Abstract

Comprehensive Identification and Characterization of Cajal Body Components

Dahyana Arias Escayola

2022

Each cell's nucleus is a safe harbor for the genome (DNA), ensuring the integrity of all of the genes that contribute to development and health of an organism. What's more, the 3-dimensional (3D) organization of the cell nucleus regulates how genes are expressed to give each cell its unique characteristics. In addition to chromosomes, the nucleus contains nuclear bodies, which are functionally distinct yet membraneless compartments that likely form through liquid-liquid phase separation (LLPS). Nuclear bodies concentrate specific proteins and RNAs on distinct chromosomal DNA regions to regulate gene expression. One of these nuclear bodies, the Cajal Body (CB), is implicated as a site of synthesis and assembly for components of the spliceosome, a molecular machine responsible for maturation of nearly every human messenger RNA (mRNA). CBs are necessary for normal development, since vertebrate embryos depleted of the CB scaffolding protein coilin do not properly process mRNA and cannot support growth. Remarkably, the composition of nuclear bodies is not well defined. This lack of knowledge is the most important current obstacle to understanding how nuclear bodies govern gene expression. This dissertation aims to discover all of the components of the CB systematically, and to determine which are necessary for CB assembly.

I have adapted and applied state-of-the-art proximity biotinylation techniques (APEX2) to obtain a comprehensive list of CB proteins by mass spectrometry. I identified 70 new CB proteins, nearly doubling the number of known constituents. Of these, IRF2BP1 is the first DNA binding protein to be identified in CBs. I have performed a screen depleting each CB protein individually and analyzed changes to CB number and shape. As a result, I found that 46 CB proteins are necessary for proper CB assembly. The siRNA screen

revealed three different phenotypes of improper CB assembly, 1) decreased number of CBs per nucleus, 2) increased number of CBs per nucleus and 3) relocalization of coilin to nucleoli. Further analysis of the increased number phenotype revealed components of the 60S large ribosomal subunit (RPL proteins) as regulators of CB assembly. This is the first study to demonstrate regulation of CBs by ribosomal proteins.

Next, I characterized the increased coilin foci upon RPL knockdown and found effects to CB morphology and composition. CBs after RPL knockdown showed a reduction in snRNP proteins and lost their distinct substructure. Cajal bodies are normally made up of two subunits, a coilin containing domain and an SMN containing domain. Upon RPL KDs, this subdomain structure was lost, instead forming one domain with both coilin and SMN. Cajal bodies are known to form on actively transcribing snRNA and histone gene loci. I performed chromatin immunoprecipitation (ChIP) with antibodies against coilin and RNA Polymerase II (Pol II) in RPL knockdowns and discovered that coilin was no longer associated to gene loci and binding of Pol II along gene bodies was reduced. These results demonstrate that ribosomal proteins regulate the assembly and morphology of CBs.

Finally, I combined proximity biotinylation with ChIP in a new method I have termed APEX-ChIP. The goal of APEX-ChIP is to understand how nuclear bodies interact with chromatin, by biotinylating molecular constituents surrounding the marker protein and thereby enhancing ChIP signals. As proof-of-principle, I show that APEX-ChIP is a viable method and use it on a well characterized nuclear body, the nucleolus. I performed nucleophosmin APEX-ChIP and show that I can detect DNA sequences corresponding to nucleolar organizing regions, while also revealing nucleophosmin at gene promoter regions.

These discoveries represent the first unbiased and comprehensive CB components list that is functionally characterized for CB assembly. This is a crucial step in understanding the role that CBs and LLPS play in regulating the expression and 3D organization of genomes.

Comprehensive Identification and Characterization of Cajal Body Components

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

by
Dahyana Arias Escayola

Dissertation Director: Karla Neugebauer

May 2022

Table of Contents

List of Figures

List of Tables

Acknowledgements

It feels impossible to summarize the number of “thank yous” owed, but I will do my best. When I started my PhD I had no idea how important forming a supportive community would be. I’m lucky to have formed strong bonds with people both in MB&B and in New Haven, but I’m especially lucky that all of these people are strong, supportive, and kind. Thank you. To all of you.

Karla, you have taught me that it is worth it to take risks if you are truly invested in the outcome. I’ve enjoyed our many discussions on ambitious experiments, even when those experiments turned out to be more difficult in practice than in theory. You gave me the opportunity to be a curious scientist, and I appreciate the freedom you give the lab in coming up with ideas and forming projects. Thank you for making the lab a great place to work. It’s your energy that pushes us forward.

From the start, the Molecular Biophysics and Biochemistry department has been a welcoming environment. The departmental seminars, events, and retreat were great opportunities to interact with the community and hold great scientific discussions. I want to thank my committee members, Dr. Susan Baserga and Dr. Matt Simon. Each committee meeting that I had pushed my project in a better direction thanks to their helpful feedback.

The Neugebauer Lab has been an incredible place to work, and I am so grateful for my colleagues. It’s reassuring going into work daily knowing that at some point in the day someone is going to make you laugh. Thank you for the shenanigans. I am happy that I can count you among my close friends. To the postdocs that have mentored me throughout my PhD- Tucker Carrocci, David Phizicky, and Jonathan Rodenfels- thank you for taking time out of your day to work through a protocol, come up with new ideas, or just reassure me that whatever I was seeing on a

gel or under the microscope was real and not in my imagination. To the OG graduate students – Tara Alpert, Edward Courchaine, Kirsten Reimer- we did it! I cannot imagine having to go through this alone, but I'm especially lucky to have you as friends both in and out of the lab. To the new(er) graduate students – Sara Gelles-Watnick, Jackson Gordon, Leo Schaeffer- I've had so much fun getting to know you recently, I am so glad that all of you chose our lab. To our lab manager, Korinna, thank you for being *the* person in the lab.

Although they were already listed under the lab thank yous I have to go back and thank them again. Tara and Korinna, I consider you my family (plus little Oli of course). All the movie nights, hikes, wine nights, long walks, spontaneous trips, conversations, laughter, and cries… they were all perfect. I've found lifelong friends in you both.

To my partner, Santi, thank you for being my rock this past year. You have had to deal with the most difficult part of my PhD. Thank you for the dinners, the snacks, the forced outside walks when I was hyperfocused on thesis writing… You are my home, and I am so excited to be moving to Boston and starting a new life with you. I love you.

Most importantly, thank you to my family, the driving force behind everything I do. To my siblings, Carolain and Sebastian. Thank you for everything that you did for me growing up, and now. I will always be the youngest, but I love growing closer to you now that we are all adults. Thank you for bringing Manu and Joha into our lives, and for the example you all provide. To my kiddos… Kevin, Amanda, Justin, Bryan, Eric, Felix, Mia, and Ella… I could write pages about each and every one of you. Your aunt loves you.

And finally, to my mom… Mami, este es tu esfuerzo. Es el amor y el trabajo que pusiste. Todo lo que soy es gracias a ti. Gracias por todo tu apoyo, y por creer en mi tanto que yo también creo. Hemos pasado por tanto juntas, pero aquí esta el resultado. Este trabajo, te lo dedico a ti.

1 Introduction

Portions of this chapter have been modified from a previously published review:

Arias Escayola, D. & Neugebauer, K. M. Dynamics and Function of Nuclear Bodies during Embryogenesis. *Biochemistry (Mosc.)* (2018). doi:10.1021/acs.biochem.7b01262

Cellular compartmentalization is a key component of maintaining efficient biochemical reactions in cells. Two main classes of organelles compartmentalize macromolecules within cells: membrane-bound and membraneless organelles (MLOs). The nucleus, a membrane-bound organelle, contains a cell's genetic material in chromosomes. Chromosomes are highly organized and regulated within the nucleus. Within the interchromatin space the nucleus is sub-compartmentalized by various nuclear MLOs, termed nuclear bodies (NBs), that organize the nucleus and regulate DNA and RNA processing events (Mao et al., 2011). NBs lack a lipid bilayer but maintain distinct boundaries and components. Without a membrane, components of nuclear bodies are free to exchange with their surroundings, making nuclear bodies highly dynamic. This means that while the structure of a nuclear body can remain stable in a cell, its components are rapidly exchanging and able to move in and out of the nuclear body. Nuclear bodies form as a collection of dynamic protein-protein and/or protein-RNA interactions. Proteins in these nuclear bodies contain higher than typical tendency to intrinsic disorder, suggesting they may arise through liquid-liquid phase separation (LLPS). Because nuclear bodies lack membranes, their assembly and morphology are responsive to stress and other biological

signals. For example, nuclear bodies disassemble at mitosis onset and rapidly form again upon exiting mitosis. Additionally, nuclear body assembly and number are tightly regulated during development. Specificity and dynamicity are key characteristics of nuclear bodies that allow for tight regulation of gene expression in the nucleus.

NBs include the nucleolus, the Cajal body (CB), the histone locus body (HLB), speckles, paraspeckles, and PML bodies. Each of these nuclear bodies compartmentalizes distinct macromolecules to enhance the efficiency of biological reactions. Except for PML bodies, all nuclear bodies are RNA-rich and contain subsets of nuclear proteins that reflect different functions in the biogenesis of polyadenylated mRNA, replication-dependent histone mRNA and ribosomes. The function of each of these nuclear bodies is dependent on its components. Comprehensive studies of nuclear body components have been historically difficult to undertake due to the dynamic nature of nuclear bodies. My thesis focuses on comprehensively identifying and characterizing the components of the Cajal Body to better understand its function.

1.1 Discovery of the Cajal Body and its components

CBs were discovered in 1903 by Santiago Ramón y Cajal, who observed a small nuclear structure by silver staining vertebrate neuronal slices (Ramón y Cajal, 1903). He called this nuclear structure an “accessory body” because of its proximity to the nucleolus. His silver staining method treated tissue samples with an aqueous silver nitrate solution that binds protein molecules and can be reduced with hydroquinone or pyrogallic acid to make the silver particles visible under a microscope (Gall, 2000). Cajal characterized the “accessory body” as a 0.5 um structure that varied in number (1-3) in the nuclei of a variety of neuronal cell types. Since the initial discovery of the “accessory body” it has been

rediscovered in the tissues of various organisms and renamed each time. In 1999, Joseph Gall proposed that this nuclear structure be renamed the Cajal Body in honor of its discoverer. In 1969, RNA was detected in CBs using EDTA staining for ribonucleoproteins, establishing the Cajal Body as an RNA-rich nuclear body (Monneron & Bernhard, 1969). While structural characterization of CBs continued, any discovery of the components of CBs stalled. In 1984, anti-Sm labeling in mouse liver tissue and WT HeLa cells established the presence of Sm proteins in CBs (Eliceiri & Ryerse, 1984; Raska et al., 1991). These findings suggested the presence of snRNPs in the CB, key components for understanding CB function.

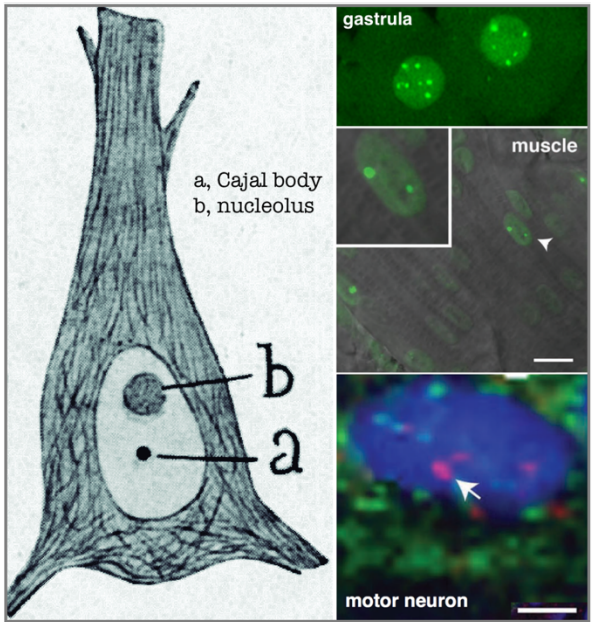


Figure 1.1 Cajal Bodies
CBs in a human pyramidal neuron, drawn by Cajal in 1895 (left). Representative images in various tissues of zebrafish embryos (right); green coilin-YFP tag in gastrula and muscle (10µm scale bar), coilin antibody staining (red) in motorneurons (3 µm scale bar). (Strzelecka, Oates, et al., 2010) Figure courtesy of Karla Neugebauer.

1.1.1 Coilin

Coilin, the main CB scaffolding protein, was initially identified in cajal bodies using human autoimmune sera and immunofluorescent staining of nuclei (Andrade et al., 1991; Raška et al., 1991). Since then, coilin has provided a molecular handle for studying cajal bodies, whether as a marker for localization in immunofluorescent stains or to deplete CBs by depleting coilin. Depletion of coilin disassembles the CB, meaning that it is necessary for the concentration of snRNAs and snRNPs in nuclear bodies. Although there is no known independent "biochemical" function for coilin, it has been shown to bind Sm proteins, RNA, and DNA (Machyna et al., 2015). Loss of coilin and therefore CBs is lethal for zebrafish embryos (Strzelecka, Trowitzsch, et al., 2010). Because viability is rescued through injection of mature human snRNPs, lethality appears to be due to insufficient production of snRNPs needed for splicing zygotic pre-mRNAs expressed in embryos at zygotic genome activation (ZGA). These findings are consistent with the coilin knockout phenotype in mice, which is "semilethal" and characterized by the reduced embryonic viability and fertility of adults (M. P. Walker et al., 2009). Conversely, in *Drosophila* and *Arabidopsis thaliana*, homozygous coilin knockouts disperse CBs but are still viable and fertile (Collier et al., 2006; J.-L. Liu et al., 2009). There is no proven explanation for why coilin seems to be essential in vertebrates and inessential in insects and plants. Coilin's N-terminal domain (NTD) mediates coilin self-interaction (Hebert & Matera, 2000). Coilin constructs lacking an NTD do not localize to Cajal bodies, suggesting that coilin-coilin interactions are necessary for CB formation (Bohmann et al., 1995; Wu et al., 1994). Additionaly, post-translational modifications (PTMs) on coilin also regulate its presence in Cajal bodies (Hebert & Poole, 2017). Coilin is hyperphosphorylated at mitosis when

CBs disassemble (Carmo-Fonseca et al., 1993; Lyon et al., 1997). Cells that lack Cajal Bodies have been found to have similar levels of hyperphosphorylated coilin as mitotic cells (Hearst et al., 2009). Taken together, these data suggest that coilin phosphorylation regulates Cajal body assembly. Coilin contains symmetrically dimethylated arginines along its RG box (a stretch of arginine/glycine repeats) (Hebert et al., 2002). These symmetrically dimethylated arginine (sDMA) modifications are necessary for localization of coilin to CBs; hypomethylated coilin will mislocalize to the nucleolus. Coilin is the canonical CB marker protein and is used throughout this study to label Cajal bodies.

1.2 SMN

Cellular localization of SMN by immunofluorescence in HeLa Cells showed its presence in the cytoplasm, CBs and a separate nuclear body called "Gems" (Q. Liu & Dreyfuss, 1996). SMN forms SMN complex along with gemin proteins; this complex is necessary for snRNP assembly in the cytoplasm (Meister et al., 2002). SMN complex binds Sm proteins and mediates formation of the Sm ring around snRNAs. After maturation of the snRNP protein, the complex is bound by snurportin-1 and the SMN/snRNP complex is transported into the nucleus where snRNPs are released from SMN complex and SMN complex localizes to either CBs or Gems. SMN's function in the nucleus and these nuclear bodies is still unknown. SMN is necessary for CB formation and has been shown to directly interact with coilin through coilin's conserved RG box (Francois-Michel Boisvert et al., 2002; Hebert et al., 2002; Hebert et al., 2001). SMN depletion results in loss of CBs in HeLa cells, suggesting that snRNP biogenesis is necessary for CB formation (Lemm et al., 2006). Recent experiments using super-resolution microscopy have demonstrated that SMN and coilin may form separate sub-compartments within Cajal bodies (Courchaine et

al., 2021; Novotný et al., 2015). More specifically, coilin and SMN only partially overlap in cells, with the larger coilin subunit partially wrapping around a smaller SMN subunit. The interaction between coilin and SMN is dependent on sDMA modifications on coilin, otherwise SMN will localize to gems, but not CBs (Francois-Michel Boisvert et al., 2002). Furthermore, coilin and SMN foci in the nucleus can merge or separate with treatment from different arginine methylation inhibitors, suggesting that DMA modifications regulate structure within the CB (Courchaine et al., 2021). This substructure with SMN and coilin in the CB may indicate one of two things; 1) CBs have substructure and SMN and coilin form separate domains or 2) the presence of SMN in CBs is the result of Gems interacting with CBs.

1.2.1 RNPs and other CB components

In addition to proteins necessary for snRNP assembly, proteins involved in other RNA processing events have also been found to localize to CBs. The Cajal body shares many components with the nucleolus, which is a site of rRNA transcription and ribosome biogenesis (Trinkle-Mulcahy & Sleeman, 2017). There has been documented protein transport between nucleoli and CBs, while HLBs and CBs have been shown to merge in some organisms. Nopp140 is a highly phosphorylated nucleolar protein that is also present in CBs (Isaac et al., 1998). Nopp140 interacts directly with coilin and is necessary for CB formation. Because Nopp140 forms a complex with snoRNPs, another group of nucleolar RNPs that can be found in CBs, it is hypothesized that it functions as a chaperone for snoRNPs between CBs and nucleoli. Some CB components can also be found in HLBs (Machyna et al., 2013b). In some organisms and cell lines, HLBs and coilin overlap and share components. As a result, proteins involved in histone mRNA maturation such as

Lsm11 and NPAT have also been found in CBs. RNA polymerases II and III have previously been identified in Cajal bodies in Xenopus eggs (Morgan et al., 2000; Murphy et al., 2002). Until recently, there was no evidence of RNA polymerase in mammalian CBs. A recent study in HCT116 cells found that RNA polymerase II condensates form near and associate with CBs, suggesting that RNA polymerase II may also be present in mammalian CBs (Imada et al., 2021). It is worth noting that the CB proteins listed here (

Table 1) have been identified in CBs mostly through immunofluorescent staining or GFP-tagged constructs. To date, there is a lack of large-scale studies to identify CB components.

Table 1. List of Known CB Proteins

Function	Gene ID	Protein Name	Reference
Transcription	EAF1	ELL-associated factor 1	(Polak et al., 2003)
	EAF2	ELL-associated factor 2	(Polak et al., 2003)
	ELL	RNA polymerase II elongation factor	(Polak et al., 2003)
	GTF2F1	General transcription factor IIF subunit 1	(Fong et al., 2013)
	ICE1	Little elongation complex subunit 1	(Smith et al., 2011)
	ICE2	Little elongation complex subunit 2	(Smith et al., 2011)
	MED26	Mediator of RNA polymerase II transcription subunit 26	(Fong et al., 2013)
	POLR2A	RPB1	(Morgan et al., 2000)
	POLR2B	RPB2	(Morgan et al., 2000)
	SERBP1	CGI-55	(Lemos & Kobarg, 2006)
	TARDBP	TAR DNA-binding protein 43	(Fong et al., 2013)
	TRIM22	E3 ubiquitin-protein ligase TRIM22	(Sivaramakrishnan et al., 2009)
	ZGPAT	Zinc finger CCCH-type with G patch domain-containing protein	(Fong et al., 2013)
	ZNF277	Zinc finger protein 277	(Fong et al., 2013)

	ZPR1	Zinc finger protein ZPR1	(Gangwani et al., 2005)
pre-mRNA splicing and cleavage	EFTUD2	snu114; 116 kDa U5 small nuclear ribonucleoprotein component	(Schaffert et al., 2004)
	FRG1	Protein FRG1	(van Koningsbruggen et al., 2004)
	NHP2L1	NHP2-like protein 1;SNU13	(Verheggen, 2002)
	PRPF3	U4/U6 small nuclear ribonucleoprotein Prp3	(Schaffert et al., 2004)
	PRPF4	U4/U6 small nuclear ribonucleoprotein Prp4	(Schaffert et al., 2004)
	SART1	U4/U6.U5 tri-snRNP-associated protein 1	(Lemm et al., 2006)
	SF3A1	Splicing factor 3A subunit 1	(Nesic et al., 2004)
	SF3A2	Splicing factor 3A subunit 2	(Nesic et al., 2004)
	SF3A3	Splicing factor 3A subunit 3	(Nesic et al., 2004)
	SF3B2	Splicing factor 3B subunit 2	(Nesic et al., 2004)
	SNRPB	Small nuclear ribonucleoprotein Sm B	(Raška et al., 1991)
	SNRPB2	Small nuclear ribonucleoprotein Sm B"	(Raška et al., 1991)
	SNRPD1	Small nuclear ribonucleoprotein Sm D1	(Raška et al., 1991)
	SNRPD2	Small nuclear ribonucleoprotein Sm D2	(Raška et al., 1991)
	SNRPD3	Small nuclear ribonucleoprotein Sm D3	(Raška et al., 1991)
	SNRPE	Small nuclear ribonucleoprotein E	(Raška et al., 1991)
	SNRPF	Small nuclear ribonucleoprotein F	(Raška et al., 1991)
	SNRPG	Small nuclear ribonucleoprotein G	(Raška et al., 1991)
	SNRPN	Small nuclear ribonucleoprotein-associated protein N	(Raška et al., 1991)
	U2AF1	Splicing factor U2AF 35 kDa subunit	(Carmo-Fonseca et al., 1992)
	U2AF2	Splicing factor U2AF 65 kDa subunit	(Carmo-Fonseca et al., 1992)
snRNA maturation and snRNP assembly	ANKS1B	AIDA-1c	(Xu & Hebert, 2005)
	COIL	Coilin	(Raška et al., 1990)
	DDX20	Gem-associated protein 3	(Hao et al., 2007)
	FAM118B	Protein FAM118B	(Fong et al., 2013)

	GEMIN2	Gem-associated protein 2	(Hao et al., 2007)
	GEMIN4	Gem-associated protein 4	(Hao et al., 2007)
	GEMIN6	Gem-associated protein 6	(Hao et al., 2007)
	GEMIN7	Gem-associated protein 7	(Hao et al., 2007)
	ISG20	Interferon-stimulated gene 20 kDa protein	(Espert et al., 2006)
	PHAX	Phosphorylated adapter RNA export protein	(Boulon et al., 2004)
	SART3	Squamous cell carcinoma antigen recognized by T-cells 3	(Staněk et al., 2003)
	SMN1	Survival motor neuron protein	(Q. Liu & Dreyfuss, 1996)
	SNUPN	Snurportin-1	(Ospina et al., 2005)
	TGS1	Trimethylguanosine synthase	(Mouaikel et al., 2003; Verheggen, 2002)
	TOE1	Target of EGR1 protein 1	(Fong et al., 2013)
	TSPYL2	hCINAP	(Santama et al., 2005)
	USH1G	SANS	(Yildirim et al., 2021)
	USPL1	SUMO-specific isopeptidase USPL1	(Schulz et al., 2012)
	XPO1	Exportin-1	(Boulon et al., 2004)
sno/scaRNP assembly	DKC1	dyskerin	(U. T. Meier & Blobel, 1994)
	CASP8AP2	FLASH	(Barcaroli et al., 2006)
	FBL	fibrillarin	(Raška et al., 1990)
	GAR1	H/ACA ribonucleoprotein complex subunit 1	(Pogacic et al., 2000)
	NOLC1	Nopp1401	(Isaac et al., 1998)
	NOP10	H/ACA ribonucleoprotein complex subunit 3	(Pogacic et al., 2000)
	NOP56	Nucleolar protein 56	(Verheggen, 2002)
	NOP58	Nucleolar protein 58	(Verheggen, 2002)

	WRAP53	TCAB1/WDR79	(Mahmoudi et al., 2010; Tycowski et al., 2009)
Histone mRNA processing	LSM10	LSm10	(Pillai, 2001)
	LSM11	LSm11	(J.-L. Liu et al., 2009)
	NPAT	Protein NPAT	(Ma et al., 2000)
	SLBP	Histone RNA hairpin-binding protein	(Abbott et al., 1999)
Signaling	CDK2	Cyclin-dependent kinase 2	(J. Liu et al., 2000)
	FGF2	Fibroblast growth factor 2	(Bruns et al., 2009)
	PIAS4	E3 SUMO-protein ligase PIAS4	(Sun et al., 2005)
	PPP1CC	Serine/threonine-protein phosphatase PP1-gamma catalytic subunit	(Moorhead et al., 2007)
	PPP1R10	Serine/threonine-protein phosphatase 1 regulatory subunit 10	(Moorhead et al., 2007)
	PSME3	PA28	(Cioce et al., 2006)
	SPOPL	Speckle-type POZ protein-like	(Fong et al., 2013)
	SUMO1	Small ubiquitin-related modifier 1	(Navascues et al., 2008)

1.3 Cajal Body function in snRNP assembly

1.3.1 snRNP assembly and recycling

After transcription by Pol II, snRNAs undergo core assembly in the cytoplasm and return to the nucleus for modification and assembly into mature snRNPs (Figure 1.2). Immature snRNPs and transient intermediates in the snRNP assembly pathway concentrate in CBs in HeLa cells, providing the first evidence that CBs are the sites of snRNP assembly (Staněk & Neugebauer, 2004). Specifically, U6 mono-snRNP is targeted to CBs by SART3; there, it undergoes secondary structure rearrangements and base pairing with U4 snRNA, creating the U4/U6 di-snRNP. Next, U4/U6-specific proteins join, facilitating U5 snRNP recruitment and formation of the splicing-competent U4/U6·U5 tri-snRNP.

Mathematical modeling and fluorescence measurements indicate that assembly in the context of concentrated precursors makes snRNP assembly more efficient (Klingauf et al., 2006; Novotný et al., 2011).

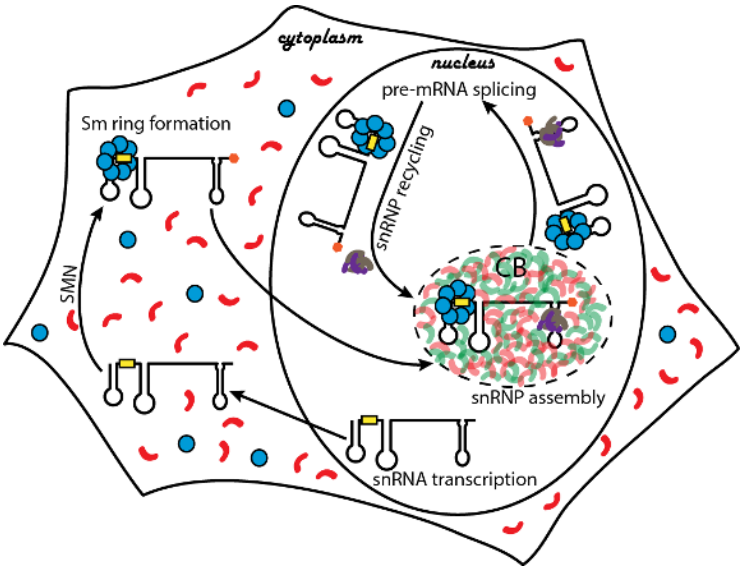


Figure 1.2 Nuclear and cytoplasmic steps of snRNP assembly
snRNAs are transcribed in the nucleus and exported to the cytoplasm. Once in the cytoplasm, SMN (red) mediates formation of the heptameric Sm ring (blue) onto all U snRNAs. snRNPs are then re-imported into the nucleus where U2, and U4/U6•5 snRNPs acquire snRNP specific proteins (grey and blue) to form splicing-competent snRNPs. During the splicing reaction, these snRNPs are disassembled and their components return to the CB for reassembly. Artwork by Valentina Botti.

