

The Texas Medical Center Library

DigitalCommons@TMC

The University of Texas MD Anderson Cancer
Center UTHealth Graduate School of
Biomedical Sciences Dissertations and Theses
(Open Access)

The University of Texas MD Anderson Cancer
Center UTHealth Graduate School of
Biomedical Sciences

8-2011

A Cell Biological Determination of Integrator Subunit Localization

Sarah B. May

Follow this and additional works at: https://digitalcommons.library.tmc.edu/utgsbs_dissertations



Part of the [Biology Commons](#), and the [Cell Biology Commons](#)

Recommended Citation

May, Sarah B., "A Cell Biological Determination of Integrator Subunit Localization" (2011). *The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access)*. 173.

https://digitalcommons.library.tmc.edu/utgsbs_dissertations/173

This Thesis (MS) is brought to you for free and open access by the The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences at DigitalCommons@TMC. It has been accepted for inclusion in The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access) by an authorized administrator of DigitalCommons@TMC. For more information, please contact digitalcommons@library.tmc.edu.

The
TMC LIBRARY
Health Sciences Resource Center

A CELL BIOLOGICAL DETERMINATION OF INTEGRATOR SUBUNIT
LOCALIZATION

By:

Sarah Beth May, B.S.

APPROVED:

Eric Wagner, Ph.D. (Supervisory Advisor)

Michael Blackburn, Ph.D.

Phillip Carpenter, Ph.D.

Joel Neilson, Ph.D.

Ambro van Hoof, Ph.D.

APPROVED:

Dean, The University of Texas
Graduate School of Biomedical Sciences at Houston

A CELL BIOLOGICAL DETERMINATION OF INTEGRATOR SUBUNIT LOCALIZATION

A

THESIS

Presented to the Faculty of

The University of Texas

Health Science Center at Houston

and

The University of Texas

M.D. Anderson Cancer Center

Graduate School of Biomedical Sciences

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

by

Sarah Beth May, B.S.

Houston, Texas

August 2011

Dedication

I dedicate this thesis to my parents, Jeffrey and Margaret, and my sister, Caryn. Their unfaltering love and support through the years has made the completion of this thesis possible.

Acknowledgements

First, I would like to thank my advisor, Eric Wagner, for his guidance and support throughout my thesis work. You took me into your lab with almost no molecular biology experience and taught me everything I needed to know to complete my Master's degree, and for that I am eternally grateful.

I would also like to thank all the members of my committee: Drs. Michael Blackburn, Phillip Carpenter, Joel Neilson, and Ambro van Hoof. Your constructive comments, ideas, and support were instrumental to the completion of my project.

Thanks to the members of the Wagner lab both past and present: Nadar Ezzeddine, Jiandong Chen, Patience Masamha, Anupama Sataluri, Natoya Peart and Todd Albrecht for teaching me many of the techniques I needed to complete this project and generally making the lab an enjoyable place. Thanks also to a collaborator of the lab, Mirek Dundr from Rosalind Franklin University in Chicago, for completing the iFRAP study used in my work.

Big thanks to Olga Chumakova in the Cytodynamic Imaging Facility for training me to use the confocal microscope, and trusting me with it after hours. Without your help, this project would not have been possible.

I am extremely grateful to be a part of the BMB department and program. The camaraderie and support I have found in this program have been wonderful. Thanks especially to the BMB staff for everything from planning social events to help with room reservations. You have truly made being a student in this program a delight.

Thanks to all of my wonderful friends, especially, Karianne New, Kari Brewer, Heather Highland, and Julie Allen for always cheering me up, giving me encouragement and listening to me complain. Without your friendship I would be lost.

Finally, thanks to my amazing family. To my parents, Jeffrey and Margaret, for all your love and support, and for questioning every grade lower than an A, and

generally never accepting anything less than my best. To my sister Caryn, for being my best friend, you're awesome. (P.S. I'm still smarter!) Last, but not least, to my furry family, especially my cat, Sherlock, for keeping my blood pressure down and my head from exploding while writing my thesis.

A cell biological determination of Integrator subunit localization

Sarah Beth May, B.S.

Supervisory Advisor: Eric Wagner, Ph.D.

Uridine-rich small nuclear (U snRNAs), with the exception of the U6 snRNA, are RNA polymerase II (RNAPII) transcripts. The mechanism of 3' cleavage of snRNAs has been unknown until recently. This area was greatly advanced when 12 of the Integrator complex subunits (IntS) were purified in 2005 through their interaction with the C-terminal domain (CTD) of the large subunit (Rpb1) of RNAPII. Subsequently, our lab performed a genome-wide RNAi screen that identified two more members of the complex that we have termed IntS13 and IntS14. We have determined that IntS9 and 11 mediate the 3' cleavage of snRNAs, but the exact function of the other subunits remains unknown. However, through the use of a U7 snRNA-GFP reporter and RNAi knockdown of the Integrator subunits in *Drosophila* S2 cells, we have shown that all subunits are required for the proper processing of snRNAs, albeit to differing degrees. Because snRNA transcription takes place in the nucleus of the cell, it is expected that all of the Integrator subunits would exhibit nuclear localization, but the knowledge of discrete subnuclear localization (i.e. to Cajal bodies) of any of the subunits could provide important clues to the function of that subunit. In this study, we used a cell biological approach to determine the localization of the 14 Integrator subunits. **We hypothesized that the majority of**

the subunits would be nuclear, however, a few would display distinct localization to the Cajal bodies, as this is where snRNA genes are localized and transcribed. The specific aims and results are: **1. To determine the subcellular localization of the 14 Integrator subunits.** To accomplish this, mCherry and GFP tagged clones were generated for each of the 14 *Drosophila* and human Integrator subunits. Confocal microscopy studies revealed that the majority of the subunits were diffuse in the nucleus, however, IntS3 formed discrete subnuclear foci. Surprisingly, two of the subunits, IntS2 and 7 were observed in cytoplasmic foci. **2. To further characterize Integrator subunits with unique subcellular localizations.** Colocalization studies with endogenous IntS3 and Cajal body marker, coilin, showed that these two proteins overlap, and from this we concluded that IntS3 localized to Cajal bodies. Additionally, colocalization studies with mCherry-tagged IntS2 and 7 and the P body marker, Dcp1, revealed that these proteins colocalize as well. IntS7, however, is more stable in cytoplasmic foci than Dcp1. It was also shown through RNAi knockdown of Integrator subunits, that the cytoplasmic localization of IntS2 and 7 is dependent on the expression of IntS1 and 11 in S2 cells.

Table of Contents

Approval Sheet.....	i
Title Page.....	ii
Dedication.....	iii
Acknowledgements.....	iv
Abstract.....	vi
Table of Contents.....	viii
List of Figures.....	xi
List of Tables.....	xiii
 Chaper 1: Introduction.....	 1
Small Nuclear RNAs.....	2
3' End Formation of Polymerase II Transcripts.....	3
The Integrator Complex.....	11
snRNP Biogenesis.....	17
Subcellular Localizations of RNPs.....	22
Cajal Bodies.....	23
Histone Locus Bodies.....	24
Nuclear Speckles.....	24
P Bodies.....	25
Stress Granules.....	27
U Bodies.....	30

Summary of Work.....	31
Chapter 2: Materials and Methods.....	33
Cell Culture.....	34
Transfections.....	34
Cloning Reactions.....	35
Generation of Tagged Integrator Clones.....	36
Generation of GFP hIntS7 Deletion Mutants.....	45
Cell Fixation and Immunofluorescence.....	45
RNAi.....	55
Western Blot.....	59
Chapter 3: Results.....	61
The Localization of Integrator Subunits.....	62
Human IntS3 Colocalizes with Cajal Body Marker Coilin in the Nucleus.....	70
Human and <i>Drosophila</i> IntS2 and 7 Colocalize with P Body Marker Dcp1 in Discrete Cytoplasmic Foci.....	73
IntS1 and 11 Are Required for IntS2 and 7 Localization to Cytoplasmic Foci in <i>Drosophila</i> S2 Cells.....	74
Deletions in Human IntS7 Gene Have Little Effect on hIntS7 Localization...90	
Depletion of Human IntS7 Leads to Increased snRNA Misprocessing.....90	
Chapter 4: Conclusion and Future Directions.....	103
The Localization of IntS3 in Cajal Bodies.....	104
Integrators 2 and 7 Display a Conserved Phenotype of P Body Localization.....	105

Domain Analysis of IntS7 with Respect to Its Cellular Localization.....	110
References.....	113
Vita.....	119

List of Figures

Figure 1: Human snRNA genes vs. mRNA genes.....	4
Figure 2: A comparison of the 3' end formation of three major RNAPII transcripts....	6
Figure 3: The Integrator complex mediates snRNA 3' end processing.....	9
Figure 4: Schematic of Integrator complex subunits.....	15
Figure 5: The biogenesis of U snRNPs.....	18
Figure 6: The mRNA cycle.....	28
Figure 7: Diagram of pUB Cherry vector.....	41
Figure 8: Schematic of <i>Drosophila</i> and human tagged Integrator clones.....	49
Figure 9: Schematic of hIntS7 deletion mutants.....	53
Figure 10: Positions of <i>Drosophila</i> dsRNA sequences.....	57
Figure 11: Confocal images of human Integrator subunit localizations.....	64
Figure 12: Confocal images of <i>Drosophila</i> Integrator subunit localizations.....	67
Figure 13: hIntS3 colocalizes with coilin.....	71
Figure 14: Cherry hIntS2 and 7 colocalize with GFP Dcp1.....	75
Figure 15: Endogenous hIntS7 localization.....	77
Figure 16: Cherry dIntS2 and 7 colocalize with GFP Dcp1.....	79
Figure 17: Effects of Integrator subunit knockdown on dIntS7 localization.....	84
Figure 18: Quantification of the effects of dIntS12 knockdown on dIntS2 and 7 localization.....	86
Figure 19: Quantification of the effects of dIntS1 knockdown on dIntS2 and 7 localization.....	88

Figure 20: Quantification of the effects of dlIntS11 knockdown on dlIntS2 and 7
localization.....91

Figure 21: Six species alignment of IntS7.....93

Figure 22: GFP tagged hIntS7 deletion mutant localizations.....95

Figure 23: Depletion of hIntS7 leads to misprocessing of snRNA in human cells..100

Figure 24: Inverse FRAP of mCherry hIntS7 and GFP Dcp1.....107

List of Tables

Table 1: Primer sets and restriction enzymes used to generate pcDNA4 N/C-terminal mCherry and GFP vectors.....	37
Table 2: Primer sets and restriction enzymes used to generate mCherry and GFP tagged human Integrator subunit clones.....	39
Table 3: Primer sets and restriction enzymes used to generate the <i>Drosophila</i> pUB Cherry vector.....	43
Table 4: Primer sets and restriction enzymes used to generate mCherry tagged <i>Drosophila</i> Integrator subunit clones.....	47
Table 5: Primer sets and restriction enzymes used to generate human IntS7 deletion fragments.....	51
Table 6: Summary of Integrator subunit knockdown experiment.....	82

Chapter 1: Introduction

Small Nuclear RNAs

Uridine-rich small nuclear (U snRNAs) are a class of ubiquitously expressed small non-coding RNAs present in all cells. These critical RNAs assemble with proteins to form small nuclear ribonucleoproteins (snRNPs) and act in concert with RNA polymerase II (RNAPII) to remove introns from pre-mRNAs and process the 3' end of histone mRNAs [1, 2]. They include the U1, U2, U4, U5, and U6 snRNAs of the major spliceosome, the U4atac, U6atac, U11, and U12 snRNAs of the minor spliceosome, and the U7 snRNA required for histone 3' end processing [3-5]. All of these RNAs, with the exception of the RNAPIII-transcribed U6 and U6atac snRNAs, are transcribed by RNAPII [6, 7].

The snRNA genes can be divided into two classes, the Sm-class consisting of the U1, U2, U4, U4atac, U5, U7, U11, and U12 snRNAs, and the Lsm-class consisting of the U6 and U6atac snRNAs. The Sm-class snRNAs have a 5'-trimethylguanosine cap, a 3' stem loop, and an Sm site that serves as the binding site for the seven Sm proteins. The Sm-class snRNAs are arranged in a similar manner to mRNA genes and a complex of proteins called the Integrator complex carries out their 3' end formation. In contrast, the Lsm-class snRNAs, U6 and U6atac, have a monomethylphosphate cap on their 5' ends and a 3' stem loop followed by a stretch of uridines that serves as the binding site for the ring of Lsm proteins. This stretch of uridines also serves as a RNAPIII transcription terminator [1].

Multiple copies of the snRNA genes are present in the genome, and they are typically found in clusters [8-10]. The genes are much simpler than their mRNA-

encoding counterparts, having no introns or 3' polyadenylation as well as lacking a TATA box and an open reading frame [11]. Instead, the promoter contains two alternative elements, a distal sequence element (DSE) that acts as a transcriptional enhancer, and a proximal sequence element (PSE) that is required for snRNA transcription [12]. At the 3' end of the gene, the 3' box, with the sequence GTTTN₀₋₃AAAPuNNAGA where N is any nucleotide and Pu is a purine, lays 9-19 nucleotides downstream of the 3' cleavage site of the mature snRNA and is required for the proper formation of the snRNA [13-15]. A comparison of a snRNA gene to a typical mRNA gene is shown in Figure 1.

3' End Formation of RNA Polymerase II Transcripts

The 3' end formation of the major RNAPII transcripts: poly(A)⁺ mRNAs, histone mRNAs, and snRNAs are similar in many ways, with a few key differences (Figure 2). The poly(A)⁺ pre-mRNAs contain two *cis*-elements defining their 3' ends. The first is the AAUAAA polyadenylation signal (PAS), and the second is the G/U rich downstream element (DSE) [16]. To form the 3' end, the Cleavage and Polyadenylation Specificity Factors (CPSFs) are recruited through the binding of CPSF160 to the PAS [17]. CPSF160 is the only member of the core CPSFs that possesses an RNA binding domain. The Cleavage-Stimulation Factor (CstF) is also recruited to the 3' end through the binding of CstF64 to DSE [18]. Once these two complexes are in place, they recruit a cleavage factor, which is comprised of CPSF73 and CPSF100 along with the large scaffolding protein, Symplekin. CPSF73 contains both a β -CASP and a β -lactamase domain that, together, cleave

Figure 1: Human snRNA genes vs. mRNA genes. mRNA genes contain a promoter, an open reading frame (ORF) and a polyadenylation signal (PAS) marking their 3' end. In contrast, snRNA genes have a distal sequence element (DSE) and a proximal sequence element (PSE) not found in typical mRNA promoters. They also lack an ORF and have a 3' box marking their 3' ends in lieu of a PAS, as snRNAs are not polyadenylated.

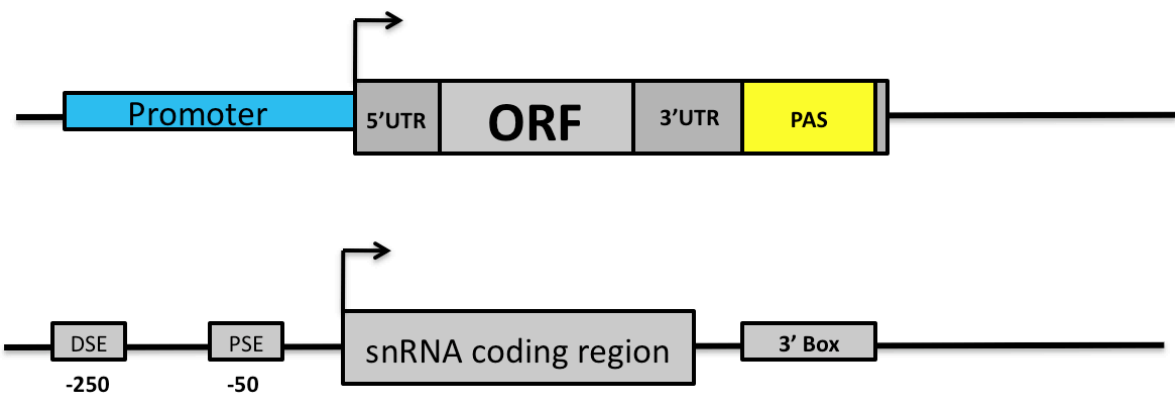
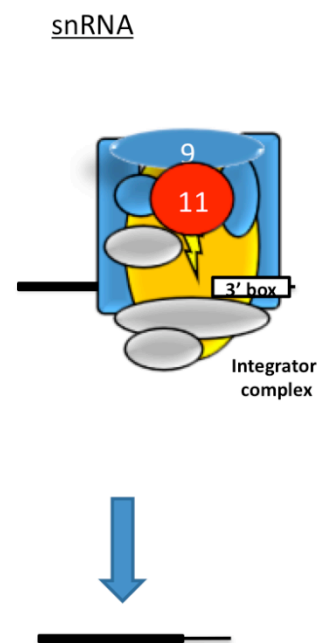
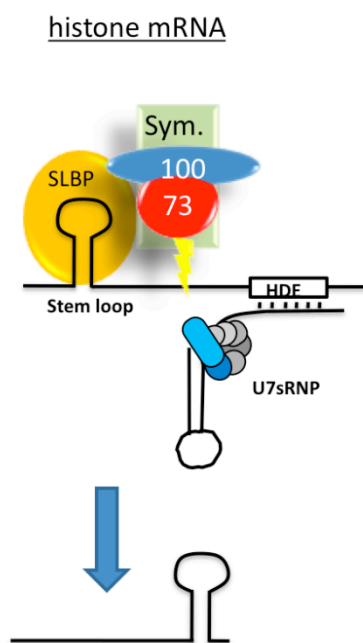
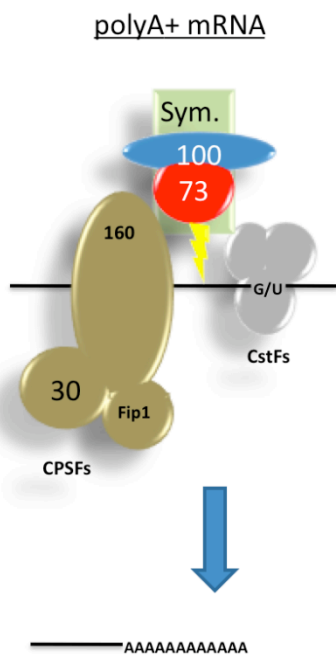


Figure 2: A comparison of the 3' end formation of three major RNAPII transcripts. Poly(A)+ mRNA and histone mRNA are both cleaved by CPSF73 and 100 to form their 3' ends. The difference between these two transcripts lies in the *cis*- and *trans*-factors that recruit the cleavage factor to the cleavage site. snRNAs, however, are cleaved by a novel complex of proteins called the Integrator Complex where the cleavage is carried out by IntS11 and 9, which show homology to CPSF73 and 100 respectively. No Integrator protein resembling Symplekin has yet been identified [2].

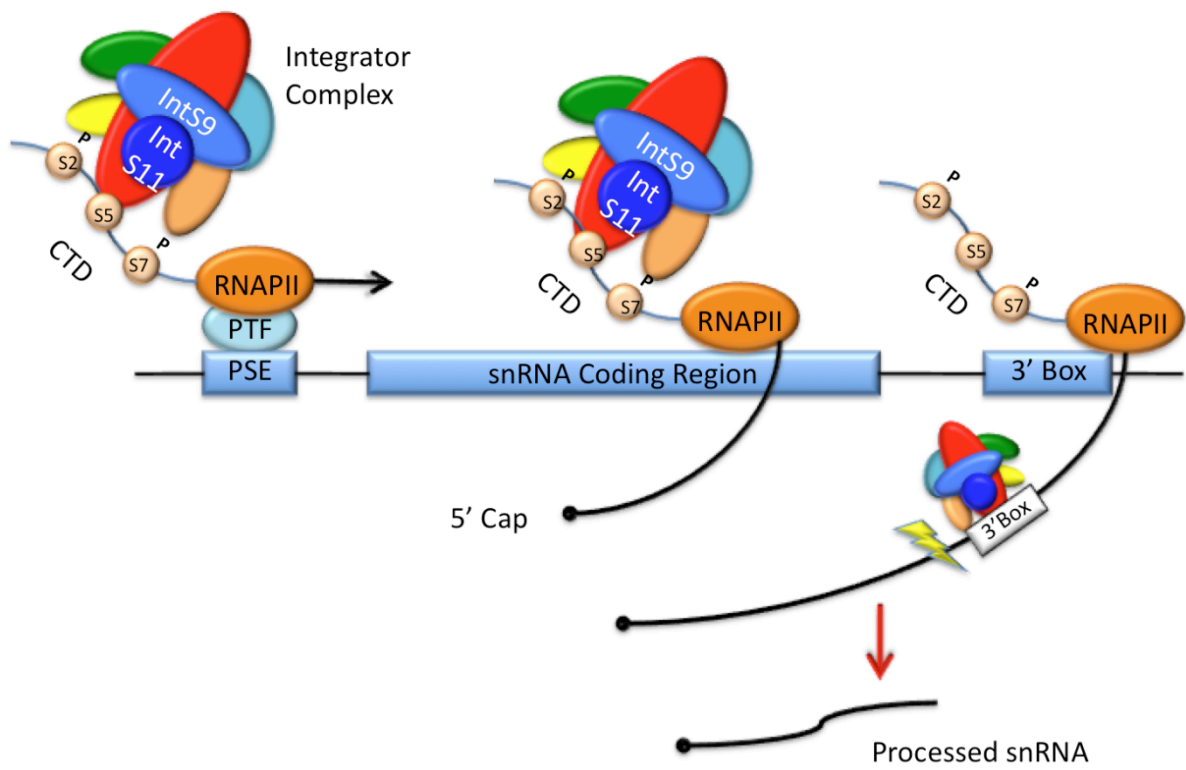


the mRNA at a CA cleavage site [16, 19, 20]. A polynucleotide tail of approximately 250 adenine residues is then added by poly(A) polymerase to produce the mature mRNA.

3' end formation of histone mRNAs is very similar to that of poly(A)⁺ mRNAs, the differences lying in the *cis*-elements that define the 3' end and the lack of a poly(A) tail [21]. Histone mRNAs have a stem loop upstream of the cleavage site that is bound by the stem-loop-binding protein (SLBP) and a histone downstream element (HDE) that base pairs with the U7 snRNA of the U7 snRNP [22, 23]. Once these factors have been recognized, the same cleavage factor that cleaves poly(A)⁺ mRNA, CPSF73/ CPSF100, and Symplekin, are recruited to cleave the histone mRNA [24-26]. Aside from the addition of the 7meGuanosine cap, this is the only mRNA processing reaction that histone mRNA's undergo as they are not polyadenylated.

Similarly to histone mRNA genes, snRNA genes also possess a stem loop upstream of the 3' cleavage site, however, in place of the HDE snRNA genes possess the poorly conserved 3' box [13]. The 3' box along with the snRNA promoter, and the C-terminal domain (CTD) of the largest subunit (Rpb1) of RNAPII are all required for proper 3' end formation [15, 27-30]. The transcription cycle of snRNA genes is highly related to standard mRNA genes with some key exceptions. When an snRNA gene is going to be transcribed (Figure 3), RNAPII is recruited to the promoter by the binding of an snRNA-specific complex of proteins termed the PSE-binding transcription factor (PTF) (also called snRNA activator protein complex (SNAPc) and PSE-binding protein (PBP)) [11]. Once recruited to the promoter, the

Figure 3: The Integrator complex mediates snRNA 3' end processing. The Integrator complex is recruited to an snRNA gene via phosphorylation of the CTD at serines 2 and 7 of RNAPII. It then rides along as RNAPII transcribes the snRNA and is transferred to the nascent snRNA at the 3' box. Upstream of the 3' box, the snRNA is cleaved by IntS11 and 9, which display homology to CPSF73 and 100 respectively [36].



CTD of RNAPII, which contains multiple repeats of the heptad YS²PTS⁵PS⁷, is phosphorylated at the serine 2 and serine 7 positions of the heptad through the activity of multiple kinases [31, 32]. The serine 2 kinase cdk9/cyclin T has been shown to be required for snRNA 3' end formation and appears likely to be one of the CTD kinases required. The identity of the serine 7 kinase is more controversial as both TFIIH and DNA-PK have been shown to possess this activity. Importantly, neither of these kinases has been shown to be functionally required for snRNA 3' end formation. This phosphorylation is thought to then recruit a recently purified complex of proteins called the Integrator complex to the CTD of RNAPII [33, 34]. The Integrator complex is speculated to associate with RNAPII as it transcribes, and is transferred to the nascent snRNA at the 3' box. Once at the 3' box, cleavage is carried out upstream by Integrator subunits (IntS) 11 and 9, which exhibit homology to CPSF73 and CPSF100 respectively [33, 35, 36]. Interestingly, it has not been shown that Symplekin or a Symplekin-like protein is a member of the Integrator complex or that any of the Integrator subunits display any homology to Symplekin, nor is it known how the Integrator complex mediates snRNA cleavage [2].

