Abstract

Discovery of new regulators of human ribosome biogenesis

Katherine Irene Farley-Barnes

2019

Ribosomes are macromolecular machines that perform the essential function of protein synthesis in human cells. The process of making a ribosome is incredibly intricate and requires all 3 RNA polymerases, a host of assembly factors, and approximately 80 different ribosomal proteins (r-proteins). Ribosome biogenesis begins in the nucleolus, proceeding outward towards the cytoplasm where the ribosomes function in protein synthesis. Despite the importance of this biological process in cellular growth and development, many questions remain regarding the synthesis of ribosomes. Still less is known about this process in human cells, as the foundational knowledge of ribosome biogenesis was developed in the budding yeast *Saccharomyces cerevisiae*. This work therefore seeks to answer key questions about the factors regulating the synthesis of ribosomes in human cells.

To identify new regulators of ribosome biogenesis in human cells, a high-throughput, genome-wide RNAi screen was performed. The screen took advantage of the fact that the structure of the nucleolus is intricately linked with its function in making ribosomes. After 72 hours of protein depletion, reductions in the number of nucleoli per cell nucleus were observed, and this readout of reduced nucleolar number was used as a proxy for changes in nucleolar function. In all, the screen identified 139 proteins required for nucleolar formation and function in human cells. Further testing of 20 high-confidence screen hits revealed functions for 18 in the nucleolar processes of rDNA transcription (7/20), pre-ribosomal RNA processing (16/20), and global protein synthesis (14/20). The protein hits from the screen function in a plethora of cellular processes including cell division, development, proliferation, and more. The screen will therefore continue to serve as a rich resource to increase our understanding of the essential biological process of making ribosomes in humans.

Additional studies of one screen hit, Paired Box 9 (PAX9), reveled a new role for this protein in human ribosome biogenesis. In human cells, depletion of PAX9 results in decreased nucleolar number as well as defects in small subunit (SSU) pre-ribosomal RNA processing and global protein synthesis. Two transcriptomics analyses, RNA-sequencing and RNA Polymerase II (RNAPII) Chromatin immunoprecipitation sequencing (ChIP-seq), revealed that PAX9 acts as a RNAPII transcription factor to drive the expression of a number of mRNAs that encode proteins critical for making ribosomes. In addition, the role of PAX9 in human ribosome biogenesis is conserved to a model organism, the African clawed frog, *Xenopus tropicalis* (*X. tropicalis*). This function for PAX9 in making ribosomes connects its role as an RNAPII transcription factor to both ribosome biogenesis and craniofacial development.

Discovery of new regulators of human ribosome biogenesis

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

by
Katherine Irene Farley-Barnes

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

May 2019

ACKNOWLEDGEMENTS

I am incredibly grateful to a host of people who have supported me throughout my time in graduate school. First, I would like to thank my advisor, Susan J. Baserga, M.D., Ph.D. for sharing her knowledge with me and helping to guide me on my journey to becoming a better scientist. I would also like to thank the members of my thesis committee, Dr. Joan Steitz and Dr. Daniel DiMaio, who have continued to be an invaluable source of knowledge and ideas.

I would also like to thank the members of the Baserga laboratory, both past and present, who have helped me along my journey: Kat McCann, Vincent Yip, Nicholas Vincent, Samuel Sondalle, Lisa Ogawa Mclean, Mason McCool, Carson Bryant, Amber Buhagiar, Sarah Ludwin-Peery, and Maya Overton. I am especially indebted to Kat McCann, who performed the original siRNA screen, without which I would not have had this thesis project. All the members of the Baserga laboratory have filled my days with joy, supporting me in bad times and laughing with me in good times.

In addition, many thanks go out to my family and friends. Without their love and encouragement, I would not be the person I am today. I would especially like to thank my husband, Bruce Barnes, for being my greatest friend and champion.

Finally, this work is dedicated to my daughter, Eleanore Amelia Barnes. May it encourage you to always ask questions, learn, and grow.

Table of Contents

Note that some of the contents of this work have been adapted from (Farley et al., 2015, Farley and Baserga, 2016, Farley-Barnes et al., 2018), a manuscript in press: Farley-Barnes, Katherine I. and Baserga, Susan J. Ribosome biogenesis and its role in cell growth and proliferation in the liver. The Liver: Biology and Pathobiology. 6 ed., *and a manuscript in preparation.*

List of Figures and Tables Page(s)

Figures for Chapter 1

Tables for Chapter 1

Figures for Chapter 2

Tables for Chapter 2

Figures for Chapter 3

Tables for Chapter 3

Figures for Chapter 4

Tables for Chapter 4

Figures for Chapter 5

Chapter 1

Human ribosome biogenesis and craniofacial development

The nucleolus and ribosome biogenesis in humans

Ribosomes are the cellular machines responsible for the synthesis of all proteins in the human cell. Thus, they are essential for organismal growth and development. The process of making a ribosome, called ribosome biogenesis, is highly energy intensive; approximately 60% of all cellular transcription is involved in ribosome creation (Woolford and Baserga, 2013). This process is also very complex; it involves all 3 RNA polymerases, 79 ribosomal proteins (r-proteins), over 200 assembly factors, and a host of small nucleolar ribonucleoproteins (snoRNPs). Making a ribosome takes place in multiple cellular compartments, beginning in the nucleolus, moving to the nucleoplasm, and concluding in the cytoplasm. Ribosome biogenesis also needs to happen quickly; to produce the 200,000 ribosomes that a rapidly growing yeast cell requires, the cell must create 2,000 new ribosomes every minute (Warner, 1999). Additionally, the cell must coordinate the production of ribosomes with cellular demand by responding to a number of external stimuli, such as nutrient availability. Understanding the intricacies of this energy intensive, complex, and fast process is essential to understanding the basic cellular processes of growth and development.

The process of making a ribosome begins at nucleolar organizing regions (NORs) with the transcription of the tandemly repeated ribosomal DNA (rDNA) (Figure 1-1). The rDNA repeats are located on the 5 acrocentric chromosomes (13, 14, 15, 21, and 22) in humans. In humans, the focus of this manuscript, each rDNA repeat is approximately 43 kb long. Approximately 30 kb of each rDNA repeat comprises an intergenic spacer sequence while 13-14 kb of this sequence specifies the 47S polycistronic pre-ribosomal RNA (pre-rRNA). The 47S pre-rRNA contains the sequences for the 18S rRNA included in the small subunit (SSU) of the ribosome as well as the 5.8S and 28S rRNAs that are included in the large subunit (LSU) of the ribosome (Figure 1-1) (Gonzalez and Sylvester, 1995).

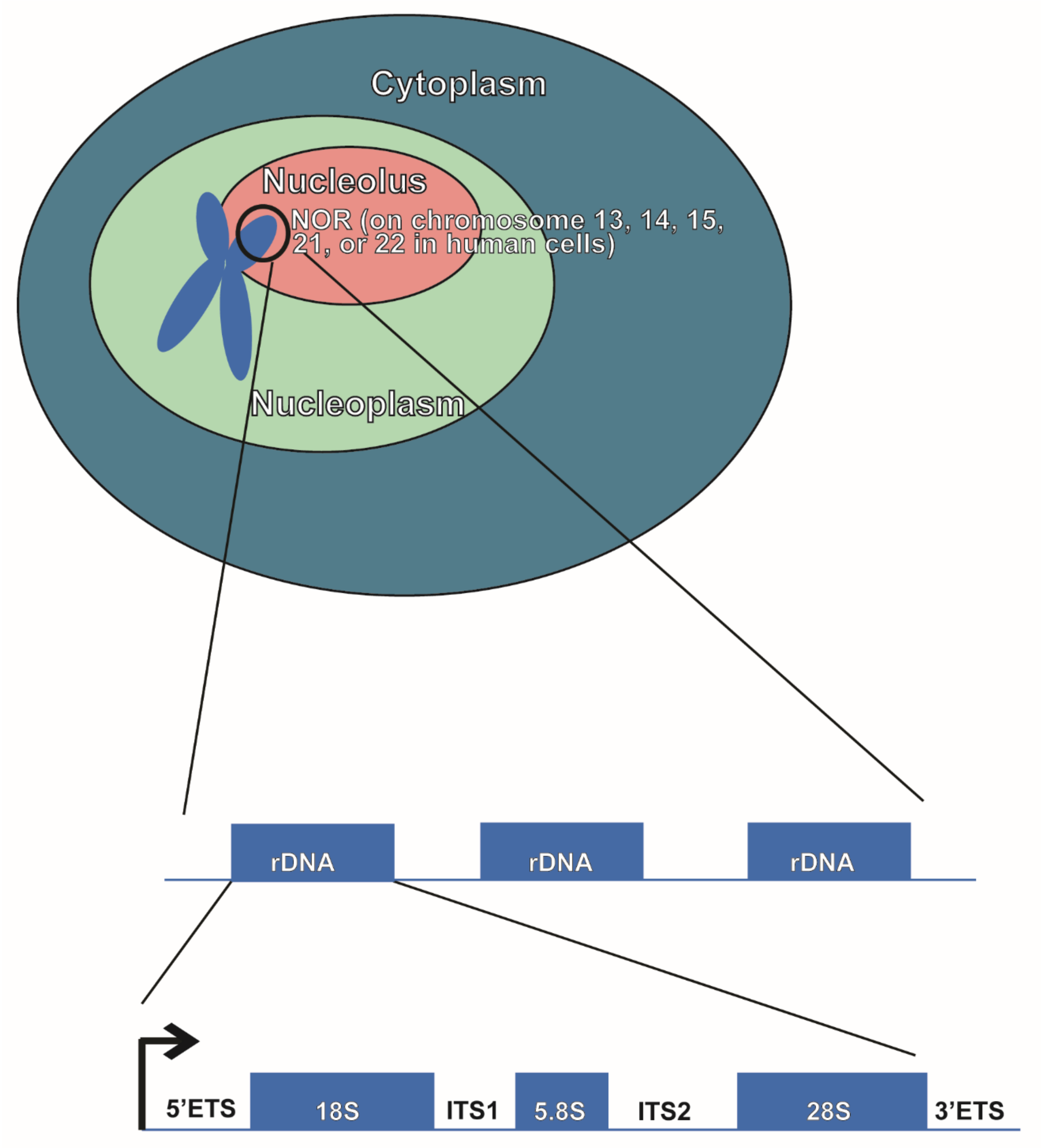

Figure 1-1. Ribosome biogenesis in human cells begins at nucleolar organizing regions (NORs), which are located on the short arms (p arms) of the 5 acrocentric chromosomes. NORs consist of tandemly repeated ribosomal DNA (rDNA) sequences. These rDNA sequences consist of a 47S polycistronic sequence that contains the 18S, 5.8S, and 28S ribosomal RNAs (rRNAs), as well as internal (ITS) and external (ETS) transcribed spacer sequences. Figure adapted from (Farley et al., 2015).

This thesis will focus on nucleolar function and formation in human cells. These cells contain approximately 400 rDNA repeats, and only about half of which are transcriptionally active (Schlesinger et al., 2009). Multiple mechanisms govern the number of active vs inactive repeats including nucleosome position, chromatin remodeling complexes (such as the activating Cockayne Syndrome protein B and NoRC repressive remodeling complex), methylation, and more [reviewed in (McStay and Grummt, 2008); (Guetg and Santoro, 2012)]. In addition to changes in the number of active vs inactive rDNA repeats per cell, the number of rDNA copies can also change greatly from person to person, and variations in this copy number have been implicated in human disease (Gibbons et al., 2014, Xu et al., 2017, Wang and Lemos, 2017). Therefore, human cells use these alterations in rDNA copy number as one mechanism for regulating ribosome biogenesis.

A second mechanism of regulation is through transcription of the rDNA by RNA Polymerase I (RNAPI). RNAPI transcription is initiated through the binding of upstream binding transcription factor (UBTF, commonly referred to as UBF) to the upstream control element of the pre-rRNA (Bell et al., 1988). UBTF then helps to recruit selectivity factor 1 (SL1), a species-specific complex that includes the TATA binding protein (TBP) and TBP associated factors (TAFs). RNAPI binds to SL1, with the help of tripartite motif containing 24 (TRIM24, also known as TIF1A), to form the pre-initiation complex (Miller et al., 2001). rDNA transcription is known to be regulated on the cellular level through a number of pathways, including mechanistic target of rapamycin (mTOR) signaling and MAPK signaling. Furthermore, rDNA transcriptional activity varies according to cell type and may change with stages of development (Haaf et al., 1991). RNAPI transcription is therefore tightly regulated at multiple points to ensure generation of an adequate number of ribosomes for cell growth or maintenance, and it is likely that there remain many more unidentified signaling pathways that regulate rDNA transcription.

The pre-rRNA is also extensively modified in coordination with rDNA transcription. These pre-rRNA modifications aid in the folding of the pre-rRNA into the correct secondary and tertiary structures. Two types of RNA modifications are primarily used: 2’-O-methylation and pseudouridylation [reviewed in (Sloan et al., 2017)]. Most of these modifications are made by snoRNPs which direct their guide RNA to base pair with the corresponding site to be modified and position the enzyme to perform the modification. Pre-rRNA modifications are generally carried out by two types of snoRNPs: box C/D, which perform the 2’-O-methylations, and box H/ACA, which perform the pseudouridylations. Both snoRNPs are named for conserved snoRNA sequences. Many of the modifications are placed at functionally important regions of the ribosome, such as the transfer RNA (tRNA) binding sites or at the interface of the SSU and LSU (Sloan et al., 2017).

In addition to being modified, the pre-rRNA must also be cleaved and processed to form the mature 18S, 5.8S, and 28S rRNAs (Figure 1-2). In humans, pre-rRNA processing occurs via multiple pathways [reviewed in (Henras et al., 2015, Aubert et al., 2018)] (Figure 1-2). The two main pathways diverge with cleavages in either internal transcribed spacer 1 (ITS1) or in the 5’ external transcribed spacer (5’ ETS). Cleavage at site 2 in ITS1 creates the 30S pre-rRNA, while cleavages in 5’ ETS first create the 41S intermediate, bypassing the 30S. Regardless of the pathway used, the end result is the mature 18S, 5.8S, and 28S rRNAs.

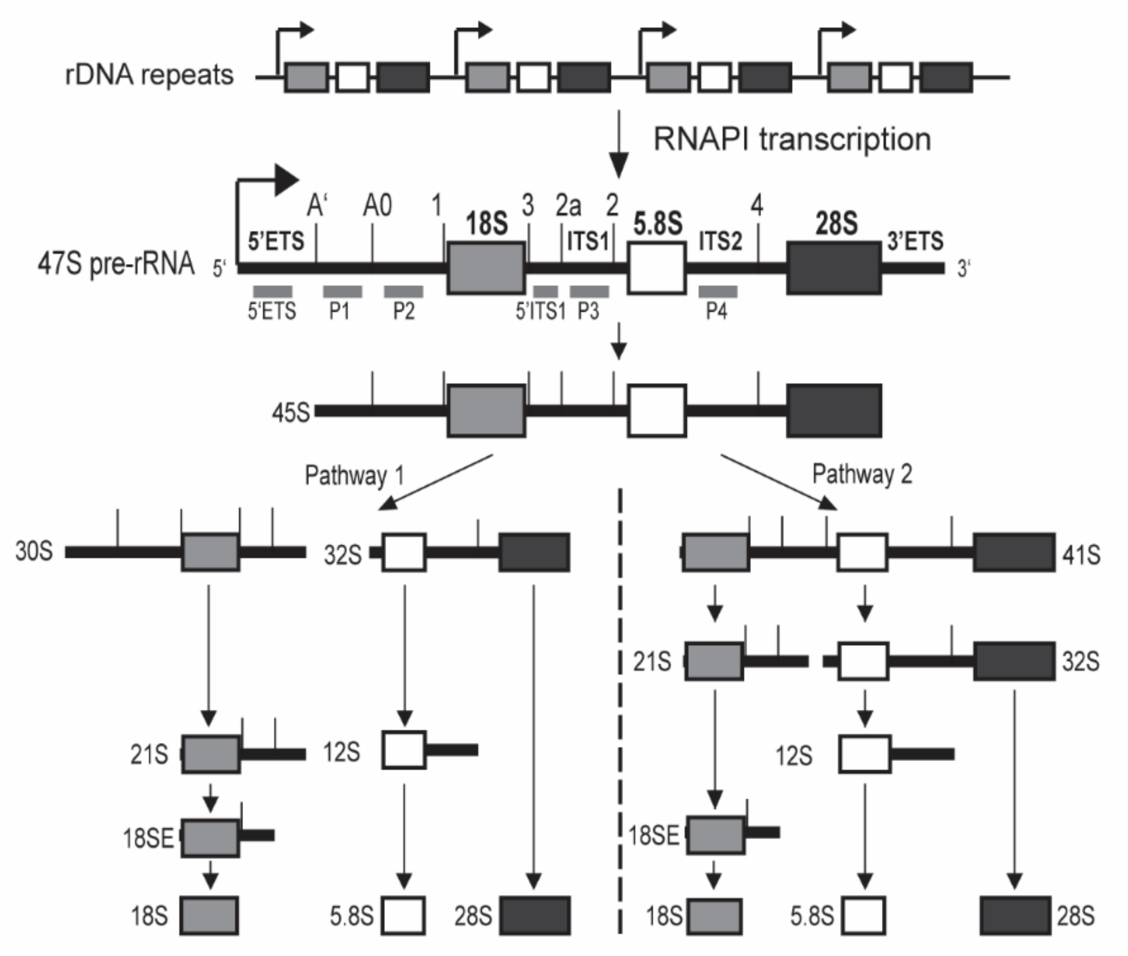

Figure 1-2. Pre-rRNA processing in human cells occurs via multiple pathways. The repeated rDNA is first transcribed by RNAPI into the 47S pre-rRNA. Then, the 47S pre-rRNA is further processed through one of two major pathways into the mature rRNAs that are incorporated into the small (18S) and large (5.8S, 28S) ribosomal subunits. Probes for northern blots (5'ETS, P1, P2, P3, 5'ITS1, and P4) are shown in grey below the 47S pre-rRNA. Adapted from (Farley-Barnes et al., 2018).

The final steps of ribosome maturation in humans occur outside of the nucleolus. The 5S rRNA is assembled into the pre-LSU (pre-60S) particle in the nucleus after being transcribed from chromosome 1 by RNA Polymerase III (RNAPIII) (Ciganda and Williams, 2011). The pre-SSU and pre-LSU are then both exported through the nuclear pore complex. Export of each subunit requires common factors as well as subunit-specific factors. For example, RIO2 aids in the export of the SSU (Zemp et al., 2009). In contrast, Exportin 1 (XPO1, or Crm1 in yeast) is one of the best characterized proteins involved in both LSU and SSU export (Zemp and Kutay, 2007). After export, the final steps of pre-rRNA processing, including the processing of the 18SE to the mature 18S rRNA at site 3, occur in the cytoplasm. There is some evidence for feedback between the final steps of pre-rRNA processing and the start of translation, possibly mediated by the protein Receptor for Activated C Kinase 1 [RACK1, (Larburu et al., 2016)]. Once ribosome biogenesis is completed in the cytoplasm, the ribosomes are able to perform their function of global protein synthesis (Schuller and Green, 2018).

The process of building a ribosome involves hundreds of factors. In addition to the snoRNAs and r-proteins, a number of trans-acting factors are also important for optimal pre-rRNA processing. Best characterized in yeast, these factors include endo- and exo-nucleases, helicases, AAA-ATPases, GTPases, and more (Woolford and Baserga, 2013). Many of the functions of these enzymes are still being defined. In yeast, over 200 proteins are involved in ribosome assembly (Woolford and Baserga, 2013), a number that is likely to be greatly increased in humans due to the increased complexity of this process. Indeed, human cells contain an expanded pre-rRNA transcript sequence with additional cleavage sites, as well as the ability to form multiple nucleoli per cell nucleus. A series of studies, including this work, are therefore striving to define all of the factors necessary for this process in humans (Farley-Barnes et al., 2018, Badertscher et al., 2015, Wild et al., 2010, Tafforeau et al., 2013).

Mechanisms dictating nucleolar formation in humans

The nucleolus is a non-membrane bound, nuclear organelle whose formation is linked with the production of functional ribosomes. First described as “un corps oviform,” this non-membrane-bound organelle is located in the nucleus, at the nucleolar organizing regions [NORs, (Fontana and Nyon, 1781, McClintock, 1934)]. When viewed by light microscopy, nucleoli feature prominently in the cell nucleus. Additionally, silver nitrate preferentially stains a group of argyrophilic proteins which localize at transcriptionally active NORs, allowing for visualization of nucleoli in cyto-histopathological samples [(Goodpasture and Bloom, 1975)]. During metaphase, these NORs which had been active in the preceding interphase appear as achromatic gaps, termed secondary constrictions, when stained with DAPI [4’,6-diamidino-2-phenylindole; (Sumner, 1982)]. The nucleoli of higher eukaryotes are comprised of three distinct subcompartments: the fibrillar center (FC), the dense fibrillar component (DFC), and the granular component (GC). These three compartments are also visible by light microscopy (Scheer and Weisenberger, 1994).

The nucleolar ultrastructure is likely a product of the functions it performs: ribosome biogenesis. Ribosome biogenesis begins in the DFC and proceeds vectorially outward towards the cytoplasm. rDNA transcription takes place at the DFC/FC interface (Cheutin et al., 2002, Koberna et al., 2002). The process of assembling the ribosome then proceeds towards the GC and eventually into the nucleus and cytoplasm. Current evidence therefore suggests that ribosome biogenesis occurs directionally away from the fibrillar center (Raska et al., 2006), and the nucleolar subcompartments are formed from the process of building a ribosome.

In addition to the three compartments, an important piece of nucleolar structure is the surrounding heterochromatin/euchromatin found both at the nucleolar periphery (perinucleolar) and penetrating into the nucleoli (intranucleolar) [reviewed in (Schöfer and Weipoltshammer, 2018)]. The perinucleolar heterochromatin is believed to hold inactive rDNA repeats (Schöfer and Weipoltshammer, 2018). Additionally, the sex chromosomes are often located adjacent to nucleoli (van Koningsbruggen et al., 2010, Barr and Bertram, 1949, Dillinger et al., 2017, Gates, 1939). Similarly, nucleolar-associated (chromatin) domains (NADs) are an active area of study. These relatively gene-poor chromatin regions include genes for olfactory receptors, zinc-finger proteins, and immunoglobulins (van Koningsbruggen et al., 2010, Németh et al., 2010). Additionally, the 5S rDNA is likely located in the euchromatic region near the nucleolus but just outside of the perinucleolar compartment (Yu and Lemos, 2018). The structure of the chromatin that surrounds the nucleolus influences nucleolar formation and thus nucleolar function. Indeed, it has been postulated that the surrounding chromatin structures of the nucleoplasm are able to sterically hinder the movement of the forming SSU (Landvogt et al., 2019). However, much remains to be learned regarding how the direct interactions between the nucleolus and surrounding chromatin influence ribosome production.

In contrast to the theory that the nucleolar ultrastructure is formed as a result of its functions in ribosome biogenesis, another theory states that the fluid, liquid-like behavior of the nucleolus is adequate to determine its size and shape. Just like an oil droplet in water, the proteins of the nucleolus separate into multiple compartments due to differences in surface tension and hydrophobicity. In support of this, Brangwynne et al. used germinal vesicles, the nuclei of amphibian oocytes, to demonstrate that the nucleolus behaves like a liquid droplet on a timescale of tens of seconds (Brangwynne et al., 2011). The surface tension of the droplet is thus responsible for the spherical shape of the nucleolus. Brangwynne et al. also observed the fusion of multiple nucleoli into a single, large nucleolus, demonstrating that volumes of nucleoli follow a power-law distribution characteristic of aggregation processes (Brangwynne et al., 2011). As additional support of the phase separation theory of nucleolar formation, Handwerger et al. used *Xenopus* GV nucleoli to suggest a “sponge model” in which the size of the molecule and the density of the subcompartment determine the movement of proteins in and out of the subcompartment (Handwerger et al., 2005). Whether or not these findings translate from *Xenopus* GV nucleoli to human nucleoli remains to be seen, however, as *Xenopus* GV nucleoli are extrachromosomal and are more numerous than human nucleoli (Wu and Gall, 1997). It should also be noted that the above *Xenopus* GV experiments were conducted using isolated nucleoli in mineral oil. The influence of such preparations on the behavior of nucleoli has not yet been determined. In more recent work, a structure similar to the DFC/GC compartments has been replicated *in vitro* by mixing fibrillarin (FBL), a box C/D small nucleolar ribonucleoprotein (snoRNP) methyltransferase that is enriched in the DFC, and nucleophosmin (NPM1), a protein with a number of functions that is enriched in the GC (Feric et al., 2016). While this droplet-like structure may look similar to the nucleolus, it is lacking the rDNA and transcription machinery required for nucleolar function as described above (Grob et al., 2014). While the mechanisms dictating nucleolar formation remain unknown, it can be concluded that the nucleolus is comprised of a specific ultrastructure that directly relates to its function in ribosome biogenesis.

There are 3 requirements for nucleolar formation: 1) at least 1 rDNA repeat, 2) recruitment of the RNAPI machinery, and 3) active RNAPI transcription. Early studies suggested that the expression of a single rDNA repeat is sufficient for nucleolar formation (Karpen et al., 1988). However, the expression system used in those early *Drosophila* experiments may have biased the results towards excessive RNAPI recruitment, because laterally amplified polytene chromosomes were used rather than a single isolated rDNA repeat (Oakes et al., 2006). More recently, the McStay laboratory has furthered our understanding of the essential components for nucleolar formation in humans through the design of synthetic nucleoli. Mais et al. were able to create “pseudo-NORs” through the insertion of UBF binding sequences called *Xenopus* Enhancer elements (XEns), normally present in the intergenic spacers of *Xenopus* rDNA, into the DNA of a human fibrosarcoma cell line (Mais et al., 2005). Visually, the “pseudo-NORs” appear the same as active NORs because they form secondary constrictions which silver stain (Mais et al., 2005). Despite apparently having the correct NOR structure, these “pseudo-NORs” lack the promoter sequence for the production of rRNA. Thus, “pseudo-NORs” are not transcriptionally active and do not form functional nucleoli which produce ribosomes (Mais et al., 2005). Therefore, the next step in the construction of synthetic nucleoli was to include RNAPI transcription. Active transcription was achieved by Grob et al. who made “neo-NORs” (Grob et al., 2014). These “neo-NORs” intersperse the XEns included in “pseudo-NORs” with human rDNA promoters, mouse pre-rRNA coding sequences, and mouse transcriptional terminators (Grob et al., 2014). The “neo-NORs” were transcriptionally active, processed pre-rRNA, produced ribosomes, and coalesced into endogenous NORs in HT1080 cells to form larger nucleoli (Grob et al., 2014). The addition of rDNA transcription units allowed “neo-NORs” to form functionally compartmentalized nucleoli, while “pseudo-NORs” only formed the FC (Grob et al., 2014). This is similar to the effect observed just after mitosis, when UBF is bound to the rDNA to denote NORs that were active during the previous interphase (Grob and McStay, 2014). Therefore, formation of functional nucleoli requires at least one rRNA gene, recruitment of the RNAPI transcription machinery including UBF and SL1, and active transcription of the rDNA.

There are therefore only a few absolute requirements for the formation of functional nucleoli, resulting in much variation in their number, shape, and size across different species and cell types. NORs are found at 10 locations in humans, on the short arms of the 5 acrocentric chromosomes (13, 14, 15, 21, and 22). This leaves the potential for human cells to form 10 nucleoli. However, cells often have far fewer nucleoli per cell nucleus. As early as the 1960s, Shea and Leblond demonstrated that nucleolar number in different mouse tissue sections ranged between one and six nucleoli per cell with an average of 2-3 nucleoli per cell (Shea and Leblond, 1966). Other studies have shown that approximately 91% of HT1080 human fibrosarcoma cells have greater than 3 nucleoli per cell (Krystosek, 1998). There are no comprehensive studies on how nucleolar number changes with differences in organism, tissue type, or over the course of development.

Nucleolar size/area can also fluctuate greatly in human cells. This parameter has been examined mainly in the context of cancer (Penzo et al., 2019). Studies using various cancer cell lines have shown that increases in nucleolar area per nucleus are directly related to increased RNAPI activity of the cell as well as to increased UBF, DNA topoisomerase I, and fibrillarin expression (Derenzini et al., 1998). Additionally, depletion of proteins responsible for controlling proliferation, such as p53 and pRb, causes an increase in nucleolar area (Treré et al., 2004). This finding is logical as increased proliferation is linked to increased production of ribosomes and therefore linked to increased rRNA transcription. However, proliferation alone does not fully explain differences in nucleolar size. For example, rapidly proliferating small-cell anaplastic lung cancer cells have a small nucleolar area per nucleus, while slower growing large-cell lung carcinoma cells have a much larger nucleolar area per nucleus (Zink et al., 2004). While these abnormal, malignant cells provide useful model systems for examining nucleolar size/area, more studies are needed to understand the mechanisms governing these parameters in normal human cells.

In the first part of my thesis work, I have identified a plethora of factors required for nucleolar formation in human cells. Kathleen L. McCann, a former graduate student in the Baserga laboratory, performed an unbiased, genome-wide siRNA screening campaign with the intent of discovering novel regulators of nucleolar number in human cells. I analyzed the data from this screen and performed several follow-up studies on a number of screen hits to delve deeper into the specific ways that these hits influence the production of ribosomes. Ultimately, I found that many (90%) of the screen hits tested are essential for the nucleolar functions of ribosomal DNA (rDNA) transcription (7/20), pre-ribosomal RNA (pre-rRNA) processing (16/20), and/or global protein synthesis (14/20). The results from this screen reveal how multiple cellular pathways converge in the regulation of human ribosome biogenesis and pave the way for new knowledge of how ribosome biogenesis is affected in human disease.

Multiple disorders of ribosome biogenesis affect craniofacial development

Although the proteins involved in regulating ribosome biogenesis are just beginning to be understood, especially in human cells, the consequences for dysregulating this process have been even less studied. Such interruptions in ribosome synthesis cause a number of human disorders, collectively called ribosomopathies. Ribosomopathies often result from haploinsufficiency or point mutation in genes that affect any of a number of steps in the process of making ribosomes. For example, North American Indian Childhood Cirrhosis (NAIC, OMIM 604901) is caused by a mutation in the *UTP4* gene (R565W) that plays a role in pre-rRNA processing at the A’, A0, 1, and 2b cleavage sites (Betard et al., 2000, Weber et al., 1981, Freed et al., 2012). Alopecia, neurologic defects, and endocrinopathy (ANE, OMIM 612079) syndrome is a second ribosomopathy that affects pre-rRNA processing, but of the pre-rRNAs responsible for the creation of the LSU (Nousbeck et al., 2008, Sloan et al., 2013b, McCann et al., 2016). The ribosomopathy Diamond Blackfan Anemia (DBA, OMIM 105650) is mainly caused by mutations in LSU or SSU r-proteins, with mutations in RPS19 being the most frequent [reviewed in (Da Costa et al., 2018)]. Additionally, X-linked dyskeratosis congenita (DC, OMIM 305000) results from mutations in dyskerin pseudouridine synthase 1 (DKC1), which plays a role in pre-rRNA modification by pseudouridylation (Heiss et al., 1998, Lafontaine et al., 1998). Traditionally, for a disease to be considered a ribosomopathy, the signs and symptoms must be directly caused by defects in any stage of ribosome biogenesis (De Keersmaecker et al., 2015). However, this definition has recently been questioned as several cancers, most notably T-cell acute lymphoblastic leukaemia (T-ALL), have also been found to have defects in proteins required for making ribosomes (De Keersmaecker et al., 2013, Novetsky et al., 2013, Wang et al., 2011). However, we do not yet know whether these ribosome biogenesis defects are the direct cause of cancer, and so the classification of these disorders as ribosomopathies remains open to debate.

Perhaps the most intriguing facet of ribosomopathies is that they result in different, tissue-specific disorders. This is contrary to the expectation that disruptions in such a ubiquitous process as ribosome biogenesis would result in an inviable organism. However, patients with ribosomopathies present with tissue-specific phenotypes that vary according to the gene mutated. For example, NAIC patients present with transient neonatal jaundice that proceeds to biliary cirrhosis (Betard et al., 2000, Drouin et al., 2000), while ANE syndrome manifests in hair loss, neurological defects, endocrine abnormalities, and short stature (Nousbeck et al., 2008). DBA patients present with anemia, low reticulocyte count, elevated erythrocyte adenosine deaminase activity, and have an increased risk of cancer (Glader et al., 1983, Halperin and Freedman, 1989, Vlachos et al., 2018, Alter et al., 2018). In DC, patients have bone marrow failure as well as an increased risk of head and neck squamous cell carcinoma, acute myeloid leukemia, and non-Hodgkin lymphoma (Dokal, 2000, Dokal, 1996, Alter et al., 2018). The variety of affected tissues presents a significant challenge in understanding the pathogenic mechanisms of these disorders.

Multiple ribosomopathies result in defects in craniofacial development, with Treacher Collins syndrome (OMIM 154500) being the most well-studied of these disorders (Table 1-1). Treacher Collins syndrome is caused by mutations in the *TCOF1* (treacle), *POLR1C*, or *POLR1D* genes (Altug Teber et al., 2004, Dauwerse et al., 2011, Bowman et al., 2012, Splendore et al., 2000). Each of these 3 genes play a role in the transcription of the rDNA (Valdez et al., 2004, Noack Watt et al., 2016). TCOF1 binds to UBF to aid in transcription initiation (Valdez et al., 2004). Additionally, TCOF1 binds to the proximal region of the promoter and 5' end of ETS1 where it may help to recruit NOP56, part of the cellular 2'-O-methylation machinery (Gonzales et al., 2005, Hayano et al., 2003). POLR1C and POLR1D are components of the RNAPI and RNAPIII polymerases, and mutations in either of these genes result in decreased rDNA transcription (Noack Watt et al., 2016). Despite the essential roles for these 3 genes in ribosome biogenesis, Treacher Collins syndrome patients are not always severely affected by their disease. Most clinical cases present with craniofacial defects, usually hypoplasia of the mandible and/or zygomatic complex [reviewed in (Trainor et al., 2008, Kadakia et al., 2014)]. Some Treacher Collins patients have dental anomalies, including a reduced number of teeth, as well as cleft palate (da Silva Dalben et al., 2006). In severe cases, perinatal death may occur due to occlusion of the airway (Edwards et al., 1996). Indeed, mouse models of Treacher Collins syndrome have severe craniofacial abnormalities and die due to respiratory arrest (Dixon et al., 2006). Therefore, the defects in rDNA transcription seen in patients with Treacher Collins syndrome can specifically affect craniofacial development.

Table 1-1. Genes with roles in both ribosome biogenesis and craniofacial development.

Gene Name	**Function in ribosome biogenesis**	**Defects in craniofacial development**	**Name of human disease**	**Reference**
DDX11	rDNA transcription	Long faces, narrow eyes, low mouths	Warsaw Breakage Syndrome (WABS)	(van der Lelij et al., 2010, Capo-Chichi et al., 2013, Sun et al., 2015)
ESF1	Pre-rRNA processing	Jaw malformations, microcephaly	n/a	(Chen et al., 2018)
NOL11	rDNA transcription and pre-rRNA processing	Microcephaly, reduced size of pharyngeal cartilages	n/a	(Griffin et al., 2015, Freed et al., 2012)
PAK1IP1	Pre-rRNA processing	Midline facial cleft	n/a	(Yu et al., 2011, Ross et al., 2013)
POLR1A	rDNA transcription	Range of mandibulofacial dystoses including downslanting palpebral fissures, eyelid clefts, and micrognathia	Acrofacial dystosis, Cincinnati type	(Weaver et al., 2015)
RPL38/eL38	Large subunit ribosomal protein	midline facial cleft, cleft palate	n/a	(Kondrashov et al., 2011)
RPS19/eS19, RPL5/uL18, RPL11/uL5, RPL35a/eL33, RPS26/eS26, RPS24/eS24, RPS17/eS17, RPS7/eS7, RPS10/eS10, RPL19/eL19, RPL26/uL24, RPS29/uS14, RPL31/eL31, RPS28/eS28, RPS20/uS10, RPL15/eL15, RPL17/uL22, GATA1, TSR2	Mainly ribosomal proteins	Cleft lip, cleft palate, flat nasal bridge, hypertelorism	Diamond Blackfan Anemia	(Lipton et al., 2006, Ball et al., 1996)

TCOF1, POLR1C, POLR1D	rDNA transcription	Hypoplasia of mandible/zygomatic complex, some dental anomalies, cleft palate	Treacher Collins Syndrome	(Trainor et al., 2008, Valdez et al., 2004, da Silva Dalben et al., 2006, Noack Watt et al., 2016)
WDR43	rDNA transcription	Reduced size of pharyngeal cartilages, hydrocephaly	Linked to 3-M syndrome	(Zhao et al., 2014, Sondalle et al., 2016)

In addition to Treacher Collins syndrome, several other ribosomopathies are known to affect craniofacial development. DBA has been linked to craniofacial abnormalities and cleft palate (Lipton et al., 2006, Ball et al., 1996). Mutations in *POLR1A*, the largest RNAPI subunit, cause the recently described ribosomopathy, Acrofacial dysostosis, Cincinnati type (OMIM 616462) (Weaver et al., 2015). Patients with this disorder present with a range of mandibulofacial dystoses, either with or without other skeletal defects. Additionally, patients with the cohesinopathy Warsaw Breakage syndrome (WABS, OMIM 613398) have several congenital abnormalities, including microcephaly, small forehead, elongated face, clinodactyly of the fifth fingers, and intellectual disability (van der Lelij et al., 2010, Capo-Chichi et al., 2013). Mutations in the helicase *DDX11* have been identified as the cause for WABS (Capo-Chichi et al., 2013, van der Lelij et al., 2010). Zebrafish with morpholino oligonucleotide (MO) knockdown of DDX11 have elongated faces, shortened torsos, small and narrow eyes, and low and protuberant mouths (Sun et al., 2015). In addition to its mitotic role in sister chromatid cohesion and segregation, DDX11 has recently been shown to bind to the rDNA and promote the rDNA association of both RNAPI and UBF, potentially classifying WABS as a ribosomopathy (Sun et al., 2015). Therefore, there is a strong link between ribosome biogenesis and craniofacial development.

There are other strong links between craniofacial development and ribosome biogenesis that have yet to be associated with a human disease. The *Rpl38(eL38)* gene is mutated in *Tail-short* mice where defects include the characteristic short and kinked tail as well as axial skeletal patterning defects, a midline facial cleft, and cleft palate (Kondrashov et al., 2011). The protein PAK1 interacting protein 1 (PAK1IP1) may also play a role in human ribosome biogenesis by localizing to nucleoli and potentially influencing pre-rRNA processing (Yu et al., 2011). *PAK1IP1* mutations in mice also cause craniofacial defects as well as smaller forebrain and hindbrain, although no corresponding mutation has yet been identified in humans (Ross et al., 2013). In addition, the protein 18S rRNA factor 1 (ESF1) has recently been implicated in craniofacial development in zebrafish (Chen et al., 2018). ESF1 is required for SSU pre-rRNA processing, and zebrafish homozygous for an *esf1* mutation die at approximately 6 dpf, highlighting the essential nature of this protein (Chen et al., 2018, Peng et al., 2004). Studies on the protein NOL11 in *X. tropicalis* have furthered our knowledge of the connection between ribosome biogenesis and the developing neural crest. Depletion of NOL11 results in defects in rDNA transcription and in the processing of the 5'ETS in both humans and *X. tropicalis* (Freed et al., 2012, Griffin et al., 2015). NOL11 depletion by MOs in *X. tropicalis* specifically disrupts cranial neural crest development (Griffin et al., 2015). The size of the Meckels, ceratohyal, and branchial cartilages are all reduced upon NOL11 knockdown (Griffin et al., 2015). Finally, mutations in *wdr43* in zebrafish (termed *fantome/fan* mutants) also cause defects in the developing neural crest, particularly to the pharyngeal arch cartilages (Zhao et al., 2014). *wdr43* (the ortholog of yeast *UTP5*) is a member of the t-UTP subcomplex, which is required for pre-rRNA transcription and processing (Gallagher et al., 2004, Prieto and McStay, 2007), and interacts with UTP4, the protein mutated in the NAIC ribosomopathy (Chagnon et al., 2002, Freed and Baserga, 2010). These examples illustrate the importance of ribosome biogenesis in the developing cranial neural crest. The precise mechanism for how some ribosomopathies specifically affect craniofacial development is an active area of investigation.

For several ribosomopathies, the pro-apoptotic protein p53 has been implicated as an important player in the signs and symptoms of the disease. When ribosome biogenesis is disrupted, the 5S ribonucleoprotein complex (5S RNP) is formed from the free uL18 (RPL5) and uL5 (RPL11) r-proteins as well as the 5S rRNA (Sloan et al., 2013a, Donati et al., 2013) (Figure 1-3). This 5S RNP is able to bind to MDM2, either with or without p14^{ARF}, sequestering it so that MDM2 can no longer act as an E3 ubiquitin ligase for p53 to trigger its degradation (Sloan et al., 2013a). Under these stress conditions, p53 is free to perform its function in arresting the cell cycle and triggering apoptosis, and this process is called the “nucleolar stress response” (Pestov et al., 2001) (Figure 1-3). When ribosome biogenesis functions correctly, however, the components of the 5S RNP are engaged in translating ribosomes, allowing the degradation of p53 via MDM2.

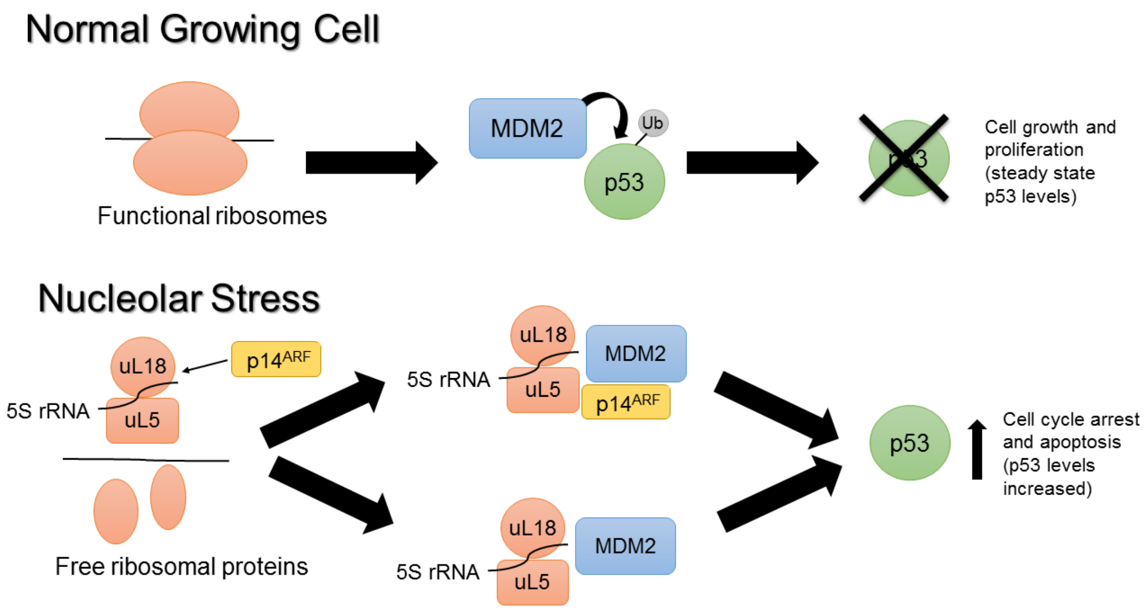

Figure 1-3. The nucleolar stress response. During normal cellular growth conditions with functional ribosomes, p53 is degraded upon ubiquitination by MDM2 (top). In contrast, under conditions of nucleolar stress, the free ribosomal proteins uL18 and uL5, along with the 5S rRNA, form the 5S RNP complex. This complex, either with or without p14ARF, is then able to sequester MDM2. Without the E3 ligase function of MDM2, p53 is not ubiquitinated and is therefore able to arrest the cell cycle and promote apoptosis. Figure from (Farley and Baserga, 2016).

p53 coordination and the nucleolar stress response have been implicated in the pathology of a number of ribosomopathies, including NAIC (Wilkins et al., 2013), DBA (Danilova et al., 2008, McGowan et al., 2008), dyskeratosis congenita (Pereboom et al., 2011, Zhang et al., 2012, Gu et al., 2008), and 5q-syndrome (Barlow et al., 2010). Perhaps the best studied ribosomopathy in regards to p53 is Treacher Collins syndrome. Mutations in *Tcof1* in certain murine backgrounds cause the craniofacial defects characteristic of Treacher Collins syndrome and result in the specific apoptosis of neural crest cells (Dixon et al., 2006). However, these defects can be rescued through inhibition of p53 (Jones et al., 2008a). This supports the argument that the cranial neural crest cells are particularly susceptible to perturbations in the levels of p53, and it is this stabilization, rather than the ribosome biogenesis defects themselves, that results in the tissue-specific phenotype. Indeed, studies of the ribosome biogenesis factors NOL11 and WDR43 also support this neural crest cell sensitivity (Griffin et al., 2015, Zhao et al., 2014). The specific molecular mechanism(s) responsible for inducing the p53-mediated apoptosis response in the cranial neural crest cells of these ribosomopathies remains to be elucidated and will be an active area for future study.

Interestingly, the nucleolar stress response is not able to fully explain all of the ribosomopathy tissue-specificities. For example, in a zebrafish model of Treacher Collins syndrome, the overall viability and small size of *polr1c*$^{-/-}$ and *polr1d*$^{-/-}$ mutant embryos is not rescued by co-depletion of p53, even though the craniofacial defects are rescued by this method (Noack Watt et al., 2016). Also, in mice, knockout of p53 in addition to the *SBDS* mutation found in Schwachman-Diamond syndrome is able to rescue the exocrine pancreatic dysfunction characteristic of the ribosomopathy (Tourlakis et al., 2015, Warren, 2018). However, p53 knockout does not rescue 80S monosome levels, overall growth, or perinatal survival (Tourlakis et al., 2015). p53 depletion was also unable to fully rescue Shwachman-Diamond syndrome defects in a zebrafish model of the disease

(Provost et al., 2012). Additionally, in zebrafish models of acrofacial dystosis, Cincinnati type, p53 mutation only postpones the craniofacial dysmorphology to a later stage of development, and the embryo's lifespan is extended only from 5 dpf to 8 dpf (Watt et al., 2018). The idea of p53-independent pathways affecting ribosomopathy signs and symptoms is also supported by models of the DBA ribosomopathy (Wan et al., 2016, Aspesi et al., 2014). Therefore, it is likely that ribosomopathies have both p53-dependent and p53-independent mechanisms of action. Therefore, although p53 and nucleolar stress can account for some tissue specific differences in ribosomopathies, it does not necessarily account for all of them.

PAX9's role in craniofacial development

Proteins of the paired box (PAX) family act as RNAPII transcription factors to control the expression of mRNAs required for organogenesis. A total of 9 PAX proteins have been discovered in vertebrates, all containing the highly conserved paired box domain that was first identified in *Drosophila* melanogaster (Bopp et al., 1986). The paired domain is comprised of amino-terminal and carboxy-terminal subdomains, both of which contact DNA in adjacent major grooves using a helix-turn-helix motif (Czerny et al., 1993, Xu et al., 1999). All 9 proteins additionally contain a c-terminal transactivation domain. The 9 PAX proteins differ depending on the presence of an additional DNA binding homeodomain and/or an octapeptide motif. The homeodomain binds DNA while the octapeptide motif interacts with repressive protein co-factors (Eberhard et al., 2000). The 9 PAX proteins are often subdivided into 4 groups based upon the presence/absence of these additional motifs (Table 1-2), which provide layers of specificity for each of the PAX proteins to modulate the transcription of certain mRNAs required for development.

Table 1-2. Paired box proteins play varying roles in human development. Data reviewed in (Monsoro-Burq, 2015, Wachtel and Schäfer, 2015, Robson et al., 2006, Chi and Epstein, 2002).

Protein Name	**Group Number**	**Octapeptide Motif**	**Homeo domain**	**Developmental targets**	**Disease Association**
PAX1	I	Yes	No	Skeleton, thymus	Otofacialcervical syndrome
PAX2	II	Yes	Yes (truncated)	Kidney, central nervous system	Renal-coloboma syndrome, Breast, ovarian, prostate, and renal carcinomas, Wilms tumor
PAX3	III	Yes	Yes	Skeletal muscle, neural crest, heart, central nervous system, craniofacial development	Waardenburg syndrome, Craniofacial-deafness-hand syndrome, alveolar rhabdomyosarcoma, biphenotypic sinonasal sarcoma
PAX4	IV	No	Yes	Pancreas	Pancreatic carcinoma
PAX5	II	Yes	Yes (truncated)	B-cell development, central nervous system	B-cell precursor acute lymphoblastic leukemia (B-ALL), chronic myelogenous leukemia (CML)
PAX6	IV	No	Yes	Ocular development, central nervous system, pancreas	Aniridia, Astrocytic glioma, glioblastoma, pancreatic carcinoma, cataracts
PAX7	III	Yes	Yes	Skeletal muscle formation, central nervous system, craniofacial development	Alveolar rhabdomyosarcoma

PAX8	II	Yes	Yes (truncated)	Thyroid, kidney, central nervous system	Congenital hypothyroidism, Follicular thyroid carcinoma, follicular-variant thyroid carcinoma, follicular adenoma, Wilms tumor
PAX9	I	Yes	No	Teeth, Craniofacial development, thymus/parathyroid, skeleton	Tooth agenesis, hypodontia, oesophageal cancer

The differing transcriptional profiles for each PAX protein are important for the specification and development of various tissue types (Table 1-2). For example, PAX6, which contains the full homeodomain but not the octapeptide motif, plays a role in ocular development (Glaser et al., 1994). However, PAX3, containing both the full homeodomain and octapeptide motif, influences the development of skeletal muscle (Barr et al., 1993, Relaix et al., 2005). Disruptions in the function of PAX proteins also influences the pathology of a number of human diseases, including Waardenburg syndrome (PAX3), hypodontia (PAX9), and B-cell precursor acute lymphoblastic leukemia (B-ALL). PAX proteins additionally play roles in a number of human cancers, either through gain-of-function (e.g. PAX3) or loss-of-function (e.g. PAX6) mutations [reviewed in (Robson et al., 2006, Wachtel and Schäfer, 2015)]. Notably, many of the PAX proteins modulate their activity through the use of alternative splicing, although these effects remain relatively understudied (Short and Holland, 2008). Subtle differences in the sequence of each PAX protein therefore may result in large differences in their expression programs, contributing to their varying roles in development and disease.

My thesis work focuses on the PAX9 protein, which specifically affects craniofacial and tooth development. In 1993, Stapleton et al. first isolated the PAX9 gene, determining that it is highly homologous to the previously identified PAX1 (Stapleton et al., 1993). PAX9 and PAX1 comprise Group I of the PAX proteins, as they contain the octapeptide motif but not the homeodomain. The expression patterns for PAX9 and PAX1 are similar; both are expressed in the pharyngeal pouches but only PAX9 is expressed in the neural crest-derived mesenchyme (Neubüser et al., 1995, Peters et al., 1998). Expression studies of PAX9 in E9.0 mice revealed PAX9 in only the pharyngeal pouches (Peters et al., 1998). However, by E13.5, PAX9 is expressed in many other tissues including the limbs, vertebral column, and facial mesenchyme

(Peters et al., 1998). By E16.5, PAX9 is expressed in the mesenchyme of all teeth, the tongue, the salivary glands, the thymus, and in other developing tissues (Peters et al., 1998). This expression pattern correlates with PAX9's function in craniofacial and tooth development.

Compared to other PAX proteins, little is known of the molecular targets of PAX9. The atomic structure of PAX9 has not been determined, although it is expected to be highly similar to the other PAX proteins according to molecular modeling (Wang et al., 2009b, Narasimhan et al., 2014). PAX9 is able to recognize two common PAX DNA binding motifs, CD19-2(A-ins) and e5, the *Drosophila even-skipped* promoter (Mensah et al., 2004, Czerny et al., 1993). This is expected, as multiple PAX proteins have been shown to bind the same DNA sequence (Epstein et al., 1994). Chromatin immunoprecipitation sequencing (ChIP-seq) of PAX9 in the vertebral columns of E12.5 mice has additionally identified a PAX9 consensus motif (5' CGCGTGACCG 3'), which is similar to the CD19-2(Ains) motif that was reported as a PAX binding site by Czerny et al. (Sivakamasundari et al., 2017, Czerny et al., 1993). As these DNA binding sequences are all similar to known PAX motifs, current analyses have revealed little about the specific PAX9 consensus motif(s) and how they relate to tooth and craniofacial development.

Regarding tooth development, a few PAX9 signaling targets have been identified (Peters and Balling, 1999) (Figure 1-4). At the initiation stage of tooth development, fibroblast growth factor 8 (FGF8) is able to upregulate expression of PAX9 (Neubuser et al., 1997). Then, at the bud stage, PAX9 upregulates the expression of bone morphogenic protein 4 (BMP4), which in turn upregulates both Msh homeobox 1 (MSX1), and lymphoid enhancer binding factor 1 (LEF1) (Peters et al., 1998). These three proteins are key to the signaling program of the enamel knot (Jernvall et al., 1998, Jernvall et al., 1994). In the cap stage of tooth development, the enamel knot directs the growth of adjacent cells and later initiates its own apoptosis [reviewed in (Peters and Balling, 1999)]. This signaling program therefore creates differential proliferation that results in the formation of tooth cusps that give the teeth their individual surfaces. Additionally, BMP4 is able to increase the expression of p21, which induces cell differentiation and cell cycle arrest (Chang et al., 2009). It has also been shown that PAX9 influences Osr2 expression, which in turn alters palatal cell proliferation and palatal shelf elevation (Zhou et al., 2013). Indeed, expression of Osr2 is able to partially rescue the cleft palates present upon PAX9 deletion (Zhou et al., 2013). PAX9 is also involved in modulating Wnt signaling, but the direct molecular mechanism for this action is not well understood (Li et al., 2017, Sivakamasundari et al., 2017). Therefore, while there is some knowledge of PAX9's role in tooth development, little is known of the molecular targets of PAX9 outside of this process.

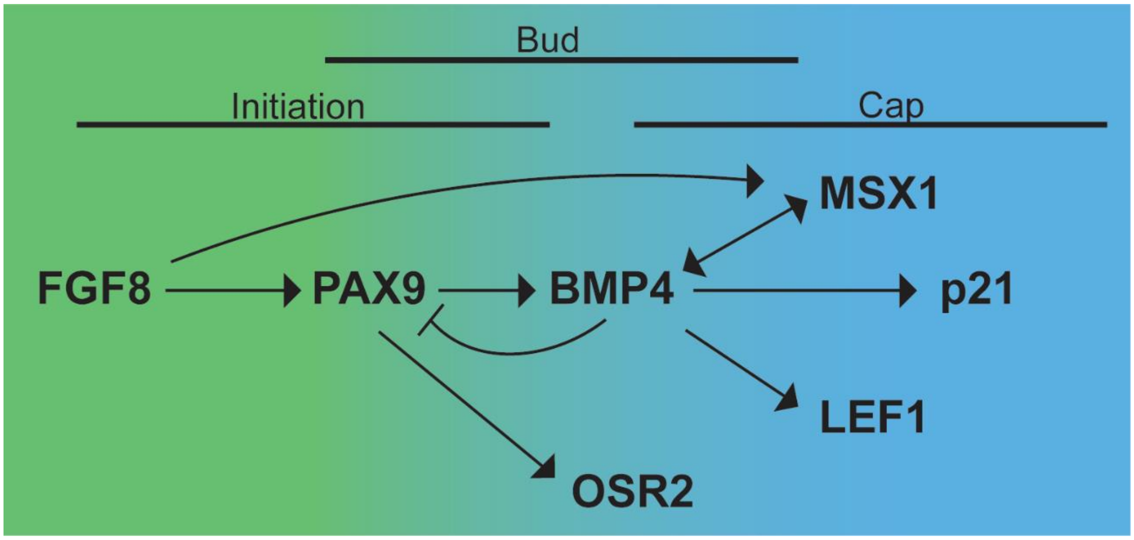

Figure 1-4. PAX9 signaling in tooth development. At the initiation stage of tooth development (green), FGF8 signaling upregulates the expression of PAX9 and MSX1. PAX9 then acts as an RNAPII transcription factor during the budding stage of tooth development to upregulate the transcription of BMP4. BMP4 then works to upregulate the expression of a number of factors, including MSX1, LEF1, and p21 at the cap stage (blue).

In humans, mutations in the *PAX9* gene result in tooth agenesis and craniofacial malformations. Indeed, for patients with nonsyndromic tooth agenesis, *PAX9* is the most commonly mutated gene (Yu et al., 2018). In 2000, Stockton et al first pinpointed an autosomal dominant frameshift mutation in the *PAX9* gene in a family presenting with oligodontia, or the loss of 6 or more teeth excluding the third molar (Stockton et al., 2000). Since then, a number of other studies have found additional *PAX9* mutations in patients with oligodontia and hypodontia (loss of 1-5 teeth, excluding the third molar) [(Nieminen et al., 2001, Frazier-Bowers et al., 2002, Šerý et al., 2015, Mostowska et al., 2013, Wong et al., 2018, Daw et al., 2017, Sarkar et al., 2017) and reviewed in (Fauzi et al., 2018)]. PAX9 mutation has also been linked to cleft lip/palate (Lee et al., 2012, Das et al., 2003, Schuffenhauer et al., 1999). To date, the Online Mendelian Inheritance in Man (OMIM) database has cited 58 PAX9 allelic variations associated with either the absence of teeth or a reduction in tooth number (www.omim.org). Most of the mutations are in the paired domain that provides PAX9's DNA-binding function. In addition to the oligodontia associated with PAX9 mutation, at least one study found hair irregularities including alopecia, hypotrichosis, and abnormalities of the hair shaft pigmentation (Mostowska et al., 2013). Interestingly, reduced PAX9 protein levels have also been implicated in both oesophageal carcinomas and epithelial dysplasias (Gerber et al., 2002, Robson et al., 2006). Investigation of PAX9's molecular function will be important in understanding craniofacial and tooth development in humans.

Several studies have been conducted in mice to probe the role of PAX9 during development. Studies in this system have been problematic, however, as mice homozygous for a mutation in which the *Pax9* start codon and paired box sequence were replaced with a functional null allele die within a few hours of birth (Peters et al., 1998). These PAX9-deficient mice exhibit multiple craniofacial abnormalities, cleft secondary palates, limb abnormalities, and have severe respiratory problems (Peters et al., 1998). Additionally, tooth development is arrested at the bud stage, and the thymus, parathyroid glands, and ultimobranchial bodies are missing (Peters et al., 1998). Notably, although PAX1 is present in the epithelium of the third pharyngeal pouches, it appears unable to compensate for the loss of Pax9, indicating a separate role for PAX9 in development of these tissues (Neubüser et al., 1995, Wallin et al., 1996). Interestingly, heterozygous *Pax9* mutant mice are phenotypically normal (Peters et al., 1998). Two recent studies have demonstrated partial rescue of the morphogenic defects in *Pax9*$^{-/-}$ mice by using small molecules that target Wnt signaling (Jia et al., 2017a, Li et al., 2017). A third study also demonstrated a partial rescue of the palatal defects in *Pax9*$^{-/-}$ mice through modulation of the related EDAR pathway (Jia et al., 2017b). Interestingly, none of these treatments were able to rescue arrested tooth development, limb defects, or lethality just after birth (Jia et al., 2017b, Jia et al., 2017a, Li et al., 2017). This suggests that there are additional unexplored targets of PAX9 signaling that contribute to the pathogenesis of the PAX9 mutation in humans.

