T, Oie S, Daitoku H, Okuwaki M, Nagata K et al. 2008. Epigenetic control of rDNA loci in response to intracellular energy status. *Cell* **133**: 627-639. doi:10.1016/j.cell.2008.03.030.

Nabet B, Roberts JM, Buckley DL, Paulk J, Dastjerdi S, Yang A, Leggett AL, Erb MA, Lawlor MA, Souza A et al. 2018. The dTAG system for immediate and target-specific protein degradation. *Nat Chem Biol* **14**: 431-441. doi:10.1038/s41589-018-0021-8.

Nagamani SC, Zhang F, Shchelochkov OA, Bi W, Ou Z, Scaglia F, Probst FJ, Shinawi M, Eng C, Hunter JV et al. 2009. Microdeletions including YWHAE in the Miller-Dieker syndrome region on chromosome 17p13.3 result in facial dysmorphisms, growth restriction, and cognitive impairment. *J Med Genet* **46**: 825-833. doi:10.1136/jmg.2009.067637.

Needleman SB, Wunsch CD. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol* **48**: 443-453. doi:10.1016/0022-2836(70)90057-4.

Nerurkar P, Altvater M, Gerhardy S, Schutz S, Fischer U, Weirich C, Panse VG. 2015. Eukaryotic Ribosome Assembly and Nuclear Export. *Int Rev Cell Mol Biol* **319**: 107-140. doi:10.1016/bs.ircmb.2015.07.002.

Niavarani A, Currie E, Reyal Y, Anjos-Afonso F, Horswell S, Griessinger E, Luis Sardina J, Bonnet D. 2015. APOBEC3A is implicated in a novel class of G-to-A mRNA editing in WT1 transcripts. *PLoS One* **10**: e0120089. doi:10.1371/journal.pone.0120089.

Nicolas E, Parisot P, Pinto-Monteiro C, de Walque R, De Vleeschouwer C, Lafontaine DL. 2016. Involvement of human ribosomal proteins in nucleolar structure and p53-dependent nucleolar stress. *Nat Commun* **7**: 11390. doi:10.1038/ncomms11390.

Norris K, Hopes T, Aspden JL. 2021. Ribosome heterogeneity and specialization in development. *Wiley Interdiscip Rev RNA* **12**: e1644. doi:10.1002/wrna.1644.

Nousbeck J, Spiegel R, Ishida-Yamamoto A, Indelman M, Shani-Adir A, Adir N, Lipkin E, Bercovici S, Geiger D, van Steensel MA et al. 2008. Alopecia, neurological defects, and endocrinopathy syndrome caused by decreased expression of RBM28, a nucleolar protein associated with ribosome biogenesis. *Am J Hum Genet* **82**: 1114-1121. doi:10.1016/j.ajhg.2008.03.014.

Nurk S, Koren S, Rhie A, Rautiainen M, Bzikadze AV, Mikheenko A, Vollger MR, Altemose N, Uralsky L, Gershman A et al. 2022. The complete sequence of a human genome. *Science* **376**: 44-53. doi:10.1126/science.abj6987.

O'Donohue MF, Choesmel V, Faubladier M, Fichant G, Gleizes PE. 2010. Functional dichotomy of ribosomal proteins during the synthesis of mammalian 40S ribosomal subunits. *J Cell Biol* **190**: 853-866. doi:10.1083/jcb.201005117.

Ogawa LM, Baserga SJ. 2017. Crosstalk between the nucleolus and the DNA damage response. *Mol Biosyst* **13**: 443-455. doi:10.1039/c6mb00740f.

Ogawa LM, Buhagiar AF, Abriola L, Leland BA, Surovtseva YV, Baserga SJ. 2021. Increased numbers of nucleoli in a genome-wide RNAi screen reveal proteins that link the cell cycle to RNA polymerase I transcription. *Mol Biol Cell* **32**: 956-973. doi:10.1091/mbc.E20-10-0670.

Olson ME, Harris RS, Harki DA. 2018. APOBEC Enzymes as Targets for Virus and Cancer Therapy. *Cell Chem Biol* **25**: 36-49. doi:10.1016/j.chembiol.2017.10.007.

Orsolic I, Jurada D, Pullen N, Oren M, Eliopoulos AG, Volarevic S. 2016. The relationship between the nucleolus and cancer: Current evidence and emerging paradigms. *Semin Cancer Biol* **37-38**: 36-50. doi:10.1016/j.semcancer.2015.12.004.

Ostergaard JR, Graakjaer J, Brandt C, Birkebaek NH. 2012. Further delineation of 17p13.3 microdeletion involving CRK. The effect of growth hormone treatment. *Eur J Med Genet* **55**: 22-26. doi:10.1016/j.ejmg.2011.09.004.

Ou L, Duan D, Wu J, Nice E, Huang C. 2012. The application of high throughput siRNA screening technology to study host-pathogen interactions. *Comb Chem High Throughput Screen* **15**: 299-305. doi:10.2174/138620712799361834.

Oughtred R, Rust J, Chang C, Breitkreutz BJ, Stark C, Willems A, Boucher L, Leung G, Kolas N, Zhang F et al. 2021. The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci* **30**: 187-200. doi:10.1002/pro.3978.

Panov KI, Friedrich JK, Russell J, Zomerdijk JC. 2006. UBF activates RNA polymerase I transcription by stimulating promoter escape. *EMBO J* **25**: 3310-3322. doi:10.1038/sj.emboj.7601221.

Panov KI, Hannan K, Hannan RD, Hein N. 2021. The Ribosomal Gene Loci-The Power behind the Throne. *Genes (Basel)* **12**. doi:10.3390/genes12050763.

Park TJ, Boyd K, Curran T. 2006. Cardiovascular and craniofacial defects in Crk-null mice. *Mol Cell Biol* **26**: 6272-6282. doi:10.1128/MCB.00472-06.

Parks MM, Kurylo CM, Dass RA, Bojmar L, Lyden D, Vincent CT, Blanchard SC. 2018. Variant ribosomal RNA alleles are conserved and exhibit tissue-specific expression. *Sci Adv* **4**: eaao0665. doi:10.1126/sciadv.aao0665.

Pawson T. 1995. Protein modules and signalling networks. *Nature* **373**: 573-580. doi:10.1038/373573a0.

Peculis B. 1997. RNA processing: pocket guides to ribosomal RNA. *Curr Biol* **7**: R480-482. doi:10.1016/s0960-9822(06)00242-9.

Peculis BA, Steitz JA. 1993. Disruption of U8 nucleolar snRNA inhibits 5.8S and 28S rRNA processing in the Xenopus oocyte. *Cell* **73**: 1233-1245. doi:10.1016/0092-8674(93)90651-6.

Pelletier J, Thomas G, Volarevic S. 2018. Ribosome biogenesis in cancer: new players and therapeutic avenues. *Nat Rev Cancer* **18**: 51-63. doi:10.1038/nrc.2017.104.

Peltonen K, Colis L, Liu H, Jaamaa S, Zhang Z, Af Hallstrom T, Moore HM, Sirajuddin P, Laiho M. 2014a. Small molecule BMH-compounds that inhibit RNA polymerase I and cause nucleolar stress. *Mol Cancer Ther* **13**: 2537-2546. doi:10.1158/1535-7163.MCT-14-0256.

Peltonen K, Colis L, Liu H, Trivedi R, Moubarek MS, Moore HM, Bai B, Rudek MA, Bieberich CJ, Laiho M. 2014b. A targeting modality for destruction of RNA polymerase I that possesses anticancer activity. *Cancer Cell* **25**: 77-90. doi:10.1016/j.ccr.2013.12.009.

Penzo M, Montanaro L, Trere D, Derenzini M. 2019. The Ribosome Biogenesis-Cancer Connection. *Cells* **8**. doi:10.3390/cells8010055.

Perez-Fernandez J, Roman A, De Las Rivas J, Bustelo XR, Dosil M. 2007. The 90S preribosome is a multimodular structure that is assembled through a hierarchical mechanism. *Mol Cell Biol* **27**: 5414-5429. doi:10.1128/MCB.00380-07.

Petljak M, Alexandrov LB, Brammeld JS, Price S, Wedge DC, Grossmann S, Dawson KJ, Ju YS, Iorio F, Tubio JMC et al. 2019. Characterizing Mutational Signatures in Human Cancer Cell Lines Reveals Episodic APOBEC Mutagenesis. *Cell* **176**: 1282-1294 e1220. doi:10.1016/j.cell.2019.02.012.

Petljak M, Dananberg A, Chu K, Bergstrom EN, Striepen J, von Morgen P, Chen Y, Shah H, Sale JE, Alexandrov LB et al. 2022. Mechanisms of APOBEC3 mutagenesis in human cancer cells. *Nature* **607**: 799-807. doi:10.1038/s41586-022-04972-y.

Phipps KR, Charette J, Baserga SJ. 2011. The small subunit processome in ribosome biogenesis-progress and prospects. *Wiley Interdiscip Rev RNA* **2**: 1-21. doi:10.1002/wrna.57.

Pianese G. 1896. Beitrag zur Histologie und Aetiologie der Carcinoma. Histologische und experimentelle Untersuchungen. *Beitr Pathol Anat Allgem Pathol* **142**: 1-193.

Pitts S, Laiho M. 2022. Regulation of RNA Polymerase I Stability and Function. *Cancers (Basel)* **14**. doi:10.3390/cancers14235776.

Ploton D, Menager M, Jeannesson P, Himber G, Pigeon F, Adnet JJ. 1986. Improvement in the staining and in the visualization of the argyrophilic proteins of the nucleolar organizer region at the optical level. *Histochem J* **18**: 5-14. doi:10.1007/BF01676192.

Poll G, Braun T, Jakovljevic J, Neueder A, Jakob S, Woolford JL, Jr., Tschochner H, Milkereit P. 2009. rRNA maturation in yeast cells depleted of large ribosomal subunit proteins. *PLoS One* **4**: e8249. doi:10.1371/journal.pone.0008249.

Popov A, Smirnov E, Kovacik L, Raska O, Hagen G, Stixova L, Raska I. 2013. Duration of the first steps of the human rRNA processing. *Nucleus* **4**: 134-141. doi:10.4161/nucl.23985.

Prieto JL, McStay B. 2007. Recruitment of factors linking transcription and processing of pre-rRNA to NOR chromatin is UBF-dependent and occurs independent of transcription in human cells. *Genes Dev* **21**: 2041-2054. doi:10.1101/gad.436707.

Raman N, Weir E, Muller S. 2016. The AAA ATPase MDN1 Acts as a SUMO-Targeted Regulator in Mammalian Pre-ribosome Remodeling. *Mol Cell* **64**: 607-615. doi:10.1016/j.molcel.2016.09.039.

Reilly CR, Shimamura A. 2022. Predisposition to myeloid malignancies in Shwachman-Diamond syndrome: Biological insights and clinical advances. *Blood*. doi:10.1182/blood.2022017739.

Reimer G, Pollard KM, Penning CA, Ochs RL, Lischwe MA, Busch H, Tan EM. 1987. Monoclonal autoantibody from a (New Zealand black x New Zealand white)F1 mouse and some human scleroderma sera target an Mr 34,000 nucleolar protein of the U3 RNP particle. *Arthritis Rheum* **30**: 793-800. doi:10.1002/art.1780300709.

Revathidevi S, Murugan AK, Nakaoka H, Inoue I, Munirajan AK. 2021. APOBEC: A molecular driver in cervical cancer pathogenesis. *Cancer Lett* **496**: 104-116. doi:10.1016/j.canlet.2020.10.004.

Risler JK, Kenny AE, Palumbo RJ, Gamache ER, Curcio MJ. 2012. Host co-factors of the retrovirus-like transposon Ty1. *Mob DNA* **3**: 12. doi:10.1186/1759-8753-3-12.

Roberts SA, Lawrence MS, Klimczak LJ, Grimm SA, Fargo D, Stojanov P, Kiezun A, Kryukov GV, Carter SL, Saksena G et al. 2013. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nat Genet* **45**: 970-976. doi:10.1038/ng.2702.

Rohrmoser M, Holzel M, Grimm T, Malamoussi A, Harasim T, Orban M, Pfisterer I, Gruber-Eber A, Kremmer E, Eick D. 2007. Interdependence of Pes1, Bop1, and WDR12 controls nucleolar localization and assembly of the PeBoW complex required for maturation of the 60S ribosomal subunit. *Mol Cell Biol* **27**: 3682-3694. doi:10.1128/MCB.00172-07.

Rosenberg BR, Hamilton CE, Mwangi MM, Dewell S, Papavasiliou FN. 2011. Transcriptome-wide sequencing reveals numerous APOBEC1 mRNA-editing targets in transcript 3' UTRs. *Nat Struct Mol Biol* **18**: 230-236. doi:10.1038/nsmb.1975.

Ross AP, Zarbalis KS. 2014. The emerging roles of ribosome biogenesis in craniofacial development. *Front Physiol* **5**: 26. doi:10.3389/fphys.2014.00026.

Rossetti S, Wierzbicki AJ, Sacchi N. 2018. Undermining ribosomal RNA transcription in both the nucleolus and mitochondrion: an offbeat approach to target MYC-driven cancer. *Oncotarget* **9**: 5016-5031. doi:10.18632/oncotarget.23579.

Rothschild D, Susanto TT, Spence JP, Genuth NR, N. S-A, Pritchard JK, Barna M. 2023. A comprehensive rRNA variation atlas in health and disease. *BioRxiv (Preprint)*.

Rubbi CP, Milner J. 2003. Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses. *EMBO J* **22**: 6068-6077. doi:10.1093/emboj/cdg579.

Russell J, Zomerdijk JC. 2006. The RNA polymerase I transcription machinery. *Biochem Soc Symp*: 203-216. doi:10.1042/bss0730203.

Russo A, Esposito D, Catillo M, Pietropaolo C, Crescenzi E, Russo G. 2013. Human rpL3 induces G(1)/S arrest or apoptosis by modulating p21 (waf1/cip1) levels in a p53-independent manner. *Cell Cycle* **12**: 76-87. doi:10.4161/cc.22963.

Saveanu C, Namane A, Gleizes PE, Lebreton A, Rousselle JC, Noaillac-Depeyre J, Gas N, Jacquier A, Fromont-Racine M. 2003. Sequential protein association with nascent 60S ribosomal particles. *Mol Cell Biol* **23**: 4449-4460. doi:10.1128/MCB.23.13.4449-4460.2003.

Schlesinger S, Selig S, Bergman Y, Cedar H. 2009. Allelic inactivation of rDNA loci. *Genes Dev* **23**: 2437-2447. doi:10.1101/gad.544509.

Schmidt EK, Clavarino G, Ceppi M, Pierre P. 2009. SUnSET, a nonradioactive method to monitor protein synthesis. *Nat Methods* **6**: 275-277. doi:10.1038/nmeth.1314.

Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**: 671-675. doi:10.1038/nmeth.2089.

Scull CE, Zhang Y, Tower N, Rasmussen L, Padmalayam I, Hunter R, Zhai L, Bostwick R, Schneider DA. 2019. Discovery of novel inhibitors of ribosome biogenesis by innovative high throughput screening strategies. *Biochem J* **476**: 2209-2219. doi:10.1042/BCJ20190207.