The Integrator Complex

In 2005, the Shiekhataar laboratory discovered the Integrator complex in a pull-down experiment to find proteins that interacted with the Deleted in Split hand/Split foot (DSS1) protein. They purified the complex and found that it was comprised of 12 subunits and associated with the CTD of the largest subunit (Rpb1) of RNAPII [33]. Subsequent purifications performed by the Shiekhataar laboratory,

in which they use flag-tagged IntS10 to pull down and purify the complex as well as a more recent purification by a separate group using specific antibodies against Integrator subunits failed to pull down DSS1, showing that its ability to pull down the Integrator complex in the first purification was serendipitous and it is not in fact a member of the Integrator complex [33, 37].

In addition to its interaction with the CTD, it was also shown that the Integrator complex mediates 3' end processing of snRNAs through the use of RNA interference (RNAi). When IntS11 or IntS1 was knocked down via RNAi, an accumulation of misprocessed snRNA, which migrates slower than normal snRNA, and is therefore longer, was observed via Northern blot. This is evidence that when the Integrator complex is compromised, primary snRNAs are not cleaved properly at the 3' end, longer misprocessed forms are allowed to accumulate, leading to the conclusion that the Integrator complex mediates snRNA 3' end processing [33].

The involvement of the Integrator complex was further demonstrated in a study by the Wagner laboratory in which a reporter system was developed in *Drosophila* S2 cells that placed the green fluorescent protein (GFP) gene and a PAS downstream of the 3' box of the U7 snRNA gene. When transfected into normal cells, any U7 snRNAs transcribed from the reporter plasmid will be cleaved upstream of the 3' box and no GFP will be expressed. However, when the reporter is transfected into cells in which the Integrator complex has been compromised through RNAi of Integrator subunits, the U7 snRNA will not be properly cleaved. As a result, RNAPII will read-through to the PAS and GFP will be expressed in these cells, indicating that snRNAs are being misprocessed. With this tool, it was shown

that depletion of each Integrator subunit causes misprocessing of snRNAs, albeit to differing degrees. It was observed that knockdown of IntS1, 4, and 9 caused strong expression of GFP and therefore high misprocessing of snRNA, knockdown of IntS5, 7 and 11 yielded an intermediate response, knockdown of IntS2, 6, 8, 10, and 12 a weak response, while the knockdown of IntS3 resulted in no detectable misprocessing. The group confirmed these results using a RT-PCR assay with forward primers designed to the U7snRNA and reverse primers designed downstream of the cleavage site [38]. In addition to the *Drosophila* assays, the Wagner laboratory has also developed similar human assays that are yielding similar results.

As mentioned previously, two Integrator subunits, IntS11 and 9 bear homology to CPSF73 and 100, the cleavage factor for poly(A)⁺ mRNAs and histone mRNAs [33, 35]. This is especially true in the β -CASP and β -lactamase domains present in both sets of proteins, where the amino acid sequences of core elements within IntS11/CPSF73 and IntS9/CPSF100 are almost exactly alike. Both IntS11 and CPSF73 contain a β -CASP (metallo- β -lactamase-associated CPSF Artemis SNM1/PSO2) β -lactamase domain, which is essentially a β -lactamase domain that has been modified to allow the endonucleolytic cleavage of nucleic acids. IntS9 and CPSF100 also contain a β -CASP β -lactamase domain, however, it has been modified so as to render it inactive [2, 35]. Baillat *et al.* determined that IntS11 was the catalytic subunit of the Integrator complex by overexpressing a mutant IntS11 that lacked catalytic activity in the β -CASP domain. However, it still interacted with other Integrator subunits and was shown to localize to snRNA genes at both the

promoter and the 3' end. When the mutant IntS11 was overexpressed following depletion of endogenous IntS11, sizable amounts of misprocessed snRNA were observed, leading to the conclusion that IntS11 is in fact the subunit responsible for snRNA cleavage [33].

When it was first purified, the Integrator complex was thought to consist of 12 subunits, however, a genome wide RNAi screen preformed in *Drosophila* S2 cells found two more subunits, bringing the total up to 14. The complex is evolutionarily conserved among metazoans; however, there are no Integrator orthologues in the yeast *Saccharomyces cerevisiae* as the Nrd1/Nab3/Sen1 complex mediates snRNA 3' end formation in this species [2, 39, 40]. A schematic of the 14 Integrator subunits mapping domains identified by Pfam analysis (in green) is shown in Figure 4. In addition to the β -CASP β -lactamase domains found in IntS9 and 11, the only other identifiable domains to be found in the complex are the armadillo repeats (ARM) found in IntS4 and 7, the HEAT repeats found in IntS4, the von Willebrand factor A domains found in IntS6 and 14, the RNA DEAD box helicase (DEAD) (likely inactive, as key residues are mutated) also in IntS6, the tetratricopeptide repeats (TPR) in Int8, and the plant homeodomain (PHD) finger in Int12. Armadillo, HEAT, and TPR repeats as well as von Willebrand factor A domains are all involved in protein-protein interactions as well as intracellular transport, and the PHD finger is a chromatin binding domain. The domain of unknown function (DUF) regions are areas that Pfam has identified as functional domains, however, the exact function of these regions is unknown. Surprisingly, none of the subunits display any known RNA binding domains [2].


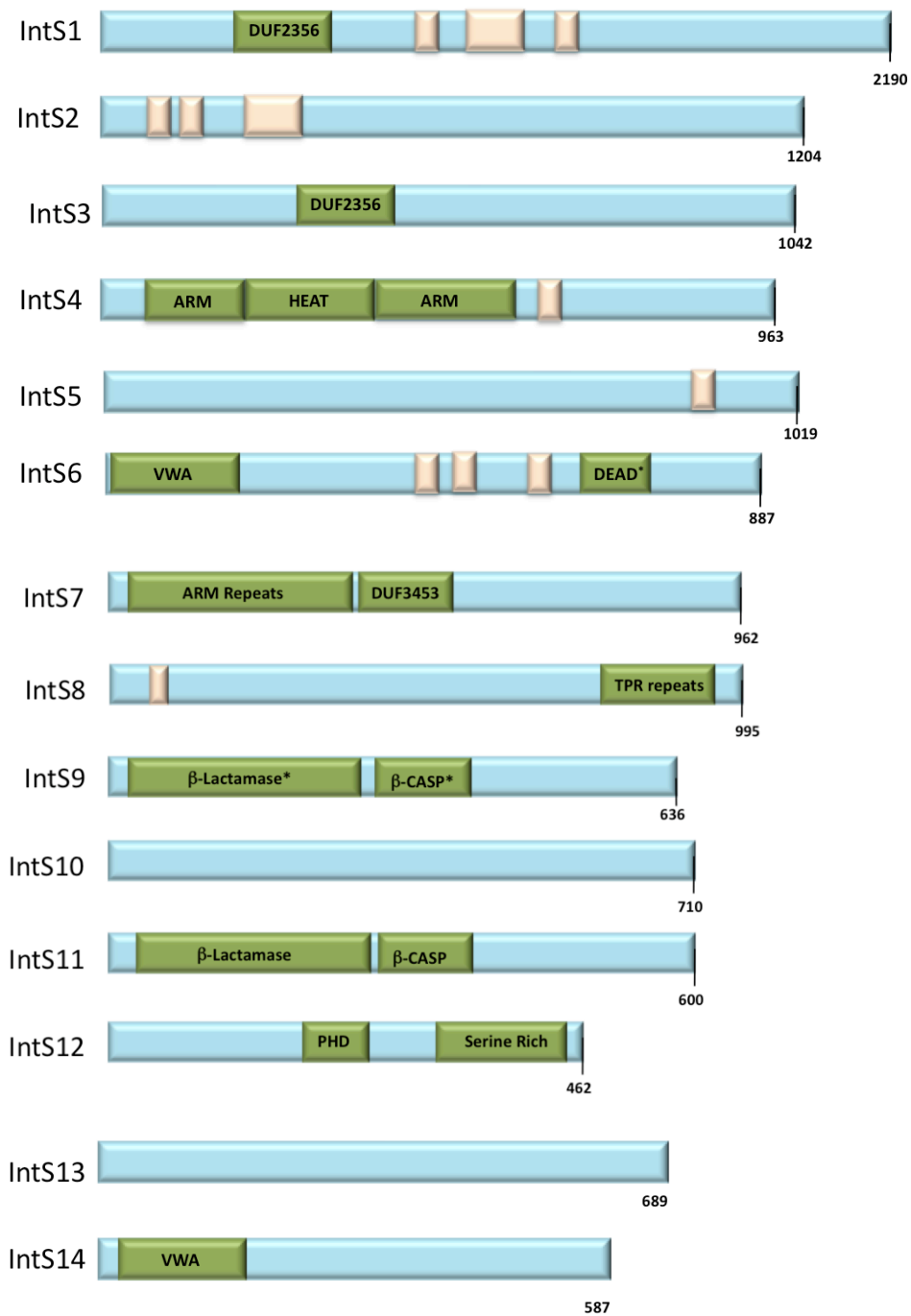


Figure 4: Schematic of Integrator complex subunits. The areas shaded in green are protein domains identified through Pfam analysis. Areas in orange are areas of high homology. Surprisingly, none of the domains identified are known RNA binding domains [2].

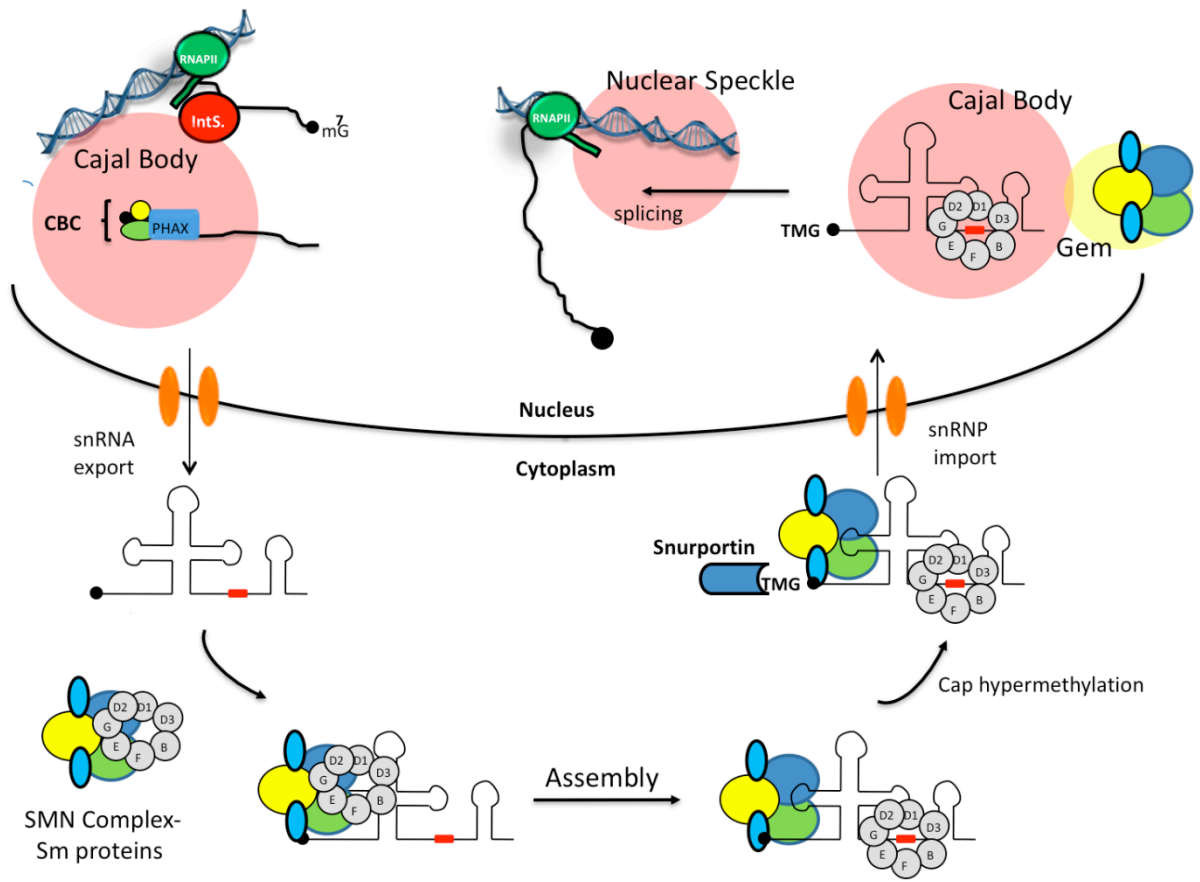


Since its purification, many members of the Integrator complex have been implicated in other studies, either for their role in snRNA expression, or for other functions, suggesting that Integrator proteins may play roles in other processes. For example, knockdown of IntS1 in mouse embryos results in lethality, and further examination of the embryos shows an increase in misprocessed snRNAs [41]. Another study has shown that when IntS5 expression is inhibited in zebrafish they fail to develop circulating blood cells. This is because the lack of functional snRNAs brought on by the disruption in the Integrator complex leads to splicing defects in the mRNAs of proteins necessary for hematopoiesis [42]. IntS3 has been shown to function as part of the sensor of single-stranded DNA (SOSS) complex responsible for the detection of double-stranded breaks in DNA, and IntS6 has been determined to be a previously described tumor suppressor gene discovered in non-small-cell-lung carcinoma, deleted in cancer 1 (DICE1), which, when overexpressed in prostate cancer cells leads to a reduction in colony formation and cell cycle arrest [43-46]. These studies suggest the possibility that some Integrator subunits may also play role in cellular processes other than snRNA expression.

snRNP Biogenesis

All U snRNAs are members of the U small nuclear ribonucleoproteins (snRNPs), the majority of which function as part of the major and minor spliceosomes that excise introns out of pre-mRNAs. The exception is the U7 snRNP, which plays a role in the 3' end formation of histone mRNAs. A brief description of snRNP biogenesis (Figure 5) begins in the nucleus with the

Figure 5: The biogenesis of U snRNPs. snRNP biogenesis begins in the nucleus with the transcription of snRNAs, usually associated with a Cajal body. The snRNA is then transported out of the nucleus to the cytoplasm where snRNP proteins are assembled onto the snRNA by the survival of motor neurons (SMN) complex. Once the proteins are properly assembled, the snRNP is then transported back to the nucleus where it first returns to the Cajal body for further processing and is then transported to the nuclear speckles to participate in splicing [47].



transcription and 3' end formation of the snRNAs. The snRNAs are then transported into the cytoplasm of the cell where a complement of proteins is assembled onto the snRNA to form a preliminary snRNP. The snRNP is then transported back to the nucleus for further processing before it assumes its function as part of the spliceosome [47]. This is the case for all of the U snRNPs with the exception of the U6 snRNP whose assembly takes place solely in the nucleus [48, 49]. A more detailed account of snRNP biogenesis is given below.

As stated previously, snRNA genes, usually found clustered adjacent to Cajal bodies [50] are transcribed by RNAPII to begin the snRNP biogenesis process. A complex of proteins called the cap-binding complex (CBC), which consists of the proteins CBP20 and CBP80, associates co-transcriptionally with the 5' 7-methylguanosine (7^mGpppN) cap of the snRNA [51]. Next, an adaptor protein termed phosphorylated adaptor for RNA export (PHAX) binds to the CBC. PHAX contains a leucine-rich nuclear export signal (NES) that is only active when PHAX is phosphorylated [52]. Once a phosphorylated PHAX is bound to the snRNA, Chromosome Region Maintenance 1 (CRM1), a nuclear export factor, binds to the exposed NES of PHAX along with the small G-protein, RanGTP, and shuttles the snRNA out of the nucleus [47, 53].

Once the complex enters the cytoplasm, PHAX is dephosphorylated by protein phosphatase 2A releasing CRM1 and RanGTP, however, it remains with the CBC/snRNA complex most likely until the monomethylated ($m_1\text{G}$) 7^mG cap is trimethylated ($m_3\text{G}$ cap) to prevent the snRNA from interacting with the translation

initiation machinery [47, 54]. The large SMN complex made up of the SMN protein, Gemin 2-8, and multiple other proteins then associates with the CBC/snRNA complex and monitors all maturation events that take place in the cytoplasmic phase of snRNP assembly [55, 56]. The first of these events is the assembly of the Sm ring around the Sm site of the snRNA. A set of seven Sm proteins, B/B' (B' is a splice variant of B), D1, D2, D3, E, F, and G form a ring around a conserved sequence called the "SM binding site" (PuAU₄₋₆GPu) upstream of the snRNA 3' stem loop [57]. By themselves, the proteins exist as dimers or trimers as they cannot form a ring in the absence of snRNA. To load the Sm proteins onto the snRNA, SMN first facilitates the formation of an open ring consisting of D1, D2, E, F, and G. The SMN complex then loads this complex on to the snRNA Sm site, and closes the ring with the addition of the B/B'-D3 dimer. After the Sm ring is in place, trimethyl guanosine synthase 1 (Tgs1), an SMN complex-associated methyltransferase, recognizes B/B' and the m₁G cap which prevents snRNAs without a complete Sm ring from being hypermethylated as the B/B'-D3 dimer is the last of the Sm proteins to be added to close up the ring. Tgs1 then transfers two methyl groups to position 2 of the m₁G cap to form the 2,2,7-trimethylguanosine (m₃G) cap. The 3' end of the snRNA then undergoes further nucleolytic trimming to form a mature length snRNA [47, 58-61]. Snurportin-1 (SPN1) along with the import receptor Importin-β (Imp β) then interacts with the m₃G cap and the Sm core, which acts as a nuclear localization signal (NLS) and imports the snRNP back into the nucleus [47].

Upon reentry to the nucleus, SPN1 and Imp β disassociate from the snRNP and return to the cytoplasm, however the SMN complex remains attached. The snRNP then returns to the Cajal body, where further modifications occur, such as pseudouridylation and 2'-O-methylation, carried out by small Cajal body RNAs (scaRNAs) [62, 63]. The SMN complex disassociates from the snRNP at an as yet unidentified stage in the nuclear maturation process and returns to the cytoplasm, and the mature snRNP is recruited to subnuclear domains called nuclear speckles where splicosomal snRNPs are stored, or, as is the case of the U7 snRNP, to histone locus bodies [3, 47].

Subcellular Localizations of RNPs

Ribonucleoproteins (RNPs) often concentrate, both in the nucleus and the cytoplasm, into subcellular foci that can be visualized under a light microscope. While the motives and mechanism behind the formation of these foci is largely unclear, it is possible that they function to concentrate factors necessary for certain RNA processing events such as, intron splicing, 3' end formation, transcription, and decay. Because the Integrator complex is responsible for the 3' end processing of snRNAs we hypothesized that some of the subunits would be present in some of these nuclear foci, and while this was shown to be the case, some of the subunits were unexpectedly shown in cytoplasmic foci as well. Possible locations of these Integrator foci include, Cajal bodies, nuclear speckles, and histone locus bodies in the nucleus, and P bodies, stress granules, and U bodies in the cytoplasm will be discussed below [64-66].

Cajal Bodies

First described in 1903 by Ramon y Cajal, Cajal bodies are dynamic, small nuclear bodies that are enriched with the protein coilin [65]. In addition to coilin, the U snRNPs are also enriched in Cajal bodies as well as the SMN complex, which point to a function in snRNP processing, and Cajal bodies contain small Cajal body-specific RNAs called scaRNAs which are responsible for carrying out 2'-O-methylation and pseudouridylation modifications on the snRNAs of snRNPs once they reenter the nucleus [62, 67]. Cajal bodies have also been shown to localize to snRNA gene clusters [50, 68, 69], and are sites of active snRNA transcription [70]. The Ohno laboratory has recently shown that Cajal bodies also play a role in snRNP assembly prior to its transport into the cytoplasm as well [71]. Their findings demonstrate that Cajal bodies monitor the export of snRNAs out of the nucleus after transcription. When PHAX, the NES-containing adaptor protein, is blocked from binding to nascent snRNAs, the snRNAs are retained in the Cajal body and prevented from exiting to the cytoplasm [71]. Despite its standard use as a marker of Cajal bodies, the cellular function of coilin is unknown. Cells lacking coilin, no longer form Cajal bodies, meaning coilin is necessary for Cajal body structure, however, snRNP modifications are carried out normally, demonstrating that Cajal bodies are not necessary for proper snRNP maturation [65]. In addition to snRNP processing, Cajal bodies are also storage sites for telomerase RNA during quiescent phases of the cell cycle. During S phase, when telomeres are being elongated, the telomerase RNA then moves to the telomerase holoenzyme at the telomeres [72-74].

Histone Locus Bodies

Histone locus bodies are so called, because they associate with the histone genes on chromosome 1 or 6 and are enriched for factors required for histone mRNA 3' end processing, such as, the U7 snRNP, SLBP, and Symplekin [75, 76]. The existence of these domains was revealed through experiments to visualize Cajal bodies in *Drosophila* cells. The sequence to *Drosophila* coilin was not available at the time, so the group used probes to the U7 snRNP and the U85 scaRNA. These probes revealed two separate types of subnuclear foci. The Cajal bodies were determined to be the ones that contained both the U7 snRNP and the U85 scaRNA, and the others, containing only the U7 snRNP were termed the histone locus bodies because of their association with histone genes [77]. As previously mentioned, histone locus bodies form around replication-dependent histone gene clusters. These genes are only active during S phase, and it is thought that these nuclear bodies serve to concentrate factors necessary for the maturation of histone mRNAs. The histone locus bodies are visible until early prometaphase and then disintegrate in metaphase. Once the new cells enter G1 phase, a few histone locus bodies reform, and their numbers increase upon the entrance of S phase [78]. Histone locus bodies are usually seen adjacent to Cajal bodies, however, the relationship between these subnuclear domains is unknown [65].

Nuclear Speckles

Nuclear speckles, also known as interchromatin granule clusters (IGCs), are sites enriched in mature U snRNPs and function in pre-mRNA splicing. In addition

to snRNPs, nuclear speckles also contain serine/arginine-rich (SR) proteins whose phosphorylation state regulates the interaction of the snRNPs with the spliceosome as well as sites of transcriptionally active genes. Each nucleus typically contains approximately 30-50 nuclear speckles, which can travel around the nucleus to associate with actively transcribing genes [64, 66].

P Bodies

The cytoplasmic foci that later became known as P bodies, were first observed by the Achsel laboratory in 2002. In this paper, it was shown that LSM1-7, an mRNA decapping activator complex, Dcp1/2, a decapping enzyme, and the exonuclease Xrn1 all co-localize to discrete cytoplasmic foci [79]. In 2003, Sheth and Parker showed that many mRNA decapping and decay factors were present in these cytoplasmic foci in yeast, including, Dcp1p, Dhh1p, and Xrn1p and hypothesized that these foci either might be storage sites for mRNA decay factors, or sites of active mRNA decapping and decay. They called these foci processing bodies, or P bodies. If P bodies were active sites of decapping and decay, then the group speculated that inhibiting mRNA decay before the decapping step should reduce the number of P bodies. This was because blocking decapping would reduce the number of mRNAs targeted for decay, and more mRNAs would remain in polysomes. Conversely, blocking mRNA decay at or after decapping would increase the number of P bodies because the mRNA could be targeted for decay, however, disruption of the decay machinery would cause targeted mRNAs to accumulate in more and more P bodies. Knockout experiments proved this to be

true. In a *ccr4Δ* strain, which is deficient in deadenylation (before decapping), P bodies are reduced, however, in *dcp1Δ* (at decapping) and *xrn1Δ* (after decapping) strains P bodies increase in both size and number [80].

In addition to decapping factors and exonucleases, proteins involved in other mRNA decay processes also accumulate in P bodies. Proteins involved in nonsense-mediated decay (NMD), which is responsible for the decay of mRNAs with premature stop codons (nonsense codons), as well as proteins involved in AU-rich element (ARE)-mediated decay have been found to accumulate in P bodies [81-85]. Additionally, the Argonaute and GW proteins involved in RNAi and microRNA (miRNA)-mediated gene silencing, and translational repressors such as eIF4E-transporter, the yeast Dhh1 and its vertebrate orthologue RCK/p54 accumulate in P bodies [86-90]. P bodies have additionally been shown to be sites of replication for RNA viruses as well as sites of host viral defense [91, 92].