In the second part of my thesis work, I have investigated a new target of PAX9 signaling: ribosome biogenesis. Through the use of the genome-wide siRNA screen described in the first part of my thesis, I identified PAX9 as a novel regulator of nucleolar number in humans. Further exploration of PAX9's function in ribosome biogenesis has revealed that PAX9 expression is required for SSU pre-rRNA processing and global protein synthesis. Using two different transcriptomics approaches, I have revealed multiple targets of PAX9 signaling that contribute to this function in making ribosomes. Additionally, this work connects PAX9's role in craniofacial development to its role in ribosome biogenesis in the model organism, *Xenopus tropicalis (X. tropicalis)*. Illuminating PAX9's function in human ribosome biogenesis will provide new understanding of the pathology of PAX9 mutations in patients with oligodontia.

Chapter 2

A genome-wide siRNA screen for human regulators of nucleolar number

INTRODUCTION

Many of the protein components involved in ribosome biogenesis were first described in the yeast, *Saccharomyces cerevisiae*, because of tractable biochemistry and genetics (Dixon et al., 2006, Weaver et al., 2015); however, there is growing evidence that ribosome biogenesis in human cells has acquired greater complexity in regulation (James et al., 2014, Rubbi and Milner, 2003, Zhang and Lu, 2009, Boulon et al., 2010, Vlatkovic et al., 2014, Golomb et al., 2014, Aubert et al., 2018). Additionally, nucleolar formation is substantially different in yeast than in humans. First, there is only one NOR in *Saccharomyces cerevisiae*, while there are 10 potential sites for nucleolar formation in humans. Second, yeast have a closed mitosis in which the nucleoli do not have to reform after each cell division, while human cells have a significantly more complex open mitosis. Finally, nucleolar ultrastructure is different between yeast and humans, with yeast having only two nucleolar compartments compared to three in human cells. Thus, it is likely that there are additional unexplored mechanisms regulating nucleolar formation and function in humans.

In humans, nucleolar number varies for unknown reasons. A few targeted studies have noted the number of nucleoli in some cell lines, but not much has been explored regarding the mechanisms regulating this number. For example, approximately 91% of HT1080 human fibrosarcoma cells have greater than 3 nucleoli per cell (Krystosek, 1998). In addition, as early as the 1960s, Shea and Leblond demonstrated that nucleolar number in different mouse tissue sections ranges between one and six nucleoli per cell with an average of 2-3 nucleoli per cell (Shea and Leblond, 1966). As with nucleolar size/area, information regarding nucleolar number in normal, non-transformed mammalian tissues is lacking. Nucleolar number therefore varies greatly among mammalian cells and between tissue types. A comprehensive account of the average nucleolar number for all tissues and cell lines is therefore needed to better understand the mechanisms governing nucleolar number determination.

Two genome-wide screens have been carried out to identify the cellular components required for nucleolar formation (Badertscher et al., 2015, Neumuller et al., 2013). Screens performed in *Drosophila melanogaster* and *S. cerevisiae* cells used nucleolar size as an endpoint (Neumuller et al., 2013), while the screen performed in HeLa cells used an assay that detects ribosome assembly and transport for the small ribosomal subunit (SSU) (Badertscher et al., 2015). However, a complete genome-wide screen for large ribosomal subunit (LSU) assembly and transport has yet to be carried out (Wild et al., 2010). Additionally, a report of an assay for nucleolar morphology in HeLa cells calculated specific parameters of abnormal-looking nucleoli, termed the iNo, but the screen was restricted only to ribosomal proteins (Nicolas et al., 2016). As the human nucleolar function screens have been limited to the aneuploid HeLa cell line, and no one screen has examined both SSU and LSU biogenesis genome-wide, there remain many open questions regarding the complex mechanisms that coordinate nucleolar morphology and function in human cells.

To enhance our understanding of the mechanisms regulating ribosome biogenesis in human cells, we embarked upon an unbiased, genome-wide siRNA screening campaign using a readout of nucleolar number. A previous graduate student, Emily Freed, determined that defective ribosome biogenesis resulting from siRNA depletion of ribosome biogenesis factors (UTP4 and NOL11) correlates with changes in nucleolar number in human cells (Freed et al., 2012). In collaboration with the Yale Center for Molecular Discovery, Kathleen L. McCann, another previous graduate student, exploited this relationship between nucleolar number and function in a genome-wide screen conducted in the “near-normal” human MCF10A breast epithelial cell line (Soule et al., 1990). All of the analyses of the screen hits were performed by myself. The screen identified 139 candidate regulators of human ribosome biogenesis whose depletion altered the number of nucleoli per nucleus from 2-3 to only 1. The identified proteins have a wide range of known functions, and likely regulate nucleolar processes from both inside and outside of the nucleolus.

The siRNA screen was published in (Farley-Barnes et al., 2018) with myself as co-first author. Kathleen L. McCann performed the original siRNA screen in collaboration with the Yale Center for Molecular Discovery. I analyzed the data generated from the screen, performed most of the follow-up experiments, and wrote the first draft of the manuscript. Additionally, the comparisons of nucleolar number among mammalian cell lines was performed by Yulia Surovtseva at the Yale Center for Molecular Discovery, and were analyzed and published in a first author review article written by me (Farley et al., 2015).

RESULTS

Nucleolar number varies between cell types

In humans, the number of nucleoli per cell nucleus varies for unknown reasons. Because there are 10 NORs located on the acrocentric chromosomes in humans, presumably a maximum of 10 nucleoli in human cells is possible. However, most human cells have far fewer active nucleoli, and many questions remain as to the mechanisms controlling nucleolar number.

To examine nucleolar number variation in mammalian cells, I collaborated with the Yale Center for Molecular Discovery to visualize nucleolar number in multiple cell lines using an antibody to the nucleolar protein fibrillarin. The number of nucleoli per nucleus was then determined via a pipeline in CellProfiler (Carpenter et al., 2006). From this analysis, we found that the number of nucleoli per cell fluctuates greatly within a given population of the same cell line. Examining a frequency distribution of the number of nucleoli per cell shows a normal curve with mean and variance that differ by cell line

(Fig 2-1). Multiple mammalian cell lines, including HeLa human cervical carcinoma cells, MCF10A human mammary epithelial cells, MDCK dog kidney cells, and CHO Chinese hamster ovary cells, contained an average of roughly 3 nucleoli per cell (Fig 2-1). However, the mean number of nucleoli per nucleus is different for other cell lines. For example, U20S human bone osteosarcoma cells have an average of 6 nucleoli per nucleus, while T98G human glioblastoma multiforme cells have 8 nucleoli per nucleus (Fig 2-1). The cell lines tested ranged in ploidy, species, and tissue type. Notably, the CHO cells used were the only tested cell line that was derived from a subclone. Future studies are necessary to determine the precise mechanisms influencing nucleolar number in these other cell/tissue types.

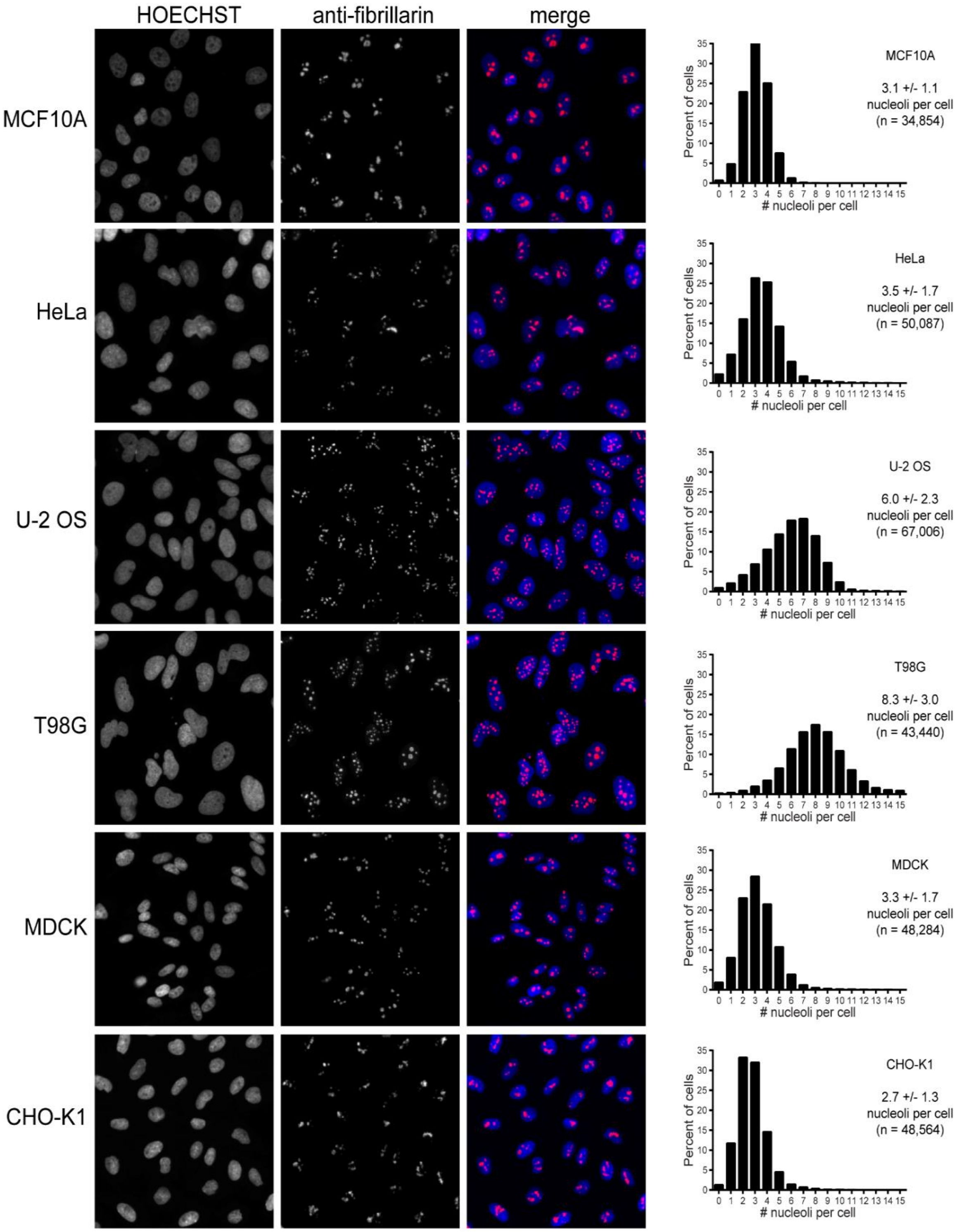

Figure 2-1. Nucleolar number varies greatly among tissue culture cell lines. The indicated cell lines were fixed with paraformaldehyde and stained with an antibody to the nucleolar protein fibrillarin [72B9; (Reimer et al., 1987b)]. Cells were also stained with

HOECHST 33342 dye for visualization of the nucleus. Images were analyzed using a CellProfiler pipeline which counts the number of nucleoli per cell (Carpenter et al., 2006). Representative images for each cell line are depicted with a frequency distribution of the number of nucleoli per nucleolus shown to the right. Figure from (Farley et al., 2015).

siRNA screening identifies 139 regulators of nucleolar number

To identify novel regulators of nucleolar number in human cells, Kathleen L. McCann, in collaboration with the Yale Center for Molecular Discovery, screened 18,107 siRNAs using the GE Healthcare Dharmacon siGENOME library (Fig 2-2, Appendix I). MCF10A breast epithelial cells were used for the screen because they are non-tumorigenic and have a near-diploid karyotype, allowing us to better infer the mechanisms regulating nucleolar number in normal human cells. MCF10A were reverse transfected in each well of a 384-well plate with a pool of 4 siRNAs targeting the same gene of interest. Nucleoli were detected by staining with an anti-fibrillarin (FBL) antibody (Reimer et al., 1987a), nuclei were detected by HOECHST 33342 staining, and the number of nucleoli per nucleus was quantified using a CellProfiler (Carpenter et al., 2006) pipeline (Fig 2-2, 2-3). In this thesis, FBL staining is therefore used as a proxy for nucleolar visualization. However, further analysis must be completed to determine the functionality of these FBL-staining bodies as well as the extent to which the other nucleolar compartments and ribosomal DNA are present in these bodies. Regardless, we hypothesized that monitoring changes in nucleolar number/FBL staining would serve as a valid method of identifying new regulators of nucleolar function.

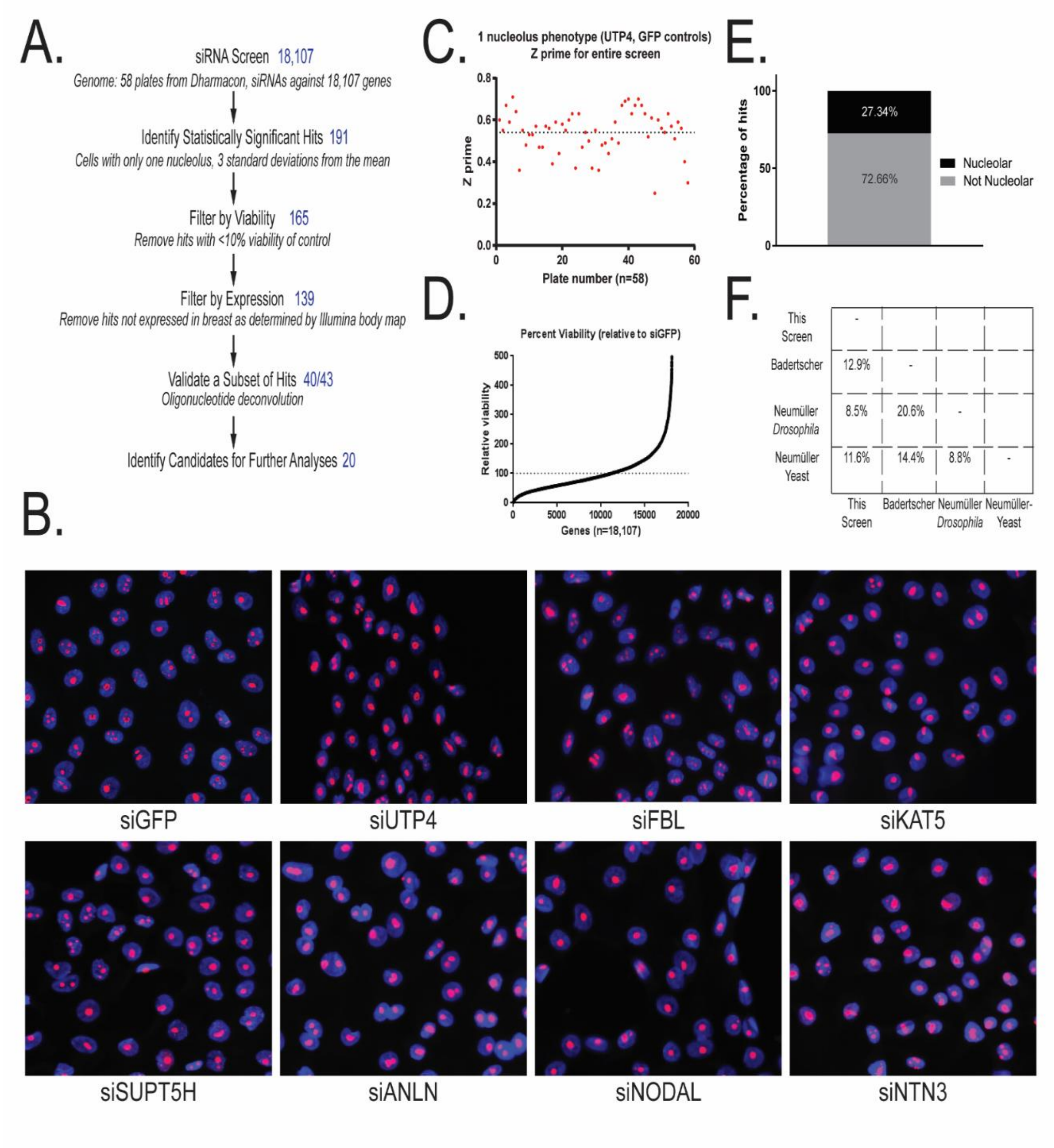

Figure 2-2. Genome-wide siRNA screen reveals 139 high-confidence regulators of nucleolar number. Figure from (Farley-Barnes et al., 2018).

A) Workflow of the screening campaign. In 384-well plates, MCF10A cells were reverse transfected with pools of siRNAs targeting 18,107 genes (each pool consisting of 4 individual siRNAs targeting a single gene). After 72 hours, the cells were stained with antibodies to fibrillarin (α-FBL; to detect nucleoli) and HOECHST 33342 (to detect nuclei). Nucleoli and nuclei were segmented and the number of nucleoli per nucleus were counted using a pipeline developed in CellProfiler (Jones et al., 2008b). Hits (191) were identified whose depletion changed the number of nucleoli from 2-3 to 1, termed the one nucleolus phenotype. These hits were filtered by viability (>10%, relative to siGFP control) and expression (FPKM>0 in Illumina Body Map (Petryszak et al., 2015)), leaving 139 high-confidence hits. Of the high-confidence hits, 40/43 were validated in a secondary oligonucleotide deconvolution screen where at least 2 of the 4 individual siRNAs gave the one nucleolus phenotype (≥50% effect). Of the validated candidates, 20 were chosen for further analyses via secondary assays.

B) Representative images from the screen showing the one nucleolus phenotype for six screen hits. The negative and positive controls are also shown (siGFP and siUTP4, respectively). Nuclei are in blue (HOECHST 33342). Nucleoli are in red (α-FBL). Histograms showing the frequency of the number of nucleoli per nucleus for the 20 candidates chosen for further analyses and the controls can be found in Figure 2-3.

C) Z’ factors for the genome-wide siRNA screen by plate relative to the siUTP4 and siGFP controls. The dashed line shows an average Z’ of 0.54.

D) Percent viability of each siRNA pool (n=18,107). Viability was assessed as the number of cells imaged over the 3 fields of view relative to the siGFP negative control. A dashed line is drawn at 100% viability.

E) Percentage of hits annotated as nucleolar (black) or not nucleolar (grey) in existing databases (Jarboui et al., 2011, Ahmad et al., 2009, Thul et al., 2017).

F) Diagram showing the percent of overlap between this screen and existing genome-wide siRNA screens for ribosome biogenesis factors (Badertscher et al., 2015, Neumuller et al., 2013) when corrected for species conservation. Badertscher refers to the screen for SSU biogenesis performed in (Badertscher et al., 2015). Neumüller *Drosophila* and Neumüller Yeast refer to the screens for nucleolar number performed in *Drosophila* and *Saccharomyces cerevisiae*, respectively, performed in (Neumuller et al., 2013).

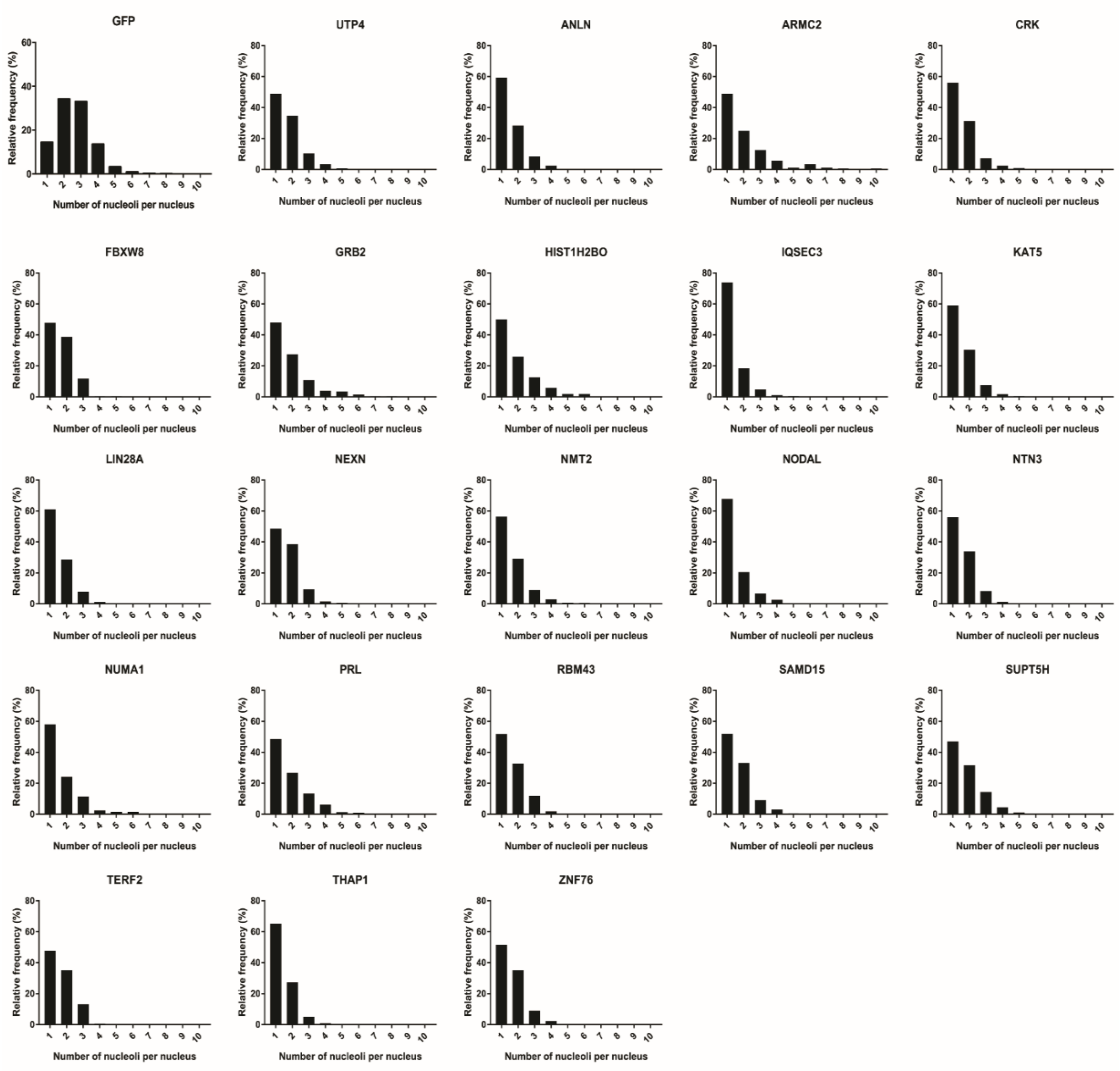

Figure 2-3. Histograms showing that the average number of nucleoli per nucleus changes from 2-3 (siGFP control) to only 1 in 20 representative high-confidence screen hits and in the siUTP4 positive control. In MCF10A cells, pools of siRNAs [indicated by HUGO Gene Nomenclature Committee (HGNC) symbol at top] were used to deplete each of the indicated hits for 72 hours. Cells were then fixed and stained with HOECHST 33342 for nuclei and with antibodies to fibrillarin to identify nucleoli. Images were acquired using the InCell 2200 Imaging System (GE Corporation). A CellProfiler (Jones et al., 2008b) pipeline was used to count the number of nucleoli per nucleus. The relative frequency of the number of nucleoli per nucleus was plotted for each siRNA target.

Histograms for the screen controls (siGFP and siUTP4) as well as the 20 high-confidence hits used in further analyses are shown here. Figure from (Farley-Barnes et al., 2018).

In the RNAi screen, the Yale Center for Molecular Discovery calculated a percent effect for the change from 2-3 nucleoli per nucleus to only 1, termed the 1 nucleolus phenotype, for each sample well. The percent effect is defined as the percentage of cells harboring 1 nucleolus normalized to the negative and positive control data from the same plate (*i.e.* mean of 16 siGFP negative control wells was set as 0% effect and the mean of 16 UTP4 positive control siRNA replicates was set as 100% effect). To monitor screen performance, coefficient of variation (CV), signal-to-background window (S/B), and Z prime (Z') statistical parameters were calculated for each screening plate using the mean and standard deviation of control samples. Dataset statistics indicate an average 6.8% CV and an average S/B of 4.7 (range 3.1-8.9). The average Z' for the screen was 0.54 (range 0.3–0.71) (Fig 2-2C). Overall, these statistics demonstrate both the low variability and the high robustness of the screen.

This siRNA campaign revealed 191 screen hits that cause the one nucleolus phenotype using a highly stringent cutoff of 3 standard deviations from the mean percent effect of the entire screening population (percent effect greater than 122%). This corresponds to a ~1% hit rate (Fig 2-2, Appendix I). In an effort to eliminate any false positives, the 191 screen hits were filtered to exclude any toxic siRNAs that conferred a viability of less than 10% of the siGFP controls (Fig 2-2A, 2-2D). Any hits not expressed in breast cells [FPKM = 0 in the Illumina Body Map (Petryszak et al., 2015)] were also removed from the dataset, leaving 139 siRNAs that gave the one nucleolus phenotype (Fig 2-2A, Appendix II).

In addition to measuring decreases in nucleolar number, information was also gathered on the extent to which depletion of the proteins increased the number of nucleoli per cell nucleus. Detailed analysis of those proteins was beyond the scope of this work. In the future, however, it will be interesting to also examine these proteins for potential roles in ribosome biogenesis. Further exploring the links between the proteins whose depletion increases vs decreases nucleolar number may provide additional insight into the mechanisms controlling nucleolar formation in humans.

Validation of the siRNA screen

The presence of nucleolar hits validates the screening approach, as it has been established that depletion of nucleolar proteins disrupts nucleolar architecture (Freed et al., 2012, Olson, 2004, Raska et al., 1990, Turner et al., 2012). Of the 139 high-confidence hits, 38 were characterized as nucleolar in 1 or more of 3 databases: the Human Protein Atlas (www.proteinatlas.org) (Thul et al., 2017), the T-cell nucleolar proteome (Jarboui et al., 2011), and the Nucleolar Proteomics Database (NOPdb) (Ahmad et al., 2009) (Fig 2-2E).

Comparison of the hit list to hit lists of existing genome-wide screens for ribosome biogenesis factors emphasizes the ability of our screening approach to discover new regulators of ribosome biogenesis. When compared to previously published genome-wide studies, I found that the screen overlap ranged from 8.5-12.9% after correcting for interspecies conservation (Fig 2-2F, Table 2-1). Although this overlap may appear low, the differences in species and tissue type, as well as the differences in screen readouts, cutoff stringency, and lack of screen saturation may contribute. Notably, this overlap was also consistent with the overlap of the previous screens to each other (Fig 2-2F, Table 2-1).

Table 2-1. List of our screen hits that were previously identified as affecting nucleolar structure in at least one of three previously published genome-wide siRNA screens (Badertscher et al., 2015, Neumuller et al., 2013). Badertscher refers to the screen for SSU biogenesis performed in (Badertscher et al., 2015). Neumüller *Drosophila* and Neumüller Yeast refer to the screens for nucleolar number performed in *Drosophila* and *Saccharomyces cerevisiae*, respectively, performed in (Neumuller et al., 2013). RPS refers to ribosomal proteins of the small subunit.

Badertscher Overlap	**Neumüller Drosophila Overlap**	**Neumüller Yeast Overlap**
AAMP	NMT2	IQSEC3
ANXA11	NOP58	KAT5
COL9A3	RPS4X	NMT2
DNTTIP2	RRN3	RRN3
MPHOSPH10	SUPT5H	SUPT5H
NOP58	SYVN1	
RPLP2	TMC2	
RPS11		
RPS13		
RPS14		
RPS16		
RPS24		
RPS28		
RPS3A		
RPS4X		
RPS5		
RPS9		
TAPBP		

A subset (43) of the 139 hits were chosen for validation by deconvolution of the pool of 4 siRNAs by Kathleen McCann, in collaboration with the Yale Center for Molecular Discovery (Fig 2-2A) (Weiss et al., 2007, Mohr et al., 2014). These hits were selected because they were undercharacterized with respect to ribosome biogenesis in the literature and/or had been implicated in human disease. Of the 43 hits, 14 are nucleolar proteins listed in existing databases and 29 are not (Ahmad et al., 2009, Jarboui et al., 2011, Thul et al., 2017). In deconvolution experiments, each of the 4 siRNAs comprising the pool used in the primary screen were tested individually. If 2 of the 4 siRNAs recapitulated the one nucleolus phenotype by having a percent effect greater than or equal to 50%, the hit was considered validated. Of the 43 tested hits, 40 were validated using this approach, resulting in a 93% validation rate (Fig 2-2A, Appendix III). I later went on to test half of the resulting 40 hits for a functional role in making ribosomes in three secondary assays (see Chapter 3).

As an additional step in validation, I performed quantitative PCR (qPCR) experiments for the siRNA pools targeting a subset of high-confidence screen hits to demonstrate that the target mRNAs are, in fact, knocked down in MCF10A cells (Fig 2-4).

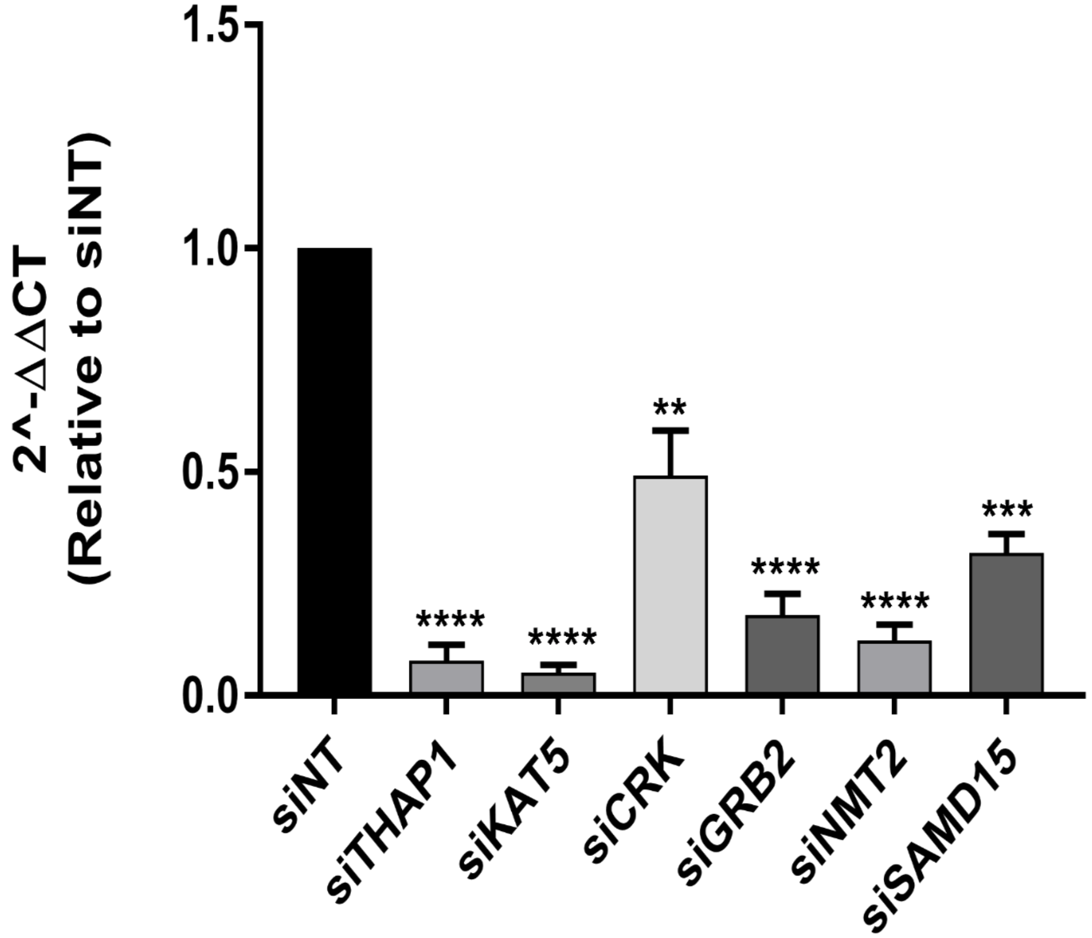

Figure 2-4. Six tested high-confidence hits show knockdown by qPCR. Data are shown as mean ± SEM, relative to siNT. N=3. Significance was calculated by One-way ANOVA with Dunnett's multiple comparisons test using GraphPad Prism. ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$. Figure from (Farley-Barnes et al., 2018).

Analysis of the 139 screen hits

To identify the functional categories associated with the one nucleolus phenotype, I explored existing algorithms that group proteins based on their GO functions and high confidence protein-protein interactions (Fig 2-5). GO analysis (database release date 12/28/2016) of the 139 hits shows significant ($p<0.05$) enrichment of 42 biological processes, including translation initiation, ribosome biogenesis, rRNA processing, rRNA metabolic process, RNA catabolic process, ribonucleoprotein complex biogenesis, translation, and RNA processing (Fig 2-5A) (Consortium, 2014). Notably, many of the top biological processes include protein targeting to the membrane or endoplasmic reticulum (ER, Fig 2-5A). This is potentially due to the enrichment of small subunit ribosomal proteins as screen hits and to the critical placement of the ribosome at the ER during the synthesis of many proteins. This STRING grouping of the 139 hits shows only one major high-confidence interaction network (interaction score > 0.7) with most of the interacting partners having known functions in ribosome biogenesis, and some proteins known to have roles in RNAPII transcription (Fig 2-5B) (Szklarczyk et al., 2015). As the hits are enriched for proteins required for ribosome biogenesis and related cellular processes, these analyses validate the screen and highlight its ability to identify regulators of ribosome biogenesis.

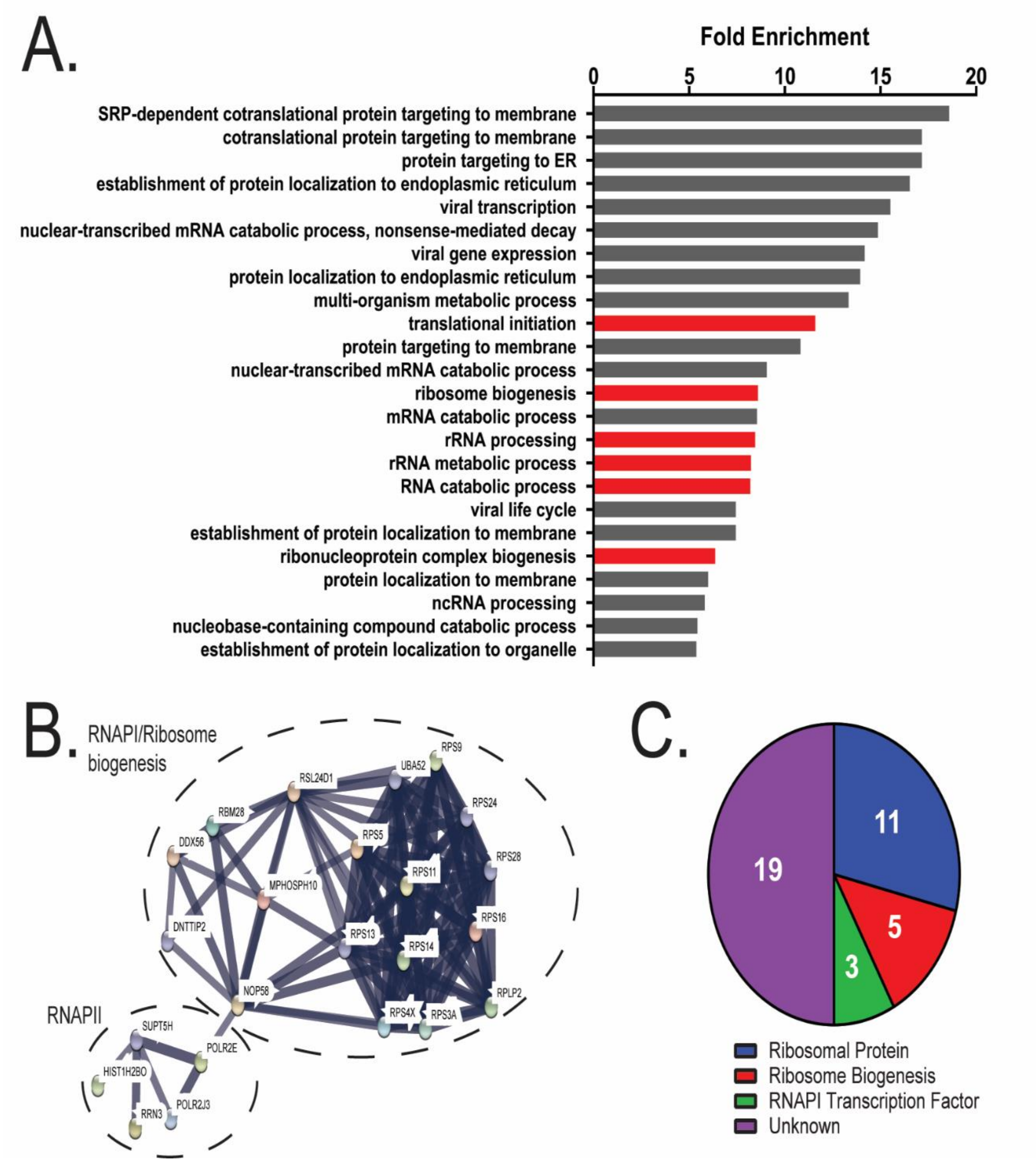

Figure 2-5. Functional analysis of the 139 high-confidence screen hits shows enrichment of ribosome biogenesis factors. Figure from (Farley-Barnes et al., 2018).

A) GO analysis (Consortium, 2014) ($p \leq 0.05$, fold enrichment ≥ 5) of the 139 screen hits reveals enrichment of 6 biological processes (red bars) related to ribosome biogenesis.

B) STRING grouping shows one large high-confidence interaction network (interaction score > 0.7) which can be further separated into two functional units (dashed circles): RNAPI/ribosome biogenesis factors and RNAPII transcription-related factors.

C) Pie chart showing that the 26 hits annotated as nucleolar in existing databases (Jarboui et al., 2011, Ahmad et al., 2009, Thul et al., 2017) can be separated into 4 functional categories: ribosomal proteins (blue), ribosome biogenesis factors (red), RNA Polymerase I (RNAPI) transcription factors (green), and unknown function (purple).

Current literature supports roles for 3 of the 38 nucleolar proteins in RNAPI transcription and 5 of the 38 nucleolar proteins in ribosome biogenesis. An additional 11 of the 38 nucleolar hits are ribosomal proteins (Fig 2-5C). This screen was therefore successful in identifying known regulators of ribosome biogenesis. For the other 19, literature searching reveals that they are undercharacterized with regard to their specific function in ribosome biogenesis.

DISCUSSION

Using our siRNA screen, we have identified, for the first time, the genetic determinants of nucleolar number in human cells genome-wide. Through examination of the hits from our genome-wide screen, I found 139 regulators of nucleolar number in human MCF10A cells. These screen hits have a wide range of known functions, including cellular development, cytoskeletal organization, proliferation, and more. Most proteins identified by this screen (69.1%) are not conserved from humans to yeast, highlighting the additional complexities of human ribosome biogenesis. The 139 screen hits include both nucleolar (27.3%) and non-nucleolar (72.7%) proteins, apparently revealing a critical contribution to the regulation and modulation of ribosome biogenesis by proteins outside of the nucleolus in human cells.

It is notable that we have found only SSU r-proteins as hits in our screen, except for RPLP2, implying that nucleolar number may be more influenced by SSU factors than LSU factors. Indeed, SSU processing defects (see Chapter 3) were the most common defects observed in my analysis of 20 screen hits. This directly contrasts with a recent siRNA screen performed only on ribosomal proteins which found that nucleolar morphology (measured by an iNo score) is more affected by depletion of LSU r-proteins than by SSU r-proteins (Nicolas et al., 2016). Such differences may be attributed to differences in screen readout and cell line used. Additionally, the iNo screen readout may not accurately reflect nucleolar function, as it has been shown that depletion of both SSU and LSU r-proteins causes clear pre-rRNA processing defects (Tafforeau et al., 2013, O'Donohue et al., 2010). It is also possible that LSU pre-rRNA processing defects have a more subtle effect on nucleolar number than SSU pre-rRNA processing defects. The screen may therefore enrich for SSU biogenesis factors because all screen hits are calculated based upon the percent effect from UTP4 depletion, which influences SSU biogenesis but not LSU biogenesis. Indeed, depletion of multiple LSU r-proteins results in a greater proportion of cells with only one nucleolus, but these effects were not large enough to fall 3 standard deviations from the mean, which was the cutoff for hits in the screen.

In addition to identifying a number of regulators of nucleolar number in humans, I have also examined how differences in nucleolar number can occur between different cell types and different organisms (Fig 2-1). I have shown here that nucleolar number is not directly related to the ploidy of the cell type (Fig 2-1), indicating the presence of additional mechanisms which control for nucleolar number. For example, both MCF10A and HeLa cells have the same average number of nucleoli per nucleus, but MCF10A cells have a normal diploid karyotype while HeLa cells are aneuploid. Also, the average number of nucleoli per cell is unchanged between multiple mammalian species such as dog, hamster, and human (Fig 2-1). However, only a single cell line of a single tissue type was tested for both the dog and hamster species. Therefore, additional studies are needed to identify differences in nucleolar number between cell lines and in primary tissue samples to help understand these additional regulatory mechanisms.

The nucleolus is a highly dynamic organelle with a complex structure which is intricately related to its primary function of ribosome biogenesis. The variations in number, size, and shape of nucleoli demonstrate the complexity of processes governing nucleolar development. Therefore, further studies are needed to elucidate the molecular mechanisms which guide the nucleolar structure/function relationship. This work is the first to identify the genetic determinants of nucleolar number on a genome-wide scale. Additionally, no genome-wide studies have yet been conducted to identify proteins which regulate nucleolar size in human cells, although this has been completed in both *Drosophila melanogaster* and *Saccharomyces cerevisiae* (Neumuller et al., 2013). Identification of the proteins involved in regulating nucleolar size and number would be the first step in elucidating the non-stochastic molecular mechanisms used by human cells to control nucleolar functions. Additionally, understanding such mechanisms could shed light on the driving role of the nucleolus in cancer progression, leading to the development of new selective therapeutics.

MATERIALS AND METHODS

Cell lines

MCF10A cells (ATCC CRL-10317) were grown in Dulbecco's modified Eagle's medium/Nutrient mixture F-12 (DMEM-F12, Gibco 1130-032) supplemented with 5% horse serum (Gibco 16050), 20 ng/mL epidermal growth factor (EGF, Sigma E4127), 0.5 µg/mL hydrocortisone (Sigma H0135), 100 ng/mL Cholera toxin (Sigma C8052), and 10 µg/mL insulin (Sigma I1882) at 37°C in a humidified incubator with 5% CO_2.

siRNA screen

In 384-well plates, MCF10A cells (3,000 cells per well) were reverse transfected with GE Healthcare Dharmacon siGENOME library SMARTpool siRNAs (20 nM final siRNA concentration per well) using Lipofectamine RNAimax. In addition to 320 library siRNAs, each screening plate contained 16 negative control wells (GFP siRNA) and 16 positive control wells (UTP4 siRNA). After 72 hours, cells were fixed with 1% paraformaldehyde in PBS for 20 min, permeabilized with 0.5% Triton X-100 for 5 min, and blocked in 10% FBS in PBS for 1 hr at room temperature before immunofluorescence staining with anti-fibrillarin antibodies (72B9 (Reimer et al., 1987a), 1:2,000 dilution) for 2 hr at room temperature and secondary Alexa Fluor 647 conjugated Goat Anti-Mouse IgG (H+L) for 1 hr at room temperature. Hoechst 33342 (1:4,000, 1 hr at room temperature) was used to detect nuclei. After washing twice with PBS, three images per well for each of the 58 library plates were acquired using the InCell 2200 Imaging System (GE Corporation). The same staining and imaging methods were also used to examine other cell lines shown in Fig 2-1.

siRNA screen image analysis and hit selection

A CellProfiler (Jones et al., 2008b) pipeline was developed for image analysis and quantitative assessment. Specifically, nuclei were segmented based on the nuclear Hoechst staining channel. Nucleoli were identified in the fibrillarin fluorescence channel using several size and intensity/contrast thresholds. Each nucleolus was mapped to a nucleus, and the image features (number nucleoli/nucleus) were quantified. Mean and standard deviation of control samples were used to calculate signal-to-background (S/B), coefficient of variation (CV), and Z prime (Z′) factors for each screening plate to ensure screening quality. The same image analysis was also used to examine nucleolar number among the cell lines shown in Fig 2-1.

The hit threshold was defined as three standard deviations above the mean of the normalized values of the entire screening population. Hits with a viability of less than 10% of the siGFP control were removed from the dataset. Hits not expressed in breast, according to Illumina Body Map data (Petryszak et al., 2015) (FPKM = 0), were also removed from the dataset. The complete list of screen hits can be found in Appendix I.

Deconvolution of siRNA pools

siRNA pool deconvolution was used to validate 43 hits from the genome-wide siRNA screen. MCF10A cells (3,000 cells per well) were transfected as above with the individual siRNAs corresponding to the selected hits (siGENOME, Dharmacon) in 384-well plates. After 72 hours, the cells were fixed, imaged, and the percent effect of each individual siRNA was calculated as in the genome-wide siRNA screen. A hit was considered validated if at least two of the four siRNAs met a minimum threshold of 50% effect.

RNAi

All siRNAs were purchased from Dharmacon (siGENOME). siNT Ca. No. D-001810-10-20 (Dharmacon).

qPCR assay

Intron-spanning primers were used whenever possible, and previously published primers were also used where such primers existed [KAT5 (Takino et al., 2016), CRK (Yue et al., 2016), and 7SL (Galiveti et al., 2010)]. Primers testing THAP1, GRB2, NMT2, and SAMD15 were developed using the online resource PrimerBank (Spandidos et al., 2010). See Table 2-2 for primer sequences.

Table 2-2. qPCR primers used, including forward and reverse primer sequences, appropriate citations or PrimerBank ID numbers (Spandidos et al., 2010).

Target Gene Name	F primer sequence (5'-3')	R primer sequence (5'-3')	Citation/Primer Bank ID
THAP1	TGCAAGAACCGCTACGACAAG	GTCTGGAGTAAAGTGCTCTGAAC	8922446a1
KAT5	CAGGACAGCTCTGATGGAATAC	AGAGGACAGGCAATGTGGTGAG	(Takino et al., 2016)
CRK	AGGGTTATCCAGAAGCGAGTC	CTTCCCACTGACCACTCACAT	(Yue et al., 2016)
GRB2	ATTCCTGCGGGACATAGAACA	GGTGACATAATTGCGGGGAAAC	156071491c2
NMT2	GGAGGCACCAAGTCAGACTC	AGCTCCATTGCTCTCTGGATA	109452600c1
SAMD15	GCCAAGGAGTCCAAGAGAGAC	CTGGATGTTTCCGACTTTACCTC	58082088c3
7SL	ATCGGGTGTCCGCACTAAGTT	CAGCACGGGAGTTTTGACCT	(Galiveti et al., 2010)

Three biological replicates, each with three technical replicates were quantified for each of the 6 tested hits as well as the non-targeting control siRNA pool (siNT) and no reverse transcriptase (RT) control reactions. All RNA used for cDNA synthesis had A_{260}/A_{230} values above 1.7. cDNA synthesis was performed using the iScript^{TM} gDNA Clear cDNA Synthesis Kit from Bio-Rad (Cat. No. 172-5035), and the SYBR Green reagent was also purchased from Bio-Rad (Cat. No. 172-5121). Melt curves were performed for each sample, ensuring the amplification of a single product. Cycling parameters using the Applied Biosystems StepOne Plus are as follows: initial denaturation at 95°C for 30s and 40 cycles of 95°C for 15s and 55°C for 30s. Melt curve analysis: 95°C for 15s, then 55°C for 1 min, and a gradual increase in temperature (0.3°/15s) to 95°C.

SAMD15 expression was low (C_T values > 35), so I utilized the SsoAdvanced PreAMP Supermix kit from Bio-Rad (Cat. No. 172-5160) according to the manufacturer's instructions to achieve appropriate C_T values for quantitation.

Analysis was completed using the comparative C_T method ($\Delta\Delta C_T$). Significance was calculated by One-way ANOVA with Dunnett's multiple comparisons test using GraphPad Prism. ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$.

Chapter 3

Analysis of 20 screen hits reveals roles for 18 in rDNA transcription, pre-rRNA processing, and/or global protein synthesis

INTRODUCTION

Our understanding of the intricacies of the regulation of human ribosome biogenesis is ever increasing. Regulation begins at the level of the rDNA with the organization and silencing of select rDNA repeats. RNAPI transcription of the rDNA is also influenced by many cellular signals. Correct processing of the pre-rRNA is controlled by a number of molecules that signal to modulate the transcription of r-proteins, snoRNPs, and other assembly and processing factors. Finally, nuclear export of the pre-ribosomal subunits can be monitored and regulated by the cell. In all, ribosome biogenesis requires all three RNA polymerases and hundreds of other factors, so there are many ways for the cell to control steps in this process. Many of the factors involved in this regulation in humans are unknown. Having identified 139 candidate regulators of nucleolar function in human cells (Chapter 2), I next set out to pinpoint the steps at which 20 screen hits affect the process of making a ribosome.

It has previously been shown that changes in the number of nucleoli can occur through multiple mechanisms that affect ribosome biogenesis. For example, in a human fibrosarcoma cell line treated with 8-chloro-cAMP, a protein kinase A agonist, the nucleoli of nondividing cells condensed from multiple nucleoli into one large nucleolus (Krystosek, 1998). This nucleolar coalescence involves a movement of the acrocentric chromosomes from being dispersed throughout the nucleus to a single central location (Krystosek, 1998). Another theory holds that the coalescence of rRNA genes is dependent on protein-protein interactions between the heterochromatin regions of the different chromosomes (Carmo-Fonseca et al., 2000). Additionally, the role of transcription in maintenance of nucleolar structure cannot be overlooked. Depletion of UBF, which is necessary for Pol I transcription *in vivo* and may also play a role in maintaining the active chromatin state, causes a coalescence of nucleoli to form one large body in mouse embryonic fibroblasts (Hamdane et al., 2014). More recent findings, including this work, suggest that there is a genetic factor in nucleolar number determination. Freed et al. showed that depletion of the ribosome biogenesis factors UTP4 and NOL11 in MCF10A cells caused a significant shift in the number of nucleoli from the average 2-3 nucleoli per cell to one nucleolus per cell (Freed et al., 2012). This is likely due to the essential functions of UTP4 and NOL11 in pre-rRNA transcription and/or processing. With the many factors that regulate nucleolar number in human cells, it is clear that nucleolar number determination is a non-stochastic process.

In this chapter, I therefore sought to identify many of the components and molecular mechanisms involved in regulating this process of nucleolar formation. To this end, I have analyzed the role of 20 high-confidence screen hits in the nucleolar functions of rDNA transcription, pre-rRNA processing, and global protein synthesis. I found that many (90%) of the protein hits are essential for completing one or more of these nucleolar functions. This validates our new screening approach to identify novel regulators of ribosome biogenesis. The results from this functional analysis reveal how multiple cellular pathways converge in the regulation of human ribosome biogenesis and pave the way for new knowledge of how ribosome biogenesis is affected in human disease.

This work was published in (Farley-Barnes et al., 2018). I performed all experiments in this chapter and wrote the first draft of the manuscript. Lisa Ogawa McLean performed the studies on 2 of the screen hits, IQSEC3 and GRB2.

RESULTS

Rationale for choosing the 20 screen hits

From the high-confidence hit list, I chose 20 hits for an in-depth analysis of their functional roles in ribosome biogenesis (Table 3-1). I chose these hits because of their novelty with respect to ribosome biogenesis, their originality compared to similar screens (only IQSEC3, KAT5, SUPT5H, and NMT2 were hits in previous screens (Table 2-1) (Badertscher et al., 2015, Neumuller et al., 2013, Tafforeau et al., 2013, Wild et al., 2010), and their disparate known roles in human development and disease. The protein hits represent a panoply of diverse cellular processes, including the chromatin state of the cells and transcription (SUPT5H, KAT5, ZNF76, HIST1H2BO, TERF2), cell division and structural organization (ANLN, NUMA1, IQSEC3), embryonic development (LIN28A, NODAL), gene expression (NTN3, THAP1), and cancer (CRK, PRL, GRB2).

Table 3-1. High-confidence screen hits chosen for further analyses. Additional studies of the PAX9 protein are discussed in Chapter 4. Table adapted from (Farley-Barnes et al., 2018).

Protein Name	HGNC Symbol	Nucleolar/ Non-nucleolar	Description
Anillin actin binding protein	ANLN	Nucleolar	Actin-binding protein with a role in cytokinesis
Armadillo repeat containing 2	ARMC2	Non-nucleolar	Armadillo repeat containing
CRK proto-oncogene, adaptor protein	CRK	Non-nucleolar	Adaptor protein involved in multiple signaling pathways, proto-oncogene
F-box and WD repeat domain containing 8	FBXW8	Non-nucleolar	F-box protein, substrate recognition component of ubiquitin protein ligase complex
Growth factor receptor bound protein 2	GRB2	Non-nucleolar	Adaptor protein that links cell receptors to the Ras signaling pathway
Histone cluster 1 H2B family member o	HIST1H2BO	Nucleolar	Core component of the nucleosome
IQ motif and Sec7 domain 3	IQSEC3	Non-nucleolar	Guanine nucleotide exchange factor
Lysine acetyltransferase 5	KAT5	Non-nucleolar	Histone acetyltransferase
Lin-28 homolog A	LIN28A	Non-nucleolar	Post-transcriptional regulator of genes needed for embryonic stem cell development, let-7 microRNA regulator
N-myristoyltransferase 2	NMT2	Non-nucleolar	N-terminal myristoyltransferase
Nodal growth differentiation factor	NODAL	Non-nucleolar	Essential for mesoderm formation and axial patterning
Netrin 3	NTN3	Non-nucleolar	May function in axon guidance during nervous system development
Nuclear mitotic apparatus protein 1	NUMA1	Nucleolar	Role in mitotic spindle formation, binds microtubules
Paired Box 9	PAX9	Non-nucleolar	RNA Polymerase II transcription factor required for craniofacial and tooth development
Prolactin	PRL	Non-nucleolar	Pituitary hormone
RNA binding motif protein 43	RBM43	Non-nucleolar	Contains RNA binding motif

Sterile alpha motif domain containing 15	SAMD15	Nucleolar	Contains sterile alpha motif
SPT5 homolog, DSIF elongation factor subunit	SUPT5H	Nucleolar	Component of DRB sensitivity-inducing factor complex, role in RNAPII transcriptional elongation and mRNA processing
Telomeric repeat binding factor 2	TERF2	Nucleolar	Binds and stabilizes telomeres
THAP domain containing 1	THAP1	Non-nucleolar	DNA-binding pro-apoptotic factor
Zinc finger protein 76	ZNF76	Nucleolar	Zinc-finger domain, transcriptional regulator

Of the 20 hits chosen, 10 are known to be nucleolar (Ahmad et al., 2009, Jarboui et al., 2011, Thul et al., 2017). Although KAT5 is not annotated as nucleolar in any of these databases, published literature supports a nucleolar localization (Halkidou et al., 2004). Additionally, LIN28A has been characterized as nucleolar, nuclear, and cytoplasmic in different cell types and at different developmental stages, but was not annotated as nucleolar in any of the databases used here (Kim et al., 2014, Chen and Carmichael, 2009, Heo et al., 2008, Balzer and Moss, 2007, Piskounova et al., 2011b, Vogt et al., 2012).

7/20 hits are required for the transcription of the rDNA

Transcription of the rDNA plays a key role in nucleolar architecture (Freed et al., 2012, Grob et al., 2014, Hamdane et al., 2014, Derenzini et al., 1998). Therefore, I investigated the effect of depletion of the 20 high-confidence screen hits on RNAPI transcription. RNAPI transcription was monitored using a dual-luciferase reporter assay that has previously been shown to be an accurate measure of rDNA transcription (Freed et al., 2012, Ghoshal et al., 2004). The ratios of firefly to *Renilla* luciferase were normalized to a control non-targeting siRNA (siNT). siRNAs targeting NOL11 were used as a positive control (Freed et al., 2012). The results show that 7 of the 20 hits are required for transcription of the rDNA by RNAPI (Fig 3-1). Notably, a decrease in rDNA transcription upon SUPT5H depletion was expected as the yeast ortholog Spt5p has previously been shown to interact with RNAPI and to associate with the rDNA (Schneider et al., 2006, Anderson et al., 2011, Leporé and Lafontaine, 2011). Additionally, while some studies have shown that both KAT5 and its yeast ortholog, Esa1, function to downregulate rDNA transcription (Koiwai et al., 2011, Clarke et al., 2006, Chang et al., 2012), our work provides evidence for the alternative hypothesis that KAT5 increases rDNA transcription (Halkidou et al., 2004). Finally, a recent paper used different assays and experimental conditions to describe NUMA1 as an enhancer of rDNA transcription, contrary to the lack of effect on transcription reported here (Jayaraman et al., 2017).

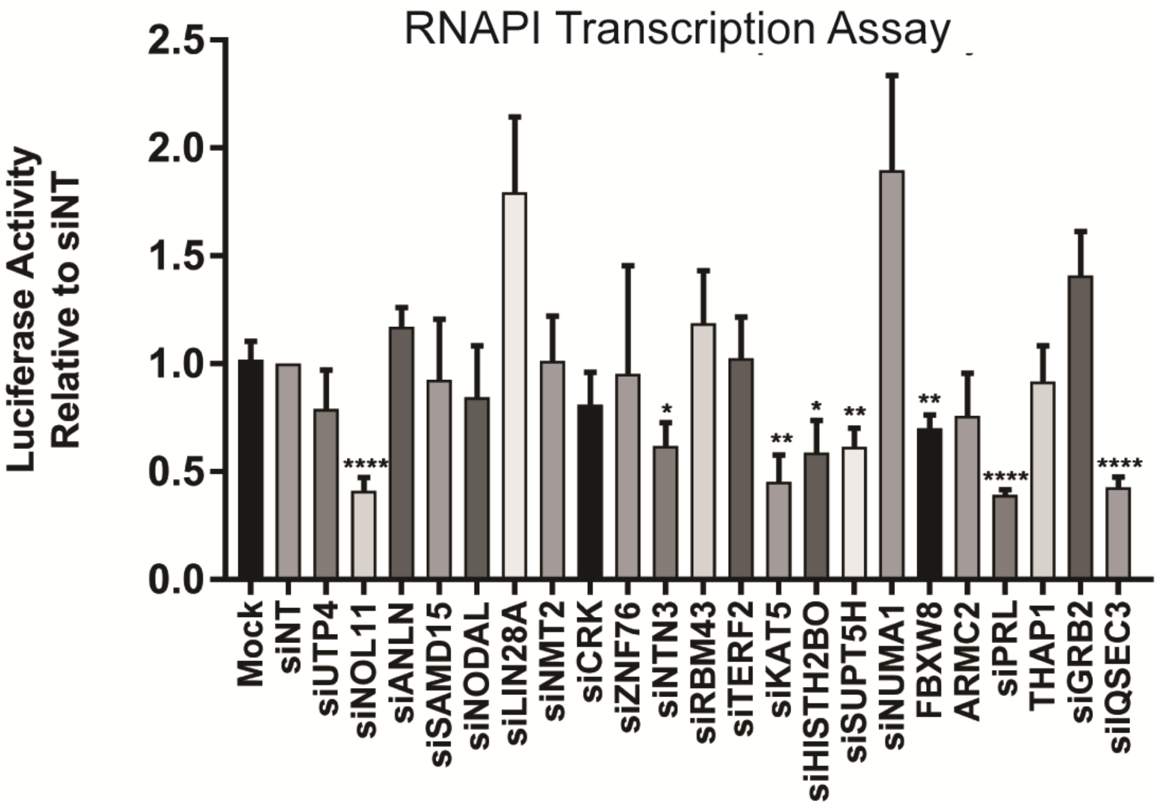

Figure 3-1. Depletion of 7/20 hits results in decreased RNAPI transcription. MCF10A cells were transfected with the indicated siRNAs. After 48 hours, two plasmids were transfected: one containing firefly luciferase under the control of the rDNA promoter (pHrD-IRES-Luc) (Ghoshal et al., 2004) and a *Renilla* luciferase transfection control (Freed et al., 2012). Twenty-four hours later, luminescence was measured using a 20/20n luminometer (Turner Biosystems) and the Dual Luciferase Reporter Assay System (Promega). The ratio of firefly to *Renilla* luciferase activity was normalized to the siNT control. Data are shown as mean ± SEM. N=3. Significance was calculated by Student's t-test using GraphPad Prism. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$.