Sekiguchi T, Hayano T, Yanagida M, Takahashi N, Nishimoto T. 2006. NOP132 is required for proper nucleolus localization of DEAD-box RNA helicase DDX47. *Nucleic Acids Res* **34**: 4593-4608. doi:10.1093/nar/gkl603.

Serrano JC, von Trentini D, Berrios KN, Barka A, Dmochowski IJ, Kohli RM. 2022. Structure-Guided Design of a Potent and Specific Inhibitor against the Genomic Mutator APOBEC3A. *ACS Chem Biol* **17**: 3379-3388. doi:10.1021/acschembio.2c00796.

Shah AN, Cadinu D, Henke RM, Xin X, Dastidar RG, Zhang L. 2011. Deletion of a subgroup of ribosome-related genes minimizes hypoxia-induced changes and confers hypoxia tolerance. *Physiol Genomics* **43**: 855-872. doi:10.1152/physiolgenomics.00232.2010.

Shao Z, Flynn RA, Crowe JL, Zhu Y, Liang J, Jiang W, Aryan F, Aoude P, Bertozzi CR, Estes VM et al. 2020. DNA-PKcs has KU-dependent function in rRNA processing and haematopoiesis. *Nature* **579**: 291-296. doi:10.1038/s41586-020-2041-2.

Sharma S, Patnaik SK, Kemer Z, Baysal BE. 2017. Transient overexpression of exogenous APOBEC3A causes C-to-U RNA editing of thousands of genes. *RNA Biol* **14**: 603-610. doi:10.1080/15476286.2016.1184387.

Sharma S, Patnaik SK, Taggart RT, Baysal BE. 2016. The double-domain cytidine deaminase APOBEC3G is a cellular site-specific RNA editing enzyme. *Sci Rep* **6**: 39100. doi:10.1038/srep39100.

Sharma S, Patnaik SK, Taggart RT, Kannisto ED, Enriquez SM, Gollnick P, Baysal BE. 2015. APOBEC3A cytidine deaminase induces RNA editing in monocytes and macrophages. *Nat Commun* **6**: 6881. doi:10.1038/ncomms7881.

Sharma S, Wang J, Alqassim E, Portwood S, Cortes Gomez E, Maguire O, Basse PH, Wang ES, Segal BH, Baysal BE. 2019. Mitochondrial hypoxic stress induces widespread RNA editing by APOBEC3G in natural killer cells. *Genome Biol* **20**: 37. doi:10.1186/s13059-019-1651-1.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J et al. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**: 539. doi:10.1038/msb.2011.75.

Sigoillot FD, Lyman S, Huckins JF, Adamson B, Chung E, Quattrochi B, King RW. 2012. A bioinformatics method identifies prominent off-targeted transcripts in RNAi screens. *Nat Methods* **9**: 363-366. doi:10.1038/nmeth.1898.

Singh S, Vanden Broeck A, Miller L, Chaker-Margot M, Klinge S. 2021. Nucleolar maturation of the human small subunit processome. *Science* **373**: eabj5338. doi:10.1126/science.abj5338.

Siriwardena SU, Chen K, Bhagwat AS. 2016. Functions and Malfunctions of Mammalian DNA-Cytosine Deaminases. *Chem Rev* **116**: 12688-12710. doi:10.1021/acs.chemrev.6b00296.

Sloan KE, Bohnsack MT, Schneider C, Watkins NJ. 2014. The roles of SSU processome components and surveillance factors in the initial processing of human ribosomal RNA. *RNA* **20**: 540-550. doi:10.1261/rna.043471.113.

Sloan KE, Warda AS, Sharma S, Entian KD, Lafontaine DLJ, Bohnsack MT. 2017. Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA Biol* **14**: 1138-1152. doi:10.1080/15476286.2016.1259781.

Smith HC, Bennett RP, Kizilyer A, McDougall WM, Prohaska KM. 2012. Functions and regulation of the APOBEC family of proteins. *Semin Cell Dev Biol* **23**: 258-268. doi:10.1016/j.semcdb.2011.10.004.

Sondalle SB, Longerich S, Ogawa LM, Sung P, Baserga SJ. 2019. Fanconi anemia protein FANCI functions in ribosome biogenesis. *Proc Natl Acad Sci U S A* **116**: 2561-2570. doi:10.1073/pnas.1811557116.

Sorino C, Catena V, Bruno T, De Nicola F, Scalera S, Bossi G, Fabretti F, Mano M, De Smaele E, Fanciulli M et al. 2020. Che-1/AATF binds to RNA polymerase I machinery and sustains ribosomal RNA gene transcription. *Nucleic Acids Res* **48**: 5891-5906. doi:10.1093/nar/gkaa344.

Soule HD, Maloney TM, Wolman SR, Peterson WD, Jr., Brenz R, McGrath CM, Russo J, Pauley RJ, Jones RF, Brooks SC. 1990. Isolation and characterization of a spontaneously immortalized human breast epithelial cell line, MCF-10. *Cancer Res* **50**: 6075-6086.

Sriram G, Birge RB. 2010. Emerging roles for crk in human cancer. *Genes Cancer* **1**: 1132-1139. doi:10.1177/1947601910397188.

Stamatopoulou V, Parisot P, De Vleeschouwer C, Lafontaine DLJ. 2018. Use of the iNo score to discriminate normal from altered nucleolar morphology, with applications in basic cell biology and potential in human disease diagnostics. *Nat Protoc* **13**: 2387-2406. doi:10.1038/s41596-018-0044-3.

Stolc V, Altman S. 1997. Rpp1, an essential protein subunit of nuclear RNase P required for processing of precursor tRNA and 35S precursor rRNA in Saccharomyces cerevisiae. *Genes Dev* **11**: 2414-2425. doi:10.1101/gad.11.18.2414.

Strezoska Z, Pestov DG, Lau LF. 2002. Functional inactivation of the mouse nucleolar protein Bop1 inhibits multiple steps in pre-rRNA processing and blocks cell cycle progression. *J Biol Chem* **277**: 29617-29625. doi:10.1074/jbc.M204381200.

Sun XX, Wang YG, Xirodimas DP, Dai MS. 2010. Perturbation of 60 S ribosomal biogenesis results in ribosomal protein L5- and L11-dependent p53 activation. *J Biol Chem* **285**: 25812-25821. doi:10.1074/jbc.M109.098442.

Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, Doncheva NT, Legeay M, Fang T, Bork P et al. 2021. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res* **49**: D605-D612. doi:10.1093/nar/gkaa1074.

Szymanski M, Zielezinski A, Barciszewski J, Erdmann VA, Karlowski WM. 2016. 5SRNAdb: an information resource for 5S ribosomal RNAs. *Nucleic Acids Res* **44**: D180-183. doi:10.1093/nar/gkv1081.

Tafforeau L, Zorbas C, Langhendries JL, Mullineux ST, Stamatopoulou V, Mullier R, Wacheul L, Lafontaine DL. 2013. The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. *Mol Cell* **51**: 539-551. doi:10.1016/j.molcel.2013.08.011.

Teng B, Burant CF, Davidson NO. 1993. Molecular cloning of an apolipoprotein B messenger RNA editing protein. *Science* **260**: 1816-1819. doi:10.1126/science.8511591.

Thorenoor N, Slaby O. 2015. Small nucleolar RNAs functioning and potential roles in cancer. *Tumour Biol* **36**: 41-53. doi:10.1007/s13277-014-2818-8.

Thul PJ, Akesson L, Wiking M, Mahdessian D, Geladaki A, Ait Blal H, Alm T, Asplund A, Bjork L, Breckels LM et al. 2017. A subcellular map of the human proteome. *Science* **356**. doi:10.1126/science.aal3321.

Tomasevic N, Peculis B. 1999. Identification of a U8 snoRNA-specific binding protein. *J Biol Chem* **274**: 35914-35920. doi:10.1074/jbc.274.50.35914.

Tomasevic N, Peculis BA. 2002. Xenopus LSm proteins bind U8 snoRNA via an internal evolutionarily conserved octamer sequence. *Mol Cell Biol* **22**: 4101-4112. doi:10.1128/MCB.22.12.4101-4112.2002.

Tsai YC, Greco TM, Boonmee A, Miteva Y, Cristea IM. 2012. Functional proteomics establishes the interaction of SIRT7 with chromatin remodeling complexes and expands its role in regulation of RNA polymerase I transcription. *Mol Cell Proteomics* **11**: 60-76. doi:10.1074/mcp.A111.015156.

Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kryukov G, Cowley GS, Gill S, Harrington WF, Pantel S, Krill-Burger JM et al. 2017. Defining a Cancer Dependency Map. *Cell* **170**: 564-576 e516. doi:10.1016/j.cell.2017.06.010.

Turi Z, Senkyrikova M, Mistrik M, Bartek J, Moudry P. 2018. Perturbation of RNA Polymerase I transcription machinery by ablation of HEATR1 triggers the RPL5/RPL11-MDM2-p53 ribosome biogenesis stress checkpoint pathway in human cells. *Cell Cycle* **17**: 92-101. doi:10.1080/15384101.2017.1403685.

Turowski TW, Tollervey D. 2015. Cotranscriptional events in eukaryotic ribosome synthesis. *Wiley Interdiscip Rev RNA* **6**: 129-139. doi:10.1002/wrna.1263.

Udquim KI, Zettelmeyer C, Banday AR, Lin SH, Prokunina-Olsson L. 2020. APOBEC3B expression in breast cancer cell lines and tumors depends on the estrogen receptor status. *Carcinogenesis* **41**: 1030-1037. doi:10.1093/carcin/bgaa002.

Upadhya R, Lee J, Willis IM. 2002. Maf1 is an essential mediator of diverse signals that repress RNA polymerase III transcription. *Mol Cell* **10**: 1489-1494. doi:10.1016/s1097-2765(02)00787-6.

Usher J, Balderas-Hernandez V, Quon P, Gold ND, Martin VJ, Mahadevan R, Baetz K. 2011. Chemical and Synthetic Genetic Array Analysis Identifies Genes that Suppress Xylose Utilization and Fermentation in Saccharomyces cerevisiae. *G3 (Bethesda)* **1**: 247-258. doi:10.1534/g3.111.000695.

Van Haute L, Hendrick AG, D'Souza AR, Powell CA, Rebelo-Guiomar P, Harbour ME, Ding S, Fearnley IM, Andrews B, Minczuk M. 2019. METTL15 introduces N4-methylcytidine into human mitochondrial 12S rRNA and is required for mitoribosome biogenesis. *Nucleic Acids Res* **47**: 10267-10281. doi:10.1093/nar/gkz735.

Vellky JE, Ricke EA, Huang W, Ricke WA. 2021. Expression, Localization, and Function of the Nucleolar Protein BOP1 in Prostate Cancer Progression. *Am J Pathol* **191**: 168-179. doi:10.1016/j.ajpath.2020.09.010.

Visintin R, Amon A. 2000. The nucleolus: the magician's hat for cell cycle tricks. *Curr Opin Cell Biol* **12**: 372-377. doi:10.1016/s0955-0674(00)00102-2.

Wallace H, Birnstiel ML. 1966. Ribosomal cistrons and the nucleolar organizer. *Biochim Biophys Acta* **114**: 296-310. doi:10.1016/0005-2787(66)90311-x.

Wallisch M, Kunkel E, Hoehn K, Grummt F. 2002. Ku antigen supports termination of mammalian rDNA replication by transcription termination factor TTF-I. *Biol Chem* **383**: 765-771. doi:10.1515/BC.2002.080.

Wang M, Anikin L, Pestov DG. 2014. Two orthogonal cleavages separate subunit RNAs in mouse ribosome biogenesis. *Nucleic Acids Res* **42**: 11180-11191. doi:10.1093/nar/gku787.

Wang Y, Liu J, Zhao H, Lu W, Zhao J, Yang L, Li N, Du X, Ke Y. 2007. Human 1A6/DRIM, the homolog of yeast Utp20, functions in the 18S rRNA processing. *Biochim Biophys Acta* **1773**: 863-868. doi:10.1016/j.bbamcr.2007.04.002.

Warner JR, Vilardell J, Sohn JH. 2001. Economics of ribosome biosynthesis. *Cold Spring Harb Symp Quant Biol* **66**: 567-574. doi:10.1101/sqb.2001.66.567.

Warren AJ. 2018. Molecular basis of the human ribosomopathy Shwachman-Diamond syndrome. *Adv Biol Regul* **67**: 109-127. doi:10.1016/j.jbior.2017.09.002.

Warren CJ, Westrich JA, Doorslaer KV, Pyeon D. 2017. Roles of APOBEC3A and APOBEC3B in Human Papillomavirus Infection and Disease Progression. *Viruses* **9**. doi:10.3390/v9080233.

Watkins NJ, Bohnsack MT. 2012. The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wires Rna* **3**: 397-414. doi:10.1002/wrna.117.

Weiss MB, Vitolo MI, Mohseni M, Rosen DM, Denmeade SR, Park BH, Weber DJ, Bachman KE. 2010. Deletion of p53 in human mammary epithelial cells causes chromosomal instability and altered therapeutic response. *Oncogene* **29**: 4715-4724. doi:10.1038/onc.2010.220.

Werner A, Iwasaki S, McGourty CA, Medina-Ruiz S, Teerikorpi N, Fedrigo I, Ingolia NT, Rape M. 2015. Cell-fate determination by ubiquitin-dependent regulation of translation. *Nature* **525**: 523-527. doi:10.1038/nature14978.

Westendorf JM, Konstantinov KN, Wormsley S, Shu MD, Matsumoto-Taniura N, Pirollet F, Klier FG, Gerace L, Baserga SJ. 1998. M phase phosphoprotein 10 is a human U3 small nucleolar ribonucleoprotein component. *Mol Biol Cell* **9**: 437-449. doi:10.1091/mbc.9.2.437.

Wild T, Horvath P, Wyler E, Widmann B, Badertscher L, Zemp I, Kozak K, Csucs G, Lund E, Kutay U. 2010. A protein inventory of human ribosome biogenesis reveals an essential function of exportin 5 in 60S subunit export. *PLoS Biol* **8**: e1000522. doi:10.1371/journal.pbio.1000522.

Willems L, Gillet NA. 2015. APOBEC3 Interference during Replication of Viral Genomes. *Viruses* **7**: 2999-3018. doi:10.3390/v7062757.

Wilm A, Aw PP, Bertrand D, Yeo GH, Ong SH, Wong CH, Khor CC, Petric R, Hibberd ML, Nagarajan N. 2012. LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Res* **40**: 11189-11201. doi:10.1093/nar/gks918.

Woolford JL, Jr., Baserga SJ. 2013. Ribosome biogenesis in the yeast Saccharomyces cerevisiae. *Genetics* **195**: 643-681. doi:10.1534/genetics.113.153197.

Woolnough JL, Atwood BL, Liu Z, Zhao R, Giles KE. 2016. The Regulation of rRNA Gene Transcription during Directed Differentiation of Human Embryonic Stem Cells. *PLoS One* **11**: e0157276. doi:10.1371/journal.pone.0157276.