P bodies require translationally repressed mRNPs in order to form, and RNAi and miRNA silencing pathways produce a significant portion of these mRNPs [93, 94]. Izaurralde and colleagues have shown that inhibition of miRNA silencing pathways and, to lesser extent, RNAi pathways also prevents the formation of visible P bodies. However, simply releasing mRNPs from polysomes with puromycin treatment, which mimics an aminoacylated tRNA and triggers release of the mRNA from the ribosome when the drug enters the A site, is not sufficient to restore P bodies in cells with inhibited silencing pathways. The mRNPs must enter silencing and decay pathways for P bodies to form. Even though P bodies require silencing pathways to form, silencing pathways do not require P bodies to function,

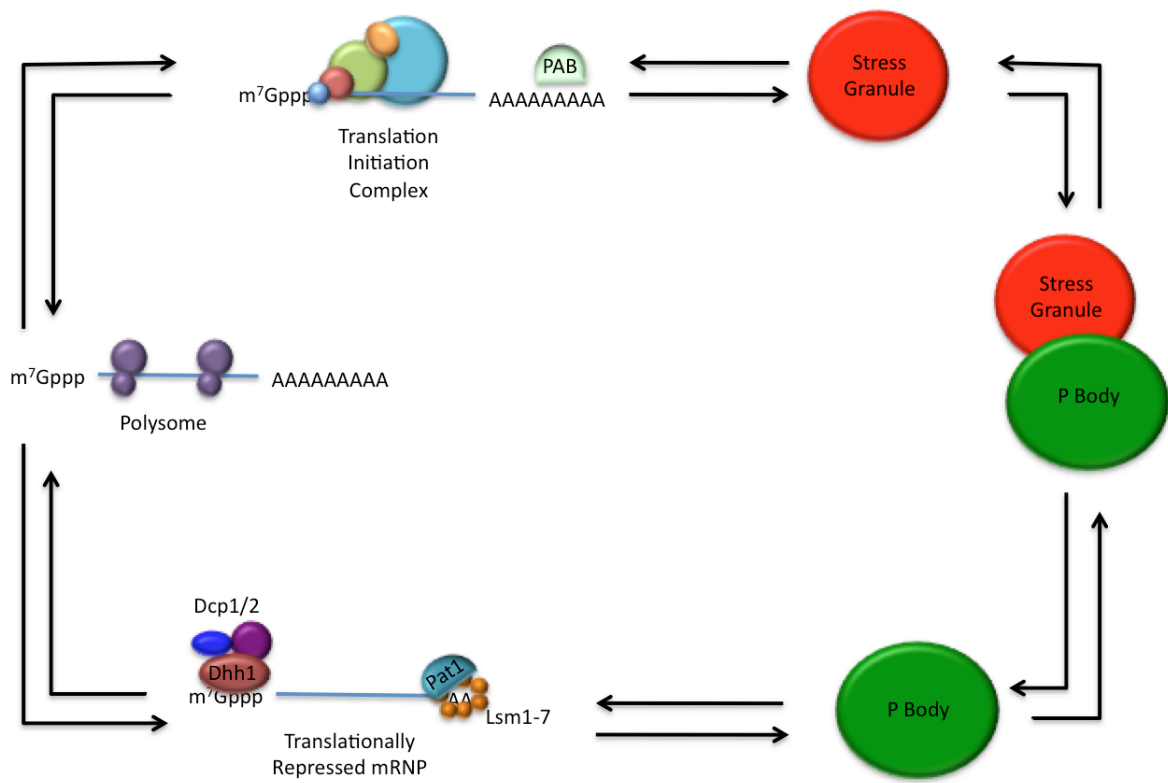
as depletion of non-silencing P body components disrupts the formation of P bodies, but not the function of miRNA silencing pathways. Therefore, P body formation is a consequence of silencing pathway function [94].

While the exact mechanism of P body assembly is unknown, many P body components contain Q/N rich prion-like domains, and it is hypothesized that protein-protein interactions between these domains play a role in P body aggregation along with other possible mechanisms [93, 95]. Additionally, other proteins with prion-like domains, such as the huntingtin (Htt) protein have also recently been shown to accumulate in P bodies [96].

Stress Granules

Non-translating mRNPs may also form into another type of cytoplasmic foci called stress granules. These granules are composed of mRNPs that are stalled in translation initiation as well as initiation factors eIF4E, eIF4G, eIF4A, eIF3 and eIF2, the small, 40S, ribosomal subunit, and the poly(A) binding protein (Pab1). Stress granules form in response to environmental stresses such as heat shock, hypoxia, glucose deprivation, and viral infection [97]. Similar to P bodies, the exact assembly mechanism of stress granule formation is unknown, though self-aggregation domains in some components are thought to play a role in this process [98, 99]. Stress granules are also often observed adjacent to P bodies, which lead Balagopal and Parker to propose a model for cytoplasmic mRNA metabolism they call the mRNA cycle (Figure 6) [93]. In this model, they suggest that mRNAs in the cytoplasm cycle between three different states: actively translating in polysomes, translationally repressed or decaying in P bodies, and stalled in initiation in stress

Figure 6: The mRNA cycle. Balagopal and Parker have proposed a model of cytoplasmic mRNA metabolism whereby mRNAs are shuttled between three sites: polysomes, P bodies, and stress granules [93].



granules. The recruitment of translational repressors, such as Dhh1 and Pat1 along with ribosomal run off transition the mRNA to a non-translating state. They can then aggregate into P bodies where they are stored to reenter a translating state in the future, decapped by Dcp1/2 and degraded, or transferred to stress granules. mRNAs stalled in initiation aggregate in stress granules. From here they can either reenter into polysomes once the stress has passed or be transferred to P bodies [93].

U Bodies

A third type of cytoplasmic foci is the recently described U body. These granules were first observed by Liu and Gall in *Drosophila* nurse cells using immunofluorescent staining and *in situ* hybridization. YFP-tagged Lsm11, an Sm core protein unique to the U7 snRNP, was expressed in transgenic flies, and in addition to the histone locus bodies found in the nucleus, the proteins were also observed in cytoplasmic foci. Immunofluorescent staining using antibodies to endogenous Lsm11 as well as Lsm10 confirmed this result. *In situ* hybridization was then performed using probes against U1, U2, U4, U5, and U6, snRNAs, and all were found to be present in these cytoplasmic foci as well. These results led to the term U bodies. U bodies were also shown to contain the SMN protein. In addition, flies that are homozygous for a mutant form of the arginine methyl transferase Dart5 lack detectable U bodies. Dart5 is responsible for the conversion of arginine residues of Sm proteins to symmetrical dimethylarginine (sDMA) leading to enhanced interaction between SMN and the Sm proteins. U bodies always occur in conjunction with P bodies although not all P bodies interact with U bodies, and

disruption of P bodies also leads to disruption of U bodies, suggesting a functional interaction between these two structures. These data led to the proposal that U bodies are either sites for certain steps in snRNP assembly or sites of cytoplasmic snRNP storage before import into the nucleus. P bodies might interact with U bodies to regulate the release of snRNPs from the U bodies to the nucleus, or they might be sites of decay for dysfunctional or unnecessary snRNPs [100].

Summary of Work

This study investigates the subcellular localizations of the Integrator subunits and seeks to determine if any exhibit unique localizations that might lead to further information about their function. **The hypothesis of this work is that the majority of subunits will display general nuclear localization, however, a few will have distinct localization to the Cajal bodies, as this is where snRNA are localized and transcribed.** To test this hypothesis, two specific aims were addressed. **The first specific aim was to determine the subcellular localization of the 14 Integrator subunits.** To do this, all Integrator subunits from both *Drosophila* and human were cloned either N-terminal or C-terminal to cDNA encoding either mCherry or GFP to generate fusion proteins. These clones were subsequently transfected into HeLa or S2 cells and their expression and localization was then visualized using confocal microscopy. The majority of the subunits were found to be diffusely expressed in the nucleus, however, IntS3 localized in discrete foci in the nucleus, and IntS2 and 7 surprisingly formed cytoplasmic foci. **The second specific aim was to further characterize those Integrator subunits that**

demonstrated a unique subcellular localization. A variety of experiments, including co-localization studies, deletions, and RNAi knockdowns were performed with IntS2, 3 and 7, and our data suggests that IntS3 localizes to the Cajal bodies, consistent with its function in snRNA 3' end processing, and while still present in the nucleus, both IntS2 and 7 also localize to cytoplasmic P bodies, suggesting they either play a role in snRNA decay or another cellular function unrelated to snRNA and snRNP biogenesis. These results have demonstrated the localizations of the Integrator subunits, revealing unique localizations for three members of this complex. Further investigation of these subunits in context of these localizations will lead to increased understanding of Integrator complex function in snRNA and snRNP biogenesis, and may lead to the discovery of Integrator involvement in other cellular processes.

Chapter 2: Materials and Methods

Cell Culture

The human cervical cancer derived cell line, HeLa, was obtained from the laboratory of Dr. Phillip Carpenter. These cells were maintained in DMEM Glutamax-1 high glucose media from Invitrogen, (Carlsbad, CA) supplemented with 10% FBS and 1% Penicillin-Streptomycin at 37°C with 5% CO₂. D.Mel-2 (S2) cells, were obtained from Invitrogen, and are derived from late-stage *Drosophila melanogaster* embryos and conditioned to grow in serum free environments. These cells were maintained in Sf-900 II SFM serum free media (Invitrogen) at 28°C. *E. coli* XL1 Blue competent cells were grown in Luria broth (LB) at 37°C, and ampicillin was added at 50 µg/mL as needed.

Transfections

Transient transfections of human cell lines were performed using the Lipofectamine 2000 transfection reagent from Invitrogen. Briefly, cells were plated in 6-well plates at 2.5×10^5 cells/well and allowed to attach overnight. To prepare the DNA for transfection, 100-500ng/well of plasmid DNA was added to 100µL/well of Opti-MEM reduced serum media (Invitrogen) in tube A. In tube B 2µL/well Lipofectamine 2000 was mixed with 100µL/well Opti-MEM. These tubes were incubated for 7 minutes after which the contents of the two tubes were mixed and allowed to incubate for another 25 minutes. 200µL of the DNA/Lipofectamine 2000 solution was then added to each well to be transfected. Transient transfections of the *Drosophila* S2 cell line were carried out using the Effectene transfection reagent from Qiagen (Hilden, Germany) according to the manufacturer's protocol.

Cloning Reactions

Insert DNA was generated through PCR by using the primer sets found in Tables 1-5. The amplified DNA was then purified using the GeneJET PCR purification system from Fermentas (now part of Thermo Scientific, Glen Burnie, MD) according to the manufacturer's protocol. Digests of the vector and insert DNA were carried out using Fermentas FastDigest restriction enzymes. To digest the vector: 1-3µg of the vector DNA was mixed with 10µL of the 10X FastDigest buffer and 1µL of each restriction enzyme. The volume of the reaction was then brought up to 100µL with dH₂O and incubated at 37°C for 10-15 minutes. To digest the insert: the PCR product eluted from the purification column (≈ 48µL) was mixed with 6µL 10X FastDigest buffer, 4µL of dH₂O, and 1µL of each restriction enzyme and incubated at 37°C for 10-15 minutes. Once the vector had been digested, 1µL of Fermentas FastAP alkaline phosphatase was added to the reaction mixture and incubated for an additional 10 minutes. To purify the digested vector and insert DNA, it was first run on a 1% agarose gel. The gel containing the DNA was then excised, and purified using the GeneJET PCR purification system. To ligate the vector and insert DNA, a ligation reaction was set up consisting of 4µL of vector DNA, 4µL of insert DNA, 1µL of 10X ligation buffer, and 1µL of T4 DNA ligase (purified by the Wagner Laboratory). A control, vector alone, reaction with dH₂O in place of the insert DNA, was also set up at this time, and these reactions were incubated at room temperature for 4-24 hours. 5µL of each ligation reaction was then transformed into XL1 Blue competent *E. coli* cells and plated on LB agar plates containing ampicillin. Approximately 16 hours later, 5mL cultures of picked colonies

were set up and allowed to grow overnight. The plasmid DNA was recovered from these cultures using the Fermentas GeneJET plasmid miniprep kit. The plasmids were then screened by restriction digest using the restriction enzymes used to clone the plasmid.

Generation of Tagged Integrator Clones

Human clones

To generate the human tagged Integrator clones, mCherry and GFP cDNAs were first cloned into the pcDNA4/TO/myc-His A vector from Invitrogen using the primers listed in Table 1. mCherry and GFP were inserted so as to create four new vectors to allow the fluorescent tags to be placed either N-terminal or C-terminal to the Integrator subunit subsequently cloned. Inserts for the Integrator subunits were then prepared using the primers listed in Table 2 and were cloned into the mCherry and GFP vectors using the restriction sites listed.

***Drosophila* clones**

Because the commercially available pIZ/V5-His vector (Invitrogen) suitable for use in insect cells yields weak protein expression, a new, more robust expression vector was created for this project. To do this, first, the strong *Drosophila* Ubiquitin 63E (Ubi63E) promoter [101] was cloned into the pUC 19 vector. Downstream of this promoter, the multiple cloning site (MCS) of the pIZ/V5-His vector was inserted. Finally, mCherry cDNA was cloned into the pIZ MCS to generate the new pUB Cherry vector (Figure 7). The primers used for the synthesis of this vector are listed in Table 3. The Integrator subunits were then cloned into

Table 1: Primers sets and restriction enzymes used to generate pcDNA4 N/C-terminal mCherry and GFP vectors.

Primer Sets	Primer Sequence (5'-3')	Restriction Site
N-GFP F N-GFP R	GGCCAAGCTTATGGTGAGCAAGGGCGAGGAG GGCCGGATCCCTTGTACAGCTCGTCCATGCC	HindIII BamHI
C-GFP F C-GFP R	GGCCTCTAGAGATGGTGAGCAAGGGCGAGGAG GGCCACTAGTTCACTTGTACAGCTCGTCCATGCC	XbaI SpeI
N-mCherry F N-mCherry R	GGCCAAGCTTATGGTGAGCAAGGGCGAGGAGGAT GGCCGGATCCCTTGTACAGCTCGTCCATGCC	HindIII BamHI
C-mCherry F C-mCherry R	GGCCTCTAGAGATGGTGAGCAAGGGCGAGGAGGAT GGCCACTAGTTTACTTGTACAGCTCGTCCATGCC	XbaI SpeI

Table 2: Primer sets and restriction enzymes used to generate mCherry and GFP tagged human Integrator subunit clones. Stop = stop codon included in this primer. Reverse primers that contain stop codons were used to generate N-terminal clones and primers lacking stop codons were used to generate C-Terminal clones.

Primer Sets	Primer Sequence (5'-3')	Restriction Site
IntS2 F IntS2 R (Stop) IntS2 R	GGCCGCGGCCGCATGACGCCCCGAGGGTACAGGC GGCCACTAGTCTTTAAATTCCACTAACACTCATGTT GGCCACTAGTCTAATTCCACTAACACTCATGTT	NotI SpeI SpeI
IntS3 F IntS3 R (Stop) IntS3 R	GGCCGAATTCATGGAGTTGCAGAAGGGAAAAG GGCCCTCGAGCGTTAGTCACTGTCAGAGCCCACTGC GGCCCTCGAGCGGTCACTGTCAGAGCCCACTGC	EcoRI XhoI XhoI
IntS4 F IntS4 R (Stop) IntS4 R	GGCCGCGGCCGCATGGCGGCGCACCTTAAGAAG GGCCTCTAGACTTTAGCGCCGTGCAGGTTTGGGCA GGCCTCTAGACTGCGCCGTGCAGGTTTGGGCA	NotI XbaI XbaI
IntS5 F IntS5 R (Stop) IntS5 R	GGCCGAATTCTATGTCCGCGCTGTGCGACCCT GGCCCTCGAGCGCTACGTCCCCTGTCTGAAGGAGAGT GGCCCTCGAGCGCGTCCCCTGTCTGAAGGAGAGT	EcoRI XhoI XhoI
IntS6 F IntS6 R (Stop) IntS6 R	GGCCGAATTCTATGAACCAGCGCAGCCATCTG GGCCTCTAGACTTTAATTGCTATTAATATGGGTGAT GGCCTCTAGACTATTGCTATTAATATGGTTGAT	EcoRI XbaI XbaI
IntS7 F IntS7 R (Stop) IntS7 R	GGCCGAATTCTATGGCGTCAAACCTCAACTAAG GGCCTCTAGACTTTAAAACCGTGTGTAGGCATT GGCCTCTAGACTAAACCGTGTGTAGGCATT	EcoRI XbaI XbaI
IntS9 F IntS9 R (Stop) IntS9 R	GGCCGCGGCCGCATGAAACTGTATTGCCTGTCA GGCCTCTAGACTTTAGAACTTTGTAAGAATTT GGCCTCTAGACTGAACCTTTGTAAGAATTT	NotI XbaI XbaI
IntS10 F IntS10 R (Stop) IntS10 R	GGCCGCGGCCGCATGTCTGCCCAGGGGGACTGC GGCCTCTAGACTTTAGGTCAGAGTCTGAAGGAG GGCCTCTAGACTGGTCAGAGTCTGAAGGAG	NotI XbaI XbaI
IntS11 F IntS11 R (Stop) IntS11 R	GGCCGGATCCATGCCTGAGATCAGAGTCACG GGCCTCTAGACTTTAGCTGGGGGCCTGGGGGAG GGCCTCTAGACTGCTGGGGGCCTGGGGGAG	BamHI XbaI XbaI
IntS12 F IntS12 R (Stop) IntS12 R	GGCCGGATCCATGGCTGCTACTGTGAACTTG GGCCTCTAGACTTTACTTCTTGAGTTTCTTTTGGGC GGCCTCTAGACTCTTCTTGAGTTTCTTTTGGGC	BamHI XbaI XbaI

Figure 7: Diagram of pUB Cherry vector. The ubiquitin 63E promoter (Ubi63E) was first cloned into a HindIII restriction site in the pUC19 vector. The multiple cloning site (MCS) from the pIZ/V5-His vector was then cloned in using HindIII and NdeI restriction sites. Finally mCherry was cloned into the piZ MCS using HindIII and BamHI restriction sites to give the pUB Cherry vector. Amp – ampicillin, PAS – polyadenylation signal. Note: the HindIII and XhoI sites that are found in the native Ubi63E promoter were destroyed to facilitate cloning.

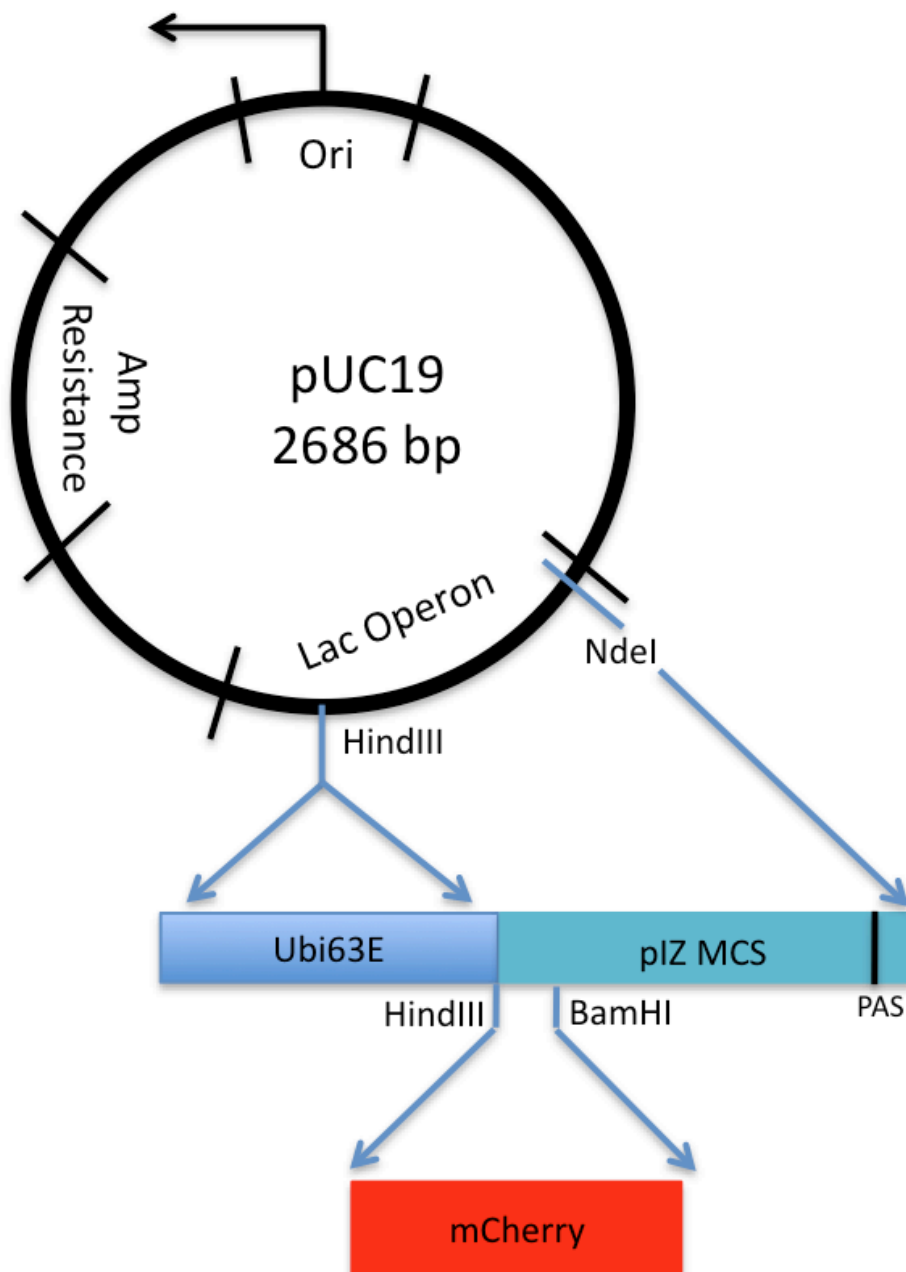


Table 3: Primer sets and restriction enzymes used to generate the *Drosophila* pUB Cherry vector. Ubi63E = ubiquitin 63 E promoter region, PizMCS = pIZ vector multiple cloning site.

Primer Sets	Primer Sequence (5'-3')	Restriction Site
Ubi63E F Ubi63E R	GGCCAAGCTTTGTGCGCCGGAAACGCAGCGACAGG GGCCAAGCTTCCTCTAGCTAGAGTCGACCTGC	HindIII HindIII
PizMCS F PizMCS R	GGCCTTTAAAGCTTATGGAGC GGCCCATATGGCTCACATGTTCTTTCC	HindIII NdeI
mCherry F mCherry R	GGCCAAGCTTATGGTGAGCAAGGGCGAGGAGATATGGTGAGCAAGGGCGAGGAG GGCCGGATCCCTCTTGTACAGCTCGTCCATGCCCTTGTACAGCTCGTCCATGCC	HindIII BamHI

this vector using the primers listed in Table 4. A schematic of the human and *Drosophila* clones is shown in Figure 8.

Generation of GFP hIntS7 Deletion Mutants

Nine N-terminal and nine C-terminal deletion mutants of hIntS7 were generated for this project. Primers were made at approximately 100 amino acid intervals from the N and C-terminal ends of hIntS7 to generate a series of fragments, NΔ1-9 and CΔ1-9 (Table 5). The reverse primer used to clone full-length hIntS7 in the fluorescent-tagged clones above was used to make the N-terminal deletion fragments, and the forward primer was used to make the C-terminal deletion fragments (Table 4). These fragments were then cloned into the N-terminal GFP vector using EcoRI and XbaI restriction sites (Figure 9).

Cell Fixation and Immunofluorescence

Human cells

HeLa cells were plated at 2.5×10^5 cells/well in 6-well plates containing cover slips (Fisher Scientific, Pittsburgh, PA) coated with a 50μg/mL poly-D-lysine solution (Sigma-Aldrich, St. Louis, MO). To coat the cover slips, the poly-D-lysine solution is pipetted onto the cover slips and allowed to sit for one hour. The solution is then aspirated and the cover slips are dried completely. For detection of tagged Integrator proteins, the cells were transfected with 100-500ng of plasmid DNA and incubated for 24 hours before fixation. For cells transfected with fluorescently

tagged proteins, the cells were washed with phosphate buffered saline (PBS) then fixed in 4% paraformaldehyde (PFM; Electron Microscopy Sciences, Hatfield, PA) for 10 minutes. The cells were then washed with PBS and the nucleus was stained with a 1:10,000 solution of DAPI in PBS for 10 minutes at 37°C. After DAPI staining, the cells were washed twice with PBS and then mounted onto glass microscope slides using an anti-fade mounting medium (Electron Microscopy Sciences).

In cells that were probed for endogenous proteins, the cells were first washed and fixed in the same manner as above. After fixing, the cells were washed once with PBS, and permeabilized with 0.5% triton X-100 in PBS for 5 minutes. The cells were then washed three times with copious amounts of PBS. After this, the cells were blocked with a 10% normal goat serum (NGS; Sigma-Aldrich) solution in PBS for 30 minutes at 37°C. Primary antibodies, mouse α -coilin (obtained from the Matera Laboratory) and rabbit α -hIntS3 and α -hIntS7 (gift from Proteintech Group, Chicago, IL) were then added at a 1:1000 dilution and the cells incubated for another 30 minutes. The cells were washed with PBS twice for five minutes at room temperature while rotating and then the fluorescently conjugated secondary antibody (α -mouse and α -rabbit AlexaFluor 555 and α -rabbit AlexaFluor 647 from Invitrogen) was added at 1:1000 for 30 minutes at 37°C. The cells were then washed with PBS twice for five minutes while rotating. DAPI staining and mounting were carried out as stated above.