16/20 hits are required for pre-rRNA processing

To ask whether depletion of any of the 20 hits affects pre-rRNA processing, I used northern blot analysis to detect and quantify the pre-rRNAs (Fig 1-2). After depletion of each hit, the pre-rRNA intermediates were observed via northern blotting with 6 different oligonucleotide probes (labeled 5'ETS, P1, P2, 5'ITS1, P3, and P4), each detecting different pre-rRNAs (Fig 1-2, 3-2, 3-3). The designation, 1°, indicates both the 47S and 45S pre-rRNA processing intermediates. Equal amounts of total RNA were loaded on each northern blot, and a probe for the 7SL RNA was used as a loading control. Ratio Analysis of Multiple Precursors (RAMP) (Wang et al., 2014) profiles were compiled for every processing intermediate observed by probes P1, P2, P3, and P4, allowing us to obtain a snapshot of processing defects in cells depleted of each protein hit (Fig 3-4). Of the 20 tested hits, 16 showed significantly processing defects by the RAMP analysis (Fig 3-4). These defects also largely correlated with the ratio of the intermediates relative to the 7SL (Fig 3-5). Additionally, 3 of the 4 hits without processing defects by RAMP had significantly decreased levels of almost all intermediates relative to the 7SL loading control (Fig 3-5). The pre-rRNA processing analysis was performed only in the MCF10A cell line, although targeted testing of several hits including CRK (data not shown) and PAX9 (Chapter 4) has shown that these effects are conserved among the additional human cell lines tested. With 3 biological replicates for each hit for the P1, P2, P3, and P4 probes, this quantitative RAMP analysis revealed the existence of 3 distinct patterns of pre-rRNA processing deficiencies (Fig 3-2C).

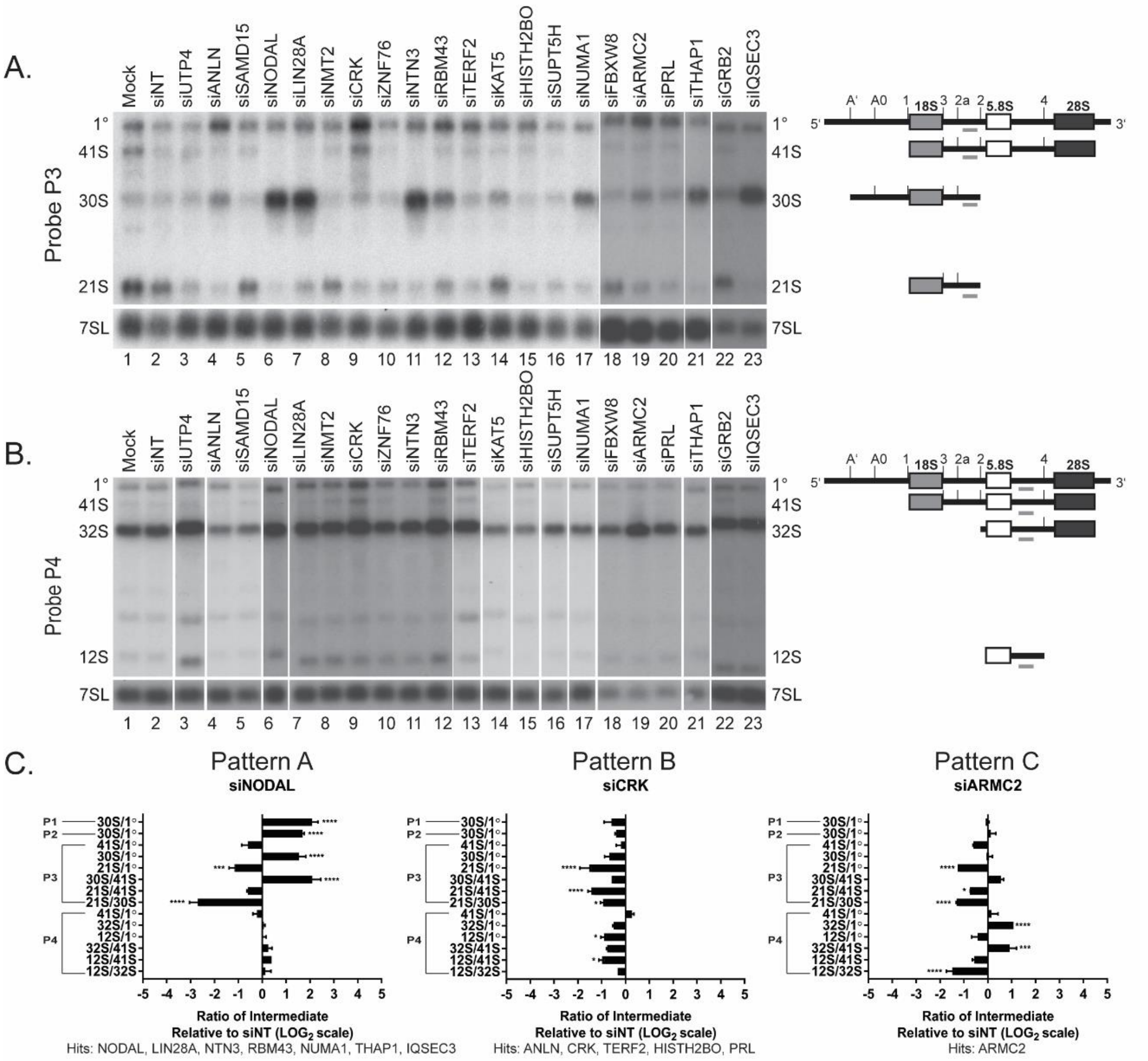

Figure 3-2. Depletion of 16/20 hits results in defective pre-rRNA processing. Figure from (Farley-Barnes et al., 2018).

A) Northern blot analysis of the 20 selected hits using a probe for the ITS1 (P3, see Fig 1-2). Pools of siRNAs targeting each protein, as well as Mock, siNT, and siUTP4 controls were transfected into MCF10A cells. Seventy-two hours later, RNA was harvested and analyzed by northern blot (probe P3). A probe for the 7SL RNA is shown as a loading control. A diagram of the pre-rRNA processing intermediates detected by probe P3 is shown to the right. 1° indicates the 47S and 45S pre-rRNA processing intermediates.

B) Northern blot analysis of the 20 selected hits using a probe for the ITS2 (P4) as in A.

C) RAMP analysis (Wang et al., 2014) reveals three major patterns of pre-rRNA processing defects among the 20 selected hits. RNAs were quantitated from the northern blots with each of the 6 probes shown in Fig 3A using a phosphorimager (Bio-Rad Personal Molecular Imager). Representative RAMP profiles for probes P1, P2, P3, and P4 are shown below each pattern, as well as a list of hits that give that pattern when siRNA-depleted. Probes used to detect each ratio are listed to the left of each RAMP profile. Data are shown as mean ± SEM, relative to siNT, on a LOG_2 scale. N=3. Significance was calculated by Two-way ANOVA with Sidak's multiple comparisons test using GraphPad Prism. * $p \leq 0.05$, **** $p \leq 0.0001$. RAMP profiles of all 20 analyzed hits, as well as controls, are shown in Fig 3-3. 1° indicates the 47S and 45S pre-rRNA processing intermediates.

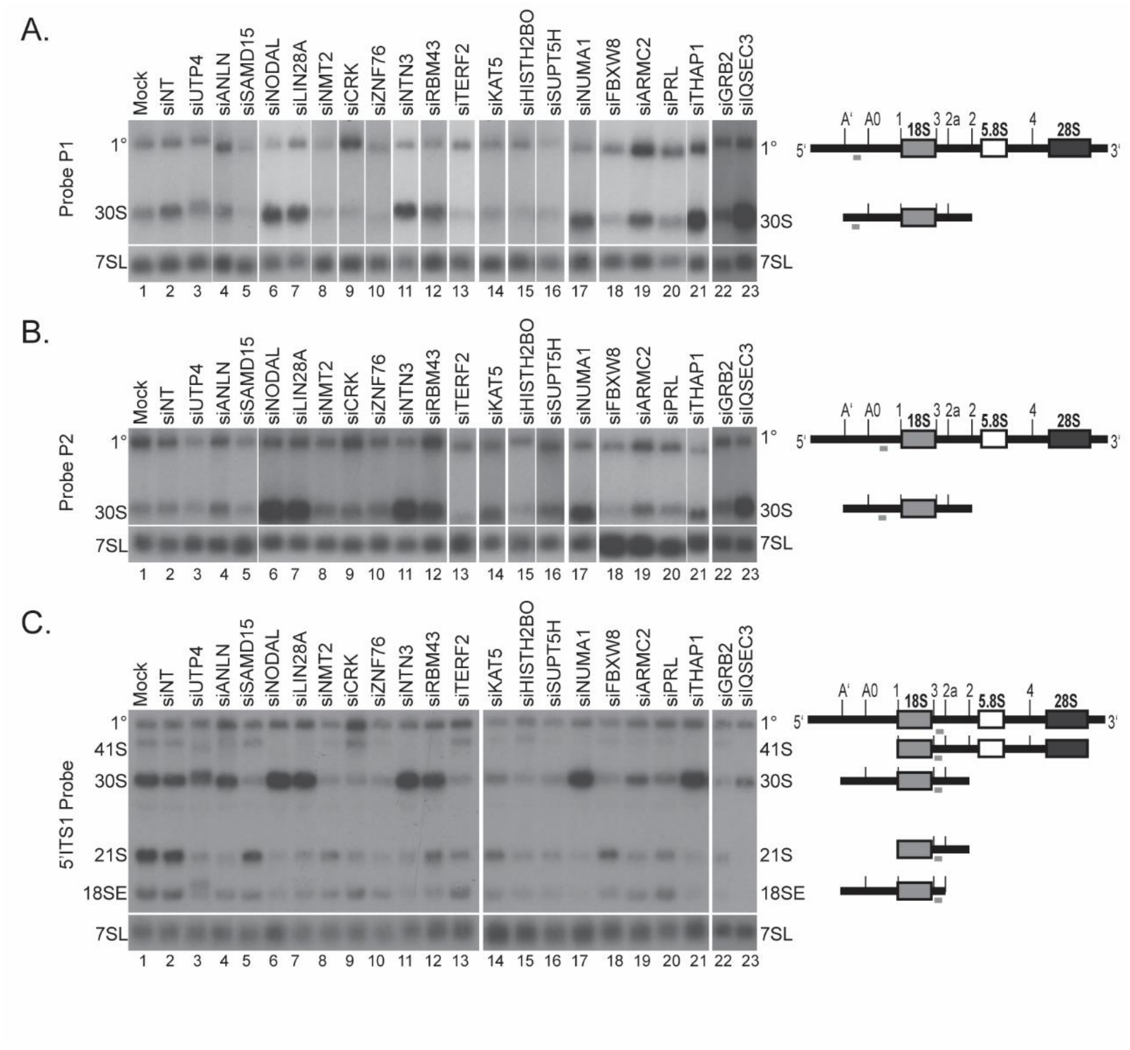

Figure 3-3. Representative northern blots for the 20 selected hits using 3 additional probes. Figure from (Farley-Barnes et al., 2018).

A) Northern blot analysis of the 20 selected hits using a probe for the 5'ETS (probe P1) as well as a 7SL loading control. A diagram of the pre-rRNA processing intermediates detected by the probe is shown to the right. Results are as described in Fig 3-2A. 1° indicates the 47S and 45S pre-rRNA processing intermediates.

B) Northern blot analysis of the 20 hits of interest using a probe for the 5'ETS (P2) as well as a 7SL loading control. A diagram of the pre-rRNA processing intermediates detected by the probe is shown to the right. Results are as described in Fig 3-2A. 1° indicates the 47S and 45S pre-rRNA processing intermediates.

C) Northern blot analysis of the 20 hits of interest using a probe for the 5’ITS1 as well as a 7SL loading control. A diagram of the pre-rRNA processing intermediates detected by the probe is shown to the right. For siGRB2 and siIQSEC3, the 7SL image is the same as that in Fig 3-2A, because the same northern blot was re-probed with the 5’ITS1 probe. 1° indicates the 47S and 45S pre-rRNA processing intermediates.

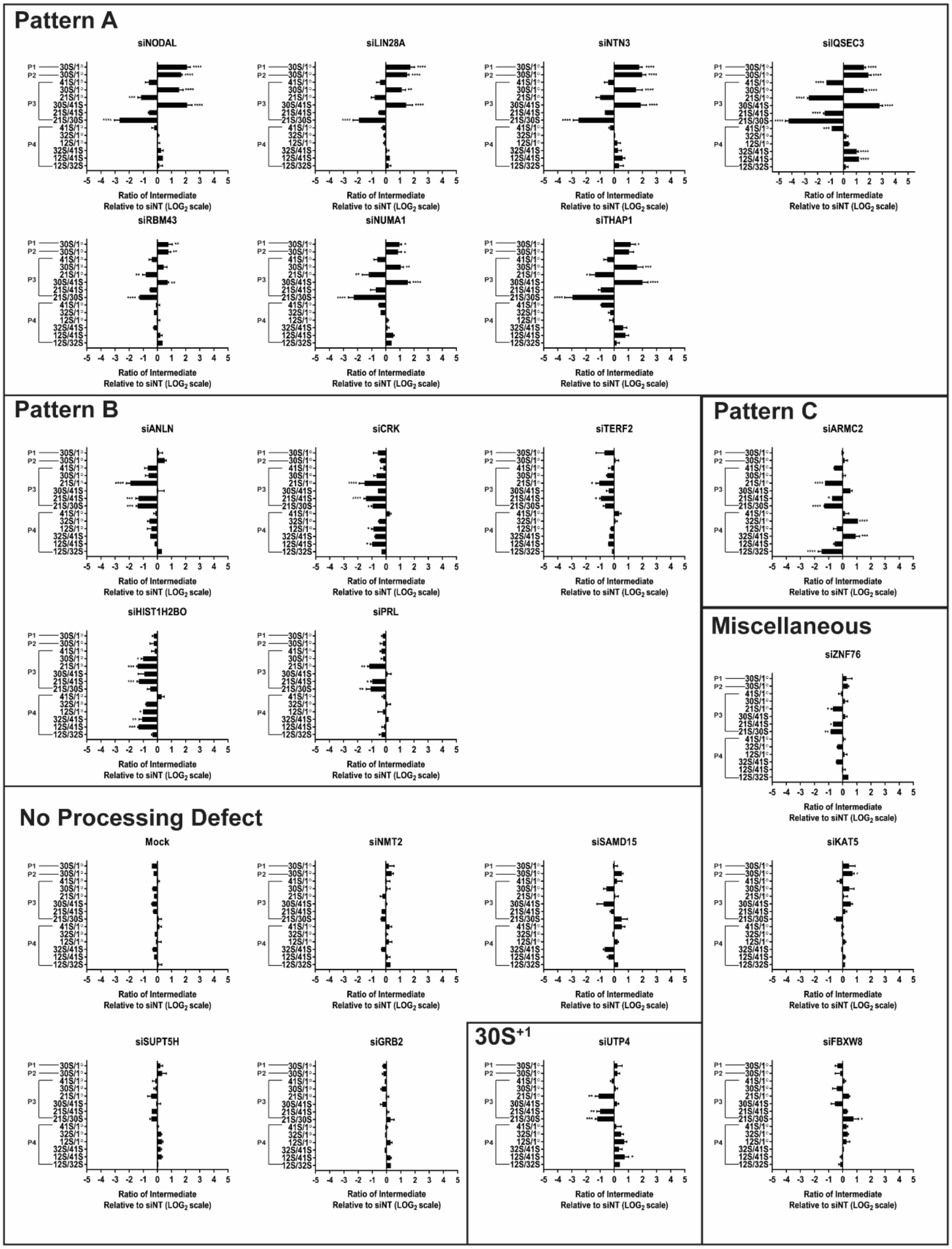

Figure 3-4. Ratio Analysis of Multiple Precursors (RAMP) quantitation of pre-rRNA levels for the 20 hits of interest and screen controls grouped by processing defect

pattern (Wang et al., 2014). For each northern blot, levels of each processing intermediate were quantitated using a phosphorimager (Bio-Rad Personal Molecular Imager). Probes used to detect each ratio are listed to the left of each RAMP profile (see Fig 1-2 for probe placement). siRNAs that gave significant pre-rRNA processing defects were grouped into 5 patterns (labeled Pattern A-C, $30S^{+1}$, or Miscellaneous). Data are shown as mean ± SEM, relative to siNT, on a LOG_2 scale. N=3. Significance was calculated by Two-way ANOVA with Sidak's multiple comparisons test using GraphPad Prism. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$. 1° indicates the 47S and 45S pre-rRNA processing intermediates. Figure from (Farley-Barnes et al., 2018).

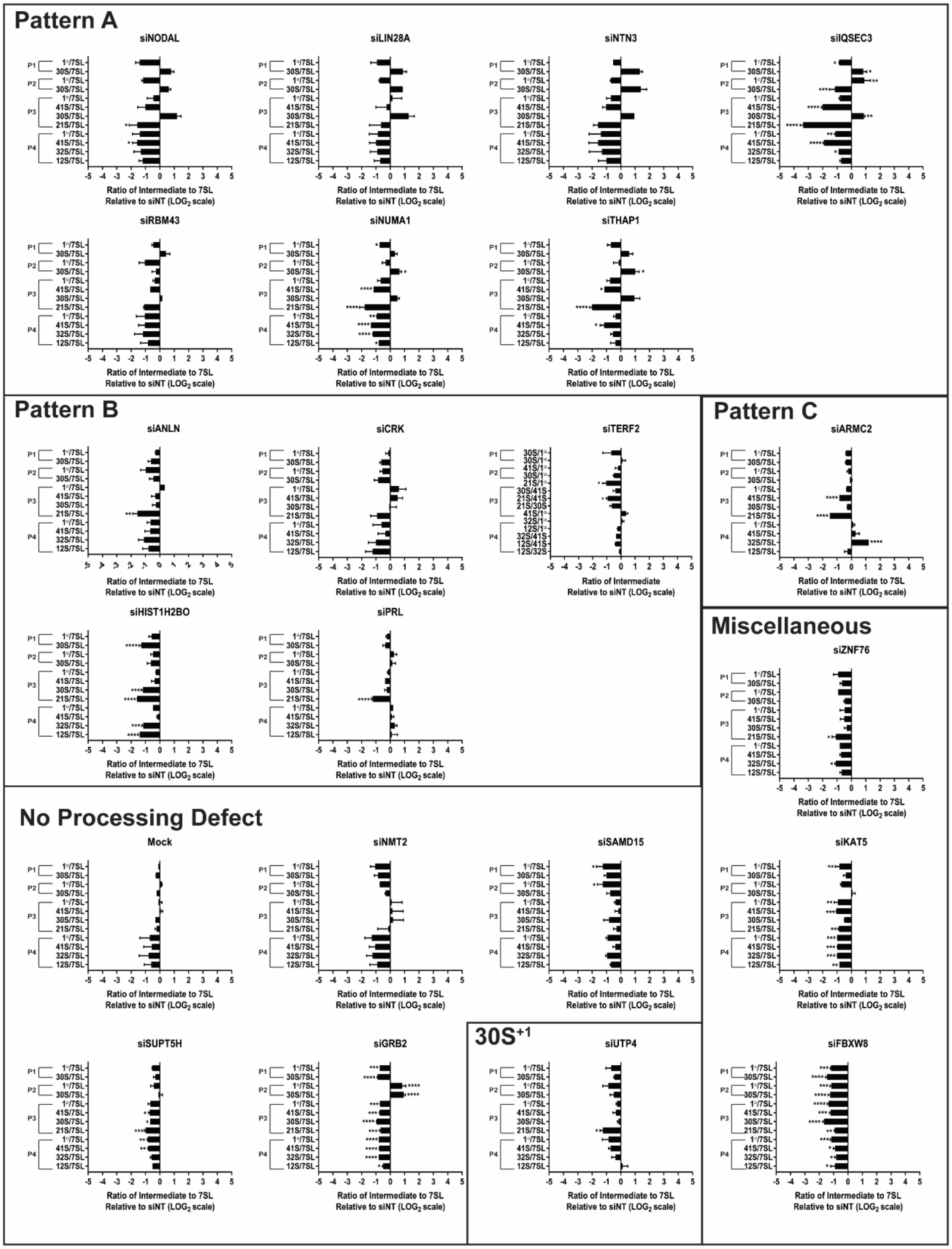

Figure 3-5. Ratio of pre-RNA processing intermediate levels relative to a 7SL control after depletion of the indicated siRNAs in MCF10A cells, grouped by the processing defect pattern shown in Fig 3-4. For each northern blot, levels of each processing

intermediate were quantitated using a phosphorimager (Bio-Rad Personal Molecular Imager). Probes used to detect each ratio are listed to the left of each RAMP profile (see Fig 1-2 for probe placement). Data are shown as mean ± SEM, relative to siNT, on a LOG_2 scale. N=3. Significance was calculated by Two-way ANOVA with Sidak's multiple comparisons test using GraphPad Prism. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$. 1° indicates the 47S and 45S pre-rRNA processing intermediates. Figure from (Farley-Barnes et al., 2018).

Pattern A was the most common pattern observed, characterized by defects in processing in the 5' External Transcribed Spacer (5'ETS) (Fig 3-2C). Pattern A is revealed by an increase in the 30S pre-rRNA and a concomitant decrease in its processing product, the 21S pre-rRNA. Hits included in this pathway are NODAL, LIN28A, NTN3, RBM43, NUMA1, THAP1, and IQSEC3 (see lanes 6, 7, 11, 12, 17, 21, and 23 of Fig 3-2A). Interestingly, this pattern has been previously seen upon the depletion of a subset of r-proteins required for 5'ETS and ITS1 processing, termed initiation ribosomal proteins of the SSU (i-RPSs) (O'Donohue et al., 2010), connecting the function of these hits to mammalian SSU biogenesis.

Pattern B involved a dramatic decrease of all intermediates relative to their precursors (Fig 3-2C). Protein hits in this category include ANLN, CRK, TERF2, HIST1H2BO, and PRL (see lanes 4, 9, 13, 15, and 20 of Fig 3-2A, B). Interestingly, I found that this pre-rRNA processing defect is not directly correlated with effects on rDNA transcription as, among these proteins, only HIST1H2BO and PRL depletion causes a transcription defect (Fig 3-1).

Pattern C includes only one analyzed hit, ARMC2, but nevertheless shows a strikingly different defect of pre-rRNA processing (Fig 3-2C and lane 19 of Fig 3-2B). In this pattern, the 32S pre-rRNA required for making the 5.8S and 28S rRNAs accumulates while the 12S pre-rRNA decreases relative to this 32S precursor. Notably, this coincides with a decrease in the 21S pre-rRNA required to make the 18S rRNA, possibly indicating a feedback mechanism at that step of the processing pathway.

Three additional hits also had minor, but significant, processing defects that did not fit into any of the above patterns: ZNF76, KAT5, and FBXW8.

Interestingly, none of the 20 tested hits showed an increase in the levels of the $30S^{+1}$ pre-rRNA, also known as the 34S pre-rRNA (Fig 1-2 and 5'ETS probe, data not shown). This stands in contrast to UTP4, which is the positive control for the siRNA screen and whose depletion had previously been shown to result in $30S^{+1}$ accumulation (Freed et al., 2012). In their study of SSU r-proteins, O'Donohue et al. also found very few changes in this transcript upon SSU r-protein depletion, leading them to postulate that the early cleavage step at site A' is uncoupled from the other steps in 18S formation (O'Donohue et al., 2010). Additionally, none of the 20 tested hits showed significant accumulation of the 18SE pre-rRNA levels relative to the primary transcript (Fig 3-3C).

14/20 hits are required for global protein synthesis

Defects in pre-rRNA transcription and/or processing are likely to lead to a reduction in protein synthesis. To measure this, I employed the use of a puromycin incorporation assay (Schmidt et al., 2009). Pulses of low doses of puromycin followed by western blotting using an anti-puromycin antibody gave a robust readout of the overall levels of protein synthesis in cells depleted of each protein hit (Fig 3-6). Of the 20 hits examined using this assay, depletion of 14 resulted in reduced levels of global protein synthesis. Expectedly, most hits whose depletion did not affect ribosome biogenesis (SAMD15 and NMT2; Fig. 3-1, 3-2) also did not affect global protein synthesis.

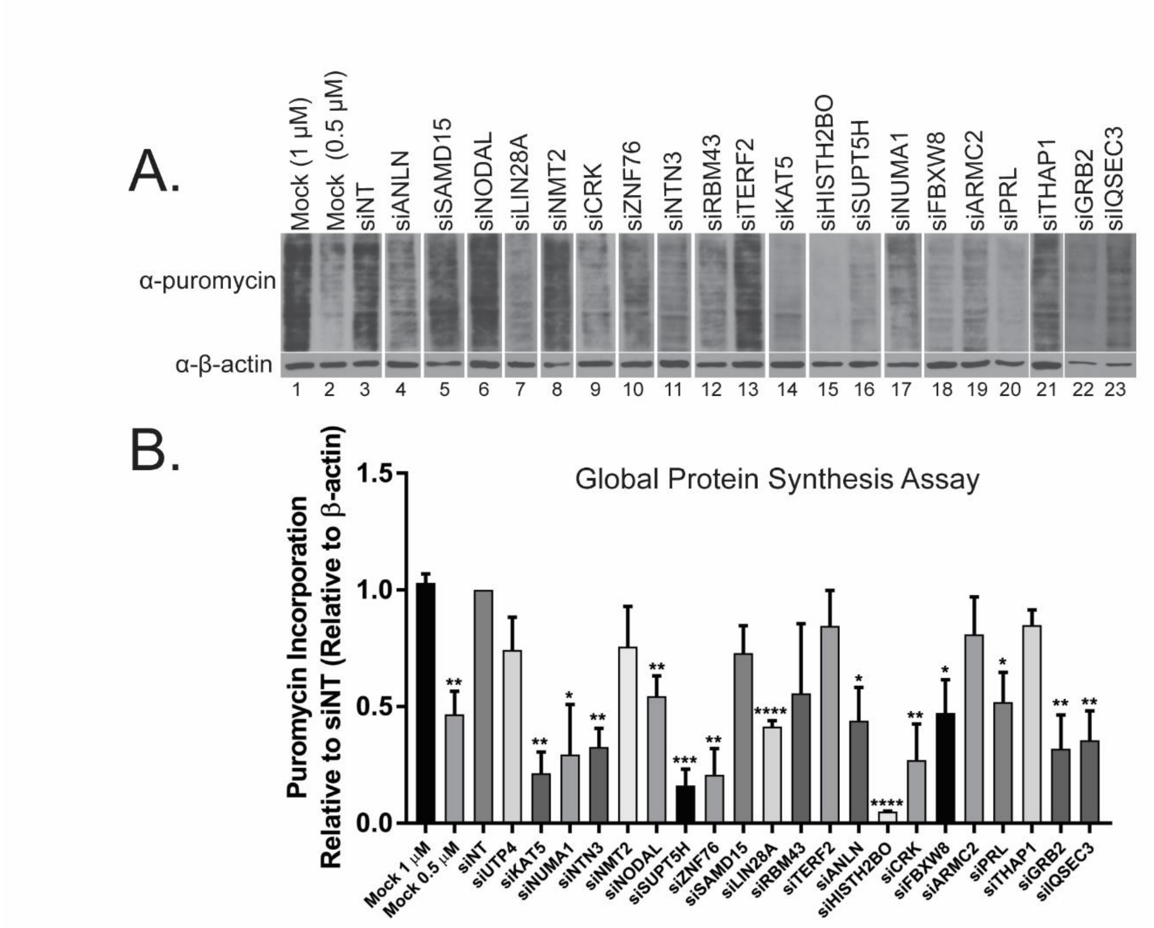

Figure 3-6. Global protein synthesis is reduced upon depletion of 14/20 hits of interest. Figure from (Farley-Barnes et al., 2018).

A) MCF10A cells were transfected with the indicated siRNAs for 72 hours. Cells were then pulsed with 1 µM puromycin (or 0.5 µM for Mock half-dose control) for 1 hour at 37°C before harvesting protein and western blotting with α-puromycin antibodies. An antibody to β-actin was used as a loading control.

B) Quantitation of global protein synthesis assay. ImageJ (Schneider et al., 2012) was used to quantify the relative puromycin incorporation for cells depleted with the indicated siRNAs, relative to siNT and the β-actin loading control. Data are shown as mean ± SEM. N=3. Significance was calculated by Student's t-test using GraphPad Prism. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$.

Ultimately, I have shown that knockdown of 18 of 20 tested screen hits results in defective rDNA transcription, pre-rRNA processing, and/or global protein synthesis (Table 3-2). For the 2 hits that did not give a phenotype in one of these assays (SAMD15 and NMT2) I performed additional qPCR experiments to demonstrate that the target mRNAs are, in fact, knocked down in MCF10A cells (Fig 2-4). Therefore, I conclude that the siRNA screen for regulators of nucleolar number in humans successfully identified novel regulators of nucleolar function.

Table 3-2. Summary of defects in pre-rRNA transcription, processing, or global protein synthesis after depletion of the 20 selected hits. Yes (Y) indicates defective transcription or global protein synthesis. Pattern A/B/C indicates the RAMP processing defect patterns shown in Fig 3-3C, 3-4.

HGNC Symbol	Transcription	Processing	Global Protein Synthesis
ANLN		Pattern B	Y
ARMC2		Pattern C	
CRK		Pattern B	Y
FBXW8	Y	Misc	Y
GRB2			Y
HIST1H2BO	Y	Pattern B	Y
IQSEC3	Y	Pattern A	Y
KAT5	Y	Misc	Y
LIN28A		Pattern A	Y
NMT2			
NODAL		Pattern A	Y
NTN3	Y	Pattern A	Y
NUMA1		Pattern A	Y
PRL	Y	Pattern B	Y
RBM43		Pattern A	
SAMD15			
SUPT5H	Y		Y
TERF2		Pattern B	
THAP1		Pattern A	
ZNF76		Misc	Y

DISCUSSION

Having identified 139 protein regulators of nucleolar number in human cells in an unbiased genome-wide screen using a stringent cut-off (see Chapter 2, Appendix II), I performed additional analyses of 20 high-confidence hits to determine their roles in ribosome biogenesis. Of the 20 representative validated hits chosen for functional analysis, I found that 7 are required for pre-rRNA transcription, 16 for pre-rRNA processing, 6 for both, and 14 to maintain normal levels of protein synthesis (Table 3-2). In all, this work integrates the varied cellular functions of the 20 screen hits with the nucleolar functions of rDNA transcription, pre-rRNA processing, and global protein synthesis.

Examination of the screen hits whose depletion decreased RNAPI transcription, gives insight into the molecular mechanisms dictating nucleolar formation in humans. For example, I have shown that HIST1H2BO plays a role in rDNA transcription, pre-rRNA processing, and global protein synthesis (Table 3-2). This is consistent with previous studies connecting several other histone variants to nucleolar organization and function [reviewed in (Schöfer and Weipoltshammer, 2018)]. Indeed, the phosphorylated form of histone variant H1.4 has been shown to colocalize with RNAPI in a UBF-dependent manner at sites of active rDNA transcription (Hamdane et al., 2014, Zheng et al., 2010). This is particularly interesting because the nucleoli in mouse ES cells coalesce into one large fibrillarin staining body upon the inactivation of UBF (Hamdane et al., 2014), and this phenotype is similar to that of the one-nucleolus phenotype seen in our screen after HIST1H2BO knockdown. Therefore, further analysis of the connection between RNAPI transcription, UBF, and nucleolar histone variants may be key to understanding the mechanisms of the one-nucleolus phenotype.

Additionally, studies of pre-rRNA processing defects after hit depletion may lend insight into the mechanisms of nucleolar formation. For example, the most common

(7/20) processing defect observed upon depletion of the 20 screen hits was Pattern A (Fig 3-2C, Table 3-2). As stated above, previous work by O’Donohue et al. has shown that depletion of i-RPS proteins results in a processing defect similar to Pattern A of this work (Fig 3-2, Table 3-2). Interestingly, all of the 10 SSU r-proteins identified as hits in our screen are designated as i-RPS, and none are classified as “progression” SSU r-proteins (p-RPS) whose depletion results in defects after the site 2 cleavage. Therefore, there may be a connection between steps in pre-rRNA processing before site 2 and nucleolar formation. However, because I have also shown that defects at other stages of ribosome biogenesis result in reductions in nucleolar number, this step in pre-rRNA processing is necessary but not sufficient for proper maintenance of nucleolar number in humans.

Ultimately, I was unable to identify a single mechanism of action through which depletion of each of these screen hits causes the one nucleolus phenotype. For example, depletion of most of the tested hits caused defects in pre-rRNA processing (Fig 3-2, 3-3, 3-4, 3-5, and Table 3-2). However, the affected points in the processing pathway differed among the hits, with some affecting pre-SSU processing, some affecting pre-LSU processing, and some affecting all intermediates. Importantly, not all steps in ribosome biogenesis were tested in these assays, including nuclear export of the ribosomal subunits. Therefore, more studies are needed to define the exact roles of each of the screen hits in both human ribosome biogenesis and nucleolar formation. Notably, depleting the vast majority of the 20 tested hits (18/20) did affect the process of making a ribosome, as was expected based on the hypothesis that nucleolar formation can be used as a readout of nucleolar function. Thus, our screen was highly successful in identifying a wide range of proteins that are functionally implicated in making ribosomes in the cell nucleolus.

MATERIALS AND METHODS

RNAi

All siRNAs were purchased from Dharmacon (siGENOME). siNT Ca. No. D-001810-10-20 (Dharmacon).

Luciferase Assays

Luciferase assays were performed as in (Freed et al., 2012). MCF10A cells were seeded at a density of 1x10^5 cells per well in 24-well plates. Twenty-four hours later, siRNAs were added (9.7 nM final) using Lipofectamine® RNAiMAX transfection reagent (ThermoFisher Scientific) as per the manufacturer's recommendations. Twenty-four hours before harvesting, the cells were transfected with 1000 ng pHrD-IRES-Luc (Ghoshal et al., 2004) and 2 ng of a *Renilla* luciferase containing plasmid (Freed et al., 2012) using Lipofectamine 2000 or 3000 (ThermoFisher Scientific) as per the manufacturer's protocol. Seventy-two hours after siRNA knockdown, luciferase activity was measured on a 20/20n luminometer (Turner Biosystems) using the Dual-Luciferase Reporter Assay System (Promega, E1910) as per the manufacturer's instructions with additional scraping to remove all cells from the plate. To control for transfection efficiency, the ratio of firefly to *Renilla* luciferase activity was calculated. Three biological replicates were performed. Significance ($p < 0.05$) was calculated by Student's t-test using GraphPad Prism, version 7.01 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com.

Northern Blotting

Northern blotting was performed as previously described (Pestov et al., 2008). MCF10A cells were seeded at 1x10^5 cells per well in 6-well tissue culture plates. Twenty-four hours later, siRNAs were added (30 nM final) using Lipofectamine® RNAiMAX transfection reagent (ThermoFisher Scientific) as per the manufacturer's recommendations. After 72 hours of depletion, RNA was harvested using TRIzol (Life Technologies 15596018) as per the manufacturer's instructions. Briefly, 3 μg of RNA was run on a 1% agarose/1.25% formaldehyde gel and transferred to a Hybond-XL membrane (GE Healthcare RPN303 S). RNA species were detected by hybridization with the following radiolabeled oligonucleotide probes:

5'ETS 5'-CCTCTCCAGCGACAGGTCGCCAGAGGACAGCGTGTCAGC-3'
P1 5'-CCCCAAGGCACGCCTCTCAGATCGCTAGAGAAGGCTTTTC-3'
P2 5'-CCACGCAAACGCGGTCGTCGGCACCGGTCACGACTCGGCA-3'
5'ITS1 5'-CCTCGCCCTCCGGGCTCCGTTAATGATC-3' (Sloan et al., 2013b)
P3 5'- AAGGGGTCTTTAAACCTCCGCGCCGGAACGCGCTAGGTAC-3'
P4 5'-CGGGAACTCGGCCCGAGCCGGCTCTCTCTTTCCCTCTCCG-3'
7SL 5'-TGCTCCGTTTCCGACCTGGGCCGGTTCACCCCTCCTT-3'.

Three biological replicates were performed for each sample. Northern blots were quantitated using a phosphorimager (Bio-Rad Personal Molecular Imager). The amount of each pre-rRNA processing intermediate was normalized to that of the non-targeting loading control. Ratio Analysis of Multiple Precursors (RAMP) was performed as previously described (Wang et al., 2014). Significance ($p < 0.05$) was calculated by Two-way ANOVA with Sidak's multiple comparisons test using GraphPad Prism, version 7.01 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com.

Puromycin incorporation assay

To assess global protein synthesis, a puromycin incorporation assay was performed as in (Schmidt et al., 2009). MCF10A cells were seeded at 1×10^5 cells per well in 6-well tissue culture plates. Twenty-four hours later, siRNAs were added (30 nM final) using Lipofectamine® RNAiMAX transfection reagent (ThermoFisher Scientific) as per the manufacturer's recommendations.

After 72 hours of depletion, puromycin (Mirus, MIR 5940) was added to the cells to a final concentration of 1 µM, 0 µM for the no puromycin control, or 0.5 µM for the half-dose control. After incubation for 1 hour at 37 °C, the cells were harvested by scraping and rinsed with PBS. Cells were lysed in 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 with vortexing. After spinning at 21000 xg for 15 minutes at 4 °C, the supernatant was removed and the protein concentration was calculated by Bradford assay (Bio-Rad).

Samples were analyzed by SDS-PAGE and puromycin incorporation was assessed by western blotting using α-puromycin antibody (1:10,000 Kerafast 3RH11) and α-β-actin as a loading control (1: 30,000 Sigma-Aldrich A1978). Secondary antibodies were α-mouse HRP conjugated (1:10,000 GE Healthcare NXA931). The amount of puromycin incorporated was determined using ImageJ 1.9 software (Schneider et al., 2012) and normalized to the β-actin control. Three biological replicates were performed for each sample. Significance ($p < 0.05$) was calculated using one-way ANOVA with Dunnett's correction for multiple comparisons in GraphPad Prism, version 7.01 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com.

Chapter 4

PAX9, a hit in the siRNA screen for nucleolar number, regulates human ribosome biogenesis

INTRODUCTION

When ribosome biogenesis is genetically disrupted in humans, a number of tissue-specific disorders called ribosomopathies arise. For example, genetic disruption of the *TCOF1*, *POLR1C*, or *POLR1D* genes in the ribosomopathy Treacher Collins syndrome results in inhibited pre-rRNA transcription (Altug Teber et al., 2004, Dauwerse et al., 2011, Bowman et al., 2012, Splendore et al., 2000). Interestingly, while these mutations each affect the global process of RNAPI transcription, patients have specific defects in craniofacial development. Treacher Collins patients present with hypoplasia of the facial bones, micrognathia with or without cleft palate, narrowing of the ear canal, and bilateral conductive hearing loss (Rovin et al., 1964, Phelps et al., 1981, Kadakia et al., 2014). Additionally, Diamond Blackfan anemia (DBA, OMIM 105650) is characterized by anemia, low reticulocyte count, and elevated erythrocyte adenosine deaminase activity (Glader et al., 1983, Halperin and Freedman, 1989). However, DBA patients also often have craniofacial anomalies and cleft palate [reviewed in (Lipton and Ellis, 2009)]. In another example, the ribosomopathy Acrofacial dysostosis, Cincinnati type (OMIM 616462) is caused by mutations in POLR1A that inhibit RNAPI activity, resulting in craniofacial defects (Weaver et al., 2015). Understanding the factors involved in human ribosome biogenesis can therefore shed light on disorders affecting of craniofacial development.

The complex development of the face is controlled by the expression of a number of proteins, including several transcription factors such as Paired Box 9 (PAX9). PAX9 belongs to a family of transcription factors that play key roles in organogenesis and neural crest cell development by controlling gene expression. In humans, mutations in PAX9 cause tooth agenesis as well as hair loss [(Stockton et al., 2000, Nieminen et al., 2001, Frazier-Bowers et al., 2002, Šerý et al., 2015, Mostowska et al., 2013, Wong et al., 2018, Daw et al., 2017, Sarkar et al., 2017) and reviewed in (Fauzi et al., 2018)].

Indeed, PAX9 mutations are the most prevalent mutation in patients with nonsyndromic tooth agenesis (Yu et al., 2018). Additionally, mice homozygous for a mutation in Pax9 have craniofacial malformations including cleft palate, skeletal abnormalities, and arrested tooth development (Peters et al., 1998). While some work has been done to identify the signaling pathways regulated by PAX9, attempts to correct the developmental defects through Wnt and EDAR signaling inhibition have been only partially successful (Jia et al., 2017a, Li et al., 2017, Jia et al., 2017b). Therefore, further studies are needed to identify all factors regulated by PAX9 in order to understand PAX9's role in craniofacial development.

This chapter ties together PAX9 and ribosome biogenesis for the first time, filling in some of the gaps in our knowledge of the many cell growth and signaling pathways influenced by PAX9 depletion. We originally identified PAX9 as a potential regulator of ribosome biogenesis in an siRNA screen for proteins required to maintain nucleolar number (Farley-Barnes et al., 2018). Further probing PAX9's specific role in making ribosomes in human tissue culture, I show here that PAX9 is required both for the pre-rRNA processing that produces the SSU 18S rRNA and for global protein synthesis. Using two genome-wide transcriptomics analyses (RNA-seq and RNAPII ChIP-seq) in human cells with and without PAX9, I was also able to define the set of mRNAs affected by PAX9 depletion. Several of the differentially expressed mRNAs were further examined to pinpoint roles for these proteins in pre-rRNA processing and global protein synthesis. In addition, this role for PAX9 is SSU biogenesis is conserved to the model organism *X. tropicalis*. These results therefore shed light on the plethora of factors whose expression is regulated by PAX9, both directly and indirectly, and opens the door to possible connections between PAX9's role in craniofacial development and human ribosome biogenesis.

I am currently preparing a manuscript reporting these results in which I will be the first author. The RNAPII ChIP-seq analysis was performed by Active Motif. The *X. tropicalis* experiments were performed by myself, Dr. John Griffin, and Dr. Engin Deniz. All other experiments in this chapter were performed by myself.

RESULTS

PAX9 depletion disrupts SSU ribosome biogenesis

In the above genome-wide siRNA screening campaign [Chapter 2 and (Farley-Barnes et al., 2018)], I identified PAX9 as one protein hit whose depletion resulted in a change in the number of nucleoli per nucleus from 2-3 to only 1 (Fig 4-1). As in (Farley-Barnes et al., 2018), cells were fixed after 72 hours of knockdown using siRNAs targeting siGFP as a negative control, siUTP4 as a positive control, or siPAX9. The MCF10A cells were then stained with an antibody to fibrillarin (FBL), a nucleolar protein, to detect nucleoli and with HOECHST 33342, to detect nuclei. A CellProfiler (Carpenter et al., 2006) pipeline was used to quantify the number of nucleoli per cell nucleus, which shifted from 2-3 in siGFP control cells to only 1 in siPAX9 and siUTP4-treated cells. As the siRNA screen served as a phenotypic readout of nucleolar function (Farley-Barnes et al., 2018), I hypothesized that PAX9 plays a role in human ribosome biogenesis through its function as a RNAPII transcription factor (Fig 4-2).

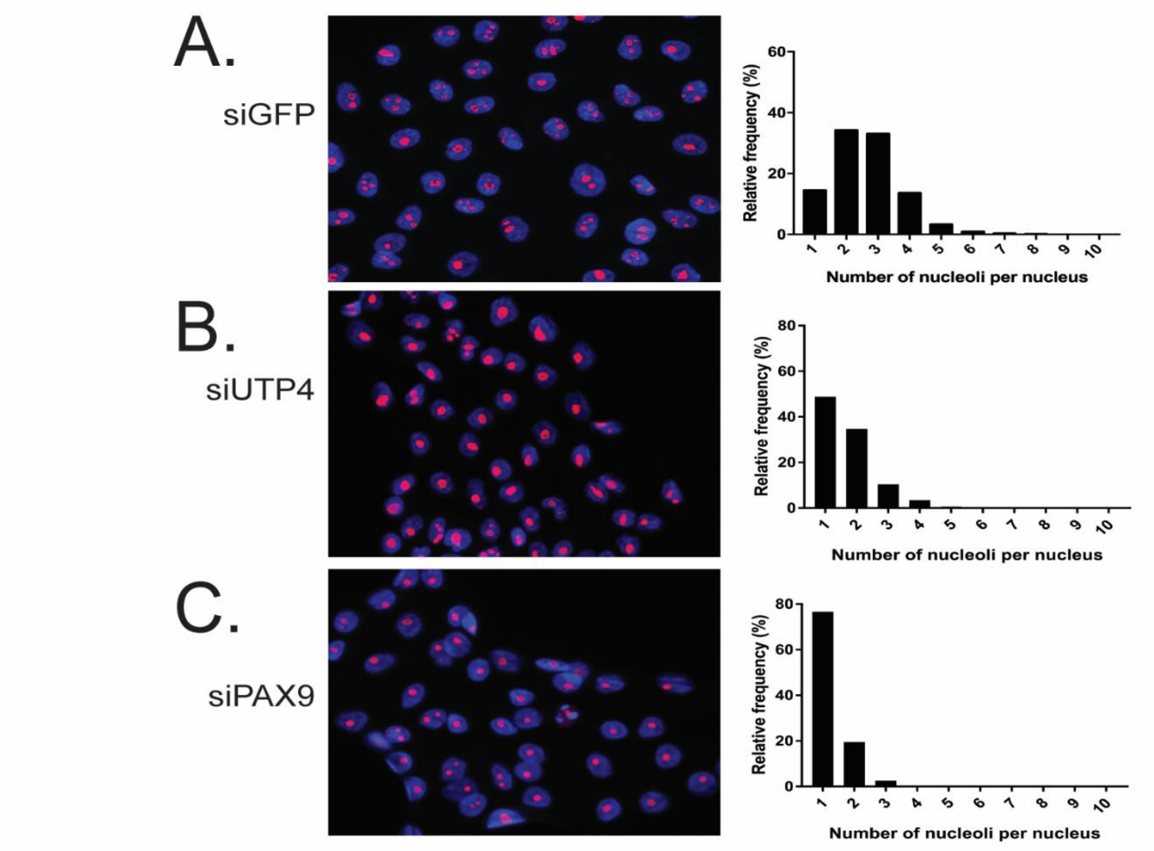

Figure 4-1. PAX9 depletion changes nucleolar number from 2-3 to only 1 in MCF10A cells. Left panel: MCF10A nuclei stained in HOECHST 33342 are shown in blue. Nucleoli are shown in pink and stained with anti-fibrillarin antibody as in (Farley-Barnes et al., 2018). Right panel: Quantitation of the number of nucleoli per nucleus for siGFP (A, 2-3 nucleoli/nucleus), siUTP4 (B, 1 nucleolus/nucleus), or siPAX9 (C, 1 nucleolus/nucleus) treated MCF10A cells after 72 hours siRNA knockdown. The data were acquired in the previously published RNAi screen (Farley-Barnes et al., 2018).

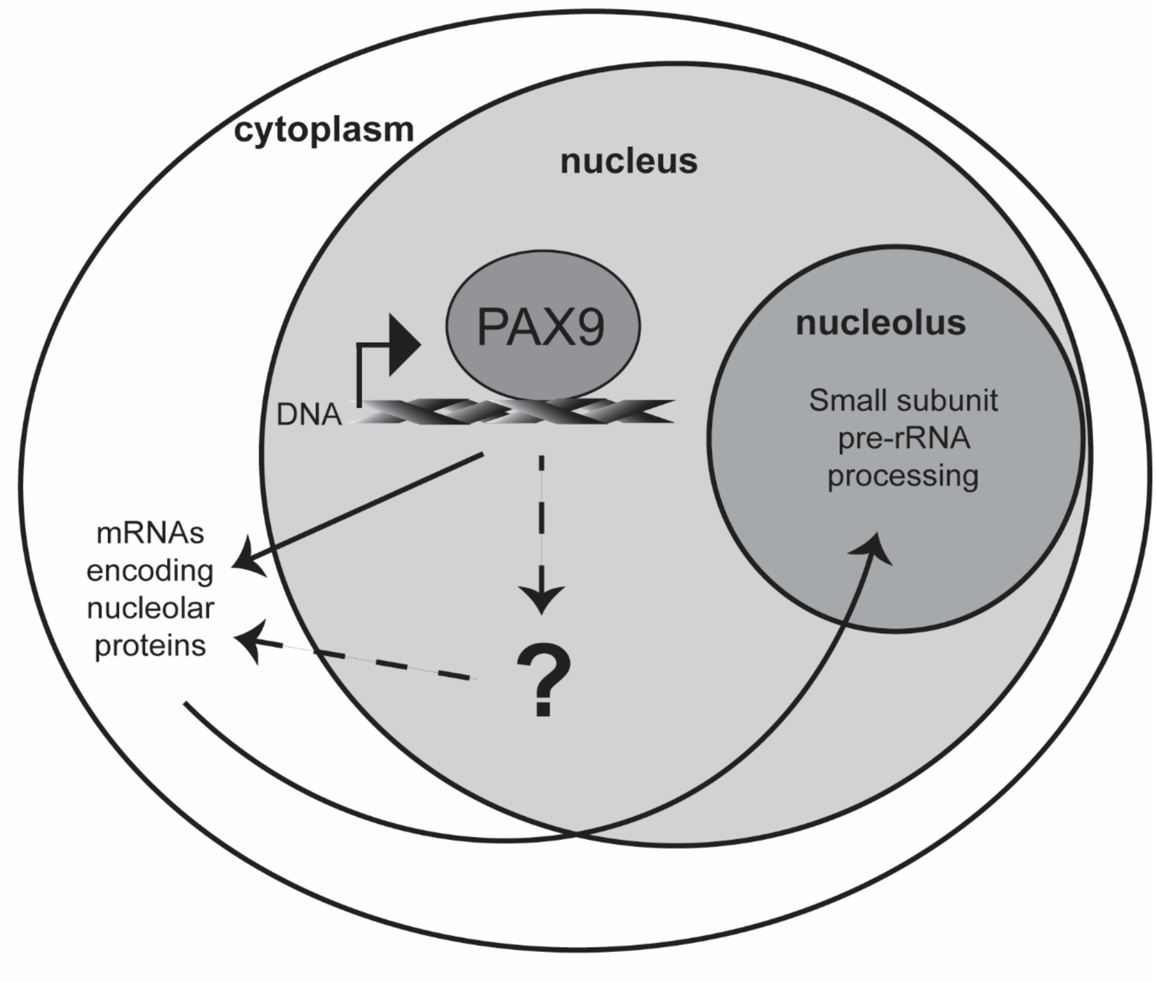

Figure 4-2. Hypothesis that PAX9 acts as a RNAPII transcription factor to influence the levels of mRNAs required for making the small subunit (SSU) of the ribosome. In the cell nucleus, PAX9 binds to DNA to affect the levels of mRNAs that encode nucleolar proteins. This action can be direct (solid arrows) with PAX9 altering the transcription of these mRNAs, or indirect (dotted arrows) with PAX9 altering the transcription of other factors which in turn alter the levels of these mRNAs. These mRNAs are translated in the cytoplasm and their corresponding proteins function in SSU pre-rRNA processing in the nucleolus.

Based on the hypothesis that PAX9 regulates nucleolar number and function through its role as an RNAPII transcription factor, I investigated the extent to which PAX9 depletion affects human ribosome biogenesis using the panel of assays performed in Chapter 3. Using a dual-luciferase reporter system previously published by our lab and others (Farley-Barnes et al., 2018, Freed et al., 2012, Ghoshal et al., 2004), I probed the role of PAX9 in human rDNA transcription. As PAX9 is a known RNAPII transcription factor, it was pertinent to test its role in RNAPI transcription as well. In this reporter assay, the ratio of firefly luciferase, which measures RNAPI transcription, was measured relative to a *Renilla* luciferase transfection control. Compared to a non-targeting control siRNA (siNT), siRNAs targeting PAX9 had no significant effect on RNAPI transcription levels after 72 hours of knockdown in MCF10A cells (Fig 4-3). A mock transfection and siRNAs targeting NOL11 were used as negative and positive controls, respectively.

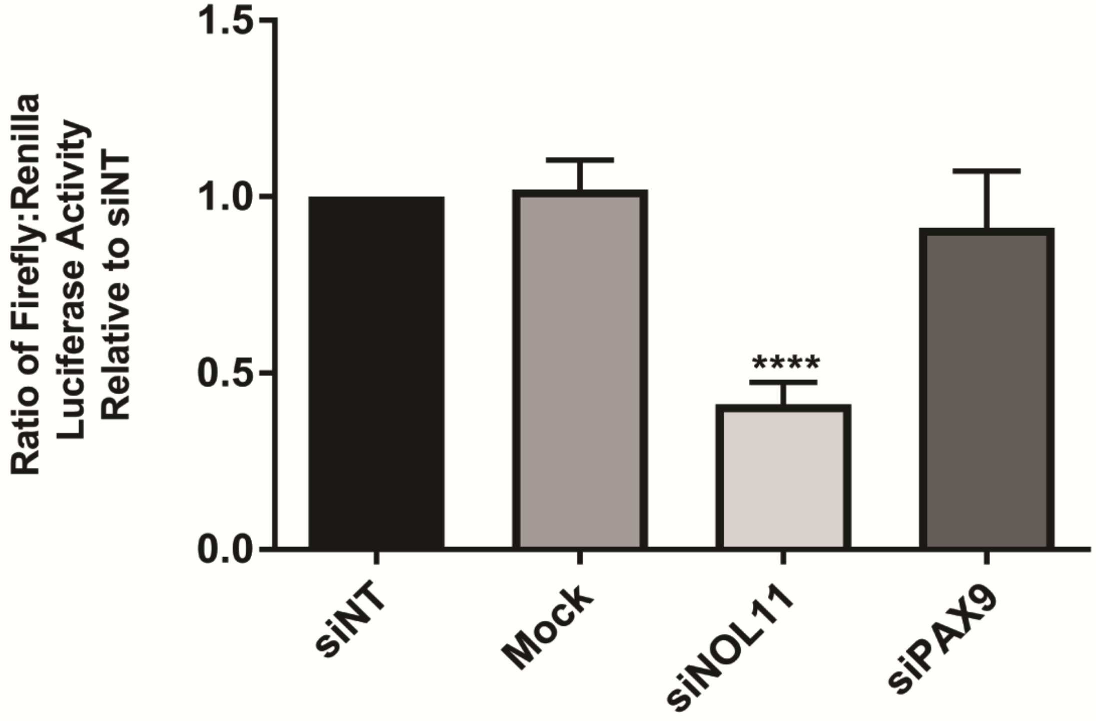

Figure 4-3. PAX9 depletion does not affect RNAPI transcription. A dual-luciferase reporter assay was used to quantify luminescence after transfection with 2 reporter plasmids, pHrD-IRES-Luc to report RNAPI transcription and a *Renilla* transfection control as in (Farley-Barnes et al., 2018). The ratio of firefly to *Renilla* luciferase was normalized to the non-targeting control siRNA (siNT). N = 4. Data were analyzed by Student’s t test using GraphPad Prism. **** $p \leq 0.0001$.

Next, I used northern blotting to observe PAX9's role in pre-rRNA processing. As in Chapter 2, I utilized 6 different probes to pinpoint any pre-rRNA processing defects occurring after PAX9 depletion in MCF10A cells (Fig 1-2). After 72 hours of knockdown, PAX9 depletion resulted in a significant increase in the 30S pre-rRNA intermediate as well as a decrease in the levels of the 21S processing product, relative to the non-targeting siRNA (siNT, Fig 4-4, and Fig 4-5). Additionally, 41S levels were decreased relative to the PTP (Fig 4-4, 4-5). Quantitation of the northern blots using RAMP (Wang et al., 2014) confirmed these results (Fig 4-4B, 4-5). These effects were also significant relative to a 7SL loading control (Fig 4-4C, 4-5). The 30S and 21S intermediates are both precursors to the 18S rRNA. Therefore, PAX9 plays a role in SSU biogenesis via pre-rRNA processing. The same pre-rRNA processing defect was also be detected in human embryonic kidney (HEK293FT) and colon carcinoma (RKO) cells, indicating that PAX9's role in ribosome biogenesis is conserved among human tissues (Fig 4-6). Notably, approximately 80% of PAX9 mRNA was depleted in MCF10A, HEK293FT, and RKO cell lines using siRNA pools targeting PAX9 (Fig 4-6). Interestingly, PAX1 depletion leads to a decrease in the 30S pre-rRNA processing intermediate and an increase in the 21S pre-rRNA processing intermediate (Fig 4-7). The defects seen upon PAX1 depletion are different than the 30S increase observed after PAX9 depletion (Fig 4-4, 4-5), even though PAX1 and PAX9 are highly similar (79%) and have similar expression patterns (Neubüser et al., 1995).

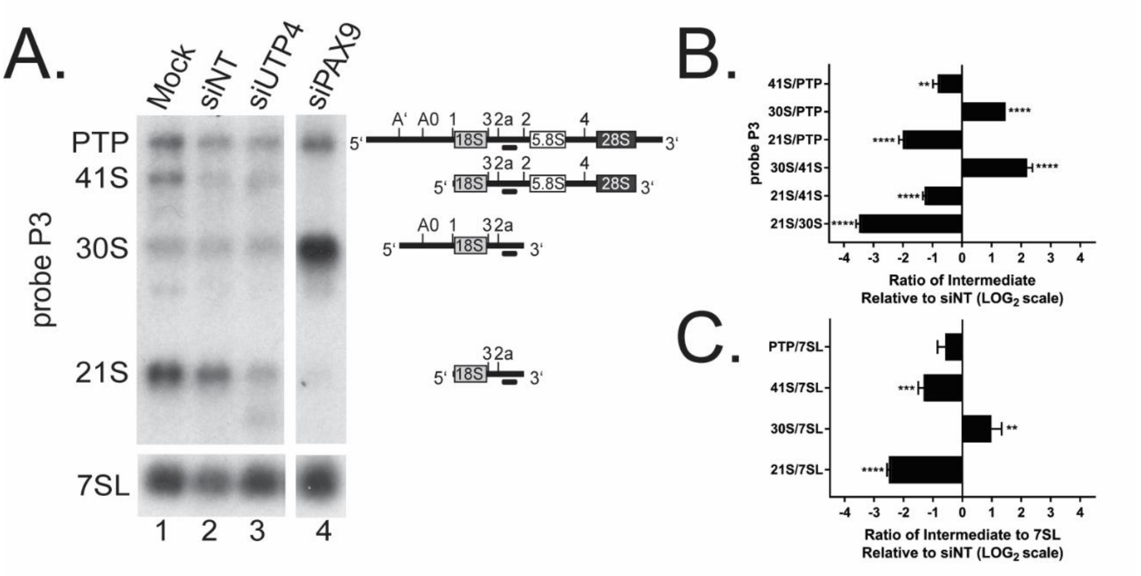

Figure 4-4. PAX9 plays a role in small subunit (SSU) pre-rRNA processing. A.) Northern blot with probe P3. A probe for the 7SL RNA was used as a loading control. Intermediates detected by probe P3 are shown to the right of the northern blot. B.) Ratio analysis of multiple precursors (RAMP) quantitation of the probe P3 northern blot. C) Quantitation of the P3 northern blots relative to the 7SL loading control. Graph is mean ± SEM. N = 3. Data were analyzed by 2-way ANOVA using GraphPad Prism. **** $p \leq 0.0001$, *** $p \leq 0.001$, ** $p \leq 0.01$, and * $p \leq 0.05$. PTP indicates the 47S, 45S, and 43S processing intermediates.