Xue S, Barna M. 2012. Specialized ribosomes: a new frontier in gene regulation and organismal biology. *Nat Rev Mol Cell Biol* **13**: 355-369. doi:10.1038/nrm3359.

Yang K, Yang J, Yi J. 2018. Nucleolar Stress: hallmarks, sensing mechanism and diseases. *Cell Stress* **2**: 125-140. doi:10.15698/cst2018.06.139.

Yang X, Boehm JS, Yang X, Salehi-Ashtiani K, Hao T, Shen Y, Lubonja R, Thomas SR, Alkan O, Bhimdi T et al. 2011. A public genome-scale lentiviral expression library of human ORFs. *Nat Methods* **8**: 659-661. doi:10.1038/nmeth.1638.

Yang Y, Isaac C, Wang C, Dragon F, Pogacic V, Meier UT. 2000. Conserved composition of mammalian box H/ACA and box C/D small nucleolar ribonucleoprotein particles and their interaction with the common factor Nopp140. *Mol Biol Cell* **11**: 567-577. doi:10.1091/mbc.11.2.567.

Ye C, Liu B, Lu H, Liu J, Rabson AB, Jacinto E, Pestov DG, Shen Z. 2020. BCCIP is required for nucleolar recruitment of eIF6 and 12S pre-rRNA production during 60S ribosome biogenesis. *Nucleic Acids Res* **48**: 12817-12832. doi:10.1093/nar/gkaa1114.

Yin Y, Zhou L, Zhan R, Zhang Q, Li M. 2018. Identification of WDR12 as a novel oncogene involved in hepatocellular carcinoma propagation. *Cancer Manag Res* **10**: 3985-3993. doi:10.2147/CMAR.S176268.

Zhang F, Lou L, Peng B, Song X, Reizes O, Almasan A, Gong Z. 2020. Nudix Hydrolase NUDT16 Regulates 53BP1 Protein by Reversing 53BP1 ADP-Ribosylation. *Cancer Res* **80**: 999-1010. doi:10.1158/0008-5472.CAN-19-2205.

Zhang XD. 2011. Illustration of SSMD, z score, SSMD*, z* score, and t statistic for hit selection in RNAi high-throughput screens. *J Biomol Screen* **16**: 775-785. doi:10.1177/1087057111405851.

Zhang Y, Forys JT, Miceli AP, Gwinn AS, Weber JD. 2011. Identification of DHX33 as a mediator of rRNA synthesis and cell growth. *Mol Cell Biol* **31**: 4676-4691. doi:10.1128/MCB.05832-11.

Zhao Z, Senturk N, Song C, Grummt I. 2018. lncRNA PAPAS tethered to the rDNA enhancer recruits hypophosphorylated CHD4/NuRD to repress rRNA synthesis at elevated temperatures. *Genes Dev* **32**: 836-848. doi:10.1101/gad.311688.118.

Zhou G, Doci CL, Lingen MW. 2010. Identification and functional analysis of NOL7 nuclear and nucleolar localization signals. *BMC Cell Biol* **11**: 74. doi:10.1186/1471-2121-11-74.

Zhou L, Hang J, Zhou Y, Wan R, Lu G, Yin P, Yan C, Shi Y. 2014. Crystal structures of the Lsm complex bound to the 3' end sequence of U6 small nuclear RNA. *Nature* **506**: 116-120. doi:10.1038/nature12803.

Zhu X, Zhang H, Mendell JT. 2020. Ribosome Recycling by ABCE1 Links Lysosomal Function and Iron Homeostasis to 3' UTR-Directed Regulation and Nonsense-Mediated Decay. *Cell Rep* **32**: 107895. doi:10.1016/j.celrep.2020.107895.

Zisi A, Bartek J, Lindstrom MS. 2022. Targeting Ribosome Biogenesis in Cancer: Lessons Learned and Way Forward. *Cancers (Basel)* **14**. doi:10.3390/cancers14092126.

Zorbas C, Nicolas E, Wacheul L, Huvelle E, Heurgue-Hamard V, Lafontaine DL. 2015. The human 18S rRNA base methyltransferases DIMT1L and WBSCR22-TRMT112 but not rRNA modification are required for ribosome biogenesis. *Mol Biol Cell* **26**: 2080-2095. doi:10.1091/mbc.E15-02-0073.

Appendix I: 5-EU nucleolar rRNA biogenesis screen results

HGNC Symbol	Original Phenotype	Mean (1 nucleolus % effect)	Mean (5+ nucleoli % effect)	Mean (% Viability)	Mean (nucleolar rRNA biogenesis % inhibition)
COPB1	1nuc	191.49	-14.39	13.84	128.02
SUOX	1nuc	3.57	3.71	21.82	125.87
FSCN1	1nuc	-28.94	8.12	18.74	124.53
SLC39A7	1nuc	-23.40	1.42	53.65	120.15
GHRL	1nuc	30.62	-6.78	20.64	119.81
MYC	1nuc	2.99	1.39	35.05	119.20
MFSD3	1nuc	-74.00	30.54	43.87	117.66
GRB2	1nuc	177.79	-11.51	12.07	117.04
EIF5A2	1nuc	149.11	-10.63	13.19	116.94
HEATR1	1nuc	43.36	4.59	31.28	116.44
FBXO33	1nuc	-27.50	4.07	15.54	115.40
CUL2	1nuc	-105.13	28.12	24.64	114.78
RRM1	5nuc	54.97	21.06	8.57	114.67
UBC	1nuc	1.26	29.63	9.22	114.62
TNRC6C	1nuc	-37.21	26.33	23.53	113.97
LTBP2	1nuc	-25.64	8.63	20.69	113.67
FGF22	1nuc	42.21	-4.41	24.70	113.42
SRP54	1nuc	-45.24	1.39	35.50	113.02
FAM110C	1nuc	169.30	-8.91	31.65	112.68
DNTTIP2	1nuc	219.58	-18.38	29.02	112.56
STARD13	1nuc	-8.87	-0.45	25.92	112.47
TRIM25	1nuc	15.67	2.04	25.93	112.30
TMEM14B	1nuc	-17.28	16.61	29.18	111.85
IQSEC3	1nuc	-27.69	-3.64	41.66	111.59
DDX19B	1nuc	51.02	2.61	2.07	111.08
CEP192	1nuc	24.49	9.53	20.10	110.88
RPL3	1nuc	274.48	-14.45	14.39	110.59
WNT4	1nuc	12.65	-1.55	28.52	110.25
KCNS2	1nuc	-31.38	12.38	30.90	109.97
FBXO44	1nuc	183.76	-12.74	18.04	109.27
HES4	1nuc	-11.56	-0.34	26.76	108.85
SPRR3	1nuc	97.64	-4.81	23.66	108.74
FAM83A	1nuc	-78.85	25.08	41.98	108.00

CILP2	1nuc	-62.47	14.13	39.85	107.49
SUPT6H	1nuc	57.89	-6.71	16.93	107.13
GLIPR2	1nuc	-23.08	13.45	57.34	106.40
PSMC6	1nuc	-26.73	20.80	31.37	106.25
RPS28	1nuc	218.33	-15.18	16.14	106.23
SUPT5H	1nuc	100.45	-6.48	23.59	105.92
SERTAD2	1nuc	-48.69	7.83	53.17	105.26
HDHD5	1nuc	-58.22	18.71	23.15	105.23
LIG3	5nuc	-9.44	2.09	64.86	105.21
POLR2C	1nuc	-17.72	7.44	32.85	105.21
RPL19	1nuc	-7.95	25.58	13.30	105.20
SFTPD	1nuc	-30.47	10.31	28.76	105.10
GGT6	1nuc	-70.18	19.37	31.88	104.86
NUP205	1nuc	-5.62	11.11	13.03	103.73
ZNF76	1nuc	-22.89	4.99	55.31	103.17
EIF4A3	1nuc	244.59	-13.51	2.53	102.98
CEBPD	1nuc	-8.46	10.35	43.86	102.77
DEAF1	1nuc	-49.11	0.50	38.95	102.59
POLR2B	1nuc	-35.70	18.34	29.57	102.37
RYBP	1nuc	43.54	8.34	29.76	102.32
IGFBP5	1nuc	-2.67	-0.70	59.90	102.16
ACTB	1nuc	201.41	-13.22	11.78	101.33
POLR2E	1nuc	-9.48	0.64	38.52	100.42
NAPA	1nuc	158.60	-13.78	26.65	100.27
SYNC	1nuc	8.76	-1.40	32.01	100.27
ZNF513	1nuc	5.43	4.20	68.06	100.08
DDX56	1nuc	26.67	-1.76	37.67	99.60
NUF2	1nuc	66.73	13.27	20.99	99.18
DAB2	1nuc	-41.47	7.29	30.51	99.02
STAB2	1nuc	-74.53	20.73	41.32	98.56
RPS4X	1nuc	258.12	-14.66	12.82	98.49
RPS6	1nuc	315.81	-18.66	8.80	98.32
DYNC2LI1	1nuc	-64.26	12.04	49.09	97.22
DNAJC1	1nuc	-60.09	28.36	36.40	96.81
INCENP	5nuc	-85.61	114.86	16.63	96.80
RPS8	1nuc	197.64	-13.91	13.34	96.80
C19orf33	1nuc	-41.01	8.15	33.87	96.34
INHBE	1nuc	-55.11	4.54	54.57	96.06
RPS3A	1nuc	147.03	-11.72	15.14	95.89
UBB	1nuc	42.93	7.24	8.82	95.78
ITFG2	1nuc	-28.75	10.17	74.80	95.49

ZNF219	5nuc	202.93	-12.39	33.86	94.75
RPA1	1nuc	60.76	6.09	12.57	94.72
CDH16	1nuc	-98.26	19.76	43.13	94.61
RPS14	1nuc	202.62	-11.15	13.06	94.49
DBNDD1	5nuc	-4.62	2.28	66.79	94.42
GNPDA1	1nuc	161.27	-9.52	20.07	94.39
SORL1	1nuc	-47.20	16.00	37.10	94.02
APBB1	1nuc	-3.28	1.22	24.55	93.88
NOL11	1nuc	9.24	9.36	44.65	93.78
KAT5	1nuc	-65.61	25.54	46.03	93.77
EEPD1		-89.76	8.93	40.87	93.33
ENTPD5	1nuc	-32.44	-3.19	39.52	93.08
CAPZB	1nuc	62.41	-6.31	26.46	92.60
FAAP100	1nuc	-35.61	15.50	21.80	92.45
LSM12	1nuc	-36.70	3.44	41.19	92.41
FAM120C	1nuc	-86.01	22.30	40.07	92.40
KCTD11	1nuc	-39.59	24.97	27.65	92.39
RPS24	1nuc	137.34	-8.79	13.86	91.96
WBP11	5nuc	-25.05	15.80	23.70	91.90
SDC1	1nuc	-73.63	29.38	54.35	91.74
TBC1D2	1nuc	-43.85	8.04	49.23	91.72
RPS5	1nuc	332.17	-17.67	14.94	91.22
NXF1	1nuc	-42.61	5.98	2.95	91.13
RPS7	1nuc	129.99	-9.40	13.72	91.08
VHL	1nuc	-38.94	6.30	40.89	91.05
MDN1	5nuc	21.17	2.46	28.62	90.81
DNAJC5	1nuc	-10.76	-1.73	60.65	90.75
SH2B1	1nuc	-4.56	6.44	33.26	90.03
RPS16	1nuc	165.41	-17.01	21.74	89.96
SYNE3	1nuc	-20.14	19.41	39.49	89.71
ESF1	1nuc	-4.32	-0.69	53.18	89.57
OR10A2	1nuc	-32.90	7.72	26.94	89.55
CFAP410	1nuc	-42.70	-3.23	45.02	89.44
CD1D	1nuc	-62.69	3.46	32.00	89.38
BOP1		-35.92	24.73	38.38	89.22
FAM217A	1nuc	-65.55	14.92	48.71	89.19
FSIP2	1nuc	-63.33	15.14	34.15	88.98
GRHPR	1nuc	101.96	-9.32	44.68	88.59
RPL18	1nuc	151.87	-13.09	11.91	88.51
MGAT4B	1nuc	-5.34	0.87	71.05	88.32
ECHDC2	5nuc	-23.85	17.14	74.35	88.27

FLVCR2	1nuc	-36.61	2.80	36.53	88.08
TMEM268	1nuc	-63.93	31.32	61.21	88.03
C12orf29	1nuc	6.05	9.40	25.55	87.82
FOXL1	1nuc	-17.01	-2.20	44.17	87.77
TOGARAM2	1nuc	-44.36	8.87	39.54	87.48
CEP250	1nuc	64.83	-5.40	30.21	87.48
COX6B1	1nuc	-44.33	13.34	64.66	87.26
SNX22	1nuc	-49.61	15.39	53.02	87.22
SLC35A5	1nuc	-36.49	13.73	29.85	86.97
RPL26	1nuc	84.28	1.09	13.38	86.94
RPL17	1nuc	102.29	-0.62	12.73	86.78
CDCA5	5nuc	-10.11	46.92	33.50	86.40
TBRG1	1nuc	-54.98	23.80	42.03	86.37
HMGN4	1nuc	-74.31	22.48	52.71	86.27
DBNL	1nuc	-66.93	-0.51	44.61	86.25
HEATR5B	1nuc	-27.83	-2.38	58.30	86.19
XPOT		-31.80	6.99	36.18	86.11
AGR2	5nuc	-57.91	19.90	41.07	86.10
IRF5	1nuc	-31.85	9.82	40.56	85.79
SLC29A1	1nuc	-53.36	13.46	55.83	85.67
CIAO2B	5nuc	26.31	-4.41	38.50	85.63
MVD	1nuc	-21.18	3.55	52.03	85.55
ZNF580	1nuc	42.52	-3.84	36.43	85.48
MBD2	1nuc	-72.50	26.15	39.55	85.41
RSL24D1	1nuc	64.73	-4.14	23.76	85.41
SBNO1	1nuc	-35.03	17.99	47.86	85.29
FAM171A1	1nuc	-86.36	17.06	58.02	84.70
EIF3K	1nuc	60.24	-6.36	30.58	84.52
RRN3	1nuc	51.11	7.23	30.34	84.51
TMBIM4	1nuc	-17.71	8.93	25.79	84.39
F8	1nuc	4.28	0.71	34.83	84.19
NANS	1nuc	10.00	-3.95	55.95	84.02
RBM47	1nuc	-33.86	24.65	61.95	83.97
UPK3B	1nuc	-19.34	8.82	44.13	83.89
RPS9	1nuc	109.35	-12.46	16.71	83.66
SLC26A7	5nuc	-56.37	15.61	62.68	83.64
ZER1	5nuc	-9.96	6.49	63.62	83.57
FAM193B	1nuc	-71.85	17.43	74.24	83.31
CXCL10	1nuc	-75.05	14.96	60.80	82.81
ATN1	1nuc	-23.20	13.98	54.66	82.75
BTN3A3	1nuc	11.09	-4.97	47.70	82.72