Table 4: Primer sets and restriction enzymes used to generate mCherry tagged Drosophila Integrator subunit clones. All Drosophila clones have the mCherry tag placed on the N-terminus of the Integrator subunit.

Primer Sets	Primer Sequence (5'-3')	Restriction Sites
dIntS2 F dIntS2 R	GGCCGGATCCAATGCCGGTGAGGATGTACGATGTATCG GGCCTCTAGACTAATACAGGTCCGATTTTTTCATGACCGCC	BamHI XbaI
dIntS3 F dIntS3 R	GGCCGGATCCAATGGAACAGCAGCAATCAAAAAATAATGCT GGCCGCGGCCGCTCAGTCAGAATCATTGTTAGCTTTTTTCC	BamHI NotI
dIntS5 F dIntS5 R	GGCCGGATCCAATGCTGCGCCAGAACCTGTTGGATCAGCTTAAG GGCCTCTAGATTAATCTATTTCAACGATCTGCAGCCGGGCC	BamHI XbaI
dIntS6 F dIntS6 R	GGCCACTAGTCATGACAATCATACTCTTCCTGGTGG GGCCCTCGAGTTAACTCTTGCGACGGCCTGCTCCG	SpeI XhoI
dIntS7 F dIntS7 R	GGCCACTAGTCATGTCTCACCTGACCGGCACCCGCGTG GGCCCTCGAGTTAAACCTCCTCGTCTGTCCCACTG	SpeI XhoI
dIntS9 F dIntS9 R	GGCCGGATCCAATGCGATTGTATTGTCTCAGCGGGGACC GGCCTCTAGATTA AAAACTCTGTAAGCATTTTCATGATGGTGTCTC	BamHI XbaI
dIntS10 F dIntS10 R	GGCCGGATCCAATGCCGAGCCAAGAGGAAAATGAGTTGTACATG GGCCTCTAGATCACTTAATCACAATCGTCTCCACGGGCTGAC	BamHI XbaI
dIntS12 F dIntS12 R	GGCCACTAGTCATGGCCGCAAATATAGCCGCC GGCCTCTAGATTACTGCTTGGATCTGCGCTT	SpeI XbaI
dIntS13 F dIntS13 R (ASU)	GGCCGAATTCATGTTTGAACGCAACCAGAAG GGCCCTCGAGTTAACTACGTACGGATTC	EcoRI XhoI
dIntS14 F dIntS14 R (CG4785)	GGCCGAATTCATGCCACCTTAATAGCGCTG GGCCCTCGAGTCAATACATGTATGCAGGAGC	EcoRI XhoI

Figure 8: Schematic of *Drosophila* and human tagged Integrator Clones. The *Drosophila* clones contain mCherry N-terminal to the Integrator subunits, while human vectors were created so that mCherry and GFP could be placed both N-terminal and C-terminal to the Integrator subunits.

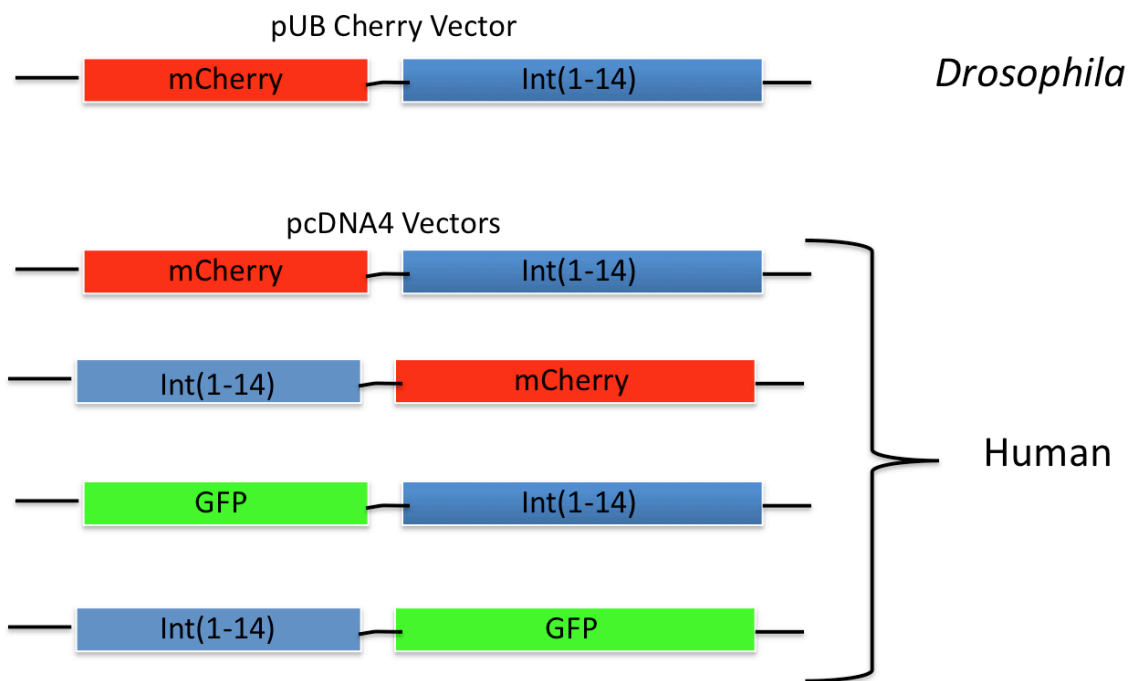
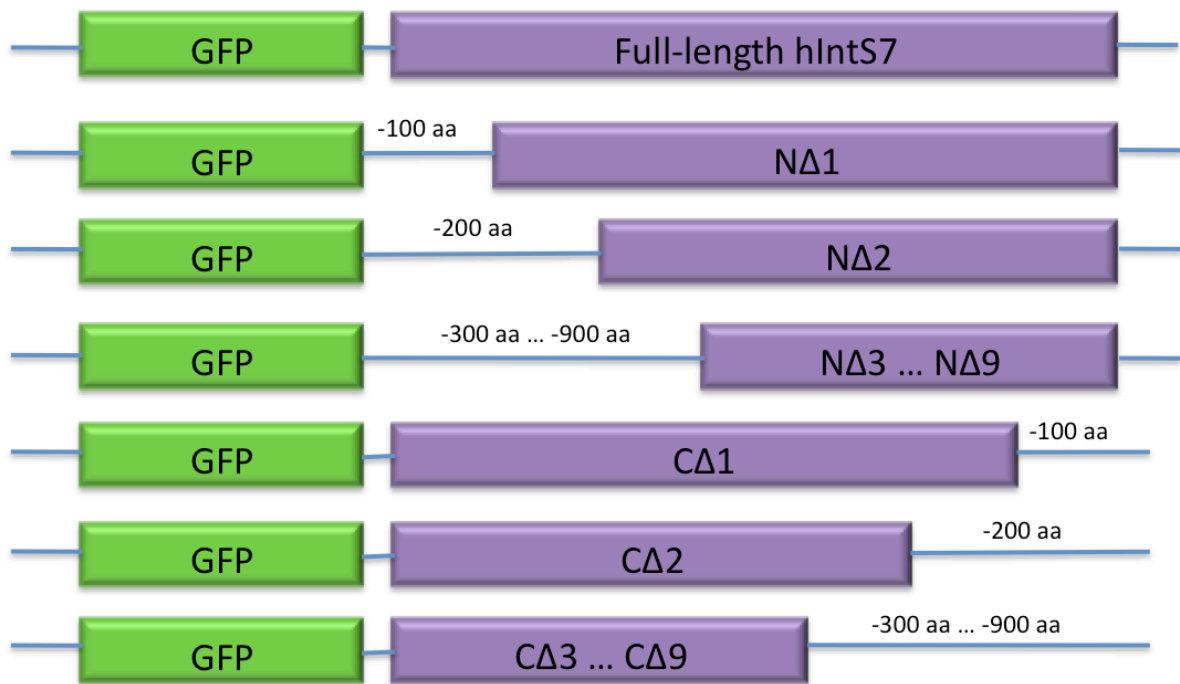


Table 5: Primers and restriction enzymes used to generate human IntS7 deletion fragments. The N Δ primers are forward primers with an EcoRI restriction site, and the hIntS7 full-length reverse primer with a stop codon in Table 2 was used with these primers. The C Δ primers are reverse primers with an XbaI restriction site, and the hIntS7 full-length forward primer in Table 2 was used with these primers.

Deletion	Primer Sequence (5'-3')
NΔ1	GGCCGAATTCTGTGGATGAATTTGTGAAGAGA
NΔ2	GGCCGAATTCTGATGCAATCTTGGCTTCCAGT
NΔ3	GGCCGAATTCTTTGTCTGTCCTTTCCACACTA
NΔ4	GGCCGAATTCTTTAAAGATTGCTCTAAACTGT
NΔ5	GGCCGAATTCTGTGATTGGACGATCAGCCACA
NΔ6	GGCCGAATTCTTCAGCACTTTCTTGCATTGCT
NΔ7	GGCCGAATTCTAAACAGTCCATGGAAGAATTT
NΔ8	GGCCGAATTCTTATATGCACACAGCATGCCTC
NΔ9	GGCCGAATTCTCAAAGGGTTGAACCTCATAAT
CΔ1	GGCCTCTAGATTACTCCATCTCATTGGTCATGTT
CΔ2	GGCCTCTAGATTAAGAAACAGGGGTATATTTCCG
CΔ3	GGCCTCTAGATTACATCTGATTGGAGATGCGACC
CΔ4	GGCCTCTAGATTAACTATAATTTTCCTCTTGCAA
CΔ5	GGCCTCTAGATTACTTGTACAGCTCCATGAGGTC
CΔ6	GGCCTCTAGATTAAGTGGCCTGAGCACCTGGACT
CΔ7	GGCCTCTAGATTACATCCCTAGTTTTAAGCTGTC
CΔ8	GGCCTCTAGATTAATGGTGCATGTGCTGTAGAAT
CΔ9	GGCCTCTAGATTAATTTAGAATCTTCTCCAAATG

Figure 9: Schematic of hIntS7 deletion mutants. Removing amino acids from each end in approximately 100 amino acid intervals created nine N-terminal and nine C-terminal deletion mutants. All deletion mutants were cloned into the pcDNA4 N-terminal GFP vector.



***Drosophila* Cells**

Drosophila S2 cells were plated at 10^6 cells/well in 24-well plates, transfected with 800ng of plasmid DNA and incubated for 24-48 hours. Cover slips (Fisher Scientific) were placed in a new 24-well plate and coated with a 50 μ g/mL concanavalin A (Sigma-Aldrich) solution. The transfected S2 cells were diluted 1:6-1:10, replated into new wells, and allowed to adhere to the concanavalin A cover slips for two hours. Concanavalin A causes the S2 cells to spread on the cover slip where they would usually loosely attach. The media was then removed and the cells were washed with PBS. The cells were fixed with 10% PFM for 10 minutes at room temperature and washed again with PBS. A 1:5000 DAPI in PBS solution was added and the cells were incubated at room temperature for 10 minutes to stain the nucleus. In my hands, increasing the concentration of DAPI and the time of incubation resulted in optimal DNA signal. Finally the cells were washed twice with PBS and mounted in the same way as the human cells.

All slides were visualized using the Nikon A1R Confocal Laser Microscope System in the Cytodynamic Imaging Facility in the Department of Integrative Biology and Pharmacology at the UT Medical School. All images were taken using a 60X plan-Apo/1.4 NA Oil objective. Images were processed using Adobe Photoshop.

RNAi

Human siRNA sequences against hIntS7, 5' GGCUAAAUAGUUUGAAGGA 3' and 5' CUCUAAACUGUAUGGUGAA 3' were purchased from Sigma-Aldrich. To perform siRNA knockdown experiments, HeLa cells were plated at 1.2×10^5

cells/well in a 24-well plate. 24 hours later, the siRNA was transfected using Lipofectamine 2000. Briefly, 3 μ L/well of siRNA and 47 μ L/well of Opti-MEM medium were mixed in tube A, and 3 μ L/well Lipofectamine 2000 and 12 μ L/well Opti-MEM were mixed in tube B and incubated for seven minutes at room temperature. The tube were then combined and incubated for 25 minutes at room temperature. After this incubation, 38 μ L/well of Opti-MEM was added to the transfection solution, and 100 μ L was added to each well.

To generate double stranded (dsRNA) for RNAi in *Drosophila* S2 cells, first, forward and reverse T7 primers with the sequence 5'-GGTAATACGACTCACTATAG-3' plus 18-30 nucleotides of gene specific sequence were designed for dIntS1-14 and Thread (positive control) at the positions shown in Figure 10. The T7 templates were amplified using PCR, and the resulting DNA was subjected to phenol-chloroform purification and ethanol precipitation for 15 minutes at -20°C and then resuspended in dH₂O at approximately 1 μ g/ μ L. These templates were then used to generate dsRNA via *in vitro* transcription. To set up the 300 μ L *in vitro* transcription reaction, 9 μ L of 62.5 mM NTPs, 30 μ L 10X transcription buffer (400mM Tris (pH8), 150mM MgCl₂, 50mM DTT, 0.5 mg/mL BSA), 10 μ L DNA template, 1-10 μ L T7 RNA polymerase and 250 μ L dH₂O were mixed together. This was then incubated at 37°C for 24 hours. 1 μ L of DNaseI was then added and the reaction incubated for 30 minutes at 37°C. After this, the reaction tubes were placed in a beaker of water and brought to a boil. The reaction was boiled for 2 minutes and then cooled to room temperature in the water. Once cooled, 1 μ L of the dsRNA was run on a 1% agarose gel with a dsRNA ladder to quantify; the top band of 2 μ L of the dsRNA

Figure 10: Positions of Drosophila dsRNA sequences. Forward and reverse T7 primers were designed to the positions shown to generate dsRNA for use in S2 knockdown experiments. Each dsRNA sequence is approximately 500 base pairs.

dIntS1



dIntS2



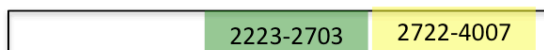
dIntS3



dIntS4



dIntS5



dIntS6



dIntS7



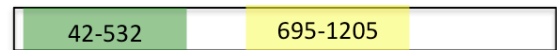
dIntS8



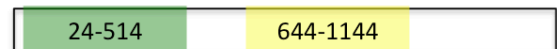
dIntS9



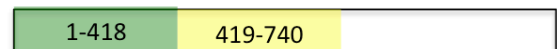
dIntS10



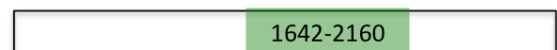
dIntS11



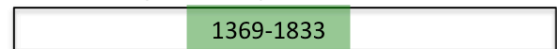
dIntS12



dIntS13 (ASU)



dIntS14 (CG4785)



Thread



ladder (New England Biolabs, Ipswich, MA) equals 140ng dsRNA. To perform RNAi knockdown in S2 cells; cells are resuspended at 10^6 cells/mL and 70 μ L of this suspension is plated in 96 well plates. The cells are then treated with 10 μ g/mL dsRNA for three days and then collected.

Western Blot

Cells were collected and lysed for Western blot by incubating them 50-250 μ L of low-salt lysis buffer for one hour. Relative protein levels were determined using the Bradford assay, and all samples were normalized to the lysate containing the least protein. To prepare the lysates for gel electrophoresis, up to 20 μ L of each lysate was mixed with an equal amount of 2X SDS loading buffer and boiled for five minutes at 95°C. The boiled samples and protein ladder (Fermentas) were loaded into a 12.5% SDS polyacrylamide gel. The samples were run at 80V through the stacking gel, and then the voltage was increased to 150V and the samples were run until the dye front ran off the bottom of the gel. The proteins were transferred to a PVDF membrane overnight at 30V. The membrane was blocked for one hour in 5% milk then probed with the primary antibody in 5% milk for one hour. The primary antibody used in this project was the mouse α -GFP antibody, JL-8 (Clontech, Mountain View, CA) used at 1:5000. The membrane was then washed three times for 10 minutes with 5% milk. Next, the membrane was probed with α -mouse horseradish peroxidase conjugated secondary antibody (Invitrogen) at 1:5000 in 5% milk for one hour, and this was followed with another set of washes. After the final

wash, the milk is rinsed off of the membrane with PBS, and the membrane is incubated for five minutes in ECL solution (Thermo Scientific, components mixed 1:1) then placed in an autoradiography cassette. Film is exposed to the membrane for <30 seconds to 10 minutes then developed.

Chapter 3: Results

The Localization of Integrator Subunits.

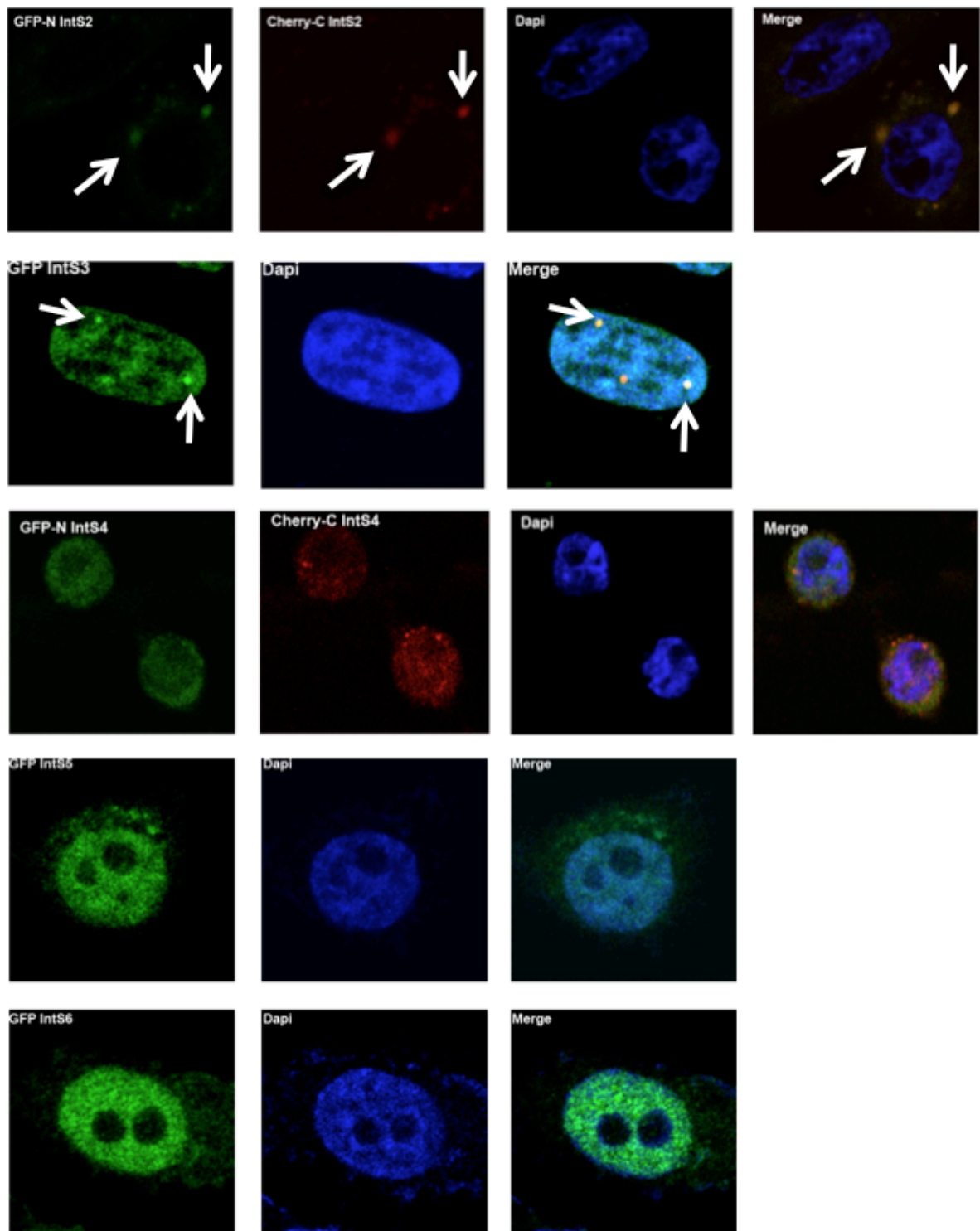
With the exception of IntS9 and 11, which serve as the cleavage factor for snRNA processing, little is known about the functions of the other Integrator subunits. Determining the subnuclear localizations of these subunits could provide valuable clues as to the function of the Integrator subunits, therefore, mCherry and GFP tagged clones were generated for all the human and *Drosophila* Integrator subunits with the exception of hIntS1 and 8 as well as the human orthologues of *Drosophila* Asunder (Asu, IntS13) and CG4785 (IntS14) as cDNAs were not available at the time of this project. In addition, dIntS1, 4, 6, 8, 10, and 11 which could not be cloned in the timeframe of this project.

To generate the expression vectors for the human Integrator subunits, mCherry and GFP cDNAs were cloned into the pcDNA4/TO/Myc-His A vector to create the four expression vectors: N-terminal GFP, C-terminal GFP, N-terminal mCherry, and C-terminal mCherry. The rationale for creating these four constructs were to test and compare localization of Integrator proteins by tagging both at the N-terminus and C-terminus to address any tag positional effects. Also, we used two fluorescence tags to address any potential artifacts associated with these fusion proteins. The Integrator subunits were then cloned into these vectors to generate Integrators that were tagged both N-terminally and C-terminally with GFP and mCherry. 500ng/well of these plasmids were transfected into HeLa cells plated on poly-D-lysine coated cover slips, which were subsequently fixed and imaged using confocal microscopy. The majority of the human Integrators, hIntS4, 5, 6, 9, 10, 11, and 12 display no remarkable subcellular localization. They are diffusely nuclear,

which we predicted given their role in snRNA 3' end processing, a nuclear event. However, while still nuclear, hIntS3 was observed in discrete foci, and surprisingly, hIntS2 and 7 were observed to form discrete cytoplasmic foci with little expression seen in the nucleus (Figure 11).

Tagged clones of *Drosophila* Integrator proteins were also created for transfection in S2 cells. cDNAs for the *Drosophila* Integrator subunits were cloned into the pUB Cherry expression vector constructed for this purpose by cloning the ubiquitin 63E promoter, pIZT/V5-His vector MCS and mCherry cDNA into the pUC19 vector. This resulted in plasmids with mCherry N-terminal to the Integrator subunits. 800ng/well of these plasmids were transfected into S2 cells, and these cells were transferred 24-48 hours later onto concanavalin A coated cover slips, fixed and imaged using confocal microscopy. Localizations for the *Drosophila* Integrator subunits are shown in Figure 12. dIntS5, 6, 9, and 12 as well as Asu (dIntS13) and CG4785 (dIntS14) demonstrate diffuse nuclear localization as was expected for most subunits, and dIntS2 and 7 localized to discrete cytoplasmic foci, similar to hIntS2 and 7 in HeLa cells. The nuclear foci seen with hIntS3 in HeLa cells were not observed with dIntS3 in S2 cells; rather, dIntS3 seems to be diffusely spread throughout the cell, suggesting *Drosophila* IntS3 does not share a similar localization with human IntS3. However, S2 cells did not tolerate overexpression of proteins well, and it is possible that the tagged form of dIntS3 does not localize to nuclear foci, but the endogenous protein does.

Figure 11: Confocal images of human Integrator subunit localizations. Most of the human Integrator subunits exhibit a diffuse nuclear localization. The exceptions are hIntS2, 3, and 7. hIntS3 forms distinct foci in the nucleus, while hIntS2 and 7 form distinct cytoplasmic foci (arrows). Some of the images show N-terminal GFP tagged Integrators and C-terminal mCherry tagged Integrators to demonstrate that the localizations of these tagged proteins overlap and the tags or their location on the protein are not the cause of the shown localizations. Cells transfected with a single tagged construct yielded similar results.