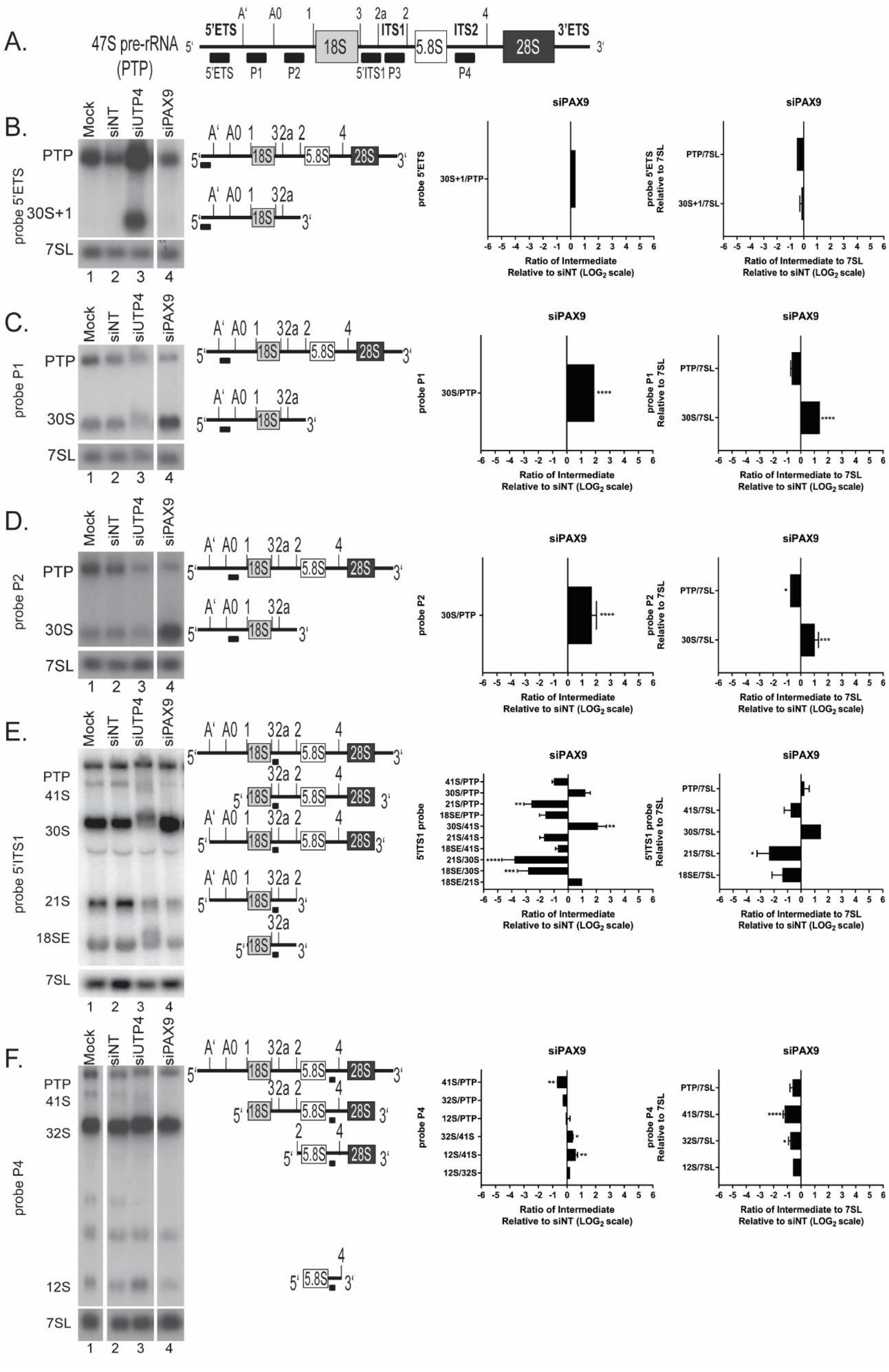
A.
47S pre-rRNA
(PTP)
5'ETS
ITS1
ITS2
3'ETS
A'
A0
1
3
2a
2
4
18S
5.8S
28S
5'ETS
P1
P2
5'ITS1
P3
P4
B.
C.
D.
E.
F.
Mock
siNT
siUTP4
siPAX9
probe 5'ETS
probe P1
probe P2
probe 5'ITS1
probe P4
PTP
30S+1
30S
41S
21S
18SE
32S
12S
7SL
1
2
3
4
siPAX9
30S+1/PTP
PTP/7SL
30S+1/7SL
30S/PTP
30S/7SL
41S/PTP
21S/PTP
18SE/PTP
30S/41S
21S/41S
18SE/41S
21S/30S
18SE/30S
18SE/21S
41S/7SL
21S/7SL
18SE/7SL
32S/PTP
12S/PTP
32S/41S
12S/41S
12S/32S
32S/7SL
12S/7SL
5'ITS1 probe
Relative to 7SL
Ratio of Intermediate
Relative to siNT (LOG2 scale)
Ratio of Intermediate to 7SL
-6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6
*
**
***
****
*****

Figure 4-5. Northern blots using other probes show small subunit (SSU) pre-rRNA processing defects after PAX9 depletion.

A. Schematic of the 47S pre-rRNA with cleavage sites indicated above. Black boxes below the pre-rRNA indicate the northern blot probes used to examine PAX9's role in pre-rRNA processing.

B. Left: Northern blot with 5'ETS probe. A probe for the 7SL RNA was used as a loading control. Intermediates detected by the 5'ETS probe are shown to the right of the northern blot. Right: Quantitation for RAMP of the 5'ETS probe (left) and 7SL (right) northern blots. Graph is mean ± SEM. N = 3. Data were analyzed by 2-way ANOVA using GraphPad Prism. PTP indicates the 47S, 45S, and 43S processing intermediates.

C. Northern blot with the P1 probe. Data shown as in (B).

D. Northern blot with the P2 probe. Data shown as in (B).

E. Northern blot with the 5'ITS1 probe. Data shown as in (B).

F. Northern blot with the P4 probe. Data shown as in (B).

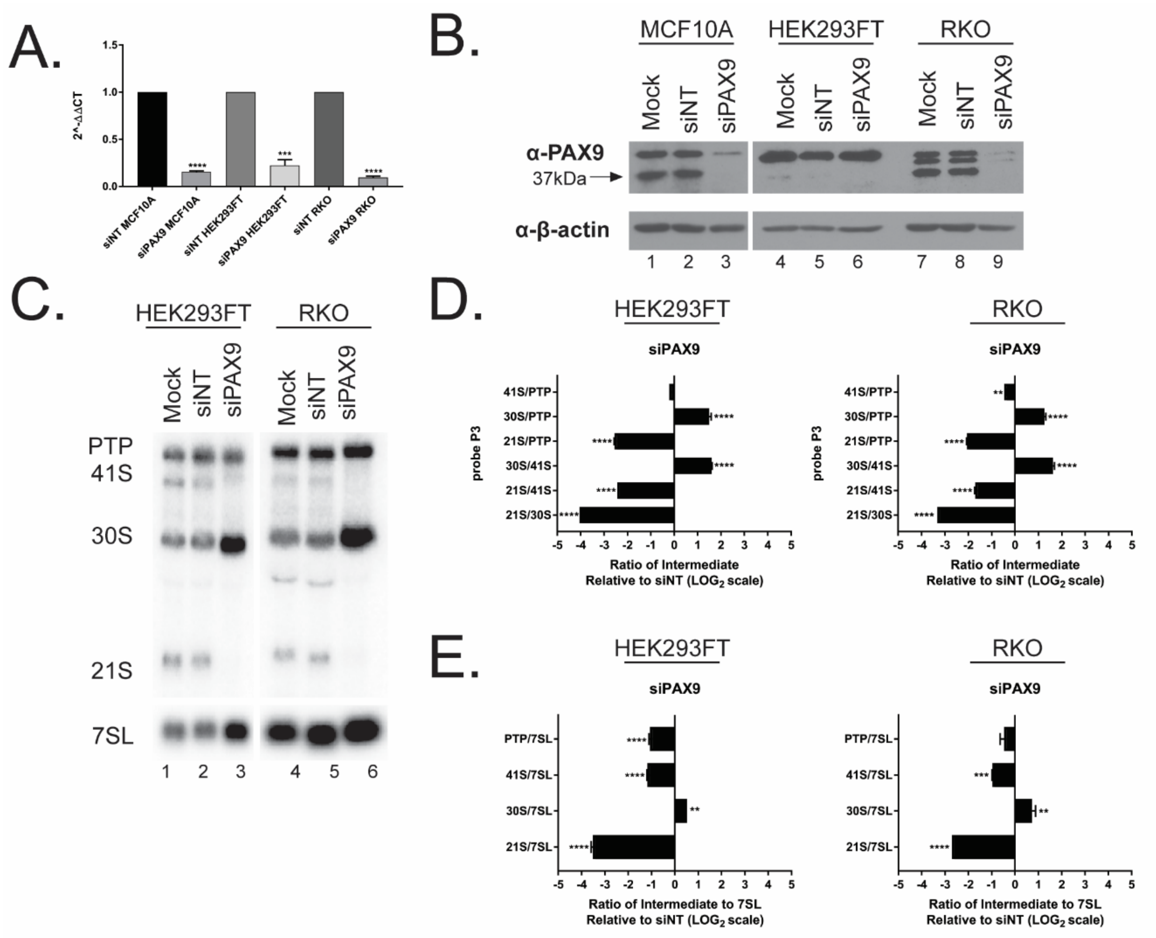

Figure 4-6. PAX9 depletion results in pre-rRNA processing defects in multiple cell lines. A. qRT-PCR confirmation of PAX9 knockdown in MCF10A, HEK294FT, and RKO cells. $2^{\wedge -\Delta\Delta Ct}$ values, relative to a siNT control and 7SL control primer, show knockdown of PAX9 by qRT-PCR using the indicated siRNAs. Data are shown as mean ± SEM. Analysis was completed by Student's t-test using GraphPad Prism where *** p ≤ 0.001 and **** p ≤ 0.0001.

B. Western blot showing depletion of PAX9 in MCF10A, HEK294FT, and RKO cells. Mock and non-targeting (siNT) siRNAs are shown as negative controls. β-actin was used as a loading control.

C. Northern blot showing depletion of PAX9 in HEK294FT and RKO cells using probe P3. A probe for the 7SL RNA was used as a loading control. Mock and siNT were used as negative controls. PTP indicates the 47S, 45S, and 43S processing intermediates.

D. Ratio analysis of multiple precursors [RAMP, (Wang et al., 2014)] data for the P3 northern blot shown in (C). N = 3. Data are shown as mean ± SEM. Significance was calculated using 2-way ANOVA in GraphPad Prism. **** $p \leq 0.0001$, *** $p \leq 0.001$, and ** $p \leq 0.01$.

E. Quantitation of the northern blot shown in (C) relative to a 7SL loading control. N = 3. Data are shown as mean ± SEM. Significance was calculated using 2-way ANOVA in GraphPad Prism. **** $p \leq 0.0001$, *** $p \leq 0.001$, and ** $p \leq 0.01$.

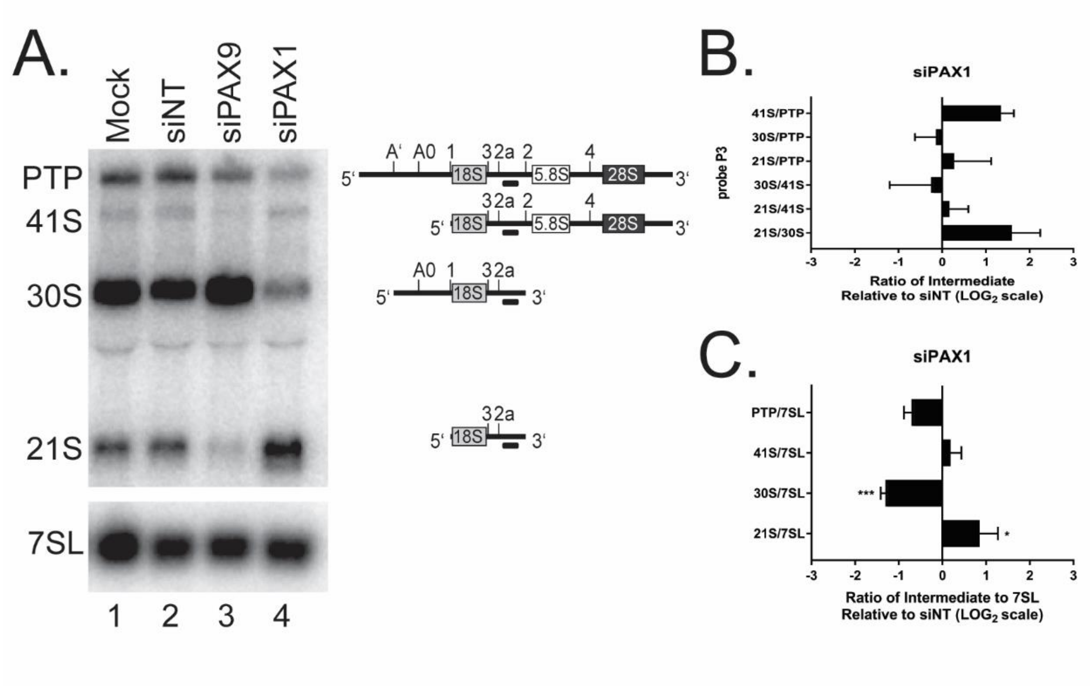

Figure 4-7. PAX1 depletion does not result in the same processing defect as PAX9 depletion.

A. Northern blot showing depletion of PAX1 in MCF10A cells using probe P3. A probe for the 7SL RNA was used as a loading control. Mock and siNT were used as negative controls. siRNAs targeting PAX9 were used as a positive control. PTP indicates the 47S, 45S, and 43S processing intermediates.

B. Ratio analysis of multiple precursors [RAMP, (Wang et al., 2014)] of the P3 northern blot shown in (A). N = 4. Data are shown as mean ± SEM. Significance was calculated using 2-way ANOVA in GraphPad Prism.

C. Quantitation of the P3 northern blot shown in (A) relative to a 7SL loading control. N = 4. Data are shown as mean ± SEM. Significance was calculated using 2-way ANOVA in GraphPad Prism. *** $p \leq 0.001$, and * $p \leq 0.05$.

Because the pre-rRNA processing defects that I observe upon PAX9 depletion indicate defects in SSU biogenesis, I next tested the extent to which PAX9 depletion affects the production of the mature 18S rRNA. Agilent BioAnalyzer quantitation shows an increase in the ratio of 28S to 18S (Fig 4-8). Combined with the northern blot results indicating defects in processing the precursors to the 18S rRNA, this result is consistent with the predicted reduction in 18S rRNA levels. Taken together, these results argue that PAX9 is required for the biogenesis of the ribosomal SSU.

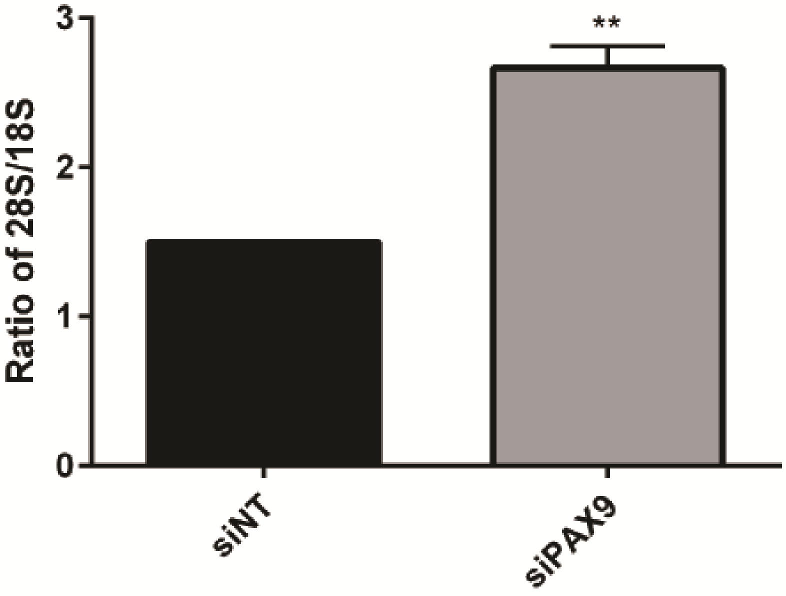

Figure 4-8. PAX9 depletion in MCF10A cells results in an increased ratio of 28S/18S by Agilent BioAnalyzer relative to a non-targeting siRNA control (siNT). Significance was calculated by Student’s t test in GraphPad Prism where ** $p \le 0.01$.

Finally, I utilized a puromycin incorporation assay to test the extent to which PAX9 depletion impacts the final product of ribosome biogenesis: global cellular translation (Fig 4-9). In this assay, western blotting for puromycin incorporated during a 1 hour pulse shows decreased protein synthesis after 72 hours of PAX9 depletion, relative to a non-targeting siRNA control (siNT) (Fig 4-9). Mock (1 µM) and Mock at a half-dose of puromycin (0.5 µM) were used as controls (Fig 4-9). These results confirm that PAX9 depletion leads to reduced ribosome function, consistent with a role for PAX9 in SSU biogenesis.

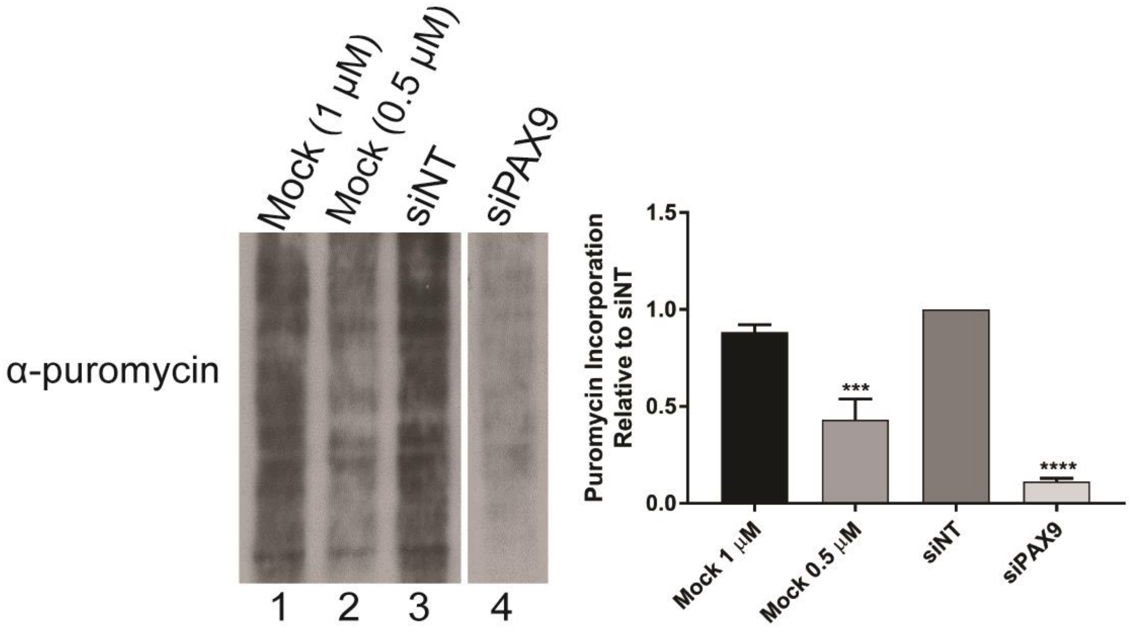

Figure 4-9. PAX9 depletion results in decreased global protein synthesis as assessed by puromycin incorporation assay. A representative western blot using an anti-puromycin antibody is shown to the left. Protein was harvested after knockdown for 72 hours using the indicated siRNAs. Quantitation of 3 biological replicates is shown to the right. Significance was calculated by One-way ANOVA in GraphPad Prism where **** $p \leq 0.0001$ and *** $p \leq 0.001$.

As disruptions in ribosome biogenesis result in interruptions in the cell cycle (Bernstein et al., 2007, Fumagalli et al., 2012), I also examined the extent to which PAX9 depletion changed the distribution of MCF10A cells within the cell cycle (Fig 4-10). Relative to a siNT control, depletion of PAX9 for 72 hours resulted in a slight increase in the proportion of cells in G1 phase of the cell cycle (80.5% vs 73.1%) and a decrease in the proportion of cells in G2/M phase (12.2% vs 19.3%) (Fig 4-10). siRNAs targeting the ribosome biogenesis factor NOL11 were used as a positive control, and depletion of this protein also resulted in an increase in the proportion of cells in G1 (Fig 4-10B). In all, these assays allowed us to conclude that PAX9 regulates human ribosome biogenesis by influencing SSU pre-rRNA processing, global protein synthesis, and movement through the cell cycle.

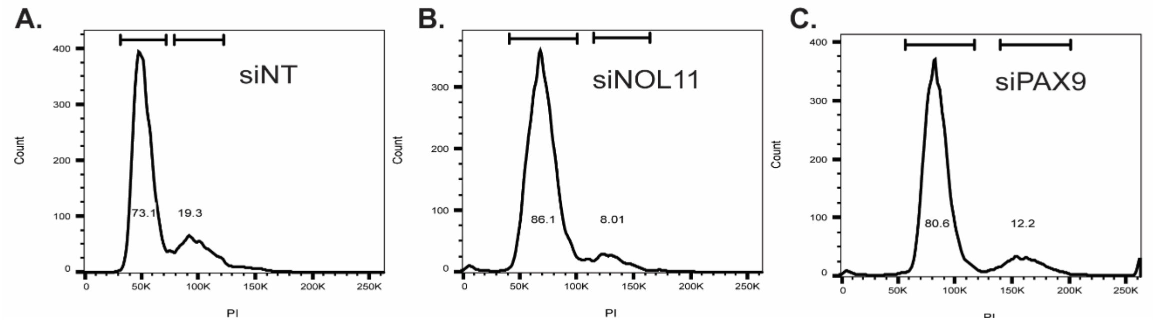

Figure 4-10. Flow cytometry shows an increased proportion of MCF10A cells in G1 phase of the cell cycle upon PAX9 knockdown. Cells were stained with propidium iodide after 72 hours knockdown with the indicated siRNAs. Live cells were analyzed using flow cytometry and the percentage of cells in G1 or G2/M was quantified as shown. N = 1.

RNA-seq analysis upon PAX9 depletion reveals decreased levels of nucleolar mRNAs responsible for SSU maturation

As PAX9 is a known RNAPII transcription factor, I hypothesized that PAX9 works in this capacity to modulate SSU biogenesis, playing an indirect role in nucleolar function (Fig 4-2). This is consistent with a nuclear but not nucleolar localization of PAX9 in 3 existing databases (Jarboui et al., 2011, Ahmad et al., 2009, Thul et al., 2017). PAX9 may act as a transcription factor for nucleolar proteins, or for proteins that affect the expression or function of nucleolar proteins. This role for PAX9 may ultimately feed back on the cell cycle as shown above (Fig 4-10). To test the hypothesis that PAX9 affects the levels of nucleolar proteins through its function as a RNAPII transcription factor, I used RNA-seq to define the set of mRNAs that were differentially expressed after PAX9 depletion. Relative to a non-targeting control siRNA, PAX9 depletion resulted in the differential expression of over 1,600 mRNAs (fold change ≥ 2 or ≤ -2, q ≤ 0.05) (Fig 4-11A). Approximately half (812) of these were reduced, consistent with the hypothesis that PAX9 acts as a transcription factor to drive their expression.

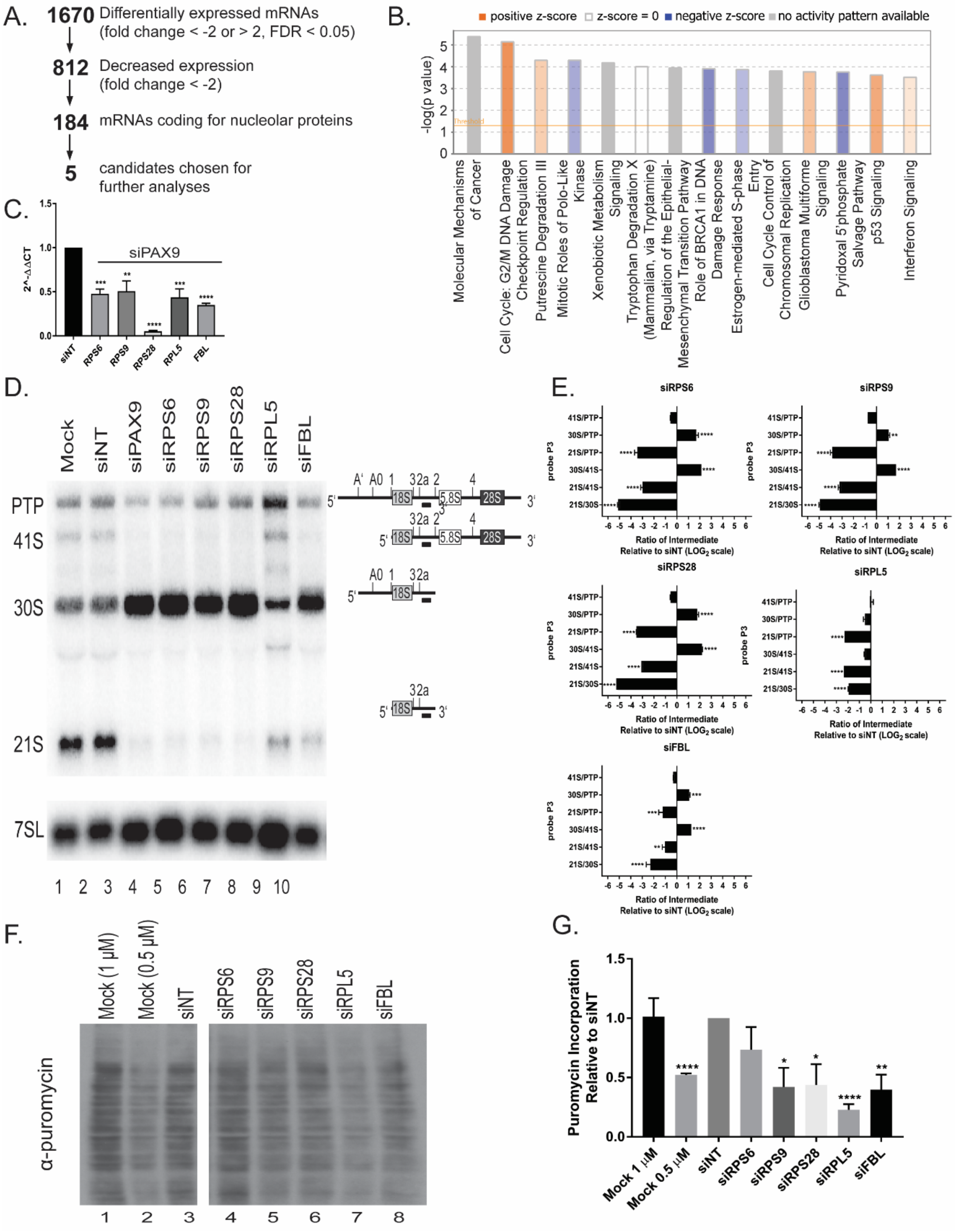

Figure 4-11. PAX9 controls the levels of several mRNAs required for ribosome biogenesis.

A. RNA-seq analysis after PAX9 depletion shows decreased levels of mRNAs encoding 184 nucleolar proteins. Relative to a non-targeting siRNA control (siNT), PAX9 depletion resulted in differential expression of 1670 mRNAs (fold change ≤ -2 or > 2 and FDR ≤ 0.05). Of these, 812 mRNAs had a decreased fold change (≤ -2) and 184 of those mRNAs code for proteins designated as nucleolar in one of three databases (Ahmad et al., 2009, Jarboui et al., 2011, Thul et al., 2017). Of the 184 mRNAs whose levels were decreased and that also code for nucleolar proteins, 5 were chosen as candidates for follow-up studies.

B. The mRNAs differentially expressed upon PAX9 depletion are enriched for genes that influence the cell cycle. The comparative pathways analysis using Ingenuity Pathways Analysis (IPA) software (QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuitypathway-analysis) reveals both upregulated (orange) and downregulated (blue) pathways enriched in the list of 1670 differentially expressed mRNAs. The comparative pathways analysis z score reflects the correlation between the observed expression of the mRNAs in each pathway and the predicted expression change based on existing literature. Pathways with a z score of 0 are shown as white and pathways with no activity pattern available in the IPA software are shown in gray. Only pathways enriched with a -log(p-value), which measures the enrichment of the pathway in the RNA-seq dataset, of ≥ 3.5 are shown.

C. qRT-PCR confirms depletion of the 5 RNA-seq candidates after PAX9 knockdown. After depletion using siRNAs targeting either PAX9 or a non-targeting control siRNA (siNT), the levels of the indicated 5 mRNAs were quantified using qRT-PCR. Data are shown as mean ± SEM. Three biological replicates, each with 3 technical replicates were performed. Significance was calculated by One-way ANOVA using GraphPad Prism where **** $p \leq 0.0001$, *** $p \leq 0.001$, and ** $p \leq 0.01$.

D. Depletion of 4 of the 5 candidate mRNAs (RPS6/eS6, RPS9/uS4, RPS28/eS28, and FBL) individually results in the same pre-rRNA processing defect as PAX9 depletion. Representative northern blot after knockdown of the indicated siRNAs using probe P3. A probe for the 7SL RNA was used as a loading control. Pre-rRNA processing intermediates detected by probe P3 are shown to the right of the northern blot. PTP indicates the 47S, 45S, and 43S processing intermediates.

E. Quantitation of three replicates of the probe P3 northern blots as shown in (D) using Ratio Analysis of Multiple Precursors [RAMP, (Wang et al., 2014)]. Graph is mean ± SEM. N = 3. Data were analyzed using 2-way ANOVA in GraphPad Prism where **** $p \leq 0.0001$, *** $p \leq 0.001$, and ** $p \leq 0.01$.

F. Depletion of some of the RNA-seq hits reduces global protein synthesis. After 72 hours of knockdown with the indicated siRNAs, MCF10A cells were pulsed with puromycin for 1 hour and protein was harvested. Western blotting with an anti-puromycin antibody was carried out. Mock at half the dose (0.5 µM) is shown as a control.

G. Quantitation of the puromycin incorporation assays as in (F). Data are shown as mean ± SEM. N = 3. Significance was calculated by Student’s t-test using GraphPad Prism where **** $p \leq 0.0001$, ** $p \leq 0.01$, and * $p \leq 0.05$.

When considered as a whole, the RNA-seq dataset is enriched for several pathways known to be regulated by PAX9. Ingenuity Pathway Analysis (Kramer et al., 2014) of the 1670 differentially expressed mRNAs reveals enrichment of both Wnt/Ca2+ signaling (differential expression of 10 pathway members, $p = 1.04 \times 10^{-2}$) (Fig 4-12), and Wnt/β-catenin signaling (differential expression of 19 pathway members, $p = 2.79 \times 10^{-2}$) (Fig 4-13). In cells depleted of PAX9, expression levels are increased for many of the mRNAs in the Wnt signaling pathway. This is consistent with previous data suggesting a role for PAX9 in the negative regulation of Wnt signaling (Jia et al., 2017a, Li et al., 2017).

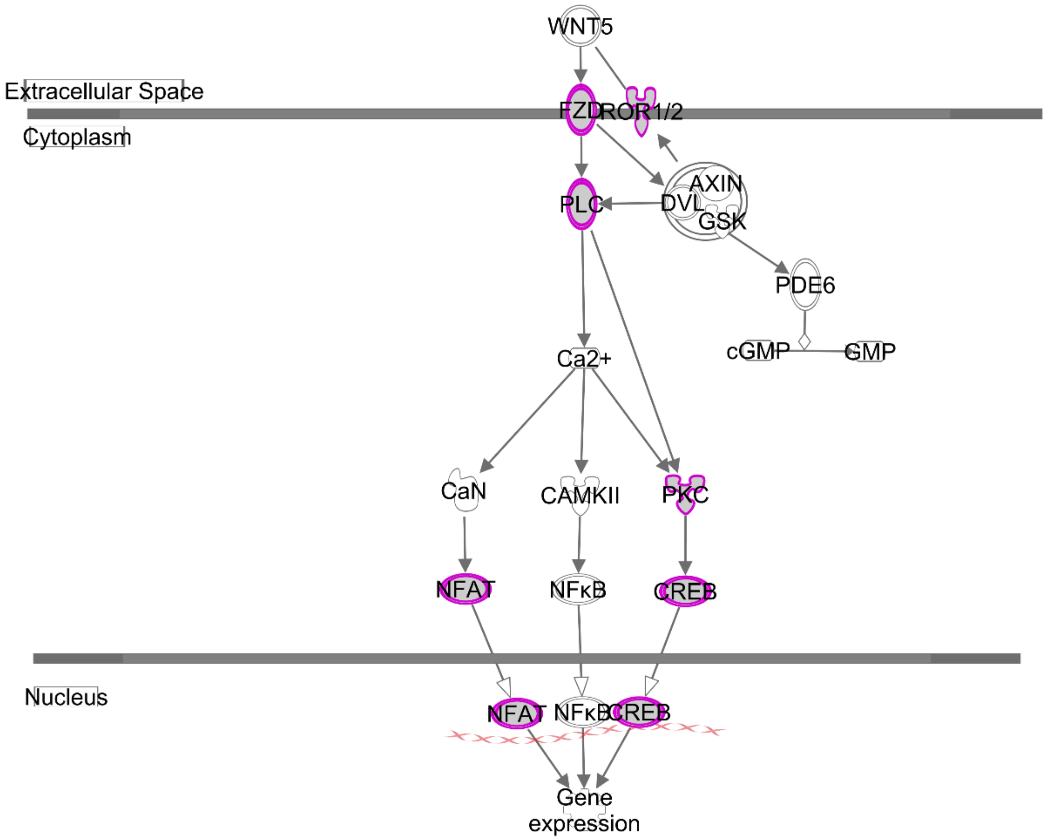

Figure 4-12. Schematic of the Wnt/Ca2+ signaling pathway. Pathway members differentially regulated (fold change ≥ 2 or ≤ -2 and FDR ≤ 0.05) after PAX9 knockdown in the RNA-seq analysis are highlighted in purple. Figure generated using IPA software (Kramer et al., 2014).

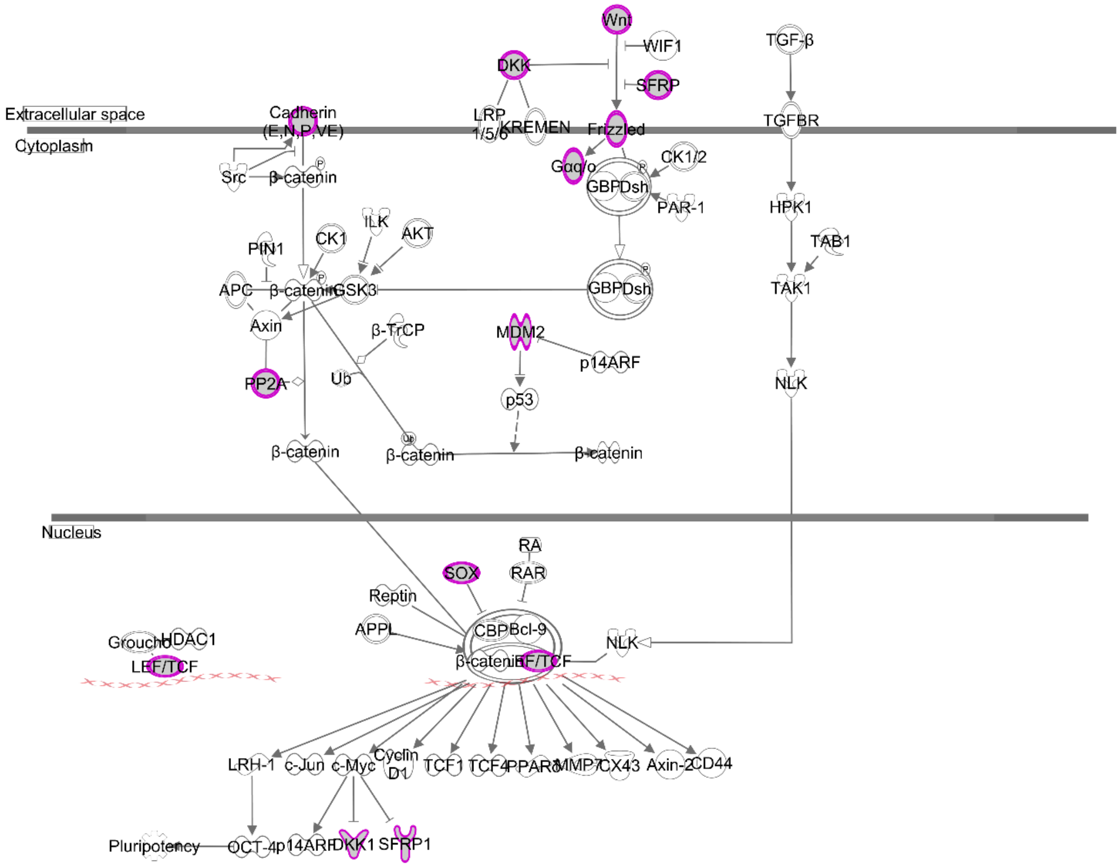

Figure 4-13. Schematic of the Wnt/β-catenin signaling pathway. Pathway members differentially regulated (fold change ≥ 2 or ≤ -2 and FDR ≤ 0.05) in the RNA-seq analysis after PAX9 knockdown are highlighted in purple. Figure generated using IPA software (Kramer et al., 2014).

Additionally validating the RNA-seq dataset, the DNA sequences 1 kb upstream of the 1670 mRNAs contain multiple potential PAX9 binding sites, making these mRNAs candidates for direct transcriptional regulation by PAX9. Scanning for the 3 known PAX9 binding sequences [5'-SGTCACGCWTGANTGMA-3', 5'-CGCGTGACCG-3' (Sivakamasundari et al., 2017) , CD19-2(A-ins) 5'-GCGTGACCA-3', and e5 5'-GCGGAACGG-3'] in the 1 kb upstream of the 1670 mRNAs using CentriMo analysis (Bailey and Machanick, 2012) reveals 3227, 2245, 2174, and 3885 potential PAX9 binding sites, respectively, in the 1 kb upstream of the 1670 mRNAs. This number of potential binding sites upstream of the 1670 mRNAs (~1.7 sites per gene) is similar to that observed in PAX9 ChIP experiments in the vertebral column of E12.5 mice [1.74 sites per gene, (Sivakamasundari et al., 2017)]. Additionally CentriMo enrichment analysis (Bailey and Machanick, 2012) of the sequence 1 kb upstream of each of the 1670 differentially expressed mRNAs reveals significant enrichment of 242 different DNA binding sequences, including the PAX3, PAX5, PAX6, and PAX7 DNA binding domains. As multiple PAX proteins can bind the same DNA sequence (Epstein et al., 1994), this provides further evidence for PAX9 regulation of these 1670 mRNAs. These analyses therefore support the hypothesis that PAX9 regulates the levels of the mRNAs identified in our RNA-seq dataset.

In the RNA-seq dataset, many of the 1670 differentially expressed mRNAs have known roles in nucleolar function. For example, 156 (9.3%) appeared in other genome-wide siRNA screens for nucleolar function (Table 4-1) (Farley-Barnes et al., 2018, Badertscher et al., 2015, Neumuller et al., 2013). Additionally, 295 (17.7%) of the 1670 differentially expressed mRNAs code for proteins designated as nucleolar in at least 1 of 3 nucleolar databases (Thul et al., 2017, Jarboui et al., 2011, Ahmad et al., 2009). Surprisingly, most of these nucleolar proteins (184/295, 62.4%) were downregulated upon PAX9 depletion. This is a significant enrichment in the expected number of nucleolar proteins, assuming that nucleolar proteins make up only 5% of the proteins in human cells (Ahmad et al., 2009). Therefore, we can conclude that PAX9 depletion alters the levels of many nucleolar mRNAs.

Table 4-1. Comparison of the RNA-seq hits to genome-wide siRNA screens for ribosome biogenesis factors [Farley-Barnes et al (Farley-Barnes et al., 2018), Badertscher et al (Badertscher et al., 2015), and Neumüller Yeast and Neumüller *Drosophila* referring to the siRNA screens in yeast and Drosophila, respectively (Neumuller et al., 2013)]. A 1 denotes a mRNA that was included as a hit in that siRNA screen, while a 0 indicates that the mRNA was not considered a hit in that screen. For the second column, a 1 denotes that the mRNA had decreased expression (fold change ≤ - 2) in the RNA-seq dataset, while a 0 indicates denotes that the mRNA had increased expression (fold change ≤ - 2) in the RNA-seq dataset.

RNA-seq all differentially expressed	Downregulated (1) or Upregulated (0)?	Human Protein Atlas	NOPdb	Gautier
TOP2A	1	1	1	1
HMGB2	1	1	1	1
NPM3	1	1	1	1
HMGA1	1	1	1	1
ALDOC	1	1	1	1
RPL22L1	1	1	1	0
KIF2C	1	0	1	1
AURKB	1	0	1	1
DEPDC1	1	1	1	0
EIF3L	1	1	1	0
PDIA6	1	0	1	1
RPL3	1	1	1	0
PSIP1	1	0	1	1
ANXA6	1	0	1	1
RSL24D1	1	1	1	0
RPL10A	1	1	1	0
CPS1	1	1	1	0
MKI67	1	1	1	0
NUSAP1	1	1	1	0
MCM2	1	0	1	1
SNRPA	1	0	1	1
LMNB1	1	0	1	1
NLE1	1	1	1	0
MCM6	1	0	1	1
MCM5	1	0	1	1

FBL	1	1	1	0
NPM1	1	1	1	0
HMGA2	1	1	1	0
NCAPD2	1	1	1	0
ERMP1	1	1	1	0
ZNRD1	1	1	1	0
CENPH	1	1	1	0
CBX5	1	0	1	1
SMC4	1	0	1	1
FANCD2	1	1	1	0
RPL13A	1	1	1	0
CDK6	1	0	1	1
RAN	1	0	1	1
NASP	1	0	1	1
DEK	1	0	1	1
HMGB3	1	1	0	1
N4BP1	0	1	0	1
PGAM2	0	1	0	1
HIST1H1C	0	1	1	0
HECW2	0	1	1	0
PHGDH	1	0	1	0
CBS	1	0	1	0
ARNT2	1	0	1	0
ALDH1L2	1	0	1	0
TBL1X	1	1	0	0
VIM	1	0	1	0
CCNA1	1	0	1	0
TRIB3	1	1	0	0
TTK	1	1	0	0
VCAN	1	0	1	0
CCNB2	1	0	1	0
PCK2	1	0	1	0
NUP210	1	0	1	0
SNX10	1	1	0	0
MRC2	1	0	1	0
KIF18B	1	0	1	0
SHMT2	1	0	1	0
MNX1	1	1	0	0
TRIM47	1	0	1	0
SERPINE2	1	0	1	0
KIF4A	1	0	0	1
HJURP	1	1	0	0
CCNB1	1	0	1	0

MAD2L1	1	0	1	0
PDK3	1	1	0	0
MYBL1	1	1	0	0
CDCA8	1	1	0	0
RNF125	1	1	0	0
CCDC34	1	1	0	0
RPL23	1	0	1	0
CDK1	1	0	1	0
RFC3	1	0	1	0
PYCR1	1	0	1	0
TAGLN2	1	0	1	0
MTHFD2	1	0	1	0
FBXO36	1	1	0	0
BRCA2	1	0	0	1
RPL12	1	0	1	0
GGT5	1	1	0	0
FHL1	1	0	1	0
KIFC1	1	0	1	0
USP13	1	0	1	0
EXO1	1	0	0	1
LMO4	1	0	1	0
KIF11	1	0	1	0
ADTRP	1	1	0	0
LRP4	1	1	0	0
RPS9	1	0	1	0
TIFA	1	1	0	0
NOB1	1	0	1	0
ARHGAP11A	1	1	0	0
RAD51AP1	1	0	1	0
IQGAP3	1	1	0	0
FAM64A	1	1	0	0
AURKA	1	0	1	0
OSMR	1	1	0	0
TPX2	1	0	0	1
TUB	1	1	0	0
HNRNPA1	1	0	1	0
EPHB6	1	0	1	0
PDE4D	1	0	1	0
HMGN2	1	1	0	0
RPS2	1	0	1	0
BLM	1	0	1	0
ATP8B2	1	1	0	0
FKBP10	1	0	1	0

SERPINH1	1	0	1	0
CBFB	1	0	1	0
HNRNPA1L2	1	0	1	0
FANCG	1	1	0	0
FPR1	1	1	0	0
LTBP1	1	1	0	0
GARS	1	0	1	0
SPC24	1	1	0	0
EIF2A	1	0	1	0
PTMS	1	1	0	0
SNRPB2	1	1	0	0
H2AFZ	1	0	1	0
TSEN15	1	1	0	0
RBMX	1	0	1	0
ANP32B	1	0	1	0
KXD1	1	1	0	0
RPS6	1	0	1	0
HELLS	1	0	0	1
RRM1	1	0	1	0
NCAPG2	1	0	1	0
ZNF367	1	0	1	0
EIF2S3	1	0	1	0
SFRP1	1	1	0	0
FGFR3	1	0	1	0
GMFB	1	0	1	0
RPSAP58	1	0	1	0
SAPCD2	1	1	0	0
ACTR3B	1	0	1	0
MRPS27	1	0	1	0
TMPO	1	0	1	0
RPSA	1	0	1	0
TBC1D1	1	1	0	0
PLP2	1	0	1	0
DTWD1	1	1	0	0
PRPS2	1	0	1	0
MRPL11	1	0	1	0
IMPDH2	1	0	1	0
IARS	1	0	1	0
MCM10	1	1	0	0
TIMM9	1	0	1	0
CA13	1	1	0	0
HSPA8	1	0	1	0
NUDT21	1	0	1	0

LONP1	1	0	1	0
LBR	1	0	0	1
FRAS1	1	1	0	0
PPIH	1	0	1	0
SIGIRR	1	1	0	0
IPO5	1	0	1	0
RPL5	1	0	1	0
LBH	1	1	0	0
LYN	1	0	1	0
SNRPF	1	0	1	0
E2F5	1	1	0	0
SLC25A6	2	0	1	0
RPS23	1	0	1	0
EIF3E	1	0	1	0
SNRPE	1	0	1	0
RPSAP9	1	0	1	0
SOX7	1	0	1	0
ANKRD18B	1	1	0	0
MTHFD1L	1	0	1	0
RTTN	1	0	1	0
CBX1	1	0	1	0
ATAD2	1	0	0	1
HADH	1	0	1	0
MAD2L2	1	1	0	0
LOC440311	1	0	1	0
ORC6	1	1	0	0
FSCN1	1	0	1	0
ATP5A1	1	0	1	0
MARS	1	0	1	0
OPTN	0	1	0	0
FAM214B	0	1	0	0
F5	0	0	1	0
SERPINF1	0	0	1	0
ZBTB43	0	1	0	0
CYB5R1	0	1	0	0
SKIL	0	0	0	1
MXRA8	0	1	0	0
MAOA	0	0	1	0
HIST2H2BE	0	0	1	0
PBXIP1	0	1	0	0
TLDC1	0	1	0	0
ZNF320	0	1	0	0
NACC2	0	1	0	0

SMARCA1	0	0	1	0
NRIP1	0	1	0	0
MAFB	0	1	0	0
TAF13	0	1	0	0
FAM198B	0	1	0	0
TOM1	0	0	1	0
FBXO10	0	0	1	0
E2F8	0	1	0	0
CFLAR	0	1	0	0
ZNF790	0	1	0	0
PRDM1	0	1	0	0
PLXNB3	0	0	1	0
KRT6C	0	0	1	0
STARD13	0	1	0	0
FOXA1	0	1	0	0
KRT6A	0	0	1	0
PIM1	0	1	0	0
GREB1L	0	1	0	0
TNFSF9	0	1	0	0
KRT6B	0	0	1	0
DENND2C	0	1	0	0
HMGN3	0	0	0	1
MC1R	0	0	1	0
HSPB1	0	0	1	0
PAX8	0	0	1	0
KRT81	0	0	1	0
FAM47E	0	0	1	0
GSN	0	0	1	0
HSPA2	0	0	1	0
HIST1H2AC	0	0	1	0
HIST1H1E	0	0	1	0
DHRS3	0	1	0	0
CDK9	0	0	1	0
HS3ST1	0	0	1	0
KRT83	0	0	1	0
GRAMD2	0	1	0	0
TTLL1	0	1	0	0
HIST1H2AD	0	0	1	0
HIST1H3E	0	0	1	0
TXNRD1	0	1	0	0
PLEKHG1	0	0	1	0
VGLL3	0	1	0	0
CDRT1	0	1	0	0

HIST1H2BC	0	0	1	0
NKD1	0	1	0	0
CTSV	0	1	0	0
GPC2	0	1	0	0
FYN	0	0	1	0
RSAD2	0	1	0	0
KRT85	0	0	1	0
COL2A1	0	0	1	0
LURAP1L	0	1	0	0
HIST1H2BG	0	0	1	0
ZEB2	0	1	0	0
TACSTD2	0	1	0	0
CLCF1	0	0	1	0
RGCC	0	1	0	0
EDN1	0	1	0	0
DDX60	0	0	0	1
CHI3L2	0	0	1	0
SP6	0	0	1	0
ZNF583	0	1	0	0
HIST1H4H	0	0	1	0
AKR1C1	0	0	1	0
KRT10	0	0	1	0
CCL28	0	0	1	0
PSCA	0	0	1	0
ALDH3A1	0	0	1	0
PCDH1	0	1	0	0
AKR1C2	0	1	0	0
GPRC5A	0	1	0	0
MAP2	0	1	0	0
COL1A1	0	0	1	0
ISG15	0	0	1	0
KRT16P2	0	0	1	0
S100A7	0	0	1	0
PLD5	0	0	1	0
KRT80	0	0	1	0
KRT16	0	0	1	0
SOWAHB	0	1	0	0
CRABP2	0	0	1	0
NCF2	0	1	0	0
STRA6	0	0	1	0
POSTN	0	0	1	0
FN1	0	0	1	0
MUC16	0	0	1	0

KLK10	0	1	0	0
HIST1H2BF	0	0	1	0
KRT4	0	0	1	0
CBSL	1	1	0	0
HIST1H2BD	0	0	1	0
HIST1H3D	0	0	1	0
PDIA5	1	1	0	0
RPL17	1	0	1	0
STMN1	1	0	1	0
ZNF117	0	1	0	0
RPL17-C18orf32	1	1	0	0
PCDHAC1	0	1	0	0

To determine the mechanism of PAX9's function in SSU biogenesis, I chose 5 candidates from the RNA-seq dataset to follow up on in greater detail. The mRNA levels for the 5 candidates (RPS6/eS6, RPS9/uS4, RPS28/eS28, RPL5/uL18, FBL) were all downregulated after PAX9 depletion and code for nucleolar proteins (Ahmad et al., 2009, Jarboui et al., 2011, Thul et al., 2017) (Fig 4-11A). Four of the candidates (RPS6/eS6, RPS9/uS4, RPS28/eS28, and FBL) were chosen on the basis of literature suggesting a role for these proteins in SSU pre-rRNA processing in HeLa cells (O'Donohue et al., 2010, Tafforeau et al., 2013). Additionally, it was pertinent to analyze RPL5/uL18, as it has a known role in the p53-dependent nucleolar stress response (Sloan et al., 2013a). Interestingly, depletion of both RPS9/uS4 and RPS28/eS28 resulted in a decrease in nucleolar number from 2-3 to only 1 in our original siRNA screen (Farley-Barnes et al., 2018). Therefore, it is possible that the mechanism through which PAX9 depletion results in decreased nucleolar number relies upon reduced expression of RPS9/uS4 and/or RPS28/eS28 (Fig 4-1).

I was able to confirm that depletion of 4 of the 5 tested candidates (RPS6/eS6, RPS9/uS4, RPS28/eS28, and FBL) in MCF10A cells resulted in pre-rRNA processing defects similar to that of PAX9 depletion (Fig 4-11D, E, 4-14). Only RPL5/uL18 did not give the 30S increase characteristic of PAX9 depletion, although this was expected given its known role in LSU pre-rRNA processing (Donati et al., 2013). Notably, knockdown of the candidate mRNAs, including RPL5/uL18, after PAX9 depletion was confirmed using qRT-PCR (Fig 4-11C). Additionally, depletion of several of the candidates individually also resulted in significantly decreased global protein synthesis by puromycin incorporation assay, similar to the effect seen after PAX9 depletion (Fig 4-11F, G). Puromycin incorporation was also reduced after RPS6/eS6 depletion, although it was not statistically significant (Fig 4-11F, G). Therefore, PAX9 may function as a transcription factor to directly or indirectly increase the expression of RPS6/eS6, RPS9/uS4, RPS28/eS28, and/or FBL (Fig 4-2). As the proteins encoded by these mRNAs are required for SSU ribosome biogenesis (Fig 4-11), their depletion after PAX9 knockdown could provide a plausible mechanism through which PAX9 regulates pre-rRNA processing and global protein synthesis.

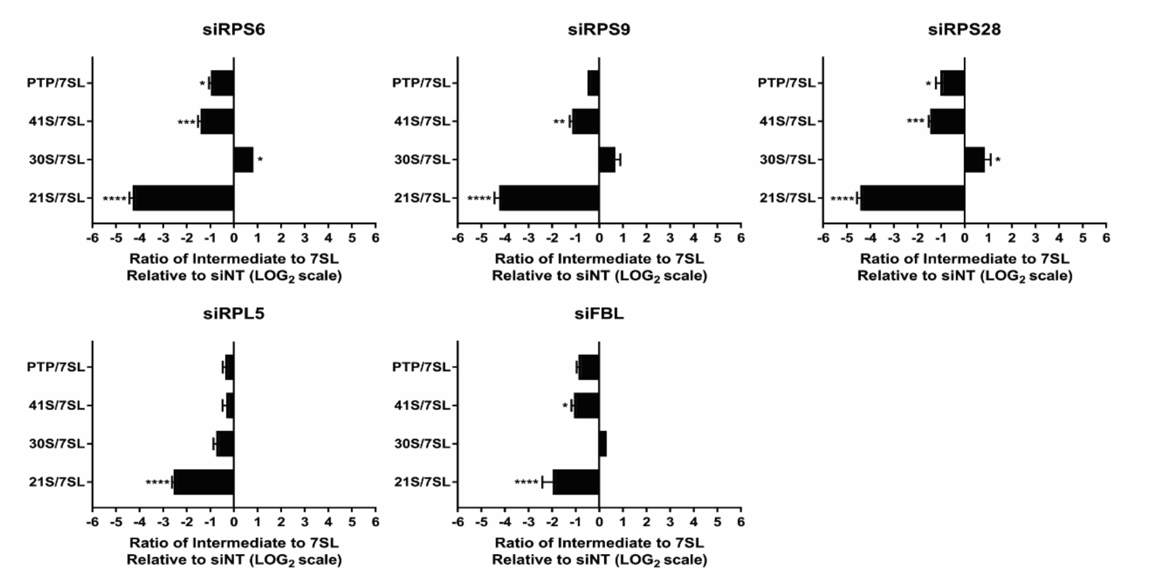

Figure 4-14. Quantitation of the northern blot ratio of each intermediate detected by probe P3 relative to the 7SL loading control for the 5 RNA-seq candidates shown in Figure 4-11D. Data are normalized to the siNT control. N = 3. Data are plotted as mean ± SEM on a LOG$_2$ scale. Statistical analysis was completed by 2-way ANOVA in GraphPad Prism where **** $p \leq 0.0001$, *** $p \leq 0.001$, ** $p \leq 0.01$, and * $p \leq 0.05$.

RNAPII ChIP-seq analysis shows decreased transcription of mRNAs encoding nucleolar proteins after PAX9 depletion

While the RNA-seq analysis confirmed that levels of many nucleolar mRNAs were decreased after PAX9 depletion (Fig 4-2, 4-11), I sought to map how RNAPII distributes on genes using RNAPII ChIP-seq as a readout of transcription. Through this approach, I aimed to untangle the effects of PAX9 depletion on RNAPII transcription from its effects on mRNA stabilization, since this technique was able to detect genome-wide changes in RNAPII density. I identified differential transcriptional regulation of 134 mRNAs after PAX9 knockdown compared to the non-targeting control siRNA (siNT) in MCF10A cells (fold change cutoff ≥ 2 or ≤ -2 and MaxTags ≥ 150) (Fig 4-15). Of these 134, 69 had decreased RNAPII occupancy, consistent with PAX9 acting as a transcriptional driver for these mRNAs.