PAQR8	1nuc	-30.29	9.78	55.46	82.65
LRP4	1nuc	-26.36	15.27	31.19	81.98
ARHGAP18	1nuc	-70.39	10.15	34.85	81.89
RPLP2	1nuc	132.47	-2.92	7.71	81.78
RPS13	1nuc	300.02	-14.63	15.90	81.74
RUSC2	1nuc	97.77	-0.89	33.70	81.72
CHRNA10	1nuc	165.73	-13.08	24.05	81.48
RPL8	1nuc	11.91	12.39	17.00	81.46
ANLN	1nuc	-88.18	70.27	27.91	81.37
FADS2	1nuc	-58.80	19.27	61.95	81.27
KIF11	5nuc	-26.24	38.88	9.53	81.25
KLRC3	1nuc	-48.14	14.65	40.27	81.11
RPL35A	1nuc	61.10	0.07	13.51	80.82
ZNF138	1nuc	71.32	-9.70	30.86	80.81
EVC2	1nuc	-54.09	10.38	44.17	80.77
RPL23A	1nuc	-7.82	-1.56	25.86	80.74
RPL24	1nuc	134.93	-7.43	10.18	80.69
BCCIP	1nuc	-3.95	6.06	31.99	80.62
CRK	1nuc	-57.29	5.72	37.53	80.33
RPL7	1nuc	16.46	1.96	15.19	80.20
COLQ	1nuc	-85.17	17.46	52.70	80.16
MRPS2	1nuc	39.97	-9.32	38.88	80.12
CASP8AP2	5nuc	-38.36	29.49	19.98	80.07
MAML3	1nuc	-92.51	38.97	58.98	79.90
SLCO5A1	1nuc	-57.45	14.14	52.96	79.90
WDR12		-39.49	10.31	36.35	79.79
RPS11	1nuc	225.25	-15.47	17.21	79.62
ST7L	1nuc	-66.96	7.23	37.78	79.53
GSS	1nuc	-42.36	9.97	47.61	79.52
C11orf16	1nuc	2.95	-0.77	78.26	79.50
TRMO	1nuc	68.23	-8.22	47.15	79.48
EGLN3	1nuc	11.87	-2.51	28.17	79.47
DOCK8	1nuc	-20.30	7.82	23.21	79.19
KCNK2	1nuc	-65.12	11.07	52.51	79.04
FBXO17	1nuc	45.49	-8.88	35.15	78.93
ATP5IF1	1nuc	-60.79	2.58	40.00	78.73
KLF16	1nuc	-75.03	22.23	54.42	78.52
ASPHD1		-41.48	8.57	24.45	78.30
TNS4	1nuc	-27.39	12.02	54.12	78.28
RHBDL1	1nuc	28.97	2.96	19.13	78.03
IL1B	1nuc	-22.37	1.25	58.44	77.97

WDR37	1nuc	-26.27	12.88	67.19	77.78
RSL24D1	1nuc	63.87	-5.43	29.35	77.76
NOP56	1nuc	267.68	-14.80	48.90	77.74
TST	1nuc	-58.71	12.55	46.41	77.71
PXMP2	1nuc	-36.59	12.57	47.87	77.41
NEPRO	1nuc	-9.29	-3.50	48.19	77.33
TSEN34	1nuc	-50.74	12.60	66.55	77.22
UTP4	1nuc	121.66	-9.42	36.91	77.17
RPS15A	1nuc	119.54	-8.71	15.10	77.09
WDR45	1nuc	-9.36	13.97	70.51	77.07
GCNT2	5nuc	-16.97	5.94	54.70	77.06
CENPJ	1nuc	-29.87	2.25	47.56	77.05
RPL37	1nuc	-28.66	9.57	66.28	76.49
UBA52	1nuc	36.36	-2.40	23.72	76.41
UGT1A7	1nuc	-33.59	12.24	74.05	76.34
TTC27	1nuc	-23.74	11.91	46.86	76.20
NTN3	1nuc	213.41	-11.57	21.54	76.18
ARSG	1nuc	-10.25	6.36	40.57	76.08
S100A10	1nuc	-28.99	26.78	27.30	75.86
OR2B11	1nuc	-21.31	2.25	38.59	75.62
UBE2F	1nuc	-38.62	7.53	70.66	75.40
WRAP53	5nuc	-50.40	19.43	51.43	75.30
NIBAN2	1nuc	-67.51	9.59	56.46	75.10
OSER1		-95.68	13.14	40.88	75.06
LAMB3	1nuc	-16.69	-0.25	60.28	75.06
TAF4	1nuc	-48.46	15.55	44.60	74.96
ASB6	1nuc	-30.47	16.58	85.66	74.89
CHCHD1	1nuc	-68.05	8.71	51.43	74.67
NUTF2	1nuc	-27.78	10.89	22.30	74.42
RFX5	1nuc	-3.55	5.18	26.22	74.39
RAB36	1nuc	-25.16	16.47	34.99	74.37
ART3	1nuc	-8.12	-1.66	60.35	74.30
ISOC2	1nuc	-90.20	28.77	50.30	74.13
IKZF4	1nuc	-13.78	13.58	66.66	74.12
CRTC1	1nuc	-36.69	7.67	73.44	74.00
EIF4E3	1nuc	-55.76	3.24	56.11	73.97
TARS2	5nuc	-29.03	10.21	66.24	73.87
SKP1	5nuc	4.02	40.82	14.27	73.34
EPN1	1nuc	-52.08	8.07	65.16	73.26
NOL7		193.77	-12.35	43.72	73.23
PHC2	1nuc	68.35	5.46	24.21	73.10

FBN2	1nuc	-45.56	10.01	69.23	73.05
USB1	1nuc	21.89	0.15	46.62	72.97
SORBS3	1nuc	-37.82	7.36	40.15	72.72
SFXN3	1nuc	27.92	-3.17	41.67	72.67
MRPL52	5nuc	-10.60	6.74	38.27	72.66
FNDC3B	1nuc	93.60	-8.79	46.42	72.20
MEPCE	1nuc	-49.25	6.73	71.25	72.04
SGO1	5nuc	5.05	31.08	14.53	71.85
CMIP	1nuc	-72.39	23.02	42.64	71.83
CAPN2	1nuc	-21.89	9.07	58.97	71.75
SYT1	1nuc	105.75	-7.41	26.02	71.70
RNPS1	1nuc	-47.66	22.19	37.78	71.56
PCDHB7	1nuc	-43.36	22.37	67.28	71.52
CREB3L4	1nuc	35.13	6.56	21.50	71.06
IRF3	1nuc	-66.44	11.80	51.70	70.98
RPP30	1nuc	-58.55	17.88	31.98	70.96
POLR2J3	1nuc	166.89	-9.71	21.27	70.70
UFSP1	1nuc	26.38	-0.21	50.81	70.62
UPF3A	1nuc	-35.67	14.85	78.61	70.29
GEMIN7	1nuc	-77.90	22.82	61.60	70.12
SLC35F2	1nuc	-65.62	8.78	50.34	69.94
TMEM260	1nuc	-47.45	14.87	58.22	69.91
EXOSC9	1nuc	-95.34	38.72	49.95	69.71
ARMC2	1nuc	-26.65	12.53	69.84	69.41
CRYL1	1nuc	-77.14	11.35	44.88	69.39
YARS2	1nuc	-29.18	6.99	48.84	69.35
PAXX	5nuc	-51.52	18.01	61.21	69.17
ZSCAN20	1nuc	-37.09	6.23	59.52	69.08
JRK	1nuc	-40.40	13.90	72.12	68.77
RIOX2	1nuc	-47.35	6.21	74.91	68.48
SUSD3	1nuc	-57.03	20.14	59.04	68.46
CYP4V2	5nuc	-27.16	1.14	82.43	68.45
CHML	1nuc	-76.31	12.77	56.12	68.17
RPL36AL	1nuc	22.12	9.59	29.19	67.93
MAPT	1nuc	-69.12	19.74	61.28	67.62
RBM7	1nuc	-5.49	6.39	56.77	67.50
RGS12	1nuc	5.39	4.61	76.74	67.35
RPS10	1nuc	-1.88	22.11	12.56	66.98
NUP35	1nuc	-33.46	3.20	54.96	66.81
FAM98A	5nuc	-8.45	4.10	91.62	66.45
NUMA1	1nuc	220.62	-12.57	20.13	66.38

ZNF708	1nuc	7.72	7.58	51.94	66.37
TPX2	5nuc	-64.81	98.89	13.84	66.31
TNFRSF8	1nuc	-28.37	7.21	54.31	66.21
C7orf61	1nuc	-2.82	19.80	31.54	66.17
CYTH1	1nuc	-58.56	13.56	30.95	66.08
STARD5	1nuc	-82.13	30.25	58.00	66.01
C7orf25	1nuc	-47.17	15.33	61.75	66.00
MPL	1nuc	57.45	-5.39	44.80	65.99
FAM20A	1nuc	-74.41	11.08	54.66	65.77
MPHOSPH10	1nuc	208.70	-13.63	51.91	65.55
ZDHHC17	5nuc	2.53	4.35	58.45	65.54
GPR156	1nuc	-66.09	27.72	43.61	65.50
C22orf23	1nuc	-80.94	12.65	60.13	65.45
ARMCX3	1nuc	-61.68	13.80	58.98	65.21
C1orf216	1nuc	-84.08	19.68	42.40	65.13
NSA2	1nuc	181.96	-9.32	22.13	64.92
RNF121	1nuc	47.97	3.09	28.35	64.78
HSPB1	1nuc	-62.78	10.77	41.97	64.77
TRMT112	1nuc	-26.09	4.33	69.04	64.72
RASGRP3	1nuc	-30.28	16.68	46.00	64.65
RPS21	1nuc	109.70	-10.28	14.66	64.56
ZNF22-AS1	1nuc	33.49	3.34	49.23	64.51
ZNF684	1nuc	-16.22	7.33	64.30	64.17
EPB41L2	1nuc	-83.80	18.64	54.76	64.05
CENPL	1nuc	-27.41	6.99	44.86	64.01
CDCA8	5nuc	-90.35	99.95	26.96	63.70
SCAF1	1nuc	-29.40	7.97	65.29	63.68
RNF169	1nuc	19.85	7.40	46.25	63.55
LPIN2	1nuc	-21.28	13.58	94.92	63.31
CCDC22	1nuc	-58.64	16.85	53.39	63.02
PES1		26.13	-6.19	43.19	62.56
TBC1D14	1nuc	-54.12	20.38	57.72	62.45
VCPKMT	1nuc	-6.80	-0.30	80.07	62.45
CD81	1nuc	-62.97	24.00	68.64	62.24
CNOT1	1nuc	-69.66	28.23	35.98	62.23
MRPL34	1nuc	-59.54	28.66	49.50	62.22
CMAS	1nuc	-74.91	15.23	63.20	62.15
KRBA1	5nuc	-7.33	5.50	53.02	62.13
HMGN2	1nuc	-89.71	25.92	45.96	61.88
SNAPC4	1nuc	-42.31	7.59	67.45	61.51
TACSTD2	1nuc	-8.79	3.72	67.99	61.46

DHCR7	1nuc	63.70	-4.58	41.69	61.43
TINF2	1nuc	-41.16	14.97	68.27	61.36
SLC4A7	1nuc	-59.98	18.66	74.76	61.10
SMOC2	5nuc	9.67	0.61	42.16	61.09
KTN1	5nuc	29.05	5.59	42.64	61.02
NOP58	1nuc	262.37	-17.46	59.92	60.94
C9orf139	1nuc	-72.11	14.07	42.34	60.91
COL8A2	1nuc	-52.27	10.22	46.17	60.80
RABGGTB	1nuc	-46.60	9.68	34.57	60.79
GLYATL2	5nuc	-16.14	4.09	58.13	60.76
TBC1D5	1nuc	-50.03	15.36	71.44	60.63
CNPY4	1nuc	-85.64	11.27	64.31	60.59
APCDD1	1nuc	-19.23	7.34	48.00	60.53
PCTP	1nuc	-14.96	2.66	60.30	60.41
ROBO3	1nuc	-31.86	11.17	74.61	60.34
SLC36A1	1nuc	-38.07	10.19	56.97	60.24
B3GAT1	1nuc	-56.30	12.35	76.78	60.24
CDS2	1nuc	-57.87	12.27	67.24	60.13
PCDHA9	1nuc	7.45	5.34	42.42	60.11
COBL	1nuc	-59.20	18.31	62.66	59.87
MZT2A	1nuc	-41.79	20.63	77.91	59.82
IQCK	1nuc	-50.17	7.60	75.55	59.79
OXNAD1	5nuc	-19.06	12.59	68.59	59.69
LYG1	1nuc	-55.99	6.27	85.62	59.53
FGFBP3	1nuc	-46.64	7.60	45.35	59.46
ISLR	1nuc	-45.40	5.33	83.54	59.32
ARC	1nuc	-45.05	5.54	63.97	59.17
CRLF1	1nuc	-37.02	6.03	76.75	58.99
UTP20	1nuc	11.95	-0.76	55.73	58.89
ZNF41	1nuc	-11.83	7.29	83.78	58.76
DMRTA1	1nuc	-57.19	11.61	73.20	58.70
MRPL23	1nuc	-67.98	15.16	62.05	58.68
ARHGEF9	1nuc	-59.98	10.01	51.74	58.49
PDE12	1nuc	-44.08	13.90	67.78	58.33
BRPF1	1nuc	-9.68	10.02	94.03	58.30
EFCAB13	1nuc	-59.08	5.24	42.31	58.29
SIGLEC15	1nuc	-24.43	14.09	67.74	58.17
LCE1E	1nuc	-26.16	0.95	104.64	58.01
DCP2		-20.20	17.03	59.87	57.95
BCAR3	1nuc	-58.48	12.57	71.44	57.92
GCLM	1nuc	-60.91	4.23	60.82	57.89

METTL26	1nuc	52.93	-8.06	35.72	57.72
YARS1		5.17	0.10	63.34	57.66
ASAH2	1nuc	-60.56	12.48	49.40	57.60
LGALS1	1nuc	-39.14	-0.01	64.95	57.56
CLHC1	1nuc	-68.12	14.98	60.15	57.54
SYT15	1nuc	-21.27	9.81	44.95	57.48
TTC22	5nuc	0.34	9.03	81.39	57.22
HNRNPA1L2	1nuc	-7.20	3.10	64.07	57.18
RPUSD2	1nuc	-12.88	4.82	62.95	57.07
GIN1	1nuc	-65.56	16.45	49.36	57.02
SAMD15	1nuc	-48.01	14.19	50.60	56.90
RASSF8	1nuc	50.06	-3.98	45.96	56.80
RRS1	1nuc	-66.83	29.49	55.52	56.68
PIK3C2B	1nuc	6.28	1.19	76.11	56.62
TAGLN	1nuc	-74.17	16.69	47.49	56.59
SELL	1nuc	-41.07	11.85	60.55	56.50
ASIC1	5nuc	20.09	3.52	74.33	56.36
FBXW8	1nuc	-49.52	15.16	80.28	56.14
EPSTI1	1nuc	-56.39	11.38	68.59	56.04
SHC3	5nuc	-61.19	6.24	51.17	55.96
ZFYVE21	1nuc	-17.99	15.72	43.99	55.66
UPF1	1nuc	1.40	-1.90	76.46	55.63
MMP23B	1nuc	-26.85	2.50	87.08	55.51
CXXC4	1nuc	-12.45	1.12	27.50	55.46
IRF2BP1	5nuc	26.39	-1.82	73.48	55.21
HAPLN3	1nuc	-57.74	11.17	74.73	55.18
SYNE1	1nuc	-22.51	5.14	57.23	54.96
APOBEC4	1nuc	60.49	-11.49	32.40	54.94
PPP1R3B	1nuc	-48.88	15.37	57.62	54.85
APOL3	1nuc	-59.89	22.23	42.67	54.82
ARHGEF35	1nuc	-69.98	10.02	71.35	54.80
IST1	1nuc	-60.94	9.34	76.35	54.69
GAST	1nuc	-81.94	28.23	54.74	54.69
H1-10	5nuc	9.24	3.13	42.59	54.68
ATG13	1nuc	-5.12	5.95	44.49	54.53
NDUFA8	1nuc	-67.47	11.26	61.98	54.47
VPREB3	1nuc	-31.79	13.45	81.10	54.45
ATAD5	5nuc	-27.34	12.99	52.00	54.31
LRRC8C	1nuc	-49.94	12.86	42.98	54.27
HSPA12A	1nuc	-25.78	9.24	83.16	54.13
RAB38	1nuc	33.99	-0.83	31.71	54.03