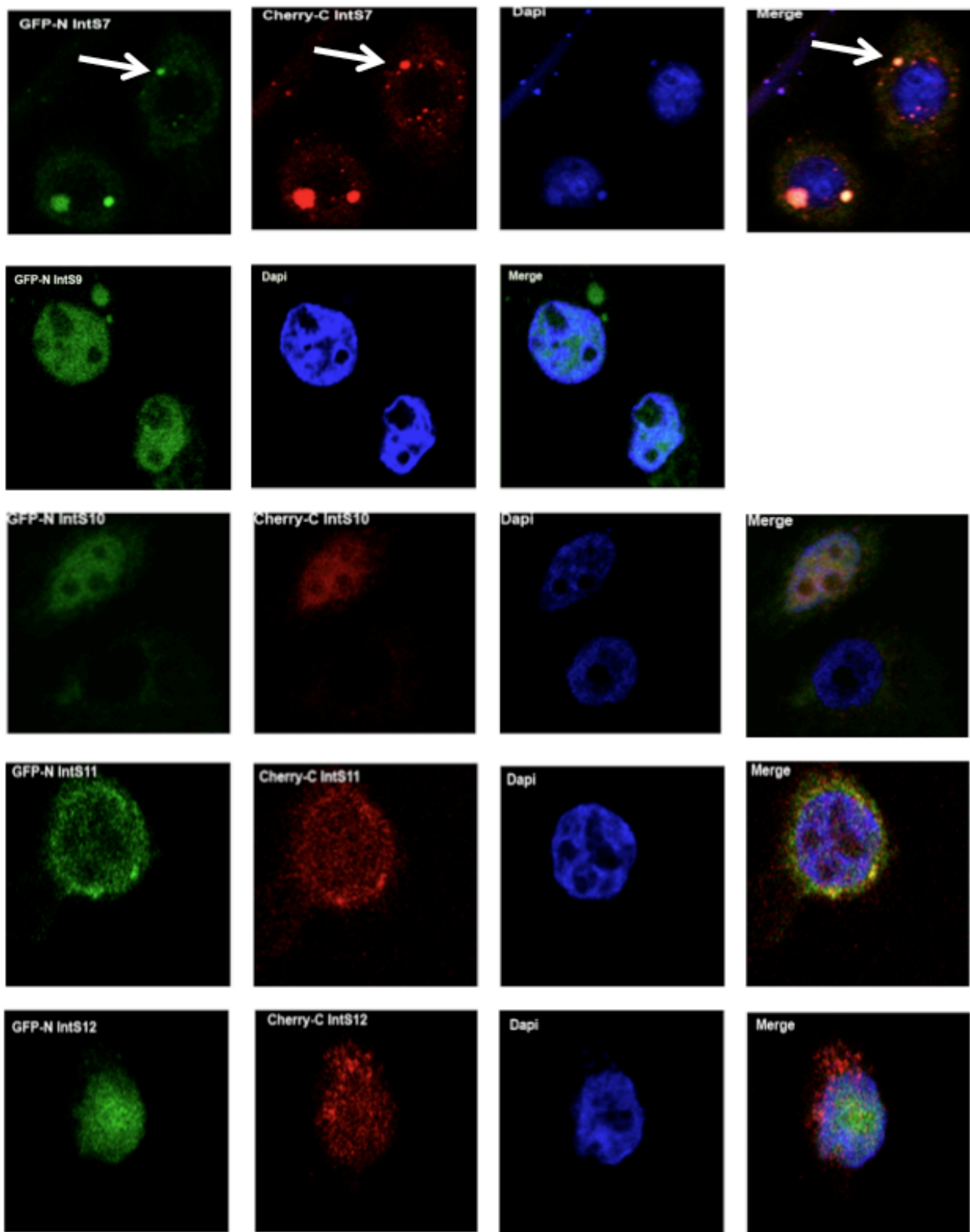
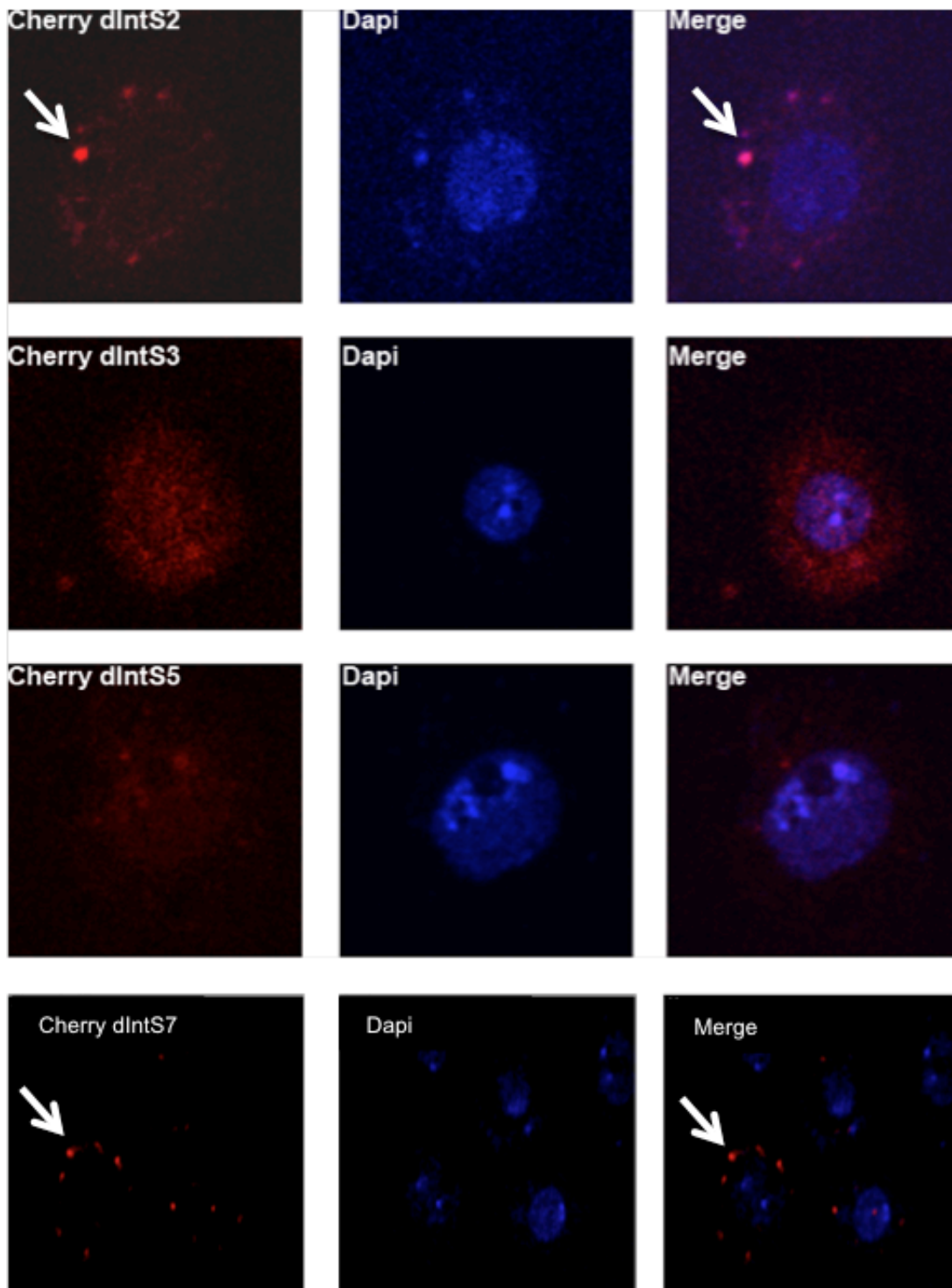
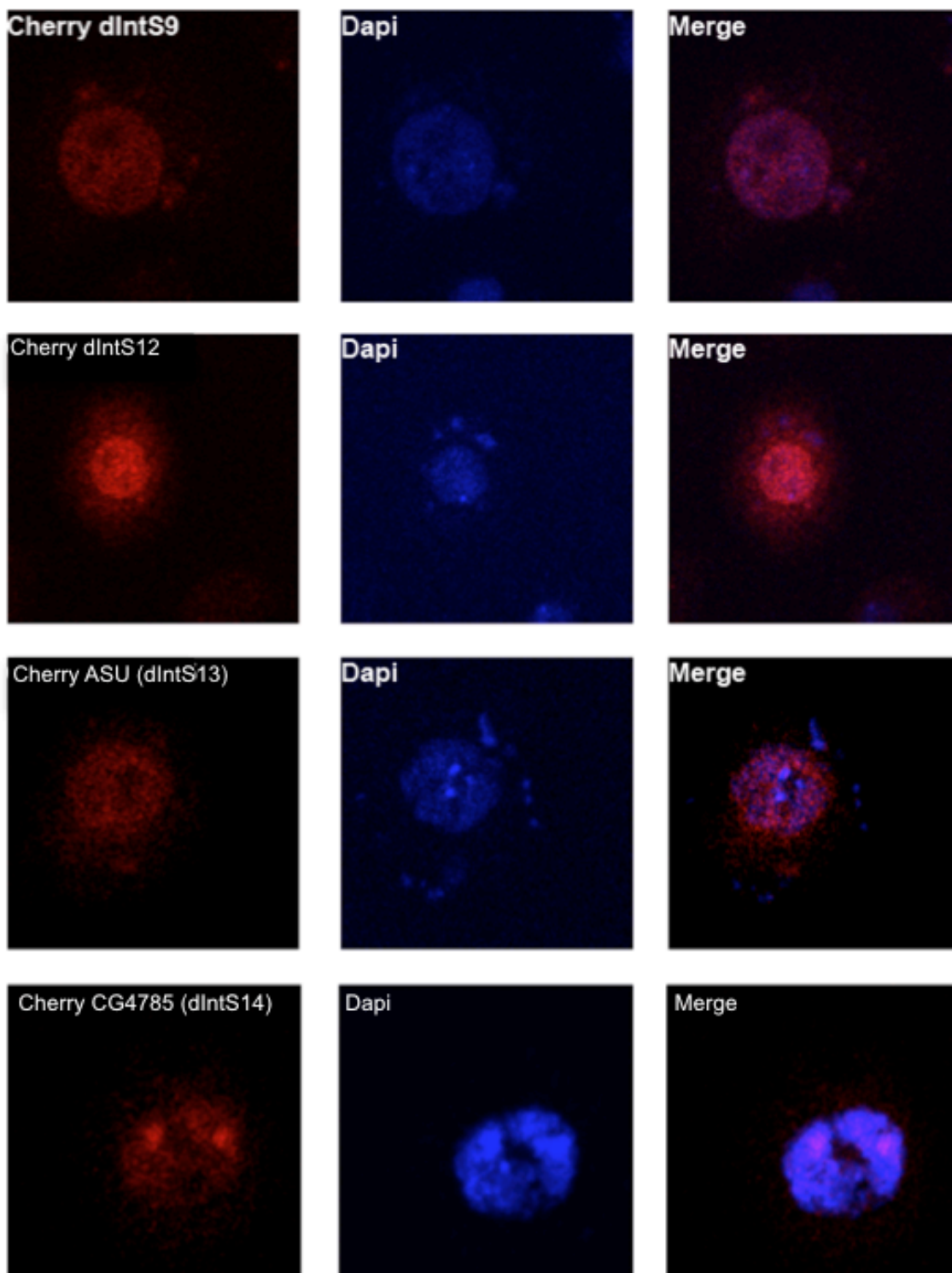


Figure 12: Confocal images of *Drosophila* Integrator subunit localizations.

Most of the *Drosophila* Integrator subunits display diffuse nuclear localization, however dIntS2 and 7 display discrete cytoplasmic foci (arrows).



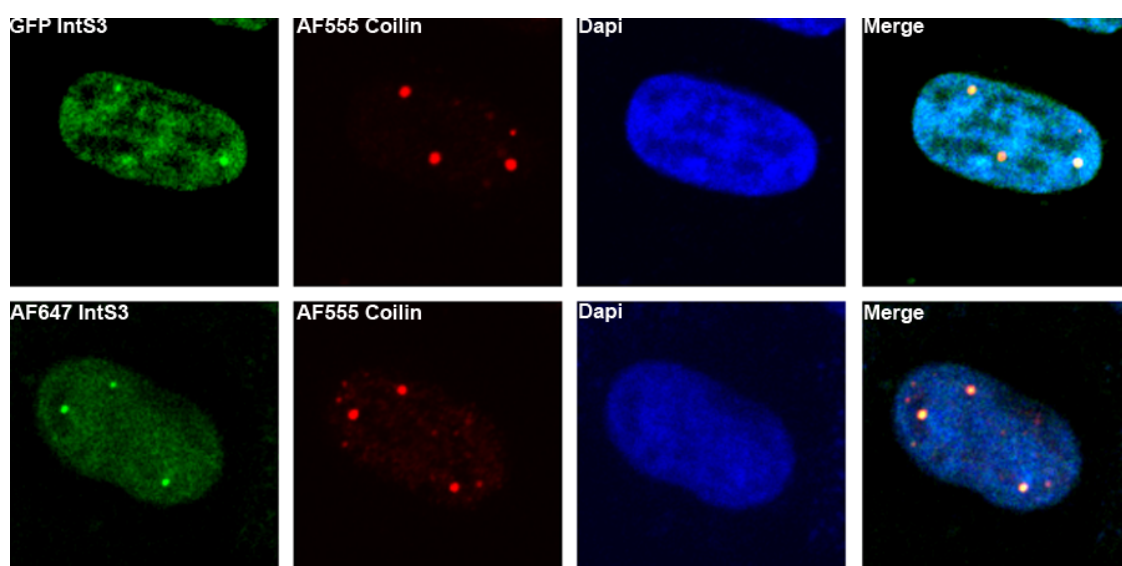


Human IntS3 Colocalizes with Cajal Body Marker Coilin in the Nucleus.

Given the localization of hIntS3 to nuclear foci we hypothesized that hIntS3 could be associating with a known nuclear body such as histone locus bodies, nuclear speckles, or Cajal bodies. Histone locus bodies form at the genes of replication dependent histones and contain biomolecules such as the U7 snRNP and other factors necessary for the 3' end processing of histone mRNA [65]. Nuclear speckles are sites of U snRNP maturation and storage; [66] while Cajal bodies are often observed adjacent to snRNA genes and are sites of snRNA transcription that have also been shown to monitor the export of snRNAs from the nucleus [71]. Since the Integrator complex is involved in snRNA transcription and 3' end processing, we hypothesized that hIntS3 localizes to the Cajal bodies.

Colocalization studies were carried out with coilin, a standard Cajal body marker to determine if our hypothesis was accurate. First, HeLa cells transfected with GFP-hIntS3 were stained for endogenous coilin (Figure 13, top panel). The GFP-hIntS3 foci colocalized completely with the coilin foci, implying that hIntS3 does localize to Cajal bodies in the nucleus. Colocalization studies staining both endogenous hIntS3 and coilin were also performed to confirm the localization of hIntS3 to Cajal bodies (Figure 13, bottom panel). Endogenous hIntS3 forms nuclear foci as well, and these foci completely colocalize with coilin foci, corroborating the results seen with GFP-tagged hIntS3. These results confirm our hypothesis that hIntS3 localizes to the Cajal bodies.

Figure 13: hIntS3 colocalizes with coilin. Both the exogenous GFP tagged hIntS3 (top panel) protein and the endogenous hIntS3 protein (bottom panel) colocalize with the Cajal body protein coilin. AF – AlexaFluor 555 (red) and 647 (green) were used to stain endogenous proteins.



Human and *Drosophila* IntS2 and 7 Colocalize with P Body Marker Dcp1 in Discrete Cytoplasmic Foci.

The most interesting and unexpected finding of this project was the localization of IntS2 and 7 to cytoplasmic foci, and as a result, these two subunits became my focus. As with hIntS3 in the nucleus, there were several cytoplasmic bodies considered to be possible locations for IntS2 and 7 in *Drosophila* and human cells, including stress granules, U bodies, and P bodies. Stress granules are sites where mRNPs stalled in translation initiation aggregate along with initiation factors, the 40S ribosomal subunit, and Pab1 [97]. U bodies are cytoplasmic foci that contain all the U snRNPs in addition to snRNP assembly factors such as SMN, and are hypothesized to be sites of U snRNP assembly [100]. Finally, P bodies are sites of storage for translationally repressed mRNPs as well as sites of mRNA decapping and decay and miRNA and siRNA gene silencing [93].

While stress granules and P bodies are both mRNA granules, U bodies are predicted to play a role in U snRNP assembly and were first thought to be a likely localization for cytoplasmic Integrator subunits. However, colocalization studies with hIntS7 and SMN performed by our collaborator, Mirek Dundr from Rosalind Franklin University in Chicago, Illinois showed that the two proteins do not colocalize (personal communication), therefore it is unlikely that IntS2 and 7 localize to U bodies. We therefore predicted IntS2 and 7 localize to P bodies because stress granules consist solely of translationally stalled mRNPs and the Integrator complex has not been implicated in translation. Additionally, it is possible that these Integrator subunits may play a role in the decay of U snRNAs.

To determine if IntS2 and 7 localize to P bodies colocalization studies with P body marker Dcp1 were performed. HeLa cells were co-transfected with 500ng of Cherry-hIntS2 or 7 and 500ng GFP-Dcp1 (plasmid obtained from the Shyu Laboratory) then fixed and imaged (Figure 14). hIntS2 and 7 foci colocalized completely with Dcp1 foci, tentatively confirming our hypothesis. Endogenous hIntS7 was also stained in HeLa cells (Figure 15, top panel) and this result confirmed the presence of hIntS7 in cytoplasmic foci. Endogenous hIntS7 was also stained in HeLa cells transfected with GFP-Dcp1 and the hIntS7 foci again colocalized with the GFP-Dcp1 foci (Figure 15, bottom panel), further demonstrating the presence of hIntS7 in P bodies. Colocalization of *Drosophila* IntS2 and 7 and Dcp1 was also performed in S2 cells co-transfected with 800ng of Cherry-dIntS2 or 7 and 800ng of *Drosophila* GFP-Dcp1 (plasmid cloned by Jiandong Chen while rotating in the Shyu laboratory) and these two subunits were also shown to colocalize with Dcp1 in this model as well (Figure 16).

IntS1 and 11 Are Required for IntS2 and 7 Localization to Cytoplasmic Foci in *Drosophila* S2 Cells.

Here, I addressed the question of the dependency of IntS2 and IntS7 cytoplasmic foci on expression of other Integrator subunits. The interdependency of dIntS2 and 7 on the expression of the other Integrator subunits was explored through systematically knocking down each Integrator subunit and monitoring the effects on the localization of dIntS2 and 7 in S2 cells. On day one, S2 cells were plated in a 96-well plate and treated with dsRNA to the 14 Integrator subunits as

Figure 14: Cherry hIntS2 and 7 colocalize with GFP Dcp1. Both mCherry tagged hIntS2 and 7 colocalize with GFP tagged P body marker Dcp1.

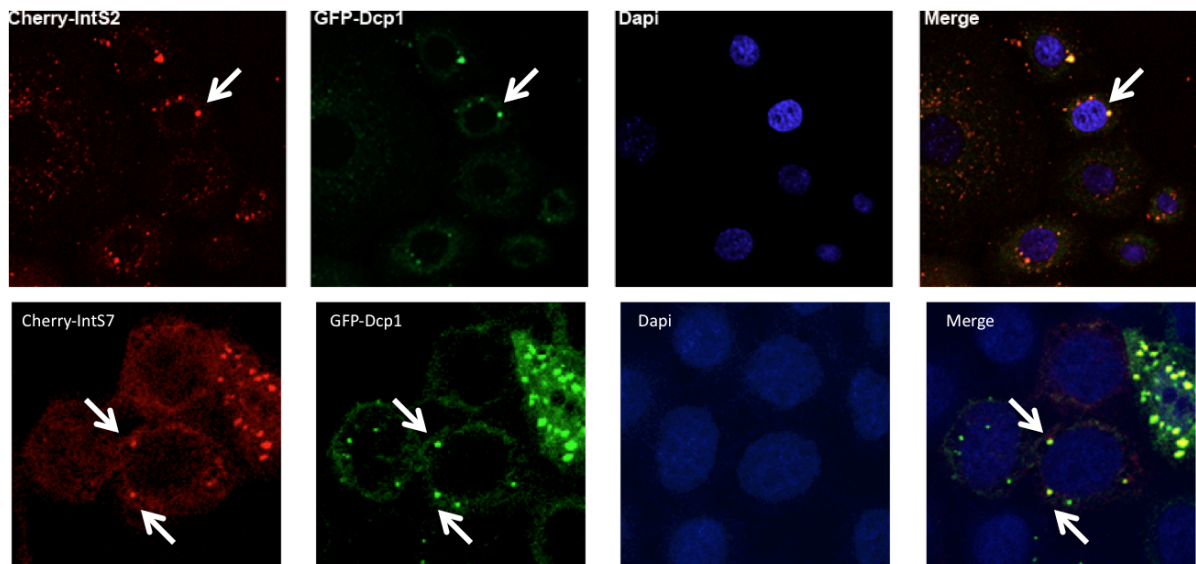


Figure 15: Endogenous hIntS7 localization. Endogenous hIntS7 localizes to discrete cytoplasmic foci (top panel) and these foci colocalize with GFP tagged Dcp1 (bottom panel). AF – AlexaFluor 555 was used to stain endogenous protein.

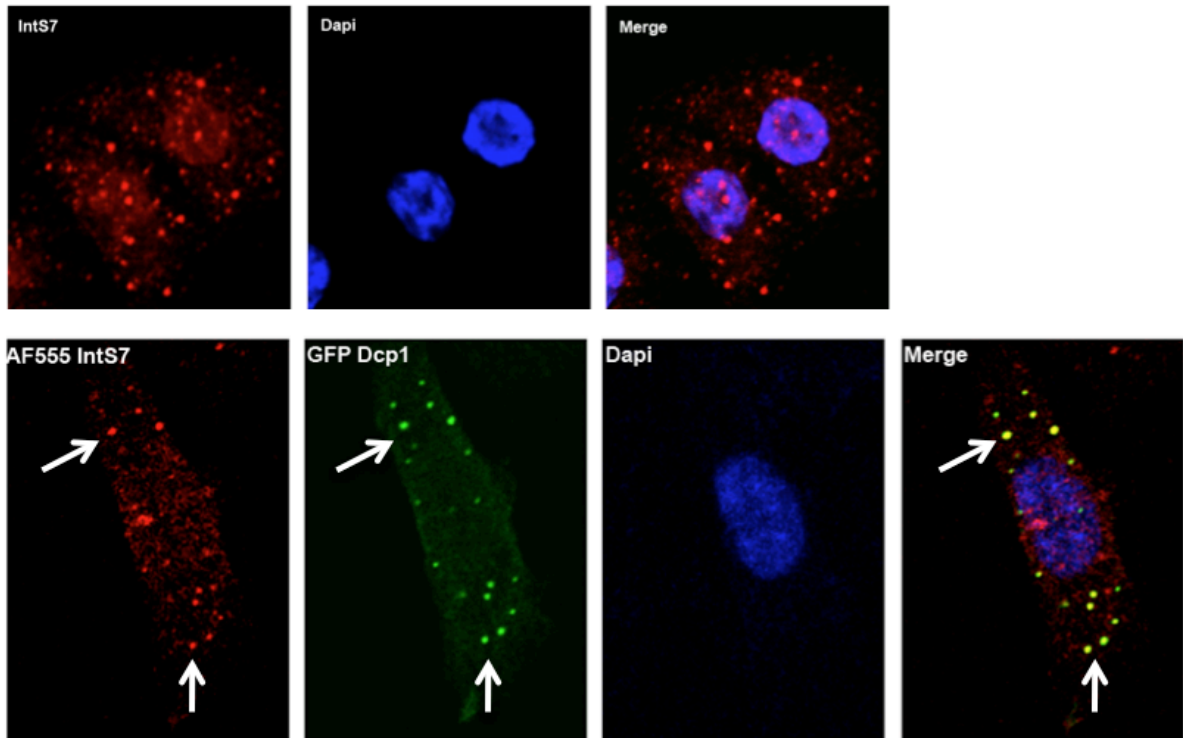
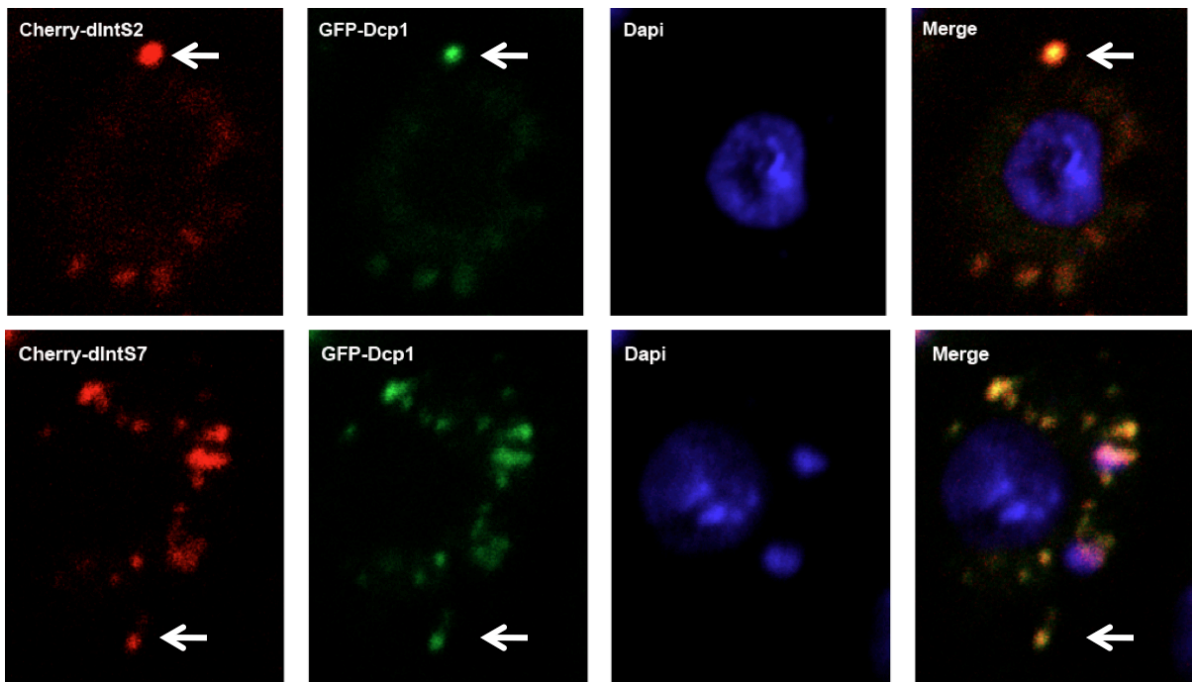


Figure 16: Cherry dIntS2 and 7 colocalize with GFP Dcp1. mCherry tagged dIntS2 and 7 colocalize with GFP tagged P body marker Dcp1 (arrows).



well as Thread, which was used as a positive control. On day two, the cells were transfected with 500ng/well of either Cherry-dIntS2 or 7, and on day three the cells were treated a second time with dsRNA. Finally, on day four the cells were fixed and visualized using confocal microscopy. It was found that the same two Integrator subunits, dIntS1 and 11 disrupted the localization of both dIntS2 and 7. When dIntS1 was knocked down dIntS2 was no longer expressed while dIntS11 knockdown cause dIntS2 to localize diffusely in the nucleus. dIntS7 was found to be either diffusely in the nucleus or diffusely in the cytoplasm when dIntS1 was knocked down, and knockdown of dIntS11 led to a diffuse cytoplasmic localization for dIntS7. These results are summarized in Table 6 and representative images for dIntS7 are shown in Figure 17.