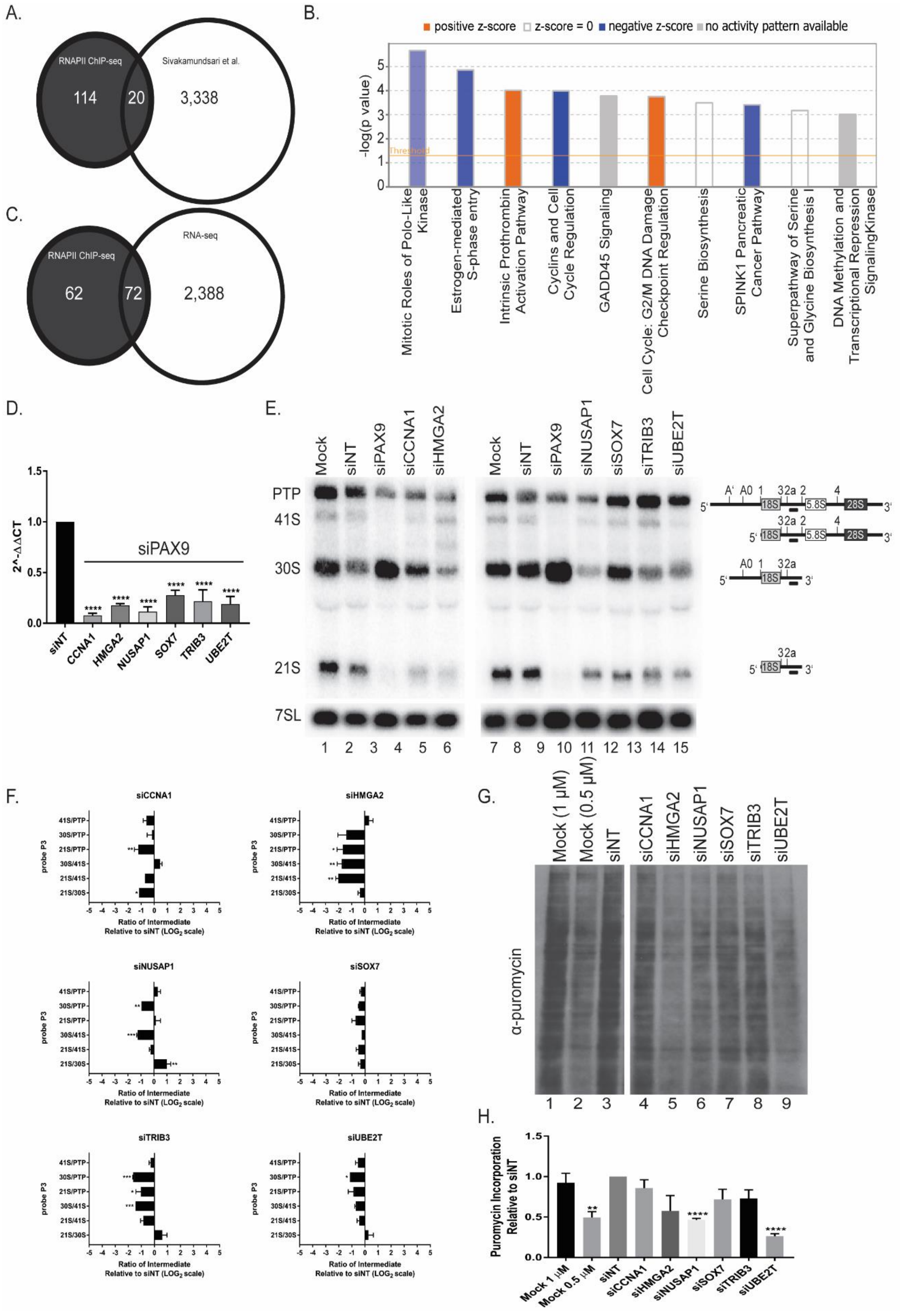
A.
RNAPII ChIP-seq
Sivakamundsari et al.
114
20
3,338
B.
positive z-score
z-score = 0
negative z-score
no activity pattern available
-log(p value)
Threshold
0
1
2
3
4
5
Mitotic Roles of Polo-Like Kinase
Estrogen-mediated S-phase entry
Intrinsic Prothrombin Activation Pathway
Cyclins and Cell Cycle Regulation
GADD45 Signaling
Cell Cycle: G2/M DNA Damage Checkpoint Regulation
Serine Biosynthesis
SPINK1 Pancreatic Cancer Pathway
Superpathway of Serine and Glycine Biosynthesis I
DNA Methylation and Transcriptional Repression SignalingKinase
C.
RNAPII ChIP-seq
RNA-seq
62
72
2,388
D.
2^-ΔΔCT
0.0
0.5
1.0
1.5
siPAX9
siNT
CCNA1
HMGA2
NUSAP1
SOX7
TRIB3
UBE2T
****
E.
Mock
siNT
siPAX9
siCCNA1
siHMGA2
siNUSAP1
siSOX7
siTRIB3
siUBE2T
PTP
41S
30S
21S
7SL
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
A' A0 1 32a 2 4
18S
5.8S
28S
5' 3'
F.
siCCNA1
siHMGA2
siNUSAP1
siSOX7
siTRIB3
siUBE2T
probe P3
41S/PTP
30S/PTP
21S/PTP
30S/41S
21S/41S
21S/30S
-5 -4 -3 -2 -1 0 1 2 3 4 5
Ratio of Intermediate Relative to siNT (LOG2 scale)
G.
Mock (1 μM)
Mock (0.5 μM)
siNT
siCCNA1
siHMGA2
siNUSAP1
siSOX7
siTRIB3
siUBE2T
α-puromycin
1 2 3 4 5 6 7 8 9
H.
Puromycin Incorporation Relative to siNT
0.0
0.5
1.0
1.5
Mock 1 μM
Mock 0.5 μM
siNT
siCCNA1
siHMGA2
siNUSAP1
siSOX7
siTRIB3
siUBE2T
**
****
****

Figure 4-15. RNAPII ChIP-seq analysis shows that PAX9 influences the transcription of several mRNAs that code for nucleolar proteins.

A. Approximately 20 of the differentially occupied genes in the RNAPII ChIP-seq analysis are bound by PAX9 in (Sivakamasundari et al., 2017).

B. RNAPII ChIP-seq differentially occupied mRNAs are enriched for cell cycle regulators. The comparative pathways analysis using Ingenuity Pathways Analysis (IPA) software (QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuitypathway-analysis) reveals both upregulated (orange) and downregulated (blue) pathways enriched in the list of 134 RNAPII ChIP-seq differentially occupied mRNAs. The comparative pathways analysis z score reflects the correlation between the observed expression of the mRNAs in each pathway and the predicted expression change based on existing literature. Pathways with a z score of 0 are shown as white and pathways with no activity pattern available are shown in gray. Only pathways enriched with a -log(p-value), which measures the enrichment of that pathway in the RNAPII ChIP-seq dataset, of ≥ 3 are shown.

C. Venn diagram showing that 72 genes overlap between the RNAPII ChIP-seq and RNA-seq datasets.

D. qRT-PCR confirms depletion of the 6 RNAPII ChIP-seq/RNA-seq candidates after PAX9 knockdown. After depletion using siRNAs targeting either PAX9 or a non-targeting control siRNA (siNT), the levels of 6 mRNAs were quantified using qRT-PCR. Data are shown as mean ± SEM. Three biological replicates, each with 3 technical replicates were performed. Significance was calculated by One-way ANOVA using GraphPad Prism where **** $p \leq 0.0001$, *** $p \leq 0.001$, and ** $p \leq 0.01$.

E. Northern blot analysis shows small subunit (SSU) pre-rRNA processing defects after depletion of the 6 RNAPII ChIP-seq/RNA-seq candidate mRNAs. Representative northern blot after knockdown of the indicated siRNAs using probe P3. A probe for the 7SL RNA was used as a loading control. Pre-rRNA processing intermediates detected by probe P3 are shown to the right of the northern blot. PTP indicates the 47S, 45S, and 43S processing intermediates.

F. Knockdown of 5 of the 6 tested RNAPII ChIP-seq/RNA-seq candidate mRNAs resulted in SSU pre-rRNA processing defects in MCF10A cells. Ratio analysis of multiple precursors [RAMP, (Wang et al., 2014)] quantitation of probe P3 northern blots in (E). Graph is mean ± SEM. N = 3. 2-way ANOVA. *** $p \leq 0.001$, ** $p \leq 0.01$, and * $p \leq 0.05$.

G. Depletion of RNAPII ChIP-seq/RNA-seq candidate mRNAs results in reduced global protein synthesis. After 72 hours knockdown with the indicated siRNAs, MCF10A cells were pulsed with puromycin for 1 hour and protein was harvested. Western blotting with an anti-puromycin antibody was completed. Mock at half the dose (0.5 μM) is shown as a control.

H. Quantitation of 3 biological replicates of the puromycin incorporation assay shown in (G). Data are shown as mean ± SEM. N = 3. Significance was calculated by Student's t-test using GraphPad Prism where **** $p \leq 0.0001$ and ** $p \leq 0.01$.

To assess the validity of the RNAPII ChIP-seq dataset, I used CentriMo to identify potential PAX9 DNA-binding sites in the 1 kb upstream of the 134 genes with differential RNAPII occupancy (Bailey and Machanick, 2012). Searching for the 5’ CGCGTGACCG 3’ PAX9 binding motif discovered in (Sivakamasundari et al., 2017) revealed 186 possible sites in the 1000 bp upstream of the 134 differentially occupied genes. Also, the known PAX9 DNA binding motifs CD19-2(A-ins) (5’-GCGTGACCA-3’) and e5 (5’-GCGGAACGG-3’) had 131 and 2 binding sites in these sequences, respectively. Additionally, Analysis of Motif Enrichment (AME) analysis (McLeay and Bailey, 2010) identified the PAX5 and PAX6 DNA binding motifs as being significantly enriched in the 1 kb of sequence upstream of the 134 genes ($p \leq 0.05$). Since multiple PAX proteins can bind the same motif (Epstein et al., 1994), this suggests that this dataset does contain mRNAs that are regulated by PAX9. Of the 134 differentially occupied genes, 20 have also been shown to be differentially regulated by PAX9 directly in PAX9 ChIP-seq experiments on E12.5 WT vertebral column murine tissue (Fig 4-15A, Table 4-2) (Sivakamasundari et al., 2017). These analyses confirm the ability of RNAPII ChIP-seq to detect changes in RNAPII-mediated transcription after PAX9 knockdown.

Table 4-2. Comparison of the RNAPII ChIP-seq differentially regulated mRNAs to genome-wide siRNA screens for ribosome biogenesis factors [Farley-Barnes et al (Farley-Barnes et al., 2018), Badertscher et al (Badertscher et al., 2015), and Neumüller Yeast and Neumüller *Drosophila* referring to the siRNA screens in yeast and Drosophila, respectively (Neumuller et al., 2013)] and to genes differentially occupied by PAX9 in the vertebral columns of E12.5 mice by PAX9 ChIP-seq [Sivakamasundari et al. (Sivakamasundari et al., 2017)]. A 1 denotes a mRNA that was included as a hit in that siRNA screen, while a 0 indicates that the mRNA was not considered a hit in that screen. For the second column, a 1 denotes that the mRNA had decreased expression (fold change ≤ - 2) in the RNAPII ChIP-seq dataset, while a 0 indicates denotes that the mRNA had increased expression (fold change ≤ - 2) in the RNAPII ChIP-seq dataset.

PAX9 RNAPII differentially occupied	Decreased occupation (1) or increased occupation (0)?	Farley-Barnes et al. Overlap	Badertscher et al. Overlap	Neumuller Drosophila Overlap	Neumuller Yeast Overlap	Sivakam-asundari et al Overlap
ADAMTS14	0	0	0	0	0	1
ALDH1A3	0	0	0	0	0	1
ANLN	1	1	0	0	0	1
AURKB	1	0	0	0	1	0
CCNA1	1	0	0	1	0	1
CDCA7	1	0	0	0	0	1
CDK1	1	0	0	0	1	0
CEP55	1	0	0	0	0	1
CLDN4	0	0	0	0	0	1
CPA4	0	0	0	1	0	0
DUSP16	0	0	0	0	0	1
FAT2	1	0	0	0	0	1
FBXO32	0	0	0	1	0	0
GPRC5D	0	0	0	0	0	1
HIST1H3B	1	0	0	0	1	0
HIST1H3F	1	0	0	0	1	0

HIST1H3G	1	0	0	0	1	0
HIST1H3I	1	0	0	0	1	0
HIST1H4D	1	0	0	0	1	0
HIST1H4K	1	0	0	0	1	0
HIST1H4L	1	0	0	0	1	0
HMGA2	1	0	0	0	0	1
HS3ST1	0	0	0	0	0	1
IL4	0	0	0	0	0	1
KIF18B	1	0	0	1	0	0
MACC1	0	0	0	0	0	1
MFAP5	0	0	0	0	0	1
MUC16	0	0	0	0	0	1
MYO5B	0	0	0	0	1	1
PCK2	1	0	0	1	0	0
PLK4	1	0	0	0	1	0
RAPSN	0	0	0	0	0	1
RRM2	1	0	0	0	0	1
SCD5	0	0	0	0	1	1
SOX7	1	0	0	1	0	0

Based on the hypothesis that PAX9 acts as an RNAPII transcription factor for regulators of nucleolar function (Fig 4-2), I expected to detect changes in the RNAPII-mediated transcription of a number of mRNAs encoding nucleolar proteins after PAX9 depletion. Indeed, mRNAs coding for nucleolar proteins were enriched, with 46 of the 134 genes with differential RNAPII occupancy (34.3%) coding for nucleolar proteins in at least one of three databases (Table 4-3) (Ahmad et al., 2009, Thul et al., 2017, Jarboui et al., 2011). This is, again, higher than would be expected, assuming that nucleolar proteins account for approximately 5% of all cellular proteins (Ahmad et al., 2009). Additionally, depletion of one gene with differential RNAPII occupancy (ANLN) resulted in decreased nucleolar number our siRNA screen, similar to the phenotype seen after PAX9 depletion (Fig 4-1B) (Farley-Barnes et al., 2018). Notably, 18 of the 134 genes with differential RNAPII occupancy (13.4%) appeared in other genome-wide screens for ribosome biogenesis factors [Table 4-2, (Badertscher et al., 2015, Neumuller et al., 2013)]. Another targeted screen investigated the effects of 2 of the 134 genes (TOP2A and CDCA8) on pre-rRNA processing when depleted by siRNA in HeLa cells (Tafforeau et al., 2013). However, depletion of neither TOP2A or CDCA8 in that screen gave the 30S pre-rRNA increase characteristic of PAX9 depletion (Tafforeau et al., 2013). In all, the RNAPII ChIP-seq analysis confirmed that PAX9 drives the transcription of mRNAs encoding proteins critical for ribosome biogenesis.

Table 4-3. Comparison of all RNAPII ChIP-seq differentially occupied genes to 3 databases of nucleolar proteins: the Human Protein Atlas (Thul et al., 2017), the Nucleolar protein database [NOPdb, (Ahmad et al., 2009)], and nucleolar proteins in human T-cells from the Gautier laboratory [Gautier, (Jarboui et al., 2011)]. For the second column, a 1 denotes that the mRNA had decreased expression (fold change ≤ -2) in the RNA-seq dataset, while a 0 indicates denotes that the mRNA had increased expression (fold change ≤ - 2) in the RNA-seq dataset.

PAX9 RNAPII differentially occupied	Decreased occupation (1) or increased occupation (0)?	Human Protein Atlas	NOPdb	Gautier
ALDH1L2	1	0	1	0
ANLN	1	0	1	0
ARHGAP11A	1	1	0	0
AURKB	1	0	1	1
CCNA1	1	0	1	0
CCNB1	1	0	1	0
CDCA8	1	1	0	0
CDK1	1	0	1	0
COL1A1	0	0	1	0
DEPDC1	1	1	1	0
FAT2	1	0	0	1
GPRC5A	0	1	0	0
HIST1H1B	1	0	1	0
HIST1H1D	1	0	1	0
HIST1H2AB	1	0	1	0
HIST1H2AI	1	0	1	0
HIST1H3B	1	0	1	0
HIST1H3F	1	0	1	0
HIST1H3G	1	0	1	0
HIST1H3I	1	0	1	0
HIST1H4D	1	0	1	0
HIST1H4K	1	0	1	0
HIST1H4L	1	0	1	0
HJURP	1	1	0	0
HMGA2	1	1	1	0
HS3ST1	0	0	1	0
KIF18B	1	0	1	0
KIF23	1	0	1	0

KLK10	0	1	0	0
KRT16	0	0	1	0
KRT17	0	0	1	0
KRT80	0	0	1	0
MAP2	0	1	0	0
MKI67	1	1	1	0
MUC16	0	0	1	0
NUSAP1	1	1	1	0
PCK2	1	0	1	0
PHGDH	1	0	1	0
PTMS	1	1	0	0
S100A2	1	1	1	0
S100A3	1	1	0	0
SOX7	1	0	1	0
SPDYA	0	0	1	0
TOP2A	1	1	1	1
TRIB3	1	1	0	0
UBE2T	1	1	0	0

In addition, both Gene Ontology (GO) enrichment (2017, Ashburner et al., 2000, Mi et al., 2017) and Ingenuity Pathway Analysis (Kramer et al., 2014) of the RNAPII ChIP-seq list show an enrichment of genes associated with cell cycle regulation (Fig 4-15B). This is consistent with PAX9 depletion resulting in a G1 accumulation as shown above (Fig 4-10) and with alterations in ribosome biogenesis resulting in cell cycle inhibition (Fumagalli et al., 2012, Bernstein et al., 2007). Therefore, the RNAPII ChIP-seq analysis confirms that PAX9 levels influence the distribution of cells within the cell cycle as would be expected from a regulator of ribosome biogenesis.

Based on our RNA-seq results above (Fig 4-11), I had hypothesized that PAX9 acts as a transcription factor for RPS6/eS6, RPS9/uS4, RPS28/eS28, and/or FBL. Interestingly, while the RNA-seq candidate mRNAs for RPS6/eS6, RPS9/uS4, RPS28/eS28, and/or FBL were decreased to a slight extent in RNAPII ChIP-seq, this depletion was not statistically significant. As only one replicate of the RNAPII ChIP-seq analysis was performed after 3 days of PAX9 depletion, it is possible that this experimental structure does not adequately capture differences in the transcription of these specific mRNAs. It is also possible that even such subtle changes in transcription levels for these essential proteins are sufficient in combination to cause the pre-rRNA processing inhibition seen upon PAX9 depletion. Additionally, the lack of significant changes in RNAPII occupancy for these mRNAs may support an indirect mechanism through which PAX9 regulates their levels, possibly by acting as an RNAPII transcription factor for proteins that affect the mRNAs' stability (Fig 4-2).

Overlap of RNAPII ChIP-seq and RNA-seq differentially expressed genes

There was significant overlap between the RNA-seq and RNAPII ChIP-seq datasets. Over half (72/134) of the RNAPII differentially occupied mRNAs were also differentially expressed in the RNA-seq analysis (Fig 4-15C and Table 4-4). Of these 72 overlapping genes, 39 genes had decreased mRNA expression in both analyses, consistent with PAX9 acting as an enhancing transcription factor for these mRNAs (Table 4-4). Of the 39 decreased and overlapping genes, 19 code for nucleolar proteins [Table 4-4, (Ahmad et al., 2009, Jarboui et al., 2011, Thul et al., 2017)]. I have designated these 19 genes as the high-confidence list of potential PAX9 targets that affect ribosome biogenesis.

Table 4-4. List of mRNAs that are differentially expressed in both the RNAPII ChIP-seq and RNA-seq datasets, mRNAs with decreased expression in both lists, and mRNAs that have decreased expression in both lists and are also nucleolar in at least 1 database (Jarboui et al., 2011, Ahmad et al., 2009, Thul et al., 2017).

Overlap between RNA-seq and RNAPII ChIP-seq lists	Decreased expression in both lists	Decreased expression in both lists and Nucleolar
ALDH1A3	CCNA1	ALDH1L2
ALDH1L2	BIRC5	ARHGAP11A
AP5B1	MKI67	AURKB
ARHGAP11A	AURKB	CCNA1
ARHGAP30	CCNB1	CCNB1
ASPM	TOP2A	CDCA8
AURKB	MT2A	CDK1
BIRC5	STC2	DEPDC1
C15orf48	LOC100288637	HJURP
CCNA1	PIF1	HMGA2
CCNB1	HJURP	KIF18B
CD24	CEP55	MKI67
CDCA7	PHGDH	NUSAP1
CDCA8	ARHGAP11A	PCK2
CDH1	PECAM1	PHGDH
CDK1	NUSAP1	PTMS
CDKN1A	TRIB3	SOX7
CDKN3	NEIL3	TOP2A
CDSN	KIF18B	TRIB3
CENPM	CHAC1	
CEP55	PRC1	
CHAC1	CENPM	
CLDN4	PCK2	
COL1A1	CDKN3	
CPA4	PSAT1	
DEPDC1	NEK2	
DSC2	CDCA7	
EIF4EBP1	PLK4	
ELF3	SOX7	
ESPL1	ESPL1	
FBXO32	ASPM	
GPRC5A	CDK1	
HJURP	PTMS	
HMGA2	KCTD15	

HS3ST1	ALDH1L2	
KCNK15	EIF4EBP1	
KCTD15	CDCA8	
KIF18B	DEPDC1	
KLK10	HMGA2	
KLK7		
KRT16		
KRT80		
LMO7		
MACC1		
MAP2		
MKI67		
MT2A		
MUC16		
NEIL3		
NEK2		
NUSAP1		
OVOL1		
PCK2		
PECAM1		
PHGDH		
PIF1		
PLAUR		
PLK4		
PRC1		
PSAT1		
PTMS		
RAB11FIP1		
SCD5		
SERPINB2		
SLC2A12		
SOX7		
SPRR2D		
STC2		
SYTL2		
TOP2A		
TRIB3		
VSTM2L		

Further analysis of the mRNAs differentially regulated in both the RNA-seq and RNAPII ChIP-seq datasets validates PAX9 as a driver of these mRNA levels. GO analysis of the 72 genes present in both datasets again revealed enrichment of genes associated with cell cycle regulation as well as cell division (2017, Mi et al., 2017, Ashburner et al., 2000). As expected based on analysis of the whole RNAPII dataset above, CentriMo analysis of the 1 kb upstream of the 72 genes that appeared in both the RNAPII ChIP-seq and RNA-seq hit lists showed enriched DNA binding motifs for both PAX5 and PAX6 (Bailey and Machanick, 2012). Therefore, these two unbiased approaches to detect changes in mRNA levels have revealed a high-confidence list of 19 nucleolar proteins for which PAX9 acts as a transcriptional enhancer.

From the high confidence list of 19 downregulated mRNAs that encode nucleolar proteins, I chose 5 (CCNA1, HMGA2, NUSAP1, SOX7, and TRIB3) to analyze in greater detail (Fig 4-15, 4-16). Both CCNA1 and HMGA2 have been shown to be directly regulated by PAX9 in a ChIP-seq study of E12.5 murine intervertebral discs (Sivakamasundari et al., 2017). Interestingly, HMGA2 deletion in rabbits resulted in dwarfism (small stature being a common phenotype in diseases of impaired ribosome biogenesis) and altered craniofacial development (Carneiro et al., 2017). Additionally, I chose to study UBE2T, as it is nucleolar and was shown to be downregulated in both datasets, although not to the same extent as the other chosen hits (Fig 4-16). UBE2T is essential for the Fanconi anemia DNA repair pathway which has been newly connected to ribosome biogenesis (Sondalle et al., 2019). We therefore hypothesized that depletion of one or multiple of these 6 mRNAs could provide the mechanism through which PAX9 regulates SSU biogenesis.

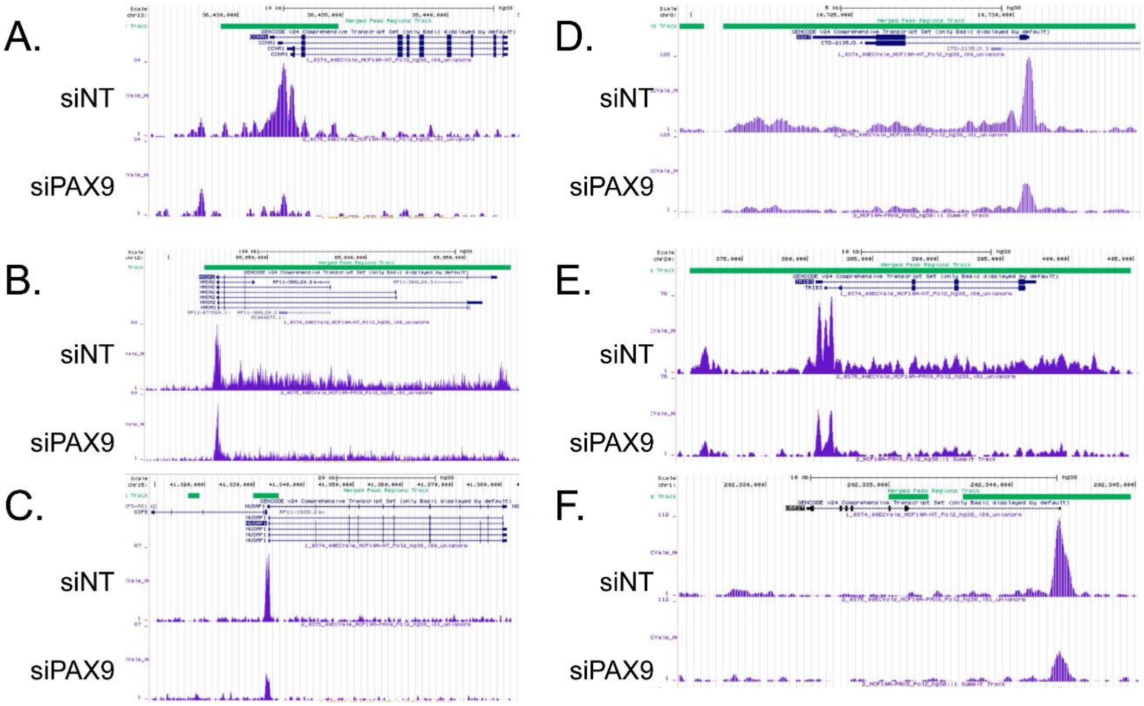

Figure 4-16. UCSC genome browser images of differentially expressed RNAPII ChIP-seq hits. Shown in green is the merged peak region used to define the indicated gene. Top is the siNT control, bottom is siPAX9. A-F are the indicated hits (CCNA1, HMGA2, NUSAP1, SOX7, TRIB3, and UBE2T, respectively).

Interestingly, while individual depletion of none of the 6 tested candidates resulted in the same pre-rRNA processing defect as PAX9 depletion, most of the candidates did affect SSU pre-rRNA processing to some extent (Fig 4-15E-F, 4-17). Depletion of CCNA1 caused a similar processing defect as PAX9 depletion (21S was significantly decreased relative to its 30S precursor and to the 7SL loading control), but the accumulation of the 30S pre-rRNA processing intermediate was not significant (Fig 4-15E-F, 4-17). Additionally, depletion of NUSAP1 and UBE2T significantly decreased global protein synthesis, with the other candidates altered protein synthesis to a lesser, non-significant extent (Fig 4-15G, H). qRT-PCR confirmed depletion of all of these candidates both when PAX9 is depleted in MCF10A cells and when each candidate is individually depleted for the northern blotting experiments (Fig 4-15D, 4-18). It is therefore unlikely that depletion of any of these 6 proteins alone results in the pre-rRNA processing defects observed after PAX9 depletion. However, reduced levels of these mRNAs in combination may contribute to the overall impairment of protein synthesis that occurs when PAX9 is depleted in human cells.

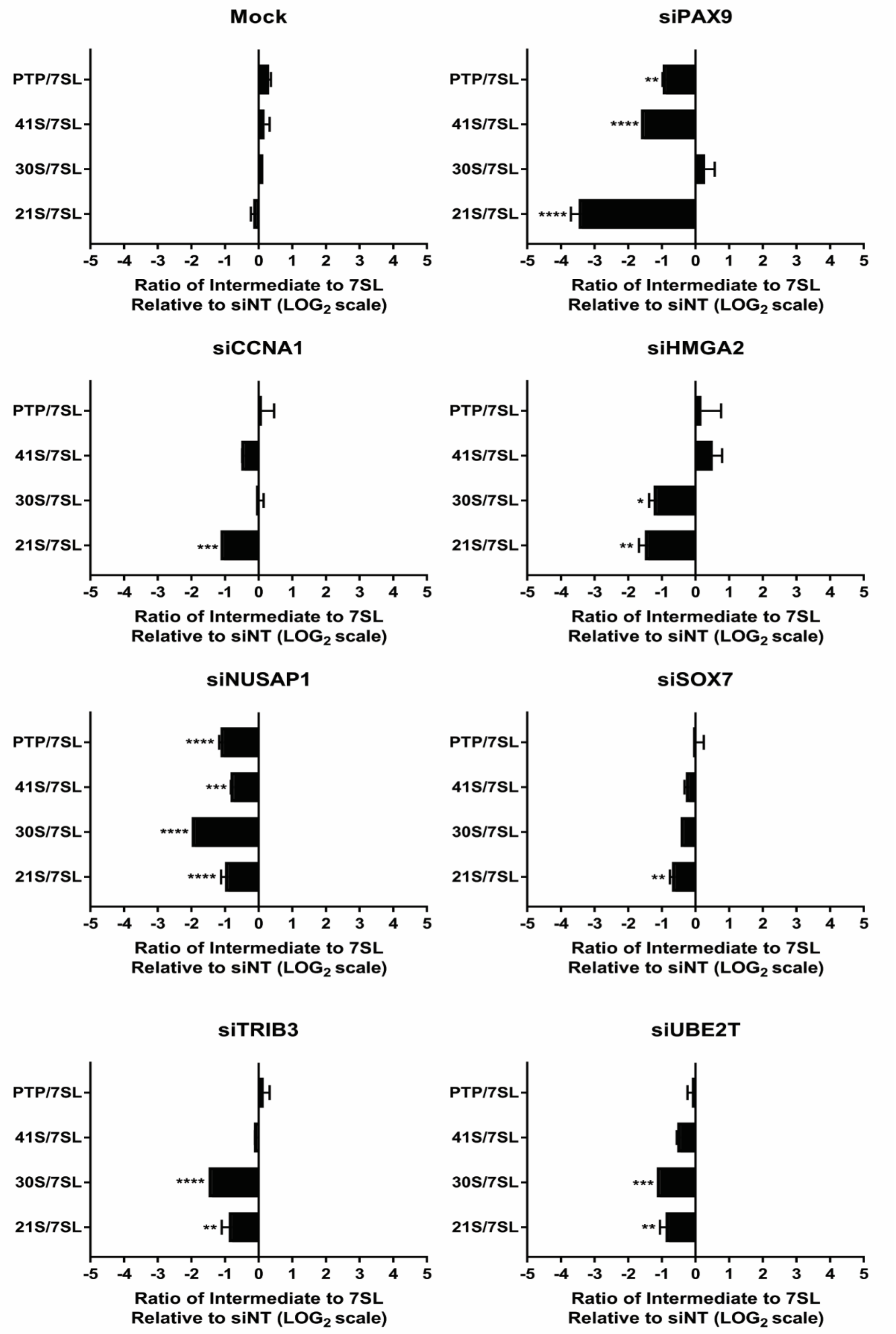

Figure 4-17. Quantitation of the northern blot ratio of each intermediate detected by probe P3 to the 7SL loading control, relative to siNT, for the 6 RNAPII ChIP-seq

candidates as well as for the Mock and siPAX9 positive and negative controls. N = 3. Data are plotted as mean ± SEM on a LOG_2 scale. Statistical analysis was completed by 2-way ANOVA in GraphPad Prism where **** $p \leq 0.0001$, *** $p \leq 0.001$, ** $p \leq 0.01$, and * $p \leq 0.05$.

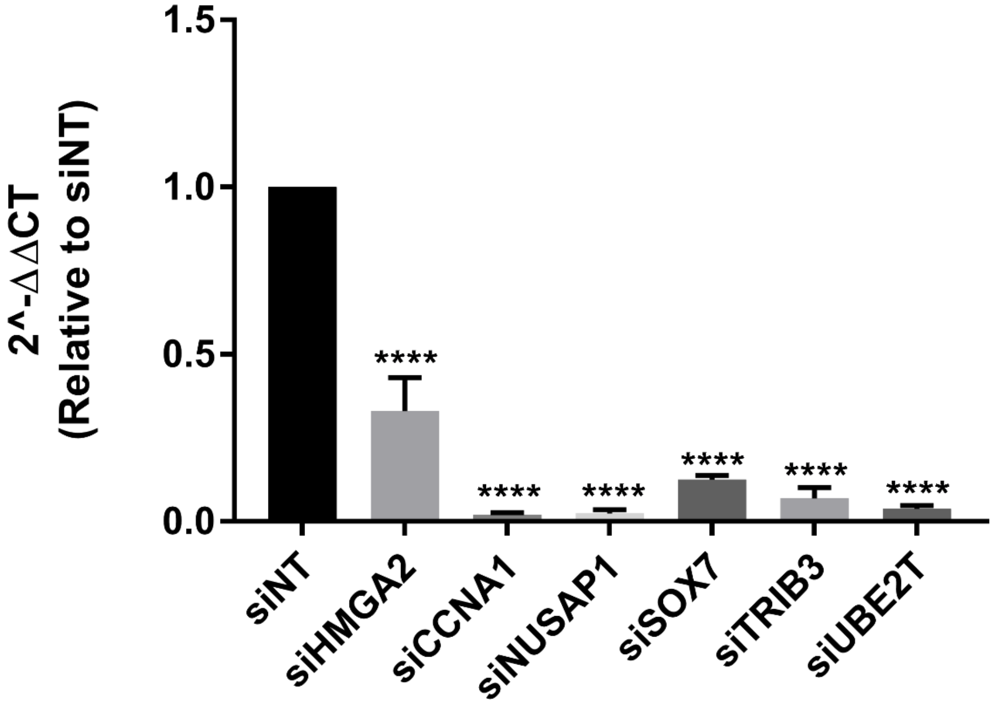

Figure 4-18. qRT-PCR after depletion of the indicated RNA-seq/RNAPII ChIP-seq hits relative to the siNT control. After depletion of each hit using siRNAs for 72 hours, qRT-PCR was performed using primers targeting that gene of interest. Data are shown as $2^{-\Delta\Delta Ct}$, relative to the siNT control. Three biological replicates, each with three technical replicates, were performed for each qRT-PCR experiment. Data are shown as mean ± SEM. Analysis was completed using One-way ANOVA using GraphPad Prism where **** $p \leq 0.0001$.

PAX9's role in ribosome biogenesis is conserved in the model organism *X. tropicalis*

Having identified a novel role for PAX9 in upregulating the transcription of mRNAs encoding nucleolar proteins in human cells, I next sought to determine whether Pax9 depletion in pipid frogs (*X. tropicalis*) leads to craniofacial defects. Given my observations on a role for PAX9 in pre-rRNA processing and global protein synthesis in human tissue culture cells, and the occurrence of craniofacial defects in many ribosomopathies (see Chapter 1), I predicted that Pax9 may play a role in the normal development of the face.

To accomplish this goal, I utilized MO knockdown in the model organism, *X. tropicalis* with the help of Dr. Engin Deniz and Dr. John Griffin in the laboratory of Dr. Mustafa Khokha. Embryos were injected with a MO targeting the translational start site of *pax9* at the two-cell stage so that only half of the organism is depleted of Pax9. Pax9 depletion resulted in microcephaly and decreased size of multiple cranial neural crest cartilages, relative to the uninjected control (UC) side of the embryo (Fig 4-19A, B). Optical coherence tomography (OCT) imaging of the cranial neural crest cartilages revealed reduced size and dysmorphology of the branchial cartilages, Meckel's, and ceratohyal cartilages (Fig 4-19B). Indeed, the Meckel's cartilage was unidentifiable in most of the embryos examined. This demonstrates a role for Pax9 in the development of the Meckel's cartilage and the cranial neural crest as a whole.

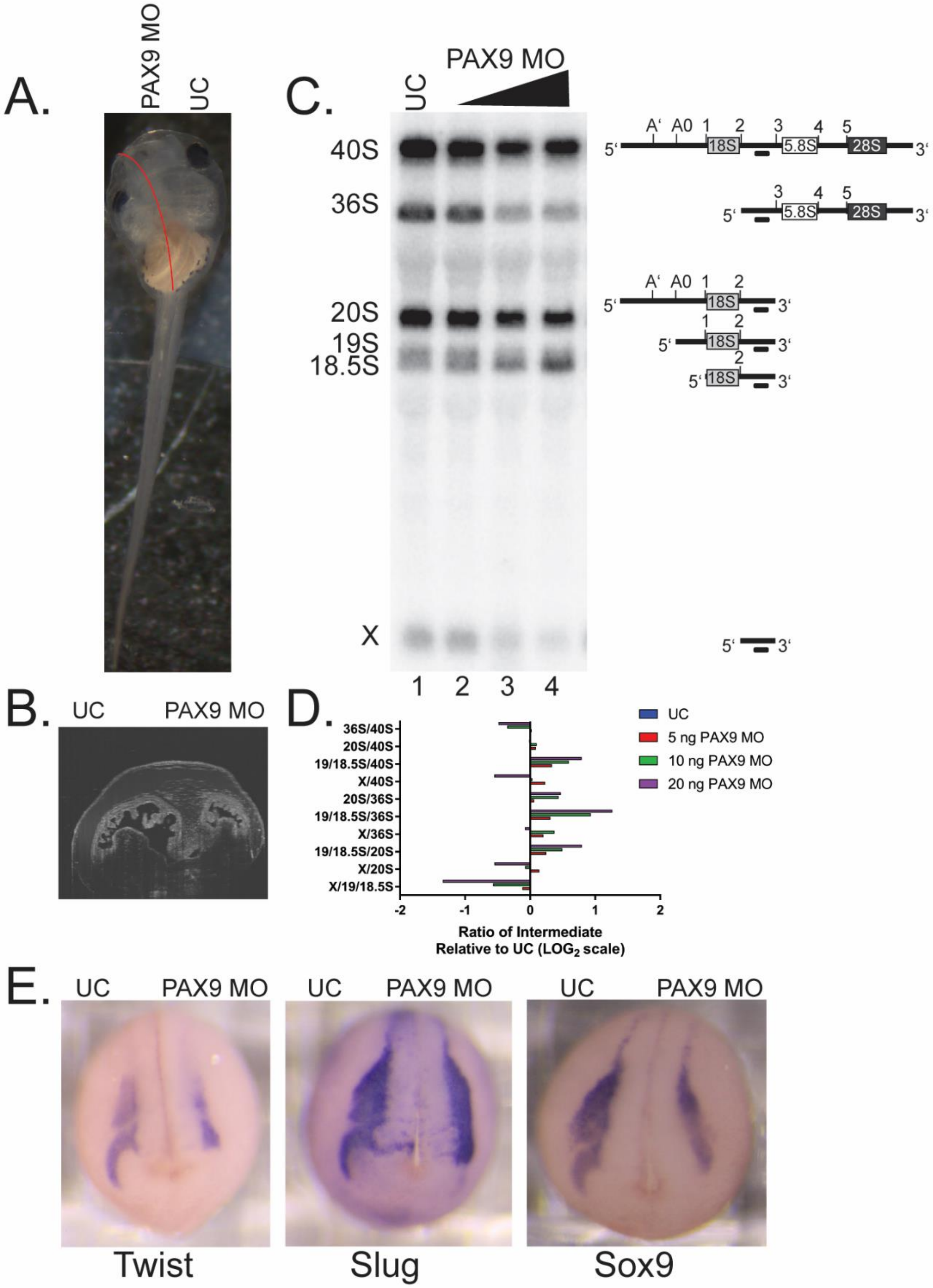

Figure 4-19. Pax9 depletion results in craniofacial defects and impaired ribosome biogenesis in *X. tropicalis*.

A. At the two-cell stage, *X. tropicalis* embryos were injected with a morpholino targeting *pax9* (left) or an uninjected control (UC, right). Image shows gross morphology of Pax9-depleted embryos at stage 45.

B. Optical coherence tomography (OCT) image of stage 45 embryos. At the two-cell stage, embryos were injected with a morpholino targeting *pax9* (right) or an uninjected control (left). OCT imaging shows decreased size of branchial cartilages upon Pax9 depletion. The Meckel's and ceratohyal also had significantly reduced size (not shown).

C. Morpholino depletion of Pax9 results in defects in small subunit (SSU) pre-rRNA processing compared to uninjected control embryos. Northern blot hybridized with probe c from (Griffin et al., 2015). Processing intermediates visualized with this probe are shown to the right. Lane 1 - uninjected control (UC); lane 2 - 5 ng *pax9* morpholino; lane 3 – 10 ng *pax9* morpholino; lane 4 – 20 ng *pax9* morpholino. RNA was harvested at stage 28.

D. Ratio Analysis of Multiple Precursors (RAMP) (Wang et al., 2014) for the northern blot shown in C. N = 1. Data are shown on a Log_2 scale, relative to the UC embryos.

E. Reduced levels of neural crest cell markers upon Pax9 depletion. At the two-cell stage, embryos were injected with a morpholino targeting *pax9* (right) or an uninjected control (left). *In situ* hybridizations were performed at stage 24-25 with the indicated cranial neural crest markers. Dorsal views are shown with the anterior at the bottom.

To test the extent to which ribosome biogenesis is affected by PAX9 depletion in *X. tropicalis*, northern blots were performed on embryos injected with varying doses of a *pax9* MO at the one-cell stage. At stage 28, RNA was harvested from those embryos as well as UC embryos. Pre-rRNA processing was analyzed using a northern blot probe targeting ITS1 (Griffin et al., 2015). Increasing doses of *pax9* MO resulted in increased levels of the 19S/18.5S processing intermediates as well as decreased levels of the 36S intermediate (Fig 4-19C, D). These SSU pre-rRNA processing defects are similar to those seen upon PAX9 depletion in humans (Fig 4-19C, 4-4). Therefore, PAX9's role in human ribosome biogenesis is conserved to the model organism *X. tropicalis*.

Given the connection between cranial neural crest development and ribosome biogenesis, we performed *in situ* hybridizations to further probe the ontogeny of the cranial neural crest defects observed upon Pax9 knockdown. Expression of the cranial neural crest master genes *twist*, *slug*, and *sox9* was decreased early in development (Fig 4-19E). These reductions were limited to the anterior portion of the embryo and coincide with known regions of high Pax9 expression [Fig 4-19E, (Sanchez and Sanchez, 2013)]. Taken together, these data suggest that Pax9's role in ribosome biogenesis is linked to its function in cranial neural crest development.

DISCUSSION

PAX9 regulates human ribosome biogenesis by acting as a transcription factor to influence the expression of multiple mRNAs required for pre-rRNA processing and global protein synthesis (Fig 4-2). Depletion of PAX9 disrupts nucleolar structure so that the number of nucleoli is reduced from 2-3 to only 1 (Fig 4-1). Additionally, PAX9 depletion results in SSU pre-rRNA processing defects and reduced global protein synthesis (Fig 4-4, 4-5). As measured by RNA-seq and RNAPII ChIP-seq, PAX9 depletion results in expression changes for a number of mRNAs (Fig 4-11, 4-15). Several of these differentially expressed mRNAs play roles in pre-rRNA processing and global protein synthesis (Fig 4-11, 4-15). Additionally, this role in ribosome biogenesis is conserved to *X. tropicalis*. This analysis therefore illuminates a critical role for the RNAPII transcription factor PAX9 in making ribosomes.

Working with the Khokha laboratory, we have made substantial progress towards defining the direct and indirect global effects of PAX9 depletion on mRNA abundance. Alignment of the RNAPII ChIP-seq dataset with existing PAX9 ChIP-seq data (Sivakamasundari et al., 2017) gives a high-confidence list of mRNAs whose transcription is directly affected by PAX9 binding (Fig 4-15A). Additionally, the alignment of the RNAPII ChIP-seq and RNA-seq datasets provides a list of mRNAs whose transcription is changed upon PAX9 depletion (Fig 4-15C, Table 4-4). These changes may be an indirect effect of PAX9 depletion through PAX9 acting as a transcription factor for mRNAs that affect the levels of nucleolar proteins. We have analyzed several proteins from both of these direct and indirect categories (Fig 4-15) to pinpoint how their individual influences on human ribosome biogenesis may contribute to the global effects of PAX9 knockdown.

Additionally, several mRNAs, including the 5 candidates further examined in Figure 4-11, appeared in only the RNA-seq list. The transcription of these mRNAs may still be regulated by one of the two mechanisms above, as neither genome-wide approach is saturating. Additionally, their stability may be affected by PAX9 depletion. Crossing the RNAPII ChIP-seq dataset with 4 RNA binding protein databases (Gerstberger et al., 2014, Brannan et al., 2016, Cook et al., 2011, Ray et al., 2013) revealed 4 potential RNA binding proteins (NEIL3, MKI67, HIST1H1B, and NUSAP1). NUSAP1 was further examined here (Fig 4-15). However, depletion of NUSAP1 did not result in the same pre-rRNA processing defect as PAX9 depletion, suggesting that this is not the singular factor controlling mRNA stability. Additionally, depletion of MKI67 has previously been reported to have no effect on pre-rRNA processing (Sobecki eLife 2016). Therefore, the mechanism(s) regulating the potential instability of these mRNAs remain undefined.

This analysis highlights the role of Wnt signaling in PAX9 function (Fig 4-12, 4-13). Dickkopf WNT Signaling Pathway Inhibitor 1 (DKK1), a negative regulator of Wnt signaling, shows increased expression after PAX9 depletion in the RNA-seq dataset presented here. Interestingly, increased expression of DKK1 was also found in a recent RNA-seq analysis of palatal shelves of E13.5 *Pax9*$^{-/-}$ mice (Jia et al., 2017a). The authors showed that inhibition of Dkk1 rescued the cleft palates of *Pax9*$^{-/-}$ mice (Jia et al., 2017a). Additionally, research has also suggested a role for the related EDAR/NF-κB signaling pathway in cleft palate formation in *Pax9*$^{-/-}$ mice. However, this pathway was not enriched in any of our datasets by Ingenuity Pathways Analysis [IPA, (Jia et al., 2017b)]. Interestingly, treatment of *Pax9*$^{-/-}$ mice with Wnt inhibitors or a monoclonal antibody against EDAR did not rescue a number of phenotypes, including arrested tooth development, limb defects, and lethality just after birth (Jia et al., 2017a, Li et al., 2017, Jia et al., 2017b). This suggests that pathways other than Wnt signaling are contributing to the defects seen in *Pax9*$^{-/-}$ mice, leaving open the possibility that ribosome biogenesis defects are responsible for these other developmental anomalies.

Notably, attempts to rescue the biochemical and morphological defects of PAX9 depletion have thus far been unsuccessful (data not shown). Using both lentiviral and HEK293FT Flp-In^{TM} T-rex^{TM} (Invitrogen) expression systems in humans, pre-rRNA processing defects remained after PAX9 depletion despite overexpression of an siRNA-resistant form of PAX9. However, the functionality of the resulting siRNA-resistant PAX9 was not assessed, and it is possible that the N-terminal HA tag used interfered with PAX9's ability to bind DNA. Additionally, we are currently attempting to rescue both the gross morphological defects and pre-rRNA processing defects in the *X. tropicalis* model system. Thus far, the human mRNA used for rescue has been too toxic to the developing embryos to complete these analyses, although we are hopeful that with further optimization this will be a powerful tool to evaluate the influence of the PAX9 human mutations on PAX9 function in both development and ribosome biogenesis.

This work suggests an unexplored mechanism through which mutation of PAX9 causes defects in craniofacial development. Future studies are needed to further elucidate PAX9's role in ribosome biogenesis in the context of craniofacial development in *X. tropicalis*. One important aspect of these studies will be to discover whether the known PAX9 mutations also result in ribosome biogenesis defects. Additionally, it will be crucial to determine the extent to which the nucleolar stress response is activated in the developing craniofacial tissues after PAX9 depletion. In the future, investigations into this role for PAX9 in regulating ribosome biogenesis will illuminate new understanding of the pathophysiology of PAX9 mutation and its contributions to craniofacial development.

MATERIALS AND METHODS

Cell Lines and Media

MCF10A cells (ATCC CRL-10317) were grown in Dulbecco's modified Eagle's medium/Nutrient mixture F-12 (DMEM-F12, Gibco 1130-032) supplemented with 5% horse serum (Gibco 16050), 20 ng/mL epidermal growth factor (Sigma E4127), 0.5 µg/mL hydrocortisone (Sigma H0135), 100 ng/mL Cholera toxin (Sigma C8052), and 10 µg/mL insulin (Sigma I1882). HEK293FT (a generous gift from P. Glazer) and RKO (ATCC CRL-2577) cells were grown in DMEM (Gibco 41965-062) with 10% FBS. All cell lines were grown at 37°C in a humidified incubator with 5% CO_2.

For all experiments, except for the original siRNA screen and luciferase assay, siRNA knockdown was performed in 6-well plates with 1×10^5 cells per well. After 24 hours, siRNAs (30 nM final) were reverse transfected for 72 hours using Lipofectamine®

RNAimax as per the manufacturer’s instructions. The original siRNA screen methods can be found in (Farley-Barnes et al., 2018). The luciferase assay was performed as in (Farley-Barnes et al., 2018). All siRNAs were purchased from Dharmacon (siGENOME, Table 4-5).

Table 4-5. Dharmacon catalog numbers (siGENOME) for each siRNA used in this study.

siRNA target	**Catalog Number**
siCCNA1	M-003204-02-0005
siFBL	M-011269-00-0005
siHMGA2	M-013495-02-0005
siNOL11	M-016695-01-0005
siNT (non-targeting)	D-001810-10-20
siNUSAP1	M-004754-04-0005
siPAX1	M-020120-01-0005
siPAX9	M-012242-01-0005
siRPL5	M-013611-01-0005
siRPS28	M-013679-01-0005
siRPS6	M-003204-01-0005
siRPS9	M-011131-01-0005
siSOX7	M-019017-00-0005
siTRIB3	M-003754-01-0005
siUBE2T	M-004898-01-0005
siUTP4	M-015011-01-0005

Luciferase Assay

The dual-luciferase reporter assay was performed as in (Farley-Barnes et al., 2018) using pHrD-IRES-Luc plasmid from (Ghoshal et al., 2004) and *Renilla* control plasmid from (Freed et al., 2012). Four biological replicates were performed. Statistical significance was calculated by Student's t test using GraphPad Prism, version 7.01 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com. **** $p \leq 0.0001$.

Northern Blots

After 72 hours knockdown, RNA was harvested using TRIzol (Life Technologies 5596018) as per the manufacturer's instructions. Further details and probe sequences are as published in (Farley-Barnes et al., 2018). Statistical significance ($p < 0.05$) was calculated by 2-way ANOVA with Sidak's multiple comparisons test using GraphPad Prism, version 7.01 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com. **** $p \leq 0.0001$, *** $p \leq 0.001$, ** $p \leq 0.01$, and * $p \leq 0.05$. Ratio analysis of multiple precursors (RAMP) was performed as previously described (Wang et al., 2014). Three biological replicates were performed for each northern blot, except Figure 4-7 where 4 replicates were performed.

Puromycin Incorporation Assay

The puromycin incorporation assay was performed as in (Farley-Barnes et al., 2018). Three biological replicates were performed. Significance was calculated by One-way ANOVA using GraphPad Prism, version 7.01 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com. **** $p \leq 0.0001$ and *** $p \leq 0.001$.

Cell Cycle Analysis

After 72 hours of knockdown as described above, cells were pelleted at 400 x g for 3 minutes at 4 °C. Samples were washed with PBS-B [PBS with 1% BSA (Sigma, A9647)] before fixation in 4% formaldehyde (J.T. Baker 2106-01) in PBS for 20 minutes at room temperature. After washing with PBS-B and centrifugation at 400 x g for 3 minutes, cells were resuspended in 90% ice-cold methanol. Cells were then incubated on ice for 30 minutes. After washing twice with PBS-B, cells were resuspended in RNase/PI (BD Pharmgen, Catalog No. 550825) and incubated for 15 minutes at room temperature. Flow data was collected on a Cytek DxP 8 Flow Cytometer and analyzed using FlowJo CE.

qRT-PCR

After 72 hours knockdown, RNA was extracted using TRIzol (Life Technologies 5596018) as per the manufacturer's instructions. cDNA synthesis was performed using the iScriptTM gDNA Clear cDNA Synthesis Kit from Bio-Rad (Cat. No. 172-5035), and the SYBR Green reagent was also purchased from Bio-Rad (Cat. No. 1725121). All $A_{260/230}$ values were above 1.7 prior to cDNA synthesis. Melt curves were performed for each sample to verify the amplification of a single product. Cycling parameters using the Applied Biosystems StepOne Plus are as follows: initial denaturation at 95°C for 30s and 40 cycles of 95 °C for 15s and 55 °C or 58 °C for 30s. Melt curve analysis: 95°C for 15s, then 55 °C or 58 °C for 1 min, and a gradual increase in temperature (0.3°/15s) to 95°C. Analysis was completed using the comparative CT method (ΔΔCT). Three biological replicates, each with three technical replicates were measured for each experiment. Significance was calculated by One-way ANOVA with Dunnett's multiple comparisons test or Student's t test, as indicated in the figure legends, using GraphPad Prism, version 7.01 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com. **** $p \leq 0.0001$, *** $p \leq 0.001$, ** $p \leq 0.01$, and * $p \leq 0.05$. Primer sequences are shown in Table 4-6. Whenever possible, published primers and intron-spanning primers were used.

Table 4-6. List of primers used in this study. For each target, the forward (F primer) and reverse (R primer) sequences are listed in the 5’-3’ direction. Where applicable, the citation or PrimerBank ID number (Spandidos et al., 2010) is also listed. N/a indicates primers designed for this study that were not previously published or included in the PrimerBank database.

Target Gene Name	**F primer (5’-3’)**	**R primer (5’-3’)**	**Citation or PrimerBank ID**
7SL	ATCGGGTGTCCGCACTAAGT	CAGCACGGGAGTTTTGACCT	(Galiveti et al., 2010)
CCNA1	TAGACACCGGCACACTCAAG	AGGAGAGATGAATCTACCAGCAT	161377466c3
FBL	CCTGCGTAATGGAGGACACT	ACTTCGGAGGCAAACACG	n/a
HMGA2	CCCAAAGGCAGCAAAAACAA	GCCTCTTGGCCGTTTTTCTC	(Rice et al., 2013)
NUSAP1	CCCTCAAGTACAGTGACCTGC	TCATTTCCTTTTCTTGCCTCA	(Gordon et al., 2017)
PAX9	GAAAGTTTCTGTCTGGGAGTGC	TCCCAGCTGGTTCACCTC	n/a
RPL5	GCCCAAGAAAGAAGTTAAAAAGAAG	CATAGAAAATTGCTGGGTTTAGC	n/a
RPS28	GGTCTGTCACAGTCTGCTCC	CATCTCAGTTACGTGTGGCG	(Wiza et al., 2014)
RPS6	CTGACGCTCTGGGTGAAGAA	AAACCTTGTTTGTCGTTCCCAC	n/a

RPS9	GCGGAGACCCTTCGAGAAAT	CCTCCAGACCTCACGTTTGT	n/a
SOX7	GCCAAGGACGAGAGGAAAC	GTTGGGGTAGTCCTGCATGT	(Guo et al., 2008)
TRIB3	GCCTTTTTCACTCGGACCCAT	CAGCGAAGACAAAGCGACAC	41327717c3
UBE2T	TTGATTCTGCTGGAAGGATTTG	CAGTTGCGATGTTGAGGGAT	(Luo et al., 2017)

Western blotting

After 72 hours of knockdown, protein was harvested by scraping, rinsed in PBS, and lysed by vortexing in AZ lysis buffer (50 mM Tris pH 7.5, 250 mM NaCl, 1% Igepal, 0.1% SDS, 5 mM EDTA pH 8.0) containing protease inhibitors (cOmpleteTM Protease Inhibitor Cocktail, Roche, 11697498001) for 15 minutes at 4 °C. The lysate was spun at 21,000 x g for 15 minutes at 4 °C. The supernatant was removed, and total protein was quantified by Bradford assay (Bio-Rad). Samples were analyzed via SDS-PAGE. Antibodies used include α-PAX9 (1:10,000, Abcam ab28538), α-β-actin (1: 30,000 Sigma-Aldrich A1978), and α-puromycin (1:10,000 Kerafast 3RH11). HRP conjugated secondary antibodies include α-mouse (1:10,000 GE Healthcare NXA931) and α-rat (1:10,000 GE Healthcare NA935V).

RNA-seq

After 72 hours of knockdown in MCF10A cells using either siRNAs targeting PAX9 or a non-targeting control (siNT), RNA was harvested using TRIZOL according to the manufacturer’s instructions. Agilent BioAnalyzer analysis revealed high quality RNA, with RIN numbers of 10, 10, and 10 for siNT samples and 10, 9.8, and 9.9 for siPAX9 samples. The Yale Center for Genomic Analysis synthesized the PolyA+ mRNA library following Illumina mRNA sample preparation guidelines and also performed the single-end RNA sequencing using the Illumina HiSeq 2500.

Total reads for each of the 6 samples ranged from approximately 19.8 million to 35.2 million, and the average read quality ranged from 36.28/40 to 36.43/40 with a read length of 76 nt. Data analysis was performed using Partek® Flow® software (version 7.0 Copyright; 2018, Partek, Inc., St. Louis, MO, USA). Reads were aligned to the genome (UCSC hg19) using Bowtie 2 (Langmead and Salzberg, 2012) and quantified to the transcriptome using Cufflinks (Trapnell et al., 2010). mRNAs were considered differentially expressed if the fold change was ≥ 2 or ≤ -2 and the q value ≤ 0.05. The functional analyses were generated through the use of IPA (QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis) (Kramer et al., 2014).

RNAPII ChIP-seq

After 72 hours of siRNA knockdown of PAX9 or a non-targeting control siRNA (siNT), MCF10A cells were fixed in 10% fixing solution (11% formaldehyde, 0.1 M NaCl, 1 mM EDTA (pH 8.0), and 50 mM HEPES (pH 7.9)) for 15 minutes at room temperature. Cross-linking was stopped with 5% v/v 2.5 M glycine at room temperature for 5 minutes. Cells were washed once in PBS-Igepal (0.5%). Cells were washed again in PBS-Igepal (0.5%) with 1 mM PMSF. The cells were centrifuged at 800 x g, the supernatant removed, and the cell pellet was flash frozen on dry ice. Further steps of the TranscriptionPath^{TM} RNAPII ChIP-seq protocol were performed by Active Motif. Chromatin immunoprecipitation was completed using an RNAPII antibody (Active Motif, Clone H4B, Cat. No. 39097). Sequencing was performed using the Illumina NextSeq 500 with 75 nt sequence reads.

Data analysis was completed by Active Motif. Total number of reads per sample ranged from 35 to 38 million. The total number of reads aligned to the hg38 reference genome, after removal of duplicates, ranged from approximately 14 million to 18 million. Tags were normalized by random sampling to the sample with the smallest number of unique alignments without duplicate reads (14,165,891). Peak calling was performed using SICER 1.1 with a FDR of $1x10^{-10}$ and a gap size of 600 base pairs. The number of intervals per sample was approximately 24 million. mRNAs were considered differentially expressed if the fold change was ≥ 2 or ≤ -2, and the max tags (5' ends of aligned reads) value was ≥ 150. The functional analyses were generated through the use of IPA

(QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis) (Kramer et al., 2014).

X. tropicalis embryos and MO knockdown

X. tropicalis were housed and cared for in the Yale aquatics facility according to protocols approved by the Yale Institutional Animal Care and Use Committee (IACUC). *X. tropicalis* embryos were produced via *in vitro* fertilization and microinjections were performed as previously described (del Viso and Khokha, 2012). Borosilicate glass needles were used to inject 16 ng of PAX9 MO (5'-TTCCCCGAAAGCGGGTTCCATTGCT-3', Gene-Tools, LLC) at either the one cell or two cell stage with mini-ruby tracer (Invitrogen), except in the northern blot experiment where the PAX9 doses were varied as indicated. After injection, embryos were left in 1/9X MR with 3% Ficoll for 1 hour before being raised to appropriate stages in 1/9X MR with 50 µg/mL gentamycin.

Optical coherence tomography (OCT)

Optical coherence tomography was performed using the Thorlabs Telesto 1325 nm spectral domain-OCT System as previously described (Deniz et al., 2017).

Northern blots on *X. tropicalis* RNA

Northern blotting was performed as described above for mammalian cells, with a few differences. At stage 28, RNA was harvested from injected embryos. Embryos were dissolved in TRIzol (Invitrogen) and total RNA was isolated as described in above in mammalian cells. For each treatment, 4 µg RNA was run on a northern blot that was probed with probe c [5’-CAG GTA CCC GGG TCG GCC TGC GGC G-3’, (Griffin et al., 2015)].

In situ hybridizations

At stage 24, *in situ* hybridizations were performed with digoxigenin-labeled antisense probes as described previously (Khokha et al., 2002, Griffin et al., 2015). Embryos were fixed in Bouin’s solution overnight, bleached under fluorescent light, and imaged as in (Khokha et al., 2002).

Chapter 5

Perspectives and future directions

INTRODUCTION

Ribosome biogenesis is a complex and intricate process that is essential for cellular function. In humans, this process is even more complex due to increased regulation and coordination of cellular stimuli as well as development and differentiation. Therefore, this work sought to identify new regulators of ribosome biogenesis in human cells. Taking advantage of the link between nucleolar structure and function, a high-throughput, genome-wide RNAi screen was performed in human MCF10A breast epithelial cells to identify proteins that, when depleted, changed the number of nucleoli per nucleus from 2-3 to only 1. Using this approach, we identified 139 proteins required for maintaining nucleolar structure (Chapter 2). These proteins are both nucleolar and non-nucleolar, reflecting the increased complexity of regulation present in human cells. Further analysis of 20 high-confidence screen hits revealed new roles for 18 in the nucleolar functions of rDNA transcription, pre-rRNA processing, and/or global cellular translation (Chapter 3). Thus, the screen will continue to serve as a vital resource for studies regarding both the mechanisms of nucleolar formation and the layers of cellular regulation of human ribosome biogenesis.

Further analysis of the screen hit PAX9 also revealed new connections between craniofacial development and ribosome biogenesis. PAX9 plays a role in SSU pre-rRNA processing and global protein synthesis in human tissue culture cells through its function as a RNAPII transcription factor for mRNAs encoding nucleolar proteins (Chapter 4). This function is conserved to the model organism *X. tropicalis*, where PAX9 expression is critical to craniofacial development (Chapter 4). Therefore, this analysis connects PAX9 to both craniofacial development and ribosome biogenesis for the first time.