Although snRNP assembly also occurs in the nucleoplasm and is not dependent on CBs, snRNP assembly is predicted to occur 11x faster in the CB than in the surrounding nucleoplasm (Klingauf et al., 2006). snRNP maturation in the nucleus occurs in two different forms: 1) after import of snRNP proteins into the nucleus and 2) after disassembly of the spliceosome following a splicing reaction. The U4/U6 di-snRNP (and consequently U4/U6·U5 tri-snRNP) falls apart after splicing and must be reformed to continue splicing (Staley & Guthrie, 1998). SART3 is necessary to reform the U4/U6 di-snRNP after a

splicing reaction (Bell, 2002). This, along with its presence in Cajal Bodies, suggests that di-snRNP reassembly can also occur in CBs. Knockdown of tri-snRNP disassembly factors hPrp22 and hNtr1 reduces free U5 snRNP in the nucleoplasm and leads to accumulation of di-snRNP in CBs, suggesting that di-snRNP intermediates accumulate in the CB if tri-snRNP assembly is stalled (Stanek et al., 2008). This suggests that tri-snRNP reassembly is occurring in CBs. The presence of snRNPs and their assembly intermediates in Cajal Bodies along with careful experiments knocking down snRNP components necessary for assembly demonstrates that CBs are sites of snRNP maturation.

1.3.2 Modification of snRNAs by snoRNPs

Proper snRNP assembly is dependent on nucleotide modification of snRNAs by small Cajal Body specific RNPs (scaRNPs) (Massenet et al., 2017). scaRNPS are protein-RNA complexes built around scaRNAs, snoRNAs that specifically localize to Cajal Bodies. snoRNAs consist of two main classes; 1) box C/D snoRNAs, which guide 2'-O-methylation of RNA and are bound by fibrillarin, NOP56, NOP58, and SNU13 and 2) box H/ACA snoRNAs, which guide pseudouridylation and are bound by dyskerin, Nhp2, Nop10, and Gar1. scaRNAs include an additional GU repeat (if box H/ACA) or Cajal body box (CAB) motif (if box C/D) that localize them to CBs (U. Thomas Meier, 2017). Most scaRNAs are box H/ACA, though some C/D and mixed (snoRNAs containing both box box H/ACA and box C/D) scaRNAs have been identified. snRNA modifications occur after nuclear import of mature snRNPs. Exogenous snRNA fragments are modified when targeted to CBs but not nucleoli, demonstrating that scaRNPs in CBs are responsible for snRNA modification (Jády et al., 2003). Knockdown of coilin does not affect snRNA

modifications, and snRNAs in cells lacking CBs are still modified (Deryusheva & Gall, 2009). This demonstrates that while modification of snRNAs occurs in CBs, it is not dependent on CBs.

1.4 Other Known Cajal Body functions

1.4.1 snoRNP assembly

Unlike snRNPs, snoRNP maturation occurs solely in the nucleus (Massenet et al., 2017). snoRNAs traffic through CBs before entering nucleoli (Machyna et al., 2014; Narayanan, 1999; Samarsky, 1998). Assembly of pre-snoRNP particles occurs cotranscriptionally (Darzacq et al., 2006; Fatica et al., 2002; Hirose et al., 2003; Richard et al., 2006; Yang et al., 2005). Box C/D snoRNAs are bound by SNU13 and NOP58 while Box H/ACA snoRNAs are bound by dyskerin, NHP2, and NOP10 through different co-transcriptional mechanisms. These pre-snoRNP particles then localize to CBs, where they undergo the final steps in their maturation before being transported to the nucleolus (or in the cased of scaRNPs, retained in CBs). CBs accumulate a short isoform of the decapping enzyme TGS1, which hypermethylates snoRNAs in the nucleoplasm much like the full length TGS1 isoform hypermethylates snRNAs in the cytoplasm (Girard et al., 2008). These long and short forms of TGS1 differentially interact with snRNAs and snoRNAs respectively (Pradet-Balade et al., 2011). Not all snoRNAs are capped, yet they still transfer to CBs. This is because other factors involved in snoRNP maturation also concentrate in Cajal bodies (Massenet et al., 2017). For example, fibrillarin and Gar1 both concentrate in CBs and bind Box C/D snoRNPs and Box H/ACA snoRNPs in the CB respectively. snoRNAs may also be modified by scaRNPs in the CB before entering the nucleolus (U. Thomas Meier, 2017). Nopp140, a snoRNP chaperone, is present in both nucleoli and CBs,

and can shuttle between the two bodies (Isaac et al., 1998). This suggests that Nopp140 transports mature snoRNPs from CBs to nucleoli. Altogether, the presence of snoRNP maturation factors and trafficking of snoRNAs through CBs suggests that CBs are also sites of snoRNP assembly.

1.4.2 Telomerase maturation

Extension of chromosomal telomeres is driven by telomerase, an RNP made of an RNA subunit (hTR), reverse transcriptase (hTERT), and box H/ACA core proteins (Shay & Wright, 2019). Like snoRNPs, telomerase maturation occurs solely inside of the nucleus. hTR contains an H/ACA RNA-like domain with a CAB box motif that causes accumulation of HTR in CBs(Jady et al., 2004). hTR is cotranscriptionally bound by H/ACA snoRNP proteins NOP10, NHP2, and NAF1 (Schmidt & Cech, 2015). hTR transit through CBs requires Wrap53 (also known as TCAB1), a protein that binds the CAB box of scaRNAs and recruits them to CBs (Venteicher et al., 2009). Wrap53 is essential for proper maturation of telomerase and consequently, telomere maintenance (Zhong et al., 2011). Once in CBs, NAF1 is replaced by GAR1 and the RNP subsequently associates with hTERT to form mature telomerase which then localizes to telomerase. While WRAP53 is necessary for telomerase maturation as well as CB assembly in some cell types, coilin (and therefore CBs) is not necessary for telomerase maturation or telomere lengthening. RNA-FISH experiments have shown distinct localization of hTR to CBs, yet low endogenous levels of hTERT have complicated analysis of mature telomerase in CBs. A recent live-cell super-resolution imaging study of telomerase showed that fewer than 10% of hTR molecules reside in CBs, yet their residence time in the CB is longer than hTERTs residence time in CBs (Laprade et al., 2020; Schmidt et al., 2016). Additionally, super

resolution images of hTR in CBs show hTR on the periphery of CBs. Together, these data suggest a model where hTERT binds hTR at the periphery of CBs. If we consider hTR to be a class of scaRNP, then its presence in CBs could be for maturation of the RNP, and CBs are not directly involved in telomere maintenance as previously suggested. This further establishes CBs as sites of RNP maturation.

1.5 Cajal Bodies in Disease and Development

CBs vary in number and size in different cell types as well as throughout different stages of development. CBs have been studied throughout development in a variety of different organisms including fruitflies (*D. melanogaster*), frogs (*X. laevis*), zebrafish (*D. rerio*), and plants (*A. thaliana*). *X. laevis* oocytes contain 50-100 CBs ranging between 1-10 μm in size (Gall, 2000). The abundance and size of CBs in Xenopus oocytes made it an ideal model for early characterization of CB components as well as the dynamics of coilin in the CB (Deryusheva & Gall, 2004; Handwerger et al., 2003). Characterization of coilin in *Xenopus* and *Drosophila* demonstrated that coilin is a component of both CBs and HLBs and that CBs and HLBs can sometimes mix. In *Drosophila* melanogaster nurse cells, CBs are abundant, yet their somatic cells only have 1 CB per nucleus. Studies in zebrafish embryos demonstrated that CB number is tightly regulated throughout development. At the one cell stage, embryos can contain up to 30 CBs, a number that is reduced until differentiated cells display only 2 CBs per cell (Strzelecka, Oates, et al., 2010). CBs are present in the absence of transcription in early zebrafish embryos and prior to bulk zygotic genome activation (ZGA) in fruitflies (Batalova et al., 2005; Heyn et al., 2017; J.-L. Liu et al., 2009). The abundance of CBs in oocytes and embryos may be to provide embryos with factors that will be required during the maternal to zygotic transition (MZT), when the

zygotic genome becomes active for the first time and when cell cycles lack prolonged growth phases for biosynthesis (R. F. Walker et al., 2015). Instead, the embryo relies on maternally provided gene products during the rapid, synchronous cell divisions of cleavage stage. In the somatic cells of zebrafish embryos, CBs were only detected in muscle and motor neurons (Strzelecka, Oates, et al., 2010). Interestingly, immunostaining of coilin in fetal human as well as fetal pig tissue shows an abundance of CBs throughout fetal tissue, but not in adult tissue (Young et al., 2001). Instead, adult neurons, muscle, and liver cells have 1-2 CBs per nucleus, while other somatic cells lack CBs entirely. One hypothesis for tissue specificity of CBs is that CBs are present in cells that are highly metabolically active. Because CBs enhance the efficiency of RNP assembly, an increased number of CBs may reflect an increased need for RNP turnover in cells. To further understand the relevance of CB tissue-specificity and whether CBs have specialized roles in tissue types, it is crucial to have a comprehensive understanding of CB components to aid further study.

CBs were originally discovered in neurons and it has since been shown that neurons are one of few somatic cell types with CBs. It is not surprising therefore, that CBs have been linked to some neuropathologies. The most common example of this, is the role of SMN in Spinal Muscular Atrophy (SMA). SMA, a genetic disease, is the result of defects in SMN1 gene that lead to low levels of SMN in cells. CBs in SMA patient motor neurons as well as in SMA mouse models are disrupted. This disruption presents itself in two different phenotypes; 1) a reduced number of CBs per cell nucleus and 2) a redistribution of coilin into perinucleolar caps. Cajal bodies have also been linked to polyglutamine (polyQ) expansion diseases. Poly-Q diseases result in the formation of phase-separated nuclear aggregates called “nuclear inclusions.” Cajal bodies have been shown to interact

with nuclear inclusions in two Poly-Q diseases, dentatorubral-pallidoluysian atrophy (DRPLA) and Machado-Joseph disease (MJS), despite CB number and morphology being unchanged (Yamada et al., 2001). Additionally, coilin has been found to interact directly with the disease product of another polyQ disease, Spinal Cerebellar Ataxia Type 1 (SCA1) (Hong et al., 2003). In a tissue culture model, coilin interacts with ATXN1, the SCA1 gene product, and co-localizes to ATXN1 nuclear aggregates. Recently, the protein VRK1 was identified as a regulator of CB formation. It binds coilin, regulates its phosphorylation, and is necessary for CB formation in motor neurons (Cantarero et al., 2015; El-Bazzal et al., 2019). Mutations in VRK1 have been identified in a variety of neuromotor diseases including SMA, amyotrophic lateral sclerosis (ALS) and distal hereditary motor neuropathies (dHMNs) (Martín-Doncel et al., 2019). At least four of these mutations lead to disassembly of Cajal body assembly (Cantarero et al., 2015; El-Bazzal et al., 2019; Marcos et al., 2020; Martín-Doncel et al., 2019). VRK1 mutations that disassemble CBs may affect motor neuron function by preventing efficient CB function.

1.6 Formation of Nuclear Bodies on actively transcribing gene loci

Nuclear bodies such as the nucleolus, Cajal body (CB), and the histone locus body (HLB) concentrate factors required for nuclear steps of RNA processing. Formation of these nuclear bodies occurs on genomic loci and is frequently associated with active sites of transcription. Whether nuclear body formation is dependent on a particular gene element, an active process such as transcription, or the nascent RNA present at gene loci is a topic of debate. Their overarching cellular roles reflect the need for the efficient expression and maturation of preribosomal subunits assembled in the nucleolus,

spliceosomal small nuclear ribonucleoprotein particles (snRNPs) assembled in the CB, and histone mRNAs processed in the HLB (Figure 1.3 and Table 2).

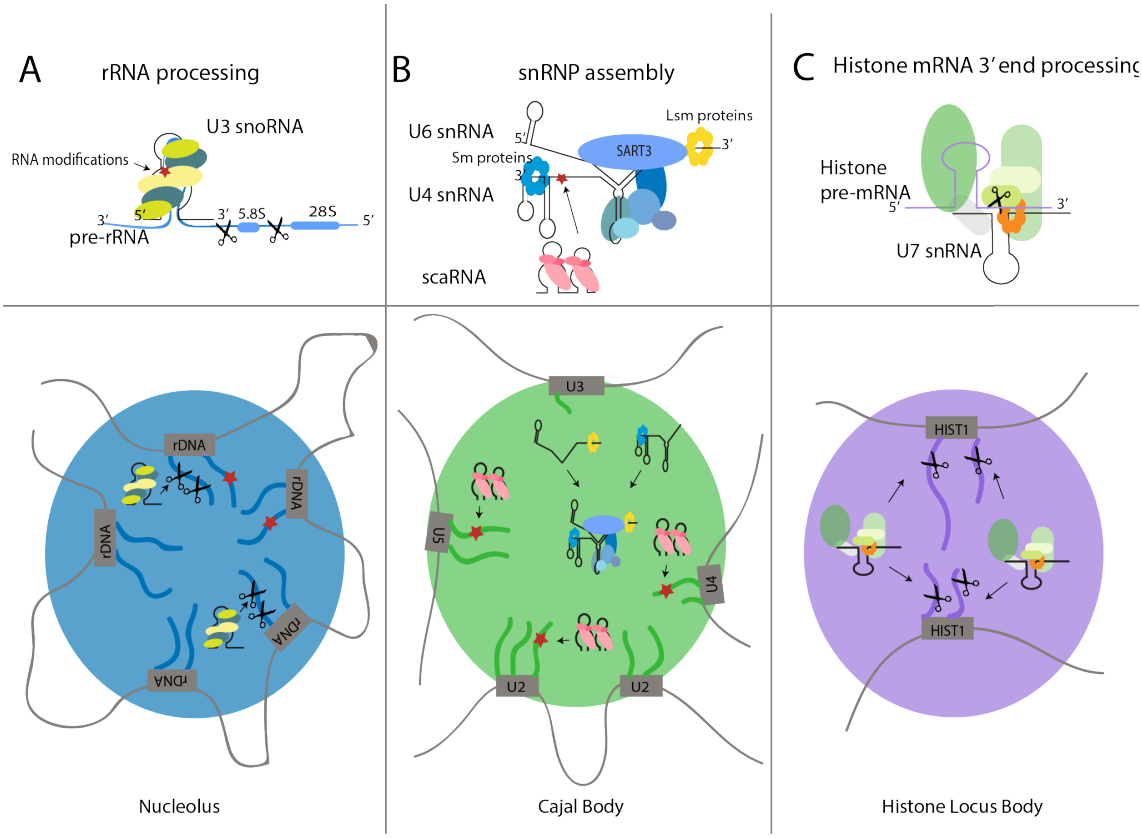


Figure 1.3 Nuclear bodies organize chromatin and concentrate components for RNA processing
(A) rRNA processing occurs in nucleoli (top), and snoRNPs in the nucleolus modify pre-rRNA (here depicted by a star). The U3 snoRNP binds nascent rRNA and promotes cleavage of nascent pre-rRNA into its 18S, 5.8S, and 28S components. Nucleoli (bottom, blue) form on rDNA repeats and concentrate factors for rRNA processing. (B) snRNP assembly occurs in CBs (top), and scaRNAs in the CB modify nascent snRNAs. Assembly of the U4/U6 snRNP occurs in CBs and is depicted here. CBs (bottom, green) form on snRNA gene loci and concentrate factors for snRNP assembly. (C) Histone mRNA 3′ end processing occurs in HLBs (top). The histone cleavage complex (HCC) binds nascent RNA at replication-dependent histone genes. U7 snRNA base pairs with the histone downstream element, and protein constituents of the snRNP guide nascent RNA cleavage in the HLB. HLBs (bottom, purple) form on histone gene clusters and concentrate factors for histone mRNA processing. Figure originally published in Arias Escayola and Neugebauer, 2018.

Table 2 Nuclear Body Function and Genomic Location

Body	Function	Genomic site

Nucleolus	pre-rRNA cleavage rRNA pseudouridylation rRNA 2’ O-methylation pre-ribosomal subunit assembly	Nucleolar Organizing Regions (NORs) rRNA gene loci
Cajal Body	snRNA transcription targeting new snRNAs for export snRNA pseudouridylation snRNA 2’ O-methylation snRNP assembly U3/U8 snoRNA 5’ end capping (7-meG) snoRNP assembly	snRNA gene loci: U1, U2, U11, U4, U5, U7 snoRNA gene loci: U3, U8
Histone Locus Body	histone gene transcription histone mRNA 3’ end processing	Replication-dependent histone gene loci: HIST1, HIST2

Gene expression and subsequent RNA processing are tightly linked to spatial organization in the nucleus (Mao et al., 2011). We have seen that nuclear bodies form on genomic loci and play an important role in nuclear organization, by concentrating RNA processing factors in each respective organelle. Nucleoli, CBs, and HLBs are separate entities within the nucleus and assemble on distinct gene loci. The formation of such distinct nuclear bodies that are specifically involved in RNA processing steps taking place on distinct pools of RNA may result from demixing of nuclear body components (Berry et al., 2015; Jain & Vale, 2017; Lin et al., 2015; Saha et al., 2016; Zhang et al., 2015). The nucleolus assembles at repeated rRNA gene loci and serves as a site of ribosome biogenesis. As rDNA is transcribed, the nucleolus concentrates ribosome maturation factors along NORs (François-Michel Boisvert et al., 2007). Not all rRNA genes are actively transcribed, and rDNA/NOR positioning in nucleoli is dependent upon transcription (Kalmárová et al., 2007). In *C. elegans*, RNA plays a modulatory role in nucleolus formation (Berry et al., 2015), explaining how enhanced rDNA transcription

caused by the ncl-1 mutation leads to larger nucleoli (Frank & Roth, 1998). Cajal bodies interact with snRNA and histone loci on separate chromosomes, facilitating intrachromosomal interactions in these otherwise distant loci (Wang et al., 2016). Cells lacking CBs do not display this chromosomal organization, and depletion of CBs correlates with lower levels of snRNAs, suggesting potential roles in transcription or stability. Coilin, the CB scaffolding protein, directly interacts with snRNA gene loci and their transcript product (Machyna et al., 2014). Histone genes occur in clusters at which transcription of the different histone genes is tightly regulated (Marzluff & Koreski, 2017). These arrays co-localize with HLBs, where efficient histone gene transcription as well as histone pre-mRNA processing occurs. Remarkably, placement of a histone gene transcription unit at an exogenous site in the genome leads to formation of a combined CB/HLB at the site in tissue culture cells that fused CBs and HLBs (Shevtsov & Dundr, 2011).

Based on the findings described above, three models are currently being considered to explain the formation of nuclear bodies at genomic sites. Two of these follow a "seeding model" but differ in the nature of the seed. The term "seed" refers to an element that stabilizes the formation of the nuclear body and promotes phase separation. The first of these models suggests that formation is dependent on a particular gene element (DNA) that acts as the seed and recruits nuclear body factors. The second is that the seed for nuclear bodies is the nascent RNA present at the gene locus. In contrast, the third existing model proposes that an active process (and its associated factors) such as transcription is required for formation. So far, most of the work done in embryos has focused on transcription as the active process in nuclear body formation. However, in tissue culture, post-translational modifications (PTMs) of proteins can modulate whether certain proteins form nuclear

bodies (Hebert & Poole, 2017). For example, methylation of coilin regulates whether coilin is in CBs or in residual bodies known as gems (Hebert et al., 2002). While active processes such as PTMs can modulate nuclear body formation, transcription is an example of this particular model. In human cells, addition of a non-transcriptionally active mouse NOR is sufficient to recruit human UBF and Pol I transcription machinery but requires a functional promoter sequence to form a functional nucleolus, suggesting that gene elements are the seed for nucleolar formation (Grob et al., 2014). Additionally, previous work in *Drosophila* embryos demonstrated that the *H3–H4* promoter sequence is sufficient to provide a scaffold for HLB factors FLASH and Mxc (Salzler et al., 2013). On the other hand, there is strong evidence that nuclear body formation is transcriptionally driven (see above). When transcription is inhibited at the onset of mitosis in somatic cells, the nucleolus disassembles (François-Michel Boisvert et al., 2007). In mammalian cells, recruitment of coilin and snRNAs to CBs is also transcription-dependent (Carmo-Fonseca et al., 1992). A reduced level of transcription of histone genes during cell cycle arrest results in a loss of HLBs (Bongiorno-Borbone et al., 2010). Most of these studies have been performed in mammalian tissue culture cells, and the correlation of nuclear body formation with transcription has relied on active tracking of the cell cycle or modulation of transcriptional activity via the use of inhibitors.

1.7 Aims of the thesis

Despite over 30 years of research in the field of Cajal bodies since the discovery of the main CB protein, coilin, there is still no comprehensive list of CB components. Though known to serve as sites of assembly for spliceosomal components (snRNPS) and transcription of regulatory RNAs (snRNAs, snoRNAs, histone RNAs), the full complexity of proteins that contribute to CB assembly and function is unknown. In my thesis, I use a combination of proteomics, microscopy, and DNA sequencing to identify new CB components and regulators of CB assembly.

2 Comprehensive Identification of the Cajal Body Proteome

2.1 Author Contributions

I prepared all cell lines and samples detailed in this chapter. Emily Nischwitz performed mass spectrometry of APEX2 samples along with mass spectrometry data analysis that contributed to Figure 2.5.

2.2 Goals and Approach

Traditionally known CB components - such as snRNP proteins and Nopp140 – were identified by immunostaining that revealed cellular co-localization of the target with coilin in nuclear foci. This is highly dependent on having good antibodies for a protein of interest, as well as a good list of candidates. One previous study attempted denovo identification of CB components by performing a whole-genome microscopy-based screen of HA-Flag tagged proteins and analyzing their localization in nuclear bodies, identifying 4 novel CB proteins in the process (Fong et al., 2013). However, expression of tagged proteins can often disrupt their localization in CBs, and this screen excludes identification of CB proteins that may also be in other nuclear bodies (such as the nucleolus) or only transiently interact with CBs. To date, there are no biochemical approaches to comprehensively identify CB proteins.

Comprehensive knowledge of CB components is crucial for understanding how CBs assemble and function. Our lab has previously immunopurified human coilin and conducted mass spectrometry (Machyna et al., 2013a). Several known coilin interactors

were absent from this preparation, and no new CB components were identified. We and others have also shown that even mild detergent treatment of cells results in the rapid loss of components from CBs in situ and in extracts (Y. W. Lam et al., 2002; Staněk et al., 2003). This is likely because interactions within the CB are transient and short-lived. Indeed, most CB proteins fully exchange with the surrounding nucleoplasm in less than one minute (Dundr et al., 2004). APEX2 is an engineered peroxidase that uses hydrogen peroxide to catalyze the formation of biotin-phenoxyl radicals from biotin-phenol (S. S. Lam et al., 2015). These biotin-phenoxyl radicals then "tag" nearby endogenous proteins at electron-rich amino acids, covalently biotinylating these proteins. Alternative methods employing BioID or other similar biotin ligases require 18 hours of labeling time; coilin molecules would cycle from the CB to the nucleoplasm repeatedly during this time, and CBs would disassemble as cells enter mitosis (Carmo-Fonseca et al., 1993; Carmo-Fonseca et al., 1992; W. Qin et al., 2021). TurboID, a similar protein ligase, only requires 10 minutes of labeling time, but this is still above the residence time of coilin and other CB components in the CB (Dundr et al., 2004; W. Qin et al., 2021). APEX2 biotinylation occurs within 20nm of APEX2 fusion protein and within one 1 minute of peroxide addition. Due to its more rapid action, I reasoned that APEX2 would allow for comprehensive identification of transient interactors in highly dynamic CBs *in situ.* In this chapter, I describe the use of APEX2 to comprehensively identify CB components, including transient components, by biotinylating proteins proximal to coilin regardless of binding affinity.

2.3 Establishing Proximity Biotinylation for CBs

The rapid exchange of Cajal body (CB) components with the surrounding nucleoplasm, has frustrated the characterization of the CB proteome by biochemical purification. To identify CB components, I chose to use APEX2, a peroxidase which uses biotin phenol and hydrogen peroxide to covalently biotinylate nearby proteins in situ in 1 minute (Figure 2.1).

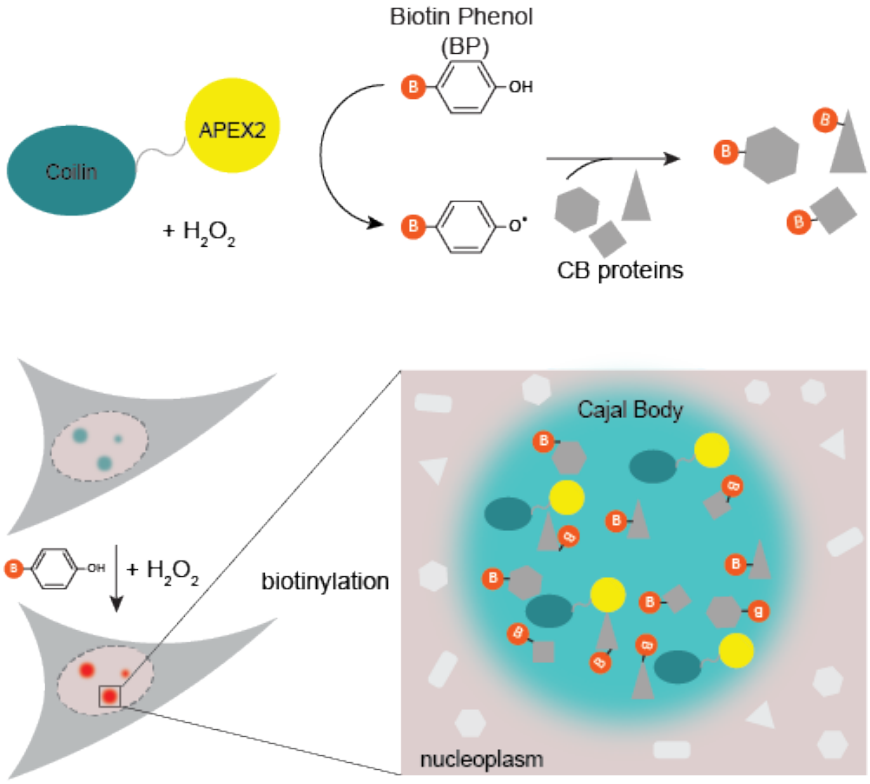


Figure 2.1 Diagram of APEX2 Biotinylation
In the presence of H_2O_2, APEX2 (yellow) catalyzes the conversion of Biotin Phenol (BP) to a biotin free radical. The radical then diffuses away and covalently attaches to nearby macromolecules. APEX2 tagged with coilin biotinylates proteins inside of the Cajal Body.

To target the Cajal body proteome, I designed a construct tagging coilin with APEX2 and a V5 linker (coilin-APEX2) (Figure 2.2). I designed two control constructs that would target the nucleoplasm when expressed in cells. The first, an APEX2-NLS

construct directly targets APEX2 to the nucleus with a nuclear localization signal (NLS). For the second, I exploited the requirement of the Coilin N-terminal domain (NTD) to target CBs and created a second coilin construct lacking the NTD (ΔNTD-APEX2).

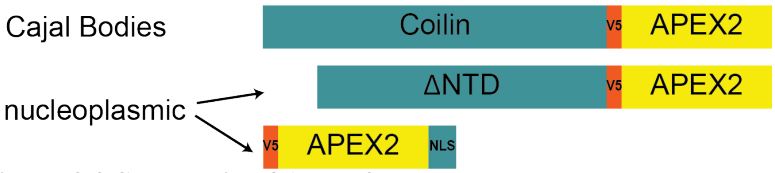


Figure 2.2 Schematic of APEX2 constructs

Transient transfection of the coilin-APEX2 construct caused an overexpression of coilin along with aberrant CB formation. To obtain close to endogenous levels of coilin-APEX2 expression, I generated cell lines via lentiviral transduction and sorted for single cells that were then screened for CBs. APEX2-NLS and ΔNTD-APEX2 expressing cell lines were generated and sorted into low-expressing pools. I confirmed that the coilin-APEX2 localized to endogenous CBs containing coilin, snRNPs, SMN, and Nopp140 by immunofluorescence (Figure 2.3A). Both APEX2-NLS and ΔNTD-APEX2 constructs localized to the nucleoplasm and did not form or associate with CBs (Figure 2.3B and 2.3C). Using these three cell lines, I can distinguish CB components from nucleoplasmic components.