The experiment was then repeated in cells transfected with Cherry dIntS2 or 7 or GFP-Dcp1, this time knocking down only dIntS1, 11, and 12. Knocking down IntS12 served as a negative control as we did not see any disruption in localization in the previous set of experiments. This experiment was performed in triplicate and the results quantified. As expected, when dIntS12 was knocked down, no change was observed in the localization of dIntS2 or 7 (Figure 18). When dIntS1 was knocked down Cherry-dIntS2 was localized to discrete foci in $56 \pm 4\%$ of cells. It was diffusely cytoplasmic in $36 \pm 4\%$ of cells and diffusely cytoplasmic in $8 \pm 4\%$ of cells. Cherry-dIntS7 was localized to discrete foci in $47 \pm 11\%$ of cells, diffusely cytoplasmic in $51 \pm 12\%$ of cells and diffusely nuclear in $1 \pm 1\%$ of cells (Figure 19). Cherry-dIntS2 was localized to discrete foci in $81 \pm 3\%$ of cells in which dIntS11 was knocked down, diffusely cytoplasmic in $12 \pm 3\%$ of cells and diffusely nuclear in

Table 6: Summary of Integrator subunit knockdown experiment. The effects of Integrator subunit knockdown on the localization of dIntS2 and 7 are listed. * dIntS2 is diffuse in the nucleus, # dIntS7 is diffuse in the nucleus or cytoplasm, ⌘ dIntS7 is diffuse in the cytoplasm.

IntS2	RNAi	Exp.	Foci
	IntS1	N	N
	IntS2	N	N
	IntS3	Y	Y
	IntS4	Y	Y
	IntS5	Y	Y
	IntS7	Y	Y
	IntS8	Y	Y
	IntS9	Y	Y
	IntS10	Y	Y
	IntS11	Y	N*
	IntS12	Y	Y
	ASU	Y	Y
	CG4785	Y	Y

IntS7	RNAi	Exp.	Foci
	IntS1	Y	N [#]
	IntS2	N	N
	IntS3	Y	Y
	IntS4	Y	Y
	IntS5	Y	Y
	IntS7	Y	Y
	IntS8	Y	Y
	IntS9	Y	Y
	IntS10	Y	Y
	IntS11	Y	N [⌘]
	IntS12	Y	Y
	ASU	Y	Y
	CG4785	Y	Y

Figure 17: Effects of Integrator subunit knockdown on dIntS7 localization.

The localization of dIntS7 exhibits no dependence on the expression of the majority of the other Integrator subunits (dIntS5 knockdown shown). However, the knockdown of dIntS1 produces dIntS7 localization that is either diffusely nuclear or cytoplasmic, and the knockdown of dIntS11 yields diffusely cytoplasmic dIntS7 localization.

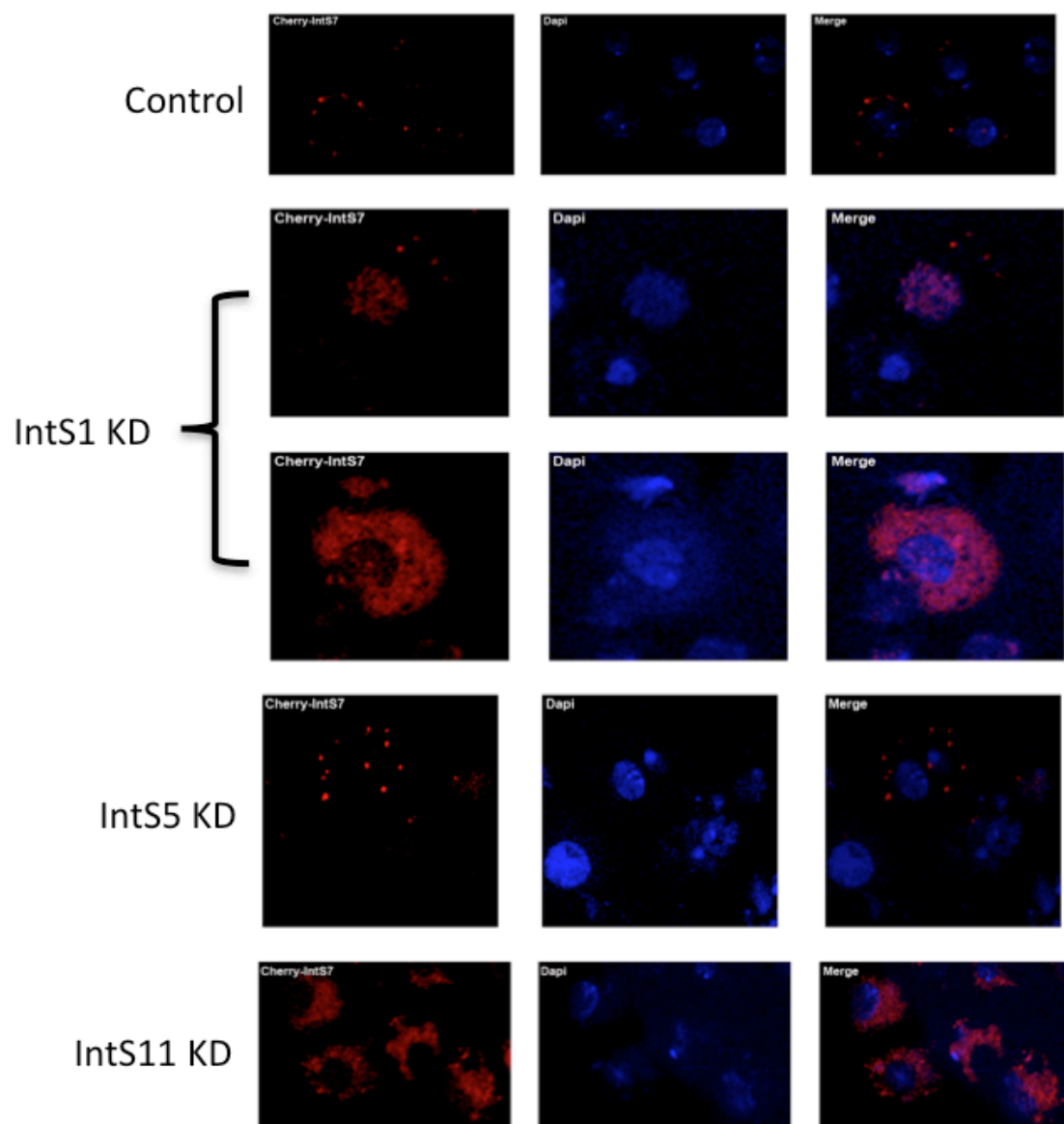
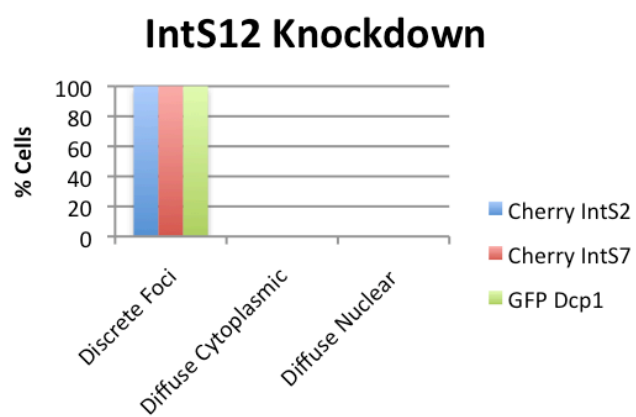


Figure 18: Quantification of the effects of dIntS12 knockdown on dIntS2 and 7 localization. A. When dIntS12 is knocked down, no change in mCherry dIntS2 and dIntS7 and GFP Dcp1 was observed. B. Representative images of localizations observed.

A.



B.

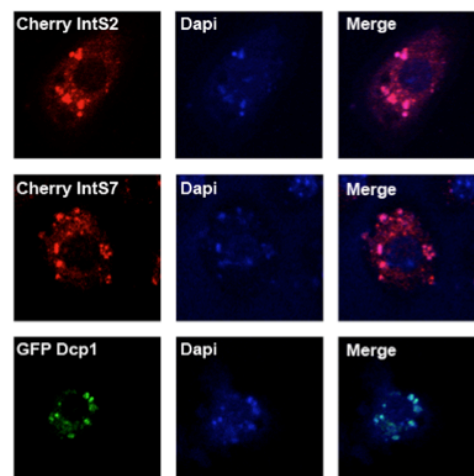
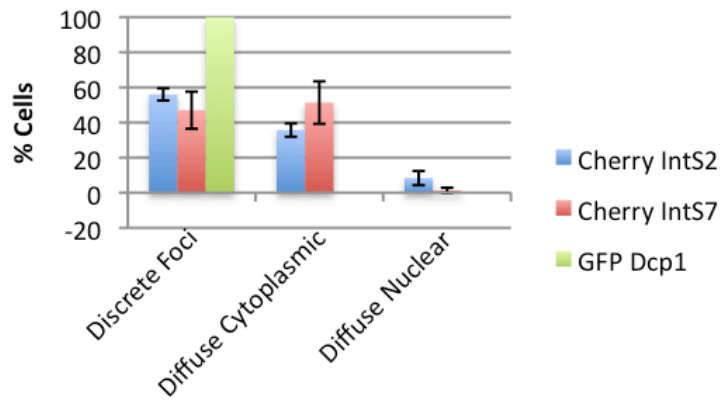


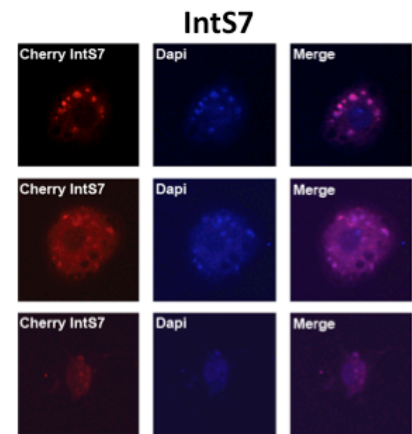
Figure 19: Quantification of the effects of dlntS1 knockdown on dlntS2 and 7 localization. A. When dlntS1 is knocked down, mCherry dlntS2 localized to discrete cytoplasmic foci in $56 \pm 4\%$ of cells, was diffusely cytoplasmic in $36 \pm 4\%$ of cells and was diffusely nuclear in $8 \pm 4\%$ of cells. mCherry dlntS7 localized to discrete foci in $47 \pm 11\%$ of cells, was diffusely cytoplasmic in $51 \pm 12\%$ of cells and diffusely nuclear in $1 \pm 1\%$ of cells. There was no change in the localization of GFP tagged Dcp1. B., C., D. Representative images of localizations observed.

A.

IntS1 Knockdown

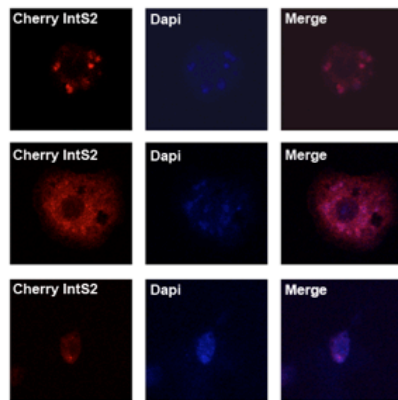


C.

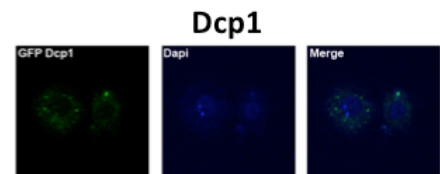


B.

IntS2



D.



7 ± 2% of cells. Cherry-dIntS7 was in discrete foci in 80 ± 1% of cells, diffusely cytoplasmic in 15 ± 2% of cells and diffusely nuclear in 5 ± 1% of cells (Figure 20). Additionally, no change was observed in the localization of GFP-Dcp1 when any of the Integrator subunits were knocked down.

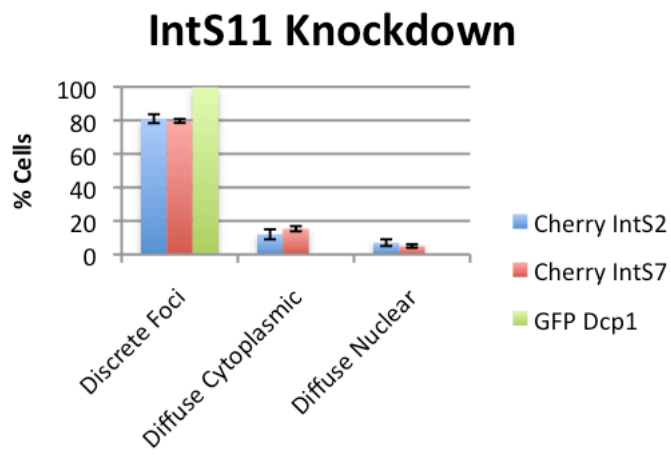
Deletions in Human IntS7 Gene Have Little Effect on hIntS7 Localization.

The amino acid alignment of IntS7 shown in Figure 21 highlights two conserved regions in the IntS7 protein. The first one (red box) corresponds to the ARM repeats which function in protein-protein interactions found in approximately the first 200 amino acids, and the second (blue box) is the DUF 3453, which is also found in Symplekin. To determine if these domains are responsible for the localization of hIntS7 to cytoplasmic foci a series of deletion mutants were created and their localizations were compared to that of the full-length hIntS7. Nine N-terminal and nine C-terminal deletion mutants were generated in the GFP-N vector (expression shown in Figure 22 C.) and these were co-transfected into HeLa cells with Cherry-hIntS7 and imaged using confocal microscopy. 500ng/well of Cherry-hIntS7 was transfected, however the amount of deletion mutant plasmid was adjusted to ensure even expression. Deletion of any part of the protein led to stronger nuclear staining where none was observed in cells transfected with the full-length protein. This localization, however, was not accompanied by a loss of discrete cytoplasmic foci (Figure 22 A. and B.).

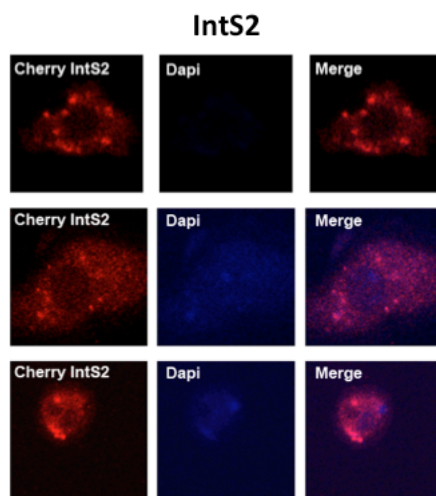
Depletion of Human IntS7 Leads to Increased snRNA Misprocessing.

Figure 20: Quantification of the effects of dlntS11 knockdown on dlntS2 and 7 localization. A. When dlntS11 is knocked down, mCherry dlntS2 localized to discrete cytoplasmic foci in $81 \pm 3\%$ of cells, was diffusely cytoplasmic in $12 \pm 3\%$ of cells and was diffusely nuclear in $7 \pm 2\%$ of cells. mCherry dlntS7 localized to discrete foci in $80 \pm 1\%$ of cells, was diffusely cytoplasmic in $15 \pm 1\%$ of cells and diffusely nuclear in $5 \pm 1\%$ of cells. There was no change in the localization of GFP tagged Dcp1. B., C., D. Representative images of localizations observed.

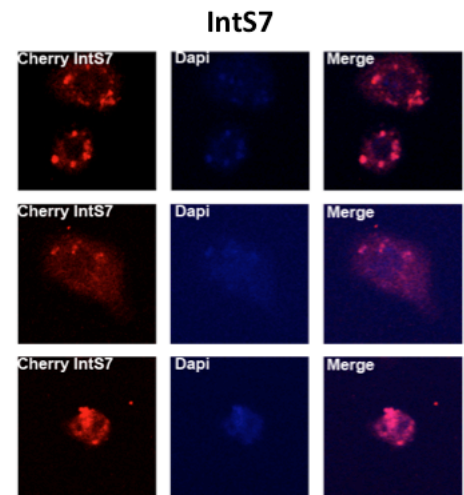
A.



B.



C.



D.

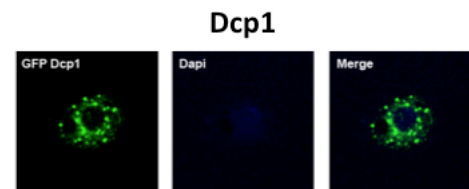
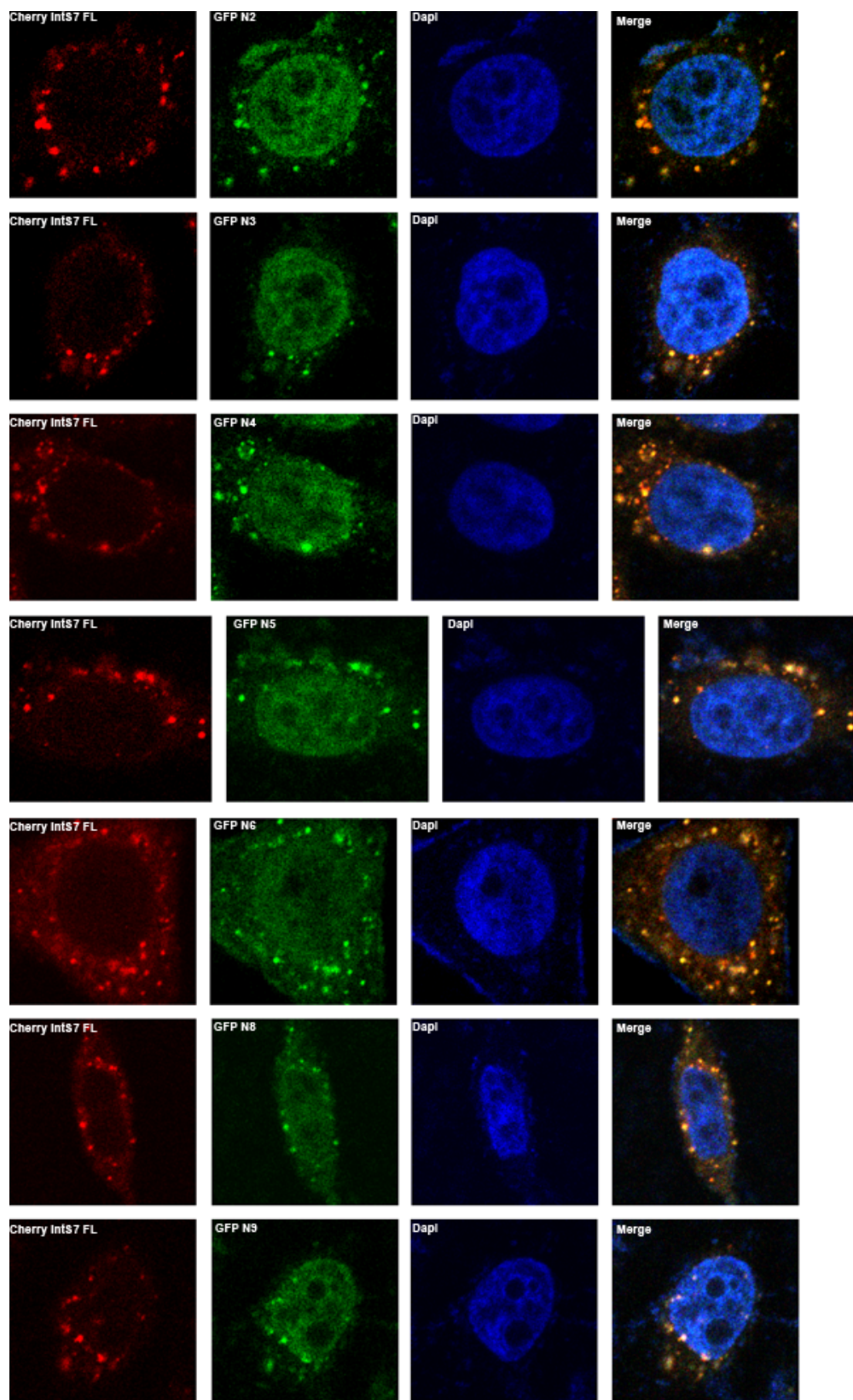


Figure 21: Six species alignment of IntS7. An alignment of the amino acid sequence of IntS7 for six species, human, cow, chicken, Xenopus, zebrafish, and Drosophila reveal two areas of high sequence similarity. The first (red box) corresponds to the ARM repeats in the first ≈ 200 amino acids. The second (blue box) corresponds to the DUF 3453.