The use of MCF10A cells revealed unexpected regulators of ribosome biogenesis

The use of MCF10A breast epithelial cells revealed several hits, such as LIN28A, that could not have been identified by screening in other cell lines. MCF10A cells are unique in that they unexpectedly express stem cell markers, including OCT4 and SOX2 (Qu et al., 2015). Similarly, LIN28A is expressed only during embryonic development and not in many human cell lines (Piskounova et al., 2011a), and thus its function in ribosome biogenesis had not been studied in detail. LIN28A localizes to the nucleolus during early mouse development and lack of LIN28A arrests murine development at the 2-4 cell stage transition (Vogt et al., 2012). Based on its ability to bind RNA, LIN28A has been postulated to play a role in pre-rRNA processing (Daley and Sung, 2014). Additionally, it has been shown to bind and enhance the translation of several ribosomal proteins (RPS11/uS17, RPS13/uS15, RPS14/uS11, RPS9/uS4) (Peng et al., 2011) whose depletion causes a pre-rRNA processing defect in HeLa cells (O'Donohue et al., 2010) similar to that of the LIN28A depletion described here (Fig 3-2, 3-3). This work therefore marks the first time that a role for LIN28A in human ribosome biogenesis has been identified, and we have shown that LIN28A depletion does, in fact, alter pre-rRNA processing by causing an accumulation of the 30S pre-rRNA (Fig 3-2, 3-3).

Prolactin (PRL) is another interesting hit that was identified due to the use of MCF10A cells. We have shown here that PRL knockdown in MCF10A cells results in decreased rDNA transcription (Fig 3-1). This unexpected role contributes to a body of work describing an extra-pituitary role for PRL, whose overexpression is associated with increased risk of multiple cancers [reviewed in (Bernard et al., 2015, Marano and Ben-Jonathan, 2014, Sethi et al., 2012)]. Current therapeutics for breast cancer, such as the dopamine-agonist bromocriptine, act at the pituitary level and so would not be effective at targeting endogenous PRL (Sethi et al., 2012). This extrapituitary role for PRL in rDNA transcription could provide insight into why current treatments targeting a secreted prolactin have been unsuccessful.

Potential mechanisms for the "one nucleolus" phenotype

While the screen was effective in identifying a plethora of factors required for human ribosome biogenesis, the precise mechanisms regulating the formation of the "one nucleolus" phenotype remain unknown. Future experiments would be designed to address key questions regarding the mechanisms that dictate the formation of these phase separated bodies.

Future studies regarding the mechanisms of nucleolar formation should focus on analyses of the cell cycle distribution upon depletion of the screen hits. Human cells undergo an open mitosis through which the nucleolus is dissolved and reformed after each cell cycle. Throughout this process, components of the Pol I transcription machinery (i.e. the RNAPI subunits, SL1, and UBTF) remain associated with the rDNA while the remaining nucleolar components diffuse throughout the cell (Roussel et al., 1996). The process of dissolution begins in prophase, when ribosome production is halted (Gebrane-Younes et al., 1997). It is then that the pre-rRNA processing components, found in the GC and DFC of active nucleoli, relocate to form the perichromosomal compartment (Gautier et al., 1992). At the end of mitosis, NORs that remained associated with the Pol I transcription machinery (previously active NORs) resume transcription through the dephosphorylation of several key proteins, including cyclin B-CDK1, by the phosphatases PP1 (Wu et al., 2009) and PP2A (Mochida et al., 2009). The remaining components necessary for nucleolar function are organized into prenucleolar bodies (PNBs). These extranucleolar PNBs contain processing proteins, snoRNAs, ribosomal proteins, and unprocessed pre-rRNAs (Azum-Gelade et al., 1994, Jimenez-Garcia et al., 1994). PNBs, however, are distinct from nucleoli as they do not contain Pol I transcriptional machinery or rDNA (Jimenez-Garcia et al., 1989). Nucleolar fusion then occurs (Savino et al., 2001) to form mature nucleoli through processes that remain largely unknown. Questions remain regarding how many NORs fuse to form one nucleolus and what properties govern the fusion process.

In preliminary studies to address whether the distribution of MCF10A cells within the cell cycle is altered upon reduction in nucleolar number, I depleted MCF10A cells of a few high-confidence screen hits. After staining with propidium iodide, the live cells were analyzed using flow cytometry and the proportion of cells in G1 or G2/M phase was quantified (Fig 5-1). Interestingly, depletion of the screen hits differentially affected the cell cycle distribution. For example, cells depleted of ANLN had an increased proportion of cells in G2/M phase (Fig 5-1). However, cells depleted of either PAX9, APBB1, or APBA1 had an increased population of cells in G1 (Fig 4-10, 5-1). Furthermore, MCF10A cells depleted of CRK had an increased population of cells in S phase (Fig 5-1). Based on these preliminary results, it is unlikely that the reduction in nucleolar number observed in our screen is solely a result of inhibition at a particular point in the cell cycle. However, to fully address this issue, cell cycle analyses will need to be performed after depletion of additional screen hits.

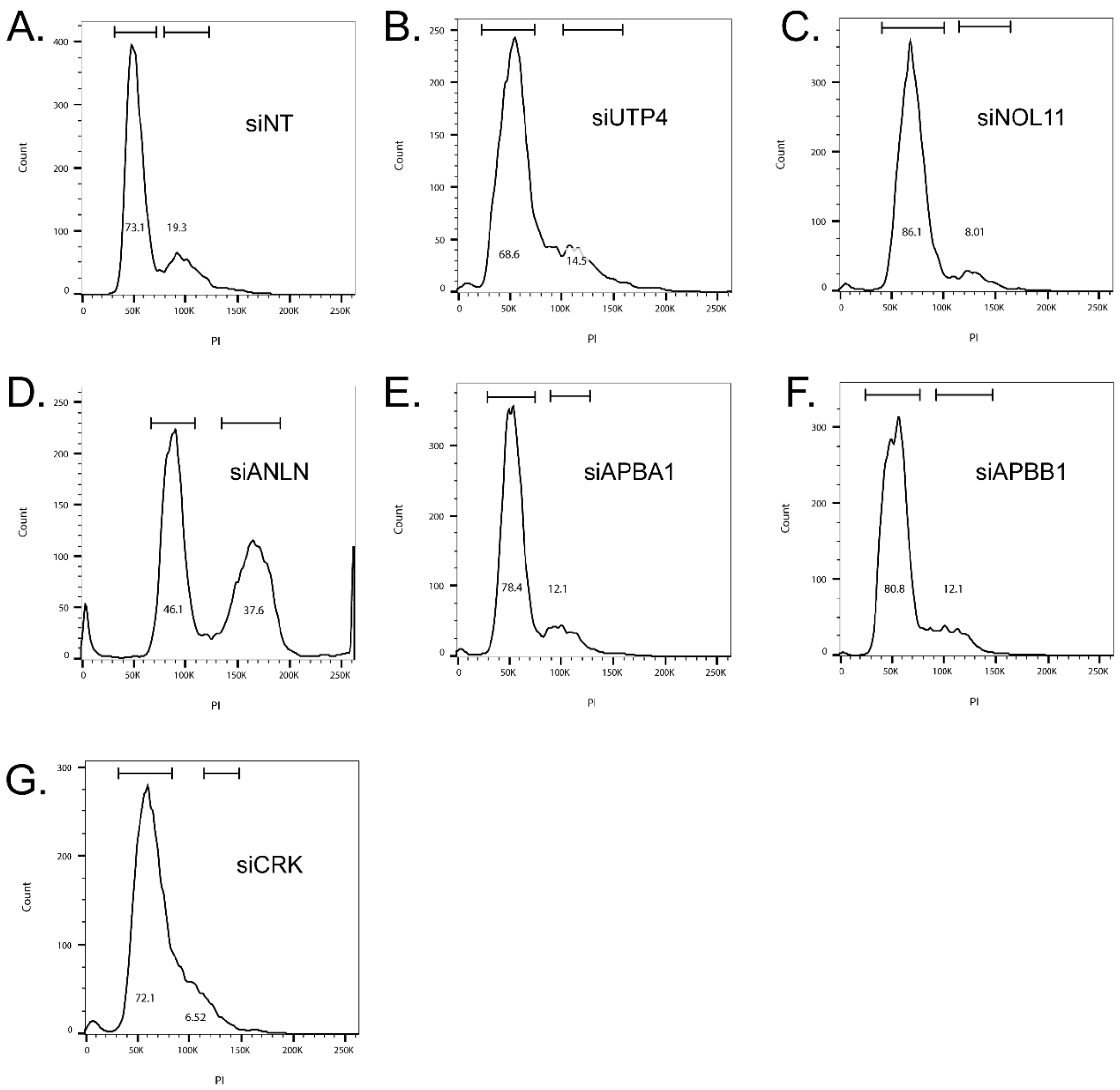

Figure 5-1. Depletion of various screen hits results in changes in the distribution of MCF10A cells within the cell cycle. MCF10A cells were depleted using the indicated siRNAs (A-G) for 72 hours. Cells were stained with propidium iodide (PI), live cells were analyzed by flow cytometry. The proportion of cells in either G1 or G2/M phase were quantified as indicated by the bars above each distribution profile. N =1. siNT and siNOL11 images are the same as in Fig 4-10, because the data were collected as part of the same experiment.

While screening for reduction in FBL-staining bodies has allowed us to identify new regulators of nucleolar function (see Chapters 2 and 3), analysis of whether or not the rDNA is located within these bodies is also essential to understanding the mechanisms of the nucleolar reductions. Throughout this manuscript, the visualization of FBL-staining bodies has been used as a proxy for nucleoli and nucleolar function. However, it is possible that there these bodies are not functional nucleoli. Similar to the single FBL-staining body seen in our screen, studies from the Moss laboratory have identified a nucleolar precursor body (NPB) in mouse embryonic stem cells lacking upstream binding factor (UBF) (Hamdane et al., 2014, Hamdane et al., 2016). Interestingly, while these NPBs contain FBL, they do not contain the rDNA and are thus not functional nucleoli. However, as many of the hits from the screen do not affect rDNA transcription (Fig 3-1), it would be unlikely that the one nucleolus phenotype is solely a measure of NPB formation. Similar to the NPB formation in mouse embryonic stem cells, IMR90 fibroblasts form a single NPM1 staining body upon senescence. However, the rDNA is maintained in the nucleolus of the senescent IMR90 cells (Dillinger et al., 2017). Therefore, future analysis of the cellular mechanisms of nucleolar formation should include fluorescence *in situ* hybridizations (FISH) to visualize the rDNA to determine its presence or absence within the FBL-staining bodies.

If the rDNA is present in these FBL-staining bodies, another pertinent question to address is whether the reduction in nucleolar number observed in our screen is due to multiple NORs coming together to form a single body or due to the inactivation of particular NORs. If it is the latter, are the same NORs being inactivated upon depletion of each hit? As all NORs consist of repeated rDNA, it has previously been difficult to identify individual NORs. However, the McStay laboratory has developed fluorescence *in situ* hybridization (FISH) probes for sequences at the distal and/or proximal junction of each NOR that can be used to address this question more fully in the future (Floutsakou et al., 2013). Additionally, live cell fluorescence microscopy of MCF10A cells with a Defective in DNA Methylation 2 (DIM2)-Snap tagged SSU can be used to observe SSU biogenesis over time after depletion of the screen hits (Landvogt et al., 2019). Finally, it will be interesting to observe the extent to which depletion of the screen hits alters nucleolar area. If multiple NORs are coming together to create a single, larger nucleolus, nucleolar area may also increase. Targeted analysis of a few screen hits, however, suggests that the nucleolar area is unchanged after knockdown (data not shown). Additionally, there is little overlap between our screen hits and those observed in a screen for changes in nucleolar size (Neumuller et al., 2013), although this may be due to differences in screen setup and species. Unfortunately, we were unable to achieve enough separation between the controls in our assay to be able to screen for this parameter in a robust manner. Future studies using these techniques will therefore be useful in addressing the effects of nucleolar reduction at the cellular level.

Cytoskeletal rearrangements provide yet another mechanism through which the cell may regulate nucleolar formation. Many of the proteins we identified in this screen have known roles in cytoskeletal organization and cell division, and those studied functionally (ANLN, NUMA1, IQSEC3) were required for pre-rRNA transcription or processing (Table 3-2). Interestingly, this connects to work by the Brangwynne laboratory that postulates that cells require a scaffolding network within the nucleus to maintain nucleolar position (Feric and Brangwynne, 2013, Shin et al., 2018). Thus, it is possible that disruption of this scaffolding network through depletion of these cytoskeletal-related protein hits causes the nucleoli to merge into one because of gravitational sedimentation forces. Further studies are needed to examine the role of phase separation via gravitational sedimentation in MCF10A cells depleted of the screen hits.

Similar to cytoskeletal organization, nucleolus associated chromatin domains (NADs) have been increasingly tied to nucleolar function. Although NADs are generally gene-poor, NADs do include genes coding for olfactory receptors (Németh et al., 2010, van Koningsbruggen et al., 2010). Interestingly, olfactory receptors comprised many of the screen hits removed from our list due to low expression in breast cells. It is possible that these hits are therefore relevant to nucleolar architecture through their function in genomic organization. Similar to the reduction in nucleolar number observed in our siRNA screen, it has been shown that a single nucleolus forms upon cellular senescence in IMR90 fibroblast cells (Dillinger et al., 2017). However, despite these large structural rearrangements of nucleoli, NADs in senescent cells are largely the same as NADs in younger cells (Dillinger et al., 2017). This argues against NAD rearrangements being causative of reduced nucleolar number. In the future, it would be interesting to analyze the differences in NAD organization between cells with 1 nucleolus, 2-3 nucleoli, and greater than 3 nucleoli per nucleus.

Defining the mechanism of action for PAX9 in ribosome biogenesis and craniofacial development

In Chapter 4 of this manuscript, I identified a new role for PAX9 in human ribosome biogenesis and showed that this role is conserved to the model organism *X. tropicalis*. Analysis of PAX9 depletion in *X. tropicalis* also highlighted Pax9's role in the developing neural crest. However, these experiments do not address whether the disruptions in Pax9's function in ribosome biogenesis are causative of the craniofacial defects. Therefore, future studies should analyze the extent to which the human *PAX9* mutations play a role in the signs and symptoms of the oligodontia and craniofacial malformations. The *X. tropicalis* system will work well for these future experiments. Using the *X. tropicalis* model organism, Pax9 levels can be decreased using the appropriate MO at the 1 cell stage before rescuing the biochemical and physiological defects by injecting the human mRNA at the 2 cell stage. Further experiments can then probe the extent to which the human *PAX9* mutations are able to rescue these phenotypes.

Despite this increased understanding of PAX9's function in ribosome biogenesis, questions remain regarding the tissue specific defects in patients with PAX9 mutations. While one might expect disruptions in such an essential process to affect all tissues, PAX9 mutations typically manifest in craniofacial defects and oligodontia. Similarly, other disruptions in ribosome biogenesis result in specific defects in craniofacial development (see Chapter 1). The most well-studied hypothesis regarding the tissue specificity of these ribosomopathies is through increased p53 levels as part of the nucleolar stress response. In preliminary studies, Maya Overton, an undergraduate student in the Baserga laboratory, analyzed p53 levels after PAX9 depletion in both MCF10A and RKO cells. In RKO cells, PAX9 depletion does not result in a significant change in p53 levels (Fig 5-2). However, no conclusions can be drawn in MCF10A cells due to lack of statistical power (data not shown). Therefore, PAX9 depletion likely does not induce the p53-mediated nucleolar stress response in human cells. This effect is similar to that of apoptosis-antagonizing transcription factor (AATF), a RNAPII transcription factor whose depletion results in a similar increase in the 30S pre-rRNA processing intermediate (Bammert et al., 2016). The SSU biogenesis defects observed upon AATF depletion in HeLa cells are independent of p53 status, despite a known role as a regulator of p53 transcription (Bammert et al., 2016). However, it is still possible that the cells may be undergoing stress and apoptosis via another stress response pathway such as p27kip1, E2F, pRB, or p21 (James et al., 2014, Yang et al., 2018). More tests are underway to identify the extent to which these other stress response pathways are activated.

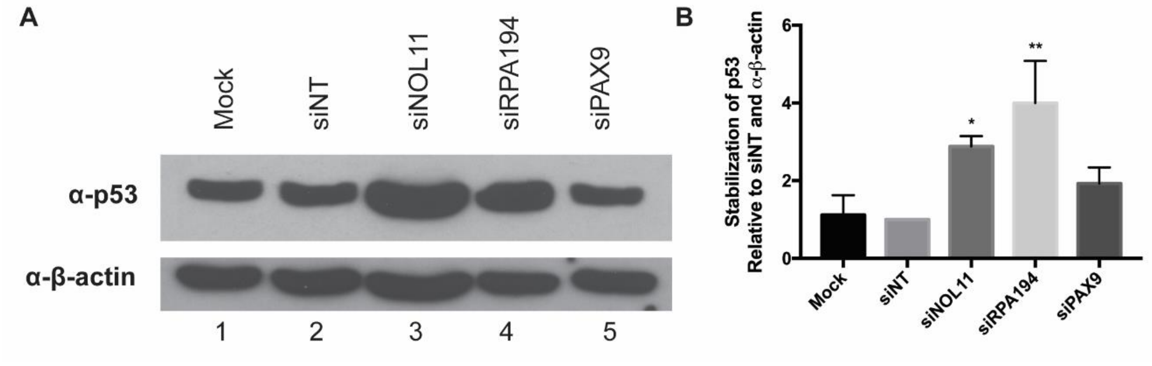

Figure 5-2. PAX9 depletion does not significantly increase p53 levels in RKO cells. A) Western blot of p53 protein levels in RKO cells treated with siRNAs targeting the indicated proteins (lanes 1-5) for 72 hours. An antibody against β-actin was used as a loading control. Mock and siNT were used as negative controls. siNOL11 and siRPA194 were used as positive controls. B) Quantitation of the northern blots shown in (A). N = 3. Data are shown as mean ± SEM, relative to the siNT control. Data were analyzed by One-way ANOVA using Graphpad Prism. * $p \leq 0.05$, ** $p \leq 0.01$.

Future studies are also needed to examine the extent to which p53 levels are affected upon Pax9 depletion in the *X. tropicalis* model system. It is possible that, despite the results in human tissue culture cells, neural crest cells depleted of PAX9 *in vivo* do undergo the p53-mediated nucleolar stress response. This could account for the craniofacial abnormalities present in *X. tropicalis* embryos depleted of PAX9. However, the phenotype seen upon Pax9 depletion affects mostly the development of the Meckel's and ceratohyal cartilages (Fig 4-19). This is different than the general reduction in size of all craniofacial cartilages seen after Nol11 depletion in *X. tropicalis* (Griffin et al., 2015). In this system, Nol11 depletion results in increased p53 levels and apoptosis of the neural crest cells. Therefore, if the ribosome biogenesis defects seen after Pax9 depletion were to also result in p53-mediated nucleolar stress response in neural crest cells, it would be expected that the phenotype would be the similar to that of Nol11 depletion. Further analyses in developing embryos of both p53 levels by western blotting and in apoptosis by TUNEL staining will therefore be critical in understanding the PAX9's role in craniofacial development.

In addition to the p53 stress response pathway, the Wnt signaling pathway provides another mechanism through which PAX9's role in ribosome biogenesis may contribute to the pathophysiology of the *PAX9* mutation. Wnt signaling influences many cell growth processes such as cell cycle progression, cell growth, and ribosome biogenesis. Ribosome biogenesis factors that are molecular targets of Wnt signaling include NPM1, Pescadillo 1 (PES1), and BOP1 (block of proliferation 1) [reviewed in (Pfister and Kühl, 2018)]. Interestingly, *TCOF1*, the gene mutated in Treacher Collins syndrome, also interacts with the Wnt signaling pathway (Dai et al., 2016). Wnt signaling is affected upon PAX9 depletion in mice (Jia et al., 2017a, Li et al., 2017) and in our analysis in human tissue culture cells (Chapter 4). Therefore, further exploration of the degree to which Wnt signaling plays a role in PAX9's function in ribosome biogenesis is needed.

PAX9, ribosome biogenesis, and cancer

The importance of the nucleolus in cancer was realized as early as 1896, when it was noted that malignant cells had large, irregular nucleoli (Pianese, 1896). With the advent of silver-staining, Ploton et al. examined human prostatic cancer cells and discovered that the malignant cells had larger nucleoli with more silver stained dots than benign hyperplastic glands and normal lymphocytes (Ploton et al., 1986). After this study, nucleolar size was found to be an accurate prognostic indicator of clinical outcome in a number of other cancers [reviewed in (Derenzini et al., 2009, Penzo et al., 2019); (Pich et al., 2000)]. Nucleolar alterations in cancer have therefore long been observed, but questions remain as to the role of the nucleolus in cancer.

Previously, changes in nucleolar structure were thought of as a byproduct of cell transformation, but could changes in the structure and function of the nucleolus drive transformation as well? Recent insights suggest that changes in proteins which affect nucleolar size/number and function can drive cancer. For example, depletion of the cell-cycle control protein ADP ribosylation factor like 2 (Arl-2) resulted in increased ribosome biogenesis, nucleolar number, nucleolar area, and tumor aggressivity (Belin et al., 2009). Bystin-like (BYSL), a protein involved in pre-18S rRNA processing, may also play a role in driving tumor formation as its inhibition has been shown to prevent tumor formation in nude mice (Wang et al., 2009a). In addition, an increased susceptibility to cancer exists in patients with many ribosomopathies, such as Diamond Blackfan anemia (Vlachos et al., 2012). The precise mechanisms describing how the proteins mutated in ribosomopathies result in cancer remain to be defined. It is also possible, that defects in these ribosome biogenesis proteins drive cancer in an indirect manner. For example, the ribosome biogenesis proteins involved may have extra-ribosomal functions which contribute to cancer development, the overall decrease in available ribosomes could alter translation of genes involved in transformation, or byproducts of ribosome biogenesis defects may cause transformation (Montanaro et al., 2008). Regardless, more information must be provided to elucidate the crucial role of the nucleolus in cancer.

This new function of PAX9 in ribosome biogenesis may also provide new insight into our understanding of the link between making ribosomes and cancer. PAX9 expression promotes cellular proliferation, and its loss triggers the induction of apoptosis (Lee et al., 2008). Therefore, one might expect that PAX9 overexpression would be associated with an increased cancer risk. Indeed, enhanced amplification frequencies of PAX9 were detected in 35% of esophageal squamous cell carcinoma tumors (Bandla et al., 2012). In addition, PAX9 expression is increased in epithelial ovarian cancer cell lines, breast cancer cell lines, lung cancer, and advanced nodular malignant melanomas (Muratovska et al., 2003, Bonds et al., 2014, Roesch et al., 2008, Kendall et al., 2007), in agreement with the results presented in Chapter 4 showing a role for PAX9 in enhancing ribosome biogenesis. Further studies are therefore needed to test the extent to which increased PAX9 expression may contribute to the development of cancer.

In contrast to the data suggesting an oncogenic role for PAX9, PAX9 may be playing a tumor suppressive role in esophageal and oral carcinomas. Increased malignancy is associated with PAX9 loss in esophageal squamous cell carcinoma and oral squamous cell carcinoma (Gerber et al., 2002, Tan et al., 2017). This contradicts what would be expected for a protein that enhances ribosome biogenesis, as the nucleolar dysmorphology observed in tumor cells is thought to reflect the rate of cellular proliferation (Derenzini et al., 1998). However, while nucleolar size does function as an accurate prognostic indicator of malignant vs benign lesions, the nucleolar size parameter is not an accurate diagnostic tool (Derenzini et al., 2009). This is because not all tumors proliferate rapidly, so many of the cells within the tumor may contain small nucleoli while the tumor is still classified as malignant (Derenzini et al., 1998). Therefore, many questions remain regarding both PAX9's role in cancer development and alterations in nucleolar function as a driver of malignancy.

PAX9, ribosome biogenesis, and alcohol exposure

Fetal alcohol spectrum disorder due to maternal alcohol abuse manifests in a number of signs and symptoms similar to those seen in ribosomopathies. Infants with fetal alcohol spectrum disorder present with neurological impairment as well as numerous congenital defects including short palpebral fissures, cardiac and skeletal malformations, conductive hearing loss, and growth deficiency (Hoyme et al., 2016). In a mechanism similar to that of several ribosomopathies (see Chapter 1), alcohol exposure during development results in increased apoptosis of neural crest cells, affecting the development of the frontonasal prominence and pharyngeal arches [reviewed in (Petrelli et al., 2019). Insights into the role of PAX9 in human ribosome biogenesis may therefore provide a mechanism through which the dysmorphologies seen in fetal alcohol syndrome occur.

PAX9 has been connected to alcohol exposure through the study of oral and esophageal cancers. PAX9 expression is downregulated in both normal and cancerous esophageal tissues of alcohol drinkers as compared to non-drinkers (Xiong et al., 2018). Indeed, PAX9 expression is decreased in the forestomach of mice exposed to 15% ethanol for 40 weeks, and this correlates with increased forestomach tumorigenesis in those mice upon addition of the carcinogen N-nitrosomethylbenzylamine (Xiong et al., 2018). However, this response was specific to the forestomach, and PAX9 expression was unchanged in the esophagus upon ethanol exposure (Xiong et al., 2018).

Additionally, questions remain as to the mechanism of action regulating the decrease in PAX9 expression after alcohol exposure.

Recent studies have probed the connection between ethanol exposure and ribosome biogenesis. Alcohol exposure in zebrafish causes craniofacial dysmorphology including cranial skeletal asymmetry, reduced ocular size, and hypotelorism (Berres et al., 2017). These defects are further enhanced upon co-depletion of ribosome biogenesis factors (Berres et al., 2017). Notably, one of the most prominent defects in zebrafish exposed to ethanol is reduced size of the Meckel’s cartilage (Berres et al., 2017), which is also affected upon PAX9 knockdown in *X. tropicalis* (Fig 4-19). In addition, many (197, 11.8%) of the mRNAs differentially expressed upon PAX9 knockdown in MCF10A cells are also differentially expressed upon alcohol exposure in zebrafish (Berres et al., 2017). While PAX9 was not significantly depleted upon alcohol exposure, several potential targets of PAX9 were differentially expressed, including RPS6/eS6, RPS28/eS28, RPL5/uL18, and UBE2T [Chapter 4 (Berres et al., 2017)]. Both RPS6/eS6 and RPS28/eS28 individually result in the same SSU pre-rRNA processing defect as PAX9 depletion (Fig 4-11). Therefore, the differential regulation of RPS6/eS6 and/or RPS28/eS28 could mechanistically link alcohol exposure, PAX9 downregulation, and the craniofacial malformations observed in fetal alcohol spectrum disorder.

REFERENCES

2017. Expansion of the Gene Ontology knowledgebase and resources. *Nucleic Acids Res,* 45**,** D331-D338. PMC5210579

Ahmad, Y., Boisvert, F. M., Gregor, P., Cobley, A. & Lamond, A. I. 2009. NOPdb: Nucleolar Proteome Database--2008 update. *Nucleic Acids Res,* 37**,** D181-4. PMC2686505

Alter, B. P., Giri, N., Savage, S. A. & Rosenberg, P. S. 2018. Cancer in the National Cancer Institute inherited bone marrow failure syndrome cohort after fifteen years of follow-up. *Haematologica,* 103**,** 30-39. PMC5777188

Altug Teber, O., Gillessen-Kaesbach, G., Fischer, S., Bohringer, S., Albrecht, B., Albert, A., Arslan-Kirchner, M., Haan, E., Hagedorn-Greiwe, M., Hammans, C., Henn, W., Hinkel, G. K., Konig, R., Kunstmann, E., Kunze, J., Neumann, L. M., Prott, E.-C., Rauch, A., Rott, H.-D., Seidel, H., Spranger, S., Sprengel, M., Zoll, B., Lohmann, D. R. & Wieczorek, D. 2004. Genotyping in 46 patients with tentative diagnosis of Treacher Collins syndrome revealed unexpected phenotypic variation. *Eur J Hum Genet,* 12**,** 879-890.

Anderson, S. J., Sikes, M. L., Zhang, Y., French, S. L., Salgia, S., Beyer, A. L., Nomura, M. & Schneider, D. A. 2011. The transcription elongation factor Spt5 influences transcription by RNA polymerase I positively and negatively. *J Biol Chem,* 286**,** 18816-24. PMC3099698

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M. & Sherlock, G. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet,* 25**,** 25-9. PMC3037419

Aspesi, A., Pavesi, E., Robotti, E., Crescitelli, R., Boria, I., Avondo, F., Moniz, H., Da Costa, L., Mohandas, N., Roncaglia, P., Ramenghi, U., Ronchi, A., Gustincich, S., Merlin, S., Marengo, E., Ellis, S. R., Follenzi, A., Santoro, C. & Dianzani, I. 2014. Dissecting the transcriptional phenotype of ribosomal protein deficiency: implications for Diamond-Blackfan Anemia. *Gene,* 545**,** 282-289.

Aubert, M., O'donohue, M.-F., Lebaron, S. & Gleizes, P.-E. 2018. Pre-Ribosomal RNA Processing in Human Cells: From Mechanisms to Congenital Diseases. *Biomolecules,* 8**,** 123.

Azum-Gelade, M. C., Noaillac-Depeyre, J., Caizergues-Ferrer, M. & Gas, N. 1994. Cell cycle redistribution of U3 snRNA and fibrillarin. Presence in the cytoplasmic nucleolus remnant and in the prenucleolar bodies at telophase. *J Cell Sci,* 107 (Pt 2)**,** 463-75.

Badertscher, L., Wild, T., Montellese, C., Alexander, L. T., Bammert, L., Sarazova, M., Stebler, M., Csucs, G., Mayer, T. U., Zamboni, N., Zemp, I., Horvath, P. & Kutay, U. 2015. Genome-wide RNAi Screening Identifies Protein Modules Required for 40S Subunit Synthesis in Human Cells. *Cell Rep,* 13**,** 2879-91.

Bailey, T. L. & Machanick, P. 2012. Inferring direct DNA binding from ChIP-seq. *Nucleic Acids Research,* 40**,** e128-e128.

Ball, S. E., Mcguckin, C. P., Jenkins, G. & Gordon-Smith, E. C. 1996. Diamond-Blackfan anaemia in the U.K.: analysis of 80 cases from a 20-year birth cohort. *Br J Haematol,* 94**,** 645-53.

Balzer, E. & Moss, E. G. 2007. Localization of the developmental timing regulator Lin28 to mRNP complexes, P-bodies and stress granules. *RNA Biol,* 4**,** 16-25.

Bammert, L., Jonas, S., Ungricht, R. & Kutay, U. 2016. Human AATF/Che-1 forms a nucleolar protein complex with NGDN and NOL10 required for 40S ribosomal subunit synthesis. *Nucleic acids research,* 44**,** 9803-9820.

Bandla, S., Pennathur, A., Luketich, J. D., Beer, D. G., Lin, L., Bass, A. J., Godfrey, T. E. & Litle, V. R. 2012. Comparative genomics of esophageal adenocarcinoma and squamous cell carcinoma. *The Annals of thoracic surgery,* 93**,** 1101-1106.

Barlow, J. L., Drynan, L. F., Hewett, D. R., Holmes, L. R., Lorenzo-Abalde, S., Lane, A. L., Jolin, H. E., Pannell, R., Middleton, A. J., Wong, S. H., Warren, A. J., Wainscoat, J. S., Boultwood, J. & Mckenzie, A. N. 2010. A p53-dependent mechanism underlies macrocytic anemia in a mouse model of human 5q-syndrome. *Nat Med,* 16**,** 59-66. PMC2803774

Barr, F. G., Galili, N., Holick, J., Biegel, J. A., Rovera, G. & Emanuel, B. S. 1993. Rearrangement of the PAX3 paired box gene in the paediatric solid tumour alveolar rhabdomyosarcoma. *Nature Genetics,* 3**,** 113-7.

Barr, M. L. & Bertram, E. G. 1949. A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. *Nature,* 163**,** 676-677.

Belin, S., Beghin, A., Solano-Gonzalez, E., Bezin, L., Brunet-Manquat, S., Textoris, J., Prats, A. C., Mertani, H. C., Dumontet, C. & Diaz, J. J. 2009. Dysregulation of ribosome biogenesis and translational capacity is associated with tumor progression of human breast cancer cells. *PLoS One,* 4**,** e7147.

Bell, S. P., Learned, R. M., Jantzen, H. M. & Tjian, R. 1988. Functional cooperativity between transcription factors UBF1 and SL1 mediates human ribosomal RNA synthesis. *Science,* 241**,** 1192-7.

Bernard, V., Young, J., Chanson, P. & Binart, N. 2015. New insights in prolactin: pathological implications. *Nat Rev Endocrinol,* 11**,** 265-275.

Bernstein, K. A., Bleichert, F., Bean, J. M., Cross, F. R. & Baserga, S. J. 2007. Ribosome biogenesis is sensed at the Start cell cycle checkpoint. *Mol Biol Cell,* 18**,** 953-64.

Berres, M. E., Garic, A., Flentke, G. R. & Smith, S. M. 2017. Transcriptome Profiling Identifies Ribosome Biogenesis as a Target of Alcohol Teratogenicity and Vulnerability during Early Embryogenesis. *PLoS One,* 12**,** e0169351. PMC5207668

Betard, C., Rasquin-Weber, A., Brewer, C., Drouin, E., Clark, S., Verner, A., Darmond-Zwaig, C., Fortin, J., Mercier, J., Chagnon, P., Fujiwara, T. M., Morgan, K., Richter, A., Hudson, T. J. & Mitchell, G. A. 2000. Localization of a recessive gene for North American Indian childhood cirrhosis to chromosome region 16q22-and identification of a shared haplotype. *Am J Hum Genet,* 67**,** 222-8. PMC1287080

Bonds, J., Pollan-White, S., Xiang, L., Mues, G. & D'souza, R. 2014. Is there a link between ovarian cancer and tooth agenesis? *European journal of medical genetics,* 57**,** 235-239.

Bopp, D., Burri, M., Baumgartner, S., Frigerio, G. & Noll, M. 1986. Conservation of a large protein domain in the segmentation gene paired and in functionally related genes of Drosophila. *Cell,* 47**,** 1033-1040.

Boulon, S., Westman, B. J., Hutten, S., Boisvert, F. M. & Lamond, A. I. 2010. The nucleolus under stress. *Mol Cell,* 40**,** 216-27. PMC2987465

Bowman, M., Oldridge, M., Archer, C., O'rourke, A., Mcparland, J., Brekelmans, R., Seller, A. & Lester, T. 2012. Gross deletions in TCOF1 are a cause of Treacher-Collins-Franceschetti syndrome. *Eur J Hum Genet,* 20**,** 769-777.

Brangwynne, C. P., Mitchison, T. J. & Hyman, A. A. 2011. Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. *Proceedings of the National Academy of Sciences,* 108**,** 4334-4339.

Brannan, K. W., Jin, W., Huelga, S. C., Banks, C. A., Gilmore, J. M., Florens, L., Washburn, M. P., Van Nostrand, E. L., Pratt, G. A., Schwinn, M. K., Daniels, D. L. & Yeo, G. W. 2016. SONAR Discovers RNA-Binding Proteins from Analysis of Large-Scale Protein-Protein Interactomes. *Mol Cell,* 64**,** 282-293. PMC5074894

Capo-Chichi, J.-M., Bharti, S. K., Sommers, J. A., Yammine, T., Chouery, E., Patry, L., Rouleau, G. A., Samuels, M. E., Hamdan, F. F., Michaud, J. L., Brosh, R. M., Jr., Mégarbane, A. & Kibar, Z. 2013. Identification and biochemical characterization of a novel mutation in DDX11 causing Warsaw breakage syndrome. *Human mutation,* 34**,** 103-107.

Carmo-Fonseca, M., Mendes-Soares, L. & Campos, I. 2000. To be or not to be in the nucleolus. *Nat Cell Biol,* 2**,** E107-E112.

Carneiro, M., Hu, D., Archer, J., Feng, C., Afonso, S., Chen, C., Blanco-Aguiar, J. A., Garreau, H., Boucher, S., Ferreira, P. G., Ferrand, N., Rubin, C. J. & Andersson, L. 2017. Dwarfism and Altered Craniofacial Development in Rabbits Is Caused by a 12.1 kb Deletion at the HMGA2 Locus. *Genetics,* 205**,** 955-965. PMC5289862

Carpenter, A. E., Jones, T. R., Lamprecht, M. R., Clarke, C., Kang, I. H., Friman, O., Guertin, D. A., Chang, J. H., Lindquist, R. A., Moffat, J., Golland, P. & Sabatini, D. M. 2006. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol,* 7**,** R100. Pmc1794559

Chagnon, P., Michaud, J., Mitchell, G., Mercier, J., Marion, J. F., Drouin, E., Rasquin-Weber, A., Hudson, T. J. & Richter, A. 2002. A missense mutation (R565W) in cirhin (FLJ14728) in North American Indian childhood cirrhosis. *Am J Hum Genet,* 71**,** 1443-9.

Chang, C. S., Clarke, A. & Pillus, L. 2012. Suppression Analysis of esa1 Mutants in Saccharomyces cerevisiae Links NAB3 to Transcriptional Silencing and Nucleolar Functions. *G3: Genes|Genomes|Genetics,* 2**,** 1223-1232.

Chang, S.-F., Chang, T.-K., Peng, H.-H., Yeh, Y.-T., Lee, D.-Y., Yeh, C.-R., Zhou, J., Cheng, C.-K., Chang, C. A. & Chiu, J.-J. 2009. BMP-4 induction of arrest and differentiation of osteoblast-like cells via p21 CIP1 and p27 KIP1 regulation. *Molecular endocrinology (Baltimore, Md.),* 23**,** 1827-1838.

Chen, J. Y., Tan, X., Wang, Z. H., Liu, Y. Z., Zhou, J. F., Rong, X. Z., Lu, L. & Li, Y. 2018. The ribosome biogenesis protein Esf1 is essential for pharyngeal cartilage formation in zebrafish. *FEBS J,* 285**,** 3464-3483.

Chen, L. L. & Carmichael, G. G. 2009. Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Mol Cell,* 35**,** 467-78. PMC2749223

Cheutin, T., O'donohue, M.-F., Beorchia, A., Vandelaer, M., Kaplan, H., Deféver, B., Ploton, D. & Thiry, M. 2002. Three-dimensional organization of active rRNA genes within the nucleolus. *Journal of Cell Science,* 115**,** 3297-3307.

Chi, N. & Epstein, J. A. 2002. Getting your Pax straight: Pax proteins in development and disease. *Trends in Genetics,* 18**,** 41-47.

Ciganda, M. & Williams, N. 2011. Eukaryotic 5S rRNA biogenesis. *Wiley interdisciplinary reviews. RNA,* 2**,** 523-533.

Clarke, A. S., Samal, E. & Pillus, L. 2006. Distinct roles for the essential MYST family HAT Esa1p in transcriptional silencing. *Mol Biol Cell,* 17**,** 1744-57. PMC1415314

Consortium, T. G. O. 2014. Gene Ontology Consortium: going forward. *Nucleic Acids Research,* 43**,** D1049-D1056.

Cook, K. B., Kazan, H., Zuberi, K., Morris, Q. & Hughes, T. R. 2011. RBPDB: a database of RNA-binding specificities. *Nucleic Acids Research,* 39**,** D301-D308.

Czerny, T., Schaffner, G. & Busslinger, M. 1993. DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site. *Genes Dev,* 7**,** 2048-61.

Da Costa, L., Narla, A. & Mohandas, N. 2018. An update on the pathogenesis and diagnosis of Diamond-Blackfan anemia. *F1000Res,* 7. PMC6117846

Da Silva Dalben, G., Costa, B. & Gomide, M. R. 2006. Prevalence of dental anomalies, ectopic eruption and associated oral malformations in subjects with Treacher Collins syndrome. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology,* 101**,** 588-592.

Dai, J., Si, J., Wang, M., Huang, L., Fang, B., Shi, J., Wang, X. & Shen, G. 2016. Tcof1-Related Molecular Networks in Treacher Collins Syndrome. *J Craniofac Surg,* 27**,** 1420-6.

Daley, J. M. & Sung, P. 2014. 53BP1, BRCA1, and the Choice between Recombination and End Joining at DNA Double-Strand Breaks. *Mol Cell Biol,* 34**,** 1380-8.

Danilova, N., Sakamoto, K. M. & Lin, S. 2008. Ribosomal protein S19 deficiency in zebrafish leads to developmental abnormalities and defective erythropoiesis through activation of p53 protein family. *Blood,* 112**,** 5228-37.

Das, P., Hai, M., Elcock, C., Leal, S. M., Brown, D. T., Brook, A. H. & Patel, P. I. 2003. Novel missense mutations and a 288-bp exonic insertion in PAX9 in families with autosomal dominant hypodontia. *Am J Med Genet A,* 118A**,** 35-42.

Dauwerse, J. G., Dixon, J., Seland, S., Ruivenkamp, C. A., Van Haeringen, A., Hoefsloot, L. H., Peters, D. J., Boers, A. C., Daumer-Haas, C., Maiwald, R., Zweier, C., Kerr, B., Cobo, A. M., Toral, J. F., Hoogeboom, A. J., Lohmann, D. R., Hehr, U., Dixon, M. J., Breuning, M. H. & Wieczorek, D. 2011. Mutations in genes encoding subunits of RNA polymerases I and III cause Treacher Collins syndrome. *Nat Genet,* 43**,** 20-2.

Daw, E. M., Saliba, C., Grech, G. & Camilleri, S. 2017. A novel PAX9 mutation causing oligodontia. *Arch Oral Biol,* 84**,** 100-105.

De Keersmaecker, K., Atak, Z. K., Li, N., Vicente, C., Patchett, S., Girardi, T., Gianfelici, V., Geerdens, E., Clappier, E., Porcu, M., Lahortiga, I., Luca, R., Yan, J., Hulselmans, G., Vranckx, H., Vandepoel, R., Sweron, B., Jacobs, K., Mentens, N., Wlodarska, I., Cauwelier, B., Cloos, J., Soulier, J., Uyttebroeck, A., Bagni, C., Hassan, B. A., Vandenberghe, P., Johnson, A. W., Aerts, S. & Cools, J. 2013. Exome sequencing identifies mutation in CNOT3 and ribosomal genes RPL5 and RPL10 in T-cell acute lymphoblastic leukemia. *Nat Genet,* 45**,** 186-190.

De Keersmaecker, K., Sulima, S. O. & Dinman, J. D. 2015. Ribosomopathies and the paradox of cellular hypo- to hyperproliferation. *Blood,* 125**,** 1377-1382.

Del Viso, F. & Khokha, M. 2012. Generating diploid embryos from Xenopus tropicalis. *Methods Mol Biol,* 917**,** 33-41.

Deniz, E., Jonas, S., Hooper, M., N. Griffin, J., Choma, M. A. & Khokha, M. K. 2017. Analysis of Craniocardiac Malformations in Xenopus using Optical Coherence Tomography. *Scientific Reports,* 7**,** 42506. PMC5307353

Derenzini, M., Montanaro, L. & Trere, D. 2009. What the nucleolus says to a tumour pathologist. *Histopathology,* 54**,** 753-62.

Derenzini, M., Trere, D., Pession, A., Montanaro, L., Sirri, V. & Ochs, R. L. 1998. Nucleolar function and size in cancer cells. *Am J Pathol,* 152**,** 1291-7. PMC1858570

Dillinger, S., Straub, T. & Nemeth, A. 2017. Nucleolus association of chromosomal domains is largely maintained in cellular senescence despite massive nuclear reorganisation. *PLoS One,* 12**,** e0178821. PMC5456395

Dixon, J., Jones, N. C., Sandell, L. L., Jayasinghe, S. M., Crane, J., Rey, J.-P., Dixon, M. J. & Trainor, P. A. 2006. Tcof1/Treacle is required for neural crest cell formation and proliferation deficiencies that cause craniofacial abnormalities. *Proceedings of the National Academy of Sciences,* 103**,** 13403-13408.

Dokal, I. 1996. Dyskeratosis congenita: an inherited bone marrow failure syndrome. *Br J Haematol,* 92**,** 775-9.

Dokal, I. 2000. Dyskeratosis congenita in all its forms. *British Journal of Haematology,* 110**,** 768-779.

Donati, G., Peddigari, S., Mercer, C. A. & Thomas, G. 2013. 5S ribosomal RNA is an essential component of a nascent ribosomal precursor complex that regulates the Hdm2-p53 checkpoint. *Cell Rep,* 4**,** 87-98.

Drouin, E., Russo, P., Tuchweber, B., Mitchell, G. & Rasquin-Weber, A. 2000. North American Indian cirrhosis in children: a review of 30 cases. *J Pediatr Gastroenterol Nutr,* 31**,** 395-404.

Eberhard, D., Jiménez, G., Heavey, B. & Busslinger, M. 2000. Transcriptional repression by Pax5 (BSAP) through interaction with corepressors of the Groucho family. *The EMBO Journal,* 19**,** 2292-303. PMC384353

Edwards, S. J., Fowlie, A., Cust, M. P., Liu, D. T., Young, I. D. & Dixon, M. J. 1996. Prenatal diagnosis in Treacher Collins syndrome using combined linkage analysis and ultrasound imaging. *J Med Genet,* 33**,** 603-6. PMC1050672

Epstein, J., Cai, J., Glaser, T., Jepeal, L. & Maas, R. 1994. Identification of a Pax paired domain recognition sequence and evidence for DNA-dependent conformational changes. *Journal of Biological Chemistry,* 269**,** 8355-8361.

Farley-Barnes, K. I., Mccann, K. L., Ogawa, L. M., Merkel, J., Surovtseva, Y. V. & Baserga, S. J. 2018. Diverse Regulators of Human Ribosome Biogenesis Discovered by Changes in Nucleolar Number. *Cell Rep,* 22**,** 1923-1934. PMC5828527

Farley, K., Surovtseva, Y., Merkel, J. & Baserga, S. 2015. Determinants of mammalian nucleolar architecture. *Chromosoma,* 124**,** 323-331.

Farley, K. I. & Baserga, S. J. 2016. Probing the mechanisms underlying human diseases in making ribosomes. *Biochem Soc Trans,* 44**,** 1035-44. PMC5360156

Fauzi, N. H., Ardini, Y. D., Zainuddin, Z. & Lestari, W. 2018. A review on non-syndromic tooth agenesis associated with PAX9 mutations. *Jpn Dent Sci Rev,* 54**,** 30-36. PMC5884223

Feric, M. & Brangwynne, C. P. 2013. A nuclear F-actin scaffold stabilizes ribonucleoprotein droplets against gravity in large cells. *Nat Cell Biol,* 15**,** 1253-1259.

Feric, M., Vaidya, N., Harmon, T. S., Mitrea, D. M., Zhu, L., Richardson, T. M., Kriwacki, R. W., Pappu, R. V. & Brangwynne, C. P. 2016. Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell,* 165**,** 1686-97. PMC5127388

Floutsakou, I., Agrawal, S., Nguyen, T. T., Seoighe, C., Ganley, A. R. & Mcstay, B. 2013. The shared genomic architecture of human nucleolar organizer regions. *Genome Res,* 23**,** 2003-12. PMC3847771

Fontana, F. & Nyon, J. L. 1781. *Traité sur le vénin de la vipere sur les poisons americains sur le laurier-cerise et sur quelques autres poisons végetaux: on y a joint des observations sur la structure primitive du corps animal : Différentes expériences sur la reproduction des nerfs et la description d'un nouveau canal de l'oeil*, chez Nyon l'Ainé.

Frazier-Bowers, S. A., Guo, D. C., Cavender, A., Xue, L., Evans, B., King, T., Milewicz, D. & D'souza, R. N. 2002. A novel mutation in human PAX9 causes molar oligodontia. *J Dent Res,* 81**,** 129-33.

Freed, E. F. & Baserga, S. J. 2010. The C-terminus of Utp4, mutated in childhood cirrhosis, is essential for ribosome biogenesis. *Nucleic Acids Res,* 38**,** 4798-806.

Freed, E. F., Prieto, J. L., Mccann, K. L., Mcstay, B. & Baserga, S. J. 2012. NOL11, implicated in the pathogenesis of North American Indian childhood cirrhosis, is required for pre-rRNA transcription and processing. *PLoS Genet,* 8**,** e1002892. PMC3420923

Fumagalli, S., Ivanenkov, V. V., 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**,** 1028-40. PMC3360559

Galiveti, C. R., Rozhdestvensky, T. S., Brosius, J., Lehrach, H. & Konthur, Z. 2010. Application of housekeeping npcRNAs for quantitative expression analysis of human transcriptome by real-time PCR. *RNA,* 16**,** 450-461.

Gallagher, J. E., Dunbar, D. A., Granneman, S., Mitchell, B. M., Osheim, Y., Beyer, A. L. & Baserga, S. J. 2004. RNA polymerase I transcription and pre-rRNA processing are linked by specific SSU processome components. *Genes Dev,* 18**,** 2506-17. PMC529538

Gates, R. 1939. Nucleoli, satellites, and sex chromosomes. *Nature***,** 794-795.

Gautier, T., Robert-Nicoud, M., Guilly, M. N. & Hernandez-Verdun, D. 1992. Relocation of nucleolar proteins around chromosomes at mitosis. A study by confocal laser scanning microscopy. *Journal of Cell Science,* 102**,** 729-737.

Gebrane-Younes, J., Fomproix, N. & Hernandez-Verdun, D. 1997. When rDNA transcription is arrested during mitosis, UBF is still associated with non-condensed rDNA. *Journal of Cell Science,* 110**,** 2429-2440.

Gerber, J. K., Richter, T., Kremmer, E., Adamski, J., Hofler, H., Balling, R. & Peters, H. 2002. Progressive loss of PAX9 expression correlates with increasing malignancy of dysplastic and cancerous epithelium of the human oesophagus. *J Pathol,* 197**,** 293-7.

Gerstberger, S., Hafner, M. & Tuschl, T. 2014. A census of human RNA-binding proteins. *Nat Rev Genet,* 15**,** 829-45.

Ghoshal, K., Majumder, S., Datta, J., Motiwala, T., Bai, S., Sharma, S. M., Frankel, W. & Jacob, S. T. 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**,** 6783-93. PMC2242730

Gibbons, J. G., Branco, A. T., Yu, S. & Lemos, B. 2014. Ribosomal DNA copy number is coupled with gene expression variation and mitochondrial abundance in humans. *Nature Communications,* 5**,** 4850.

Glader, B. E., Backer, K. & Diamond, L. K. 1983. Elevated erythrocyte adenosine deaminase activity in congenital hypoplastic anemia. *N Engl J Med,* 309**,** 1486-90.

Glaser, T., Jepeal, L., Edwards, J. G., Young, S. R., Favor, J. & Maas, R. L. 1994. PAX6 gene dosage effect in a family with congenital cataracts, aniridia, anophthalmia and central nervous system defects. *Nat Genet,* 7**,** 463-71.

Golomb, L., Volarevic, S. & Oren, M. 2014. p53 and ribosome biogenesis stress: the essentials. *FEBS Lett,* 588**,** 2571-9.

Gonzales, B., Henning, D., So, R. B., Dixon, J., Dixon, M. J. & Valdez, B. C. 2005. The Treacher Collins syndrome (TCOF1) gene product is involved in pre-rRNA methylation. *Human Molecular Genetics,* 14**,** 2035-2043.

Gonzalez, I. L. & Sylvester, J. E. 1995. Complete Sequence of the 43-kb Human Ribosomal DNA Repeat: Analysis of the Intergenic Spacer. *Genomics,* 27**,** 320-328.

Goodpasture, C. & Bloom, S. E. 1975. Visualization of nucleolar organizer regions im mammalian chromosomes using silver staining. *Chromosoma,* 53**,** 37-50.

Gordon, C. A., Gong, X., Ganesh, D. & Brooks, J. D. 2017. NUSAP1 promotes invasion and metastasis of prostate cancer. *Oncotarget,* 8**,** 29935-29950.

Griffin, J. N., Sondalle, S. B., Del Viso, F., Baserga, S. J. & Khokha, M. K. 2015. The Ribosome Biogenesis Factor Nol11 Is Required for Optimal rDNA Transcription and Craniofacial Development in *Xenopus*. *PLoS Genet,* 11**,** e1005018.

Grob, A., Colleran, C. & Mcstay, B. 2014. Construction of synthetic nucleoli in human cells reveals how a major functional nuclear domain is formed and propagated through cell division. *Genes Dev,* 28**,** 220-30. Pmc3923965

Grob, A. & Mcstay, B. 2014. Construction of synthetic nucleoli and what it tells us about propagation of sub-nuclear domains through cell division. *Cell Cycle,* 13**,** 2501-2508.

Gu, B.-W., Bessler, M. & Mason, P. J. 2008. A pathogenic dyskerin mutation impairs proliferation and activates a DNA damage response independent of telomere length in mice. *Proceedings of the National Academy of Sciences,* 105**,** 10173-10178.

Guetg, C. & Santoro, R. 2012. Formation of nuclear heterochromatin: the nucleolar point of view. *Epigenetics,* 7**,** 811-4. PMC3427276

Guo, L., Zhong, D., Lau, S., Liu, X., Dong, X.-Y., Sun, X., Yang, V. W., Vertino, P. M., Moreno, C., Varma, V., Jin-Tang, D. & Wei, Z. 2008. Sox7 is an independent checkpoint for β-catenin function in prostate and colon epithelial cells. *Molecular cancer research : MCR,* 6**,** 1421-1430.

Haaf, T., Hayman, D. L. & Schmid, M. 1991. Quantitative determination of rDNA transcription units in vertebrate cells. *Experimental Cell Research,* 193**,** 78-86.

Halkidou, K., Logan, I. R., Cook, S., Neal, D. E. & Robson, C. N. 2004. Putative involvement of the histone acetyltransferase Tip60 in ribosomal gene transcription. *Nucleic Acids Res,* 32**,** 1654-65. PMC390321

Halperin, D. S. & Freedman, M. H. 1989. Diamond-blackfan anemia: etiology, pathophysiology, and treatment. *Am J Pediatr Hematol Oncol,* 11**,** 380-94.

Hamdane, N., Stefanovsky, V. Y., Tremblay, M. G., Nemeth, A., Paquet, E., Lessard, F., Sanij, E., Hannan, R. & Moss, T. 2014. Conditional inactivation of upstream binding factor reveals its epigenetic functions and the existence of a somatic nucleolar precursor body. *PLoS Genet,* 10**,** e1004505. PMC4133168

Hamdane, N., Tremblay, M. G., Dillinger, S., Stefanovsky, V. Y., Nemeth, A. & Moss, T. 2016. Disruption of the UBF gene induces aberrant somatic nucleolar bodies and disrupts embryo nucleolar precursor bodies. *Gene,* 612**,** 5-11.

Handwerger, K. E., Cordero, J. A. & Gall, J. G. 2005. Cajal bodies, nucleoli, and speckles in the Xenopus oocyte nucleus have a low-density, sponge-like structure. *Mol Biol Cell,* 16**,** 202-11. PMC539164

Hayano, T., Yanagida, M., Yamauchi, Y., Shinkawa, T., Isobe, T. & Takahashi, N. 2003. Proteomic analysis of human Nop56p-associated pre-ribosomal ribonucleoprotein complexes. Possible link between Nop56p and the nucleolar protein treacle responsible for Treacher Collins syndrome. *J Biol Chem,* 278**,** 34309-19.

Heiss, N. S., Knight, S. W., Vulliamy, T. J., Klauck, S. M., Wiemann, S., Mason, P. J., Poustka, A. & Dokal, I. 1998. X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nat Genet,* 19**,** 32-8.

Henras, A. K., Plisson-Chastang, C., O'donohue, M.-F., Chakraborty, A. & Gleizes, P.-E. 2015. An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdisciplinary Reviews. RNA,* 6**,** 225-242.

Heo, I., Joo, C., Cho, J., Ha, M., Han, J. & Kim, V. N. 2008. Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. *Mol Cell,* 32**,** 276-84.

Hoyme, H. E., Kalberg, W. O., Elliott, A. J., Blankenship, J., Buckley, D., Marais, A.-S., Manning, M. A., Robinson, L. K., Adam, M. P., Abdul-Rahman, O., Jewett, T., Coles, C. D., Chambers, C., Jones, K. L., Adnams, C. M., Shah, P. E., Riley, E. P., Charness, M. E., Warren, K. R. & May, P. A. 2016. Updated Clinical Guidelines for Diagnosing Fetal Alcohol Spectrum Disorders. *Pediatrics,* 138**,** e20154256.

James, A., Wang, Y., Raje, H., Rosby, R. & Dimario, P. 2014. Nucleolar stress with and without p53. *Nucleus,* 5**,** 402-26. PMC4164484

Jarboui, M. A., Wynne, K., Elia, G., Hall, W. W. & Gautier, V. W. 2011. Proteomic profiling of the human T-cell nucleolus. *Mol Immunol,* 49**,** 441-52.

Jayaraman, S., Chittiboyina, S., Bai, Y., Abad, P. C., Vidi, P.-A., Stauffacher, C. V. & Lelièvre, S. A. 2017. The nuclear mitotic apparatus protein NuMA controls rDNA transcription and mediates the nucleolar stress response in a p53-independent manner. *Nucleic Acids Research***,** gkx782-gkx782.

Jernvall, J., Aberg, T., Kettunen, P., Keranen, S. & Thesleff, I. 1998. The life history of an embryonic signaling center: BMP-4 induces p21 and is associated with apoptosis in the mouse tooth enamel knot. *Development,* 125**,** 161-169.

Jernvall, J., Kettunen, P., Karavanova, I., Martin, L. B. & Thesleff, I. 1994. Evidence for the role of the enamel knot as a control center in mammalian tooth cusp formation: non-dividing cells express growth stimulating Fgf-4 gene. *Int J Dev Biol,* 38**,** 463-9.

Jia, S., Zhou, J., Fanelli, C., Wee, Y., Bonds, J., Schneider, P., Mues, G. & Souza, R. N. 2017a. Small-molecule Wnt agonists correct cleft palates in Pax9 mutant mice in utero. *Development,* 144**,** 3819-3828. PMC5675451

Jia, S., Zhou, J., Wee, Y., Mikkola, M. L., Schneider, P. & D'souza, R. N. 2017b. Anti-EDAR Agonist Antibody Therapy Resolves Palate Defects in Pax9-/- Mice. *Journal of Dental Research,* 96**,** 1282-1289.

Jimenez-Garcia, L. F., Rothblum, L. I., Busch, H. & Ochs, R. L. 1989. Nucleologenesis: use of non-isotopic in situ hybridization and immunocytochemistry to compare the localization of rDNA and nucleolar proteins during mitosis. *Biol Cell,* 65**,** 239-46.

Jimenez-Garcia, L. F., Segura-Valdez, M. D. L., Ochs, R. L., Rothblum, L. I., Hannan, R. & Spector, D. L. 1994. Nucleologenesis: U3 snRNA-containing prenucleolar bodies move to sites of active pre-rRNA transcription after mitosis. *Molecular Biology of the Cell,* 5**,** 955-966.

Jones, N. C., Lynn, M. L., Gaudenz, K., Sakai, D., Aoto, K., Rey, J. P., Glynn, E. F., Ellington, L., Du, C., Dixon, J., Dixon, M. J. & Trainor, P. A. 2008a. Prevention of the neurocristopathy Treacher Collins syndrome through inhibition of p53 function. *Nat Med,* 14**,** 125-33. PMC3093709

Jones, T. R., Kang, I. H., Wheeler, D. B., Lindquist, R. A., Papallo, A., Sabatini, D. M., Golland, P. & Carpenter, A. E. 2008b. CellProfiler Analyst: data exploration and analysis software for complex image-based screens. *BMC Bioinformatics,* 9**,** 482. PMC2614436

Kadakia, S., Helman, S. N., Badhey, A. K., Saman, M. & Ducic, Y. 2014. Treacher Collins Syndrome: The genetics of a craniofacial disease. *International Journal of Pediatric Otorhinolaryngology,* 78**,** 893-898.