MASTL	5nuc	-60.44	63.98	32.01	53.88
STK24	5nuc	-16.88	4.66	62.48	53.78
NUDT16		18.07	-0.34	77.53	53.72
EPS15L1	1nuc	-73.99	19.33	74.62	53.68
WFDC3	1nuc	-18.86	10.24	77.60	53.66
TXNRD2	1nuc	-42.19	11.81	63.64	53.60
ZNF678	5nuc	38.38	12.35	23.85	53.55
ELK3	1nuc	-79.71	18.36	72.17	53.40
CCDC9	1nuc	-68.85	14.88	72.16	53.29
PFDN2	1nuc	-21.89	8.41	81.38	53.24
GALK2	1nuc	-76.74	21.41	63.44	53.23
TBL1X	1nuc	-70.72	19.48	57.27	53.20
SRPRA	1nuc	-25.73	2.66	48.01	53.02
S100A12	1nuc	-75.53	20.80	65.86	52.84
CHAF1A	1nuc	-33.78	20.69	39.63	52.75
PLOD3	1nuc	-16.39	12.67	52.95	52.46
SNX21	5nuc	-6.46	5.98	83.94	52.44
VKORC1L1	1nuc	-17.14	11.20	79.49	52.43
NDUFS3	1nuc	-44.47	2.90	58.45	52.16
SHARPIN	1nuc	-12.70	12.90	39.24	52.14
LRRC4	1nuc	-33.00	19.92	98.02	52.01
SNX13	1nuc	-49.61	15.73	67.90	51.78
OVOL2	1nuc	18.93	-10.61	48.66	51.67
CARMIL1	1nuc	0.50	2.60	85.67	51.63
FAM131C	1nuc	-87.22	16.75	41.78	51.55
TPST1	1nuc	-13.72	4.56	76.39	51.53
C2orf16	1nuc	-70.27	17.09	47.84	51.47
OXSM	1nuc	9.53	2.83	66.16	51.47
C20orf96	1nuc	-58.84	10.92	63.41	51.42
NOLC1	1nuc	-28.82	4.74	74.87	51.33
PIGX	1nuc	-21.71	9.21	79.44	51.28
BICDL1	1nuc	-73.89	21.70	60.98	51.13
ZNF502	1nuc	4.47	6.83	58.37	50.97
RPP38-DT	1nuc	-20.63	6.77	78.98	50.91
GABPB2	1nuc	-96.01	20.88	69.32	50.90
MAP3K11	1nuc	-99.12	26.86	49.45	50.90
ZNF777	1nuc	-52.85	13.41	73.75	50.84
TRMT10A	1nuc	-11.47	8.55	66.79	50.79
MEGF6	1nuc	-29.47	6.29	71.53	50.57
RIMS3	5nuc	14.21	8.33	76.15	50.44
ITGA3	1nuc	-28.02	5.13	64.87	50.36

ACD	1nuc	-64.45	15.97	45.90	50.33
CCDC125	1nuc	-26.81	6.60	78.00	50.30
MCM6	5nuc	-15.36	9.21	50.26	50.14
H4C8	1nuc	-63.50	6.29	75.20	50.10
LHFPL6	1nuc	8.14	2.29	32.61	50.05
SLC46A2	1nuc	-46.36	13.49	65.76	50.04
FAM221A	5nuc	-5.45	5.94	45.27	49.87
ASCC2	1nuc	-17.81	7.58	72.88	49.85
MVP	1nuc	15.87	-0.08	74.74	49.82
KCTD20	1nuc	-12.24	1.69	90.25	49.73
CUL1	5nuc	-22.31	17.14	37.49	49.22
APOBEC3A	1nuc	299.38	-18.53	21.38	49.06
UBE4B	1nuc	-22.54	14.39	44.81	48.98
PPP1R9B	1nuc	-25.83	7.06	41.43	48.87
INKA1	5nuc	-16.69	6.85	85.62	48.85
EBF3	5nuc	-40.01	14.58	76.59	48.78
LMBR1	1nuc	-20.25	4.99	76.14	48.55
KCTD4	1nuc	-34.09	14.56	72.08	48.54
CRYBG3	1nuc	-39.96	2.29	48.08	48.52
PCOLCE2	5nuc	-1.50	3.11	56.89	48.51
RAB11FIP2	1nuc	36.40	0.84	43.35	48.50
STXBP2	1nuc	-55.73	13.89	78.46	48.41
ZZZ3	1nuc	-50.04	17.43	68.05	48.38
MCOLN2	1nuc	-37.94	7.00	77.81	48.32
NUP58	1nuc	-4.03	-0.09	57.75	48.18
LRRC32	1nuc	-32.47	4.96	82.70	48.16
FGD4	5nuc	83.20	-6.15	65.24	48.12
GJC3	1nuc	8.11	-5.53	53.60	48.11
SPACA4	1nuc	-33.50	8.36	71.10	48.08
MUSTN1	1nuc	-46.81	17.63	73.83	48.04
SLC16A8	1nuc	-50.65	20.27	53.52	47.95
NODAL	1nuc	-34.02	10.25	53.93	47.91
NMT2	1nuc	-45.08	7.22	82.55	47.91
PARP3	1nuc	-19.85	0.19	49.64	47.85
NOL8	1nuc	26.02	-2.79	69.13	47.82
SYCP3	1nuc	-50.31	19.08	78.26	47.69
TSPO	1nuc	-57.71	8.39	47.57	47.57
CTF1	5nuc	-45.07	19.65	50.32	47.27
SMG5	5nuc	18.75	0.61	70.11	47.25
RACGAP1	5nuc	-67.42	68.94	31.85	47.13
EAF1	1nuc	-58.16	12.13	55.69	47.07

MINAR1	1nuc	-70.92	17.35	80.51	46.89
MMP25	1nuc	1.64	-1.56	82.17	46.88
CA14	1nuc	-44.26	11.03	65.23	46.84
FAM177A1	1nuc	-92.06	29.25	69.52	46.84
CLCN7	1nuc	-86.79	17.61	55.28	46.75
BAK1	1nuc	-51.93	7.12	62.88	46.65
LDB1	5nuc	-30.47	11.51	84.43	46.64
GSKIP	1nuc	-24.24	8.08	48.34	46.62
LUC7L	5nuc	-5.39	-1.19	83.28	46.60
OLFML3	1nuc	-15.74	9.99	87.12	46.51
METTL15	1nuc	15.94	-0.25	87.79	46.47
ZSWIM3	1nuc	-22.86	7.20	66.89	46.42
SCMH1	1nuc	-63.22	19.77	58.11	46.39
TMEM138	1nuc	-20.18	5.86	74.91	46.33
BBS10	1nuc	-77.32	19.30	69.46	46.32
GIGYF2	1nuc	-56.78	16.34	86.55	46.28
SUV39H1	5nuc	-30.68	14.05	61.23	46.09
PROCA1	1nuc	-16.11	6.42	84.11	45.98
PARD6A	1nuc	-18.86	6.90	71.76	45.86
SHROOM2	5nuc	16.85	-0.74	77.30	45.85
PTHLH	1nuc	-16.86	11.39	72.71	45.85
DCLRE1B	1nuc	-43.54	1.55	71.37	45.76
ICA1	1nuc	-37.91	12.08	89.47	45.70
RBFOX2	1nuc	-30.15	9.22	73.05	45.60
LRRC42	1nuc	-23.51	9.95	94.02	45.57
PUS1	1nuc	-6.62	6.06	97.98	45.49
JHY	5nuc	-1.43	6.67	94.84	45.48
TSGA10	1nuc	-36.12	9.73	65.09	45.48
HOXC5	1nuc	-40.27	11.68	76.83	45.26
FOXM1	1nuc	-88.53	30.66	53.86	45.06
ANXA11	1nuc	-41.51	8.37	63.52	44.78
C1QL1	1nuc	-67.15	8.63	60.51	44.77
TM2D1	5nuc	-3.74	9.29	76.55	44.69
ZNF225	1nuc	-14.16	8.82	93.70	44.67
TOPBP1	5nuc	-53.11	29.31	38.98	44.66
SP100	1nuc	-49.17	9.55	63.89	44.63
RPS23	1nuc	-55.29	15.28	57.23	44.42
ZNF555	1nuc	9.83	3.20	98.69	44.41
THAP1	1nuc	-19.17	7.46	67.73	44.41
XRCC5	5nuc	-14.50	3.26	62.33	44.37
ENY2	5nuc	-18.63	11.97	99.30	44.15

TUFT1	1nuc	-37.00	16.60	79.51	43.96
THRB	1nuc	-40.13	7.31	51.56	43.76
METTL27	5nuc	-6.29	3.94	71.87	43.66
ZNF747	1nuc	-9.16	8.37	81.25	43.65
CCDC34	1nuc	-61.43	4.54	48.13	43.60
STAG1	1nuc	-62.80	21.92	70.96	43.28
PAX9	1nuc	-4.30	6.60	74.50	43.12
RBM28	1nuc	7.75	-1.30	46.87	43.04
KCNAB2	1nuc	-10.47	5.26	88.39	42.97
ABCB10	1nuc	-56.62	32.75	59.00	42.93
SPARCL1	1nuc	-48.53	14.66	47.35	42.91
CES2	1nuc	-22.10	8.90	54.85	42.84
ATP6AP1L	1nuc	-42.12	7.47	46.99	42.75
DYNC1H1	5nuc	-43.70	20.84	54.26	42.73
ZMYND12	1nuc	-33.37	9.58	93.61	42.60
RANGAP1	1nuc	25.54	-3.00	48.40	42.36
COX20	1nuc	-30.81	10.15	70.93	42.36
RBM43	1nuc	-29.66	11.50	51.02	42.30
NDUFAF1	1nuc	-13.59	1.59	71.03	42.26
TERF2	1nuc	-39.34	6.73	46.63	42.26
NAA10	1nuc	-15.82	9.80	80.90	42.03
THAP6	1nuc	-56.37	18.34	60.14	41.39
BZW1	1nuc	-23.11	5.91	87.23	41.33
TMEM220	1nuc	-60.30	11.12	68.55	41.23
LIN28A	1nuc	-34.14	6.66	68.34	41.13
ZNF649	1nuc	33.72	-3.44	87.79	41.11
ALDH8A1	1nuc	-47.92	8.05	74.15	41.07
SLC38A1	5nuc	29.66	-3.70	64.82	40.98
HDAC9	1nuc	-59.50	20.14	59.19	40.95
TRIM4	1nuc	-20.37	8.31	96.04	40.73
PLAA	1nuc	227.77	-14.31	23.47	40.66
SLC25A17	1nuc	-42.53	10.05	68.88	40.61
CMPK2	5nuc	15.39	2.58	71.60	40.35
CCDC82	1nuc	-73.47	18.13	63.74	40.34
BTBD16	1nuc	24.22	-6.72	70.88	40.17
CCDC81	5nuc	-18.01	22.32	69.88	40.16
CAMK2N1	5nuc	-27.17	12.53	76.95	40.11
ALAD	1nuc	-43.59	12.17	69.57	39.48
GEN1	5nuc	-11.32	7.37	86.92	39.47
RREB1	1nuc	0.10	-4.45	103.80	39.34
C4orf46	1nuc	-20.04	6.28	91.92	39.29

KCNAB1	1nuc	-9.79	-1.48	97.36	39.14
SH3TC2	1nuc	-30.56	8.64	79.59	39.11
NANP	1nuc	-20.05	0.16	76.54	39.04
RGMB	1nuc	-7.52	9.49	80.63	38.69
AHDC1	1nuc	-37.82	4.84	74.88	38.48
MTRF1L	1nuc	-23.15	9.84	100.96	38.41
PCDHAC1	1nuc	-4.78	-1.68	55.35	38.35
RASAL2	1nuc	108.32	-9.69	36.91	38.32
MTBP	1nuc	-68.36	23.01	57.22	38.12
SMAD5	5nuc	-11.51	9.00	77.37	38.10
ANKEF1	5nuc	-42.51	17.57	75.13	37.99
SCN1B	1nuc	-46.76	9.96	88.97	37.73
NGRN	5nuc	-4.01	-3.00	69.10	37.64
PPP4R4	1nuc	23.56	-1.13	75.36	37.63
HPRT1	1nuc	-34.41	9.82	90.15	37.63
PHF13	1nuc	22.29	1.62	91.54	37.63
RAI1	1nuc	-21.63	2.81	85.33	37.17
PRKD3	1nuc	-43.85	17.06	79.15	37.01
IL6R	1nuc	-36.83	5.40	78.28	36.95
STX6	1nuc	-39.21	12.53	57.65	36.93
SOWAHC	1nuc	-36.74	11.47	80.09	36.50
ABCE1	5nuc	-20.73	9.24	60.05	36.43
AIMP1	1nuc	-28.08	-5.59	78.24	36.31
PCED1A	1nuc	-34.56	3.82	69.31	36.30
NAXE	1nuc	-19.02	8.86	76.00	35.94
PRUNE1	5nuc	-18.09	7.26	68.27	35.46
TNFRSF21	1nuc	-23.04	10.38	68.80	35.39
PRR19	1nuc	-16.87	8.85	86.22	35.32
CHN2	1nuc	-54.75	17.48	93.42	35.23
NLRC5	5nuc	-20.39	9.13	54.72	35.12
PIK3C2A	1nuc	-25.54	6.54	57.13	35.11
PRKAA1	1nuc	7.57	2.78	64.90	35.00
ITGA1	1nuc	-35.30	8.78	92.48	34.90
CYCS	1nuc	-74.98	3.91	54.57	34.65
ARL14EP	1nuc	-16.26	10.09	81.45	34.62
PEDS1	1nuc	-29.09	8.47	77.24	34.56
MLEC	1nuc	-20.50	5.03	106.96	34.52
RAP2C	5nuc	-12.41	1.87	57.33	34.46
TMEM101	1nuc	-22.92	10.11	78.30	34.40
TMEM120A	1nuc	-7.76	8.51	80.01	34.14
TECPR2		-13.47	5.56	65.27	34.05