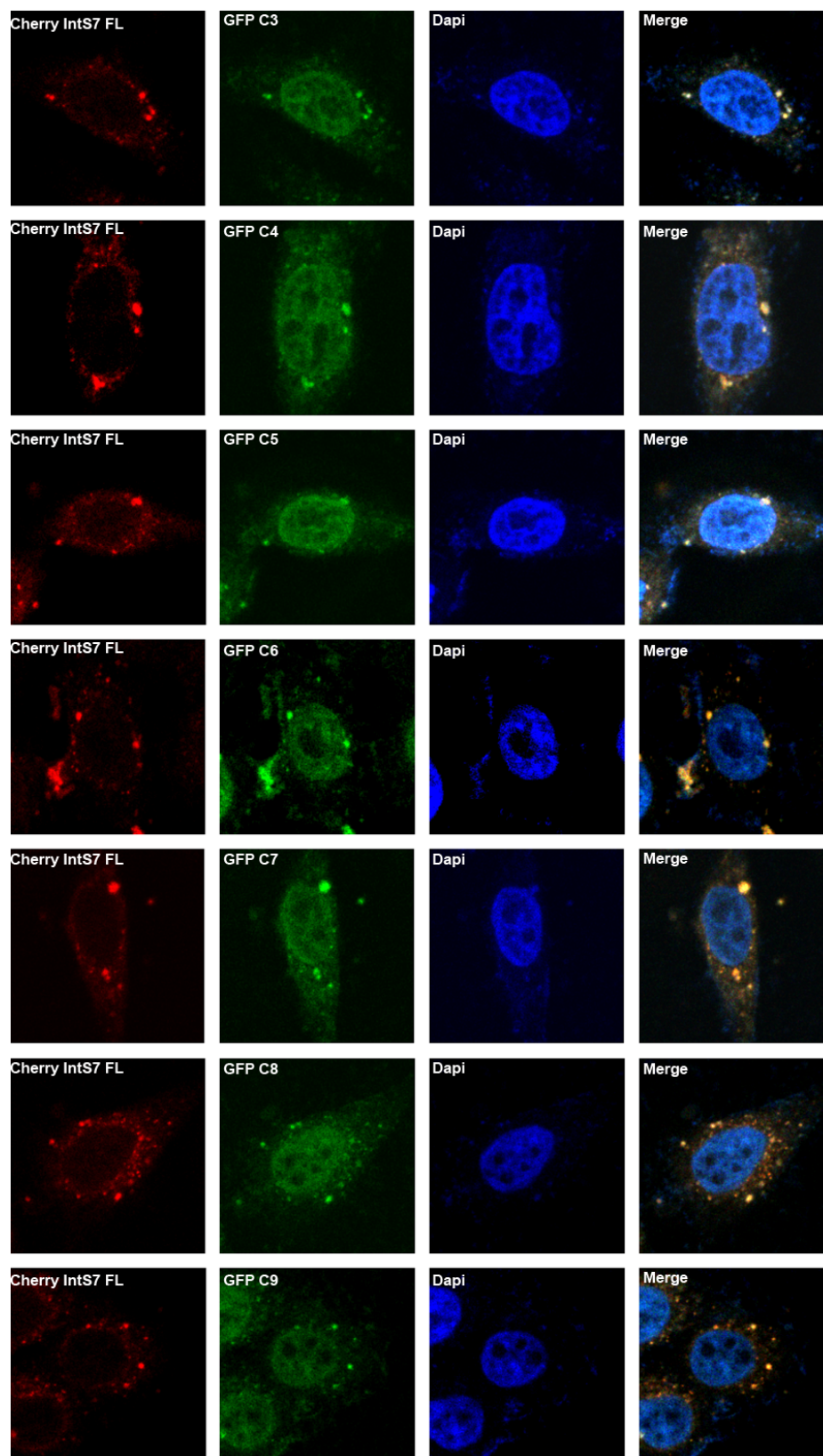
	(1)	1	10	20	30	40	50	60	70	80	90	100	110	120	130	140	152
human IntS7	(1)	---	MA	NR	SR	FL	AD	AG	---	---	---	---	---	---	---	---	---
cow IntS7	(1)	---	MA	NR	SR	FL	AD	AG	---	---	---	---	---	---	---	---	---
Chid.en IntS7	(1)	---	MA	NR	SR	FL	AD	AG	---	---	---	---	---	---	---	---	---
Xenopus IntS7	(1)	---	MA	NR	SR	FL	AD	AG	---	---	---	---	---	---	---	---	---
zebrafish IntS7	(1)	---	MA	NR	SR	FL	AD	AG	---	---	---	---	---	---	---	---	---
osphiha IntS7	(1)	---	MA	NR	SR	FL	AD	AG	---	---	---	---	---	---	---	---	---
Consensus	(1)	---	MA	NR	SR	FL	AD	AG	---	---	---	---	---	---	---	---	---
human IntS7	(153)	153	160	170	180	190	200	210	220	230	240	250	260	270	280	290	304
cow IntS7	(153)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Chid.en IntS7	(153)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Xenopus IntS7	(153)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
zebrafish IntS7	(153)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
osphiha IntS7	(153)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Consensus	(153)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
human IntS7	(304)	304	310	320	330	340	350	360	370	380	390	400	410	420	430	440	452
cow IntS7	(304)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Chid.en IntS7	(304)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Xenopus IntS7	(304)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
zebrafish IntS7	(304)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
osphiha IntS7	(304)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Consensus	(304)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
human IntS7	(456)	456	460	470	480	490	500	510	520	530	540	550	560	570	580	590	607
cow IntS7	(456)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Chid.en IntS7	(456)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Xenopus IntS7	(456)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
zebrafish IntS7	(456)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
osphiha IntS7	(456)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Consensus	(456)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
human IntS7	(608)	608	620	630	640	650	660	670	680	690	700	710	720	730	740	750	759
cow IntS7	(608)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Chid.en IntS7	(608)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Xenopus IntS7	(608)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
zebrafish IntS7	(608)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
osphiha IntS7	(608)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Consensus	(608)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
human IntS7	(760)	760	770	780	790	800	810	820	830	840	850	860	870	880	890	900	911
cow IntS7	(760)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Chid.en IntS7	(760)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Xenopus IntS7	(760)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
zebrafish IntS7	(760)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
osphiha IntS7	(760)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Consensus	(760)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
human IntS7	(883)	883	890	900	910	920	930	940	950	960	970	980	990	1000	1010	1020	1034
cow IntS7	(883)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Chid.en IntS7	(883)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Xenopus IntS7	(883)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
zebrafish IntS7	(883)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
osphiha IntS7	(883)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
Consensus	(883)	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Figure 22: GFP tagged hIntS7 deletion mutant localizations. Deletion of any part of hIntS7 leads to stronger nuclear staining, however, discrete cytoplasmic foci are still detectable. A. N-terminal deletions, B. C-terminal deletions, C. Western blot showing expression of deletion fragments.

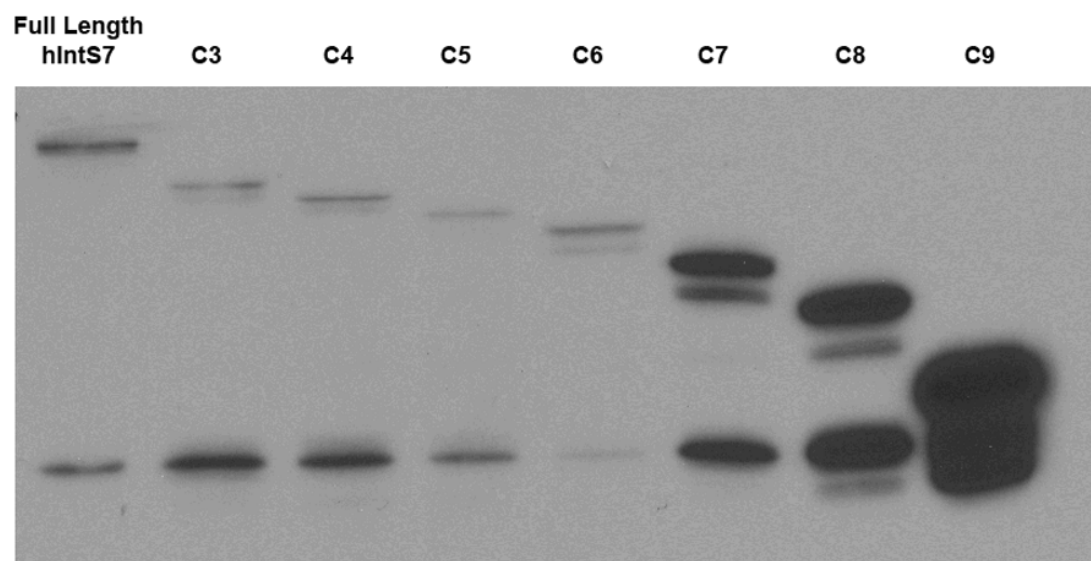
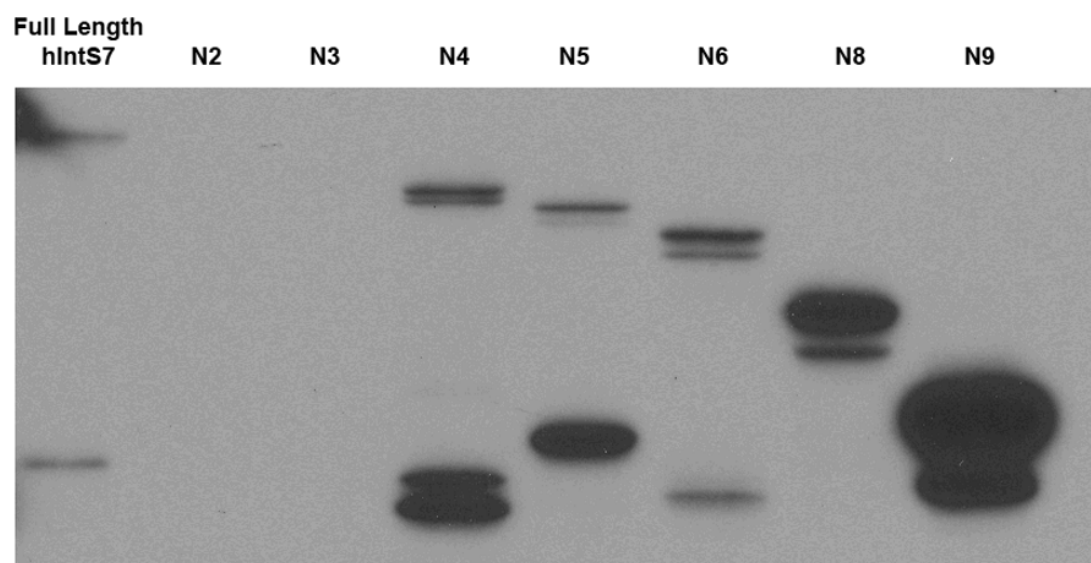
A.



B.

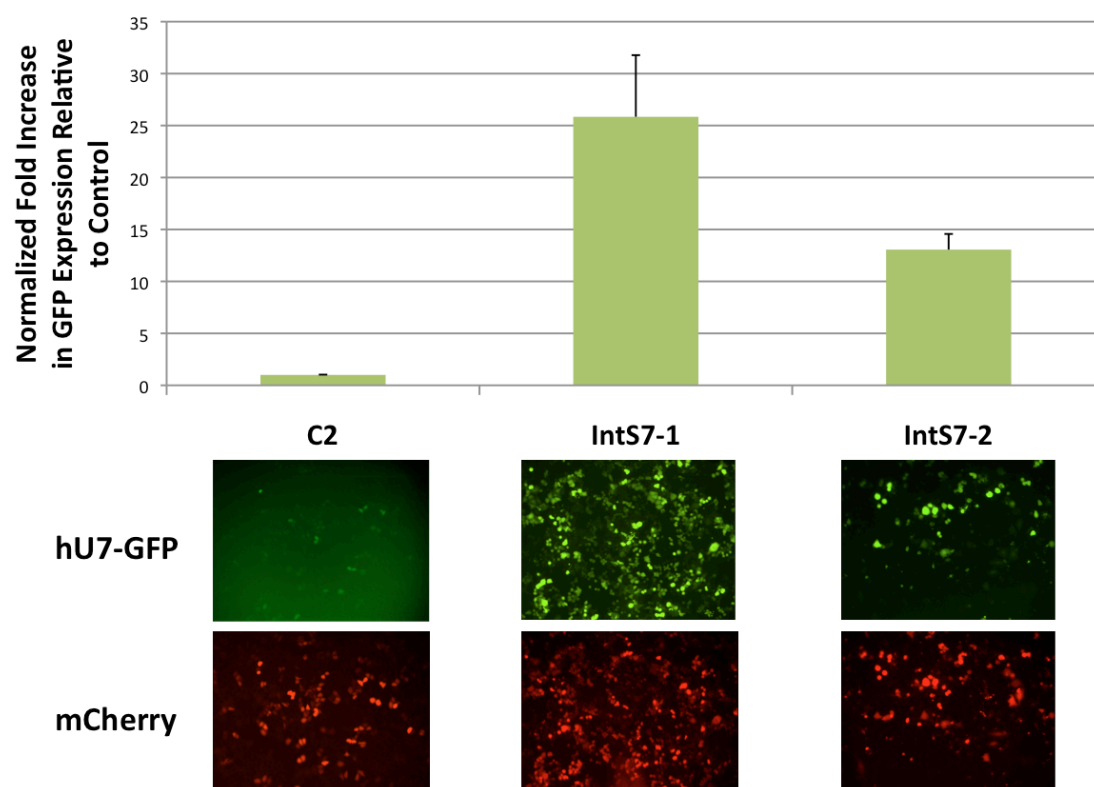


C.



While human and *Drosophila* IntS2 and 7 display a cytoplasmic localization, it has been shown in S2 cells that depletion of these subunits leads to accumulation of misprocessed snRNAs, suggesting that IntS2 and 7 also play a role in the nucleus [38]. A humanized version of the GFP-U7 snRNA reporter used in [38] was developed to determine if depletion of hIntS7 effects snRNA processing in HeLa cells. Two hIntS7 siRNA sequences were used in this experiment as well as a control siRNA sequence and the experiment was done in triplicate. Briefly, cells were plated in 24-well plates and allowed to attach overnight. Then, the cells were treated with siRNA once and again a second time 24 hours later. 24 hours after the last siRNA treatment, the cells were transfected with the GFP-U7 reporter plasmid and mCherry plasmid. The next day the amount of GFP and mCherry fluorescence was measured in multiple locations in each well and the results quantified. mCherry expression was used as a control readout and GFP expression was normalized to the amount of mCherry expression. The normalized GFP expression of the test siRNA wells was then normalized to the control siRNA wells to derive the fold increase in GFP expression in hIntS7 depleted cells (Figure 23). A 26 ± 6 fold increase in GFP expression in hIntS7 depleted cells (Figure 23). A 26 ± 6 fold increase in GFP expression was seen in cells treated with hIntS7 siRNA 7-1 and a 13 ± 2 fold increase in GFP expression was seen in cells treated with the 7-2 siRNA. From these results, we concluded that depletion of hIntS7 leads to increased misprocessing of snRNAs in human cells as well as *Drosophila* cells. These data also suggest that IntS7 is required for the processing occurring in nucleus and likely resides there as well as the cytoplasm. Additionally, in cells

Figure 23: Depletion of hIntS7 leads to misprocessing of snRNA in human cells. Misprocessing of snRNA in HeLa cells increases 25.84 ± 5.93 fold over control cells (C2) when hIntS7 is knocked down with siRNA 7-1 and 13.06 ± 1.51 fold when knocked down with siRNA 7-2 (graph). Bottom panel shows examples expression of hU7-GFP reporter and mCherry for each case.



stained for endogenous hIntS7, staining can be observed in the nucleus along with the cytoplasmic foci, further confirming this conclusion (Figure 15).

Chapter 4:

Conclusion and Future Directions

The purpose of this study was to paint a picture of Integrator subunit localization in the cell. Between the human and *Drosophila* models, localizations for all of the subunits, with the exception of IntS1 and 8 were revealed, and localizations seen in one model generally held for the other. From these images we conclude that the majority of the Integrators are diffusely spread across the nucleus, an expected result given the major function of the Integrator complex is to facilitate the proper 3' end processing of snRNAs. We also hypothesized that some of the subunits would localize to discrete subnuclear domains such as Cajal bodies, which are found adjacent to snRNA genes and are sites of snRNA transcription [50]. Confirming this, hIntS3 was seen in discrete nuclear foci that colocalized with Cajal body marker coilin. Unexpectedly, however, IntS2 and 7, in both human and *Drosophila* cells was observed in discrete cytoplasmic foci, a localization incompatible with the known nuclear function of the Integrator complex. These foci colocalized with Dcp1 suggesting that these subunits are present in P bodies.

The Localization of IntS3 in Cajal Bodies.

The colocalization of endogenous hIntS3 and coilin confirms the presence of hIntS3 in Cajal bodies, however, further biochemical experiments, such as immunoprecipitation (IP) of IntS3 and coilin as well as IP/mass spectrometry (MS) of Cajal body components would provide additional support for this conclusion [37]. While the role of IntS3 in Cajal bodies is unknown, it might function to recruit Cajal bodies and/or other Integrators to snRNA genes. Interestingly, it has been shown that depletion of IntS3 in S2 cells has no effect on snRNA processing [38]. These

data suggest that while IntS3 might facilitate the recruitment of these factors to snRNA genes, processing of snRNAs is not compromised when IntS3 is depleted. The role of IntS3 in snRNA processing in human cells has not been formally tested and thus it may behave differently in human cells playing a more dominant role in processing. One alternative hypothesis is that IntS3 is indeed required for processing of snRNA but it may also be required for transcription of snRNA. Therefore, depletion of IntS3, unlike other Integrator subunits, would not result in accumulation of misprocessed snRNA but rather a depletion of processed only. It would be interesting to determine if cells in which IntS3 is depleted produce less snRNA than cells with functioning IntS3. Additionally, it would be useful to determine the fate of Cajal bodies when IntS3 is depleted. If IntS3 is responsible for the recruitment of Cajal bodies to snRNA genes, depletion of IntS3 could prevent the localization of Cajal bodies to these sites or possibly prevent the formation of Cajal bodies all together.

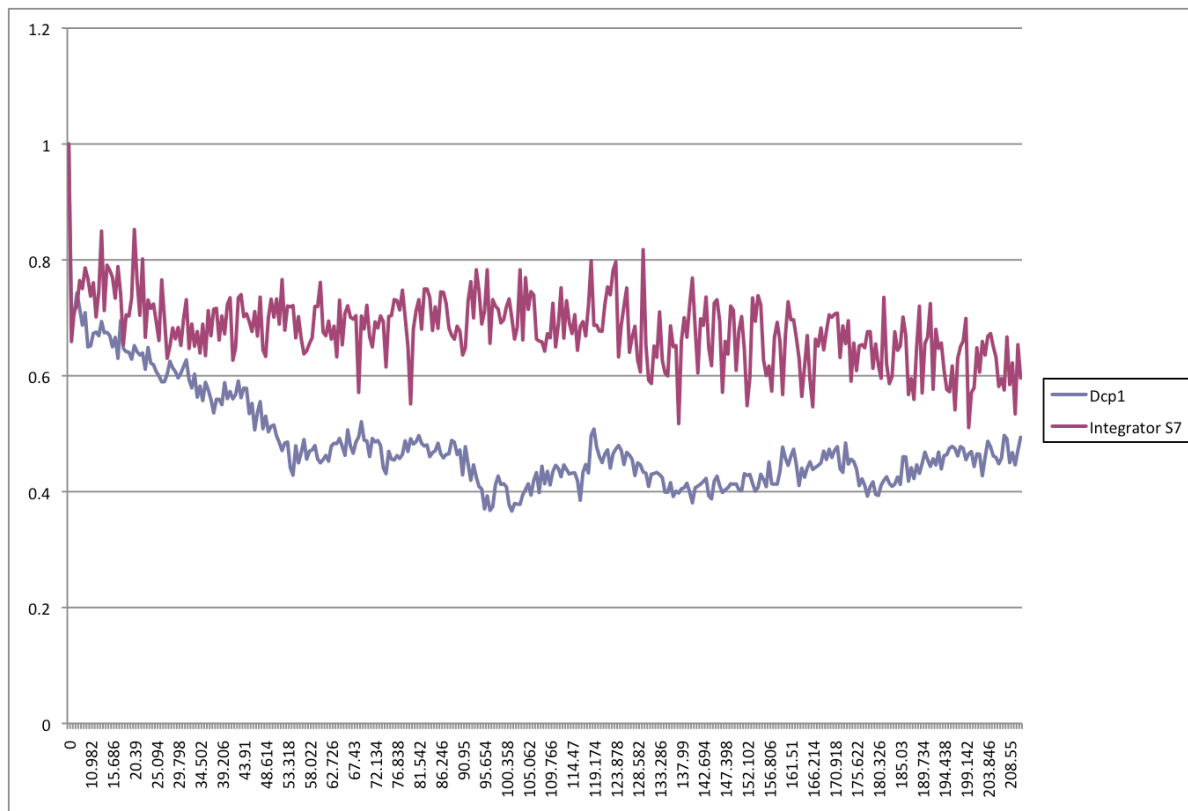
Integrators 2 and 7 Display a Conserved Phenotype of P Body Localization.

Here, we show that mCherry tagged IntS2 and 7 colocalize with GFP tagged Dcp1 in both HeLa and S2 cells. This finding was further confirmed through the detection of endogenous hIntS7 in P bodies. As with IntS3, further studies should be completed to confirm the presence of IntS2 and 7 in P bodies that are biochemical in nature. These include IP of IntS2/7 with Dcp1 and IP/MS of P body components as mentioned previously.

P bodies have been shown to be dynamic structures with molecules such as Dcp1 constantly entering and leaving [93]. This was determined using a technique called fluorescence recovery after photobleaching (FRAP) where a particular region is photobleached using a confocal microscope and the rate of the return of fluorescent molecules that that region is monitored. This experimental technique has been mastered by the Dundr laboratory at Rosalind Franklin University and was performed with our tagged Integrator constructs by Dr. Mirek Dundr. To determine if hIntS7 displays a dynamic association with P bodies similar to Dcp1 a similar technique called inverse (iFRAP) was used whereby the area surrounding a particular region is photobleached and the rate that the fluorescent molecules leave that region is monitored. Cells were transfected with Cherry-hIntS7 and GFP-Dcp1, and iFRAP was performed on several hIntS7 and GFP foci. We predicted that the association of hIntS7 with the cytoplasmic foci would be dynamic with hIntS7 proteins slowly migrating away from the targeted foci, similar to what is seen with Dcp1 foci. However, this was not the case. While the Dcp1 foci were dynamic in this experiment, hIntS7 was very stable in the foci with very little protein moving away (Figure 24).

Although IntS2 and 7 appear to colocalize with Dcp1 in P bodies, the Integrator subunits behave differently in these structures than Dcp1 evidenced by the results of the iFRAP experiment described above, which demonstrates that hIntS7 is more stable within its foci while Dcp1 is more dynamic. Testing the effects of drugs, such as, cycloheximide and RNase A, which have been shown to disrupt the formation of P bodies in cells, on IntS2 and 7 foci as well as the effects of

Figure 24: Inverse FRAP of mCherry hIntS7 and GFP Dcp1. The Y-axis represents the percentage of original fluorescence immediately after iFRAP. The percentage decreases over time as the GFP or mCherry-fusion protein diffuses out of the foci. The X-axis is a timescale of the experiment in seconds.



depletion of Dcp1 or other P body components would provide further insight to the behavior of these proteins in the cytoplasm.

The localization of these two proteins in P bodies was not expected. Aside from these structures residing in the cytoplasm where snRNA processing does not occur, P bodies contain many factors involved in mRNA destabilization or translational repression. Thus, it is not clear what the function of Integrator subunits would be while in P bodies. It is possible that they have a distinct cytoplasmic-specific function that is somehow relevant to mRNA decay. Alternatively, one function that fits nicely with P body localization is snRNA decay. While IntS2 and 7 do not seem to play a role in the decay of misprocessed snRNA, it is possible that they target snRNAs not properly assembled into snRNPs for decay in P bodies. One hypothesis is that IntS2 and 7 remain associated with the snRNA after it is cleaved and exported to the cytoplasm. Once in the cytoplasm, IntS2 and 7 monitor the assembly of the snRNP, and shuttles misassembled snRNP to P bodies for decay. Blocking nuclear export with leptomycin B, for example, would be one way to test this hypothesis, however, the iFRAP study suggests against the shuttling of IntS7 between the nucleus and the cytoplasm. It is still possible that these Integrators play a role in the decay of misassembled snRNPs or snRNAs exported to the cytoplasm, but not assembled into snRNPs. To investigate this theory, SMN or one of the Sm proteins could be knocked down via RNAi and the size of the IntS2 and 7 foci could be monitored. Larger foci would suggest that IntS2 and 7 are shuttling snRNAs unable to be assembled into snRNP to the P bodies for decay. Additionally, *in situ* hybridization could be used to determine if snRNAs can be

detected in P bodies. It is also possible that IntS2 and 7 in P bodies function in a role that is completely separate from snRNA 3' end processing and decay.

While it is unknown if the disruption of Dcp1 or other P body components affects the expression of IntS2 and 7 in cytoplasmic foci, the depletion of Ints1 and 11 in *Drosophila* S2 cells appears to reduce the number of cells with IntS2 or 7 foci. IntS11 is the catalytic subunit of the Integrator complex, responsible for the 3' cleavage of the nascent snRNA. Its depletion produces misprocessed snRNA in the cell [38] suggesting that properly processed snRNA is necessary for the localization of IntS2 and 7 to cytoplasmic foci. IntS1 causes abundant misprocessing when depleted as well, supporting this hypothesis. However, the depletion of other Integrator subunits, which also cause misprocessing when disrupted, had no effect on the localization of IntS2 and 7. This discrepancy implies the disruption of the localization of IntS2 and 7 through the depletion of IntS1 and 11 could be unrelated to the 3' end processing of snRNA and these four Integrator subunits form a sub complex that performs an entirely different function.

Domain Analysis of IntS7 with Respect to Its Cellular Localization.

In this study we attempted to define the domain responsible for the localization of Ints7 to cytoplasmic foci by designing GFP tagged deletion mutants of hIntS7. We hypothesized that the highly conserved ARM repeat region at the N-terminus of the protein may be involved in the aggregation of the protein in cytoplasmic foci considering ARM repeats function in protein-protein interactions. However, this was not the case, as cytoplasmic foci could still be detected in

mutants in which this region was deleted. In fact, cytoplasmic foci could be observed in all deletion mutants made, and we were unable to pinpoint an exact region responsible for the formation of foci using this method. These foci colocalized with full-length hIntS7 foci; however, the deletion mutants had a strong nuclear presence, where the full-length protein was not seen in the nucleus. This suggests that deletion of any part of the IntS7 protein causes it to be retained in the nucleus to a greater extent than the full-length protein. One method in which P bodies are thought to assemble is through the interaction of glutamine/asparagine (Q/N) rich prion-like domains found in many P body components [102]. IntS7 contains a glutamine-rich stretch at its C-terminus, which could be responsible for its apparent aggregation into P bodies; however, no effect was observed with the removal of this region in our experiment.