Karpen, G. H., Schaefer, J. E. & Laird, C. D. 1988. A Drosophila rRNA gene located in euchromatin is active in transcription and nucleolus formation. *Genes Dev,* 2**,** 1745-63.

Kendall, J., Liu, Q., Bakleh, A., Krasnitz, A., Nguyen, K. C., Lakshmi, B., Gerald, W. L., Powers, S. & Mu, D. 2007. Oncogenic cooperation and coamplification of developmental transcription factor genes in lung cancer. *Proc Natl Acad Sci U S A,* 104**,** 16663-8. PMC2034240

Khokha, M. K., Chung, C., Bustamante, E. L., Gaw, L. W., Trott, K. A., Yeh, J., Lim, N., Lin, J. C., Taverner, N., Amaya, E., Papalopulu, N., Smith, J. C., Zorn, A. M., Harland, R. M. & Grammer, T. C. 2002. Techniques and probes for the study of Xenopus tropicalis development. *Dev Dyn,* 225**,** 499-510.

Kim, S.-K., Lee, H., Han, K., Kim, S. C., Choi, Y., Park, S.-W., Bak, G., Lee, Y., Choi, J. K., Kim, T.-K., Han, Y.-M. & Lee, D. 2014. SET7/9 methylation of the pluripotency factor LIN28A is a nucleolar localization mechanism that blocks let-7 biogenesis in human ESCs. *Cell stem cell,* 15**,** 735-749.

Koberna, K., Malinsky, J., Pliss, A., Masata, M., Vecerova, J., Fialova, M., Bednar, J. & Raska, I. 2002. Ribosomal genes in focus: new transcripts label the dense fibrillar components and form clusters indicative of "Christmas trees" in situ. *J Cell Biol,* 157**,** 743-8. PMC2173423

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**,** 748-64.

Kondrashov, N., Pusic, A., Stumpf, C. R., Shimizu, K., Hsieh, A. C., Ishijima, J., Shiroishi, T. & Barna, M. 2011. Ribosome mediated specificity in Hox mRNA translation and vertebrate tissue patterning. *Cell,* 145**,** 383-397.

Kramer, A., Green, J., Pollard, J., Jr. & Tugendreich, S. 2014. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics,* 30**,** 523-30. PMC3928520

Krystosek, A. 1998. Repositioning of human interphase chromosomes by nucleolar dynamics in the reverse transformation of HT1080 fibrosarcoma cells. *Exp Cell Res,* 241**,** 202-9.

Lafontaine, D. L., Bousquet-Antonelli, C., Henry, Y., Caizergues-Ferrer, M. & Tollervey, D. 1998. The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. *Genes Dev,* 12**,** 527-37. PMC316522

Landvogt, L., Ruland, J. A., Montellese, C., Siebrasse, J. P., Kutay, U. & Kubitscheck, U. 2019. Observing and tracking single small ribosomal subunits in vivo. *Methods,* 153**,** 63-70.

Langmead, B. & Salzberg, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Meth,* 9**,** 357-359.

Larburu, N., Montellese, C., O'donohue, M.-F., Kutay, U., Gleizes, P.-E. & Plisson-Chastang, C. 2016. Structure of a human pre-40S particle points to a role for RACK1 in the final steps of 18S rRNA processing. *Nucleic acids research,* 44**,** 8465-8478.

Lee, J. C., Sharma, M., Lee, Y. H., Lee, N. H., Kim, S. Y., Yun, J. S., Nam, S. Y., Hwang, P. H., Jhee, E. C. & Yi, H. K. 2008. Pax9 mediated cell survival in oral squamous carcinoma cell enhanced by c-myb. *Cell Biochem Funct,* 26**,** 892-9.

Lee, J. K., Park, J. W., Kim, Y. H. & Baek, S. H. 2012. Association between PAX9 single-nucleotide polymorphisms and nonsyndromic cleft lip with or without cleft palate. *J Craniofac Surg,* 23**,** 1262-6.

Leporé, N. & Lafontaine, D. L. J. 2011. A Functional Interface at the rDNA Connects rRNA Synthesis, Pre-rRNA Processing and Nucleolar Surveillance in Budding Yeast. *PLOS ONE,* 6**,** e24962.

Li, C., Lan, Y., Krumlauf, R. & Jiang, R. 2017. Modulating Wnt Signaling Rescues Palate Morphogenesis in Pax9 Mutant Mice. *Journal of Dental Research,* 96**,** 1273-1281.

Lipton, J. M., Atsidaftos, E., Zyskind, I. & Vlachos, A. 2006. Improving clinical care and elucidating the pathophysiology of Diamond Blackfan anemia: An update from the Diamond Blackfan Anemia Registry. *Pediatric Blood & Cancer,* 46**,** 558-564.

Lipton, J. M. & Ellis, S. R. 2009. Diamond-Blackfan anemia: diagnosis, treatment, and molecular pathogenesis. *Hematol Oncol Clin North Am,* 23**,** 261-82. PMC2886591

Luo, C., Yao, Y., Yu, Z., Zhou, H., Guo, L., Zhang, J., Cao, H., Zhang, G., Li, Y. & Jiao, Z. 2017. UBE2T knockdown inhibits gastric cancer progression. *Oncotarget,* 8**,** 32639-32654.

Mais, C., Wright, J. E., Prieto, J. L., Raggett, S. L. & Mcstay, B. 2005. UBF-binding site arrays form pseudo-NORs and sequester the RNA polymerase I transcription machinery. *Genes Dev,* 19**,** 50-64. Pmc540225

Marano, R. J. & Ben-Jonathan, N. 2014. Minireview: Extrapituitary Prolactin: An Update on the Distribution, Regulation, and Functions. *Molecular Endocrinology,* 28**,** 622-633.

Mccann, K. L., Teramoto, T., Zhang, J., Tanaka Hall, T. M. & Baserga, S. J. 2016. The molecular basis for ANE syndrome revealed by the large ribosomal subunit processome interactome. *Elife,* 5. PMC4859800

Mcclintock, B. 1934. The relation of a particular chromosomal element to the development of the nucleoli in Zea mays. *Zeitschrift für Zellforschung und Mikroskopische Anatomie,* 21**,** 294-326.

Mcgowan, K. A., Li, J. Z., Park, C. Y., Beaudry, V., Tabor, H. K., Sabnis, A. J., Zhang, W., Fuchs, H., De Angelis, M. H., Myers, R. M., Attardi, L. D. & Barsh, G. S. 2008. Ribosomal mutations cause p53-mediated dark skin and pleiotropic effects. *Nat Genet,* 40**,** 963-70.

Mcleay, R. C. & Bailey, T. L. 2010. Motif Enrichment Analysis: a unified framework and an evaluation on ChIP data. *BMC Bioinformatics,* 11**,** 165. PMC2868005

Mcstay, B. & Grummt, I. 2008. The epigenetics of rRNA genes: from molecular to chromosome biology. *Annu Rev Cell Dev Biol,* 24**,** 131-57.

Mensah, J. K., Ogawa, T., Kapadia, H., Cavender, A. C. & D'souza, R. N. 2004. Functional Analysis of a Mutation in PAX9 Associated with Familial Tooth Agenesis in Humans. *Journal of Biological Chemistry,* 279**,** 5924-5933.

Mi, H., Huang, X., Muruganujan, A., Tang, H., Mills, C., Kang, D. & Thomas, P. D. 2017. PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res,* 45**,** D183-D189. PMC5210595

Miller, G., Panov, K. I., Friedrich, J., Trinkle-Mulcahy, L., Lamond, A. I. & Zomerdijk, J. C. B. M. 2001. hRRN3 is essential in the SL1-mediated recruitment of RNA Polymerase I to rRNA gene promoters. *The EMBO Journal,* 20**,** 1373-1382.

Mochida, S., Ikeo, S., Gannon, J. & Hunt, T. 2009. Regulated activity of PP2A-B55 delta is crucial for controlling entry into and exit from mitosis in Xenopus egg extracts. *Embo j,* 28**,** 2777-85. PMC2750019

Mohr, S. E., Smith, J. A., Shamu, C. E., Neumuller, R. A. & Perrimon, N. 2014. RNAi screening comes of age: improved techniques and complementary approaches. *Nat Rev Mol Cell Biol,* 15**,** 591-600.

Monsoro-Burq, A. H. 2015. PAX transcription factors in neural crest development. *Semin Cell Dev Biol,* 44**,** 87-96.

Montanaro, L., Trere, D. & Derenzini, M. 2008. Nucleolus, ribosomes, and cancer. *Am J Pathol,* 173**,** 301-10.

Mostowska, A., Zadurska, M., Rakowska, A., Lianeri, M. & Jagodzinski, P. P. 2013. Novel PAX9 mutation associated with syndromic tooth agenesis. *Eur J Oral Sci,* 121**,** 403-11.

Muratovska, A., Zhou, C., He, S., Goodyer, P. & Eccles, M. R. 2003. Paired-Box genes are frequently expressed in cancer and often required for cancer cell survival. *Oncogene,* 22**,** 7989-97.

Narasimhan, K., Hilbig, A., Udayasuryan, B., Jayabal, S., Kolatkar, P. R. & Jauch, R. 2014. Crystallization and preliminary X-ray diffraction analysis of the Pax9 paired domain bound to a DC5 enhancer DNA element. *Acta crystallographica. Section F, Structural biology communications,* 70**,** 1357-1361.

Németh, A., Conesa, A., Santoyo-Lopez, J., Medina, I., Montaner, D., Péterfia, B., Solovei, I., Cremer, T., Dopazo, J. & Längst, G. 2010. Initial genomics of the human nucleolus. *PLoS genetics,* 6**,** e1000889-e1000889.

Neubüser, A., Koseki, H. & Balling, R. 1995. Characterization and Developmental Expression of Pax9, a Paired-Box-Containing Gene Related to Pax1. *Developmental Biology,* 170**,** 701-716.

Neubuser, A., Peters, H., Balling, R. & Martin, G. R. 1997. Antagonistic interactions between FGF and BMP signaling pathways: a mechanism for positioning the sites of tooth formation. *Cell,* 90**,** 247-55.

Neumuller, R. A., Gross, T., Samsonova, A. A., Vinayagam, A., Buckner, M., Founk, K., Hu, Y., Sharifpoor, S., Rosebrock, A. P., Andrews, B., Winston, F. & Perrimon, N. 2013. Conserved regulators of nucleolar size revealed by global phenotypic analyses. *Sci Signal,* 6**,** ra70. Pmc3964804

Nicolas, E., Parisot, P., Pinto-Monteiro, C., De Walque, R., De Vleeschouwer, C. & Lafontaine, D. L. 2016. Involvement of human ribosomal proteins in nucleolar structure and p53-dependent nucleolar stress. *Nat Commun,* 7**,** 11390. PMC4897761

Nieminen, P., Arte, S., Tanner, D., Paulin, L., Alaluusua, S., Thesleff, I. & Pirinen, S. 2001. Identification of a nonsense mutation in the PAX9 gene in molar oligodontia. *Eur J Hum Genet,* 9**,** 743-6.

Noack Watt, K. E., Achilleos, A., Neben, C. L., Merrill, A. E. & Trainor, P. A. 2016. The Roles of RNA Polymerase I and III Subunits Polr1c and Polr1d in Craniofacial Development and in Zebrafish Models of Treacher Collins Syndrome. *PLoS Genet,* 12**,** e1006187.

Nousbeck, J., Spiegel, R., Ishida-Yamamoto, A., Indelman, M., Shani-Adir, A., Adir, N., Lipkin, E., Bercovici, S., Geiger, D., Van Steensel, M. A., Steijlen, P. M., Bergman, R., Bindereif, A., Choder, M., Shalev, S. & Sprecher, E. 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**,** 1114-21. PMC2427309

Novetsky, A. P., Zighelboim, I., Thompson, D. M., Powell, M. A., Mutch, D. G. & Goodfellow, P. J. 2013. Frequent Mutations in the RPL22 Gene and its Clinical and Functional Implications. *Gynecologic oncology,* 128**,** 10.1016/j.ygyno.2012.10.026.

O'donohue, M. F., Choesmel, V., Faubladier, M., Fichant, G. & Gleizes, P. E. 2010. Functional dichotomy of ribosomal proteins during the synthesis of mammalian 40S ribosomal subunits. *J Cell Biol,* 190**,** 853-66. PMC2935573

Oakes, M. L., Johzuka, K., Vu, L., Eliason, K. & Nomura, M. 2006. Expression of rRNA Genes and Nucleolus Formation at Ectopic Chromosomal Sites in the Yeast Saccharomyces cerevisiae. *Molecular and Cellular Biology,* 26**,** 6223-6238.

Olson, M. O. 2004. Sensing cellular stress: another new function for the nucleolus? *Sci STKE,* 2004**,** pe10.

Peng, S., Chen, L. L., Lei, X. X., Yang, L., Lin, H., Carmichael, G. G. & Huang, Y. 2011. Genome-wide studies reveal that Lin28 enhances the translation of genes important for growth and survival of human embryonic stem cells. *Stem Cells,* 29**,** 496-504.

Peng, W. T., Krogan, N. J., Richards, D. P., Greenblatt, J. F. & Hughes, T. R. 2004. ESF1 is required for 18S rRNA synthesis in Saccharomyces cerevisiae. *Nucleic Acids Res,* 32**,** 1993-9. PMC390369

Penzo, M., Montanaro, L., Treré, D. & Derenzini, M. 2019. The Ribosome Biogenesis-Cancer Connection. *Cells,* 8**,** e55. PMC6356843

Pereboom, T. C., Van Weele, L. J., Bondt, A. & Macinnes, A. W. 2011. A zebrafish model of dyskeratosis congenita reveals hematopoietic stem cell formation failure resulting from ribosomal protein-mediated p53 stabilization. *Blood,* 118**,** 5458-65.

Pestov, D. G., Lapik, Y. R. & Lau, L. F. 2008. Assays for ribosomal RNA processing and ribosome assembly. *Curr Protoc Cell Biol,* Chapter 22**,** Unit 22 11.

Pestov, D. G., Strezoska, Z. & Lau, L. F. 2001. Evidence of p53-dependent cross-talk between ribosome biogenesis and the cell cycle: effects of nucleolar protein Bop1 on G(1)/S transition. *Mol Cell Biol,* 21**,** 4246-55.

Peters, H. & Balling, R. 1999. Teeth: where and how to make them. *Trends in Genetics,* 15**,** 59-65.

Peters, H., Neubuser, A., Kratochwil, K. & Balling, R. 1998. Pax9-deficient mice lack pharyngeal pouch derivatives and teeth and exhibit craniofacial and limb abnormalities. *Genes Dev,* 12**,** 2735-47. PMC317134

Petrelli, B., Bendelac, L., Hicks, G. G. & Fainsod, A. 2019. Insights into retinoic acid deficiency and the induction of craniofacial malformations and microcephaly in fetal alcohol spectrum disorder. *Genesis,* 57**,** e23278.

Petryszak, R., Keays, M., Tang, Y. A., Fonseca, N. A., Barrera, E., Burdett, T., Füllgrabe, A., Fuentes, A. M.-P., Jupp, S., Koskinen, S., Mannion, O., Huerta, L., Megy, K., Snow, C., Williams, E., Barzine, M., Hastings, E., Weisser, H., Wright, J., Jaiswal, P., Huber, W., Choudhary, J., Parkinson, H. E. & Brazma, A. 2015. Expression Atlas update—an integrated database of gene and protein expression in humans, animals and plants. *Nucleic Acids Research,* 44**,** D746-D752.

Pfister, A. S. & Kühl, M. 2018. Chapter Four - Of Wnts and Ribosomes. *In:* LARRAÍN, J. & OLIVARES, G. (eds.) *Progress in Molecular Biology and Translational Science.* Academic Press.

Phelps, P. D., Poswillo, D. & Lloyd, G. A. 1981. The ear deformities in mandibulofacial dysostosis (Treacher Collins syndrome). *Clin Otolaryngol Allied Sci,* 6**,** 15-28.

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

Pich, A., Chiusa, L. & Margaria, E. 2000. Prognostic relevance of AgNORs in tumor pathology. *Micron,* 31**,** 133-41.

Piskounova, E., Polytarchou, C., Thornton, J. E., Hagan, J. P., Lapierre, R. J., Pothoulakis, C., Iliopoulos, D. & Gregory, R. I. 2011a. Oncogenic Lin28A and Lin28B inhibit let-7 microRNA biogenesis by distinct mechanisms. *Cell,* 147**,** 1066-1079.

Piskounova, E., Polytarchou, C., Thornton, J. E., Lapierre, R. J., Pothoulakis, C., Hagan, J. P., Iliopoulos, D. & Gregory, R. I. 2011b. Lin28A and Lin28B inhibit let-7 microRNA biogenesis by distinct mechanisms. *Cell,* 147**,** 1066-79. PMC3227872

Ploton, D., Menager, M., Jeannesson, P., Himber, G., Pigeon, F. & Adnet, J. J. 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.

Prieto, J. L. & 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**,** 2041-54.

Provost, E., Wehner, K. A., Zhong, X., Ashar, F., Nguyen, E., Green, R., Parsons, M. J. & Leach, S. D. 2012. Ribosomal biogenesis genes play an essential and p53-independent role in zebrafish pancreas development. *Development,* 139**,** 3232-41. PMC3413166

Qu, Y., Han, B., Yu, Y., Yao, W., Bose, S., Karlan, B. Y., Giuliano, A. E. & Cui, X. 2015. Evaluation of MCF10A as a Reliable Model for Normal Human Mammary Epithelial Cells. *PLOS ONE,* 10**,** e0131285.

Raska, I., Ochs, R. L., Andrade, L. E., Chan, E. K., Burlingame, R., Peebles, C., Gruol, D. & Tan, E. M. 1990. Association between the nucleolus and the coiled body. *J Struct Biol,* 104**,** 120-7.

Raska, I., Shaw, P. J. & Cmarko, D. 2006. Structure and function of the nucleolus in the spotlight. *Curr Opin Cell Biol,* 18**,** 325-34.

Ray, D., Kazan, H., Cook, K. B., Weirauch, M. T., Najafabadi, H. S., Li, X., Gueroussov, S., Albu, M., Zheng, H., Yang, A., Na, H., Irimia, M., Matzat, L. H., Dale, R. K., Smith, S. A., Yarosh, C. A., Kelly, S. M., Nabet, B., Mecenas, D., Li, W., Laishram, R. S., Qiao, M., Lipshitz, H. D., Piano, F., Corbett, A. H., Carstens, R. P., Frey, B. J., Anderson, R. A., Lynch, K. W., Penalva, L. O. F., Lei, E. P., Fraser, A. G., Blencowe, B. J., Morris, Q. D. & Hughes, T. R. 2013. A compendium of RNA-binding motifs for decoding gene regulation. *Nature,* 499**,** 172-7. PMC3929597

Reimer, G., Pollard, K. M., Penning, C. A., Ochs, R. L., Lischwe, M. A., Busch, H. & Tan, E. M. 1987a. Monoclonal antibody 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 and Rheumatism,* 30**,** 793-800.

Reimer, G., Pollard, K. M., Penning, C. A., Ochs, R. L., Lischwe, M. A., Busch, H. & Tan, E. M. 1987b. 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.

Relaix, F., Rocancourt, D., Mansouri, A. & Buckingham, M. 2005. A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature,* 435**,** 948-53.

Rice, S. J., Lai, S.-C., Wood, L. W., Helsley, K. R., Runkle, E. A., Winslow, M. M. & Mu, D. 2013. MicroRNA-33a Mediates the Regulation of High Mobility Group AT-Hook 2 Gene (HMGA2) by Thyroid Transcription Factor 1 (TTF-1/NKX2–1). *Journal of Biological Chemistry,* 288**,** 16348-16360.

Robson, E. J. D., He, S.-J. & Eccles, M. R. 2006. A PANorama of PAX genes in cancer and development. *Nature Reviews Cancer,* 6**,** 52-62.

Roesch, A., Mueller, A. M., Stempfl, T., Moehle, C., Landthaler, M. & Vogt, T. 2008. RBP2-H1/JARID1B is a transcriptional regulator with a tumor suppressive potential in melanoma cells. *Int J Cancer,* 122**,** 1047-57.

Ross, A. P., Mansilla, M. A., Choe, Y., Helminski, S., Sturm, R., Maute, R. L., May, S. R., Hozyasz, K. K., Wójcicki, P., Mostowska, A., Davidson, B., Adamopoulos, I. E., Pleasure, S. J., Murray, J. C. & Zarbalis, K. S. 2013. A Mutation in Mouse Pak1ip1 Causes Orofacial Clefting while Human PAK1IP1 Maps to 6p24 Translocation Breaking Points Associated with Orofacial Clefting. *PLoS ONE,* 8**,** e69333.

Roussel, P., Andre, C., Comai, L. & Hernandez-Verdun, D. 1996. The rDNA transcription machinery is assembled during mitosis in active NORs and absent in inactive NORs. *J Cell Biol,* 133**,** 235-46. PMC2120807

Rovin, S., Dachi, S. F., Borenstein, D. B. & Cotter, W. B. 1964. MANDIBULOFACIAL DYSOSTOSIS, A FAMILIAL STUDY OF FIVE GENERATIONS. *J Pediatr,* 65**,** 215-21.

Rubbi, C. P. & Milner, J. 2003. Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses. *EMBO J,* 22**,** 6068-77. PMC275437

Sanchez, R. S. & Sanchez, S. S. 2013. Characterization of pax1, pax9, and uncx sclerotomal genes during Xenopus laevis embryogenesis. *Dev Dyn,* 242**,** 572-9.

Sarkar, T., Bansal, R. & Das, P. 2017. A novel G to A transition at initiation codon and exon-intron boundary of PAX9 identified in association with familial isolated oligodontia. *Gene,* 635**,** 69-76.

Savino, T. M., Gébrane-Younès, J., De Mey, J., Sibarita, J.-B. & Hernandez-Verdun, D. 2001. Nucleolar Assembly of the Rrna Processing Machinery in Living Cells. *The Journal of Cell Biology,* 153**,** 1097-1110.

Scheer, U. & Weisenberger, D. 1994. The nucleolus. *Curr Opin Cell Biol,* 6**,** 354-9.

Schlesinger, S., Selig, S., Bergman, Y. & Cedar, H. 2009. Allelic inactivation of rDNA loci. *Genes Dev,* 23**,** 2437-47. PMC2764490

Schmidt, E. K., Clavarino, G., Ceppi, M. & Pierre, P. 2009. SUnSET, a nonradioactive method to monitor protein synthesis. *Nat Meth,* 6**,** 275-277.

Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat Meth,* 9**,** 671-675.

Schneider, D. A., French, S. L., Osheim, Y. N., Bailey, A. O., Vu, L., Dodd, J., Yates, J. R., Beyer, A. L. & Nomura, M. 2006. RNA polymerase II elongation factors Spt4p and Spt5p play roles in transcription elongation by RNA polymerase I and rRNA processing. *Proceedings of the National Academy of Sciences,* 103**,** 12707-12712.

Schöfer, C. & Weipoltshammer, K. 2018. Nucleolus and chromatin. *Histochemistry and cell biology,* 150**,** 209-225.

Schuffenhauer, S., Leifheit, H., Lichtner, P., Peters, H., Murken, J. & Emmerich, P. 1999. De novo deletion (14)(q11.2q13) including PAX9: clinical and molecular findings. *Journal of Medical Genetics,* 36**,** 233-236.

Schuller, A. P. & Green, R. 2018. Roadblocks and resolutions in eukaryotic translation. *Nature reviews. Molecular cell biology,* 19**,** 526-541.

Šerý, O., Bonczek, O., Hloušková, A., Černochová, P., Vaněk, J., Míšek, I., Krejčí, P. & Izakovičová Hollá, L. 2015. A screen of a large Czech cohort of oligodontia patients implicates a novel mutation in the PAX9 gene. *European Journal of Oral Sciences,* 123**,** 65-71.

Sethi, B. K., Chanukya, G. V. & Nagesh, V. S. 2012. Prolactin and cancer: Has the orphan finally found a home? *Indian Journal of Endocrinology and Metabolism,* 16**,** S195-S198.

Shea, J. R. & Leblond, C. P. 1966. Number of nucleoli in various cell types of the mouse. *Journal of Morphology,* 119**,** 425-433.

Shin, Y., Chang, Y.-C., Lee, D. S. W., Berry, J., Sanders, D. W., Ronceray, P., Wingreen, N. S., Haataja, M. & Brangwynne, C. P. 2018. Liquid Nuclear Condensates Mechanically Sense and Restructure the Genome. *Cell,* 175**,** 1481-1491.e13.

Short, S. & Holland, L. Z. 2008. The Evolution of Alternative Splicing in the Pax Family: The View from the Basal Chordate Amphioxus. *Journal of Molecular Evolution,* 66**,** 605-20.

Sivakamasundari, V., Kraus, P., Sun, W., Hu, X., Lim, S. L., Prabhakar, S. & Lufkin, T. 2017. A developmental transcriptomic analysis of Pax1 and Pax9 in embryonic intervertebral disc development. *Biol Open,* 6**,** 187-199. PMC5312110

Sloan, Katherine e., Bohnsack, Markus t. & Watkins, Nicholas j. 2013a. The 5S RNP Couples p53 Homeostasis to Ribosome Biogenesis and Nucleolar Stress. *Cell Reports,* 5**,** 237-247.

Sloan, K. E., Mattijssen, S., Lebaron, S., Tollervey, D., Pruijn, G. J. & Watkins, N. J. 2013b. Both endonucleolytic and exonucleolytic cleavage mediate ITS1 removal during human ribosomal RNA processing. *J Cell Biol,* 200**,** 577-88. PMC3587827

Sloan, K. E., Warda, A. S., Sharma, S., Entian, K.-D., Lafontaine, D. L. J. & Bohnsack, M. T. 2017. Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA Biology,* 14**,** 1138-1152.

Sondalle, S. B., Baserga, S. J. & Yelick, P. C. 2016. The Contributions of the Ribosome Biogenesis Protein Utp5/WDR43 to Craniofacial Development. *J Dent Res,* 95**,** 1214-20. PMC5076753

Sondalle, S. B., Longerich, S., Ogawa, L. M., Sung, P. & Baserga, S. J. 2019. Fanconi anemia protein FANCI functions in ribosome biogenesis. *Proceedings of the National Academy of Sciences,* 116**,** 2561-2570.

Soule, H. D., Maloney, T. M., Wolman, S. R., Peterson, W. D., Jr., Brenz, R., Mcgrath, C. M., Russo, J., Pauley, R. J., Jones, R. F. & Brooks, S. C. 1990. Isolation and characterization of a spontaneously immortalized human breast epithelial cell line, MCF-10. *Cancer Res,* 50**,** 6075-86.

Spandidos, A., Wang, X., Wang, H. & Seed, B. 2010. PrimerBank: a resource of human and mouse PCR primer pairs for gene expression detection and quantification. *Nucleic Acids Research,* 38**,** D792-D799.

Splendore, A., Silva, E. O., Alonso, L. G., Richieri-Costa, A., Alonso, N., Rosa, A., Carakushanky, G., Cavalcanti, D. P., Brunoni, D. & Passos-Bueno, M. R. 2000. High mutation detection rate in TCOF1 among Treacher Collins syndrome patients reveals clustering of mutations and 16 novel pathogenic changes. *Hum Mutat,* 16**,** 315-22.

Stapleton, P., Weith, A., Urbánek, P., Kozmik, Z. & Busslinger, M. 1993. Chromosomal localization of seven PAX genes and cloning of a novel family member, PAX-9. *Nature Genetics,* 3**,** 292-8.

Stockton, D. W., Das, P., Goldenberg, M., D'souza, R. N. & Patel, P. I. 2000. Mutation of PAX9 is associated with oligodontia. *Nat Genet,* 24**,** 18-19.

Sumner, A. T. 1982. The nature and mechanisms of chromosome banding. *Cancer Genetics and Cytogenetics,* 6**,** 59-87.

Sun, X., Chen, H., Deng, Z., Hu, B., Luo, H., Zeng, X., Han, L., Cai, G. & Ma, L. 2015. The Warsaw breakage syndrome-related protein DDX11 is required for ribosomal RNA synthesis and embryonic development. *Human Molecular Genetics,* 24**,** 4901-4915.

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K. P., Kuhn, M., Bork, P., Jensen, L. J. & Von Mering, C. 2015. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res,* 43**,** D447-52. PMC4383874

Tafforeau, L., Zorbas, C., Langhendries, J. L., Mullineux, S. T., Stamatopoulou, V., Mullier, R., Wacheul, L. & Lafontaine, D. L. 2013. The Complexity of Human Ribosome Biogenesis Revealed by Systematic Nucleolar Screening of Pre-rRNA Processing Factors. *Mol Cell,* 51**,** 539-51.

Takino, T., Nakada, M., Li, Z., Yoshimoto, T., Domoto, T. & Sato, H. 2016. Tip60 regulates MT1-MMP transcription and invasion of glioblastoma cells through NF-κB pathway. *Clinical & Experimental Metastasis,* 33**,** 45-52.

Tan, B., Wang, J., Song, Q., Wang, N., Jia, Y., Wang, C., Yao, B., Liu, Z., Zhang, X. & Cheng, Y. 2017. Prognostic value of PAX9 in patients with esophageal squamous cell carcinoma and its prediction value to radiation sensitivity. *Molecular medicine reports,* 16**,** 806-816.

Thul, P. J., Akesson, L., Wiking, M., Mahdessian, D., Geladaki, A., Ait Blal, H., Alm, T., Asplund, A., Bjork, L., Breckels, L. M., Backstrom, A., Danielsson, F., Fagerberg, L., Fall, J., Gatto, L., Gnann, C., Hober, S., Hjelmare, M., Johansson, F., Lee, S., Lindskog, C., Mulder, J., Mulvey, C. M., Nilsson, P., Oksvold, P., Rockberg, J., Schutten, R., Schwenk, J. M., Sivertsson, A., Sjostedt, E., Skogs, M., Stadler, C., Sullivan, D. P., Tegel, H., Winsnes, C., Zhang, C., Zwahlen, M., Mardinoglu, A., Ponten, F., Von Feilitzen, K., Lilley, K. S., Uhlen, M. & Lundberg, E. 2017. A subcellular map of the human proteome. *Science,* 356**,** eaal3321.

Tourlakis, M. E., Zhang, S., Ball, H. L., Gandhi, R., Liu, H., Zhong, J., Yuan, J. S., Guidos, C. J., Durie, P. R. & Rommens, J. M. 2015. In Vivo Senescence in the Sbds-Deficient Murine Pancreas: Cell-Type Specific Consequences of Translation Insufficiency. *PLoS Genet,* 11**,** e1005288. PMC4461263

Trainor, P. A., Dixon, J. & Dixon, M. J. 2008. Treacher Collins syndrome: etiology, pathogenesis and prevention. *European Journal Of Human Genetics,* 17**,** 275-83. PMC2986179

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, M. J., Salzberg, S. L., Wold, B. J. & Pachter, L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotech,* 28**,** 511-515.

Treré, D., Ceccarelli, C., Montanaro, L., Tosti, E. & Derenzini, M. 2004. Nucleolar Size and Activity Are Related to pRb and p53 Status in Human Breast Cancer. *Journal of Histochemistry & Cytochemistry,* 52**,** 1601-1607.

Turner, A. J., Knox, A. A. & Watkins, N. J. 2012. Nucleolar disruption leads to the spatial separation of key 18S rRNA processing factors. *RNA Biol,* 9**,** 175-86.

Valdez, B. C., Henning, D., So, R. B., Dixon, J. & Dixon, M. J. 2004. The Treacher Collins syndrome (TCOF1) gene product is involved in ribosomal DNA gene transcription by interacting with upstream binding factor. *Proc Natl Acad Sci U S A,* 101**,** 10709-14. PMC489999

Van Der Lelij, P., Chrzanowska, K. H., Godthelp, B. C., Rooimans, M. A., Oostra, A. B., Stumm, M., Zdzienicka, M. Z., Joenje, H. & De Winter, J. P. 2010. Warsaw Breakage Syndrome, a Cohesinopathy Associated with Mutations in the XPD Helicase Family Member DDX11/ChlR1. *The American Journal of Human Genetics,* 86**,** 262-266.

Van Koningsbruggen, S., Gierlinski, M., Schofield, P., Martin, D., Barton, G. J., Ariyurek, Y., Den Dunnen, J. T. & Lamond, A. I. 2010. High-resolution whole-genome sequencing reveals that specific chromatin domains from most human chromosomes associate with nucleoli. *Molecular biology of the cell,* 21**,** 3735-3748.

Vlachos, A., Rosenberg, P. S., Atsidaftos, E., Alter, B. P. & Lipton, J. M. 2012. *Incidence of neoplasia in Diamond Blackfan anemia: a report from the Diamond Blackfan Anemia Registry*.

Vlachos, A., Rosenberg, P. S., Atsidaftos, E., Kang, J., Onel, K., Sharaf, R. N., Alter, B. P. & Lipton, J. M. 2018. Increased risk of colon cancer and osteogenic sarcoma in Diamond-Blackfan anemia. *Blood,* 132**,** 2205-2208.

Vlatkovic, N., Boyd, M. T. & Rubbi, C. P. 2014. Nucleolar control of p53: a cellular Achilles' heel and a target for cancer therapy. *Cell Mol Life Sci,* 71**,** 771-91.

Vogt, E. J., Meglicki, M., Hartung, K. I., Borsuk, E. & Behr, R. 2012. Importance of the pluripotency factor LIN28 in the mammalian nucleolus during early embryonic development. *Development,* 139**,** 4514-23.

Wachtel, M. & Schäfer, B. W. 2015. Unpeaceful roles of mutant PAX proteins in cancer. *Seminars in Cell & Developmental Biology,* 44**,** 126-134.

Wallin, J., Eibel, H., Neubuser, A., Wilting, J., Koseki, H. & Balling, R. 1996. Pax1 is expressed during development of the thymus epithelium and is required for normal T-cell maturation. *Development,* 122**,** 23-30.

Wan, Y., Zhang, Q., Zhang, Z., Song, B., Wang, X., Zhang, Y., Jia, Q., Cheng, T., Zhu, X., Leung, A. Y., Yuan, W., Jia, H. & Fang, X. 2016. Transcriptome analysis reveals a ribosome constituents disorder involved in the RPL5 downregulated zebrafish model of Diamond-Blackfan anemia. *BMC Med Genomics,* 9**,** 13. PMC4785739

Wang, H., Xiao, W., Zhou, Q., Chen, Y., Yang, S., Sheng, J., Yin, Y., Fan, J. & Zhou, J. 2009a. Bystin-like protein is upregulated in hepatocellular carcinoma and required for nucleologenesis in cancer cell proliferation. *Cell Res,* 19**,** 1150-64.

Wang, K., Kan, J., Yuen, S. T., Shi, S. T., Chu, K. M., Law, S., Chan, T. L., Kan, Z., Chan, A. S., Tsui, W. Y., Lee, S. P., Ho, S. L., Chan, A. K., Cheng, G. H., Roberts, P. C., Rejto, P. A., Gibson, N. W., Pocalyko, D. J., Mao, M., Xu, J. & Leung, S. Y. 2011. Exome sequencing identifies frequent mutation of ARID1A in molecular subtypes of gastric cancer. *Nat Genet,* 43**,** 1219-23.

Wang, M., Anikin, L. & Pestov, D. G. 2014. Two orthogonal cleavages separate subunit RNAs in mouse ribosome biogenesis. *Nucleic Acids Research,* 42**,** 11180-11191.

Wang, M. & Lemos, B. 2017. Ribosomal DNA copy number amplification and loss in human cancers is linked to tumor genetic context, nucleolus activity, and proliferation. *PLoS Genet,* 13**,** e1006994. PMC5605086

Wang, Y., Groppe, J. C., Wu, J., Ogawa, T., Mues, G., D'souza, R. N. & Kapadia, H. 2009b. Pathogenic mechanisms of tooth agenesis linked to paired domain mutations in human PAX9. *Human Molecular Genetics,* 18**,** 2863-2874.

Warner, J. R. 1999. The economics of ribosome biosynthesis in yeast. *Trends in Biochemical Sciences,* 24**,** 437-440.

Warren, A. J. 2018. Molecular basis of the human ribosomopathy Shwachman-Diamond syndrome. *Adv Biol Regul,* 67**,** 109-127.

Watt, K. E. N., Neben, C. L., Hall, S., Merrill, A. E. & Trainor, P. A. 2018. tp53-dependent and independent signaling underlies the pathogenesis and possible prevention of Acrofacial Dysostosis–Cincinnati type. *Human Molecular Genetics,* 27**,** 2628-2643.

Weaver, K. N., Watt, K. E., Hufnagel, R. B., Navajas Acedo, J., Linscott, L. L., Sund, K. L., Bender, P. L., Konig, R., Lourenco, C. M., Hehr, U., Hopkin, R. J., Lohmann, D. R., Trainor, P. A., Wieczorek, D. & Saal, H. M. 2015. Acrofacial Dysostosis, Cincinnati Type, a Mandibulofacial Dysostosis Syndrome with Limb Anomalies, Is Caused by POLR1A Dysfunction. *Am J Hum Genet,* 96**,** 765-74. PMC4570288

Weber, A. M., Tuchweber, B., Yousef, I., Brochu, P., Turgeon, C., Gabbiani, G., Morin, C. L. & Roy, C. C. 1981. Severe familial cholestasis in North American Indian children: a clinical model of microfilament dysfunction? *Gastroenterology,* 81**,** 653-62.

Weiss, W. A., Taylor, S. S. & Shokat, K. M. 2007. Recognizing and exploiting differences between RNAi and small-molecule inhibitors. *Nat Chem Biol,* 3**,** 739-44. PMC2924165

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. PMC2964341

Wilkins, B. J., Lorent, K., Matthews, R. P. & Pack, M. 2013. p53-Mediated Biliary Defects Caused by Knockdown of cirh1a, the Zebrafish Homolog of the Gene Responsible for North American Indian Childhood Cirrhosis. *PLoS One,* 8**,** e77670.

Wiza, C., Chadt, A., Blumensatt, M., Kanzleiter, T., Herzfeld De Wiza, D., Horrighs, A., Mueller, H., Nascimento, E. B. M., Schürmann, A., Al-Hasani, H. & Ouwens, D. M. 2014. Over-expression of PRAS40 enhances insulin sensitivity in skeletal muscle. *Archives of Physiology and Biochemistry,* 120**,** 64-72.

Wong, S. W., Han, D., Zhang, H., Liu, Y., Zhang, X., Miao, M. Z., Wang, Y., Zhao, N., Zeng, L., Bai, B., Wang, Y. X., Liu, H., Frazier-Bowers, S. A. & Feng, H. 2018. Nine Novel PAX9 Mutations and a Distinct Tooth Agenesis Genotype-Phenotype. *J Dent Res,* 97**,** 155-162. PMC5784472

Woolford Jr., J. L. & Baserga, S. J. 2013. Ribosome Biogenesis in the Yeast *Saccaromyces cerevisiae*. *Genetics* [Online], 195.

Wu, J. Q., Guo, J. Y., Tang, W., Yang, C.-S., Freel, C. D., Chen, C., Nairn, A. C. & Kornbluth, S. 2009. PP1-mediated dephosphorylation of phosphoproteins at mitotic exit is controlled by inhibitor-1 and PP1 phosphorylation. *Nat Cell Biol,* 11**,** 644-651.

Wu, Z. & Gall, J. G. 1997. "Micronucleoli" in the Xenopus germinal vesicle. *Chromosoma,* 105**,** 438-43.

Xiong, Z., Ren, S., Chen, H., Liu, Y., Huang, C., Zhang, Y. L., Odera, J. O., Chen, T., Kist, R., Peters, H., Garman, K., Sun, Z. & Chen, X. 2018. PAX9 regulates squamous cell differentiation and carcinogenesis in the oro-oesophageal epithelium. *The Journal of pathology,* 244**,** 164-175.

Xu, B., Li, H., Perry, J. M., Singh, V. P., Unruh, J., Yu, Z., Zakari, M., Mcdowell, W., Li, L. & Gerton, J. L. 2017. Ribosomal DNA copy number loss and sequence variation in cancer. *PLoS Genet,* 13**,** e1006771. PMC5480814

Xu, H. E., Rould, M. A., Xu, W., Epstein, J. A., Maas, R. L. & Pabo, C. O. 1999. Crystal structure of the human Pax6 paired domain–DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding. *Genes & Development,* 13**,** 1263-1275.

Yang, K., Yang, J. & Yi, J. 2018. Nucleolar Stress: hallmarks, sensing mechanism and diseases. *Cell Stress,* 2**,** 125-140.

Yu, M., Wong, S.-W., Han, D. & Cai, T. 2018. Genetic analysis: Wnt and other pathways in nonsyndromic tooth agenesis. *Oral Diseases***,** 1-6. PMC6318069

Yu, S. & Lemos, B. 2018. The long-range interaction map of ribosomal DNA arrays. *PLoS genetics,* 14**,** e1007258-e1007258.

Yu, W., Qiu, Z., Gao, N., Wang, L., Cui, H., Qian, Y., Jiang, L., Luo, J., Yi, Z., Lu, H., Li, D. & Liu, M. 2011. PAK1IP1, a ribosomal stress-induced nucleolar protein, regulates cell proliferation via the p53–MDM2 loop. *Nucleic Acids Research,* 39**,** 2234-2248.

Yue, S., Shi, H., Han, J., Zhang, T., Zhu, W. & Zhang, D. 2016. Prognostic value of microRNA-126 and CRK expression in gastric cancer. *OncoTargets and therapy,* 9**,** 6127-6135.

Zemp, I. & Kutay, U. 2007. Nuclear export and cytoplasmic maturation of ribosomal subunits. *FEBS Lett,* 581**,** 2783-93.

Zemp, I., Wild, T., O'donohue, M.-F., Wandrey, F., Widmann, B., Gleizes, P.-E. & Kutay, U. 2009. Distinct cytoplasmic maturation steps of 40S ribosomal subunit precursors require hRio2. *The Journal of Cell Biology,* 185**,** 1167-1180.

Zhang, Y. & Lu, H. 2009. Signaling to p53: ribosomal proteins find their way. *Cancer Cell,* 16**,** 369-77.

Zhang, Y., Morimoto, K., Danilova, N., Zhang, B. & Lin, S. 2012. Zebrafish models for dyskeratosis congenita reveal critical roles of p53 activation contributing to hematopoietic defects through RNA processing. *PLoS One,* 7**,** e30188. PMC3267717

Zhao, C., Andreeva, V., Gibert, Y., Labonty, M., Lattanzi, V., Prabhudesai, S., Zhou, Y., Zon, L., Mccann, K. L., Baserga, S. & Yelick, P. C. 2014. Tissue Specific Roles for the Ribosome Biogenesis Factor Wdr43 in Zebrafish Development. *PLoS Genet,* 10**,** e1004074.

Zheng, Y., John, S., Pesavento, J. J., Schultz-Norton, J. R., Schiltz, R. L., Baek, S., Nardulli, A. M., Hager, G. L., Kelleher, N. L. & Mizzen, C. A. 2010. Histone H1 phosphorylation is associated with transcription by RNA polymerases I and II. *J Cell Biol,* 189**,** 407-15. PMC2867294

Zhou, J., Gao, Y., Lan, Y., Jia, S. & Jiang, R. 2013. Pax9 regulates a molecular network involving Bmp4, Fgf10, Shh signaling and the Osr2 transcription factor to control palate morphogenesis. *Development (Cambridge, England),* 140**,** 4709-4718.

Zink, D., Fischer, A. H. & Nickerson, J. A. 2004. Nuclear structure in cancer cells. *Nat Rev Cancer,* 4**,** 677-687.

APPENDIX I

Statistics for the full list of high-confidence screen hits that give the one nucleolus phenotype, including percent effect, viability relative to siGFP, and nucleolar classification

1 Nucleolus PERCENT EFFECT	PERCENT VIABILITY relative to siGFP	HGNC Symbol	Description	HGNC ID	Ensemble ID	Present in NOPDB? (1=YES, 0=NO)	Present in Gautier? (1=YES, 0=NO)	Nucleolar in HumanAtlas? (1=YES, 0=NO)	Sum of Gautier, NORDB and Human Atlas (0-not nucleolar, 1 - nucleolar in one database; 2 - nucleolar in two databases; 3 - nucleolar in all three databases	1 Nucleolus hit? (1-yes, 0-no)
207.36	2.15	UBC	ubiquitin C	HGNC:12468	ENSG00000150991	0	0	0	0	1
127.35	2.25	OR10A2	olfactory receptor family 10 subfamily A member 2	HGNC:8161	ENSG00000170790	0	0	0	0	1
274.19	3.10	GLRA2	glycine receptor alpha 2	HGNC:4327	ENSG00000101958	0	0	0	0	1
176.34	3.26	PXMP2	peroxisomal membrane protein 2 22kDa	HGNC:9716	ENSG00000176894	0	0	0	0	1
132.04	3.52	RAI1	retinoic acid induced 1	HGNC:9834	ENSG00000108557	0	1	0	1	1

171.87	3.89	PIK3R6	phosphoinositide-3-kinase regulatory subunit 6	HGNC:27101	ENSG00000174083	0	0	0	0	1
150.37	4.05	TMBIM4	transmembrane BAX inhibitor motif containing 4	HGNC:24257	ENSG00000155957	0	0	0	0	1
149.80	4.24	OR2T12	olfactory receptor family 2 subfamily T member 12	HGNC:19592	ENSG00000177201	0	0	0	0	1
206.32	4.27	NOX1	NADPH oxidase 1	HGNC:7889	ENSG00000007952	0	0	0	0	1
147.01	4.80	CNOT1	CCR4-NOT transcription complex subunit 1	HGNC:7877	ENSG00000125107	0	0	0	0	1
163.48	5.01	RPL8	ribosomal protein L8	HGNC:10368	ENSG00000161016	1	1	1	3	1
203.78	5.06	DDX19B	DEAD (Asp-Glu-Ala-Asp) box polypeptide 19B	HGNC:2742	ENSG00000157349	0	0	0	0	1
139.36	5.14	ST8SIA2	ST8 alpha-N-acetyl-neuraminide alpha-	HGNC:10870	ENSG00000140557	0	0	0	0	1

			28-sialyltransferase 2							
200.47	5.52	CYTH1	cytohesin 1	HGNC:9501	ENSG00000108669	0	0	0	0	1
144.47	5.76	NXF1	nuclear RNA export factor 1	HGNC:8071	ENSG00000162231	0	0	0	0	1
183.15	5.77	KCNAB1	potassium voltage-gated channel shaker-related subfamily beta member 1	HGNC:6228	ENSG00000169282	0	0	0	0	1
192.57	6.02	UBB	ubiquitin B	HGNC:12463	ENSG00000170315	0	0	0	0	1
222.28	6.35	PIK3C2A	phosphatidylinositol-4-phosphate 3-kinase, catalytic subunit type 2 alpha	HGNC:8971	ENSG00000011405	0	0	0	0	1
127.95	6.63	POLR2C	polymerase (RNA) II (DNA directed) polypeptide C 33kDa	HGNC:9189	ENSG00000102978	0	0	0	0	1
131.93	6.70	HMGN2	high mobility group nucleosomal	HGNC:4986	ENSG00000198830	0	0	1	1	1

			binding domain 2							
205.75	6.89	SPAG1	sperm associated antigen 1	HGNC:11212	ENSG00000104450	0	0	0	0	1
264.59	7.01	SYNE1	spectrin repeat containing nuclear envelope 1	HGNC:17089	ENSG00000131018	0	1	0	1	1
152.86	7.12	COPB1	coatomer protein complex subunit beta 1	HGNC:2231	ENSG00000129083	0	0	0	0	1
152.64	7.91	KCNK16	potassium channel subfamily K member 16	HGNC:14464	ENSG00000095981	0	0	0	0	1
128.08	8.50	ACTG1	actin gamma 1	HGNC:144	ENSG00000184009	0	0	0	0	1
131.81	9.70	GCLM	glutamate-cysteine ligase modifier subunit	HGNC:4312	ENSG00000023909	0	0	0	0	1
152.24	10.04	AVP	arginine vasopressin	HGNC:894	ENSG00000101200	0	0	0	0	1
124.05	10.32	ONECUT2	one cut homeobox 2	HGNC:8139	ENSG00000119547	0	0	0	0	1
130.51	11.01	TRPM3	transient receptor potential cation channel	HGNC:17992	ENSG00000083067	0	0	0	0	1

			subfamily M member 3							
139.83	11.56	HIST1H2BO	histone cluster 1 H2bo	HGNC:4758	ENSG00000196331	1	0	0	1	1
128.31	15.07	CBR1	carbonyl reductase 1	HGNC:1548	ENSG00000159228	0	0	0	0	1
160.79	15.28	RPS28	ribosomal protein S28	HGNC:10418	ENSG00000233927	1	0	0	1	1
138.81	15.38	NUMA1	nuclear mitotic apparatus protein 1	HGNC:8059	ENSG00000137497	1	1	0	2	1
135.08	16.79	RPS16	ribosomal protein S16	HGNC:10396	ENSG00000105193	1	1	0	2	1
122.99	18.48	ARMC2	armadillo repeat containing 2	HGNC:23045	ENSG00000118690	0	0	0	0	1
140.12	19.26	RPS4X	ribosomal protein S4 X-linked	HGNC:10424	ENSG00000198034	1	1	0	2	1
129.64	20.26	F8	coagulation factor VIII procoagulant component	HGNC:3546	ENSG00000185010	0	0	0	0	1
155.62	22.79	HES4	hairy and enhancer of split 4 (Drosophila)	HGNC:24149	ENSG00000188290	0	0	0	0	1

155.77	22.94	PPP1R3B	protein phosphatase 1 regulatory subunit 3B	HGNC:14942	ENSG00000173281	0	0	0	0	1
122.49	24.76	CNPY4	canopy 4 homolog (zebrafish)	HGNC:28631	ENSG00000166997	0	0	0	0	1
122.01	25.57	GRB2	growth factor receptor-bound protein 2	HGNC:4566	ENSG00000177885	0	0	0	0	1
137.74	25.88	PDE6H	phosphodiesterase 6H cGMP-specific cone gamma	HGNC:8790	ENSG00000139053	0	0	0	0	1
135.64	26.03	C16orf54	chromosome 16 open reading frame 54	HGNC:26649	ENSG00000185905	0	0	0	0	1
125.81	26.47	UBE4B	ubiquitination factor E4B	HGNC:12500	ENSG00000130939	0	0	0	0	1
128.91	26.89	RASAL2	RAS protein activator like 2	HGNC:9874	ENSG00000075391	0	0	0	0	1
126.80	27.84	SUPT5H	suppressor of Ty 5 homolog (S. cerevisiae)	HGNC:11469	ENSG00000196235	0	1	0	1	1
135.84	28.17	RUSC2	RUN and SH3 domain containing 2	HGNC:23625	ENSG00000198853	0	0	0	0	1

134.64	28.46	TMC2	transmembrane channel-like 2	HGNC:16527	ENSG00000149488	0	0	0	0	1
137.34	28.66	NUDT13	nudix (nucleoside diphosphate linked moiety X)-type motif 13	HGNC:18827	ENSG00000166321	0	0	0	0	1
124.16	28.74	TNFRSF8	tumor necrosis factor receptor superfamily member 8	HGNC:11923	ENSG00000120949	0	0	0	0	1
125.70	29.02	OR2A4	olfactory receptor family 2 subfamily A member 4	HGNC:14729	ENSG00000180658	0	0	0	0	1
137.19	29.13	AP2M1	adaptor-related protein complex 2 mu 1 subunit	HGNC:564	ENSG00000161203	0	0	0	0	1
124.01	29.90	FAM193B	family with sequence similarity 193 member B	HGNC:25524	ENSG00000146067	0	0	0	0	1
163.18	29.96	RPS14	ribosomal protein S14	HGNC:10387	ENSG00000164587	1	1	0	2	1
126.84	29.97	UBA52	ubiquitin A-52 residue ribosom	HGNC:12458	ENSG00000221983	0	0	0	0	1

			al protein fusion product 1							
126.68	30.04	FAM129B	family with sequence similarity 129 member B	HGNC:25282	ENSG00000136830	0	0	0	0	1
127.52	30.93	CRYBG3	beta-gamma crystallin domain containing 3	HGNC:34427	ENSG00000080200	0	0	0	0	1
125.28	31.11	GIN1	gypsy retrotransposon integrase 1	HGNC:25959	ENSG00000145723	0	0	0	0	1
122.37	31.27	EPSTI1	epithelial stromal interaction 1 (breast)	HGNC:16465	ENSG00000133106	0	0	0	0	1
130.98	31.49	ANXA11	annexin A11	HGNC:535	ENSG00000122359	0	0	0	0	1
195.78	31.51	NODAL	nodal growth differentiation factor	HGNC:7865	ENSG00000156574	0	0	0	0	1
129.35	32.32	KCTD11	potassium channel tetramerisation domain containing 11	HGNC:21302	ENSG00000213859	0	0	0	0	1
122.97	32.64	ASCC2	activating signal cointegrator 1	HGNC:24103	ENSG00000100325	0	0	0	0	1

			complex subunit 2							
172.11	33.28	TNFRSF21	tumor necrosis factor receptor superfamily member 21	HGNC:13469	ENSG00000146072	0	0	0	0	1
199.66	33.34	APBB1	amyloid beta (A4) precursor protein-binding family B member 1 (Fe65)	HGNC:581	ENSG00000166313	0	0	0	0	1
129.19	33.36	C21orf2	chromosome 21 open reading frame 2	HGNC:1260	ENSG00000160226	0	0	0	0	1
132.68	33.38	HLA-B	major histocompatibility complex class I B	HGNC:4932	ENSG00000206450	0	0	0	0	1
130.03	33.51	NKAIN1	Na+/K+ transporting ATPase interacting 1	HGNC:25743	ENSG00000084628	0	0	0	0	1
123.82	34.32	TMEM132E	transmembrane protein 132E	HGNC:26991	ENSG00000181291	0	0	0	0	1
156.44	34.97	RGN	regucalcin (senescence marker protein-30)	HGNC:9989	ENSG00000130988	0	0	0	0	1

130.01	35.18	SFXN3	sideroflexin 3	HGNC:16087	ENSG00000107819	0	0	0	0	1
161.14	35.21	RPLP2	ribosomal protein large P2	HGNC:10377	ENSG00000177600	1	1	0	2	1
127.95	35.44	PRL	prolactin	HGNC:9445	ENSG00000172179	0	0	0	0	1
147.70	35.72	FBXO33	F-box protein 33	HGNC:19833	ENSG00000165355	0	0	0	0	1
129.05	35.92	NEXN	nexilin (F actin binding protein)	HGNC:29557	ENSG00000162614	0	0	0	0	1
122.12	35.97	CCDC22	coiled-coil domain containing 22	HGNC:28909	ENSG00000101997	0	0	0	0	1
130.37	35.99	FBXW8	F-box and WD repeat domain containing 8	HGNC:13597	ENSG00000174989	0	0	0	0	1
129.24	36.07	KLF16	Kruppel-like factor 16	HGNC:16857	ENSG00000129911	0	0	0	0	1
143.42	36.75	LSM14A	LSM14A SCD6 homolog A (S. cerevisiae)	HGNC:24489	ENSG00000257103	0	0	0	0	1
180.93	37.39	RPS3A	ribosomal protein S3A	HGNC:10421	ENSG00000145425	0	1	0	1	1
215.14	37.63	RPS11	ribosomal protein S11	HGNC:10384	ENSG00000142534	1	1	0	2	1

137.34	37.69	C2orf57	chromosome 2 open reading frame 57	HGNC:28563	ENSG00000177673	0	0	0	0	1
127.94	37.72	CLCN7	chloride channel voltage-sensitive 7	HGNC:2025	ENSG00000103249	0	0	0	0	1
128.38	37.83	MMP25	matrix metallopeptidase 25	HGNC:14246	ENSG00000008516	0	0	0	0	1
133.32	38.02	COL9A3	collagen type IX alpha 3	HGNC:2219	ENSG00000092758	0	0	0	0	1
129.80	38.47	HIST1H2AL	histone cluster 1 H2al	HGNC:4730	ENSG00000198374	0	0	0	0	1
141.50	38.74	EPB41L2	erythrocyte membrane protein band 4.1-like 2	HGNC:3379	ENSG00000079819	0	0	0	0	1
126.42	38.82	TAPBP	TAP binding protein (tapasin)	HGNC:11566	ENSG00000112493	0	0	0	0	1
127.90	38.90	OPTN	optineurin	HGNC:17142	ENSG00000123240	0	0	0	0	1
173.70	39.21	PAX9	paired box 9	HGNC:8623	ENSG00000198807	0	0	0	0	1
152.88	39.53	SYVN1	synovial apoptosis inhibitor 1 synoviolin	HGNC:20738	ENSG00000162298	0	0	0	0	1