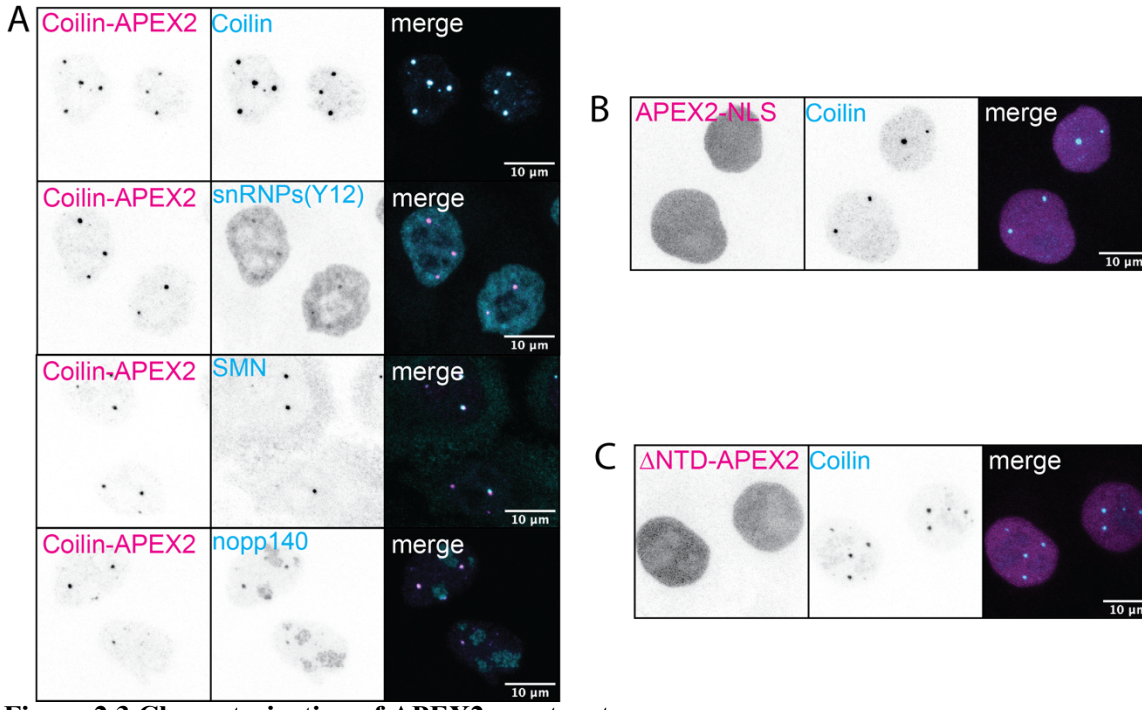


Figure 2.3 Characterization of APEX2 constructs
A) Immunofluorescent staining of Coilin-AEPX2 labeled by anti-V5 antibody (magenta) and anti-coilin, anti-Sm (Y12), anti-SMN, or anti-Nopp140 antibody (cyan). B and C) Immunofluorescent staining of APEX2-NLS (B) and ΔNTD-APEX2 (C) labeled y anti-V5 antibody (magenta) and anti-coilin (cyan).

I performed protein biotinylation using biotin phenol and hydrogen peroxide on all APEX2-expressing cell lines, then fixed and stained cells with streptavidin to characterize the sites of biotinylation. Coilin-APEX2 reactivity was restricted to CBs while APEX2-NLS and ΔNTD-APEX2 reactivities were present throughout the entire nucleus without showing enrichment in any nuclear body (Figure 2.4A). All three cell lines showed faint levels of biotinylation in the cytoplasm arising from endogenously biotinylated proteins as well as endogenous peroxidases that increase background signal regardless of whether cells have been exposed to biotin phenol and hydrogen peroxide. To exclude this background signal from our analysis, biotinylated cells were fractionated into nuclear and cytoplasmic compartments and only the nuclear fractions were analyzed. To confirm successful

depletion of biotinylated cytoplasmic proteins, I performed western blotting. Biotinylated proteins were seen in the cytoplasmic fraction of both control and experimental samples, consistent with endogenously biotinylated proteins (Figure 2.4B). Robust biotinylation of newly appearing proteins was only seen in the nuclear fraction, corresponding to localization of the APEX-tagged proteins. After biotinylation and fractionation of cells, I therefore enriched biotinylated proteins from nuclear fractions using streptavidin magnetic beads and confirmed successful enrichment via western blot (Figure 2.4C).

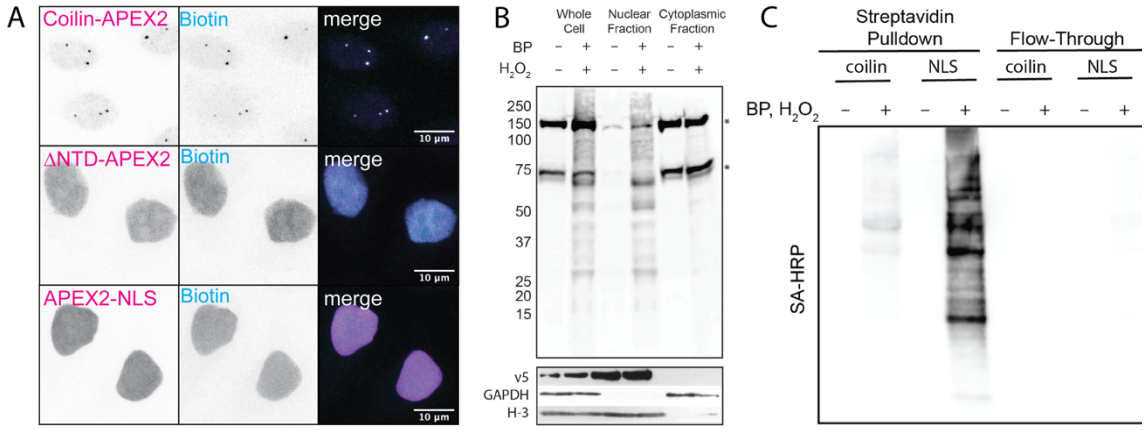


Figure 2.4 Enrichment of CB proteins using APEX2
A) Immunofluorescent staining of Coilin-APEX2, ΔNTD-APEX2, and APEX2-NLS after addition of biotin phenol (BP) (30 min) and hydrogen peroxide (1 min) stained with anti-V5 antibody (magenta) and Fluorescent Streptavidin to label for biotin (cyan). B) Streptavidin Blot of fractionated coilin-APEX2 cells +/- BP and H_2O_2 showing endogenously biotinylated proteins (*) and biotinylated proteins after treatment (top). Primary antibodies showing localization of proteins in fractions; anti-V5 antibody is a marker for Coilin-APEX2, anti-GAPDH antibody is a marker for cytoplasm, anti-H3 antibody is a marker for nucleus. C) Streptavidin Blot of biotinylated proteins pulled down from nuclear lysates +/- BP and H_2O_2 treatment.

2.4 Identification of New Cajal Body Components

Enriched proteins from nuclear fractions were identified using liquid chromatography followed by tandem mass spectrometry (LC MS-MS). In total, I identified

100 proteins significantly enriched in Coilin-APEX2 compared to ΔNTD-APEX2 (43 hits) and APEX2-NLS (77 hits) conditions (Table 3 and Figure 2.5 A and B).

Table 3 List of Proteins enriched in Coilin-APEX2

Function	Proteins enriched in Coilin-APEX2
Cell Cycle Progression	ANLN, ANXA11, BUB3, KIF4A, NCAPD3, NUMA1, ZNF207
DNA Replication and Repair	ATAD2, POLD1, RFC5, RECQL, RIF1, USP7
Transcription	CDC73, ERCC3, FUBP1, GTF2F2, GTF3C5, HTATSF1, IRF2BP1, MED1, MED14, MED17, POLR2C, SMARCA4, SMARCC1, SSRP1, SUPT5H, SUPT6H, UBR5, **ELL, ICE1, ICE2, POLR2A, POLR2B**
Pre-mRNA splicing and cleavage	CLP1, CPSF6, DHX38, IK, KHSRP, NUDT21, PPAN, RBM10, SON, SRRM2, **EFTUD2, NHP2L1, PRPF3, PRPF4, SART1, SF3A1, SF3B2, SNRPB, SNRPD1, SNRPD2, SNRPD3, SNRPE, TARDBP**
snRNA maturation and snRNP assembly	**COIL, DDX20, SART3, TGS1, TOE1**
rRNA processing and ribosome biogenesis	GNL3L, GTPBP4, NSA2, TEX10, **DKC1, GAR1, NOLC1, NOP56, NOP58**
Histone mRNA processing	**LSM11, NPAT**
miRNA processing	SRRT
RNA helicase	DDX18
RNA exosome	EXOSC2
Ribosomal Protein	RPL13, RPL14, RPL15, RPL24, RPL3, RPL34, RPL7A, RPL8, RPS2, RPS24
Other	AP2A2, ATP6V1A, CSNK2A1, CSNK2B, CYB5R1, HIST1H2AJ, NUP54, OXA1L, PPIL4, SFXN3, SLC25A13, SPCS3, **PSME3**

Importantly, there was a large overlap in proteins identified, with 30 of these enriched proteins shared between the two data sets. As expected, there were fewer proteins enriched

when compared to the ΔNTD-APEX2 control, since the coilin truncation can interact with CB proteins in the nucleoplasm. Of the 100 total enriched proteins, 30 were previously known CB proteins, with 17 of these appearing in both datasets, demonstrating specificity in enriching CB proteins (Figure 2.5C). 30% of significantly enriched hits were CB components (Figure 2.5E). A previously published study using BioID identified 342 coilin interacting proteins (Go et al., 2021); thirty-one of these proteins were also enriched in our Coilin-APEX2 dataset (Figure 2.5D). Gene ontology (GO) analysis of the 100 enriched proteins reveals strong association with functions and complexes known to be in CBs such as snRNA binding, snoRNA binding, and Sm-like protein family complex (Figure 2.5F and G).

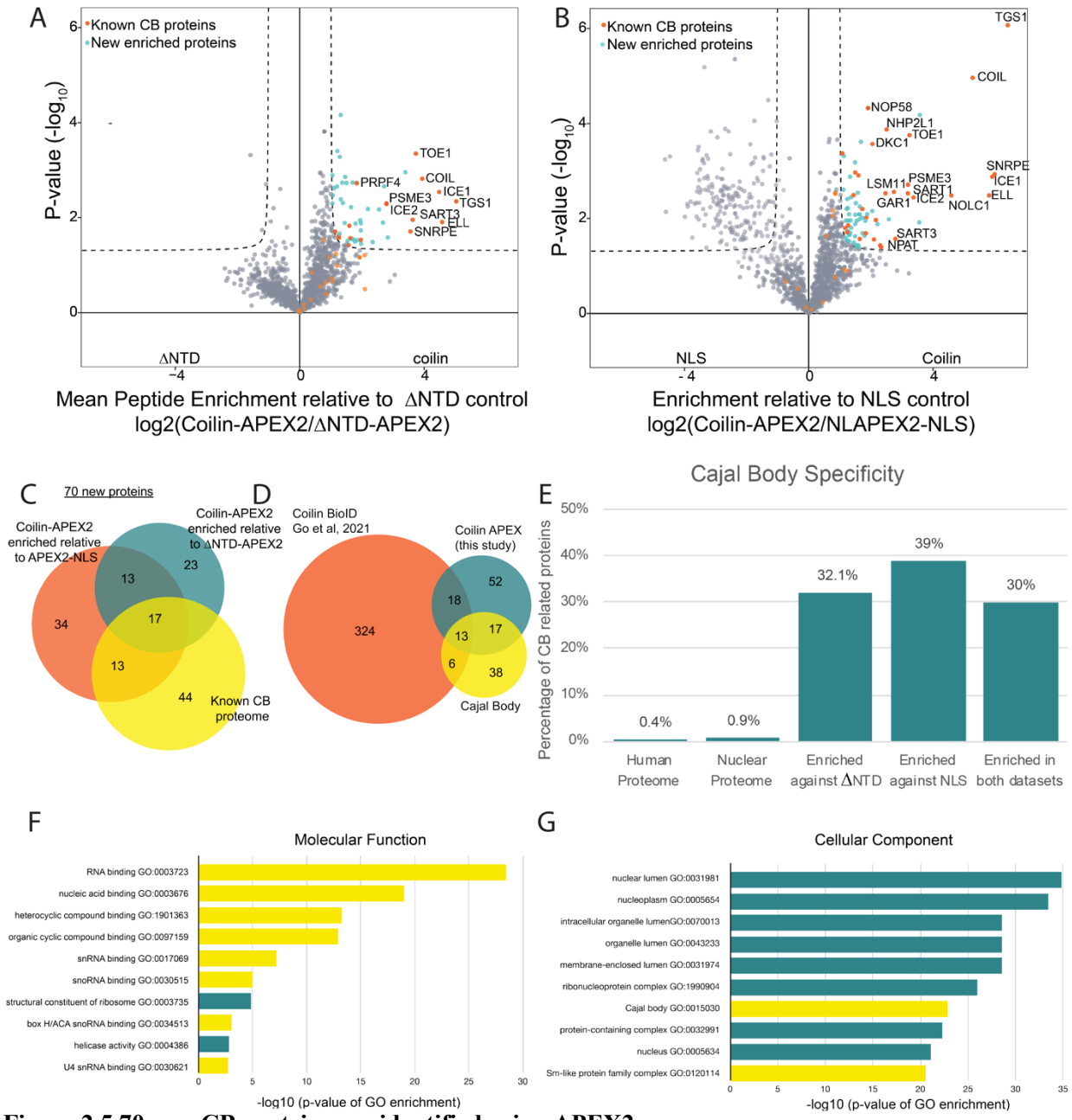


Figure 2.5 70 new CB proteins are identified using APEX2.

A) and B) Label free quantification (LFQ) of CB proteome identification using Coilin-APEX2 against ΔNTD-APEX2 (A) and APEX2-NLS (B) controls, data obtained from four biological replicates. Known CB proteins are labeled and colored orange, new, significant hits are colored cyan. C) Venn diagram of statistically enriched CB proteins enriched against ΔNTD-APEX2 (teal) and APEX2-NLS (orange) compared to known CB proteome (yellow). D) Venn diagram of statistically enriched CB proteins against both controls (teal) compared to Coilin-BioID (orange)(Go et al., 2021) and known CB proteome (yellow) E) Cajal body proteins are enriched in the 100 hits. The percent of proteins in the human and nuclear proteome and of the enriched hits that localize to CBs. F and G) GO-term analysis of proteins enriched in Coilin-APEX2. The 10 most significant hits along with the -log10 of their p-value are displayed, CB associated GO-terms are highlighted in yellow.

2.5 The transcription factor IRF2BP1 is highly enriched in the Cajal Body

To validate the new "hits" enriched in coilin-APEX2 as CB proteins co-immunofluoresence was performed with antibodies against the new proteins and coilin. The highly enriched, highly significant hits, such as the new hit IRF2BP1, strongly co-localized with coilin in CBs (Figure 2.7A). Other hits, such as SRRT, showed diffuse nuclear staining but were not excluded from CBs (Figure 2.7B). This pattern was also true with previously known CB components (TOE1 and EFTUD2).

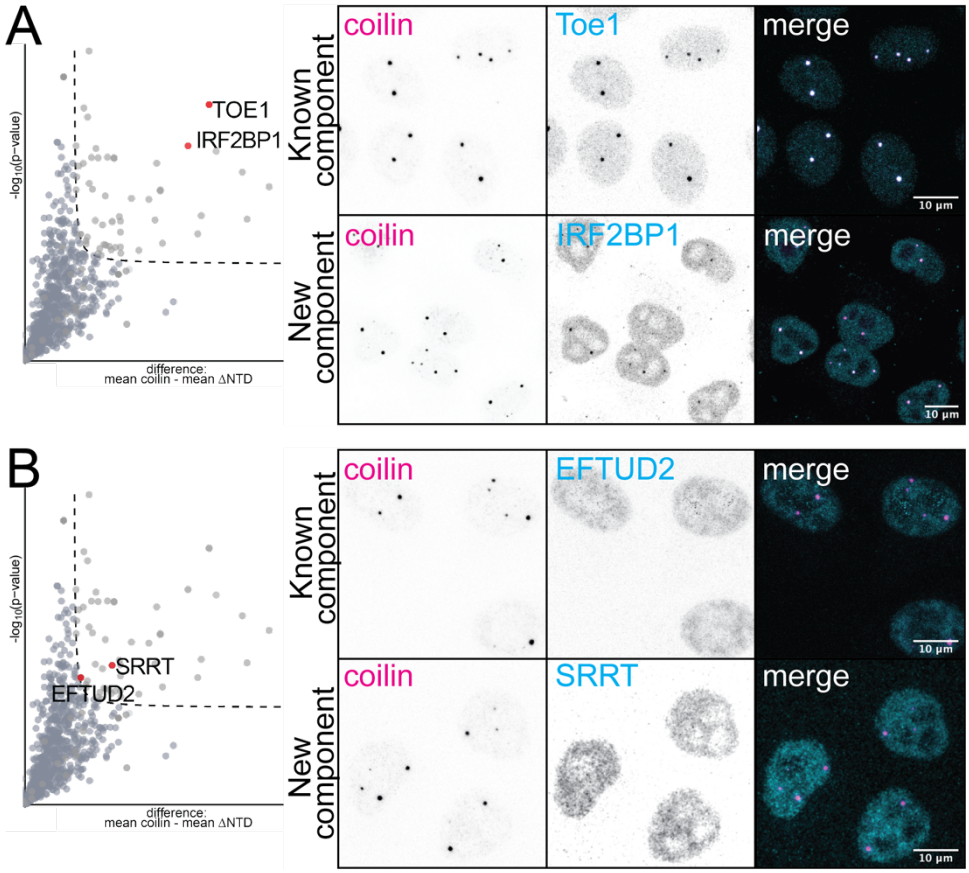


Figure 2.6 Validation of novel CB components
A) Volcano plot showing location of TOE1 (known component) and IRF2BP1 (new component) (left). Immunofluorescent staining of highly enriched, highly significant components in CBs using anti-coilin (magenta) and anti-Toe1 or anti-IRF2BP1 antibodies (cyan) (right). B) Volcano plot showing location of EFTUD2 (known component) and SRRT (new component) (left). Immunofluorescent staining of low enrichment, low significance components in CBs unsing anti-coilin (magenta) and anti-EFTUD2 or anti-SRRT antibodies (cyan) (right).

Of the new hits, Interferon Regulatory Factor Binding Protein 1 (IRF2BP1) was the most significant and highly enriched protein. Therefore, I further characterized its localization in CBs by immunofluorescence. Although IRF2BP1 showed strong colocalization with coilin, it did not co-localize with the SMN-containing portion of CBs (Figure 2.7A and B). Line graphs drawn through CBs in cells stained with IRF2BP1 showed that IRF2BP1 signal completely overlaps with coilin signal, while it is slightly offset from the SMN signal (Figure 2.7B). Knockdown of coilin leads to diffuse IRF2BP1 signal, demonstrating that IRF2BP1 bodies are CBs and dependent on coilin expression (Figure 2.7C). Interestingly, knockdown of IRF2BP1 does not affect CB number or recruitment of snRNPs to CBs (Figure 2.7D). Therefore, IRF2BP1 is not an essential component for CB formation.

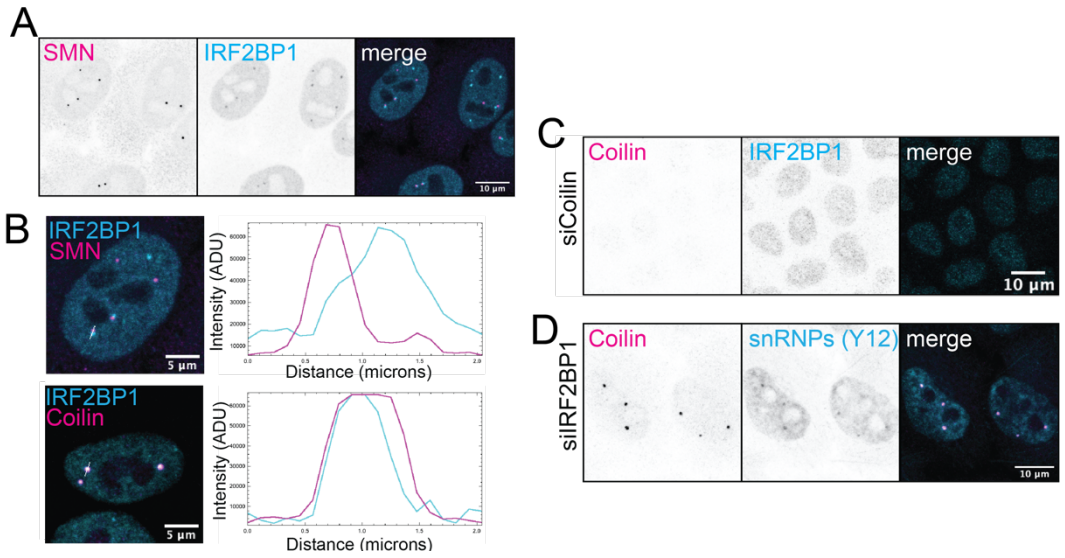


Figure 2.7 IRF2BP1 is a novel CB component
A) Immunofluorescent staining of SMN (magenta) and IRF2BP1 (cyan). B) Representative single nucleus immunofluorescent staining of SMN (magenta) and IRF2BP1 (cyan) or Coilin (magenta) and IRF2BP1 (cyan) with accompanying line profile plots showing intensities through a single CB. C) Immunofluorescent staining of IRF2BP1(cyan) and coilin (magenta) after siRNA knockdown of coilin. D) Immunofluorescent staining of snRNPs (cyan) and coilin (magenta) after siRNA knockdown of IRF2BP1.

2.6 Discussion

I identified 70 new CB proteins using APEX2 biotinylation, expanding the known CB proteome and including transient CB interactors. We chose APEX2 biotinylation over other robust proximity biotinylation techniques, such as BioID or TurboID, due to its short labeling time. We were able to achieve robust biotinylation of CBs with one minute labeling times, just below the 2-3 minute residence time of coilin in CBs (Dundr et al., 2004). For comparison, a recent study by the Gingras lab used coilin as a bait protein in a BioID-based map of the human cell to identify proteins that concentrate in nuclear bodies (Go et al., 2021). Their study identified more prey proteins (385) but was overall less specific to CB proteins. Using BioID, biotinylation was robust throughout the nucleus and not restricted to CBs. In contrast, I showed that biotinylation signal was dramatically concentrated in CBs. This demonstrates that our approach is more specific for CB-localized proteins. For example, their study does not identify IRF2BP1 as a CB protein, yet I clearly show IRF2BP1 as a highly enriched CB protein. Interestingly, some of the same ribosomal proteins I discovered and studied were also detected by BioID. Taken together, these results demonstrate that using proximity biotinylation with APEX2 is a robust and specific method for identifying transient components of nuclear bodies.

A key step in sample preparation is the addition of a nuclear isolation step. Mammalian cells contain ~5 carboxylases that are endogenously biotinylated and localized to the cytosol and mitochondria (Tong, 2013). These biotinylated carboxylases are known contaminants in proximity biotinylation mass spectrometry (Papageorgiou et al., 2013). APEX2 biotinylation was first tested on enrichment of mitochondrial and endoplasmic reticulum membrane proteins, meaning that background signal from mitochondrial carboxylases was not problematic (S. S. Lam et al., 2015). Additionally, many early APEX2 experiments used transient transfection of APEX2 fusion proteins, leading to overexpression and high enough levels of biotinylation by APEX2 to overcome any background signal from endogenously biotinylated proteins. Enrichment of Coilin-APEX2 biotinylated proteins without nuclear isolation led to an abundance of carboxylases and mitochondrial proteins as top hits (data not shown). Because CBs are in the nucleus and

coilin expression had to be tightly regulated to avoid aberrant CBs, the addition of a nuclear isolation step is essential to avoid background signal from endogenously biotinylated proteins.

The higher and more enriched proteins identified using APEX2 have previously been shown to have longer residence times in the CB than other components. For example, coilin and TGS1 both have similar residence times in the Cajal body and are among the top hits in the enriched in coilin-APEX2. Similarly, snRNP proteins and SART3 have previously been shown to have shorter residence times in CBs than coilin and show less enrichment/significance in our results. One of the only new hits to show strong enrichment against both APEX2-NLS and dNTD-APEX2 was IRF2BP1. IRF2BP1 is an understudied protein originally identified as an interactor of Interferon Regulatory Factor 2 (IRF2) (Childs, 2003). Its localization has not been previously studied and this is the first study demonstrating its presence in CBs. IRF2BP1 has zinc finger and C3HC4 RING domains at the N- and C-terminus, respectively, and has been shown to inhibit transcription of various genes (Barysch et al., 2021; Faresse et al., 2008; Kimura, 2008; Yeung et al., 2011). I did not test the function of IRF2BP1 in this study, but due to its role as a transcriptional regulator I speculate that it may be acting as a transcriptional regulator of CB associated genes.

Forty-five out of the 100 proteins enriched in Coilin-APEX2 are involved in mRNA processing. Many of these are involved in transcriptional termination and mRNA cleavage. Histone RNA processing and cleavage proteins have mainly been found in CBs in cell lines and organisms where CBs and HLBs are the same body. However, the proteins identified here (CLP1, NUDT21, and CPSF6) are not part of the histone 3' end processing complex. It is possible that mRNA termination and cleavage occurs in CBs. CBs have previously been shown to form near actively transcribing snRNA gene loci. The presence of cleavage factors in the CB may be due to transcription and processing of snRNA genes actively occurring in the CB.

My results are the first to show multiple ribosomal proteins components of CBs. The only prior record of a ribosomal protein localized in CBs came shortly after the discovery of coilin as a CB marker. An immunofluorescent stain of ribosomal protein S6 and coilin in HeLa cells shows light S6 signal in CBs, and that rRNA did not co-localize with S6 in CBs, suggesting that it is not present as a complex in CBs (Jiménez-García et al., 1994). I do not detect S6 in my hits; out of the eight ribosomal proteins detected, two are from the small 40S subunit (RPS2 and RPS24) while the rest are components of the large 60S subunit (RPL proteins). These ribosomal proteins are short and disordered outside of the ribosomal subunit, suggesting that they may co-localize with Cajal bodies through LLPS interactions, and not as entire subunits. While nonribosomal nucleolar proteins frequently associate with CBs and can even be integral to their assembly, ribosomal subunits have not previously been shown to reside in Cajal bodies (Trinkle-Mulcahy & Sleeman, 2017). This may be because their residence time in the CB is short-lived. It has previously been shown that nucleolar components that also localize in CBs have a shorter residence time in CBs than snRNPs and other CB specific components. Additionally, their residence time in the nucleolus is longer than in the CB. Ribosomal proteins may also transit between nucleoli and CBs, but their residence time may be short enough that they have not previously been detected by other methods. More recently, ribosomes have been shown to associate with SMN (Fabio Lauria et al., 2020), it is possible that this association with SMN also occurs inside the nucleus and can be detected here.

Recently, a published study used APEX2 to identify nuclear body associated transcripts using RNA-sequencing and nuclear body proteins using mass spectrometry (Barutcu et al., 2022). This study largely focused on nuclear speckles but also included other nuclear bodies, including CBs. They used SMN and Wrap53 to target CBs, and were unable to get a Coilin-APEX2 cell line that properly localized to CBs. The results of their mass spectrometry experiments are not yet publicly available but will be interesting to compare to my list of CB components. Because of the distinct substructure between SMN and Coilin in CBs, it is possible that their list will reveal new CB components that are only present in the SMN containing subunit of CBs. This would indicate that Coilin-APEX2 is

specific for only the Coilin containing subunit of CBs. Additionally, their results using Wrap53-APEX2 may reveal more about the interplay between CBs and telomeres, as Wrap53 is a telomerase component. By using multiple CB markers to identify new components, we can obtain a more robust understanding of CB components and where they reside within CBs.