SRSF12	1nuc	-47.06	14.05	72.21	33.91
IL22RA1	1nuc	26.89	-4.97	67.91	33.35
SCRN3	5nuc	-26.34	13.51	86.61	32.76
TSSK6	1nuc	-12.55	4.78	85.21	32.58
NUSAP1	1nuc	-22.10	7.47	42.04	32.54
ZNF552	1nuc	13.63	4.00	97.74	32.51
IL33	1nuc	-38.53	5.43	74.61	32.19
AAMP	1nuc	86.07	-4.16	30.85	31.60
ZNF630	1nuc	-3.99	5.06	69.79	31.60
RAB19	1nuc	-18.56	11.51	58.38	31.03
FLRT1	1nuc	-71.46	22.92	74.28	30.75
ARL11	1nuc	-6.27	4.04	91.67	30.73
OAZ3	1nuc	-11.09	0.56	58.22	30.72
SPAG1	1nuc	175.27	-11.34	22.66	30.56
CRYBG2	1nuc	-62.21	9.33	79.57	30.50
MPST	1nuc	178.67	-9.54	35.93	30.19
PATL1	1nuc	-16.56	8.42	81.10	30.19
NNT	1nuc	-48.63	3.66	87.10	30.16
LAPTM5	5nuc	14.03	-3.23	100.81	30.15
SLC2A12	5nuc	17.83	-2.11	84.22	30.13
LSM14A	1nuc	-16.82	4.07	89.73	30.09
MTFR2	1nuc	-54.38	17.10	79.76	30.01
SYVN1	1nuc	47.04	-3.07	88.28	29.93
KRT6A	1nuc	-32.00	-1.44	69.08	29.90
OSBP2	5nuc	-23.87	6.89	78.10	29.83
NFATC4	1nuc	36.26	-2.88	100.22	29.42
RAPH1	5nuc	-7.58	9.11	89.98	29.18
BCAR1	1nuc	-8.04	5.14	42.61	28.62
ZC2HC1A	1nuc	-30.52	10.76	96.67	28.53
RFFL	5nuc	-21.58	12.87	81.15	28.24
C11orf52	1nuc	-52.84	12.10	66.56	27.94
ONECUT2	1nuc	11.48	-1.87	32.12	27.85
HYPK	5nuc	-6.23	-1.74	79.75	27.71
TTF1	1nuc	19.26	2.84	70.19	27.40
ACTG1	1nuc	-27.36	7.46	65.54	27.31
CBR1	1nuc	-3.97	-3.88	59.61	27.06
AKT1	1nuc	50.04	-7.62	44.09	27.00
CCN4	5nuc	-28.11	13.03	79.61	25.04
KRT222	5nuc	17.34	0.33	98.34	24.97
OPTN	1nuc	-13.08	3.61	72.67	24.89
MSL2	1nuc	-41.23	5.79	89.70	24.89

MPV17L2	5nuc	-6.29	-3.89	66.66	24.79
PSORS1C2	1nuc	-25.27	0.48	61.29	24.71
MAF1	1nuc	-25.61	7.01	82.67	24.49
NUP43	1nuc	-4.30	1.00	80.41	24.32
PRRT2	5nuc	-14.27	5.85	75.28	24.29
RNF19B	1nuc	-12.62	11.27	65.93	24.28
MAP4K5	5nuc	6.31	4.92	69.07	24.04
RBBP8	5nuc	8.08	14.77	32.58	23.97
PMM2	5nuc	-28.27	8.82	67.49	23.82
SYNJ2BP	1nuc	-44.64	15.79	62.98	23.29
GTF2IRD1	5nuc	-15.28	10.15	80.65	23.09
AQP11	1nuc	-17.48	5.58	86.45	22.63
PHACTR4	1nuc	-2.33	3.06	91.67	22.50
PSCA	1nuc	42.91	-5.53	58.23	22.22
TSPAN5	1nuc	26.15	-3.28	92.21	22.14
ZSCAN5A	1nuc	21.10	0.54	81.71	22.11
GLTP	5nuc	11.71	-0.49	93.67	21.83
ARHGAP28	1nuc	-12.02	6.51	94.23	19.86
PLSCR4	1nuc	-5.26	3.61	95.46	19.51
NUDT13	1nuc	-32.15	6.24	80.28	19.42
DDAH1	5nuc	-23.04	5.67	101.41	19.32
KRAS	1nuc	-36.41	0.56	75.92	17.48
IFT88	5nuc	13.05	-2.25	103.88	17.43
ZNF571	1nuc	27.00	-2.57	108.18	17.20
NEXN	1nuc	-17.65	5.31	93.97	17.14
POLR1D	1nuc	-14.63	3.81	101.42	16.69
A1BG	1nuc	-42.06	14.85	67.53	16.66
R3HDM1	1nuc	-9.71	12.36	86.96	16.35
HSD11B2	5nuc	43.57	-1.40	78.31	16.30
MIA	5nuc	-18.55	10.94	61.07	15.25
MAN1A1	5nuc	-4.59	3.02	79.69	15.17
ARHGAP19	1nuc	-31.24	2.33	83.83	13.39
RHBDF2	1nuc	-4.06	3.75	79.00	12.70
AIFM3	1nuc	-6.65	2.02	56.78	12.61
NDUFB6	1nuc	-25.04	-2.50	85.32	12.30
AIF1L	1nuc	-17.73	3.76	81.87	12.27
SBDS		21.14	-2.72	85.70	12.20
NKAIN1	1nuc	-11.84	2.64	100.71	12.16
WASF2	1nuc	63.91	-4.52	108.05	11.74
TYMP	5nuc	-11.86	9.91	77.13	11.46
TMEM107	1nuc	-12.87	7.79	89.41	11.42

NFYB	5nuc	-34.40	10.51	59.17	11.30
LCORL	5nuc	30.01	-2.19	73.28	10.97
TAF1D	5nuc	-12.58	8.75	81.54	10.76
NOMO1	1nuc	18.52	-6.26	109.18	10.42
COQ8A	1nuc	-53.08	7.17	71.48	8.84
NAP1L4	1nuc	42.67	-12.04	84.14	8.05
SPOP	1nuc	-34.30	8.96	88.01	7.59
ABHD15	1nuc	-12.91	2.19	53.85	7.04
ANKHD1-EIF4EBP3	1nuc	3.05	-0.35	66.68	6.54
GZF1	5nuc	9.10	-1.33	77.88	5.49
MFSD4A	5nuc	-0.25	6.94	67.43	5.02
AP2M1	1nuc	-42.66	11.55	67.86	4.79
POC5	1nuc	-8.61	13.06	69.60	4.41
RFC1	5nuc	-33.88	16.88	62.42	3.38
KMT2C	5nuc	1.90	0.77	90.00	-5.31
NOX1	1nuc	-10.69	1.08	67.12	-5.55
ADNP2	1nuc	-60.44	-4.19	57.13	-25.86

Appendix II: C to U editing on the 47S pre-rRNA (NR_145144.1) in siNT, siAPOBEC3A pool, and siAPOBEC3A #1 treated cells. (N = 2).

siNT Replicate 1

Chr	Position	Ref	Alts	Read depth	Allele Frequency
NR_146144.1	49	C	T	1232	0.00974
NR_146144.1	800	C	T	20152	0.004913
NR_146144.1	927	C	T	96687	0.015783
NR_146144.1	1098	C	T	75665	0.003489
NR_146144.1	1279	C	T	80815	0.0074
NR_146144.1	1514	C	T	32412	0.006201
NR_146144.1	1685	C	T	16129	0.010292
NR_146144.1	2224	C	T	31034	0.020945
NR_146144.1	2604	C	T	26573	0.006699
NR_146144.1	2813	C	T	6500	0.004462
NR_146144.1	3949	C	T	494633	0.001822
NR_146144.1	5926	C	T	52681	0.214707
NR_146144.1	7091	C	T	76513	0.004548
NR_146144.1	7821	C	T	56618	0.004363
NR_146144.1	8160	C	T	993038	0.006955
NR_146144.1	9193	C	T	114785	0.011387
NR_146144.1	10409	C	T	967131	0.003289
NR_146144.1	10598	C	T	911741	0.001638
NR_146144.1	10939	C	T	42000	0.029214
NR_146144.1	11772	C	N/A	N/A	N/A

siAPOBEC3A pool Replicate 1

Chr	Position	Ref	Alts	Read depth	Allele Frequency
NR_146144.1	49	C	N/A	N/A	N/A
NR_146144.1	800	C	N/A	N/A	N/A
NR_146144.1	927	C	T	57077	0.015856
NR_146144.1	1098	C	T	35988	0.003696
NR_146144.1	1279	C	T	33353	0.006986
NR_146144.1	1514	C	T	17253	0.00823
NR_146144.1	1685	C	N/A	N/A	N/A
NR_146144.1	2224	C	T	13727	0.020616
NR_146144.1	2604	C	N/A	N/A	N/A
NR_146144.1	2813	C	N/A	N/A	N/A
NR_146144.1	3949	C	T	542184	0.001948
NR_146144.1	5926	C	T	25709	0.201291
NR_146144.1	7091	C	T	73512	0.003904
NR_146144.1	7821	C	T	71541	0.004249
NR_146144.1	8160	C	T	991078	0.006536
NR_146144.1	9193	C	T	101445	0.010853
NR_146144.1	10409	C	T	928610	0.00325
NR_146144.1	10598	C	T	N/A	N/A
NR_146144.1	10939	C	T	34632	0.028038
NR_146144.1	11772	C	T	825205	0.00201

siAPOBEC3A #1 Replicate 1

Chr	Position	Ref	Alts	Read depth	Allele Frequency
NR_146144.1	49	C	N/A	N/A	N/A
NR_146144.1	800	C	N/A	N/A	N/A
NR_146144.1	927	C	T	54385	0.016512
NR_146144.1	1098	C	T	38806	0.003737
NR_146144.1	1279	C	T	34649	0.008023
NR_146144.1	1514	C	T	16438	0.006205
NR_146144.1	1685	C	N/A	N/A	N/A
NR_146144.1	2224	C	T	14495	0.017661
NR_146144.1	2604	C	T	9305	0.006448
NR_146144.1	2813	C	N/A	N/A	N/A
NR_146144.1	3949	C	T	490264	0.001493
NR_146144.1	5926	C	T	29623	0.208588
NR_146144.1	7091	C	T	69500	0.004201
NR_146144.1	7821	C	T	71026	0.003858
NR_146144.1	8160	C	T	992004	0.006414
NR_146144.1	9193	C	T	91582	0.010854
NR_146144.1	10409	C	T	951270	0.003162
NR_146144.1	10598	C	T	807637	0.001488
NR_146144.1	10939	C	T	40043	0.02827
NR_146144.1	11772	C	T	846986	0.001933

Combined Replicate 1

Chr	Position	Ref	Alts	Percent Difference siNT vs. siAPOBEC3A pool	Percent Difference siNT vs. siAPOBEC3A #1
NR_146144.1	49	C	T	-100	-100
NR_146144.1	800	C	T	-100	-100
NR_146144.1	927	C	T	0.462522968	4.618893746
NR_146144.1	1098	C	T	5.932932072	7.108053884
NR_146144.1	1279	C	T	-5.594594595	8.418918919
NR_146144.1	1514	C	T	32.72052895	0.064505725
NR_146144.1	1685	C	T	-100	-100
NR_146144.1	2224	C	T	-1.570780616	-15.6791597
NR_146144.1	2604	C	T	-100	-3.746827885
NR_146144.1	2813	C	T	-100	-100
NR_146144.1	3949	C	T	6.915477497	-18.05708013
NR_146144.1	5926	C	T	-6.248515419	-2.84993037
NR_146144.1	7091	C	T	-14.16007036	-7.629727353
NR_146144.1	7821	C	T	-2.612881045	-11.57460463
NR_146144.1	8160	C	T	-6.024442847	-7.778576564
NR_146144.1	9193	C	T	-4.689558268	-4.680776324
NR_146144.1	10409	C	T	-1.185770751	-3.861356035
NR_146144.1	10598	C	T	-100	-9.157509158
NR_146144.1	10939	C	T	-4.025467242	-3.231327446
NR_146144.1	11772	C	N/A	N/A	N/A

siNT Replicate 2

Chr	Position	Ref	Alts	Read depth	Allele Frequency
NR_146144.1	49	C	N/A	N/A	N/A
NR_146144.1	800	C	N/A	N/A	N/A
NR_146144.1	927	C	T	84601	0.015721
NR_146144.1	1098	C	T	49145	0.008953
NR_146144.1	1279	C	T	25323	0.006911
NR_146144.1	1514	C	N/A	N/A	N/A
NR_146144.1	1685	C	N/A	N/A	N/A
NR_146144.1	2224	C	T	32854	0.017836
NR_146144.1	2604	C	N/A	N/A	N/A
NR_146144.1	2813	C	N/A	N/A	N/A
NR_146144.1	3949	C	N/A	N/A	N/A
NR_146144.1	5926	C	T	34134	0.253179
NR_146144.1	7091	C	T	71663	0.005414
NR_146144.1	7821	C	T	50429	0.004065
NR_146144.1	8160	C	T	852607	0.005183
NR_146144.1	9193	C	T	79628	0.011566
NR_146144.1	10409	C	T	966539	0.003562
NR_146144.1	10598	C	N/A	N/A	N/A
NR_146144.1	10939	C	N/A	N/A	N/A
NR_146144.1	11772	C	N/A	N/A	N/A

siAPOBEC3A pool Replicate 2

Chr	Position	Ref	Alts	Read depth	Allele Frequency
NR_146144.1	49	C	N/A	N/A	N/A
NR_146144.1	800	C	N/A	N/A	N/A
NR_146144.1	927	C	T	58006	0.017671
NR_146144.1	1098	C	N/A	N/A	N/A
NR_146144.1	1279	C	T	34004	0.007823
NR_146144.1	1514	C	N/A	N/A	N/A
NR_146144.1	1685	C	N/A	N/A	N/A
NR_146144.1	2224	C	T	21469	0.016675
NR_146144.1	2604	C	N/A	N/A	N/A
NR_146144.1	2813	C	N/A	N/A	N/A
NR_146144.1	3949	C	N/A	N/A	N/A
NR_146144.1	5926	C	T	20402	0.243849
NR_146144.1	7091	C	T	81719	0.004515
NR_146144.1	7821	C	T	65016	0.004291
NR_146144.1	8160	C	T	843472	0.004722
NR_146144.1	9193	C	T	71546	0.010427
NR_146144.1	10409	C	T	970533	0.003346
NR_146144.1	10598	C	N/A	N/A	N/A
NR_146144.1	10939	C	N/A	N/A	N/A
NR_146144.1	11772	C	N/A	N/A	N/A

siAPOBEC3A #1 Replicate 2

Chr	Position	Ref	Alts	Read depth	Allele Frequency
NR_146144.1	49	C	N/A	N/A	N/A
NR_146144.1	800	C	N/A	N/A	N/A
NR_146144.1	927	C	T	54793	0.01606
NR_146144.1	1098	C	N/A	N/A	N/A
NR_146144.1	1279	C	T	32016	0.007403
NR_146144.1	1514	C	N/A	N/A	N/A
NR_146144.1	1685	C	N/A	N/A	N/A
NR_146144.1	2224	C	T	20391	0.018783
NR_146144.1	2604	C	T	10843	0.006732
NR_146144.1	2813	C	N/A	N/A	N/A
NR_146144.1	3949	C	N/A	N/A	N/A