168.19	39.66	LGI1	leucine-rich glioma inactivated 1	HGNC:6572	ENSG00000108231	0	0	0	0	1
131.84	40.38	SLC39A7	solute carrier family 39 (zinc transporter) member 7	HGNC:4927	ENSG00000112473	0	0	0	0	1
183.71	40.72	MRPL34	mitochondrial ribosomal protein L34	HGNC:14488	ENSG00000130312	0	0	0	0	1
137.97	41.27	CEBPD	CCAAT/enhancer binding protein (C/EBP) delta	HGNC:1835	ENSG00000221869	0	0	0	0	1
139.22	41.30	SPRR3	small proline-rich protein 3	HGNC:11268	ENSG00000163209	0	0	0	0	1
133.19	41.43	CRK	v-crk sarcoma virus CT10 oncogene homolog (avian)	HGNC:2362	ENSG00000167193	0	0	0	0	1
152.51	41.94	ZZZ3	zinc finger ZZ-type containing 3	HGNC:24523	ENSG00000036549	0	0	0	0	1
137.29	42.07	ZNF684	zinc finger protein 684	HGNC:28418	ENSG00000117010	0	0	0	0	1
155.75	42.15	WNT4	wingless-type MMTV integrati	HGNC:12783	ENSG00000162552	0	0	0	0	1

			on site family member 4							
141.05	42.19	ZAN	zonadhesin	HGNC:12857	ENSG00000146839	0	0	0	0	1
192.72	42.47	RPS5	ribosomal protein S5	HGNC:10426	ENSG00000083845	1	1	0	2	1
217.24	43.07	IQSEC3	IQ motif and Sec7 domain 3	HGNC:29193	ENSG00000120645	0	0	0	0	1
174.89	43.62	OR9I1	olfactory receptor family 9 subfamily I member 1	HGNC:14718	ENSG00000172377	0	0	0	0	1
122.89	43.89	TERF2	telomeric repeat binding factor 2	HGNC:11729	ENSG00000132604	0	1	0	1	1
147.14	44.16	OR6X1	olfactory receptor family 6 subfamily X member 1	HGNC:14737	ENSG00000221931	0	0	0	0	1
144.24	44.21	OR5I1	olfactory receptor family 5 subfamily I member 1	HGNC:8347	ENSG00000167825	0	0	0	0	1
124.57	44.80	CHN2	chimerin 2	HGNC:1944	ENSG00000106069	0	0	0	0	1
131.36	45.00	LRRTM1	leucine rich repeat transme	HGNC:19408	ENSG00000162951	0	0	0	0	1

			mbrane neuronal 1							
123.78	45.01	ZSCAN26	zinc finger and SCAN domain containing 26	HGNC:12978		0	0	0	0	1
122.87	45.09	LRRC4	leucine rich repeat containing 4	HGNC:15586	ENSG00000128594	0	0	0	0	1
125.14	45.31	INA	internexin neuronal intermediate filament protein alpha	HGNC:6057	ENSG00000148798	0	0	0	0	1
169.72	45.51	RAB38	RAB38 member RAS oncogene family	HGNC:9776	ENSG00000123892	0	0	0	0	1
153.87	45.62	SCN1B	sodium channel, voltage-gated, type I, beta subunit	HGNC:10586	ENSG00000105711	0	0	0	0	1
131.02	45.70	NDUFS3	NADH dehydrogenase (ubiquinone) Fe-S protein 3 30kDa (NADH-coenzyme Q reductase)	HGNC:7710	ENSG00000213619	0	0	0	0	1
165.93	45.81	RPS9	ribosomal	HGNC:	ENSG00000	1	1	0	2	1

			protein S9	10442	170889					
139.41	46.05	KAT5	K(lysine) acetyltransferase 5	HGNC:5275	ENSG00000172977	0	0	0	0	1
125.63	46.17	SNX13	sorting nexin 13	HGNC:21335	ENSG00000071189	0	0	0	0	1
133.16	46.19	CMIP	c-Maf inducing protein	HGNC:24319	ENSG00000153815	0	0	0	0	1
132.58	46.95	TSGA10	testis specific 10	HGNC:14927	ENSG00000135951	0	0	0	0	1
160.66	47.27	UPK3B	uroplakin 3B	HGNC:21444	ENSG00000243566	0	0	0	0	1
133.34	47.33	KRTAP13-2	keratin associated protein 13-2	HGNC:18923	ENSG00000182816	0	0	0	0	1
127.21	47.52	LSM12	LSM12 homolog (S. cerevisiae)	HGNC:26407	ENSG00000161654	0	0	0	0	1
122.04	47.63	CCDC34	coiled-coil domain containing 34	HGNC:25079	ENSG00000109881	0	0	0	0	1
127.24	48.20	POLR2J3	polymerase (RNA) II (DNA directed) polypeptide J3	HGNC:33853	ENSG00000168255	0	0	0	0	1
154.21	48.25	CHAF1A	chromatin assembly factor 1	HGNC:1910	ENSG00000167670	0	0	0	0	1

			subunit A (p150)							
126.38	48.35	TCP10	t-complex 10	HGNC:11656	ENSG00000203690	0	0	0	0	1
125.33	48.44	XCL2	chemokine (C motif) ligand 2	HGNC:10646	ENSG00000143185	0	0	0	0	1
124.43	50.08	OR2T11	olfactory receptor family 2 subfamily T member 11	HGNC:19574	ENSG00000183130	0	0	0	0	1
132.73	50.25	PDE12	phosphodiesterase 12	HGNC:25386	ENSG00000174840	0	0	0	0	1
126.07	50.32	LRRC16A	leucine rich repeat containing 16A	HGNC:21581	ENSG00000079691	0	0	0	0	1
158.01	50.33	RPS24	ribosomal protein S24	HGNC:10411	ENSG00000138326	1	1	0	2	1
141.97	50.33	SPOP	speckle-type POZ protein	HGNC:11254	ENSG00000121067	0	0	0	0	1
164.21	50.36	RPS13	ribosomal protein S13	HGNC:10386	ENSG00000110700	1	1	0	2	1
150.49	50.47	RBM28	RNA binding motif protein 28	HGNC:21863	ENSG00000106344	1	1	1	3	1
139.90	50.74	LRRC71	leucine rich repeat containing 71	HGNC:26556	ENSG00000160838	0	0	0	0	1
137.41	50.80	RRN3	RRN3 RNA	HGNC:	ENSG00000	1	0	1	2	1

			polymerase I transcription factor homolog (S. cerevisiae)	30346	085721					
135.95	50.82	TBC1D5	TBC1 domain family member 5	HGNC:19166	ENSG00000131374	0	0	0	0	1
143.17	51.02	OR2AK2	olfactory receptor family 2 subfamily AK member 2	HGNC:19569	ENSG00000187080	0	0	0	0	1
138.80	51.09	OR52A5	olfactory receptor family 52 subfamily A member 5	HGNC:19580	ENSG00000171944	0	0	0	0	1
127.47	51.21	SLC22A16	solute carrier family 22 (organic cation/carnitine transporter) member 16	HGNC:20302	ENSG00000004809	0	0	0	0	1
140.03	51.40	ZNF76	zinc finger protein 76	HGNC:13149	ENSG00000065029	0	0	1	1	1
129.80	51.52	SCAF1	SR-related CTD-associated factor 1	HGNC:30403	ENSG00000126461	0	0	0	0	1

146.44	51.65	DLGAP1	discs large (Drosophila) homolog-associated protein 1	HGNC:2905	ENSG00000170579	0	0	0	0	1
126.61	51.73	SELL	selectin L	HGNC:10720	ENSG00000188404	0	0	0	0	1
125.09	51.76	TECTB	tectorin beta	HGNC:11721	ENSG00000119913	0	0	0	0	1
143.64	51.96	RNF169	ring finger protein 169	HGNC:26961	ENSG00000166439	0	0	0	0	1
131.25	51.98	MOG	myelin oligodendrocyte glycoprotein	HGNC:7197	ENSG00000137345	0	0	0	0	1
127.04	52.35	ATP13A5	ATPase type 13A5	HGNC:31789	ENSG00000187527	0	0	0	0	1
152.39	52.97	NTN3	netrin 3	HGNC:8030	ENSG00000162068	0	0	0	0	1
130.09	53.31	APOBEC4	apolipoprotein B mRNA editing enzyme catalytic polypeptide-like 4 (putative)	HGNC:32152	ENSG00000173627	0	0	0	0	1
147.28	54.38	OR6B3	olfactory receptor family 6 subfamily B	HGNC:15042	ENSG00000178586	0	0	0	0	1

			member 3							
134.77	54.48	PRR19	proline rich 19	HGNC:33728	ENSG00000188368	0	0	0	0	1
135.73	55.15	PCTP	phosphatidylcholine transfer protein	HGNC:8752	ENSG00000141179	0	0	0	0	1
133.41	55.28	POLR2E	polymerase (RNA) II (DNA directed) polypeptide E 25kDa	HGNC:9192	ENSG00000099817	1	0	0	1	1
129.27	56.28	SAMD15	sterile alpha motif domain containing 15	HGNC:18631	ENSG00000100583	0	0	1	1	1
155.58	56.54	RSL24D1	ribosomal L24 domain containing 1	HGNC:18479	ENSG00000137876	1	0	0	1	1
146.64	58.08	ARL11	ADP-ribosylation factor-like 11	HGNC:24046	ENSG00000152213	0	0	0	0	1
156.00	58.16	GJC3	gap junction protein gamma 3 30.2kDa	HGNC:17495	ENSG00000176402	0	0	0	0	1
122.07	59.44	FAM110C	family with sequence similarity 110 member C	HGNC:33340	ENSG00000184731	0	0	0	0	1

136.49	59.68	ANKHD1-EIF4EBP3	ANKHD1-EIF4EBP3 readthrough	HGNC:33530	ENSG00000254996	0	0	0	0	1
144.67	60.23	APBA1	amyloid beta (A4) precursor protein-binding family A member 1	HGNC:578	ENSG00000107282	0	0	0	0	1
178.95	60.32	ARHGAP28	Rho GTPase activating protein 28	HGNC:25509	ENSG00000088756	0	0	0	0	1
133.13	61.13	DNTTIP2	deoxynucleotidyltransferase terminal interacting protein 2	HGNC:24013	ENSG00000067334	1	1	1	3	1
125.17	61.25	RBM43	RNA binding motif protein 43	HGNC:24790	ENSG00000184898	0	0	0	0	1
144.81	61.57	CES2	carboxyl esterase 2	HGNC:1864	ENSG00000172831	0	0	0	0	1
131.05	61.64	IL1B	interleukin 1 beta	HGNC:5992	ENSG00000125538	0	0	0	0	1
122.06	61.64	NMT2	N-myristoyltransferase 2	HGNC:7858	ENSG00000152465	0	0	0	0	1
137.27	61.92	TUFT1	tuftelin 1	HGNC:12422	ENSG00000143367	0	0	0	0	1

156.52	62.21	LYG1	lysozyme G-like 1	HGNC:27014	ENSG00000144214	0	0	0	0	1
133.24	62.58	METTL21D	methyltransferase like 21D	HGNC:20352	ENSG00000100483	0	0	0	0	1
134.95	63.08	ATP6AP1L	ATPase, H+ transporting, lysosomal accessory protein 1-like	HGNC:28091	ENSG00000205464	0	0	0	0	1
126.10	63.43	C10orf62	chromosome 10 open reading frame 62	HGNC:23294	ENSG00000203942	0	0	0	0	1
157.96	63.43	SP7	Sp7 transcription factor	HGNC:17321	ENSG00000170374	0	0	0	0	1
162.87	63.56	NOMO1	NODAL modulator 1	HGNC:30060	ENSG00000103512	0	0	0	0	1
122.10	63.81	CILP2	cartilage intermediate layer protein 2	HGNC:24213	ENSG00000160161	0	0	0	0	1
128.32	63.88	IKZF4	IKAROS family zinc finger 4 (Eos)	HGNC:13179	ENSG00000123411	0	0	0	0	1
186.42	63.90	THAP1	THAP domain containing apoptosis associat	HGNC:20856	ENSG00000131931	0	0	0	0	1

			ed protein 1							
153.70	64.25	ZNF658	zinc finger protein 658	HGNC:25226	ENSG00000196409	0	0	0	0	1
127.70	64.97	SEPT12	septin 12	HGNC:26348	ENSG00000140623	0	0	0	0	1
152.02	65.49	OR9K2	olfactory receptor family 9 subfamily K member 2	HGNC:15339	ENSG00000170605	0	0	0	0	1
139.89	67.40	AAMP	angio-associated migratory cell protein	HGNC:18	ENSG00000127837	0	0	0	0	1
126.71	68.86	SELV	selenoprotein V	-	ENSG00000186838	0	0	0	0	1
162.50	75.72	ANLN	anillin actin binding protein	HGNC:14082	ENSG00000011426	1	0	0	1	1
176.24	76.67	THAP6	THAP domain containing 6	HGNC:23189	ENSG00000174796	0	0	0	0	1
169.29	82.69	LIN28A	lin-28 homolog A (C. elegans)	HGNC:15986	ENSG00000131914	0	0	0	0	1
136.44	82.98	FAM179A	family with sequence similarity 179 member A	HGNC:33715	ENSG00000189350	0	0	0	0	1
127.46	83.13	TCL1A	T-cell leukemi	HGNC:	ENSG00000	0	0	0	0	1

			a/lymphoma 1A	11648	100721					
136.78	84.47	MPHOSPH10	M-phase phosphoprotein 10 (U3 small nucleolar ribonucleoprotein)	HGNC:7213	ENSG00000124383	1	1	1	3	1
125.78	91.56	DDX56	DEAD (Asp-Glu-Ala-Asp) box helicase 56	HGNC:18193	ENSG00000136271	1	1	0	2	1
151.45	110.64	NOP58	NOP58 ribonucleoprotein	HGNC:29926	ENSG00000055044	1	1	1	3	1

APPENDIX II

Filtered list of 139 high-confidence screen hits after removal of hits with low viability (<10% relative to siGFP control) and low expression (FPKM = 0 in breast cells, Illumina Body Map)

1 Nucleolus PERCENT EFFECT	PERCENT VIABILITY relative to siGFP	HGNC Symbol	Description	HGNC ID	Ensemble ID	Present in NOPDB? (1=YES, 0=NO)	Present in Gautier? (1=YES, 0=NO)	Nucleolar in HumanAtlas? (1=YES, 0=NO)	Sum of Gautier, NOPDB and HumanAtlas (0-not nucleolar, 1 - nucleolar in one database; 2 - nucleolar in two databases; 3 - nucleolar in all three databases
139.89	67.40	AAMP	angio-associated migratory cell protein	HGNC:18	ENSG00000127837	0	0	0	0
136.49	59.68	ANKHD1-EIF4EBP3	ANKHD1-EIF4EBP3 readthrough	HGNC:33530	ENSG00000254996	0	0	0	0
162.50	75.72	ANLN	anillin actin binding protein	HGNC:14082	ENSG00000011426	1	0	0	1
130.98	31.49	ANXA11	annexin A11	HGNC:535	ENSG00000122359	0	0	0	0

137.19	29.13	AP2M1	adaptor-related protein complex 2 mu 1 subunit	HGNC: 564	ENSG00000161203	0	0	0	0
144.67	60.23	APBA1	amyloid beta (A4) precursor protein-binding family A member 1	HGNC: 578	ENSG00000107282	0	0	0	0
199.66	33.34	APBB1	amyloid beta (A4) precursor protein-binding family B member 1 (Fe65)	HGNC: 581	ENSG00000166313	0	0	0	0
178.95	60.32	ARHGAP28	Rho GTPase activating protein 28	HGNC: 25509	ENSG00000088756	0	0	0	0

146.64	58.08	ARL11	ADP-ribosylation factor-like 11	HGNC: 24046	ENSG00000152213	0	0	0	0
122.99	18.48	ARMC2	armadillo repeat containing 2	HGNC: 23045	ENSG00000118690	0	0	0	0
122.97	32.64	ASCC2	activating signal cointegrator 1 complex subunit 2	HGNC: 24103	ENSG00000100325	0	0	0	0
127.04	52.35	ATP13A5	ATPase type 13A5	HGNC: 31789	ENSG00000187527	0	0	0	0
134.95	63.08	ATP6AP1L	ATPase, H+ transporting, lysosomal accessory protein 1-like	HGNC: 28091	ENSG00000205464	0	0	1	1
152.24	10.04	AVP	arginine vasopressin	HGNC: 894	ENSG00000101200	0	0	0	0

135.64	26.03	C16orf54	chromosome 16 open reading frame 54	HGNC:26649	ENSG00000185905	0	0	0	0
129.19	33.36	C21orf2	chromosome 21 open reading frame 2	HGNC:1260	ENSG00000160226	0	0	0	0
128.31	15.07	CBR1	carbonyl reductase 1	HGNC:1548	ENSG00000159228	0	0	0	0
122.12	35.97	CCDC22	coiled-coil domain containing 22	HGNC:28909	ENSG00000101997	0	0	0	0
122.04	47.63	CCDC34	coiled-coil domain containing 34	HGNC:25079	ENSG00000109881	0	0	1	1
137.97	41.27	CEBPD	CCAAT/enhancer binding protein (C/EB	HGNC:1835	ENSG00000221869	0	0	0	0

			P) delta						
144.81	61.57	CES2	carboxylesterase 2	HGNC:1864	ENSG00000172831	0	0	0	0
154.21	48.25	CHAF1A	chromatin assembly factor 1 subunit A (p150)	HGNC:1910	ENSG00000167670	0	0	0	0
124.57	44.80	CHN2	chimerin 2	HGNC:1944	ENSG00000106069	0	0	0	0
122.10	63.81	CILP2	cartilage intermediate layer protein 2	HGNC:24213	ENSG00000160161	0	0	0	0
127.94	37.72	CLCN7	chloride channel voltage-sensitive 7	HGNC:2025	ENSG00000103249	0	0	0	0
133.16	46.19	CMIP	c-Maf inducing protein	HGNC:24319	ENSG00000153815	0	0	0	0
122.49	24.76	CNPY4	canopy 4 homolog	HGNC:28631	ENSG00000166997	0	0	0	0

			(zebrafish)						
133.32	38.02	COL9A3	collagen type IX alpha 3	HGNC:2219	ENSG00000092758	0	0	0	0
133.19	41.43	CRK	v-crk sarcoma virus CT10 oncogene homolog (avian)	HGNC:2362	ENSG00000167193	0	0	0	0
127.52	30.93	CRYBG3	beta-gamma crystallin domain containing 3	HGNC:34427	ENSG00000080200	0	0	0	0
125.78	91.56	DDX56	DEAD (Asp-Glu-Ala-Asp) box helicase 56	HGNC:18193	ENSG00000136271	1	1	0	2
146.44	51.65	DLGAP1	discs large (Drosophila) homolog-assoc	HGNC:2905	ENSG00000170579	0	0	0	0

			iated protein 1						
133.13	61.13	DNTTIP2	deoxynucleotidyltransferase terminal interacting protein 2	HGNC: 24013	ENSG00000067334	1	1	1	3
141.50	38.74	EPB41L2	erythrocyte membrane protein band 4.1-like 2	HGNC: 3379	ENSG00000079819	0	0	0	0
122.37	31.27	EPSTI1	epithelial stromal interaction 1 (breast)	HGNC: 16465	ENSG00000133106	0	0	0	0
129.64	20.26	F8	coagulation factor VIII procoagulant component	HGNC: 3546	ENSG00000185010	0	0	0	0

122.07	59.44	FAM 110C	family with sequence similarity 110 member C	HGNC: 33340	ENSG 0000 0184 731	0	0	0	0
126.68	30.04	FAM 129B	family with sequence similarity 129 member B	HGNC: 25282	ENSG 0000 0136 830	0	0	0	0
136.44	82.98	FAM 179A	family with sequence similarity 179 member A	HGNC: 33715	ENSG 0000 0189 350	0	0	0	0
124.01	29.90	FAM 193B	family with sequence similarity 193 member B	HGNC: 25524	ENSG 0000 0146 067	0	0	1	1
147.70	35.72	FBXO 33	F-box protein 33	HGNC: 19833	ENSG 0000 0165 355	0	0	0	0
130.37	35.99	FBX W8	F-box and WD	HGNC: 13597	ENSG 0000	0	0	0	0

			repeat domain containing 8		0174989				
125.28	31.11	GIN1	gypsy retrotransposon integrase 1	HGNC: 25959	ENSG00000145723	0	0	1	1
156.00	58.16	GJC3	gap junction protein gamma 3 30.2kDa	HGNC: 17495	ENSG00000176402	0	0	0	0
122.01	25.57	GRB2	growth factor receptor-bound protein 2	HGNC: 4566	ENSG00000177885	0	0	1	1
155.62	22.79	HES4	hairy and enhancer of split 4 (Drosophila)	HGNC: 24149	ENSG00000188290	0	0	0	0

139.83	11.56	HIST1H2BO	histone cluster 1 H2bo	HGNC:4758	ENSG00000196331	1	0	0	1
132.68	33.38	HLA-B	major histocompatibility complex class I B	HGNC:4932	ENSG00000206450	0	0	0	0
128.32	63.88	IKZF4	IKAROS family zinc finger 4 (Eos)	HGNC:13179	ENSG00000123411	0	0	0	0
131.05	61.64	IL1B	interleukin 1 beta	HGNC:5992	ENSG00000125538	0	0	0	0
217.24	43.07	IQSEC3	IQ motif and Sec7 domain 3	HGNC:29193	ENSG00000120645	0	0	0	0
139.41	46.05	KAT5	K(lysine) acetyltransferase 5	HGNC:5275	ENSG00000172977	0	0	0	0
129.35	32.32	KCTD11	potassium channel tetramerisation domain	HGNC:21302	ENSG00000213859	0	0	0	0

			containing 11						
129.24	36.07	KLF16	Kruppel-like factor 16	HGNC:16857	ENSG00000129911	0	0	0	0
169.29	82.69	LIN28A	lin-28 homolog A (C. elegans)	HGNC:15986	ENSG00000131914	0	0	0	0
122.87	45.09	LRRC4	leucine rich repeat containing 4	HGNC:15586	ENSG00000128594	0	0	0	0
139.90	50.74	LRRC71	leucine rich repeat containing 71	HGNC:26556	ENSG00000160838	0	0	0	0
127.21	47.52	LSM12	LSM12 homolog (S. cerevisiae)	HGNC:26407	ENSG00000161654	0	0	0	0
143.42	36.75	LSM14A	LSM14A SCD6 homolog A (S. cerevisiae)	HGNC:24489	ENSG00000257103	0	0	0	0

156.52	62.21	LYG1	lysozyme G-like 1	HGNC:27014	ENSG00000144214	0	0	0	0
128.38	37.83	MMP25	matrix metallopeptidase 25	HGNC:14246	ENSG00000008516	0	0	0	0
131.25	51.98	MOG	myelin oligodendrocyte glycoprotein	HGNC:7197	ENSG00000137345	0	0	0	0
136.78	84.47	MPHOSPH10	M-phase phosphoprotein 10 (U3 small nucleolar ribonucleoprotein)	HGNC:7213	ENSG00000124383	1	1	1	3
183.71	40.72	MRPL34	mitochondrial ribosomal protein L34	HGNC:14488	ENSG00000130312	0	0	0	0
131.02	45.70	NDUFS3	NADH dehydrogenas	HGNC:7710	ENSG00000213619	0	0	0	0

			e (ubiquinone) Fe-S protein 3 30kDa (NADH-coenzyme Q reductase)						
129.05	35.92	NEXN	nexilin (F actin binding protein)	HGNC: 29557	ENSG00000162614	0	0	0	0
130.03	33.51	NKAIN1	Na+/K+ transporting ATPase interacting 1	HGNC: 25743	ENSG00000084628	0	0	0	0
122.06	61.64	NMT2	N-myristoyltransferase 2	HGNC: 7858	ENSG00000152465	0	0	0	0
195.78	31.51	NODAL	nodal growth differentiation factor	HGNC: 7865	ENSG00000156574	0	0	0	0

162.87	63.56	NOMO1	NODAL modulator 1	HGNC:30060	ENSG00000103512	0	0	0	0
151.45	110.64	NOP58	NOP58 ribonucleoprotein	HGNC:29926	ENSG00000055044	1	1	1	3
152.39	52.97	NTN3	netrin 3	HGNC:8030	ENSG00000162068	0	0	0	0
137.34	28.66	NUDT13	nudix (nucleoside diphosphate linked moiety X)-type motif 13	HGNC:18827	ENSG00000166321	0	0	0	0
138.81	15.38	NUMA1	nuclear mitotic apparatus protein 1	HGNC:8059	ENSG00000137497	1	1	0	2
127.90	38.90	OPTN	optineurin	HGNC:17142	ENSG00000123240	0	0	1	1
173.70	39.21	PAX9	paired box 9	HGNC:8623	ENSG00000198807	0	0	0	0
135.73	55.15	PCTP	phosphati	HGNC:8752	ENSG0000	0	0	1	1

			dylcholine transfer protein		0141179				
132.73	50.25	PDE12	phosphodiesterase 12	HGNC: 25386	ENSG00000174840	0	0	0	0
137.74	25.88	PDE6H	phosphodiesterase 6H cGMP-specific cone gamma	HGNC: 8790	ENSG00000139053	0	0	0	0
133.41	55.28	POLR2E	polymerase (RNA) II (DNA directed) polypeptide E 25kDa	HGNC: 9192	ENSG00000099817	1	0	0	1
127.24	48.20	POLR2J3	polymerase (RNA) II (DNA directed) polypeptide J3	HGNC: 33853	ENSG00000168255	0	0	0	0

155.77	22.94	PPP1R3B	protein phosphatase 1 regulatory subunit 3B	HGNC:14942	ENSG00000173281	0	0	0	0
127.95	35.44	PRL	prolactin	HGNC:9445	ENSG00000172179	0	0	0	0
134.77	54.48	PRR19	proline rich 19	HGNC:33728	ENSG00000188368	0	0	1	1
169.72	45.51	RAB38	RAB38 member RAS oncogene family	HGNC:9776	ENSG00000123892	0	0	1	1
128.91	26.89	RASAL2	RAS protein activator like 2	HGNC:9874	ENSG00000075391	0	0	0	0
150.49	50.47	RBM28	RNA binding motif protein 28	HGNC:21863	ENSG00000106344	1	1	1	3
125.17	61.25	RBM43	RNA binding motif protein 43	HGNC:24790	ENSG00000184898	0	0	1	1
156.44	34.97	RGN	regucalcin	HGNC:9989	ENSG0000	0	0	0	0

			(senescence marker protein-30)		0130988				
143.64	51.96	RNF169	ring finger protein 169	HGNC:26961	ENSG00000166439	0	0	0	0
161.14	35.21	RPLP2	ribosomal protein large P2	HGNC:10377	ENSG00000177600	1	1	0	2
215.14	37.63	RPS11	ribosomal protein S11	HGNC:10384	ENSG00000142534	1	1	1	3
164.21	50.36	RPS13	ribosomal protein S13	HGNC:10386	ENSG00000110700	1	1	0	2
163.18	29.96	RPS14	ribosomal protein S14	HGNC:10387	ENSG00000164587	1	1	0	2
135.08	16.79	RPS16	ribosomal protein S16	HGNC:10396	ENSG00000105193	1	1	0	2
158.01	50.33	RPS24	ribosomal protein S24	HGNC:10411	ENSG00000138326	1	1	0	2
160.79	15.28	RPS28	ribosomal	HGNC:10418	ENSG0000	1	0	0	1

			protein S28		0233927				
180.93	37.39	RPS3A	ribosomal protein S3A	HGNC:10421	ENSG00000145425	0	1	1	2
140.12	19.26	RPS4X	ribosomal protein S4 X-linked	HGNC:10424	ENSG00000198034	1	1	0	2
192.72	42.47	RPS5	ribosomal protein S5	HGNC:10426	ENSG00000083845	1	1	0	2
165.93	45.81	RPS9	ribosomal protein S9	HGNC:10442	ENSG00000170889	1	1	0	2
137.41	50.80	RRN3	RRN3 RNA polymerase I transcription factor homolog (S. cerevisiae)	HGNC:30346	ENSG00000085721	1	0	1	2
155.58	56.54	RSL24D1	ribosomal L24 domain containing 1	HGNC:18479	ENSG00000137876	1	0	1	2

135.84	28.17	RUSC2	RUN and SH3 domain containing 2	HGNC:23625	ENSG00000198853	0	0	0	0
129.27	56.28	SAMD15	sterile alpha motif domain containing 15	HGNC:18631	ENSG00000100583	0	0	1	1
129.80	51.52	SCAF1	SR-related CTD-associated factor 1	HGNC:30403	ENSG00000126461	0	0	0	0
153.87	45.62	SCN1B	sodium channel, voltage-gated, type I, beta subunit	HGNC:10586	ENSG00000105711	0	0	0	0
126.61	51.73	SELL	selectin L	HGNC:10720	ENSG00000188404	0	0	0	0
130.01	35.18	SFXN3	sideroflexin 3	HGNC:16087	ENSG00000107819	0	0	0	0

127.47	51.21	SLC22A16	solute carrier family 22 (organic cation/carnitine transporter) member 16	HGNC: 20302	ENSG00000004809	0	0	0	0
131.84	40.38	SLC39A7	solute carrier family 39 (zinc transporter) member 7	HGNC: 4927	ENSG00000112473	0	0	0	0
125.63	46.17	SNX13	sorting nexin 13	HGNC: 21335	ENSG00000071189	0	0	0	0
141.97	50.33	SPOP	speckle-type POZ protein	HGNC: 11254	ENSG00000121067	0	0	0	0
126.80	27.84	SUPT5H	suppressor of Ty 5 homolog (S.	HGNC: 11469	ENSG00000196235	0	1	0	1

			cerevisiae)						
152.88	39.53	SYVN1	synovial apoptosis inhibitor 1 synoviolin	HGNC:20738	ENSG00000162298	0	0	0	0
126.42	38.82	TAPBP	TAP binding protein (tapasin)	HGNC:11566	ENSG00000112493	0	0	0	0
135.95	50.82	TBC1D5	TBC1 domain family member 5	HGNC:19166	ENSG00000131374	0	0	0	0
127.46	83.13	TCL1A	T-cell leukemia/lymphoma 1A	HGNC:11648	ENSG00000100721	0	0	0	0
125.09	51.76	TECTB	tectorin beta	HGNC:11721	ENSG00000119913	0	0	0	0
122.89	43.89	TERF2	telomeric repeat binding factor 2	HGNC:11729	ENSG00000132604	0	1	0	1
186.42	63.90	THAP1	THAP domain conta	HGNC:20856	ENSG00000131931	0	0	1	1

			ining apoptosis associated protein 1						
176.24	76.67	THAP6	THAP domain containing 6	HGNC:23189	ENSG00000174796	0	0	0	0
134.64	28.46	TMC2	transmembrane channel-like 2	HGNC:16527	ENSG00000149488	0	0	0	0
123.82	34.32	TMEM132E	transmembrane protein 132E	HGNC:26991	ENSG00000181291	0	0	0	0
172.11	33.28	TNFRSF21	tumor necrosis factor receptor superfamily member 21	HGNC:13469	ENSG00000146072	0	0	0	0
124.16	28.74	TNFRSF8	tumor necrosis factor	HGNC:11923	ENSG00000120949	0	0	0	0

			receptor superfamily member 8						
130.51	11.01	TRPM3	transient receptor potential cation channel subfamily M member 3	HGNC:17992	ENSG00000083067	0	0	0	0
132.58	46.95	TSGA10	testis specific 10	HGNC:14927	ENSG00000135951	0	0	0	0
137.27	61.92	TUFT1	tuftelin 1	HGNC:12422	ENSG00000143367	0	0	0	0
126.84	29.97	UBA52	ubiquitin A-52 residue ribosomal protein fusion product 1	HGNC:12458	ENSG00000221983	0	0	0	0
125.81	26.47	UBE4B	ubiquitination	HGNC:12500	ENSG00000130939	0	0	0	0

			factor E4B						
160.66	47.27	UPK3B	uroplakin 3B	HGNC:21444	ENSG00000243566	0	0	0	0
155.75	42.15	WNT4	wingless-type MMTV integration site family member 4	HGNC:12783	ENSG00000162552	0	0	0	0
125.33	48.44	XCL2	chemokine (C motif) ligand 2	HGNC:10646	ENSG00000143185	0	0	0	0
153.70	64.25	ZNF658	zinc finger protein 658	HGNC:25226	ENSG00000196409	0	0	0	0
137.29	42.07	ZNF684	zinc finger protein 684	HGNC:28418	ENSG00000117010	0	0	0	0
140.03	51.40	ZNF76	zinc finger protein 76	HGNC:13149	ENSG00000065029	0	0	1	1
123.78	45.01	ZSCAN26	zinc finger and SCAN dom	HGNC:12978		0	0	0	0

			ain containing 26						
152.51	41.94	ZZZ3	zinc finger ZZ-type containing 3	HGNC: 24523	ENSG00000036549	0	0	1	1

APPENDIX III

Results from three replicates of the oligonucleotide deconvolution validation screen showing percent effect relative to the siGFP and siUTP4 controls as well as viability relative to siGFP

HGNC Symbol	Description	HGNC ID	Ensemble ID	% VIABILITY Repl 1	% VIABILITY Repl 2	% VIABILITY Repl 3	% EFFECT Repl 1	% EFFECT Repl 2	% EFFECT Repl 3	Mean % effect	Validated? (50% effect cutoff) 1 = Yes, 0 = No
AAMP	angio-associated, migratory cell protein	HGNC:18	ENSG00000127837	77.02	38.55	60.42	43.23	61.37	47.55	50.71	1
AAMP	angio-associated, migratory cell protein	HGNC:18	ENSG00000127837	70.84	57.57	64.32	44.99	71.86	77.76	64.87	1
AAMP	angio-associated, migratory cell protein	HGNC:18	ENSG00000127837	178.22	125.15	102.50	28.57	35.31	31.18	31.69	0
AAMP	angio-associated, migratory cell protein	HGNC:18	ENSG00000127837	137.74	134.54	120.59	21.21	10.93	32.16	21.43	0
ANLN	anillin, actin binding protein	HGNC:14082	ENSG00000011426	50.46	52.38	55.15	92.26	63.72	77.23	77.74	1
ANLN	anillin, actin	HGNC:14082	ENSG00000011426	75.61	73.01	68.34	54.09	37.38	30.87	40.78	0

	binding protein										
ANLN	anillin, actin binding protein	HGNC: 14082	ENSG00000011426	103.44	95.99	91.07	-2.83	-0.11	10.27	2.44	0
ANLN	anillin, actin binding protein	HGNC: 14082	ENSG00000011426	49.19	53.62	31.66	153.95	151.02	151.03	152.00	1
ARHGAP28	Rho GTPase activating protein 28	HGNC: 25509	ENSG00000088756	98.24	84.63	53.51	3.64	8.83	14.33	8.93	0
ARHGAP28	Rho GTPase activating protein 28	HGNC: 25509	ENSG00000088756	57.62	50.53	57.03	272.62	240.16	270.85	261.21	1
ARHGAP28	Rho GTPase activating protein 28	HGNC: 25509	ENSG00000088756	10.82	7.54	10.17	58.51	79.76	68.11	68.79	1
ARHGAP28	Rho GTPase activating protein 28	HGNC: 25509	ENSG00000088756	59.17	77.59	63.81	34.36	60.02	77.77	57.38	1
ARL11	ADP-ribosylation factor-like 11	HGNC: 24046	ENSG00000152213	169.08	36.45	40.07	1.16	29.17	-3.56	8.92	0
ARL11	ADP-ribosylation factor-like 11	HGNC: 24046	ENSG00000152213	59.73	68.07	50.87	200.68	158.86	215.92	191.82	1
ARL11	ADP-ribosylation	HGNC: 24046	ENSG00000152213	149.54	124.90	93.84	49.14	44.74	34.57	42.82	0

	factor-like 11										
ARL11	ADP-ribosylation factor-like 11	HGNC:24046	ENSG00000152213	34.15	60.04	50.62	86.11	79.32	90.97	85.47	1
ARMC2	armadillo repeat containing 2	HGNC:23045	ENSG00000118690	37.39	8.40	5.53	91.60	337.28	381.93	270.27	1
ARMC2	armadillo repeat containing 2	HGNC:23045	ENSG00000118690	123.96	14.70	35.80	13.89	131.56	70.12	71.86	1
ARMC2	armadillo repeat containing 2	HGNC:23045	ENSG00000118690	84.61	59.42	39.19	46.07	51.72	74.21	57.33	1
ARMC2	armadillo repeat containing 2	HGNC:23045	ENSG00000118690	12.79	14.08	16.96	195.41	201.73	180.20	192.45	1
CCDC22	coiled-coil domain containing 22	HGNC:28909	ENSG00000101997	106.68	83.02	82.53	27.30	25.69	11.84	21.61	0
CCDC22	coiled-coil domain containing 22	HGNC:28909	ENSG00000101997	28.11	18.04	20.10	223.46	204.42	216.11	214.66	1
CCDC22	coiled-coil domain containing 22	HGNC:28909	ENSG00000101997	290.93	254.38	250.98	26.34	27.63	29.49	27.82	0
CCDC22	coiled-coil domain containing 22	HGNC:28909	ENSG00000101997	123.82	101.43	103.13	22.97	20.76	15.50	19.74	0

CHAF1A	chromatin assembly factor 1, subunit A (p150)	HGNC: 1910	ENSG00000167670	50.74	64.24	72.23	95.63	139.61	156.71	130.65	1
CHAF1A	chromatin assembly factor 1, subunit A (p150)	HGNC: 1910	ENSG00000167670	96.28	79.81	51.50	17.68	47.62	16.80	27.36	0
CHAF1A	chromatin assembly factor 1, subunit A (p150)	HGNC: 1910	ENSG00000167670	130.29	103.90	113.31	-16.27	-17.85	-13.51	-15.88	0
CHAF1A	chromatin assembly factor 1, subunit A (p150)	HGNC: 1910	ENSG00000167670	74.35	59.05	63.06	124.29	129.92	144.87	133.03	1
CHN2	chimerin 2	HGNC: 1944	ENSG00000106069	68.59	55.47	60.80	47.93	55.09	62.41	55.14	1
CHN2	chimerin 2	HGNC: 1944	ENSG00000106069	183.98	158.88	171.47	18.69	11.16	11.39	13.75	0
CHN2	chimerin 2	HGNC: 1944	ENSG00000106069	321.71	260.80	288.79	3.93	15.43	10.33	9.90	0
CHN2	chimerin 2	HGNC: 1944	ENSG00000106069	39.35	39.16	37.18	132.67	143.02	126.37	134.02	1
CRK	v-crk sarcoma virus	HGNC: 2362	ENSG00000167193	44.69	32.12	33.54	152.07	96.62	113.64	120.78	1

	CT10 oncogene homolog (avian)										
CRK	v-crk sarcoma virus CT10 oncogene homolog (avian)	HGNC: 2362	ENSG00000167193	57.77	59.55	65.82	196.95	195.32	216.71	202.99	1
CRK	v-crk sarcoma virus CT10 oncogene homolog (avian)	HGNC: 2362	ENSG00000167193	105.27	90.80	74.87	64.06	73.20	85.01	74.09	1
CRK	v-crk sarcoma virus CT10 oncogene homolog (avian)	HGNC: 2362	ENSG00000167193	53.55	61.52	52.76	182.09	124.09	170.83	159.01	1
DDX56	DEAD (Asp-Glu-Ala-Asp) box helicase 56	HGNC: 18193	ENSG00000136271	36.81	60.40	51.43	64.19	34.88	31.68	43.58	0
DDX56	DEAD (Asp-Glu-Ala-Asp) box helicase 56	HGNC: 18193	ENSG00000136271	110.86	102.32	114.04	15.71	6.01	-0.10	7.21	0

DDX56	DEAD (Asp-Glu-Ala-Asp) box helicase 56	HGNC: 18193	ENSG00000136271	58.32	45.38	77.09	90.47	74.41	55.62	73.50	1
DDX56	DEAD (Asp-Glu-Ala-Asp) box helicase 56	HGNC: 18193	ENSG00000136271	111.84	96.59	118.95	63.37	76.30	58.47	66.05	1
EPSTI1	epithelial stromal interaction 1 (breast)	HGNC: 16465	ENSG00000133106	97.96	66.10	45.10	66.47	68.62	52.53	62.54	1
EPSTI1	epithelial stromal interaction 1 (breast)	HGNC: 16465	ENSG00000133106	250.04	212.62	136.17	0.00	18.08	13.68	10.59	0
EPSTI1	epithelial stromal interaction 1 (breast)	HGNC: 16465	ENSG00000133106	70.84	44.23	57.53	25.29	83.63	53.48	54.13	1
EPSTI1	epithelial stromal interaction 1 (breast)	HGNC: 16465	ENSG00000133106	26.56	20.51	26.38	89.32	59.45	43.74	64.17	1
FBXW8	F-box and WD repeat domain containing 8	HGNC: 13597	ENSG00000174989	95.15	37.06	22.86	36.75	31.04	57.99	41.93	0
FBXW8	F-box and WD repeat	HGNC: 13597	ENSG00000174989	65.78	49.42	36.68	21.12	50.79	79.22	50.38	1

	domain containing 8										
FBXW8	F-box and WD repeat domain containing 8	HGNC:13597	ENSG00000174989	17.71	20.14	16.08	128.73	154.50	142.06	141.76	1
FBXW8	F-box and WD repeat domain containing 8	HGNC:13597	ENSG00000174989	87.98	54.24	68.08	136.97	96.93	79.03	104.31	1
GRB2	growth factor receptor-bound protein 2	HGNC:4566	ENSG00000177885	26.84	3.09	6.28	74.92	177.10	130.68	127.57	1
GRB2	growth factor receptor-bound protein 2	HGNC:4566	ENSG00000177885	71.68	47.19	18.34	32.25	32.71	16.01	26.99	0
GRB2	growth factor receptor-bound protein 2	HGNC:4566	ENSG00000177885	60.30	45.83	48.36	43.30	86.62	107.88	79.27	1
GRB2	growth factor receptor-bound protein 2	HGNC:4566	ENSG00000177885	18.55	27.67	28.77	68.19	92.23	46.94	69.12	1
HIST1H2AL	histone cluster 1, H2al	HGNC:4730	ENSG00000198374	45.00	59.53	74.24	118.83	127.63	109.27	118.58	1
HIST1H2AL	histone cluster 1, H2al	HGNC:4730	ENSG00000198374	134.12	156.01	161.60	68.63	61.22	47.08	58.98	1

HIST1H2AL	histone cluster 1, H2al	HGNC:4730	ENSG00000198374	71.76	57.15	60.56	25.36	44.07	74.05	47.83	0
HIST1H2AL	histone cluster 1, H2al	HGNC:4730	ENSG00000198374	42.16	46.67	44.71	38.09	38.16	51.08	42.44	0
HIST1H2BO	histone cluster 1, H2bo	HGNC:4758	ENSG00000196331	15.60	30.64	26.38	187.94	160.80	186.27	178.34	1
HIST1H2BO	histone cluster 1, H2bo	HGNC:4758	ENSG00000196331	105.83	87.22	108.41	24.58	22.65	12.94	20.06	0
HIST1H2BO	histone cluster 1, H2bo	HGNC:4758	ENSG00000196331	53.27	32.62	12.56	78.49	77.65	30.91	62.35	1
HIST1H2BO	histone cluster 1, H2bo	HGNC:4758	ENSG00000196331	1.97	8.90	6.91	345.46	378.56	313.90	345.97	1
IQSEC3	IQ motif and Sec7 domain 3	HGNC:29193	ENSG00000120645	158.54	59.05	62.56	19.17	11.12	7.19	12.49	0
IQSEC3	IQ motif and Sec7 domain 3	HGNC:29193	ENSG00000120645	113.42	58.31	82.66	23.64	14.58	17.05	18.42	0
IQSEC3	IQ motif and Sec7 domain 3	HGNC:29193	ENSG00000120645	57.91	47.56	32.79	284.34	260.95	270.29	271.86	1
IQSEC3	IQ motif and Sec7 domain 3	HGNC:29193	ENSG00000120645	78.00	64.00	59.04	75.22	67.79	58.29	67.10	1
KAT5	K(lysine) acetyltransferase 5	HGNC:5275	ENSG00000172977	71.68	74.00	76.63	59.51	70.72	54.95	61.73	1
KAT5	K(lysine) acetyltr	HGNC:5275	ENSG00000172977	43.29	42.25	31.15	121.38	124.60	140.18	128.72	1

	ansferase 5										
KAT5	K(lysine) acetyltransferase 5	HGNC: 5275	ENSG00000172977	47.93	52.38	44.97	66.00	63.72	63.96	64.56	1
KAT5	K(lysine) acetyltransferase 5	HGNC: 5275	ENSG00000172977	69.85	70.30	67.21	102.20	70.14	106.81	93.05	1
LIN28A	lin-28 homolog A (C. elegans)	HGNC: 15986	ENSG00000131914	31.76	32.12	23.49	188.22	225.60	169.85	194.55	1
LIN28A	lin-28 homolog A (C. elegans)	HGNC: 15986	ENSG00000131914	134.65	106.49	133.28	167.32	141.95	131.64	146.97	1
LIN28A	lin-28 homolog A (C. elegans)	HGNC: 15986	ENSG00000131914	175.40	131.82	130.14	18.60	19.05	34.40	24.02	0
LIN28A	lin-28 homolog A (C. elegans)	HGNC: 15986	ENSG00000131914	75.05	71.90	76.88	32.52	37.00	15.49	28.34	0
METTL21D	methyltransferase like 21D	HGNC: 20352	ENSG00000100483	165.85	18.04	51.50	-10.15	11.48	-13.62	-4.10	0
METTL21D	methyltransferase like 21D	HGNC: 20352	ENSG00000100483	119.89	110.70	95.09	18.79	-2.59	12.38	9.53	0
METTL21D	methyltransferase like 21D	HGNC: 20352	ENSG00000100483	115.53	107.36	98.99	71.92	49.95	84.67	68.85	1
METTL21D	methyltransferase like 21D	HGNC: 20352	ENSG00000100483	63.11	63.75	87.81	62.61	69.93	33.88	55.47	1
NEXN	nexilin (F actin	HGNC: 29557	ENSG00000162614	56.36	49.17	47.48	33.35	78.20	64.59	58.71	1

	binding protein)										
NEXN	nexilin (F actin binding protein)	HGNC: 29557	ENSG00000162614	71.54	75.73	69.21	16.80	62.41	-3.64	25.19	0
NEXN	nexilin (F actin binding protein)	HGNC: 29557	ENSG00000162614	39.63	29.28	46.60	92.67	83.22	92.28	89.39	1
NEXN	nexilin (F actin binding protein)	HGNC: 29557	ENSG00000162614	138.72	141.95	126.50	7.16	-4.83	-0.36	0.66	0
NMT2	N-myristoyltransferase 2	HGNC: 7858	ENSG00000152465	66.34	62.76	45.85	122.60	61.06	35.83	73.16	1
NMT2	N-myristoyltransferase 2	HGNC: 7858	ENSG00000152465	54.53	52.88	70.85	122.98	134.85	147.45	135.10	1
NMT2	N-myristoyltransferase 2	HGNC: 7858	ENSG00000152465	284.19	225.47	241.31	18.20	19.57	13.16	16.98	0
NMT2	N-myristoyltransferase 2	HGNC: 7858	ENSG00000152465	51.72	41.76	87.68	175.26	90.80	149.07	138.38	1
NODAL	nodal growth differentiation factor	HGNC: 7865	ENSG00000156574	94.59	70.91	101.25	8.46	10.78	13.60	10.95	0
NODAL	nodal growth differentiation factor	HGNC: 7865	ENSG00000156574	107.94	72.15	94.46	48.44	83.45	34.68	55.52	1
NODAL	nodal growth differentiation factor	HGNC: 7865	ENSG00000156574	15.18	13.47	8.67	180.61	171.69	124.61	158.97	1

NODAL	nodal growth differentiation factor	HGNC:7865	ENSG00000156574	45.82	55.35	45.10	29.97	79.25	74.77	61.33	1
NOMO1	NODAL modulator 1	HGNC:30060	ENSG00000103512	163.74	127.37	144.21	26.09	33.17	21.49	26.92	0
NOMO1	NODAL modulator 1	HGNC:30060	ENSG00000103512	468.87	393.98	364.04	36.46	36.79	41.12	38.12	0
NOMO1	NODAL modulator 1	HGNC:30060	ENSG00000103512	136.47	114.03	138.56	29.49	16.25	11.24	18.99	0
NOMO1	NODAL modulator 1	HGNC:30060	ENSG00000103512	105.83	107.61	96.85	49.63	35.89	48.79	44.77	0
NTN3	netrin 3	HGNC:8030	ENSG00000162068	103.30	70.67	80.90	10.74	0.79	10.74	7.42	0
NTN3	netrin 3	HGNC:8030	ENSG00000162068	45.54	24.34	38.06	166.82	155.40	199.70	173.97	1
NTN3	netrin 3	HGNC:8030	ENSG00000162068	126.91	132.56	148.60	9.23	8.32	16.03	11.19	0
NTN3	netrin 3	HGNC:8030	ENSG00000162068	22.91	29.16	28.39	183.80	201.20	204.94	196.65	1
NUDT13	nudix (nucleoside diphosphate linked moiety X)-type motif 13	HGNC:18827	ENSG00000166321	51.88	54.56	53.71	62.72	61.50	49.71	57.98	1
NUDT13	nudix (nucleoside diphosphate linked moiety X)-type	HGNC:18827	ENSG00000166321	41.39	59.32	49.04	40.70	52.56	16.03	36.43	0

	motif 13										
NUDT13	nudix (nucleoside diphosphate linked moiety X)-type motif 13	HGNC: 18827	ENSG00000166321	29.93	36.09	45.73	25.58	21.61	65.11	37.43	0
NUDT13	nudix (nucleoside diphosphate linked moiety X)-type motif 13	HGNC: 18827	ENSG00000166321	83.12	76.93	77.66	32.99	11.41	19.40	21.27	0
NUMA1	nuclear mitotic apparatus protein 1	HGNC: 8059	ENSG00000137497	14.62	5.07	4.77	36.50	235.12	239.91	170.51	1
NUMA1	nuclear mitotic apparatus protein 1	HGNC: 8059	ENSG00000137497	58.75	9.51	16.08	55.12	43.78	1.76	33.55	0
NUMA1	nuclear mitotic apparatus protein 1	HGNC: 8059	ENSG00000137497	7.03	12.60	2.26	292.97	41.39	145.09	159.82	1
NUMA1	nuclear mitotic apparatus protein 1	HGNC: 8059	ENSG00000137497	51.58	33.48	29.65	320.62	277.61	308.33	302.19	1

PDE12	phosphodiesterase 12	HGNC: 25386	ENSG00000174840	47.65	53.99	105.90	104.73	121.07	95.48	107.10	1
PDE12	phosphodiesterase 12	HGNC: 25386	ENSG00000174840	51.16	44.48	36.55	52.19	91.65	79.66	74.50	1
PDE12	phosphodiesterase 12	HGNC: 25386	ENSG00000174840	120.31	96.36	120.97	41.77	33.56	45.89	40.41	0
PDE12	phosphodiesterase 12	HGNC: 25386	ENSG00000174840	142.38	102.17	113.43	31.31	32.77	27.33	30.47	0
PPP1R3B	protein phosphatase 1, regulatory subunit 3B	HGNC: 14942	ENSG00000173281	11.52	15.32	16.83	191.49	211.29	181.91	194.89	1
PPP1R3B	protein phosphatase 1, regulatory subunit 3B	HGNC: 14942	ENSG00000173281	54.95	57.94	58.66	41.69	37.80	28.00	35.83	0
PPP1R3B	protein phosphatase 1, regulatory subunit 3B	HGNC: 14942	ENSG00000173281	74.49	64.74	42.71	63.08	78.53	68.47	70.03	1
PPP1R3B	protein phosphatase 1, regulatory subunit 3B	HGNC: 14942	ENSG00000173281	317.22	194.71	254.63	11.88	16.39	15.08	14.45	0
PRL	prolactin	HGNC: 9445	ENSG00000172179	29.66	37.43	14.95	44.76	14.13	43.32	34.07	0
PRL	prolactin	HGNC: 9445	ENSG00000172179	131.13	106.37	88.43	61.24	34.67	59.51	51.81	1

PRL	prolactin	HGNC:9445	ENSG00000172179	70.56	54.48	48.86	71.06	53.73	30.60	51.80	1
PRL	prolactin	HGNC:9445	ENSG00000172179	151.93	140.84	142.83	10.00	12.69	8.57	10.42	0
RAB38	RAB38, member RAS oncogene family	HGNC:9776	ENSG00000123892	31.20	31.50	37.56	116.37	150.10	124.61	130.36	1
RAB38	RAB38, member RAS oncogene family	HGNC:9776	ENSG00000123892	69.01	74.13	70.97	42.31	77.99	56.16	58.82	1
RAB38	RAB38, member RAS oncogene family	HGNC:9776	ENSG00000123892	67.74	51.76	43.21	212.87	204.72	201.97	206.52	1
RAB38	RAB38, member RAS oncogene family	HGNC:9776	ENSG00000123892	186.51	175.56	173.98	-3.50	24.47	31.77	17.58	0
RASAL2	RAS protein activator like 2	HGNC:9874	ENSG00000075391	80.82	57.20	65.95	124.59	131.48	104.08	120.05	1
RASAL2	RAS protein activator like 2	HGNC:9874	ENSG00000075391	59.73	83.15	59.29	34.78	34.23	34.59	34.54	0
RASAL2	RAS protein activator like 2	HGNC:9874	ENSG00000075391	82.08	81.29	81.40	60.77	50.82	59.64	57.08	1
RASAL2	RAS protein activator like 2	HGNC:9874	ENSG00000075391	85.73	70.79	60.55	89.64	89.67	80.47	86.59	1

RBM28	RNA binding motif protein 28	HGNC: 21863	ENSG00000106344	165.58	109.12	149.85	36.67	37.88	27.44	34.00	0
RBM28	RNA binding motif protein 28	HGNC: 21863	ENSG00000106344	34.73	33.39	25.43	81.98	63.15	57.18	67.44	1
RBM28	RNA binding motif protein 28	HGNC: 21863	ENSG00000106344	243.45	221.16	228.77	61.40	44.60	54.68	53.56	1
RBM28	RNA binding motif protein 28	HGNC: 21863	ENSG00000106344	70.34	57.48	77.09	17.12	7.32	4.06	9.50	0
RBM43	RNA binding motif protein 43	HGNC: 24790	ENSG00000184898	102.18	117.00	87.05	73.49	42.75	72.75	63.00	1
RBM43	RNA binding motif protein 43	HGNC: 24790	ENSG00000184898	142.09	121.81	92.71	16.24	15.99	23.42	18.55	0
RBM43	RNA binding motif protein 43	HGNC: 24790	ENSG00000184898	71.12	88.58	58.03	-2.46	24.22	28.84	16.86	0
RBM43	RNA binding motif protein 43	HGNC: 24790	ENSG00000184898	52.85	35.70	49.49	185.11	152.78	176.47	171.45	1
RGN	regucalcin (senescence marker protein-30)	HGNC: 9989	ENSG00000130988	18.69	5.93	7.41	86.00	344.40	424.59	285.00	1

RGN	regucalcin (senescence marker protein-30)	HGNC: 9989	ENSG00000130988	80.96	57.08	47.86	49.30	19.58	23.11	30.66	0
RGN	regucalcin (senescence marker protein-30)	HGNC: 9989	ENSG00000130988	69.43	34.96	44.85	117.16	147.34	157.90	140.80	1
RGN	regucalcin (senescence marker protein-30)	HGNC: 9989	ENSG00000130988	211.52	147.39	167.45	1.86	27.32	6.85	12.01	0
RPS23	ribosomal protein S23	HGNC: 10410	ENSG00000186468	50.57	38.03	57.71	64.27	105.35	106.19	91.94	1
RPS23	ribosomal protein S23	HGNC: 10410	ENSG00000186468	36.04	33.49	32.27	84.80	119.89	110.27	104.99	1
RPS23	ribosomal protein S23	HGNC: 10410	ENSG00000186468	37.24	31.12	51.43	47.87	62.90	46.76	52.51	1
RPS23	ribosomal protein S23	HGNC: 10410	ENSG00000186468	80.39	29.17	41.85	119.50	111.83	114.68	115.33	1
RSL24D1	ribosomal L24 domain containing 1	HGNC: 18479	ENSG00000137876	40.96	128.25	88.50	24.00	14.05	34.42	24.16	0
RSL24D1	ribosomal L24 domain containing 1	HGNC: 18479	ENSG00000137876	40.63	35.01	43.34	71.15	54.97	42.54	56.22	1

RSL24D1	ribosomal L24 domain containing 1	HGNC:18479	ENSG00000137876	50.79	65.58	62.04	98.15	89.93	91.82	93.30	1
RSL24D1	ribosomal L24 domain containing 1	HGNC:18479	ENSG00000137876	46.64	24.31	48.47	107.89	56.39	105.60	89.96	1
SAMD15	sterile alpha motif domain containing 15	HGNC:18631	ENSG00000100583	64.65	46.95	43.21	30.90	31.90	27.95	30.25	0
SAMD15	sterile alpha motif domain containing 15	HGNC:18631	ENSG00000100583	52.57	49.79	47.86	19.07	24.61	17.87	20.52	0
SAMD15	sterile alpha motif domain containing 15	HGNC:18631	ENSG00000100583	39.21	42.13	52.51	209.82	165.72	196.94	190.83	1
SAMD15	sterile alpha motif domain containing 15	HGNC:18631	ENSG00000100583	55.94	34.96	62.43	83.84	93.62	136.79	104.75	1
SLC39A7	solute carrier family 39 (zinc transporter), member 7	HGNC:4927	ENSG00000112473	91.22	86.60	57.03	39.50	36.47	34.60	36.86	0
SLC39A7	solute carrier family 39 (zinc transpo	HGNC:4927	ENSG00000112473	62.97	66.34	86.93	-4.76	7.44	21.74	8.14	0

	rter), member 7										
SLC39A7	solute carrier family 39 (zinc transporter), member 7	HGNC: 4927	ENSG00000112473	45.26	42.75	53.64	228.27	219.64	212.79	220.23	1
SLC39A7	solute carrier family 39 (zinc transporter), member 7	HGNC: 4927	ENSG00000112473	67.88	57.70	36.30	64.31	105.14	121.98	97.14	1
SUPT5H	suppressor of Ty 5 homolog (S. cerevisiae)	HGNC: 11469	ENSG00000196235	61.00	40.65	32.91	64.03	71.21	71.05	68.76	1
SUPT5H	suppressor of Ty 5 homolog (S. cerevisiae)	HGNC: 11469	ENSG00000196235	78.14	66.71	61.68	89.29	51.08	105.53	81.97	1
SUPT5H	suppressor of Ty 5 homolog (S. cerevisiae)	HGNC: 11469	ENSG00000196235	42.59	32.86	36.30	84.80	128.89	113.35	109.01	1
SUPT5H	suppressor of Ty 5 homolog (S. cerevisiae)	HGNC: 11469	ENSG00000196235	30.08	45.96	25.12	115.43	142.77	98.26	118.82	1

TERF2	telomeric repeat binding factor 2	HGNC: 11729	ENSG00000132604	18.27	18.41	23.62	54.64	136.53	81.11	90.76	1
TERF2	telomeric repeat binding factor 2	HGNC: 11729	ENSG00000132604	179.90	149.24	173.35	11.20	11.34	6.76	9.77	0
TERF2	telomeric repeat binding factor 2	HGNC: 11729	ENSG00000132604	101.34	74.74	70.47	23.52	26.78	2.67	17.65	0
TERF2	telomeric repeat binding factor 2	HGNC: 11729	ENSG00000132604	41.88	42.00	89.56	73.64	78.21	105.02	85.62	1
THAP1	THAP domain containing, apoptosis associated protein 1	HGNC: 20856	ENSG00000131931	65.21	52.40	51.89	59.04	109.41	67.55	78.67	1
THAP1	THAP domain containing, apoptosis associated protein 1	HGNC: 20856	ENSG00000131931	59.53	57.91	57.71	16.45	26.98	29.75	24.39	0
THAP1	THAP domain containing, apoptosis associat	HGNC: 20856	ENSG00000131931	84.32	57.26	70.94	81.24	74.96	89.91	82.04	1

	ed protein 1										
THAP1	THAP domain containing, apoptosis associated protein 1	HGNC: 20856	ENSG00000131931	154.00	120.90	132.52	64.10	48.78	48.80	53.89	1
UBA52	ubiquitin A-52 residue ribosomal protein fusion product 1	HGNC: 12458	ENSG00000221983	60.30	14.70	11.18	89.60	90.23	102.43	94.09	1
UBA52	ubiquitin A-52 residue ribosomal protein fusion product 1	HGNC: 12458	ENSG00000221983	54.11	65.23	71.73	58.51	39.54	62.92	53.66	1
UBA52	ubiquitin A-52 residue ribosomal protein fusion product 1	HGNC: 12458	ENSG00000221983	58.05	45.22	36.05	166.93	148.17	100.58	138.56	1
UBA52	ubiquitin A-52 residue ribosomal protein fusion	HGNC: 12458	ENSG00000221983	114.41	90.56	105.39	41.35	33.24	35.52	36.70	0

	product 1										
UTP4	UTP4, small subunit processome component (formerly Cirhin)	HGNC: 1983	ENSG00000141076	37.90	69.47	53.26	78.88	54.36	54.06	62.43	1
UTP4	UTP4, small subunit processome component (formerly Cirhin)	HGNC: 1983	ENSG00000141076	80.61	109.12	118.95	56.13	55.80	60.10	57.34	1
UTP4	UTP4, small subunit processome component (formerly Cirhin)	HGNC: 1983	ENSG00000141076	159.68	196.21	213.38	68.89	59.47	83.22	70.53	1
UTP4	UTP4, small subunit processome component (formerly Cirhin)	HGNC: 1983	ENSG00000141076	135.00	114.85	113.82	77.83	60.74	83.44	74.00	1
ZNF76	zinc finger protein 76	HGNC: 13149	ENSG00000065029	90.51	74.62	66.83	35.52	15.78	12.04	21.12	0
ZNF76	zinc finger protein 76	HGNC: 13149	ENSG00000065029	28.53	41.39	30.27	111.88	134.66	143.60	130.04	1

ZNF76	zinc finger protein 76	HGNC: 13149	ENSG00000065029	38.93	32.86	21.73	166.85	142.34	75.09	128.09	1
ZNF76	zinc finger protein 76	HGNC: 13149	ENSG00000065029	444.55	391.14	388.28	7.77	13.47	4.03	8.42	0