This study was done in HeLa cells and establishes the use of proximity biotinylation for identification of nuclear body components. This approach can now be expanded for use in other cell lines, tissue types, and organisms where CB composition and function may vary. For example, APEX2 biotinylation has been successfully established in developing zebrafish embryos and we know that CB number changes in early zebrafish embryogenesis (Ariotti et al., 2015; Strzelecka, Oates, et al., 2010). Using this approach, we can discover differences in CB components before and after MZT in zebrafish embryos. We could speculate that CBs prior to transcription initiation in embryos do not have all CB components (especially transcription related ones) and are instead proto-CBs (much like proto-HLBs and proto-nucleoli in early embryogenesis) (Arias Escayola & Neugebauer, 2018). Additionally, later developmental stages in zebrafish can be used to look at CB components in different tissue types. Our lab has previously shown CBs in zebrafish neuron and muscle tissue, but it is unknown whether these CBs have the same components and function in different tissue types (Strzelecka, Oates, et al., 2010). Furthermore, cell type specific APEX2 proximity labeling has recently been achieved in a mouse model (Dumrongprechachan et al., 2021), and this could also be expanded to look at tissue specificity of CBs as well as changes in CB composition in disease states such as SMA. Because Coilin-APEX2 biotinylation is performed prior to any fixation or harvesting of cells, it can also be used to analyze changes to CBs in perturbed cells. For example, our lab has seen changes in CBs after induction of different stresses (i.e. heat shock, osmotic shock). Using this technique, we can ask how CB proteins change during stress without having to probe each protein individually. This study not only provides a list of proteins to further study in CB function, but also demonstrates the use of a novel technique to better understand CBs in different contexts.

3 Cajal Body specific siRNA screen reveals RPLs as regulators of Cajal Bodies

3.1 Author Contributions

This work was done in collaboration with the Yale Center for Molecular Discovery (YCMD) and Korinna Straube, I used the YCMD facilities and performed all immunostaining for the siRNA screen and the high-content imaging was performed by Yulia Surovtseva and Laura Abriola. I designed the image analysis pipeline and performed analyses myself. Lisa Ogawa and Carson Bryant from the Baserga lab provided valuable input for performing the siRNA screen and analyzing data. I performed all other microscopy experiments and analyses in this chapter. Korinna prepared all samples for the ChIP-seq experiments and I analyzed the ChIP-seq data.

3.2 Goals and Approach

Studies on CB assembly have largely targeted two canonical CB proteins: coilin and SMN. Both proteins are structurally required for the assembly of CBs. Other known regulatory CB proteins are few, and the overall effect of protein-loss on CB integrity can range from partial loss of CBs (reduced CB number) to complete obliteration of CBs in cells (Arias Escayola & Neugebauer, 2018; Sawyer et al., 2017). Much like the discovery of CB proteins, there have not been any comprehensive studies on CB assembly.

CB assembly occurs in a densely packed nucleus, yet is specific for particular proteins, RNAs, and genomic loci. CB assembly in the nucleus requires many separate molecular events, including spatiotemporal regulation throughout the cell cycle, nucleation

of CBs by transcription of non-coding RNAs, and oligomerization of structural components (Arias Escayola & Neugebauer, 2018; Sawyer et al., 2017). Current understanding of CB assembly is highly dependent on characterizing singular CB components and their role in regulating CB assembly. For example, CB substructure is regulated by interactions between SMN and coilin. SMN and coilin occupy different CB subdomains with an interacting coilin-SMN interface (Courchaine et al., 2021; Novotný et al., 2011). Our lab recently characterized the formation of CBs by studying the phase separation characteristics of coilin and SMN domains. This study also revealed that DMA modifications can alter CB substructure, but the DMA ligands responsible for this regulation are still uncharacterized. Studies of individual CB protein domains such as this one are important for understanding the mechanistic details of CB assembly, but larger screens can reveal more targets.

Previous imaging-based siRNA screens have addressed regulation of coilin and SMN nuclear bodies by phosphatases and kinases. In a study from Lucas Pelkmans lab, regulators of six membrane-less organelles (MLOs) were identified using an siRNA screen (Berchtold et al., 2018). However, this screen focused on knockdown of kinases and phosphatases to identify pathways that regulate MLO assembly. This study revealed connections between regulation of nucleoli and CBs, as well as a link between increased CB number and increased cell area, demonstrating that CBs can be regulated through separate signaling pathways. Two separate screens have looked at SMN condensation after knocking down phosphatases and kinases and identified SMN phosphorylation sites that regulate its assembly in CBs (Husedzinovic et al., 2015; Schilling et al., 2021). These screens have been identified PTMs on SMN that regulate its localization to CB. There are

few phosphatases and kinases that have been identified as CB proteins, the enzymes identified in past screens regulate CB assembly but are not localized to the CB. In this chapter, I outline a CB specific siRNA screen to identify regulators of CB assembly.

3.3 siRNA screen of CB components reveals 46 CB regulators

To analyze the importance of each CB protein on the assembly and/or maintenance of CBs, I established a CB-specific siRNA screen to look at changes in CB number. I compiled a list of 144 total CB proteins – including the 70 new proteins identified by mass spectrometry – and generated a CB specific siRNA library. I then "reverse transfected" HeLa cells to knockdown CB proteins and assayed for changes in CB morphology by staining with an anti-coilin antibody to quantify the number of CBs per nucleus (Figure 3.1).

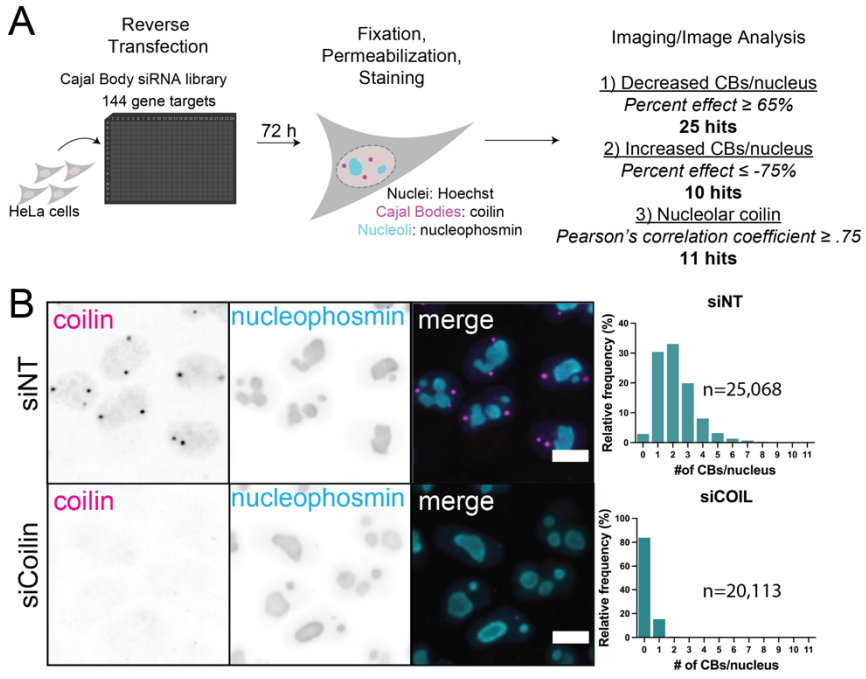


Figure 3.1 siRNA screen of CB components identifies 46 CB regulators
A) Schematic of microscopy-based CB specific siRNA screen B) Representative images of control conditions (siNT and siCoilin). Immunofluorescent images show anti-coilin antibody in magenta and anti-nucleophosmin antibody in cyan. Histograms for each siRNA condition quantifying the number of CBs per nucleus from their respective well and plate in the screen (12 fields of view/well with Representative hit is shown in teal with siNT in light gray

Coilin has previously been shown to interact with nucleoli under stress and in disease states (Trinkle-Mulcahy & Sleeman, 2017). To analyze the relationship between CBs and nucleoli, I also stained for nucleoli with an antibodyspecific for nucleophosmin. For each sample, I calculated the average number of CBs per nucleus and the intensity of coilin staining in the nucleolus. Non-targeting siRNA (siNT) and coilin siRNA (siCOIL) were used as negative and positive controls, respectively, for inhibition of CB assembly (Figure 3.1). In total, 46 proteins (32% of CB proteins) were found to affect CB formation (Table 4).

Table 4. List of proteins altering CB number and coilin localization

Change in CB number				**Change in Coilin localization**	
Decrease in CBs	**Percent Effect**	**Increase in CBs**	**Percent Effect**	**Nucleolar Coilin**	**Pearson's R**
NPAT	120	ANLN	-179	MED14	0.87
SLBP	115	RPL14	-95	SNRPG	0.84
CASP8AP2	104	RPL24	-94	POLR2B	0.84
POLR2A	101	NUMA1	-90	POLR2A	0.84
COIL	100	RPL13	-80	POLR2C	0.83
SNRPD3	99	RPL8	-77	ICE1	0.83
SNU13	98	NSA2	-76	USPL1	0.82
SNRPB	94	EFTUD2	-75	CYB5R1	0.81
SMN1	94	RPL7A	-75	UBR5	0.78
IK	93			MED17	0.77
TRIM22	91			SNRPD1	0.76
SNRPD2	90				
NOLC1	90				
SUPT6H	86				
NOP58	83				
CDC73	77				
CSNK2B	77				
SNRPF	74				
PRPF4	68				
TGS1	66				
EAF1	66				
SNRPG	65				
SON	65				

I measured re-localization of coilin to nucleoli by measuring the Pearson's Correlation Coefficient between the nucleophosmin and coilin intensities. I identified 11 siRNAs that caused coilin to relocalize to nucleophosmin and prevented assembly of CBs (Figure 3.2).

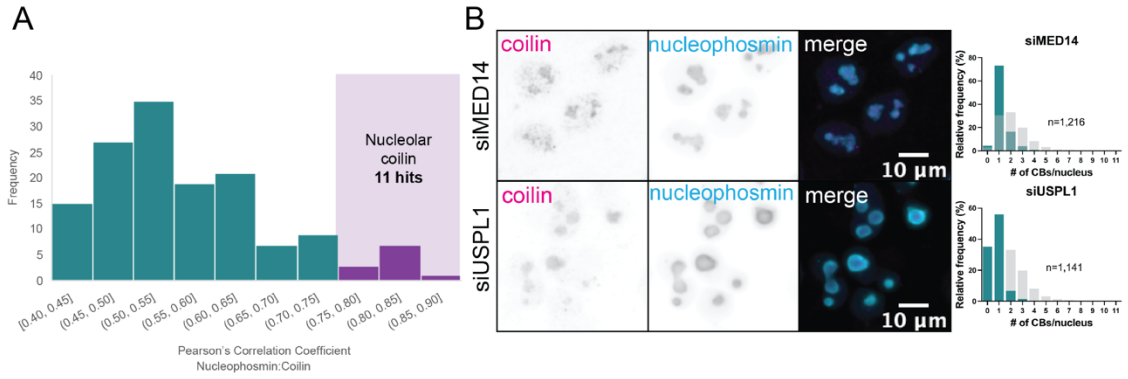


Figure 3.2 11 siRNAs cause re-localization of coilin to nucleoli upon KD
A) Frequency histogram of Pearson's Correlation Coefficient between Nucleophosmin and Coilin fluorescent signal per siRNA in screen. Values of siRNAs that re-localize coilin to nucleoli are highlighted in purple (Pearson's > .75). B) Representative images of hits showing re-localization of coilin to nucleoli (siMED14 and siUSPL1). Immunofluorescent images show anti-coilin antibody in magenta and anti-nucleophosmin antibody in cyan. Histograms for each hit quantifying the number of CBs per nucleus from their respective well and plate in the screen (12 fields of view).

I calculated percent effect (PE) of CB knockdown by setting the average CB/nucleus of siNT to 0 and siCOIL to 100. By these measures, 35 siRNAs regulate the number of CBs in a cell (Figure 3.3A). Specifically, 25 of these hits decreased the number of CBs per nucleus (PE > 60). Of the newly identified CB proteins, SPT6 and TRIM22 had the strongest effect in reducing CB number (Figure 3.3B). Surprisingly, I detected 10 hits that showed a negative percent effect (PE < -60), indicating an increase in CB number (Figure 3.3A).

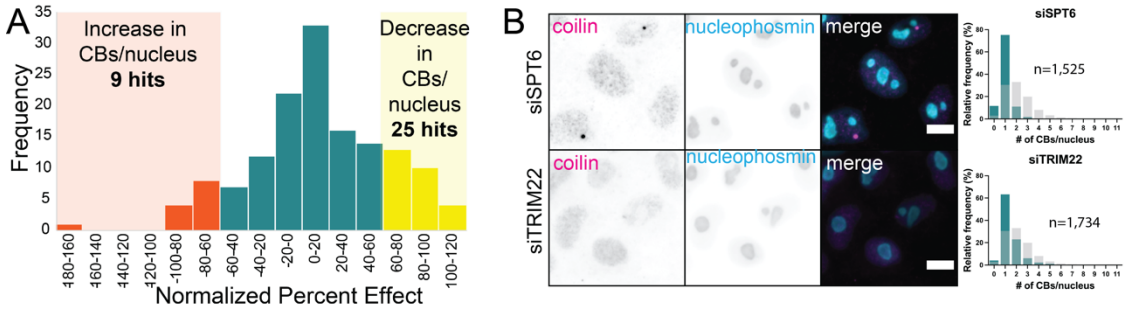


Figure 3.3 siRNA screen of CB components identifies 25 proteins that decrease CB count
Frequency histogram of Percent Effect per siRNA in screen. Values of siRNAs that decreased the number of CBs per nucleus are highlighted yellow (PE>60). Values of siRNAs that increased the number of CBs per nucleus are highlighted in orange (PE<-60). C) Representative images of hits showing a decrease in CBs per nucleus (siSPT6 and siTRIM22). Immunofluorescent images show anti-coilin antibody in magenta and anti-nucleophosmin antibody in cyan. Histograms for each hit quantifying the number of CBs per nucleus from their respective well and plate in the screen (12 fields of view).

3.4 Depletion of 60S ribosomal proteins increase CB numbers and change their structure

Representative examples of the top three hits increasing CB number are shown in Figure 3.4 (ANLN, RPL14, and RPL24). Of the ten hits that showed an increase in CB number, five were ribosomal proteins found in the large 60S subunit and two are involved in biogenesis of the 60S subunit (Table 4). All of these proteins were new hits from the APEX2 dataset. Thus, I wanted to further investigate the role of these ribosomal proteins in CB assembly and set out to characterize these increased CBs.

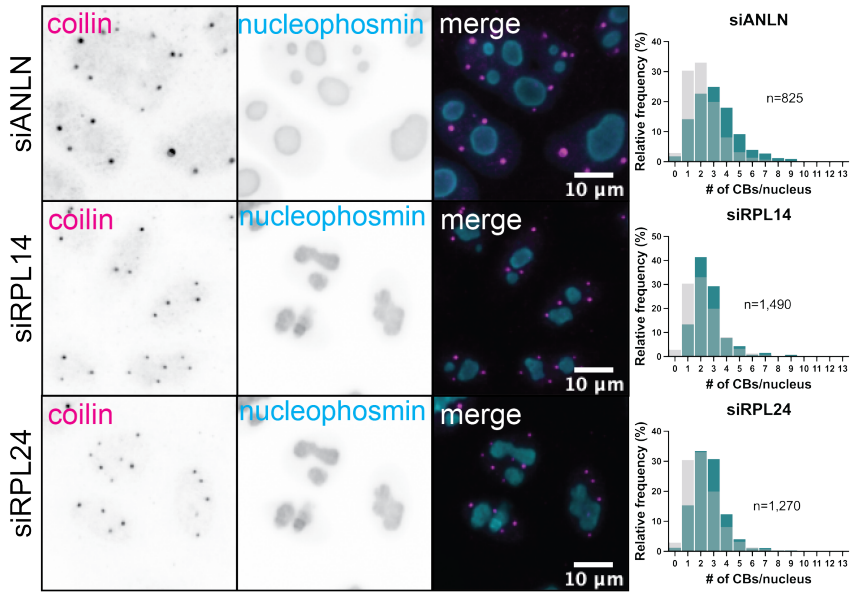


Figure 3.4 Top three siRNA hits that increase CBs/nucleus
A) Representative images of top three hits showing an increase in CBs per nucleus (siANLN, siRPL14, and siRPL24). Immunofluorescent images show anti-coilin antibody in magenta and anti-nucleophosmin antibody in cyan. Histograms for each hit quantifying the number of CBs per nucleus from their respective well and plate in the screen (12 fields of view).

Depletion of ribosomal proteins can lead to defects in cell cycle progression and cell ploidy (Bhavsar et al., 2010; Warner & McIntosh, 2009).To test whether changes in cell ploidy were causing the increase in CBs, I used the Hoechst stain from the siRNA screen to analyze the DNA content of all proteins that increased CB count. Anillin, one of the non-ribosomal proteins that increased CBs affected cytokinesis during the cell cycle and led to an accumulation of cells with 4DN genomes (Figure 3.4 and Figure 3.5). However, none of the ribosomal proteins analyzed showed an effect in cell ploidy or cells in different cell cycle stages (

Table 5 and Figure 3.5)

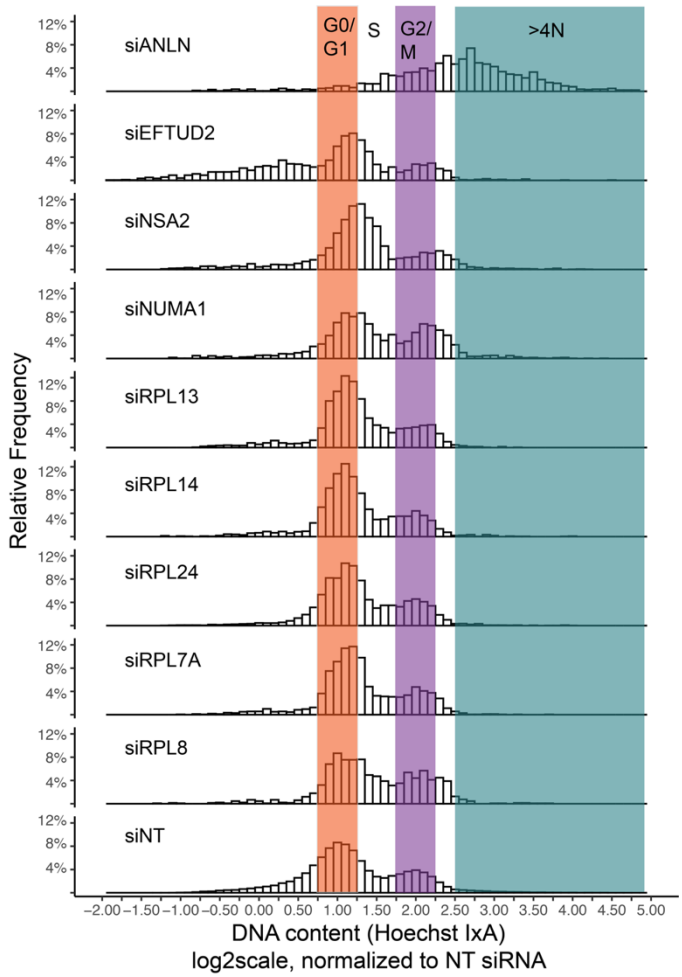


Figure 3.5 RPL knockdown does not affect DNA content in HeLa cells
DNA content analysis. Representative histograms of Hoechst log2 integrated intensity in siNT, siANLN, and siRPL14 normalized to siNT control. G0/G1 phase (0.75-1.25) highlighted and quantified in orange, S phase (1.25-1.75) in white/black font, G2/M phase (1.75-2.25) highlighted and quantified in yellow, >4N (>2.5) highlighted and quantified in teal.

Table 5 Cell Cycle Analysis of siRNAs that increase CB number

siRNA	G0/G1	S	G2/M	>4N
NT	44.3	18.5	18.7	5
ANLN	3.9	3.8	18.3	58.7
EFTUD2	42.9	20.9	17.9	2.6
NSA2	38.8	29.5	14.4	4.7
NUMA1	32.6	18.9	25.5	6.5
RPL13	52.5	18.2	20.1	1.3
RPL14	54.8	15.9	20.8	1.8
RPL24	49.5	15.9	22.7	1.9
RPL7A	50.2	18.6	22	1.1
RPL8	38.9	17.6	26.6	2.7

To validate the identity of the increased coilin foci as CBs, I stained for common CB components (snRNPs, Nopp140, and dyskerin) upon knockdown of the top two ribosomal hits (RPL14 and RPL24). I found that the more numerous CBs in RPL KDs still contain Nopp140 and dyskerin, components that bind snoRNPs and shuttle between CBs and the nucleolus (Figure 3.6A and B). While snRNPs still co-localize with coilin when RPLs are depleted, this co-localization is less distinct and snRNPs show stronger nucleoplasmic signal (Figure 3.7). This suggests that the proportion of snRNPs that concentrate in CBs is decreased when ribosomal proteins are depleted.

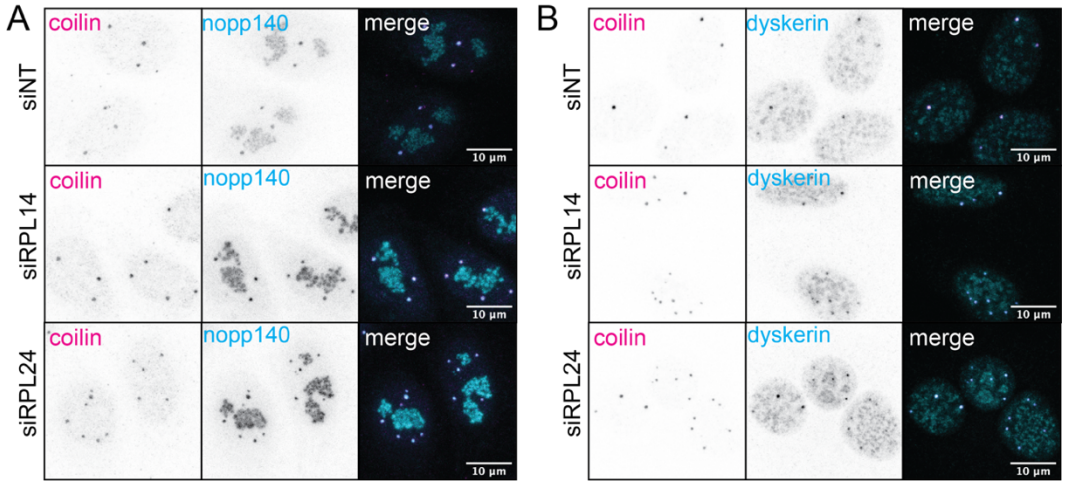


Figure 3.6 Nucleolar CB components remain in CBs after RPL knockdown
Immunofluorescent images showing the presence of Nopp140 (A) and dyskerin (B) (cyan) in CBs (anti-coilin antibody in magenta).

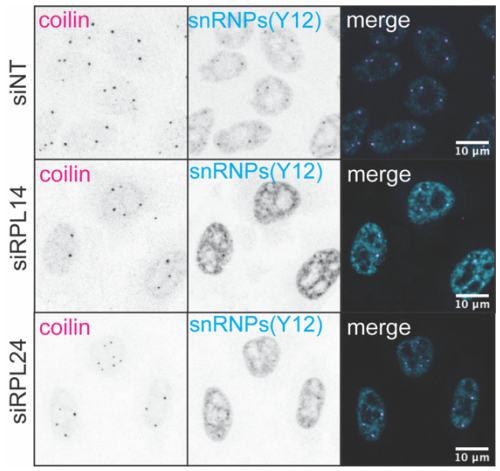


Figure 3.7 snRNP residence in CBs is affected by knockdown of RPLs
Immunofluorescent images showing the presence of snRNPS (anti-Sm Y12 antibody in cyan) in CBs (anti-coilin antibody in magenta).

I then wanted to assess whether CB substructure was altered in RPL knockdowns. To do this, I immunostained for SMN and coilin in ribosomal protein knockdowns. Surprisingly, SMN completely overlaps with coilin when RPL14 and RPL24 are knocked down, abolishing the distinct coilin/SMN substructure that normally characterizes the CB (Figure 3.8A). I validated this using STED microscopy to quantify the offset between coilin and SMN foci as well as the percent overlap in the two intensities between the knockdowns (Figure 3.8B). The distance between the centers of mass in coilin and SMN decreased in both RPL14 and RPL24 knockdowns (Figure 3.8C). Additionally, the percent of coilin overlapping with SMN increased in these same knockdowns (Figure 3.8D). These data indicate that CB substructure is abolished upon RPL KD without affecting the overall makeup of CBs.

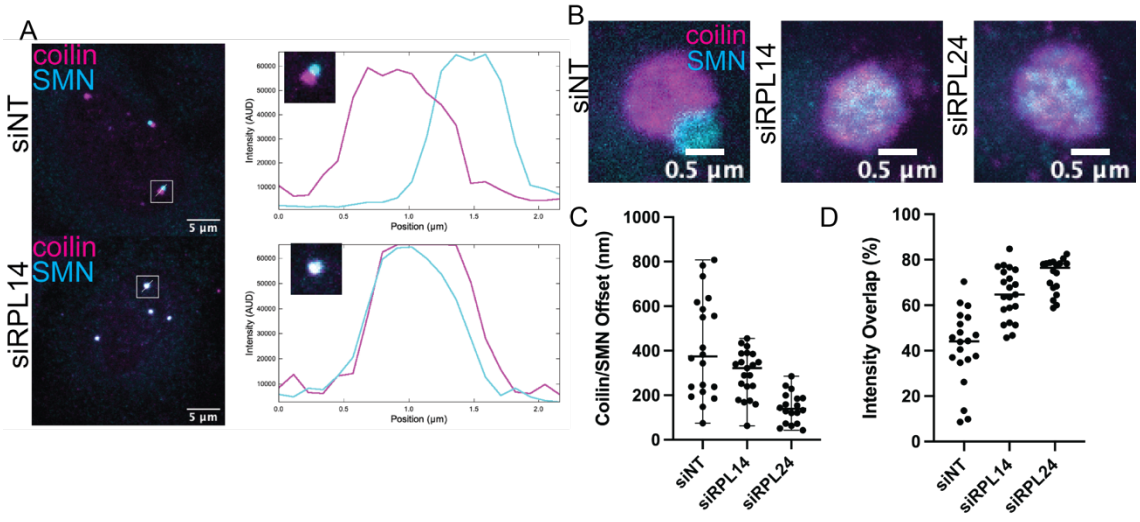


Figure 3.8 A subset of RPL proteins alter CB structure upon KD
A) Representative single nucleus immunofluorescent staining of coilin (magenta) and SMN (cyan) in siNT and siRPL14 with accompanying line profile plots showing intensities through a single CB. B) Representative single CB immunofluorescent staining of coilin (magenta) and SMN (cyan) using STED microscopy in siNT, siRPL14, and siRPL24 cells. C) Offset between coilin and SMN fluorescent intensity-weighted center of mass for CBs from siNT, siRPL14, and siRPL24 cells (n= 20 CBs per condition). D) Percent of intensity weighted overlap between coilin and SMN fluorescence for CBs from siNT, siRPL14, and siRPL24 cells (n= 20 CBs per condition).

Because KD of ribosomal subunits can lead to defects in protein synthesis and/or ribosome biogenesis, I inhibited these two processes and assayed for the increased CB/merged CB subunit phenotype. To inhibit mRNA translation, I treated cells with 50 µm of rapamycin, which blocks mTOR signaling and thus prevents translation (Nandagopal & Roux, 2015). To inhibit ribosome biogenesis, I treated cells with BMH-21, a small molecule that inhibits ribosome biogenesis by directly inhibiting PolI mediated transcription of rDNA (T. Wei et al., 2018). After inhibition with either rapamycin or BMH-21, there was no effect on CB morphology (Figure 3.9). This suggests that the effect of RPL KD on CB morphology is specific and not due to larger processes being disrupted.