NR_146144.1	5926	C	T	19302	0.248627
NR_146144.1	7091	C	T	84879	0.004489
NR_146144.1	7821	C	T	66027	0.004574
NR_146144.1	8160	C	T	957138	0.004475
NR_146144.1	9193	C	T	78864	0.010486
NR_146144.1	10409	C	T	977895	0.003433
NR_146144.1	10598	C	N/A	N/A	N/A
NR_146144.1	10939	C	N/A	N/A	N/A
NR_146144.1	11772	C	N/A	N/A	N/A

Combined Replicate 2

Chr	Position	Ref	Alts	Percent Difference siAPOBEC3A pool vs. siNT	Percent Difference siAPOBEC3A #1 vs. siNT
NR_146144.1	49	C	T	N/A	N/A
NR_146144.1	800	C	T	N/A	N/A
NR_146144.1	927	C	T	12.40379111	2.156351377
NR_146144.1	1098	C	T	-100	-100
NR_146144.1	1279	C	T	-12.62146766	-17.31263264
NR_146144.1	1514	C	T	N/A	N/A
NR_146144.1	1685	C	T	N/A	N/A
NR_146144.1	2224	C	T	-6.50930702	5.309486432
NR_146144.1	2604	C	T	N/A	N/A
NR_146144.1	2813	C	T	N/A	N/A
NR_146144.1	3949	C	T	N/A	N/A
NR_146144.1	5926	C	T	-3.685139763	-1.797937428
NR_146144.1	7091	C	T	-16.60509789	-17.08533432
NR_146144.1	7821	C	T	5.559655597	12.52152522
NR_146144.1	8160	C	T	-8.894462666	-13.66004245
NR_146144.1	9193	C	T	-9.847829846	-9.337713989
NR_146144.1	10409	C	T	-6.064008984	-3.621560921
NR_146144.1	10598	C	T	N/A	N/A
NR_146144.1	10939	C	T	N/A	N/A
NR_146144.1	11772	C	T	N/A	N/A

Combined Results (Detected in all samples, both replicates)

Chr	Position	Ref	Alts	Avg. Percent Difference siAPOBEC3A pool vs. siNT	Standard Deviation (Percent Difference siAPOBEC3A pool vs. siNT)	Avg. Percent Difference siAPOBEC3A #1 vs. siNT	Standard Deviation (Percent Difference siAPOBEC3A #1 vs. siNT)
NR_146144.1	927	C	T	6.433	8.444	3.388	1.741
NR_146144.1	1279	C	T	-9.108	4.969	-4.447	18.195
NR_146144.1	2224	C	T	-4.040	3.492	-5.185	14.841
NR_146144.1	5926	C	T	-4.967	1.813	-2.324	0.744
NR_146144.1	7091	C	T	-15.383	1.729	-12.358	6.686
NR_146144.1	7821	C	T	1.473	5.779	0.473	17.039
NR_146144.1	8160	C	T	-7.459	2.029	-10.719	4.159
NR_146144.1	9193	C	T	-7.269	3.647	-7.009	3.293
NR_146144.1	10409	C	T	-3.625	3.449	-3.741	0.170

Appendix III: APOBEC3A candidate edit sites on nuclear RNAs

Chr	Position	Ref / Alts (strand)	Type	Gene ID	siNT Allele Frequency (Avg.)	siAPOBEC3A #1 Allele Frequency (Avg.)	siAPOBEC3A pool Allele Frequency (Avg.)	siNT/siAPOBEC3A (both) Ratio (Avg.)
2	240465606	C/T (+)	Non-coding RNA	GPC1	0.670	0.355	0.280	2.179
6	41965816	G/A (-)	Intron	CCND3	0.624	0.323	0.274	2.151
6	111374684	G/A (-)	Missense	REV3L	0.617	0.270	0.423	2.057
1	26700833	C/T (+)	Intron	ARID1A	0.529	0.295	0.380	2.028
2	44320435	G/A (-)	3-prime UTR	PREPL	0.896	0.502	0.565	1.991
10	97680760	G/A (-)	Intron	AVPI1	0.476	0.327	0.190	1.988
7	130891504	G/A (-)	Intron	MIR29A	0.571	0.361	0.254	1.975
6	150990351	C/T (+)	Intron	MTHFD1L	0.588	0.542	0.238	1.828
2	151799138	G/A (-)	3-prime UTR	ARL5A	0.688	0.525	0.317	1.820
9	94167287	C/T (+)	Intron	RP11-2B6.3	0.542	0.294	0.378	1.810
19	54460050	C/T (+)	3-prime UTR	LENG8	0.443	0.187	0.394	1.783
13	109194076	C/T (+)	Intron	MYO16	0.474	0.270	0.268	1.780
11	13461680	G/A (-)	Intron	BTBD10	0.528	0.394	0.338	1.773
2	69749292	C/T (+)	Intron	ANXA4	0.550	0.302	0.340	1.738
1	47377105	C/T (+)	3-prime UTR	CMPK1	0.668	0.306	0.600	1.719

10	97671727	C/T (+)	Intron	PI4K2A	0.512	0.353	0.417	1.712
1	183144984	C/T (+)	3-prime UTR	LAMC1	0.390	0.160	0.386	1.660
6	149889549	C/T (+)	Intron	RAET1E-AS1	0.416	0.283	0.252	1.654
19	24059854	C/T (+)	Intron	CTD-2017D11.1	0.800	0.511	0.611	1.653
5	132345364	G/A (-)	Intron	AC034220.3	0.572	0.272	0.531	1.649
10	27111932	G/A (-)	3-prime UTR	YME1L1	0.607	0.406	0.400	1.636
7	111417254	G/A (-)	Intron	IMMP2L	0.357	0.245	0.330	1.623
11	1836532	C/T (+)	Non-coding RNA	SYT8	0.428	0.331	0.267	1.610
14	35762820	G/A (-)	Intron	RALGAPA1	0.518	0.340	0.330	1.603
11	110125219	C/T (+)	Intron	ZC3H12C	0.612	0.478	0.538	1.591
22	15825736	C/T (+)	Intron	DUXAP8	0.638	0.442	0.410	1.578
16	68298844	G/A (-)	Non-coding RNA	SLC7A6OS	0.622	0.508	0.398	1.570
17	79940313	G/A (-)	3-prime UTR	TBC1D16	0.518	0.443	0.319	1.567
11	96298313	G/A (-)	Intron	MAML2	0.380	0.290	0.214	1.558
7	75985688	C/T (+)	Synonymous	POR	0.867	0.554	0.744	1.524
3	47083895	G/A (-)	Missense	SETD2	0.655	0.394	0.646	1.502
7	45017729	C/T (+)	Intron	CCM2	0.431	0.290	0.310	1.499
22	15817491	C/T (+)	Intron	DUXAP8	0.399	0.314	0.221	1.495

21	37424426	C/T (+)	Intron	DYRK1A	0.418	0.321	0.300	1.487
11	110124748	C/T (+)	Intron	ZC3H12C	0.625	0.384	0.518	1.486
14	24306812	C/T (+)	3-prime UTR	NOP9	0.639	0.581	0.345	1.486
9	137039568	G/A (-)	3-prime UTR	NPDC1	0.522	0.490	0.348	1.479
6	10730985	C/T (+)	Non-coding RNA	TMEM14C	0.604	0.487	0.365	1.477
20	44171741	G/A (-)	Intron	JPH2	0.566	0.423	0.329	1.475
21	44770495	G/A (-)	Non-coding RNA	UBE2G2	0.736	0.543	0.493	1.472
7	64991180	G/A (-)	Intron	ZNF117	0.864	0.587	0.636	1.461
11	130907991	G/A (-)	Synonymous	SNX19	0.686	0.418	0.563	1.461
19	23358866	G/A (-)	3-prime UTR	ZNF91	0.474	0.409	0.344	1.436
17	72417112	G/A (-)	Intron	LINC00511	0.307	0.225	0.209	1.434
5	6600037	G/A (-)	Non-coding RNA	NSUN2	0.608	0.474	0.437	1.434
6	29942973	C/T (+)	Intron	HLA-A	0.414	0.227	0.394	1.431
10	3129953	C/T (+)	Synonymous	PFKP	0.660	0.432	0.572	1.424
20	1326316	C/T (+)	Intron	SDCBP2-AS1	0.417	0.245	0.367	1.424
13	33271167	G/A (-)	Intron	STARD13	0.614	0.407	0.410	1.422
14	64210033	C/T (+)	Missense	SYNE2	0.557	0.344	0.463	1.418
7	130885731	G/A (-)	Intron	MIR29A	0.595	0.490	0.408	1.416

7	6392059	C/T (+)	Intron	RAC1	0.488	0.432	0.354	1.416
1	147178710	C/T (+)	Intron	RP11-337C18.8	0.880	0.711	0.623	1.411
4	40938914	G/A (-)	Intron	APBB2	0.516	0.445	0.349	1.405
11	119054629	G/A (-)	Non-coding RNA	HYOU1	0.694	0.471	0.611	1.399
11	20080877	C/T (+)	Intron	NAV2	0.260	0.248	0.157	1.396
8	11179525	G/A (-)	Intron	XKR6	0.523	0.491	0.375	1.394
11	118898175	G/A (-)	3-prime UTR	BCL9L	0.497	0.389	0.401	1.391
10	3129929	C/T (+)	Synonymous	PFKP	0.648	0.452	0.586	1.390
9	555772	C/T (+)	Intron	KANK1	0.297	0.275	0.228	1.388
12	66934292	G/A (-)	Intron	RP11-123O10.4	0.468	0.333	0.317	1.376
11	130210122	C/T (+)	3-prime UTR	ST14	0.641	0.464	0.488	1.374
15	64952588	C/T (+)	Intron	ANKDD1A	0.625	0.483	0.433	1.374
11	90210053	G/A (-)	Intron	CHORDC1	0.537	0.422	0.491	1.365
8	119832542	G/A (-)	Missense	TAF2	0.431	0.357	0.281	1.352
2	65268941	C/T (+)	3-prime UTR	ACTR2	0.636	0.457	0.481	1.349
14	104948292	G/A (-)	Non-coding RNA	AHNAK2	0.348	0.262	0.253	1.348
12	32473592	C/T (+)	Intron	FGD4	0.500	0.354	0.375	1.343
19	29952872	C/T (+)	Intron	URI1	0.891	0.681	0.682	1.337

2	101030287	G/A (-)	Intron	TBC1D8	0.763	0.577	0.567	1.334
2	68204127	G/A (-)	Intron	RP11-474G23.1	0.876	0.605	0.724	1.332
3	150417399	C/T (+)	Intron	TSC22D2	0.821	0.655	0.643	1.330
8	116846150	G/A (-)	3-prime UTR	RAD21	0.510	0.487	0.332	1.323
14	103682568	C/T (+)	Intron	RP11-73M18.2	0.382	0.256	0.350	1.320
6	159680813	G/A (-)	Intron	SOD2	0.357	0.312	0.284	1.318
7	130896461	G/A (-)	Intron	MIR29A	0.936	0.853	0.613	1.316
1	225042906	C/T (+)	Synonymous	DNAH14	0.630	0.463	0.489	1.311
X	3823028	G/A (-)	Intron	RP11-706O15.1	0.452	0.412	0.302	1.307
5	146182580	G/A (-)	5-prime UTR	LARS	0.825	0.626	0.647	1.299
2	85326957	G/A (-)	Intron	TGOLN2	0.470	0.342	0.423	1.298
15	83281134	G/A (-)	Intron	BNC1	0.920	0.687	0.775	1.296
2	172487375	C/T (+)	Synonymous	ITGA6	0.442	0.298	0.433	1.295
19	36236463	G/A (-)	Intron	ZNF565	0.683	0.624	0.568	1.294
12	124015335	C/T (+)	3-prime UTR	ZNF664	0.542	0.432	0.412	1.293
8	11176228	G/A (-)	Intron	XKR6	0.577	0.402	0.519	1.286
8	51845027	G/A (-)	Intron	PCMTD1	0.879	0.655	0.754	1.286

6	43784799	C/T (+)	Non-coding RNA	VEGFA	0.505	0.405	0.406	1.285
10	5693972	C/T (+)	Intron	FAM208B	0.317	0.227	0.295	1.285
8	102254199	G/A (-)	3-prime UTR	UBR5	0.677	0.636	0.498	1.282
6	142930032	G/A (-)	Intron	HIVEP2	0.774	0.610	0.608	1.275
14	104939605	G/A (-)	Non-coding RNA	AHNAK2	0.594	0.464	0.475	1.273
7	87159937	C/T (+)	Intron	DMTF1	0.854	0.723	0.673	1.272
2	38112662	C/T (+)	Intron	CYP1B1-AS1	0.477	0.358	0.470	1.272
7	66693723	C/T (+)	Intron	KCTD7	0.675	0.602	0.494	1.270
8	11130457	G/A (-)	Intron	XKR6	0.568	0.447	0.500	1.268
1	23961942	C/T (+)	3-prime UTR	PNRC2	0.518	0.381	0.485	1.264
11	33707713	G/A (-)	Intron	CD59	0.628	0.544	0.528	1.260
20	17614156	G/A (-)	3-prime UTR	RRBP1	0.348	0.269	0.281	1.258
13	113232079	C/T (+)	Intron	CUL4A	0.518	0.495	0.381	1.257
7	158733663	G/A (-)	3-prime UTR	ESYT2	0.697	0.663	0.488	1.254
6	30299971	G/A (-)	Intron	HCG18	0.912	0.721	0.815	1.254
5	93683578	G/A (-)	Intron	FAM172A	0.536	0.452	0.467	1.249
9	2828224	G/A (-)	Intron	PUM3	0.359	0.306	0.264	1.248
8	11179410	G/A (-)	Intron	XKR6	0.665	0.506	0.567	1.246
14	69054628	G/A (-)	Synonymous	DCAF5	0.496	0.420	0.394	1.246
12	65904399	C/T (+)	Intron	HMGA2	0.341	0.294	0.275	1.245