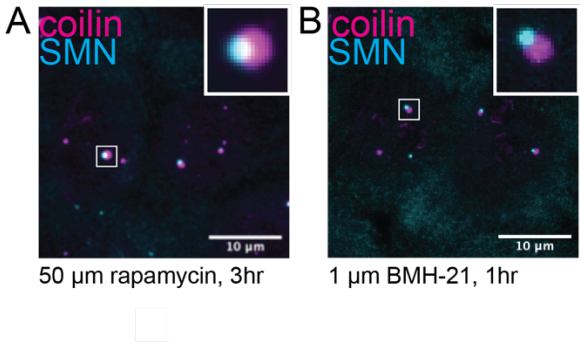


Figure 3.9 Inhibition of protein synthesis and ribosome biogenesis does not dirupt CB morphology
A) Representative immunofluorescent staining of coilin (magenta) and SMN (cyan) after inhibition of protein signaling with 50 µm rapamycin for three hours. B) Representative immunofluorescent staining of coilin (magenta) and SMN (cyan) after inhibition of ribosome biogenesis with 1 µm BMH-21 for one hour.

CBs are known to form on actively transcribing histone and snRNA loci. Our lab has used chromatin immunoprecipitation followed by next generation sequencing (ChIPseq) to identify DNA loci associated with CBs by performing coilin-ChIPseq and found that the most prominent genes identified were histone and snRNA genes (Machyna et al., 2014).

To further understand the environmental context of the increased CB phenotype, I performed coilin ChIP-seq in RPL14 and RPL24 knockdowns. In the non-targeting control, coilin-IP revealed significant peaks at snRNA and histone gene clusters. As expected, peaks were highest along histone gene bodies and enrichment at snRNA genes was subtle, but still significant to be called by the peak caller MACS2. This signal was greatly reduced in RPL14 and RPL24 knockdowns (Figure 3.10A and B). The peaks caller did not detect any peaks along snRNA or histone gene loci in the knockdown conditions. Peaks called in the knockdown conditions did not correspond to gene bodies. In fact, most peaks called in the knockdown conditions were eliminated after comparing to "blacklist" regions of the genome, demonstrating that these peaks were background artifact. Taken together, these data show that CBs in RPL KDs do not form on genomic loci. As a proxy for transcription of these genes, I also performed Pol II ChIP to measure binding of Pol II at snRNA and histone loci. As expected, there were strong peaks along all snRNA and histone genes. Upon knockdown of RPL14 and RPL24, the intensity of these peaks was diminished >10-fold (Figure 3.10). This was also true for coding genes, demonstrating an overall decrease in Pol II binding to chromatin upon RPL KD. This suggests that Pol II transcription is downregulated upon RPL KD, preventing assembly of CBs on chromatin.

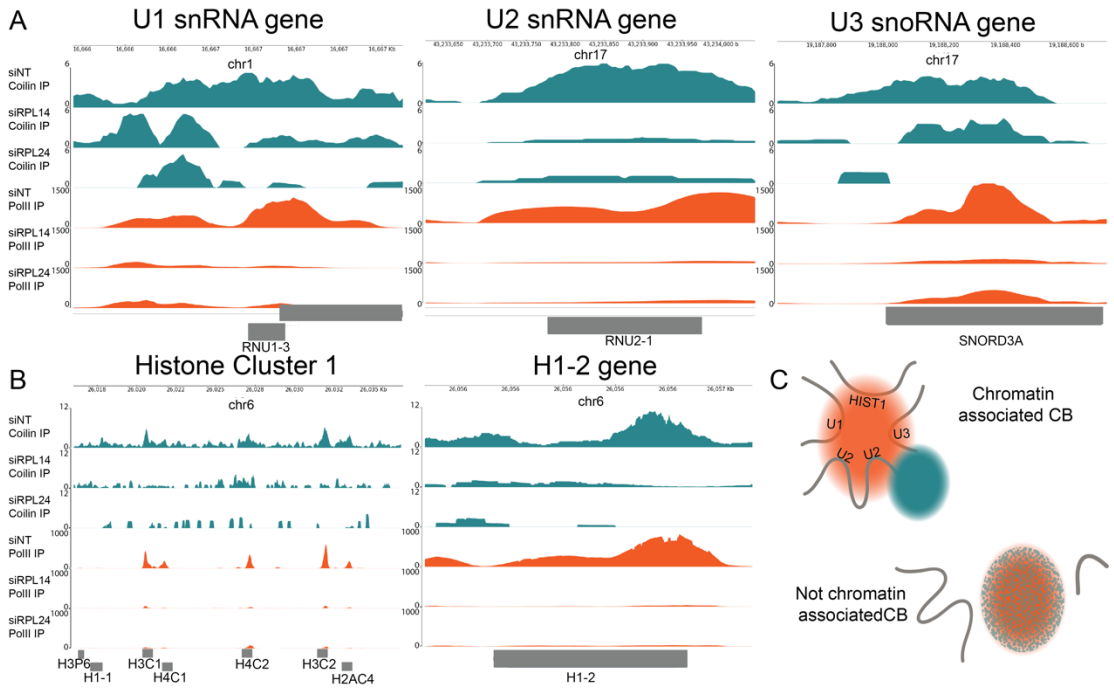


Figure 3.10 Coilin and Pol II peaks are diminished along snRNA and histone gene loci

Genome browser tracks of Coilin-ChIP and fold enrichment (FE) at snRNA and histone gene loci for siNT, siRPL14, and siRPL24 cells. Coilin-IP signal is shown in teal while PolII signal is in orange. Note different scales for Coilin-ChIP and PolII-ChIP. A) Genome browser tracks of U1 (RNU1-3), U2 (RNU2-1), and U3 (SNORD3A) snRNA and snoRNA genes. B) Genome browser tracks of histone gene cluster 1 (HIST1) on human chromosome 6 (left) and individual histone 1 gene (H1-2, right). C) Model figure of chromatin associated CB on snRNA and histone genes with Coilin subunit in orange and SMN subunit in teal losing association with chromatin when Coilin and SMN subunits merge.

3.5 Discussion

Using a CB-specific siRNA imaging screen, I identified novel regulators of CB assembly. Initially, I expected that this screen would reveal proteins that were required for CB assembly, and that knockdown of these proteins would result in a decrease in CB number. Surprisingly, the screen itself revealed two other phenotypes that identify proteins that regulate CB assembly. Additionally, known regulators of CB assembly were hits in the screen, further validating this approach for studying CB assembly.

The top three hits for decreased CBs per nucleus (*NPAT, SLBP, CASP8AP2*) are canonical HLB proteins. This is likely because CBs and HLBs are the same body in HeLa cells. It is, however, surprising that there are no residual coilin foci upon knockdown of these proteins, as HLBs and CBs are separate entities in other cell lines and organisms. These data indicate that CBs in HeLa cells require HLB proteins, and their assembly is potentially dependent on proper histone RNA processing. Most of the other hits that decrease the number of CBs per nucleus have already previously been identified in the literature (i.e. snRNP proteins, TGS1, Nopp140). Of the new hits, most are involved in transcription, further establishing the relationship between CB assembly and transcription. One surprising hit, *CSNK2B*, is a novel component identified by coilin-APEX. This gene encodes for the beta subunit of Casein Kinase 2 (CK2). It has not previously been shown to reside in CBs yet has been identified as a regulator of two structural components of CBs, coilin and Nopp140. CK2 phosphorylates Nopp140, and the interaction between the beta subunit and Nopp140 has been extensively characterized (Lee et al., 2013; D. Li et al., 1997; Na et al., 2016). The interaction between coilin and CK2 is not well studied, though CK2 has been shown to phosphorylate coilin *in vitro* (Hebert & Matera, 2000). This screen is the first to identify a direct effect of CK2 on CB assembly. One recent study has shown that phosphorylation of Nopp140 by CK2 is necessary for localization of Nopp140 to CBs (Bizarro et al., 2021). Nopp140 is an essential structural component of CBs, KD of Nopp140 abolishes CBs (this screen and unpublished data). Phosphorylation of coilin also dictates when it forms CBs throughout the cell cycle (Hearst et al., 2009). It is possible that CK2 regulates CB assembly by altering phosphorylation of Nopp140, and possibly coilin itself. CK2 is a ubiquitous kinase, and its effect on CB assembly may be further upstream

than its interaction with Nopp140 and coilin. Further studies must be done to better understand phosphorylation of these proteins by CK2 to regulate CBs.

The nucleolar phenotype was identified because of the parameters established for counting CBs in my analysis pipeline. Because CBs were identified as increased areas of fluorescent intensity over the nucleoplasm, coilin that localized to nucleoli was counted as a CB initially. This led to the pipeline calling nucleoplasmic coilin as hits with >10fold increase in CB number. Further analysis of these wells revealed that there was no increase in CB number, but a relocalization of coilin. I adjusted the analysis pipeline to include these as separate phenotypes by measuring the correlation between the fluorescence intensity in the nucleophosmin channel with the coilin channel. Of the 11 hits that caused relocalization of coilin to the nucleolus, 8 are directly involved in transcription and two are Sm subunits of snRNPs. This indicates that transcription is necessary for coilin to form CBs and without it, coilin can be mislocalized. It is not clear why coilin re-localizes to nucleoli. It has previously been shown that the hypomethylation state of coilin regulates its localization, and that lack of methylation leads to coilin accumulating in nucleoli Tapia (Tapia et al., 2010). However, my screen did not identify any direct regulators of coilin methylation as nucleolar coilin hits. The coilin methylation study suggests that PTMs on coilin determine whether it localizes to nucleoli or CBs, and that this implies a constant flux of coilin between CBs and nucleoli. However, coilin and CBs are usually excluded from nucleoli unless in diseased or stress states. It is possible that it is not the methylation of coilin itself that matters, but that cells that lack the ability to properly methylate proteins are under stress, leading to relocalizstion of nuclear body components such as coilin. The results from this siRNA screen support that, as cells lacking essential proteins involved in transcription

and splicing are in a constant state of stress. Localization of coilin to nucleoli in this case may be due to either interactions between coilin and nucleolar proteins being stronger when coilin cannot nucleate on actively transcribing loci, or that the intrinsic disordered regions of coilin make it more likely to go to another phase separated body when lacking CB specific interactions.

The enrichment of RPL proteins in the increased CB phenotype was surprising. As discussed in Chapter 2, RPL proteins had not been identified in Cajal bodies previously. Two small subunit proteins were included in this screen, yet neither had a significant effect on CB number in cells. Of the six RPL proteins assayed, only RPL15 did not show an effect on CB number upon knockdown. Within the 60S subunit, RPL15 is the only of the RPL proteins assayed that is not found on the surface (see Appendix Figure 7.1). This suggests that the phenotype seen here may be due to loss of some interaction with the ribosomal subunit along its surface, though more careful experiments are required. Another hit for increased CB assembly TINP1 - a protein encoded by *NSA2* gene- is a ribosome assembly factor that binds pre-60S particles and is released from the 60S subunit prior to nuclear export (W. Li et al., 2013; Paternoga et al., 2020). Knockdown of NuMA protein has previously been shown to affect rDNA transcription as well as pre-rRNA processing (Farley-Barnes et al., 2018; Jayaraman et al., 2017). The effect of TINP1 and NuMA knockdown on CB number would suggest that ribosome biogenesis plays a role in the phenotype observed. This is not supported by my data inhibiting ribosome biogenesis using BMH-21. It is possible that this is due to the short treatment time. The siRNA knockdown occurred over 72 hours while BMH-21 treatment was only one hour long. However, in this one hour treatment, I could already observe some coilin re-localization to the nucleolus,

suggesting that longer treatment time would result in a nucleolar coilin phenotype and not increased CBs. It seems likely that the increased CB phenotype is not due to inhibition of a larger processing pathway, but instead due to the loss of specific protein interactions. NuMA protein has also been found to bind RPL24-one of the top increased CB hits- in mammalian cells (Jayaraman et al., 2017). It is possible that this interaction is somehow necessary for proper CB assembly, as both proteins have the same effect on CB assembly. One study detected NuMA protein in CBs following changes in transcription during lens cell differentiation (Gribbon et al., 2002). In transcriptionally inactive epithelial cells, NuMA could be found in some CBs. Upon upregulation of transcription in differentiating cells, CB number increased and NuMA no longer localized to these increased CB foci. The data for this interaction is sparse but supports the results from this screen. It will be crucial to further probe the interaction between CB proteins and ribosomal proteins to better understand the role of ribosomal proteins in regulating CB assembly.

The changes in CB composition shown in the RPL KDs suggest that protein interactions within the CB are being regulated following KD. It was surprising to find that snoRNP components Nopp140 and dyskerin were still present in the coilin foci after RPL KD, but snRNPs were greatly reduced. This suggests that RPL KD is affecting a specific subset of CB proteins. Because both snRNP and SMN components are affected upon RPL KD, this would suggest that components involved in snRNP assembly, but not snoRNP/scaRNP maturation are being regulated by RPL proteins. This is surprising, as these are components that are *not* shared with the nucleolus, the site of ribosome biogenesis. This further supports the idea that the phenotype is not related to ribosome biogenesis. The merging of the coilin and SMN containing domains of CBs was previously

seen upon inhibition of asymmetric dimethylation (aDMA) in HeLa cells (Courchaine et al., 2021). Surprisingly, inhibition of symmetric dimethylation (sDMA) has the opposite effect, with both subunits completely separating and losing the coilin/SMN interface in CBs. This implies an equilibrium between aDMA and sDMA modifications in HeLa cells that regulates CB composition. Because RPL KD causes merging of coilin and SMN domains in CBs, it is possible that RPLs regulate CB composition via regulation of DMA modifications. SMN binds DMA modifications through its tudor domain. In CBs, the SMN tudor domain binds sDMAs on coilin. These sDMA modifications are necessary for recruitment of SMN to CBs. RPL KD mimics the phenotype seen upon inhibition of asymmetric dimethylation. Although SMN tudor domain has a high affinity for sDMA (K_d = 0.476mM) it also recognizes aDMA (K_d = 1.025mM) (Tripsianes et al., 2011). RPL KD may shift the availability of specific sDMA and aDMA modifications, leading to a higher affinity between coilin sDMA and SMN that causes merging of CB subdomains. Because Sm proteins are reduced in the CB upon RPL KD and the SMN binds, it is possible that Sm proteins and SMN are competing for binding on coilin. snRNP Sm proteins contain sDMA modifications that are also bound by SMN during snRNP assembly. Although the interactions between SMN/coilin, SMN/Sm, and coilin/Sm proteins have been characterized, there are no studies looking at competition amongst these proteins. While SMN binds coilin at its RG domain, this domain is not necessary for coilin binding Sm proteins (Xu et al., 2005). In fact, the methylation state of coilin is irrelevant on Sm binding to coilin. SMN and Sm proteins both bind coilin at its C-terminus, but they have distinct binding sites. Further studies of RPL KDs should address the methylation state of CB proteins, as well as binding amongst SMN/coilin/Sm proteins.

It is possible that other methylated proteins are the cause of the merged phenotype. A recent proteomic study showed that ribosomal proteins themselves are dimethylated, and mutations in their methylation sites can affect ribosome biogenesis, protein translation, and cell proliferation (H.-H. Wei et al., 2021). This study characterized the interactome of several protein arginine methyltransferases (PRMTs) and found high enrichment of ribosomal proteins. Of the proteins tested in my siRNA screen, all RPLs except for RPL7A and RPL24 were found to interact with PRMTs (see Appendix Table 1 for interacting PRMTs and methylation sites). RPL24 and RPL7A also contain predicted methylation sites, but these were not confirmed in this study. The interacting PRMTs include PRMTs from all three types of methyltransferases. SMN has recently been demonstrated to bind ribosomes, though the exact binding proteins are uncharacterized (F. Lauria et al., 2020). It is possible that binding of SMN to ribosomes is mediated through ribosomal DMA modifications, and that this interaction is also important to CB assembly in the nucleus. This supports the idea that DMA modified proteins regulate CB composition and indicates that RPLs specifically may play a role.

Recently, another link between SMN, ribosomal proteins, and protein translation was discovered as the result of an SMN specific siRNA microscopy screen. In this screen, knockdown of phosphatases revealed that a ribosomal kinase (RPSK6) regulated phosphorylation of SMN and SMN's ability to phase separate (Schilling et al., 2021). This study also found that mTOR signaling regulated SMN's ability to phase separate, and that inhibition of the mTOR pathway reduced the number of CBs with SMN. This is not supported by my data using rapamycin to look at CB morphology, however the treatment timeline for my experiment was much shorter than in their study. Surprisingly, they also

found that knockdown of RPSK6 led to activated mTOR being detected at SMN foci. It is unclear whether these are gems or CBs, as they did not stain for CBs.

CB assembly on chromosomal loci is dependent on transcription. I used coilin-ChIPseq to determine changes in the chromosomal environment of CBs upon RPL KD. I expected that the increased number of CBs would result in either an increase in signal at coilin-associated loci or the discovery of new coilin-associated loci. I was surprised that neither result proved true. Instead, coilin signal was greatly diminished in RPL KDs, suggesting that CBs no longer form on chromosomes under KD conditions (Figure 3.10C). Additionally, the diminished Pol II peaks along gene bodies suggest that DNA transcription is also downregulated in these cells. This suggests that knockdown of these ribosomal proteins can lead to changes in gene expression that prevent formation of CBs on chromosomes. Although these RPLs have not previously been shown to affect transcription, a recent study demonstrated that knockdown of ribosomal protein eL29 can cause changes in transcription without an effect on cell viability, ribosome biogenesis, or global translation (Gopanenko et al., 2020). The RPL proteins identified in this screen may also play a role in transcription outside of their main function in the ribosome. One question that remains is whether the change in composition of CBs are due to changes in the chromosomal environment of CBs, or vice versa. Though coilin has been shown to bind DNA *in vitro* SMN has not. Additionally, nuclear gems do not require transcription to form and are not known to form on any gene loci. It is possible that when the SMN and coilin subdomains of CBs merge, SMN prohibits binding of coilin to DNA. This question could be answered by superresolution microscopy in the future to look at DNA contacts of CB subdomains.

Using an siRNA screen to look at all CB components allowed for discovery of new CB assembly regulators. This siRNA screen can be adapted for use with other CB markers, measuring changes in CB components on a CB proteome-wide scale. Additionally, this screen could also be combined with RNA metabolic labeling to look for changes in global transcription levels that correlate with changes in CB number. This screen is the beginning of larger-scale studies to study CBs.

4 APEX-ChIP: A proximity biotinylation method to define the genomic landscape of nuclear bodies

4.1 Author Contributions

This chapter is a collaborative effort amongst multiple current and former members of the lab. I came up with the original concept for the method and wrote the grant that funded the project with Karla Neugebauer. I carried out the molecular cloning of constructs and generated all cell lines used. I worked closely with Korinna Straube to optimize the protocol used and she generated the final samples sent out for sequencing along with help from undergraduate Crystal Xu. Analyses of the sequencing data was carried out by Martin Machyna, a member of the Simon lab.

4.2 Goals and Approach

How nuclear bodies and compartments organize the genome is a central question in molecular cell biology. Current methods designed to analyze chromosome topology in relation to nuclear bodies (e.g. chromosomal contacts near or in nuclear speckles, nucleoli, and Cajal bodies) include PLAC-seq, TSA-seq and SPRITE[2–4]. These are "high-end" methods requiring significant reagent development, resources, training, and/or bioinformatic analysis of the data. Moreover, nuclear compartments are analyzed by proximity to a single protein marker in these assays, which runs the risk of detecting DNA interactions outside of the compartment itself. To overcome these limitations, this project leverages a commonly used, engineered peroxidase APEX2 fused to nuclear body proteins to specifically biotinylate and identify regions of the genome proximal to nuclear bodies

and compartments. The intersection of the data obtained with multiple protein markers of the same nuclear body provides specificity.

A significant portion of the genome is organized preferentially around nuclear bodies. I am interested in identifying chromosomal contacts that occur at nuclear bodies, such as nucleoli, nuclear speckles, and Cajal bodies, and understanding how these contacts change upon perturbation. Several excellent methods for determining 3D chromosomal organization within or near nuclear bodies have been developed. In general, these methods are technically challenging, laborious, and require significant reagent development, resources and optimization. APEX-ChIP, the method outlined in this chapter, is an alternative and complementary approach to current methods being used to map out chromosomal interactions within nuclei. In this chapter, I outline preliminary data using APEX-ChIP to look at nucleoli.

4.3 Generation of APEX-ChIP cell lines

To target APEX2 to nuclear bodies, I designed various constructs to tag nuclear body proteins with APEX2, a V5 linker, and a GFP tag for visualization of nuclear bodies (Figure 4.1). I chose two proteins per nuclear body so that the results from each protein could be combined to create a list of genomic loci interacting with nuclear bodies instead of one protein. Nopp140 is found in both CBs and nucleoli and should show peaks at loci associated with both bodies. I then transduced K562 cells with lentiviral constructs, sorted for GFP positive cells, and confirmed GFP-positive nuclear body formation under the microscope. These stable cell lines were then used to optimize biotinylation and streptavidin-ChIP for APEX-ChIP (Workflow outlined in Figure 4.2).

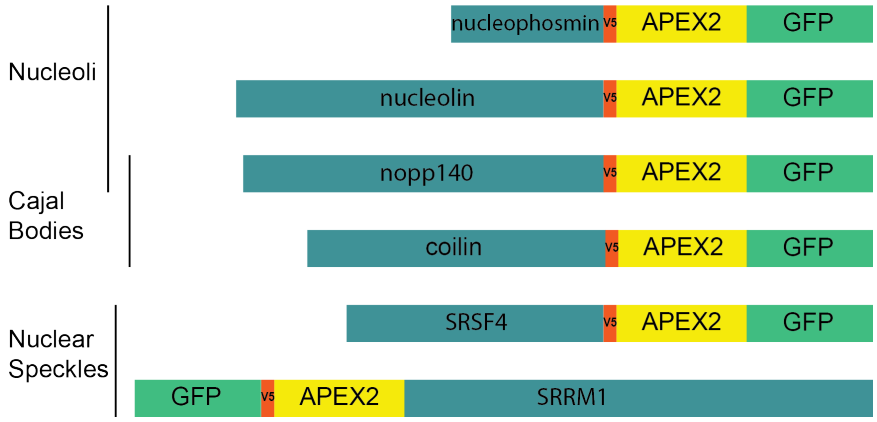


Figure 4.1 Schematic of APEX-ChIP constructs
Schematic of constructs used in APEX-ChIP. Nucleolar proteins include nucleophosmin, nucleolin, and Nopp140. CB proteins include coilin and Nopp140. Nuclear speckle proteins include SRSF4 and SRRM1

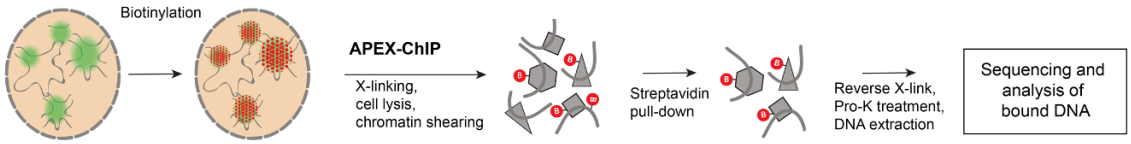


Figure 4.2 APEX-ChIP workflow
Biotinylation of nuclear proteins occurs upon addition of hydrogen peroxide and biotin phenol. Cells are then cross-linked, lysed, and the chromatin sheared. Chromatin bound to biotinylated proteins is then pulled down using streptavidin beads, DNA is extracted and sequenced.

4.4 Nucleophosmin-APEX reveals nucleophosmin at gene promoters

To validate the use of APEX-ChIP for identifying new gene loci enriched at nuclear bodies, I chose to use nucleolar APEX. Nucleoli are known to form at NORs, providing a simple genomic target to look for enriched peaks after streptavidin pulldown. K562 cells are grown in suspension, therefore the biotinylation, quenching, and formaldehyde crosslinking steps had to be done in quick succession with centrifugation steps in between. In traditional APEX biotinylation, media containing biotin-phenol and hydrogen peroxide is suctioned off the plate and quencher solution is added to the plate to quickly quench the

reaction. Because we were using suspension cells, we added quenching solution directly to the media before spinning down cells, then immediately washed cells in quencher solution again. After biotinylation and crosslinking, the chromatin was sheared and we performed a streptavidin pulldown. Pulldown with streptavidin beads had to be optimized for the relative amount of biotinylation (data not shown). After isolation of DNA bound to biotinylated proteins, we sequenced the DNA and normalized to input DNA. Peaks were called using MACS2 peak caller. As expected, we found peaks at NORs (Figure 4.3A and B). This confirmed that nucleophosmin APEX-ChIP properly identified genomic loci associated with nucleoli. Surprisingly, we also found that proteins biotinylated in nucleophosmin APEX-ChIP were enriched along gene promoters (Figure 4.3C), indicating that APEX-ChIP may have identified either new genomic loci associated with nucleoli or with nucleophosmin itself.

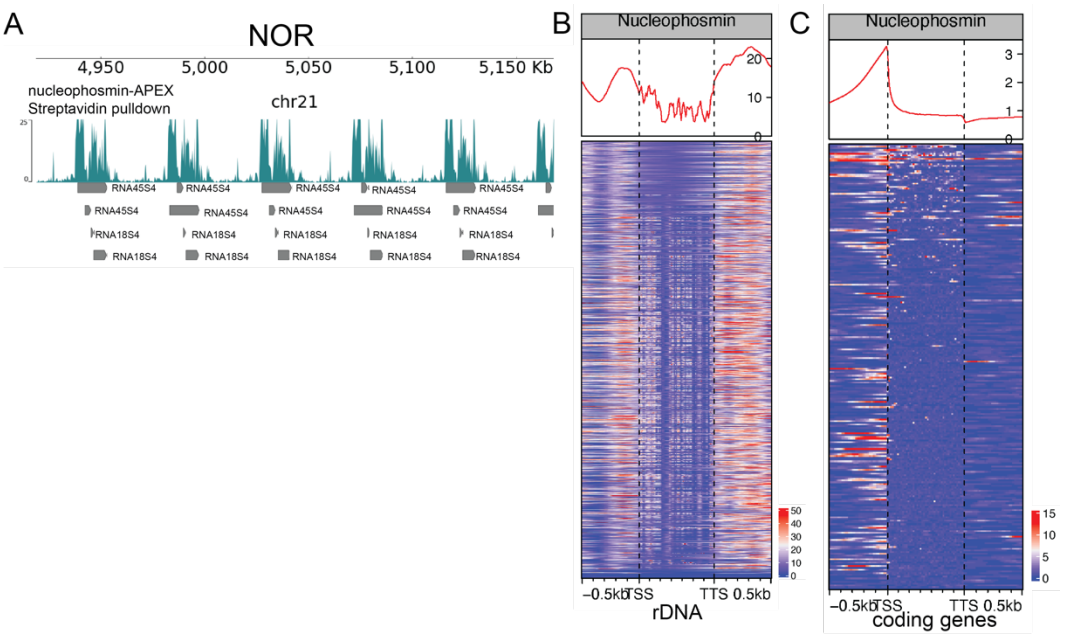


Figure 4.3 APEX-ChIP reveals association of nucleophosmin with gene promoters
A) Fold enrichment tracks of streptavidin pulldown ChIP from nucleophosmin-APEX cells along nucleolar organizing regions (NORs). B and C) Heatmap of nucleophosmin enrichment at rDNA genes (B) and coding genes (C) (individually aligned at bottom, grouped at top graph).

4.5 Discussion

One of the barriers to performing ChIP-seq experiments is the availability of good ChIP antibodies. The use of streptavidin beads to pull down on biotinylated proteins removes the need to test different antibodies for each protein of interest, which is often one of the more time-consuming steps in developing ChIP protocols. In this chapter, I have demonstrated the use of APEX-ChIP to identify the chromosome environments associated with nuclear bodies. As proof-of-principle, I used nucleophosmin-APEX to detect peaks at NORs. My results show that nucleophosmin APEX-ChIP detects rDNA repeats robustly. Future directions include comparing the peak intensity to ChIP using an antibody for nucleophosmin and expanding this protocol for use in smaller nuclear bodies.

I was surprised to also see peaks at promoter regions of coding genes. Because APEX-ChIP is specific to a compartment and not a particular protein, I cannot attribute these peaks to nucleophosmin alone. Because nucleolar components can be found in the nucleoplasm as well, this result would have to be repeated with other nucleolar proteins before attributing these peaks to nucleoli. However, several papers have identified nucleophosmin as binding regulatory promoter regions, supporting the idea that these peaks are due to nucleophosmin. Nucleophosmin has been found to bind specifically at DNA regions forming G-quadruplex structures, and has been identified to bind to promoter regions of c-MYC, SODS, and PD-L1 (Gallo et al., 2012; G. Qin et al., 2020). It was recently found that acetylated nucleophosmin (acNPM1) does not localize to nucleoli (Senapati et al., 2021). ChIP-seq experiments using antibodies against acNPM1 detected acNMP1 at gene promoters in oral tumorigenesis. The peaks detected at gene promoters in nucleophosmin APEX-ChIP may be from acetylated nucleophosmin, as the

nucleophosmin-APEX protein can be modified in cells. This further supports the idea that two nuclear body proteins are required before ascribing a peak to a nuclear body, as nuclear body proteins can exchange with the nucleoplasm and perform different functions.

A similar method to APEX-ChIP, ALAP-seq, was used to identify PML body-associated regions (Kurihara et al., 2020). The main difference in their workflow and the one outlined here is the duration of labeling time with APEX. This supports the idea that APEX and ChIP-sequencing can be combined to detect nuclear body associated regions. APEX-ChIP requires identification of a region with at least two nuclear body proteins before calling a region nuclear body associated instead of protein associated. This project is ongoing, but the preliminary results in this chapter along with the recent demonstration of ALAP-seq support its viability for use in comprehensive characterization of nuclear bodies.

One of the benefits of APEX-ChIP is that it does not require fixation of cells prior to biotinylation. A similar approach, TSA-seq, has been extensively used to characterize nuclear speckles (Y. Chen et al., 2018). TSA-seq is based on the diffusion of free radicals produced by HRP, which form covalent bonds with nearby macromolecules. TSA-Seq previously identified of a subset of transcriptionally active zones close to nuclear speckles. This sensitive technique utilizes primary and HRP-conjugated secondary antibodies, limiting the proteins of interest and/or requiring antibody development and testing. The major difference is that TSA-seq biotinylates DNA directly, and DNA is purified from cells separately from bound proteins. As a result, TSA-seq has been used to measure the distance of a gene to a nuclear speckle by relating the peak intensity to a distance measurement. This approach could potentially be applied to APEX-ChIP to measure changes after

perturbations to the cell, since biotinylation with APEX does not require fixation and can be quenched. APEX has not yet been shown to directly biotinylate DNA, though it acts through the same mechanism as HRP, suggesting that it is possible. In the future, APEX-ChIP can be adapted to directly isolate DNA rather than biotinylated protein-DNA complexes.

5 Concluding Remarks and Outlook

In my thesis, I developed and adapted new methods to comprehensively characterize the CB and its components. By using approaches that targeted multiple CB components, I was able to identify new CB components and characterize new regulators of CB assembly. I focused on the specific effect of large ribosomal subunits and demonstrated that knockdown of RPL proteins affects CB assembly. This opens a new venue of studying CBs and their interaction with RNP complexes, this time through interactions with ribosomal proteins. Although I chose to focus on RPLs, both the mass spectrometry done in Chapter 2 and the siRNA screen in Chapter 3 contain more CB components that can be further studied. These methods were developed in HeLa cells but can be expanded to other cell lines to better understand regulation of CBs in different tissues. The approaches used here can be applied to other nuclear bodies, further expanding our understanding of nuclear organization and the role of nuclear bodies.

My thesis shows the first link between ribosomal proteins and CBs, yet the molecular mechanisms that regulate this interaction remain undetermined. Are ribosomal proteins directly interacting with canonical CB proteins? If so, where are these interactions occurring and are they necessary for CB assembly and function? I have shown that RPL knockdown changes CB structure and assembly on chromatin. In the future, it will be important to understand whether these altered CBs function differently from canonical CBs. If so, are there populations of normal CBs that perform different functions? How are these regulated in cells and are there differential functions for CBs in different tissue types?

We still do not fully understand the relationship between CBs and transcription. It will be important for the field to understand whether transcription regulates CBs or if CBs themselves can directly affect transcription. The changes in Pol II binding upon RPL KD despite increased CB number may provide a molecular tool to further question the relationship between CBs and transcription. Furthermore, these results may indicate a role for ribosomal proteins in Pol II transcription. Further studies are needed to parse out whether these results can be attributed to individual ribosomal subunits or the assembled ribosome itself. The abundance of transcriptional proteins identified by Coilin-APEX2 strongly suggests that CBs are near sites of transcription. Together, these results provide a list of proteins to study in order to better understand the role of CBs in transcription.

The list of CB components and subsequent screen to identify which of these regulate CB formation can now be used to perturb CB assembly. By having multiple proteins that can be targeted within CBs, we can begin to probe CBs and attribute results to loss of CBs rather than loss of a particular component. We still lack the tools as a field to get rid of CBs without getting rid of the components themselves. By having a more robust list of CB components, we can now begin to think about processes that can be perturbed to get rid of CBs without always depleting coilin.

The most important next steps in studying CBs will be to develop functional assays as readouts of CB activity. This is a considerably difficult step given the many functions attributed to CBs. In my thesis, I have characterized the components of CBs. Moving forward, we will need functional characterization of these components within CBs. Together, this will provide a full picture of the CBs role in the cell nucleus.

6 Methods

6.1 Cell culture

HeLa cells were grown in DMEM GlutaMAX medium (Gibco) supplemented with 10% heat-inactivated FBS (Gibco) and 1% penicillin-streptomycin (Gibco). Cells were incubated in humidified 5% CO_2 at 37°C.

K562 cells were grown in suspension in RPMI 1640 medium (Gibco) supplemented with 10% heat-inactivated FBS (Gibco) and 1% penicillin-streptomycin (Gibco). Cells were incubated in humidified 5% CO_2 at 37°C.

6.2 APEX2 cell line generation

Coilin_APEX2, APEX2_NLS, and ΔNTD_NLS constructs were generated with the InFusion HD kit (Takara) by inserting sequences after the EF-1α promoter into plasmid backbone pWPI (generated by Didier Trono, Addgene #12254). Lentiviruses from these constructs were prepared by transfecting confluent HEK293FT cells with pWPI containing the desired insert along with pMD2.G, and pCMV R8.74 (generated by Dider Trono, Addgene #12250 and #22036 respectively) using Fugene HD reagent (Promega). After 72h, viral supernatant was harvested, filtered through a 0.45 μm filter and used to transduce HeLa and K562 cells. After 3 passages, GFP positive cells were sorted using FACSAria II. Coilin_APEX2 expressing cells were sorted into single cells and validated by staining for V5 and coilin to ensure that CB formation was not altered. APEX2_NLS and ΔNTD_NLS were sorted into low-GFP expressing pools and validated by staining for V5 to confirm nuclear expression of the construct.

6.3 APEX2 labeling

APEX2 labeling was performed as described in Hung et al (Hung et al., 2016). Briefly, confluent HeLa cells were incubated with 500 uM biotinyl tyramide (Chemodex) for 30 minutes at 37°C with 5% CO_2. H_2O_2 was then added to a final concentration of 1mM for 1 minute. The reaction was immediately quenched by washing 3 times with 'quencher solution' (10mM sodium ascorbate, 10mM sodium azide, and 5 mM Trolox in 1XPBS). Cells were then either fixed for imaging or pelleted for nuclei isolation and enrichment with streptavidin.

6.4 Nuclei isolation and streptavidin enrichment

Cells from 5 15 cm plates (~$1x10^8$ cells) were labeled as previously described and scraped from plates using quencher solution. Cells were centrifuged at 3000g for 10 minutes at 4°C. Cells were then washed with cold PBS, transferred to a fresh tube, and spun down at 250g for 5 minutes at 4°C. Cells were then gently resuspended in 5mL of 'Buffer A' (10mM Hepes pH 7.4, 10mM KCl, 340mM sucrose, 10% glycerol, 4 mM MgCl2, 1 mM DTT, 1X protease inhibitors (Roche), 10 mM beta-GP, .1% RNAse OUT (Invitrogen)). 5mL of 'Buffer B' (Buffer A + 0.2% TritonX-100) were added and samples were incubated on ice for 10 minutes and centrifuged at 1200g for 5 minutes at 4°C. A fraction of the supernatant (cytoplasmic fraction) was saved for downstream analysis while the rest was discarded. Cells were resuspended in 5 mL Nuclear Resuspension Buffer (10mM Hepes pH 7.4, 50% glycerol, 75mM NaCl, 1 mM DTT) and centrifuged at 1000g for 5 minutes at 4°C 2X. The supernatant was discarded, and the remaining pellet (nuclei) lysed in 1mL RIPA lysis buffer (50mM Tris pH 7.5, 150mM NaCL, 0.1% SDS, 0.5%

sodium deoxycholate, 1% Triton X-100) supplemented with 1X protease inhibitors (Roche) and 1mM PMSF. Nuclei were sonicated at 30%AMP, 10sON 20sOFF 30X. Lysates were clarified by centrifuging at 15000g for 10 min at 4°C. Protein concentration was quantified using Pierce 660-nm Protein Assay Reagant (Thermo Scientific). Streptavidin enrichment was performed as previously described in Hung et al (Hung et al., 2016)4mg of protein were incubated with 50 uL of magnetic streptavidin beads (Pierce) for 3 hours at 4°C. Beads were pelleted and the "flow-through" was set apart for downstream analysis. Then, beads were sequentially washed 2X with RIPA lysis buffer, 1X with 1M KCl, 1X with 0.1M Na_2CO_3, 1X 2M urea in 10mM Tris-HCl pH8.0, and 2X with RIPA lysis buffer. Biotinylated proteins were eluted from beads by boiling for 10 minutes in 4X NuPAGE LDS Sample Buffer (Invitrogen) supplemented with 2mM biotin and 20mM DTT.

6.5 Western blotting

Proteins were resolved on 4-12% Bis-Tris polyacrylamide gels (Invitrogen) and transferred to .45um nitrocellulose membrane (BioRad). After blocking, membranes were probed with the antibodies in Appendix Table 2.

6.6 Fixed cell imaging

Cells were grown in No 1.5 coverslips (Zeiss) in either 6 or 24 well plates, fixed in 4% paraformaldehyde, and permeabilized in 0.5% Triton-X100 (American Bioanalytical) in 1XPBS. Cells were then blocked in 3% Bovine Serum Albumin (Sigma) and probed with antibodies in blocking buffer see (Appendix Table 2) Cell nuclei were stained with .25 ug/mL Hoechst 34580 (Thermo Fisher Scientific) in 1XPBS and mounted using

Prolong Diamond Antifade Mountant (Thermo Fisher Scientific). Imaging was done using a Leica SP8 Laser Scanning Confocal. STED imaging was carried out on the same confocal microscope enabled with a 775nm depletion laser.

6.7 Mass Spectrometry

6.7.1 Experimental procedure

Samples were separated on a 4%–12% NOVEX NuPage gradient SDS gel (Thermo) for 10 minutes at 180 V in 1X MES buffer (Thermo). Proteins were fixated and stained with coomassie G250 brilliant blue (Carl Roth). The gel lanes were cut, and each lane was minced into approximately 1x1 mm pieces. Gel pieces were destained with a 50% ethanol/50 mM ammoniumbicarbonate (ABC) solution. Proteins were reduced in 10 mM DTT (Sigma-Aldrich) for 1 hour at 56°C and then alkylated with 50 mM iodoacetamide (Sigma-Aldrich) for 45 min at room temperature. Proteins were digested with mass spec grade trypsin (Sigma) overnight at 37°C. Peptides were extracted from the gel by two incubations with 30% ABC/acetonitrile and three subsequent incubations with pure acetonitrile. The acetonitrile was subsequently evaporated in a concentrator (Eppendorf) and loaded on StageTips (Z. A. Chen et al., 2016) for desalting and storage.

6.7.2 Analysis

For mass spectrometric analysis, peptides were separated on a 20 cm self-packed column with 75-µm inner diameter filled with ReproSil-Pur 120 C_{18}-AQ (Dr. Maisch GmbH) mounted to an EASY HPLC 1000 (Thermo Fisher) and sprayed online into an Q Exactive Plus mass spectrometer (Thermo Fisher). We used a 94 min gradient from 2% to 40% acetonitrile in 0.1% formic acid at a flow of 225 nL/min. The mass spectrometer was operated with a top 10 MS/MS data-dependent acquisition scheme per MS full scan. Mass

spectrometry raw data were searched using the Andromeda search (Cox et al., 2011) integrated into MaxQuant suite 1.5.2.8 engine (Cox & Mann, 2008) using the Uniprot Homo sapiens database (January 2020; 42,338 entries). In all analyses, carbamidomethylation at cysteine was set as fixed modification while methionine oxidation and protein N-acetylation were considered as variable modifications. Match between run option was activated. Prior to bioinformatics analysis, reverse hits, proteins only identified by site, protein groups based on one unique peptide, and known contaminants were removed.

For the further bioinformatics analysis, the LFQ values were log2 transformed and the median across the replicates was calculated. This enrichment was plotted against the – log 10 transformed p-value (Welch t-test) using the ggplot2 package in the R environment.

6.8 STED imaging and analysis

STED imaging was performed as described previously(Courchaine et al., 2021). Analysis was done using FIJI software (Schindelin et al., 2012). STED and confocal image stacks of Coilin and SMN were segmented using a linear-bin Otsu threshold applied to the entire volume of each channel after smoothing by a Gaussian filter (sigma of 40 nm in x and y, 43.87 nm in z). The STED mask in both colors was taken to be the intersection of the smoothed and thresholded STED channels and their paired confocal channel masks. The original STED images were then analyzed using these masks. The intensity-weighted and masked center of mass was calculated in 3D for both Coilin and SMN, and the Euclidean distance separation calculated as the offset. To calculate the fractional intensity overlap, the intersection of Coilin and SMN masks were taken, and the intensity in this

region was summed over both channels and divided by the combined sum of the intensity in each of the individual channels within their own mask.

6.9 ChIPseq

6.9.1 Experimental procedure

For the siRNA knockdown ChIP-seq experiments using Hela cells, 2 plates for each condition were grown to confluence on 15 cm dishes to obtain ~10^8 cells/ plate. Cells were crosslinked for 10 minutes by adding 37% formaldehyde solution directly to the medium for a final concentration of 1%. The medium was aspirated and cells were washed twice with 5 ml cold 1x PBS with the addition of a protease inhibitor (1:100, Roche). Cells from 2 dishes of the same condition were then scraped of the plate with 10 ml of the PBS/ protease inhibitor solution and transferred to a Falcon tube (20 ml total volume) to be pelleted for 5 min at 2500 x g at 4°C. Pellets were frozen at -80°C until further continuation.

For cell lysis pellets were thawed on ice and 1 ml SDS lysis buffer was added containing 1x protease inhibitor. The solution was transferred to a 15 ml tube and incubated on ice for 10 min. To shear the DNA cells were sonicated using a tapered microtip (1/8" diameter, Branson) optimized and set to 30% amplitude, 30 cycles of 10 s pulses with 20 s pauses between pulses (optimal product size ~250 nt). The lysate was transferred to a new 1.5 ml tube and centrifuged for 10 min at 14.000rpm at 4°C. The supernatant of the ChIP lysate was transferred to a new tube and 200 µl were diluted in 1.8 ml ChIP Dilution Buffer + 1x PI (1x final) for a total volume of 2 ml/ immunoprecipitation (IP) reaction. 50 µl of the sonicated solution were used as Input material and diluted in 450 µl ChIP Dilution Buffer + PI (1x final) and frozen at -20°C. Lysates were immunoprecipitated over night on

a rotary wheel at 4°C with the addition of 5 µg of the respective immunoprecipitating antibody (anti-Coilin, anti-Pol2). Antibodies specific to their DNA-protein complexes were bound by adding 18 µl Dynabeads protein G (Life technologies) to each IP and incubating for 1 hour on a rotary wheel at 4°C. Beads were captured on a magnetic rack and washed for 4 min on a rotary wheel by addition of 1 ml of the following buffers:

Low salt immune complex wash buffer

High salt immune complex wash buffer

Lithium chloride immune complex wash buffer

Tris-EDTA buffer

After the last wash protein/DNA complexes were eluted of the beads by adding 2x 250 µl elution buffer and incubating for 15 minutes on rotary wheel. Eluates were combined to a total volume of 500 µl and IPs as well as frozen inputs were uncrosslinked and Proteinase K digested for 6 hours at 65°C with addition of 20 µl of 5M NaCl, 10 µl of 0.5M EDTA, 20 µl 1M Tris-HCl pH 6.5 and 10 µl of 10 mg/ml Proteinase K. DNA was extracted using phenol-chloroform. Dry pellets were resuspended in 50 µl water. For ChIP-Seq, 2 IPs for Pol2 and 3 IPs for Coilin were combined for each condition for a total of 3 libraries for each condition (Input, Coilin-IP, Pol2-IP). Quality control and library preparations were performed by the Yale Center for Genome Analysis (YCGA). Libraries were sequenced on the Novaseq high throughput sequencer (paired-end, 50 million reads/ sample). For buffer recipes, see Appendix Table 3.

APEX-ChIP

K562 suspension cells were grown to confluence in a T170 flask containing 40 ml medium to obtain 2 x 10^7 cells/ flask. Cells were pre-incubated with biotin phenol by

directly adding biotin phenol (final concentration of 500 µM) to the medium and incubating for 30 min at 37°C. Cells are then transferred to a 50 ml Falcon tube. To catalyze proximity biotinylation hydrogen peroxide was added to the cells (final concentration of 1 mM) and incubated for exactly 1 minute at room temperature. The reaction was quenched by adding 425 µl of 1 M sodium ascorbate (final concentration of 10 mM) and 2.125 ml of 100 mM Trolox (final concentration of 5 mM) followed by centrifugation for 5 minutes at 1100 g at 4°C. The supernatant was aspirated and pellets were washed with quencher solution. Cells were crosslinked by adding 20 ml quencher solution 2 containing 1% formaldehyde for 10 minutes at room temperature. The reaction was quenched by addition of Glycine (0.125M final concentration) for 10 minutes. Cells were washed with 1x PBS and pelleted for 5 min at 2500 x g at 4°C. Pellets were frozen at -80°C until further continuation.

For cell lysis pellets were thawed on ice and 1 ml lysis buffer was added containing 1x protease inhibitor. The solution was transferred to a 15 ml tube and incubated on ice for 10 min. Cells were pelleted for 5 min at 2500 x g at 4°C and washed twice with wash buffer containing 1x protease inhibitor. To shear the DNA, 1.5 ml shearing buffer was carefully added and cells were washed twice for 5 min at 500 g, at 4˚C. Pellets were resuspended in 1 ml shearing buffer containing 1x protease inhibitor followed by sonication using a tapered microtip (1/8" diameter, Branson) optimized and set to 30% amplitude, 30 cycles of 10 s pulses with 20 s pauses between pulses (optimal product size ~250 nt). The lysate was transferred to a new 1.5 ml tube and centrifuged for 10 min at 14.000rpm at 4°C. The supernatant of the ChIP lysate was transferred to a new tube and 200 µl were diluted in 1.8 ml ChIP Dilution Buffer + 1x PI (1x final) for a total volume of 2 ml/ immunoprecipitation

(IP) reaction. 50 µl of the sonicated solution were used as Input material and frozen at -20°C.

Lysates were immunoprecipitated over night on a rotary wheel at 4°C with the addition of 60 µl Streptavidin beads (Pierce, 88817). Beads were captured on a magnetic rack and washed for 4 min on a rotary wheel by addition of 1 ml of the following buffers:

Low salt immune complex wash buffer

High salt immune complex wash buffer

Lithium chloride immune complex wash buffer (twice)

Tris-EDTA buffer

After the last wash proteins were digested by adding 500 µl PK digestion containing 10 µl 10mg/ml Proteinase K for 2 hours at 50˚C in a Thermomixer set to 1000 rpm.

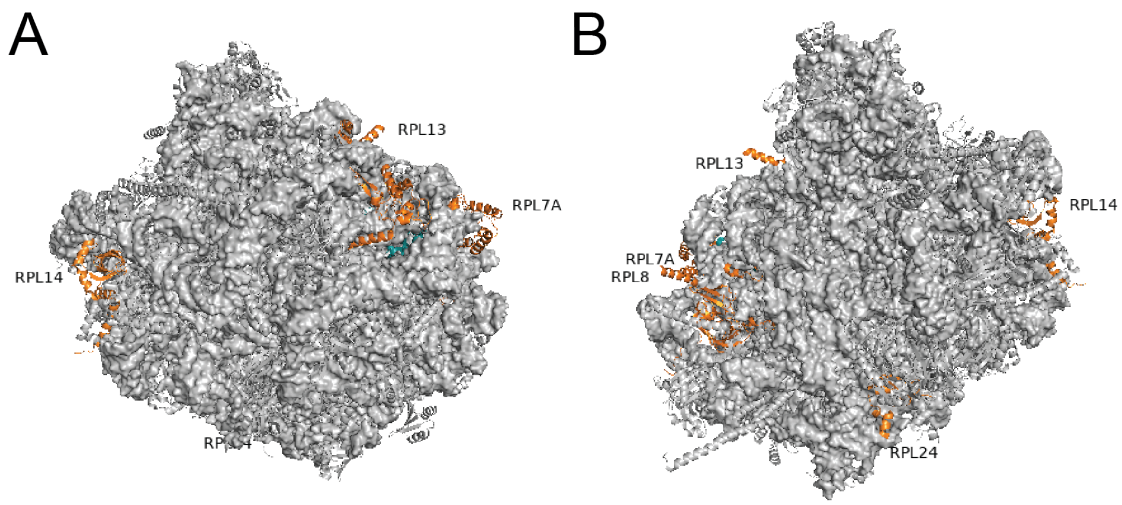
For uncrosslinking Inputs were thawed and 450 µl of PK Digestion buffer with 20 µl of 5M NaCl and 2 µl of RNaseA (10 mg/ml stock) was added. The IP samples were captured on beads after Proteinasek K treatment and supernatants were transferred to a new 1.5 ml tube and 20 µl of 5M NaCl and 2 µl of RNaseA (10 mg/ml stock) were added like for the Input samples. All samples were uncrosslinked at 65˚C over night.

DNA was extracted using phenol-chloroform. Dry pellets were resuspended in 50 µl water. For Sequencing, 4 IPs were combined. Quality control and library preparations (IPs and Input) were performed by the Yale Center for Genome Analysis (YCGA). Libraries were sequenced on the Novaseq high throughput sequencer (paired-end, 50 million reads/ sample).

6.9.2 ChIP-seq Data Analysis

ChIPseq analysis was done as described previously in (Machyna et al., 2014). Reads were first quality checked using FastQC ("FastQC," 2015) , cleaned from adapter sequences with Cutadapt (Martin, 2011), and aligned to reference genome hg38 using Bowtie2 aligner (Langmead & Salzberg, 2012). If necessary, SAM format files were converted to BAM format with Samtools (H. Li et al., 2009). Peaks were then called with MACS2 using input as a noIP control. Tracks were visualized using pyGenometracks (Lopez-Delisle et al., 2021).

7 Appendix



Appendix Figure 7.1 RPL proteins identified in screen on 60S subunit structure
Structure of *T. thermophilia* 60S subunit in complex with eIF6. Views of the solvent exposed (A) and 40S binding side (B). Color-coded ribosomal proteins are shown as ribbons. Proteins found to effect CB structure (RPL13, 14, 24, 7A, and 8) are colored orange. Proteins that did not show an effect (RPL15) on CB structure upon knockdown are colored teal.

Appendix Table 1

Ribosomal Protein	Interacting PRMT	Protein sequence (putative methyl-R labeled with red)
RPL13	PRMT8 (Type I, aDMA)	MAPSRNGMVLKPHFHKDWQRRVATWFNQPARKIR RRKARQAKARRIAPRPASGPIRPIVRCPTVRYHTKV RAGRGFSLEELRVAGIHKKVARTIGISVDPRRRNKST ESLQANVQRLKEYRSKLILFPRKPSAPKKGDSSAEEL KLATQLTGPVMPVRNVYKKEKARVITEEEKNFKAF ASLRMARANARLFGIRAKRAKEAAEQDVEKKK
RPL14	PRMT5 (Type II, sDMA)	MVFRRFVEVGRVAYVSFGPHAGKLVAIVDVIDQNR ALVDGPCTQVRRQAMPFKCMQLTDFILKFPHSAHQ KYVRQAWQKADINTKWAATRWAKKIEARERKAK MTDFDRFKVMKAKKMRNRIIKNEVKKLQKAALLK ASPKKAPGTKGTAAAAAAAAAAKVPAKKITAASK KAPAQKVPAQKATGQKAAPAPKAQKGQKAPAQKA PAPKASGKKA
RPL24	N/A	MKVELCSFSGYKIYPGHGRRYARTDGKVFQFLNAK CESAFLSKRNPRQINWTVLYRRKHKKGQSEEIQKKR TRRAVKFQRAITGASLADIMAKRNQKPEVRKAQRE QAIRAAKEAKKAKQASKKTAMAAAKAPTKAAPKQ KIVKPVKVSAPRVGGKR
RPL7A	N/A	MEGVEEKKKEVPAVPETLKKKRRNFAELKIKRLRK KFAQKMLRKARRKLIYEKAKHYHKEYRQMYRTEIR MARMARKAGNFYVPAEPKLAFVIRIRGINGVSPKVR KVLQLLRLRQIFNGTFVKLNKASINMLRIVEPYIAW GYPNLKSVNELIYKRGYGKINKKRIALTDNALIARS LGKYGIICMEDLIHEIYTVGKRFKEANNFLWPFKLSS PRGGMKKKTTHFVEGGDAGNREDQINRLIRRMN
RPL8	PRMT5 (Type II, sDMA) PRMT7 (Type III, monomethyl guanidinium)	MGRVIRGQRKGAGSVFRAHVKHRKGAARLRAVDF AERHGYIKGIVKDIIHDPGRGAPLAKVVFRDPYRFK KRTELFIAAEGIHTGQFVYCGKKAQLNIGNVLPVGT MPEGTIVCCLEEKPGDRGKLARASGNYATVISHNPE TKKTRVKLPSGSKKVISSANRAVVGVVAGGGRIDK PILKAGRAYHKYKAKRNCWPRVRGVAMNPVEHPF GGGNHQHIGKPSTIRRDAPAGRKVGLIAARRTGRLR GTKTVQEKEN

Appendix Table 2 List of Antibodies Used in this Study

Name	Source	Application and Dilution
Streptavidin-HRP	Jackson Immunoresearch	WB: 1:20,000 IF: 1:200
GAPDH	Santa Cruz	WB: 1:2000
COILIN	Abcam	WB: 1:1000 IF 1:2000 IP: 5ug
SMN	Abcam	IF 1:500
V5	Invitrogen	IF 1:500
H3	Santa Cruz	WB 1:10,000
Y12	Gift from Joan Steitz	WB: 1:3 IF 1:16
Nopp140	Abcam	IF: 1:1000
Dyskerin	Sigma Aldrich	Sigma Aldrich
IRF2BP1	Sigma Aldrich	IF: 1:500
SRRT	Sigma Aldrich	IF 1:500
EFTUD2	Sigma Aldrich	IF 1:500

Appendix Table 3 ChIP Buffer Recipes

Component	Final concentration in water
SDS Lysis Buffer	
SDS	1%
EDTA	10 mM
Tris-HCl, pH 8.1	50 mM
Add protease inhibitor before use	
ChIP Dilution Buffer	
SDS	0.01%
Triton X-100	1.1%
EDTA	1.2 mM
Tris-HCl, pH 8.1	16.7 mM
NaCl	167 mM
Add protease inhibitor before use	
Low Salt Immune Complex Wash Buffer	
SDS	0.1%
Triton X-100	1%
EDTA	2 mM
Tris-HCl, pH 8.1	20 mM
NaCl	150 mM
High Salt Immune Complex Wash Buffer	
SDS	0.1%
Triton X-100	1%
EDTA	2 mM
Tris-HCl, pH 8.1	20 mM
NaCl	500 mM
Lithium Chloride Immune Complex Wash Buffer	
LiCl	0.25 M
NP-40	1%
deoxycholic acid (sodium salt)	1%
EDTA	1 mM
Tris-HCl, pH 8.1	10 mM
Elution Buffer	
SDS	1%
$NaHCO_3$	0.1 M
Tris-EDTA Buffer	
EDTA	1 mM
Tris-HCl, pH 8.1	10 mM

Appendix Table 4 APEX-ChIP Buffer Recipes

Component	Final concentration in water
Quencher solution (in 1XPBS)	
Sodium Ascorbate	10 mM
Trolox	5 mM
Lysis Buffer	
HEPES pH 7.9	50 mM
NaCl	140 mM
EDTA	1 mM
Glycerol	10%
NP-40	0.5%
Triton X-100	0.25%
Add protease inhibitor before use	
Wash Buffer	
Tris-HCl, pH 8.1	10 mM
NaCl	200 mM
EDTA, pH 8	1 mM
EGTA pH 8	0.5 mM
Shearing Buffer	
SDS	0.1%
EDTA	1 mM
Tris-HCl, pH 8.1	10 mM
Proteinase K Digest Buffer	
HEPES, pH 7.9	20 mM
EDTA	1 mM
SDS	0.5 %

References

Abbott, J., Marzluff, W. F., & Gall, J. G. (1999). The stem-loop binding protein (SLBP1) is present in coiled bodies of the Xenopus germinal vesicle. *Molecular biology of the cell, 10*(2), 487-499.