5	179833153	C/T (+)	Synonymous	SQSTM1	0.389	0.335	0.291	1.243
X	73826253	C/T (+)	Non-coding RNA	TSIX	0.165	0.133	0.149	1.235
10	12029398	G/A (-)	Synonymous	UPF2	0.474	0.421	0.351	1.235
1	112912477	G/A (-)	3-prime UTR	SLC16A1	0.520	0.455	0.403	1.234
22	25194628	C/T (+)	3-prime UTR	KIAA1671	0.538	0.451	0.457	1.234
18	12316780	C/T (+)	Intron	TUBB6	0.727	0.611	0.574	1.233
12	64716472	G/A (-)	3-prime UTR	GNS	0.695	0.605	0.559	1.233
7	130909810	G/A (-)	Intron	MIR29A	0.646	0.576	0.520	1.233
9	131475956	C/T (+)	Intron	PRRC2B	1.000	0.883	0.776	1.229
8	22620072	C/T (+)	3-prime UTR	CCAR2	0.752	0.650	0.584	1.224
3	196049583	G/A (-)	3-prime UTR	TFRC	0.574	0.489	0.468	1.216
9	4735079	G/A (-)	Intron	AK3	0.546	0.528	0.393	1.213
1	220207785	G/A (-)	Intron	RAB3GAP2	0.513	0.477	0.392	1.212
14	76197008	C/T (+)	Intron	GPATCH2L	0.786	0.726	0.589	1.212
7	116257830	C/T (+)	3-prime UTR	TES	0.669	0.538	0.609	1.210
22	44851174	C/T (+)	Intron	PRR5-ARHGAP8	0.451	0.352	0.428	1.208
15	98961208	C/T (+)	3-prime UTR	IGF1R	0.451	0.430	0.355	1.207

8	128101210	C/T (+)	Non-coding RNA	PVT1	0.599	0.462	0.568	1.206
4	82835511	G/A (-)	Intron	SEC31A	0.800	0.699	0.638	1.204
1	226991247	G/A (-)	3-prime UTR	CDC42BPA	0.850	0.706	0.785	1.200
12	96257481	C/T (+)	Intron	ELK3	0.867	0.737	0.764	1.199
6	142939463	G/A (-)	Intron	HIVEP2	0.720	0.640	0.568	1.197
20	2965113	C/T (+)	Missense	PTPRA	0.614	0.576	0.570	1.195
12	105184332	G/A (-)	Intron	APPL2	0.707	0.614	0.598	1.195
6	36681197	C/T (+)	Intron	CDKN1A	0.766	0.673	0.646	1.192
19	38701093	C/T (+)	Intron	ACTN4	0.416	0.403	0.309	1.192
7	130901314	G/A (-)	Intron	MIR29A	0.965	0.879	0.761	1.190
2	180990735	C/T (+)	Intron	UBE2E3	0.700	0.609	0.577	1.187
8	86558500	C/T (+)	Non-coding RNA	CPNE3	0.691	0.592	0.575	1.185
5	178210881	G/A (-)	Intron	PHYKPL	0.323	0.269	0.293	1.184
16	83678339	C/T (+)	Synonymous	CDH13	1.000	0.860	0.864	1.182
9	81649349	G/A (-)	Intron	TLE1	0.864	0.745	0.733	1.180
6	149410768	C/T (+)	3-prime UTR	TAB2	0.636	0.589	0.583	1.179
1	1055393	C/T (+)	3-prime UTR	AGRN	0.532	0.446	0.475	1.177
1	153645273	C/T (+)	Non-coding RNA	CHTOP	0.731	0.629	0.618	1.176
1	147179337	C/T (+)	Intron	RP11-337C18.8	0.586	0.516	0.486	1.176

7	842031	C/T (+)	Non-coding RNA	SUN1	0.706	0.623	0.581	1.175
3	190306399	G/A (-)	3-prime UTR	CLDN1	0.656	0.511	0.629	1.174
4	57023511	C/T (+)	Non-coding RNA	POLR2B	0.462	0.370	0.410	1.174
16	28111905	G/A (-)	Synonymous	XPO6	0.758	0.633	0.712	1.173
7	158743643	G/A (-)	Non-coding RNA	ESYT2	0.578	0.561	0.458	1.171
13	113232108	C/T (+)	Intron	CUL4A	0.514	0.427	0.454	1.170
8	118072678	G/A (-)	Intron	EXT1	0.290	0.235	0.247	1.169
10	5695422	C/T (+)	Intron	FAM208B	0.728	0.636	0.613	1.168
12	132808278	G/A (-)	Non-coding RNA	GOLGA3	0.606	0.563	0.523	1.166
8	17232060	G/A (-)	Intron	CNOT7	0.741	0.700	0.625	1.163
14	104950766	G/A (-)	Non-coding RNA	AHNAK2	0.401	0.380	0.326	1.162
13	33279001	G/A (-)	Intron	STARD13	0.859	0.779	0.705	1.162
2	231708591	C/T (+)	Non-coding RNA	PTMA	0.506	0.489	0.400	1.161
22	32862230	G/A (-)	Intron	SYN3	0.571	0.442	0.565	1.160
14	34773622	G/A (-)	Synonymous	BAZ1A	0.638	0.531	0.612	1.157
3	155987978	C/T (+)	Non-coding RNA	SETP14	1.000	0.786	0.962	1.157
4	186588166	G/A (-)	3-prime UTR	FAT1	0.632	0.612	0.523	1.153
11	1010876	C/T (+)	Non-coding RNA	AP2A2	0.472	0.444	0.397	1.151

7	158732475	G/A (-)	3-prime UTR	ESYT2	0.696	0.581	0.638	1.150
18	12318643	C/T (+)	Intron	TUBB6	0.657	0.641	0.517	1.150
6	43785314	C/T (+)	Non-coding RNA	VEGFA	0.587	0.541	0.554	1.148
17	76564478	C/T (+)	Non-coding RNA	SNHG16	0.426	0.394	0.351	1.145
Y	5338698	C/T (+)	Intron	PCDH11Y	0.632	0.525	0.618	1.145
7	87160110	C/T (+)	Intron	DMTF1	0.794	0.773	0.684	1.145
18	12315676	C/T (+)	Intron	TUBB6	0.832	0.731	0.728	1.144
1	211974403	G/A (-)	Intron	INTS7	0.802	0.656	0.758	1.142
12	53479582	C/T (+)	Non-coding RNA	PCBP2	0.441	0.422	0.382	1.141
7	77083311	C/T (+)	Intron	RP11-467H10.2	0.481	0.415	0.438	1.140
11	3364238	G/A (-)	Intron	ZNF195	0.927	0.778	0.855	1.139
3	121694790	G/A (-)	Synonymous	GOLGB1	0.633	0.579	0.541	1.138
5	149852650	C/T (+)	3-prime UTR	PPARGC1B	0.771	0.621	0.742	1.136
16	227459	G/A (-)	Non-coding RNA	LUC7L	0.539	0.511	0.477	1.130
16	14305936	C/T (+)	Intron	MIR193BHG	1.000	0.912	0.864	1.129
10	1693349	G/A (-)	Intron	ADARB2	0.782	0.632	0.774	1.127
11	985547	C/T (+)	Synonymous	AP2A2	1.000	0.978	0.842	1.127
16	9105995	C/T (+)	Non-coding RNA	RP11-473I1.9	0.922	0.750	0.917	1.126

11	70407369	C/T (+)	Non-coding RNA	CTTN	0.613	0.539	0.560	1.123
12	123326326	G/A (-)	Synonymous	SBNO1	1.000	0.929	0.862	1.123
11	130889029	G/A (-)	Intron	SNX19	0.464	0.445	0.388	1.122
7	101090992	C/T (+)	3-prime UTR	TRIM56	0.458	0.399	0.433	1.121
20	17614084	G/A (-)	3-prime UTR	RRBP1	0.709	0.590	0.699	1.116
7	105373482	G/A (-)	Intron	SRPK2	0.631	0.572	0.560	1.115
6	142942004	G/A (-)	Intron	HIVEP2	0.717	0.605	0.679	1.114
5	32442389	G/A (-)	Intron	ZFR	0.772	0.758	0.646	1.113
11	18269312	C/T (+)	Missense	SAA1	0.651	0.542	0.645	1.112
7	158731198	G/A (-)	3-prime UTR	ESYT2	0.719	0.639	0.687	1.110
6	41984820	G/A (-)	Intron	CCND3	0.881	0.809	0.798	1.105
17	16441561	C/T (+)	Non-coding RNA	LRRC75A-AS1	0.490	0.421	0.468	1.103
15	63071432	C/T (+)	3-prime UTR	TPM1	0.645	0.604	0.633	1.101
14	21392201	G/A (-)	Intron	CHD8	0.457	0.416	0.428	1.101
3	195692672	C/T (+)	Intron	LINC00969	0.631	0.599	0.545	1.099
17	44352876	C/T (+)	3-prime UTR	GRN	0.499	0.486	0.440	1.097
8	51845101	G/A (-)	Intron	PCMTD1	0.858	0.772	0.803	1.097
8	128016010	C/T (+)	Intron	PVT1	0.815	0.727	0.764	1.094
1	204221288	G/A (-)	3-prime UTR	PLEKHA6	0.686	0.636	0.635	1.092

7	64992758	G/A (-)	Intron	ZNF117	1.000	0.917	0.929	1.092
15	60391750	G/A (-)	Intron	ANXA2	0.628	0.543	0.605	1.091
11	130196440	C/T (+)	Synonymous	ST14	1.000	0.888	0.958	1.091
8	51838246	G/A (-)	Intron	PCMTD1	0.859	0.780	0.824	1.088
7	157368878	C/T (+)	Intron	DNAJB6	0.928	0.906	0.813	1.088
9	4739151	G/A (-)	Intron	AK3	0.905	0.893	0.782	1.087
22	38483490	G/A (-)	3-prime UTR	DDX17	0.578	0.501	0.568	1.087
15	85580693	C/T (+)	Synonymous	AKAP13	1.000	0.986	0.880	1.086
5	139283535	C/T (+)	Intron	MATR3	0.317	0.292	0.302	1.086
5	177542624	G/A (-)	Intron	FAM193B	0.895	0.808	0.849	1.085
6	159680891	G/A (-)	Intron	SOD2	0.446	0.415	0.407	1.084
19	11116926	C/T (+)	Missense	LDLR	0.519	0.505	0.457	1.084
4	186617176	G/A (-)	Synonymous	FAT1	0.539	0.514	0.502	1.079
11	62523404	G/A (-)	Synonymous	AHNAK	0.625	0.595	0.569	1.078
17	7514346	C/T (+)	3-prime UTR	POLR2A	1.000	0.929	0.938	1.077
15	85579788	C/T (+)	Missense	AKAP13	1.000	0.981	0.897	1.075
8	127995166	C/T (+)	Intron	PVT1	0.689	0.653	0.635	1.073
21	42976841	C/T (+)	Intron	PKNOX1	0.960	0.875	0.923	1.073
9	33625227	C/T (+)	Non-coding RNA	ANXA2P2	0.913	0.897	0.824	1.069

11	73677454	G/A (-)	3-prime UTR	RAB6A	0.531	0.479	0.522	1.069
3	197614675	G/A (-)	Intron	AC024560.3	0.506	0.463	0.482	1.067
5	107523783	G/A (-)	Intron	EFNA5	0.280	0.268	0.278	1.065
7	158732370	G/A (-)	3-prime UTR	ESYT2	0.705	0.637	0.701	1.060
X	148052736	C/T (+)	Non-coding RNA	FTH1P8	0.906	0.864	0.870	1.058
1	112913787	G/A (-)	3-prime UTR	SLC16A1	0.465	0.423	0.461	1.057
1	153032377	C/T (+)	Missense	SPRR1B	0.960	0.900	0.917	1.057
20	35740352	G/A (-)	Non-coding RNA	RBM39	0.771	0.698	0.768	1.055
X	74276052	G/A (-)	Intron	FTX	0.971	0.938	0.923	1.052
6	41981005	G/A (-)	Intron	CCND3	0.702	0.664	0.683	1.051
6	42023999	G/A (-)	Intron	CCND3	0.880	0.864	0.812	1.050
15	75940641	C/T (+)	3-prime UTR	FBXO22	1.000	0.972	0.941	1.048
14	103683232	C/T (+)	Intron	RP11-73M18.2	0.756	0.694	0.754	1.045
17	50089616	C/T (+)	Non-coding RNA	ITGA3	0.486	0.458	0.479	1.045
15	26645502	G/A (-)	Intron	GABRB3	0.960	0.938	0.900	1.045
1	155009232	C/T (+)	Intron	ZBTB7B	0.657	0.620	0.654	1.042
7	64978289	G/A (-)	Nonsense	ZNF117	1.000	0.952	0.978	1.038
6	108680839	C/T (+)	3-prime UTR	FOXO3	1.000	0.985	0.946	1.038

2	121231913	G/A (-)	Synonymous	TFCP2L1	1.000	0.976	0.954	1.036
1	114750181	G/A (-)	5-prime UTR	CSDE1	0.504	0.462	0.500	1.036
12	6598360	G/A (-)	Synonymous	CHD4	1.000	0.966	0.968	1.036
5	53051539	C/T (+)	Synonymous	ITGA2	1.000	0.955	0.981	1.035
1	180197955	C/T (+)	3-prime UTR	QSOX1	0.736	0.721	0.732	1.035
1	86447582	C/T (+)	Synonymous	CLCA2	0.543	0.522	0.527	1.035
1	224387671	G/A (-)	3-prime UTR	WDR26	1.000	0.955	0.982	1.034
5	151563408	G/A (-)	Missense	FAT2	0.947	0.904	0.932	1.033
12	6599893	G/A (-)	Synonymous	CHD4	1.000	0.978	0.960	1.033
17	28468211	C/T (+)	Non-coding RNA	RPS7P1	1.000	0.952	0.989	1.032
2	238855353	C/T (+)	Intron	TWIST2	0.806	0.783	0.792	1.031
16	31183958	C/T (+)	Non-coding RNA	FUS	1.000	0.977	0.967	1.030
1	36286832	C/T (+)	Missense	THRAP3	1.000	0.967	0.977	1.029
5	177546150	G/A (-)	Intron	FAM193B	0.859	0.832	0.846	1.024
1	201969867	C/T (+)	Non-coding RNA	TIMM17A	1.000	0.975	0.980	1.024
9	128598471	C/T (+)	Non-coding RNA	SPTAN1	0.988	0.968	0.965	1.021
1	21864961	G/A (-)	Missense	HSPG2	1.000	0.987	0.974	1.021

2	218274850	G/A (-)	Non-coding RNA	TMBIM1	1.000	0.986	0.975	1.021
12	68846494	C/T (+)	3-prime UTR	MDM2	1.000	0.979	0.983	1.019
20	30668968	G/A (-)	Intron	ANKRD20A21P	1.000	0.979	0.988	1.017
5	179699089	C/T (+)	Intron	CANX	0.949	0.931	0.942	1.014
11	20080158	C/T (+)	Synonymous	NAV2	1.000	0.986	0.989	1.013
5	134925215	G/A (-)	Non-coding RNA	MTND5P11	0.318	0.305	0.306	1.012
6	42013133	G/A (-)	Intron	CCND3	0.855	0.841	0.850	1.012
9	128617688	C/T (+)	Synonymous	SPTAN1	1.000	0.988	0.990	1.012
X	74278613	G/A (-)	Non-coding RNA	FTX	0.975	0.963	0.969	1.012
20	30668988	G/A (-)	Intron	ANKRD20A21P	1.000	0.989	0.991	1.011
13	24446725	G/A (-)	Missense	PARP4	0.989	0.972	0.987	1.011
11	18402904	C/T (+)	Non-coding RNA	LDHA	1.000	0.986	0.995	1.010
X	74275170	G/A (-)	Intron	FTX	0.970	0.958	0.968	1.008