Andrade, L. E., Chan, E. K., Raska, I., Peebles, C. L., Roos, G., & Tan, E. M. (1991). Human autoantibody to a novel protein of the nuclear coiled body: immunological characterization and cDNA cloning of p80-coilin. *The Journal of experimental medicine, 173*(6), 1407-1419.

Arias Escayola, D., & Neugebauer, K. M. (2018). Dynamics and Function of Nuclear Bodies during Embryogenesis. *Biochemistry*. doi:10.1021/acs.biochem.7b01262

Ariotti, N., Rae, J., Leneva, N., Ferguson, C., Loo, D., Okano, S., Hill, M. M., Walser, P., Collins, B. M., & Parton, R. G. (2015). Molecular Characterization of Caveolin-induced Membrane Curvature. *The Journal of biological chemistry, 290*(41), 24875-24890. doi:10.1074/jbc.M115.644336

Barcaroli, D., Dinsdale, D., Neale, M. H., Bongiorno-Borbone, L., Ranalli, M., Munarriz, E., Sayan, A. E., McWilliam, J. M., Smith, T. M., Fava, E., Knight, R. A., Melino, G., & De Laurenzi, V. (2006). FLASH is an essential component of Cajal bodies. *Proceedings of the National Academy of Sciences of the United States of America, 103*(40), 14802-14807. doi:10.1073/pnas.0604225103

Barutcu, A. R., Wu, M., Braunschweig, U., Dyakov, B. J. A., Luo, Z., Turner, K. M., Durbic, T., Lin, Z. Y., Weatheritt, R. J., Maass, P. G., Gingras, A. C., & Blencowe, B. J. (2022). Systematic mapping of nuclear domain-associated transcripts reveals speckles and lamina as hubs of functionally distinct retained introns. *Molecular cell, 82*(5), 1035-1052.e1039. doi:10.1016/j.molcel.2021.12.010

Barysch, S. V., Stankovic‐Valentin, N., Miedema, T., Karaca, S., Doppel, J., Nait Achour, T., Vasudeva, A., Wolf, L., Sticht, C., Urlaub, H., & Melchior, F. (2021). Transient deSUMOylation of IRF2BP proteins controls early transcription in EGFR signaling. *EMBO Reports, 22*(3). doi:10.15252/embr.201949651

Batalova, F. M., Stepanova, I. S., Skovorodkin, I. N., Bogolyubov, D. S., & Parfenov, V. N. (2005). Identification and dynamics of Cajal bodies in relation to karyosphere formation in scorpionfly oocytes. *Chromosoma, 113*(8), 428-439. doi:10.1007/s00412-004-0328-y

Bell, M. (2002). p110, a novel human U6 snRNP protein and U4/U6 snRNP recycling factor. *The EMBO Journal, 21*(11), 2724-2735. doi:10.1093/emboj/21.11.2724

Berchtold, D., Battich, N., & Pelkmans, L. (2018). A Systems-Level Study Reveals Regulators of Membrane-less Organelles in Human Cells. *Molecular Cell, 72*(6), 1035-1049.e1035. doi:10.1016/j.molcel.2018.10.036

Berry, J., Weber, S. C., Vaidya, N., Haataja, M., & Brangwynne, C. P. (2015). RNA transcription modulates phase transition-driven nuclear body assembly. *Proceedings of the National Academy of Sciences, 112*(38), E5237-E5245. doi:10.1073/pnas.1509317112

Bhavsar, R. B., Makley, L. N., & Tsonis, P. A. (2010). The other lives of ribosomal proteins. *Human Genomics, 4*(5), 327. doi:10.1186/1479-7364-4-5-327

Bizarro, J., Deryusheva, S., Wacheul, L., Gupta, V., Ernst, F. G. M., Lafontaine, D. L. J., Gall, J. G., & Meier, U. T. (2021). Nopp140-chaperoned 2′ -O-methylation of small nuclear RNAs in Cajal bodies ensures splicing fidelity. *Genes & Development, 35*(15-16), 1123-1141. doi:10.1101/gad.348660.121

Bohmann, K., Ferreira, J. A., & Lamond, A. I. (1995). Mutational analysis of p80 coilin indicates a functional interaction between coiled bodies and the nucleolus. *The Journal of cell biology, 131*(4), 817-831.

Boisvert, F.-M., Cote, J., Boulanger, M.-C., Cleroux, P., Bachand, F., Autexier, C., & Richard, S. (2002). Symmetrical dimethylarginine methylation is required for the localization of SMN in Cajal bodies and pre-mRNA splicing. *The Journal of cell biology, 159*(6), 957-969. doi:10.1083/jcb.200207028

Boisvert, F.-M., Koningsbruggen, S. v., Navascués, J., & Lamond, A. I. (2007). The multifunctional nucleolus. *Nature Reviews Molecular Cell Biology, 8*(7), 574. doi:10.1038/nrm2184

Bongiorno-Borbone, L., Cola, A. D., Barcaroli, D., Knight, R. A., Ilio, C. D., Melino, G., & Laurenzi, V. D. (2010). FLASH degradation in response to UV-C results in histone locus bodies disruption and cell-cycle arrest. *Oncogene, 29*(6), 802. doi:10.1038/onc.2009.388

Boulon, S., Verheggen, C., Jady, B. E., Girard, C., Pescia, C., Paul, C., Ospina, J. K., Kiss, T., Matera, A. G., Bordonné, R., & Bertrand, E. (2004). PHAX and CRM1 Are Required Sequentially to Transport U3 snoRNA to Nucleoli. *Molecular Cell, 16*(5), 777-787. doi:10.1016/j.molcel.2004.11.013

Bruns, A.-F., van Bergeijk, J., Lorbeer, C., Nolle, A., Jungnickel, J., Grothe, C., & Claus, P. (2009). Fibroblast growth factor-2 regulates the stability of nuclear bodies. *Proceedings of the National Academy of Sciences of the United States of America, 106*(31), 12747-12752. doi:10.1073/pnas.0900122106

Cantarero, L., Sanz-Garcia, M., Vinograd-Byk, H., Renbaum, P., Levy-Lahad, E., & Lazo, P. A. (2015). VRK1 regulates Cajal body dynamics and protects coilin from proteasomal degradation in cell cycle. *Scientific Reports, 5*, 10543. doi:10.1038/srep10543

Carmo-Fonseca, M., Ferreira, J., & Lamond, A. I. (1993). Assembly of snRNP-containing coiled bodies is regulated in interphase and mitosis--evidence that the coiled body is a kinetic nuclear structure. *Journal of Cell Biology, 120*(4), 841-852. doi:10.1083/jcb.120.4.841

Carmo-Fonseca, M., Pepperkok, R., Carvalho, M. T., & Lamond, A. I. (1992). Transcription-dependent colocalization of the U1, U2, U4/U6, and U5 snRNPs in coiled bodies. *The Journal of Cell Biology, 117*(1), 1-14. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1532583

Chen, Y., Zhang, Y., Wang, Y., Zhang, L., Brinkman, E. K., Adam, S. A., Goldman, R., van Steensel, B., Ma, J., & Belmont, A. S. (2018). Mapping 3D genome organization relative to nuclear compartments using TSA-Seq as a cytological ruler. *The Journal of Cell Biology, 217*(11), 4025-4048. doi:10.1083/jcb.201807108

Chen, Z. A., Fischer, L., Cox, J., & Rappsilber, J. (2016). Quantitative Cross-linking/Mass Spectrometry Using Isotope-labeled Cross-linkers and MaxQuant. *Molecular & Cellular Proteomics, 15*(8), 2769-2778. doi:10.1074/mcp.m115.056481

Childs, K. S. (2003). Identification of novel co-repressor molecules for Interferon Regulatory Factor-2. *Nucleic Acids Research, 31*(12), 3016-3026. doi:10.1093/nar/gkg431

Cioce, M., Boulon, S., Matera, A. G., & Lamond, A. I. (2006). UV-induced fragmentation of Cajal bodies. *The Journal of cell biology, 175*(3), 401-413. doi:10.1083/jcb.200604099

Collier, S., Pendle, A., Boudonck, K., van Rij, T., Dolan, L., & Shaw, P. (2006). A distant coilin homologue is required for the formation of cajal bodies in Arabidopsis. *Molecular Biology of the Cell, 17*(7), 2942-2951. doi:10.1091/mbc.E05-12-1157

Courchaine, E. M., Barentine, A. E. S., Straube, K., Lee, D.-R., Bewersdorf, J., & Neugebauer, K. M. (2021). DMA-tudor interaction modules control the specificity of in vivo condensates. *Cell, 184*(14), 3612-3625.e3617. doi:10.1016/j.cell.2021.05.008

Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology, 26*(12), 1367-1372. doi:10.1038/nbt.1511

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., & Mann, M. (2011). Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *Journal of Proteome Research, 10*(4), 1794-1805. doi:10.1021/pr101065j

Darzacq, X., Kittur, N., Roy, S., Shav-Tal, Y., Singer, R. H., & Meier, U. T. (2006). Stepwise RNP assembly at the site of H/ACA RNA transcription in human cells. *Journal of Cell Biology, 173*(2), 207-218. doi:10.1083/jcb.200601105

Deryusheva, S., & Gall, J. G. (2004). Dynamics of coilin in Cajal bodies of the Xenopus germinal vesicle. *Proceedings of the National Academy of Sciences of the United States of America, 101*(14), 4810-4814. doi:10.1073/pnas.0401106101

Deryusheva, S., & Gall, J. G. (2009). Small Cajal Body–specific RNAs of Drosophila Function in the Absence of Cajal Bodies. *Molecular Biology of the Cell, 20*(24), 5250-5259. doi:10.1091/mbc.E09-09-0777

Dumrongprechachan, V., Salisbury, R. B., Soto, G., Kumar, M., MacDonald, M. L., & Kozorovitskiy, Y. (2021). Cell-type and subcellular compartment-specific APEX2 proximity labeling reveals activity-dependent nuclear proteome dynamics in the striatum. *Nature Communications, 12*(1), 4855. doi:10.1038/s41467-021-25144-y

Dundr, M., Hebert, M. D., Karpova, T. S., Stanek, D., Xu, H., Shpargel, K. B., Meier, U. T., Neugebauer, K. M., Matera, A. G., & Misteli, T. (2004). In vivo kinetics of Cajal body components. *J Cell Biol, 164*(6), 831-842. Retrieved from files/1303/PMC1630494.html

files/1302/831.html

El-Bazzal, L., Rihan, K., Bernard-Marissal, N., Castro, C., Chouery-Khoury, E., Desvignes, J.-P., Atkinson, A., Bertaux, K., Koussa, S., Lévy, N., Bartoli, M., Mégarbané, A., Jabbour, R., & Delague, V. (2019). Loss of Cajal bodies in motor neurons from patients with novel mutations in VRK1. *Human Molecular Genetics, 28*(14), 2378-2394. doi:10.1093/hmg/ddz060

Eliceiri, G. L., & Ryerse, J. S. (1984). Detection of intranuclear clusters of Sm antigens with monoclonal anti-Sm antibodies by immunoelectron microscopy. *Journal of cellular physiology, 121*(2), 449-451. doi:10.1002/jcp.1041210226

Espert, L., Eldin, P., Gongora, C., Bayard, B., Harper, F., Chelbi-Alix, M. K., Bertrand, E., Degols, G., & Mechti, N. (2006). The exonuclease ISG20 mainly localizes in the nucleolus and the Cajal (Coiled) bodies and is associated with nuclear SMN protein-containing complexes. *Journal of cellular biochemistry, 98*(5), 1320-1333. doi:10.1002/jcb.20869

Faresse, N., Colland, F., Ferrand, N., Prunier, C., Bourgeade, M.-F., & Atfi, A. (2008). Identification of PCTA, a TGIF antagonist that promotes PML function in TGF-β signalling. *The EMBO Journal, 27*(13), 1804-1815. doi:10.1038/emboj.2008.109

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 Reports, 22*(7), 1923-1934. doi:https://doi.org/10.1016/j.celrep.2018.01.056

FastQC. (2015). In.

Fatica, A., Dlakić, M., & Tollervey, D. (2002). Naf1 p is a box H/ACA snoRNP assembly factor. *RNA (New York, N.Y.), 8*(12), 1502-1514. Retrieved from https://pubmed.ncbi.nlm.nih.gov/12515383

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1370356/

Fong, K.-W., Li, Y., Wang, W., Ma, W., Li, K., Qi, R. Z., Liu, D., Songyang, Z., & Chen, J. (2013). Whole-genome screening identifies proteins localized to distinct nuclear bodies. *Journal of Cell Biology, 203*(1), 149-164. doi:10.1083/jcb.201303145

Frank, D. J., & Roth, M. B. (1998). ncl-1 is required for the regulation of cell size and ribosomal RNA synthesis in Caenorhabditis elegans. *The Journal of Cell Biology, 140*(6), 1321-1329. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9508766

Gall, J. G. (2000). Cajal Bodies: The First 100 Years. *Annual Review of Cell and Developmental Biology, 16*(1), 273-300. doi:10.1146/annurev.cellbio.16.1.273

Gallo, A., Lo Sterzo, C., Mori, M., Di Matteo, A., Bertini, I., Banci, L., Brunori, M., & Federici, L. (2012). Structure of nucleophosmin DNA-binding domain and analysis of its complex with a G-quadruplex sequence from the c-MYC promoter. *The Journal of biological chemistry, 287*(32), 26539-26548. doi:10.1074/jbc.M112.371013

Gangwani, L., Flavell, R. A., & Davis, R. J. (2005). ZPR1 is essential for survival and is required for localization of the survival motor neurons (SMN) protein to Cajal bodies. *Molecular and cellular biology, 25*(7), 2744-2756. doi:10.1128/MCB.25.7.2744-2756.2005

Girard, C., Verheggen, C., Neel, H., Cammas, A., Vagner, S., Soret, J., Bertrand, E., & Bordonné, R. (2008). Characterization of a short isoform of human Tgs1 hypermethylase associating with small nucleolar ribonucleoprotein core proteins and produced by limited proteolytic processing. *The Journal of biological chemistry, 283*(4), 2060-2069. doi:10.1074/jbc.M704209200

Go, C. D., Knight, J. D. R., Rajasekharan, A., Rathod, B., Hesketh, G. G., Abe, K. T., Youn, J.-Y., Samavarchi-Tehrani, P., Zhang, H., Zhu, L. Y., Popiel, E., Lambert, J.-P., Coyaud, É., Cheung, S. W. T., Rajendran, D., Wong, C. J., Antonicka, H.,

Pelletier, L., Palazzo, A. F., Shoubridge, E. A., Raught, B., & Gingras, A.-C. (2021). A proximity-dependent biotinylation map of a human cell. *Nature, 595*(7865), 120-124. doi:10.1038/s41586-021-03592-2

Gopanenko, A. V., Kolobova, A. V., Meschaninova, M. I., Venyaminova, A. G., Tupikin, A. E., Kabilov, M. R., Malygin, A. A., & Karpova, G. G. (2020). Knockdown of the Ribosomal Protein eL29 in Mammalian Cells Leads to Significant Changes in Gene Expression at the Transcription Level. *Cells, 9*(5), 1228. doi:10.3390/cells9051228

Gribbon, C., Dahm, R., Prescott, A. R., & Quinlan, R. A. (2002). Association of the nuclear matrix component NuMA with the Cajal body and nuclear speckle compartments during transitions in transcriptional activity in lens cell differentiation. *European Journal of Cell Biology, 81*(10), 557-566. doi:https://doi.org/10.1078/0171-9335-00275

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 & Development*, 1-11. doi:10.1101/gad.234591.113

Handwerger, K. E., Murphy, C., & Gall, J. G. (2003). Steady-state dynamics of Cajal body components in the Xenopus germinal vesicle. *The Journal of cell biology, 160*(4), 495-504. doi:10.1083/jcb.200212024

Hao, L. t., Fuller, H. R., Lam, L. T., Le, T. T., Burghes, A. H. M., & Morris, G. E. (2007). Absence of gemin5 from SMN complexes in nuclear Cajal bodies. *BMC cell biology, 8*, 28. doi:10.1186/1471-2121-8-28

Hearst, S. M., Gilder, A. S., Negi, S. S., Davis, M. D., George, E. M., Whittom, A. A., Toyota, C. G., Husedzinovic, A., Gruss, O. J., & Hebert, M. D. (2009). Cajal-body formation correlates with differential coilin phosphorylation in primary and transformed cell lines. *Journal of cell science, 122*(Pt 11), 1872-1881. doi:10.1242/jcs.044040

Hebert, M. D., & Matera, A. G. (2000). Self-association of coilin reveals a common theme in nuclear body localization. *Molecular biology of the cell, 11*(12), 4159-4171.

Hebert, M. D., & Poole, A. R. (2017). Towards an understanding of regulating Cajal body activity by protein modification. *RNA Biology, 14*(6), 761-778. doi:10.1080/15476286.2016.1243649

Hebert, M. D., Shpargel, K. B., Ospina, J. K., Tucker, K. E., & Matera, A. G. (2002). Coilin Methylation Regulates Nuclear Body Formation. *Developmental Cell, 3*(3), 329-337. doi:10.1016/S1534-5807(02)00222-8

Hebert, M. D., Szymczyk, P. W., Shpargel, K. B., & Matera, A. G. (2001). Coilin forms the bridge between Cajal bodies and SMN, the spinal muscular atrophy protein. *Genes & development, 15*(20), 2720-2729. doi:10.1101/gad.908401

Heyn, P., Salmonowicz, H., Rodenfels, J., & Neugebauer, K. M. (2017). Activation of transcription enforces the formation of distinct nuclear bodies in zebrafish embryos. *RNA Biology, 14*(6), 752-760. doi:10.1080/15476286.2016.1255397

Hirose, T., Shu, M.-D., & Steitz, J. A. (2003). Splicing-Dependent and -Independent Modes of Assembly for Intron-Encoded Box C/D snoRNPs in Mammalian Cells. *Molecular Cell, 12*(1), 113-123. doi:10.1016/s1097-2765(03)00267-3

Hong, S., Ka, S., Kim, S., Park, Y., & Kang, S. (2003). p80 coilin, a coiled body-specific protein, interacts with ataxin-1, the SCA1 gene product. *Biochimica et biophysica acta, 1638*(1), 35-42. doi:10.1016/s0925-4439(03)00038-3

Hung, V., Udeshi, N. D., Lam, S. S., Loh, K. H., Cox, K. J., Pedram, K., Carr, S. A., & Ting, A. Y. (2016). Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nature Protocols, 11*(3), 456-475. doi:10.1038/nprot.2016.018

Husedzinovic, A., Neumann, B., Reymann, J., Draeger-Meurer, S., Chari, A., Erfle, H., Fischer, U., & Gruss, O. J. (2015). The catalytically inactive tyrosine phosphatase HD-PTP/PTPN23 is a novel regulator of SMN complex localization. *Molecular Biology of the Cell, 26*(2), 161-171. doi:10.1091/mbc.e14-06-1151

Imada, T., Shimi, T., Kaiho, A., Saeki, Y., & Kimura, H. (2021). RNA polymerase II condensate formation and association with Cajal and histone locus bodies in living human cells. *Genes to Cells, 26*(5), 298-312. doi:10.1111/gtc.12840

Isaac, C., Yang, Y., & Thomas Meier, U. (1998). Nopp140 Functions as a Molecular Link Between the Nucleolus and the Coiled Bodies. *Journal of Cell Biology, 142*(2), 319-329. doi:10.1083/jcb.142.2.319

Jady, B. E., Bertrand, E., & Kiss, T. (2004). Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. *The Journal of cell biology, 164*(5), 647-652. doi:10.1083/jcb.200310138

Jády, B. E., Darzacq, X., Tucker, K. E., Matera, A. G., Bertrand, E., & Kiss, T. (2003). Modification of Sm small nuclear RNAs occurs in the nucleoplasmic Cajal body following import from the cytoplasm. *The EMBO journal, 22*(8), 1878-1888. doi:10.1093/emboj/cdg187

Jain, A., & Vale, R. D. (2017). RNA phase transitions in repeat expansion disorders. *Nature, 546*(7657), 243-247. doi:10.1038/nature22386

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, 45*(20), 11725-11742. doi:10.1093/nar/gkx782

Jiménez-García, L. F., Segura-Valdez, M. 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*(9), 955-966. doi:10.1091/mbc.5.9.955

Kalmárová, M., Smirnov, E., Mašata, M., Koberna, K., Ligasová, A., Popov, A., & Raška, I. (2007). Positioning of NORs and NOR-bearing chromosomes in relation to nucleoli. *Journal of Structural Biology, 160*(1-3), 49-56. doi:10.1016/j.jsb.2007.06.012

Kimura, M. (2008). IRF2-binding protein-1 is a JDP2 ubiquitin ligase and an inhibitor of ATF2-dependent transcription. *FEBS Letters, 582*(19), 2833-2837. doi:10.1016/j.febslet.2008.07.033

Klingauf, M., Stanek, D., & Neugebauer, K. M. (2006). Enhancement of U4/U6 small nuclear ribonucleoprotein particle association in Cajal bodies predicted by mathematical modeling. *Molecular Biology of the Cell, 17*(12), 4972-4981. doi:10.1091/mbc.E06-06-0513

Kurihara, M., Kato, K., Sanbo, C., Shigenobu, S., Ohkawa, Y., Fuchigami, T., & Miyanari, Y. (2020). Genomic Profiling by ALaP-Seq Reveals Transcriptional Regulation by PML Bodies through DNMT3A Exclusion. *Molecular Cell, 78*(3), 493-505.e498. doi:https://doi.org/10.1016/j.molcel.2020.04.004

Lam, S. S., Martell, J. D., Kamer, K. J., Deerinck, T. J., Ellisman, M. H., Mootha, V. K., & Ting, A. Y. (2015). Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nature Methods, 12*(1), 51-54. doi:10.1038/nmeth.3179

Lam, Y. W., Lyon, C. E., & Lamond, A. I. (2002). Large-scale isolation of Cajal bodies from HeLa cells. *Molecular biology of the cell, 13*(7), 2461-2473. doi:10.1091/mbc.02-03-0034

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods, 9*(4), 357-359. doi:10.1038/nmeth.1923

Laprade, H., Querido, E., Smith, M. J., Guérit, D., Crimmins, H., Conomos, D., Pourret, E., Chartrand, P., & Sfeir, A. (2020). Single-Molecule Imaging of Telomerase RNA Reveals a Recruitment-Retention Model for Telomere Elongation. *Molecular Cell, 79*(1), 115-126.e116. doi:10.1016/j.molcel.2020.05.005

Lauria, F., Bernabo, P., Tebaldi, T., Groen, E. J. N., Perenthaler, E., Maniscalco, F., Rossi, A., Donzel, D., Clamer, M., Marchioretto, M., Omersa, N., Orri, J., Dalla Serra, M., Anderluh, G., Quattrone, A., Inga, A., Gillingwater, T. H., & Viero, G. (2020). SMN-primed ribosomes modulate the translation of transcripts related to spinal muscular atrophy. *Nat Cell Biol, 22*(10), 1239-1251. doi:10.1038/s41556-020-00577-7

Lauria, F., Bernabò, P., Tebaldi, T., Groen, E. J. N., Perenthaler, E., Maniscalco, F., Rossi, A., Donzel, D., Clamer, M., Marchioretto, M., Omersa, N., Orri, J., Dalla Serra, M., Anderluh, G., Quattrone, A., Inga, A., Gillingwater, T. H., & Viero, G. (2020). SMN-primed ribosomes modulate the translation of transcripts related to spinal muscular atrophy. *Nature Cell Biology, 22*(10), 1239-1251. doi:10.1038/s41556-020-00577-7

Lee, W.-K., Son, S. H., Jin, B.-S., Na, J.-H., Kim, S.-Y., Kim, K.-H., Kim, E. E., Yu, Y. G., & Lee, H. H. (2013). Structural and functional insights into the regulation mechanism of CK2 by IP₆ and the intrinsically disordered protein Nopp140. *Proceedings of the National Academy of Sciences, 110*(48), 19360. doi:10.1073/pnas.1304670110

Lemm, I., Girard, C., Kuhn, A. N., Watkins, N. J., Schneider, M., Bordonne, R., & Luhrmann, R. (2006). Ongoing U snRNP biogenesis is required for the integrity of Cajal bodies. *Molecular biology of the cell, 17*(7), 3221-3231. doi:10.1091/mbc.E06-03-0247

Lemos, T. A., & Kobarg, J. (2006). CGI-55 Interacts With Nuclear Proteins and Co-Localizes to p80-Coilin Positive-Coiled Bodies in the Nucleus. *Cell Biochemistry and Biophysics, 44*(3), 463-474. doi:10.1385/cbb:44:3:463

Li, D., Meier, U. T., Dobrowolska, G., & Krebs, E. G. (1997). Specific interaction between casein kinase 2 and the nucleolar protein Nopp140. *The Journal of biological chemistry, 272*(6), 3773-3779. doi:10.1074/jbc.272.6.3773

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools.

Bioinformatics (Oxford, England), 25(16), 2078-2079. doi:10.1093/bioinformatics/btp352

Li, W., Song, A. P., Zhao, F., Hu, Y. M., & Hua, M. (2013). A novel human TINP1 gene promotes cell proliferation through inhibition of p53 and p21 expression. *Oncol Rep, 30*(4), 1848-1852. doi:10.3892/or.2013.2647

Lin, Y., Protter, D. S. W., Rosen, M. K., & Parker, R. (2015). Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Molecular Cell, 60*(2), 208-219. doi:10.1016/j.molcel.2015.08.018

Liu, J., Hebert, M. D., Ye, Y., Templeton, D. J., Kung, H., & Matera, A. G. (2000). Cell cycle-dependent localization of the CDK2-cyclin E complex in Cajal (coiled) bodies. *Journal of cell science, 113 (Pt 9)*, 1543-1552.

Liu, J.-L., Wu, Z. a., Nizami, Z., Deryusheva, S., Rajendra, T. K., Beumer, K. J., Gao, H., Matera, A. G., Carroll, D., & Gall, J. G. (2009). Coilin Is Essential for Cajal Body Organization in Drosophila melanogaster. *Molecular Biology of the Cell, 20*(6), 1661-1670. doi:10.1091/mbc.E08-05-0525

Liu, Q., & Dreyfuss, G. (1996). A novel nuclear structure containing the survival of motor neurons protein. *The EMBO Journal, 15*(14), 3555-3565. doi:10.1002/j.1460-2075.1996.tb00725.x

Lopez-Delisle, L., Rabbani, L., Wolff, J., Bhardwaj, V., Backofen, R., Grüning, B., Ramírez, F., & Manke, T. (2021). pyGenomeTracks: reproducible plots for multivariate genomic datasets *Bioinformatics (Oxford, England), 37*(3), 422-423. doi:10.1093/bioinformatics/btaa692

Lyon, C. E., Bohmann, K., Sleeman, J., & Lamond, A. I. (1997). Inhibition of protein dephosphorylation results in the accumulation of splicing snRNPs and coiled bodies within the nucleolus. *Experimental Cell Research, 230*(1), 84-93. doi:10.1006/excr.1996.3380

Ma, T., Van Tine, B. A., Wei, Y., Garrett, M. D., Nelson, D., Adams, P. D., Wang, J., Qin, J., Chow, L. T., & Harper, J. W. (2000). Cell cycle-regulated phosphorylation of p220(NPAT) by cyclin E/Cdk2 in Cajal bodies promotes histone gene transcription. *Genes & development, 14*(18), 2298-2313.

Machyna, M., Heyn, P., & Neugebauer, K. M. (2013a). Cajal bodies: where form meets function. *Wiley interdisciplinary reviews. RNA, 4*(1), 17-34. doi:10.1002/wrna.1139

Machyna, M., Heyn, P., & Neugebauer, K. M. (2013b). Cajal bodies: where form meets function. *Wiley Interdisciplinary Reviews: RNA, 4*(1), 17-34. doi:10.1002/wrna.1139

Machyna, M., Kehr, S., Straube, K., Kappei, D., Buchholz, F., Butter, F., Ule, J., Hertel, J., Stadler, P. F., & Neugebauer, K. M. (2014). The coilin interactome identifies hundreds of small noncoding RNAs that traffic through Cajal bodies. *Molecular Cell, 56*(3), 389-399. doi:10.1016/j.molcel.2014.10.004

Machyna, M., Neugebauer, K. M., & Staněk, D. (2015). Coilin: The first 25 years. *RNA Biology, 12*(6), 590-596. doi:10.1080/15476286.2015.1034923

Mahmoudi, S., Henriksson, S., Weibrecht, I., Smith, S., Söderberg, O., Strömblad, S., Wiman, K. G., & Farnebo, M. (2010). WRAP53 Is Essential for Cajal Body

Formation and for Targeting the Survival of Motor Neuron Complex to Cajal Bodies. *PLoS Biology, 8*(11). doi:10.1371/journal.pbio.1000521

Mao, Y. S., Zhang, B., & Spector, D. L. (2011). Biogenesis and function of nuclear bodies. *Trends in genetics: TIG, 27*(8), 295-306. doi:10.1016/j.tig.2011.05.006

Marcos, A. T., Martín‐Doncel, E., Morejón‐García, P., Marcos‐Alcalde, I., Gómez‐Puertas, P., Segura‐Puimedon, M., Armengol, L., Navarro‐Pando, J. M., & Lazo, P. A. (2020). VRK1 (Y213H) homozygous mutant impairs Cajal bodies in a hereditary case of distal motor neuropathy. *Annals of Clinical and Translational Neurology, 7*(5), 808-818. doi:10.1002/acn3.51050

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal, 17*(1), 10. doi:10.14806/ej.17.1.200

Martín-Doncel, E., Rojas, A. M., Cantarero, L., & Lazo, P. A. (2019). VRK1 functional insufficiency due to alterations in protein stability or kinase activity of human VRK1 pathogenic variants implicated in neuromotor syndromes. *Scientific Reports, 9*(1). doi:10.1038/s41598-019-49821-7

Marzluff, W. F., & Koreski, K. P. (2017). Birth and Death of Histone mRNAs. *Trends in Genetics, 33*(10), 745-759. doi:10.1016/j.tig.2017.07.014

Massenet, S., Bertrand, E., & Verheggen, C. (2017). Assembly and trafficking of box C/D and H/ACA snoRNPs. *RNA Biology, 14*(6), 680-692. doi:10.1080/15476286.2016.1243646

Meier, U. T. (2017). RNA modification in Cajal bodies. *RNA Biology, 14*(6), 693-700. doi:10.1080/15476286.2016.1249091

Meier, U. T., & Blobel, G. (1994). NAP57, a mammalian nucleolar protein with a putative homolog in yeast and bacteria. *Journal of Cell Biology, 127*(6), 1505-1514. doi:10.1083/jcb.127.6.1505

Meister, G., Eggert, C., & Fischer, U. (2002). SMN-mediated assembly of RNPs: a complex story. *Trends in Cell Biology, 12*(10), 472-478. doi:10.1016/s0962-8924(02)02371-1

Monneron, A., & Bernhard, W. (1969). Fine structural organization of the interphase nucleus in some mammalian cells. *Journal of Ultrastructure Research, 27*(3), 266-288. doi:https://doi.org/10.1016/S0022-5320(69)80017-1

Moorhead, G. B. G., Trinkle-Mulcahy, L., & Ulke-Lemée, A. (2007). Emerging roles of nuclear protein phosphatases. *Nature Reviews Molecular Cell Biology, 8*(3), 234-244. doi:10.1038/nrm2126

Morgan, G. T., Doyle, O., Murphy, C., & Gall, J. G. (2000). RNA polymerase II in Cajal bodies of amphibian oocytes. *Journal of structural biology, 129*(2-3), 258-268. doi:10.1006/jsbi.2000.4231

Mouaikel, J., Narayanan, U., Verheggen, C., Matera, A. G., Bertrand, E., Tazi, J., & Bordonne, R. (2003). Interaction between the small-nuclear-RNA cap hypermethylase and the spinal muscular atrophy protein, survival of motor neuron. *EMBO Reports, 4*(6), 616-622. doi:10.1038/sj.embor.embor863

Murphy, C., Wang, Z., Roeder, R. G., & Gall, J. G. (2002). RNA polymerase III in Cajal bodies and lampbrush chromosomes of the Xenopus oocyte nucleus. *Molecular biology of the cell, 13*(10), 3466-3476. doi:10.1091/mbc.E02-05-0281

Na, J. H., Lee, W. K., Kim, Y., Jeong, C., Song, S. S., Cha, S. S., Han, K. H., Shin, Y. K., & Yu, Y. G. (2016). Biophysical characterization of the structural change of Nopp140, an intrinsically disordered protein, in the interaction with CK2α. *Biochemical and biophysical research communications, 477*(2), 181-187. doi:10.1016/j.bbrc.2016.06.040

Nandagopal, N., & Roux, P. P. (2015). Regulation of global and specific mRNA translation by the mTOR signaling pathway. *Translation (Austin, Tex.), 3*(1), e983402-e983402. doi:10.4161/21690731.2014.983402

Narayanan, A. (1999). Nucleolar localization signals of Box H/ACA small nucleolar RNAs. *The EMBO Journal, 18*(18), 5120-5130. doi:10.1093/emboj/18.18.5120

Navascues, J., Bengoechea, R., Tapia, O., Casafont, I., Berciano, M. T., & Lafarga, M. (2008). SUMO-1 transiently localizes to Cajal bodies in mammalian neurons. *Journal of structural biology, 163*(2), 137-146. doi:10.1016/j.jsb.2008.04.013

Nesic, D., Tanackovic, G., & Kramer, A. (2004). A role for Cajal bodies in the final steps of U2 snRNP biogenesis. *Journal of cell science, 117*(Pt 19), 4423-4433. doi:10.1242/jcs.01308

Novotný, I., Blažíková, M., Staněk, D., Herman, P., & Malinsky, J. (2011). In vivo kinetics of U4/U6·U5 tri-snRNP formation in Cajal bodies. *Molecular Biology of the Cell, 22*(4), 513-523. doi:10.1091/mbc.E10-07-0560

Novotný, I., Malinová, A., Stejskalová, E., Matějů, D., Klimešová, K., Roithová, A., Švéda, M., Knejzlík, Z., & Staněk, D. (2015). SART3-Dependent Accumulation of Incomplete Spliceosomal snRNPs in Cajal Bodies. *Cell Reports, 10*(3), 429-440. doi:10.1016/j.celrep.2014.12.030

Ospina, J. K., Gonsalvez, G. B., Bednenko, J., Darzynkiewicz, E., Gerace, L., & Matera, A. G. (2005). Cross-Talk between Snurportin1 Subdomains. *Molecular Biology of the Cell, 16*(10), 4660-4671. doi:10.1091/mbc.e05-04-0316

Papageorgiou, D. N., Demmers, J., & Strouboulis, J. (2013). NP-40 reduces contamination by endogenous biotinylated carboxylases during purification of biotin tagged nuclear proteins. *Protein Expression and Purification, 89*(1), 80-83. doi:https://doi.org/10.1016/j.pep.2013.02.015

Paternoga, H., Früh, A., Kunze, R., Bradatsch, B., Baßler, J., & Hurt, E. (2020). Mutational Analysis of the Nsa2 N-Terminus Reveals Its Essential Role in Ribosomal 60S Subunit Assembly. *International journal of molecular sciences, 21*(23), 9108. doi:10.3390/ijms21239108

Pillai, R. S. (2001). Purified U7 snRNPs lack the Sm proteins D1 and D2 but contain Lsm10, a new 14 kDa Sm D1-like protein. *The EMBO Journal, 20*(19), 5470-5479. doi:10.1093/emboj/20.19.5470

Pogacic, V., Dragon, F. O., & Filipowicz, W. (2000). Human H/ACA Small Nucleolar RNPs and Telomerase Share Evolutionarily Conserved Proteins NHP2 and NOP10. *Molecular and Cellular Biology, 20*(23), 9028-9040. doi:10.1128/mcb.20.23.9028-9040.2000

Polak, P. E., Simone, F., Kaberlein, J. J., Luo, R. T., & Thirman, M. J. (2003). ELL and EAF1 are Cajal body components that are disrupted in MLL-ELL leukemia. *Molecular biology of the cell, 14*(4), 1517-1528. doi:10.1091/mbc.E02-07-0394

Pradet-Balade, B., Girard, C., Boulon, S., Paul, C., Azzag, K., Bordonné, R., Bertrand, E., & Verheggen, C. (2011). CRM1 controls the composition of nucleoplasmic pre-snoRNA complexes to licence them for nucleolar transport. *The EMBO journal, 30*(11), 2205-2218. doi:10.1038/emboj.2011.128

Qin, G., Wang, X., Ye, S., Li, Y., Chen, M., Wang, S., Qin, T., Zhang, C., Li, Y., Long, Q., Hu, H., Shi, D., Li, J., Zhang, K., Zhai, Q., Tang, Y., Kang, T., Lan, P., Xie, F., Lu, J., & Deng, W. (2020). NPM1 upregulates the transcription of PD-L1 and suppresses T cell activity in triple-negative breast cancer. *Nature Communications, 11*(1), 1669. doi:10.1038/s41467-020-15364-z

Qin, W., Cho, K. F., Cavanagh, P. E., & Ting, A. Y. (2021). Deciphering molecular interactions by proximity labeling. *Nature Methods, 18*(2), 133-143. doi:10.1038/s41592-020-01010-5

Ramón y Cajal, S. (1903). Un sencillo metodo de coloracion seletiva del reticulo protoplasmatico y sus efectos en los diversos organos nerviosos de vertebrados e invertebrados. *Trab. Lab. Invest. Biol., 12*, 491-495.

Raška, I., Andrade, L. E., Ochs, R. L., Chan, E. K., Chang, C. M., Roos, G., & Tan, E. M. (1991). Immunological and ultrastructural studies of the nuclear coiled body with autoimmune antibodies. *Experimental Cell Research, 195*(1), 27-37. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2055273

Raska, I., Michel, L. S., Jarnik, M., Dundr, M., Fakan, S., Gasser, S., Gassmann, M., Hubscher, U., Izaurralde, E., & Martinez, E. (1991). Ultrastructural cryoimmunocytochemistry is a convenient tool for the study of DNA replication in cultured cells. *Journal of electron microscopy technique, 18*(2), 91-105. doi:10.1002/jemt.1060180202

Raška, 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. *Journal of Structural Biology, 104*(1-3), 120-127. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2088441

Richard, P., Kiss, A. M., Darzacq, X., & Kiss, T. S. (2006). Cotranscriptional Recognition of Human Intronic Box H/ACA snoRNAs Occurs in a Splicing-Independent Manner. *Molecular and Cellular Biology, 26*(7), 2540-2549. doi:10.1128/mcb.26.7.2540-2549.2006

Saha, S., Weber, C. A., Nousch, M., Adame-Arana, O., Hoege, C., Hein, M. Y., Osborne-Nishimura, E., Mahamid, J., Jahnel, M., Jawerth, L., Pozniakovski, A., Eckmann, C. R., Jülicher, F., & Hyman, A. A. (2016). Polar Positioning of Phase-Separated Liquid Compartments in Cells Regulated by an mRNA Competition Mechanism. *Cell, 166*(6), 1572-1584.e1516. doi:10.1016/j.cell.2016.08.006

Salzler, H. R., Tatomer, D. C., Malek, P. Y., McDaniel, S. L., Orlando, A. N., Marzluff, W. F., & Duronio, R. J. (2013). A Sequence in the Drosophila H3-H4 Promoter Triggers Histone Locus Body Assembly and Biosynthesis of Replication-Coupled Histone mRNAs. *Developmental Cell, 24*(6), 623-634. doi:10.1016/j.devcel.2013.02.014

Samarsky, D. A. (1998). The snoRNA box C/D motif directs nucleolar targeting and also couples snoRNA synthesis and localization. *The EMBO Journal, 17*(13), 3747-3757. doi:10.1093/emboj/17.13.3747

Santama, N., Ogg, S. C., Malekkou, A., Zographos, S. E., Weis, K., & Lamond, A. I. (2005). Characterization of hCINAP, a novel coilin-interacting protein encoded by a transcript from the transcription factor TAFIID32 locus. *The Journal of biological chemistry, 280*(43), 36429-36441. doi:10.1074/jbc.M501982200

Sawyer, I. A., Hager, G. L., & Dundr, M. (2017). Specific genomic cues regulate Cajal body assembly. *RNA Biology, 14*(6), 791-803. doi:10.1080/15476286.2016.1243648

Schaffert, N., Hossbach, M., Heintzmann, R., Achsel, T., & Luhrmann, R. (2004). RNAi knockdown of hPrp31 leads to an accumulation of U4/U6 di-snRNPs in Cajal bodies. *The EMBO journal, 23*(15), 3000-3009. doi:10.1038/sj.emboj.7600296

Schilling, M., Prusty, A. B., Boysen, B., Oppermann, F. S., Riedel, Y. L., Husedzinovic, A., Rasouli, H., König, A., Ramanathan, P., Reymann, J., Erfle, H., Daub, H., Fischer, U., & Gruss, O. J. (2021). TOR signaling regulates liquid phase separation of the SMN complex governing snRNP biogenesis. *Cell Reports, 35*(12), 109277. doi:https://doi.org/10.1016/j.celrep.2021.109277

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods, 9*(7), 676-682. doi:10.1038/nmeth.2019

Schmidt, J. C., & Cech, T. R. (2015). Human telomerase: biogenesis, trafficking, recruitment, and activation. *Genes & Development, 29*(11), 1095-1105. doi:10.1101/gad.263863.115

Schmidt, J. C., Zaug, A. J., & Cech, T. R. (2016). Live Cell Imaging Reveals the Dynamics of Telomerase Recruitment to Telomeres. *Cell, 166*(5), 1188-1197.e1189. doi:10.1016/j.cell.2016.07.033

Schulz, S., Chachami, G., Kozaczkiewicz, L., Winter, U., Stankovic‐Valentin, N., Haas, P., Hofmann, K., Urlaub, H., Ovaa, H., Wittbrodt, J., Meulmeester, E., & Melchior, F. (2012). Ubiquitin‐specific protease‐like 1 (USPL1) is a SUMO isopeptidase with essential, non‐catalytic functions. *EMBO Reports, 13*(10), 930-938. doi:10.1038/embor.2012.125

Senapati, P., Bhattacharya, A., Das, S., Dey, S., Sudarshan, D., G, S., Vishwakarma, J., Sudevan, S., Ramachandran, R., Maliekal, T. T., & Kundu, T. K. (2021). Histone chaperone Nucleophosmin regulates transcription of key genes involved in oral tumorigenesis. *Molecular and Cellular Biology*, Mcb0066920. doi:10.1128/mcb.00669-20

Shay, J. W., & Wright, W. E. (2019). Telomeres and telomerase: three decades of progress. *Nature Reviews Genetics, 20*(5), 299-309. doi:10.1038/s41576-019-0099-1

Shevtsov, S. P., & Dundr, M. (2011). Nucleation of nuclear bodies by RNA. *Nature Cell Biology, 13*(2), 167. doi:10.1038/ncb2157

Sivaramakrishnan, G., Sun, Y., Tan, S. K., & Lin, V. C. L. (2009). Dynamic localization of tripartite motif-containing 22 in nuclear and nucleolar bodies. *Experimental Cell Research, 315*(8), 1521-1532. doi:10.1016/j.yexcr.2009.01.028

Smith, E. R., Lin, C., Garret, A., Thornton, J., Mohaghegh, N., Hu, D., Jackson, J., Saraf, A., Swanson, S., Seidel, C., Florens, L., Washburn, M., Eissenberg, J., &

Shilatifard, A. (2011). The Little Elongation Complex Regulates Small Nuclear RNA Transcription. *Molecular Cell, 44*(6), 954-965. doi:10.1016/j.molcel.2011.12.008

Staley, J. P., & Guthrie, C. (1998). Mechanical Devices of the Spliceosome: Motors, Clocks, Springs, and Things. *Cell, 92*(3), 315-326. doi:10.1016/s0092-8674(00)80925-3

Staněk, D., & Neugebauer, K. M. (2004). Detection of snRNP assembly intermediates in Cajal bodies by fluorescence resonance energy transfer. *The Journal of cell biology, 166*(7), 1015-1025. Retrieved from files/1300/PMC2172029.html

files/1296/1015.html

Stanek, D., Pridalová-Hnilicová, J., Novotný, I., Huranová, M., Blazíková, M., Wen, X., Sapra, A. K., & Neugebauer, K. M. (2008). Spliceosomal small nuclear ribonucleoprotein particles repeatedly cycle through Cajal bodies. *Molecular Biology of the Cell, 19*(6), 2534-2543. doi:10.1091/mbc.E07-12-1259

Staněk, D., Rader, S. D., Klingauf, M., & Neugebauer, K. M. (2003). Targeting of U4/U6 small nuclear RNP assembly factor SART3/p110 to Cajal bodies. *The Journal of Cell Biology, 160*(4), 505-516. doi:10.1083/jcb.200210087

Strzelecka, M., Oates, A. C., & Neugebauer, K. M. (2010). Dynamic control of Cajal body number during zebrafish embryogenesis. *Nucleus (Austin, Tex.), 1*(1), 96-108. doi:10.4161/nucl.1.1.10680

Strzelecka, M., Trowitzsch, S., Weber, G., Lührmann, R., Oates, A. C., & Neugebauer, K. M. (2010). Coilin-dependent snRNP assembly is essential for zebrafish embryogenesis. *Nature Structural & Molecular Biology, 17*(4), 403-409. doi:10.1038/nsmb.1783

Sun, J., Xu, H., Subramony, S. H., & Hebert, M. D. (2005). Interactions between coilin and PIASy partially link Cajal bodies to PML bodies. *Journal of cell science, 118*(Pt 21), 4995-5003. doi:10.1242/jcs.02613

Tapia, O., Bengoechea, R., Berciano, M. T., & Lafarga, M. (2010). Nucleolar targeting of coilin is regulated by its hypomethylation state. *Chromosoma, 119*(5), 527-540. doi:10.1007/s00412-010-0276-7

Tong, L. (2013). Structure and function of biotin-dependent carboxylases. *Cellular and molecular life sciences : CMLS, 70*(5), 863-891. doi:10.1007/s00018-012-1096-0

Trinkle-Mulcahy, L., & Sleeman, J. E. (2017). The Cajal body and the nucleolus: "In a relationship" or "It's complicated"? *RNA Biology, 14*(6), 739-751. doi:10.1080/15476286.2016.1236169

Tripsianes, K., Madl, T., Machyna, M., Fessas, D., Englbrecht, C., Fischer, U., Neugebauer, K. M., & Sattler, M. (2011). Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins. *Nature Structural & Molecular Biology, 18*(12), 1414-1420. doi:10.1038/nsmb.2185

Tycowski, K. T., Shu, M.-D., Kukoyi, A., & Steitz, J. A. (2009). A conserved WD40 protein binds the Cajal body localization signal of scaRNP particles. *Molecular Cell, 34*(1), 47-57. doi:10.1016/j.molcel.2009.02.020

van Koningsbruggen, S., Dirks, R. W., Mommaas, A. M., Onderwater, J. J., Deidda, G., Padberg, G. W., Frants, R. R., & van der Maarel, S. M. (2004). FRG1P is localised

in the nucleolus, Cajal bodies, and speckles. *Journal of medical genetics, 41*(4), e46.

Venteicher, A. S., Abreu, E. B., Meng, Z., McCann, K. E., Terns, R. M., Veenstra, T. D., Terns, M. P., & Artandi, S. E. (2009). A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis. *Science (New York, N.Y.), 323*(5914), 644-648. doi:10.1126/science.1165357

Verheggen, C. (2002). Mammalian and yeast U3 snoRNPs are matured in specific and related nuclear compartments. *The EMBO Journal, 21*(11), 2736-2745. doi:10.1093/emboj/21.11.2736

Walker, M. P., Tian, L., & Matera, A. G. (2009). Reduced Viability, Fertility and Fecundity in Mice Lacking the Cajal Body Marker Protein, Coilin. *PLOS ONE, 4*(7). doi:10.1371/journal.pone.0006171

Walker, R. F., Liu, J. S., Peters, B. A., Ritz, B. R., Wu, T., Ophoff, R. A., & Horvath, S. (2015). Epigenetic age analysis of children who seem to evade aging. *Aging, 7*(5), 334-339. doi:10.18632/aging.100744

Wang, Q., Sawyer, I. A., Sung, M.-H., Sturgill, D., Shevtsov, S. P., Pegoraro, G., Hakim, O., Baek, S., Hager, G. L., & Dundr, M. (2016). Cajal bodies are linked to genome conformation. *Nature Communications, 7*, 10966. doi:10.1038/ncomms10966

Warner, J. R., & McIntosh, K. B. (2009). How Common Are Extraribosomal Functions of Ribosomal Proteins? *Molecular Cell, 34*(1), 3-11. doi:https://doi.org/10.1016/j.molcel.2009.03.006

Wei, H.-H., Fan, X.-J., Hu, Y., Tian, X.-X., Guo, M., Mao, M.-W., Fang, Z.-Y., Wu, P., Gao, S.-X., Peng, C., Yang, Y., & Wang, Z. (2021). A systematic survey of PRMT interactomes reveals the key roles of arginine methylation in the global control of RNA splicing and translation. *Science Bulletin, 66*(13), 1342-1357. doi:https://doi.org/10.1016/j.scib.2021.01.004

Wei, T., Najmi, S. M., Liu, H., Peltonen, K., Kucerova, A., Schneider, D. A., & Laiho, M. (2018). Small-Molecule Targeting of RNA Polymerase I Activates a Conserved Transcription Elongation Checkpoint. *Cell Reports, 23*(2), 404-414. doi:10.1016/j.celrep.2018.03.066

Wu, Z., Murphy, C., & Gall, J. G. (1994). Human p80-coilin is targeted to sphere organelles in the amphibian germinal vesicle. *Molecular biology of the cell, 5*(10), 1119-1127.

Xu, H., & Hebert, M. D. (2005). A novel EB-1/AIDA-1 isoform, AIDA-1c, interacts with the Cajal body protein coilin. *BMC cell biology, 6*(1), 23. doi:10.1186/1471-2121-6-23

Xu, H., Pillai, R. S., Azzouz, T. N., Shpargel, K. B., Kambach, C., Hebert, M. D., Schumperli, D., & Matera, A. G. (2005). The C-terminal domain of coilin interacts with Sm proteins and U snRNPs. *Chromosoma, 114*(3), 155-166. doi:10.1007/s00412-005-0003-y

Yamada, M., Sato, T., Shimohata, T., Hayashi, S., Igarashi, S., Tsuji, S., & Takahashi, H. (2001). Interaction between Neuronal Intranuclear Inclusions and Promyelocytic Leukemia Protein Nuclear and Coiled Bodies in CAG Repeat Diseases. *The American Journal of Pathology, 159*(5), 1785-1795. doi:10.1016/s0002-9440(10)63025-8

Yang, P. K., Hoareau, C., Froment, C., Monsarrat, B., Henry, Y., & Chanfreau, G. (2005). Cotranscriptional Recruitment of the Pseudouridylsynthetase Cbf5p and of the RNA Binding Protein Naf1p during H/ACA snoRNP Assembly. *Molecular and Cellular Biology, 25*(8), 3295-3304. doi:10.1128/mcb.25.8.3295-3304.2005

Yeung, K. T., Das, S., Zhang, J., Lomniczi, A., Ojeda, S. R., Xu, C.-F., Neubert, T. A., & Samuels, H. H. (2011). A Novel Transcription Complex That Selectively Modulates Apoptosis of Breast Cancer Cells through Regulation of FASTKD2. *Molecular and Cellular Biology, 31*(11), 2287-2298. doi:10.1128/mcb.01381-10

Yildirim, A., Mozaffari-Jovin, S., Wallisch, A.-K., Schäfer, J., Ludwig, S. E. J., Urlaub, H., Lührmann, R., & Wolfrum, U. (2021). SANS (USH1G) regulates pre-mRNA splicing by mediating the intra-nuclear transfer of tri-snRNP complexes. *Nucleic Acids Research, 49*(10), 5845-5866. doi:10.1093/nar/gkab386

Young, P. J., Le, T. T., Dunckley, M., Nguyen, T. M., Burghes, A. H., & Morris, G. E. (2001). Nuclear gems and Cajal (coiled) bodies in fetal tissues: nucleolar distribution of the spinal muscular atrophy protein, SMN. *Experimental Cell Research, 265*(2), 252-261. doi:10.1006/excr.2001.5186

Zhang, H., Elbaum-Garfinkle, S., Langdon, E. M., Taylor, N., Occhipinti, P., Bridges, Andrew A., Brangwynne, Clifford P., & Gladfelter, Amy S. (2015). RNA Controls PolyQ Protein Phase Transitions. *Molecular Cell, 60*(2), 220-230. doi:10.1016/j.molcel.2015.09.017

Zhong, F., Savage, S. A., Shkreli, M., Giri, N., Jessop, L., Myers, T., Chen, R., Alter, B. P., & Artandi, S. E. (2011). Disruption of telomerase trafficking by TCAB1 mutation causes dyskeratosis congenita. *Genes & Development, 25*(1), 11-16. doi:10.1101/gad.2006411