In conclusion, this study has demonstrated the cellular localization of the majority of the Integrator complex subunits. As expected, we have found that one of the subunits, IntS3 localizes to the Cajal bodies in the nucleus. Unexpectedly, however, we have shown that two of the subunits, IntS2 and 7 localize to discrete cytoplasmic foci that colocalize to P bodies. This was the first time any Integrator subunits have been associated with these cytoplasmic structures. These localizations provide clues as to the function of these Integrator subunits in snRNA biogenesis and snRNP assembly as well as suggest the possibility of function in other cellular processes. In addition to this, this study has also generated many useful tools, such as the tagged Integrator clones for both human and *Drosophila*

cells, that can be used to further the understanding of the Integrator complex and its functions in the future.

References

1. Matera, A.G., R.M. Terns, and M.P. Terns, *Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs*. Nat Rev Mol Cell Biol, 2007. **8**(3): p. 209-20.
2. Chen, J. and E.J. Wagner, *snRNA 3' end formation: the dawn of the Integrator complex*. Biochem Soc Trans. **38**(4): p. 1082-7.
3. Neuenkirchen, N., A. Chari, and U. Fischer, *Deciphering the assembly pathway of Sm-class U snRNPs*. FEBS Lett, 2008. **582**(14): p. 1997-2003.
4. Schumperli, D. and R.S. Pillai, *The special Sm core structure of the U7 snRNP: far-reaching significance of a small nuclear ribonucleoprotein*. Cell Mol Life Sci, 2004. **61**(19-20): p. 2560-70.
5. Marzluff, W.F., *Metazoan replication-dependent histone mRNAs: a distinct set of RNA polymerase II transcripts*. Curr Opin Cell Biol, 2005. **17**(3): p. 274-80.
6. Kunkel, G.R., R.L. Maser, J.P. Calvet, and T. Pederson, *U6 small nuclear RNA is transcribed by RNA polymerase III*. Proc Natl Acad Sci U S A, 1986. **83**(22): p. 8575-9.
7. Reddy, R., D. Henning, G. Das, M. Harless, and D. Wright, *The capped U6 small nuclear RNA is transcribed by RNA polymerase III*. J Biol Chem, 1987. **262**(1): p. 75-81.
8. Matera, A.G., A.M. Weiner, and C.W. Schmid, *Structure and evolution of the U2 small nuclear RNA multigene family in primates: gene amplification under natural selection?* Mol Cell Biol, 1990. **10**(11): p. 5876-82.

9. Card, C.O., G.F. Morris, D.T. Brown, and W.F. Marzluff, *Sea urchin small nuclear RNA genes are organized in distinct tandemly repeating units*. Nucleic Acids Res, 1982. **10**(23): p. 7677-88.
10. Westin, G., J. Zabielski, K. Hammarstrom, H.J. Monstein, C. Bark, and U. Pettersson, *Clustered genes for human U2 RNA*. Proc Natl Acad Sci U S A, 1984. **81**(12): p. 3811-5.
11. Hernandez, N., *Small nuclear RNA genes: a model system to study fundamental mechanisms of transcription*. J Biol Chem, 2001. **276**(29): p. 26733-6.
12. Carbon, P., S. Murgo, J.P. Ebel, A. Krol, G. Tebb, and L.W. Mattaj, *A common octamer motif binding protein is involved in the transcription of U6 snRNA by RNA polymerase III and U2 snRNA by RNA polymerase II*. Cell, 1987. **51**(1): p. 71-9.
13. Hernandez, N., *Formation of the 3' end of U1 snRNA is directed by a conserved sequence located downstream of the coding region*. Embo J, 1985. **4**(7): p. 1827-37.
14. Yuo, C.Y., M. Ares, Jr., and A.M. Weiner, *Sequences required for 3' end formation of human U2 small nuclear RNA*. Cell, 1985. **42**(1): p. 193-202.
15. Uguen, P. and S. Murphy, *The 3' ends of human pre-snRNAs are produced by RNA polymerase II CTD-dependent RNA processing*. Embo J, 2003. **22**(17): p. 4544-54.
16. Mandel, C.R., Y. Bai, and L. Tong, *Protein factors in pre-mRNA 3'-end processing*. Cell Mol Life Sci, 2008. **65**(7-8): p. 1099-122.

17. Murthy, K.G. and J.L. Manley, *The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation*. Genes Dev, 1995. **9**(21): p. 2672-83.
18. MacDonald, C.C., J. Wilusz, and T. Shenk, *The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location*. Mol Cell Biol, 1994. **14**(10): p. 6647-54.
19. Callebaut, I., D. Moshous, J.P. Mornon, and J.P. de Villartay, *Metallo-beta-lactamase fold within nucleic acids processing enzymes: the beta-CASP family*. Nucleic Acids Res, 2002. **30**(16): p. 3592-601.
20. Mandel, C.R., S. Kaneko, H. Zhang, D. Gebauer, V. Vethantham, J.L. Manley, and L. Tong, *Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease*. Nature, 2006. **444**(7121): p. 953-6.
21. Marzluff, W.F., E.J. Wagner, and R.J. Duronio, *Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail*. Nat Rev Genet, 2008. **9**(11): p. 843-54.
22. Wang, Z.F., M.L. Whitfield, T.C. Ingledue, 3rd, Z. Dominski, and W.F. Marzluff, *The protein that binds the 3' end of histone mRNA: a novel RNA-binding protein required for histone pre-mRNA processing*. Genes Dev, 1996. **10**(23): p. 3028-40.
23. Mowry, K.L. and J.A. Steitz, *Identification of the human U7 snRNP as one of several factors involved in the 3' end maturation of histone premessenger RNA's*. Science, 1987. **238**(4834): p. 1682-7.

24. Dominski, Z., X.C. Yang, and W.F. Marzluff, *The polyadenylation factor CPSF-73 is involved in histone-pre-mRNA processing*. Cell, 2005. **123**(1): p. 37-48.
25. Kolev, N.G., T.A. Yario, E. Benson, and J.A. Steitz, *Conserved motifs in both CPSF73 and CPSF100 are required to assemble the active endonuclease for histone mRNA 3'-end maturation*. EMBO Rep, 2008. **9**(10): p. 1013-8.
26. Sullivan, K.D., M. Steiniger, and W.F. Marzluff, *A core complex of CPSF73, CPSF100, and Symplekin may form two different cleavage factors for processing of poly(A) and histone mRNAs*. Mol Cell, 2009. **34**(3): p. 322-32.
27. de Vegvar, H.E., E. Lund, and J.E. Dahlberg, *3' end formation of U1 snRNA precursors is coupled to transcription from snRNA promoters*. Cell, 1986. **47**(2): p. 259-66.
28. Hernandez, N. and A.M. Weiner, *Formation of the 3' end of U1 snRNA requires compatible snRNA promoter elements*. Cell, 1986. **47**(2): p. 249-58.
29. Jacobs, E.Y., I. Ogiwara, and A.M. Weiner, *Role of the C-terminal domain of RNA polymerase II in U2 snRNA transcription and 3' processing*. Mol Cell Biol, 2004. **24**(2): p. 846-55.
30. Medlin, J.E., P. Uguen, A. Taylor, D.L. Bentley and S. Murphy, *The C-terminal domain of pol II and a DRB-sensitive kinase are required for 3' processing of U2 snRNA*. Embo J, 2003. **22**(4): p. 925-34.
31. Egloff, S., D. O'Reilly, R.D. Chapman, A. Taylor, K. Tanzhaus L. Pitts, D. Eick, and S. Murphy, *Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression*. Science, 2007. **318**(5857): p. 1777-9.

32. Egloff, S., S.A. Szczepaniak, M. Dienstbier, A. Taylor, S. Knight, and S. Murphy, *The integrator complex recognizes a new double mark on the POL II CTD*. J Biol Chem.
33. Baillat, D., M.A. Hakimi, A.M. Naar, A. Shiatifard, N. Cooch, and R. Shiekhattar, *Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II*. Cell, 2005. **123**(2): p. 265-76.
34. Egloff, S. and S. Murphy, *Role of the C-terminal domain of RNA polymerase II in expression of small nuclear RNA genes*. Biochem Soc Trans, 2008. **36**(Pt 3): p. 537-9.
35. Dominski, Z., X.C. Yang, M. Purdy, E.J. Wagner, and W.F. Marzluff, *A CPSF-73 homologue is required for cell cycle progression but not cell growth and interacts with a protein having features of CPSF-100*. Mol Cell Biol, 2005. **25**(4): p. 1489-500.
36. Egloff, S., D. O'Reilly, and S. Murphy, *Expression of human snRNA genes from beginning to end*. Biochem Soc Trans, 2008. **36**(Pt 4): p. 590-4.
37. Malovannaya, A., Y. Li, Y. Bulynko, S.Y. Jung, Y. Wang, R.B. Lanz, B.W. O'Malley, and J. Qin, *Streamlined analysis schema for high-throughput identification of endogenous protein complexes*. Proc Natl Acad Sci U S A, 2010. **107**(6): p. 2431-6.
38. Ezzeddine, N., J. Chen, B. Waltenspiel, B. Burch, T. Albrecht, M. Zhuo, W.D. Warren, W.F. Marzluff, and E.J. Wagner, *A subset of Drosophila integrator*

- proteins is essential for efficient U7 snRNA and spliceosomal snRNA 3'-end formation. Molecular and cellular biology, 2011. 31(2): p. 328-41.*
39. Steinmetz, E.J. and D.A. Brow, *Repression of gene expression by an exogenous sequence element acting in concert with a heterogeneous nuclear ribonucleoprotein-like protein, Nrd1, and the putative helicase Sen1. Mol Cell Biol, 1996. 16(12): p. 6993-7003.*
 40. Steinmetz, E.J., N.K. Conrad, D.A. Brow, and J.L. Corden, *RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. Nature, 2001. 413(6853): p. 327-31.*
 41. Hata, T. and M. Nakayama, *Targeted disruption of the murine large nuclear KIAA1440/Ints1 protein causes growth arrest in early blastocyst stage embryos and eventual apoptotic cell death. Biochim Biophys Acta, 2007. 1773(7): p. 1039-51.*
 42. Tao, S., Y. Cai, and K. Sampath, *The Integrator subunits function in hematopoiesis by modulating Smad/BMP signaling. Development, 2009. 136(16): p. 2757-65.*
 43. Skaar, J.R., D.J. Richar, A. Saraf, A. Toschi, E. Bolderson, L. Florens, M.P. Wasburn, K.K. Khanna, and W. Wang, *INTS3 controls the hSSB1-mediated DNA damage response. J Cell Biol, 2009. 187(1): p. 25-32.*
 44. Filleur, S., J. Hirsch, A. Wille, M. Schon, C. Sell, M.H. Shearer, T. Nelius, and I. Wieland, *INTS6/DICE1 inhibits growth of human androgen-independent prostate cancer cells by altering the cell cycle profile and Wnt signaling. Cancer Cell Int, 2009. 9: p. 28.*

45. Wieland, I., K.C. Arden, D. Michels, L. Klein-Hitpass, M. Bohm, C.S. Viars, and U.H. Weidle, *Isolation of DICE1: a gene frequently affected by LOH and downregulated in lung carcinomas*. *Oncogene*, 1999. **18**(32): p. 4530-7.
46. Wieland, I., C. Sell, U.H. Weidle, and P. Wieacker, *Ectopic expression of DICE1 suppresses tumor cell growth*. *Oncol Rep*, 2004. **12**(2): p. 207-11.
47. Patel, S.B. and M. Bellini, *The assembly of a spliceosomal small nuclear ribonucleoprotein particle*. *Nucleic Acids Res*, 2008. **36**(20): p. 6482-93.
48. Kiss, T., *Biogenesis of small nuclear RNPs*. *J Cell Sci*, 2004. **117**(Pt 25): p. 5949-51.
49. Will, C.L. and R. Luhrmann, *Spliceosomal UsnRNP biogenesis, structure, and function*. *Curr. Opin. Cell Biol.*, 2001(13): p. 290-301.
50. Frey, M.R. and A.G. Matera, *Coiled bodies contain U7 small nuclear RNA and associate with specific DNA sequences in interphase human cells*. *Proc Natl Acad Sci U S A*, 1995. **92**(13): p. 5915-9.
51. Gornemann, J., K.M. Kotovic, K. Hujer, and K.M. Neugebauer, *Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex*. *Mol Cell*, 2005. **19**(1): p. 53-63.
52. Ohno, M., A. Segref, A. Bachi, M. Wilm, and L.W. Mattaj, *PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation*. *Cell*, 2000. **101**(2): p. 187-98.
53. Askjaer, P., A. Bachi, M. Wilm, F.R. Bischoff, D. L. Weeks, V. Ogniewski, M. Ohno, C. Niehrs, K. Kjems, I.W. Mattaj, and M. Fornerod, *RanGTP-regulated*

- interactions of CRM1 with nucleoporins and a shuttling DEAD-box helicase.* Mol Cell Biol, 1999. **19**(9): p. 6276-85.
54. Kitao, S., A. Segref, J. Kast, M. Wilm, I.W. Mattaj, and M. Ohno, A *compartmentalized phosphorylation/dephosphorylation system that regulates U snRNA export from the nucleus.* Mol Cell Biol, 2008. **28**(1): p. 487-97.
 55. Kolb, S.J., D.J. Battle, and G. Dreyfuss, *Molecular functions of the SMN complex.* J Child Neurol, 2007. **22**(8): p. 990-4.
 56. Battle, D.J., M. Kasim, J. Yong, F. Lotti, C.K. Lau, J. Mouaikel, Z. Zhang, K. Han, L. Wan, and G. Dreyfuss, *The SMN complex: an assembly machine for RNPs.* Cold Spring Harb Symp Quant Biol, 2006. **71**: p. 313-20.
 57. Stark, H., P. Dube, R. Luhrmann, and B. Kastner, *Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle.* Nature, 2001. **409**(6819): p. 539-42.
 58. Mouaikel, J., C. Verheggen, E. Bertrand, J. Tazi, and R. Bordonne, *Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus.* Mol Cell, 2002. **9**(4): p. 891-901.
 59. Plessel, G., U. Fischer, and R. Luhrmann, *m3G cap hypermethylation of U1 small nuclear ribonucleoprotein (snRNP) in vitro: evidence that the U1 small nuclear RNA-(guanosine-N2)-methyltransferase is a non-snRNP cytoplasmic protein that requires a binding site on the Sm core domain.* Mol Cell Biol, 1994. **14**(6): p. 4160-72.

60. Huang, Q. and T. Pederson, *A human U2 RNA mutant stalled in 3' end processing is impaired in nuclear import*. Nucleic Acids Res, 1999. **27**(4): p. 1025-31.
61. Kleinschmidt, A.M. and T. Pederson, *Accurate and efficient 3' processing of U2 small nuclear RNA precursor in a fractionated cytoplasmic extract*. Mol Cell Biol, 1987. **7**(9): p. 3131-7.
62. Darzacq, X., B.E. Jady, C. Verheggen, A.M. Kiss, E. Bertrand, and T. Kiss, *Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs*. Embo J, 2002. **21**(11): p. 2746-56.
63. Jady, B.E., X. Darzacq, K.E. Tucker, A.G. Matera, E. Bertrand, and T. Kiss, *Modification of Sm small nuclear RNAs occurs in the nucleoplasmic Cajal body following import from the cytoplasm*. Embo J, 2003. **22**(8): p. 1878-88.
64. Zhao, R., M.S. Bodnar, and D.L. Spector, *Nuclear neighborhoods and gene expression*. Curr Opin Genet Dev, 2009. **19**(2): p. 172-9.
65. Nizami, Z., S. Deryusheva, and J.G. Gall, *The Cajal body and histone locus body*. Cold Spring Harb Perspect Biol, 2010. **2**(7): p. a000653.
66. Handwerger, K.E. and J.G. Gall, *Subnuclear organelles: new insights into form and function*. Trends Cell Biol, 2006. **16**(1): p. 19-26.
67. Kiss, A.M., B.E. Jady, X. Darzacq, C. Verheggen, E. Bertrand, and T. Kiss, *A Cajal body-specific pseudouridylation guide RNA is composed of two box H/ACA snoRNA-like domains*. Nucleic Acids Res, 2002. **30**(21): p. 4643-9.
68. Jacobs, E.Y., M.R. Frey, W. Wu, T.C. Ingledue, T.C. Gebuhr, L. Gao, W.F. Marzluff, and A.G. Matera, *Coiled bodies preferentially associate with U4,*

- U11, and U12 small nuclear RNA genes in interphase HeLa cells but not with U6 and U7 genes.* Mol Biol Cell, 1999. **10**(5): p. 1653-63.
69. Frey, M.R. and A.G. Matera, *RNA-mediated interaction of Cajal bodies and U2 snRNA genes.* J Cell Biol, 2001. **154**(3): p. 499-509.
70. Dundr, M., J.K. Ospina, M.H. Sung, S. John, M. Upender, T. Reid, G.L. Hager, and A.G. Matera, *Actin-dependent intranuclear repositioning of an active gene locus in vivo.* J Cell Biol, 2007. **179**(6): p. 1095-103.
71. Suzuki, T., H. Izumi, and M. Ohno, *Cajal body surveillance of U snRNA export complex assembly.* J Cell Biol, 2010. **190**(4): p. 603-12.
72. Jady, B.E., P. Richard, E. Bertrand, and T. Kiss, *Cell cycle-dependent recruitment of telomerase RNA and Cajal bodies to human telomeres.* Mol Biol Cell, 2006. **17**(2): p. 944-54.
73. Tomlinson, R.L., E.B. Abren, T. Ziegler, H. Ly, C.M. Connter, R.M. Terns, and M.P. Terns, *Telomerase reverse transcriptase is required for the localization of telomerase RNA to cajal bodies and telomeres in human cancer cells.* Mol Biol Cell, 2008. **19**(9): p. 3793-800.
74. Tomlinson, R.L., T.D. Ziegler, T. Supakorudej, R.M. Terns, and M.P. Terns, *Cell cycle-regulated trafficking of human telomerase to telomeres.* Mol Biol Cell, 2006. **17**(2): p. 955-65.
75. Bongiorno-Borbone, L., A. De Cola, P. Vemole, L. Finos, D. Barcaroli, R.A. Knight, G. Melino, and V. De Laurenzi, *FLASH and NPAT positive but not Coilin positive Cajal Bodies correlate with cell ploidy.* Cell cycle, 2008. **7**(15): p. 2357-67.

76. Ghule, P.N., Z. Dominski, X.C. Yang, W.F. Marzluff, K.A. Becker, J.W. Harper, J.B. Lian, J.L. Stein, A.J. van Wijnen, and G.S. Stein, *Staged assembly of histone gene expression machinery at subnuclear foci in the abbreviated cell cycle of human embryonic stem cells*. Proc Natl Acad Sci U S A, 2008. **105**(44): p. 16964-9.
77. Liu, J.L., C. Murphy, M. Buszczak, S. Clatterbuck, B. Goodman, and J.G. Gall, *The Drosophila melanogaster Cajal body*. J Cell Biol, 2006. **172**(6): p. 875-84.
78. Dundr, M. and T. Misteli, *Biogenesis of nuclear bodies*. Cold Spring Harb Perspect Biol, 2010. **2**(12): p. a000711.
79. Ingelfinger, D., D.J. Arndt-Jovin, R. Luhrmann, and T. Achsel, *The human LSm1-7 proteins colocalize with the mRNA-degrading enzymes Dcp1/2 and Xrn1 in distinct cytoplasmic foci*. RNA, 2002. **8**(12): p. 1489-501.
80. Sheth, U. and R. Parker, *Decapping and decay of messenger RNA occur in cytoplasmic processing bodies*. Science, 2003. **300**(5620): p. 805-8.
81. Stoecklin, G., T. Mayo, and P. Anderson, *ARE-mRNA degradation requires the 5'-3' decay pathway*. EMBO Rep, 2006. **7**(1): p. 72-7.
82. Fukuhara, N., J. Ebert, L. Unterholzner, D. Lindner, E. Izaurralde, and E. Conti, *SMG7 is a 14-3-3-like adaptor in the nonsense-mediated mRNA decay pathway*. Mol Cell, 2005. **17**(4): p. 537-47.
83. Sheth, U. and R. Parker, *Targeting of aberrant mRNAs to cytoplasmic processing bodies*. Cell, 2006. **125**(6): p. 1095-109.

84. Unterholzner, L. and E. Izaurralde, *SMG7 acts as a molecular link between mRNA surveillance and mRNA decay*. Mol Cell, 2004. **16**(4): p. 587-96.
85. Lykke-Andersen, J. and E. Wagner, *Recruitment and activation of mRNA decay enzymes by two ARE-mediated decay activation domains in the proteins TTP and BRF-1*. Genes Dev, 2005. **19**(3): p. 351-61.
86. Liu, J., F.V. Rivas, J. Wohlschlegel, J.R. Yates III, R. Parker, and G.J. Hannon, *A role for the P-body component GW182 in microRNA function*. Nat Cell Biol, 2005. **7**(12): p. 1261-6.
87. Liu, J., M.A. Valencia-Sanchez, G.J. Hannon, and R. Parker, *MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies*. Nat Cell Biol, 2005. **7**(7): p. 719-23.
88. Sen, G.L. and H.M. Blau, *Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies*. Nat Cell Biol, 2005. **7**(6): p. 633-6.
89. Andrei, M.A., D. Ingelfinger, R. Heintzmann, T. Achsel, R. Rivera-Pomar, and R. Luhrmann, *A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies*. RNA, 2005. **11**(5): p. 717-27.
90. Ferraiuolo, M.A., S. Basak, J. Dostie, E.L. Murray, D.R. Schoenberg, and N. Sonenberg, *A role for the eIF4E-binding protein 4E-T in P-body formation and mRNA decay*. J Cell Biol, 2005. **170**(6): p. 913-24.
91. Yedavalli, V.S., C. Neuveut, Y.H. Chi, L. Kleiman, and K.T. Jeang, *Requirement of DDX3 DEAD box RNA helicase for HIV-1 Rev-RRE export function*. Cell, 2004. **119**(3): p. 381-92.

92. Beckham, C.J. and R. Parker, *P bodies, stress granules, and viral life cycles*. Cell Host Microbe, 2008. **3**(4): p. 206-12.
93. Balagopal, V. and R. Parker, *Polysomes, P bodies and stress granules: states and fates of eukaryotic mRNAs*. Curr Opin Cell Biol, 2009. **21**(3): p. 403-8.
94. Eulalio, A., I. Behm-Ansmant, D. Schweizer, and E. Izaurralde, *P-body formation is a consequence, not the cause, of RNA-mediated gene silencing*. Mol Cell Biol, 2007. **27**(11): p. 3970-81.
95. Decker, C.J., D. Teixeira, and R. Parker, *Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in Saccharomyces cerevisiae*. J Cell Biol, 2007. **179**(3): p. 437-49.
96. Savas, J.N., A. Makusky, S. Ottosen, D. Baillat, F. Then, D. Krainc, R. Shiekhattar, S.P. Markey, and N. Tanese, *Huntington's disease protein contributes to RNA-mediated gene silencing through association with Argonaute and P bodies*. Proc Natl Acad Sci U S A, 2008. **105**(31): p. 10820-5.
97. Anderson, P. and N. Kedersha, *Stress granules: the Tao of RNA triage*. Trends Biochem Sci, 2008. **33**(3): p. 141-50.
98. Gilks, N., N. Kedersha, M. Ayodele, L. Shen, G. Stoecklin, L.M. Dember, and P. Anderson, *Stress granule assembly is mediated by prion-like aggregation of TIA-1*. Mol Biol Cell, 2004. **15**(12): p. 5383-98.
99. Tourriere, H., K. Chebli, L. Zekri, B. Courselaud, J.M. Blanchard, E. Bertrand, and J. Tazi, *The RasGAP-associated endoribonuclease G3BP assembles stress granules*. J Cell Biol, 2003. **160**(6): p. 823-31.

100. Liu, J.L. and J.G. Gall, *U bodies are cytoplasmic structures that contain uridine-rich small nuclear ribonucleoproteins and associate with P bodies.* Proc Natl Acad Sci U S A, 2007. **104**(28): p. 11655-9.
101. Akbari, O.S., D. Oliver, K. Eyer, and C.Y. Pai, *An Entry/Gateway cloning system for general expression of genes with molecular tags in Drosophila melanogaster.* BMC Cell Biol, 2009. **10**: p. 8.
102. Reijns, M.A., R.D. Alexander, M.P. Spiller, and J.D. Beggs, *A role for Q/N-rich aggregation-prone regions in P-body localization.* J Cell Sci, 2008. **121**(Pt 15): p. 2463-72.

Vita

Sarah Beth May was born on November 16, 1984 in Waco, Texas to Jeffrey and Margaret May. She grew up in Mexia, Texas, and in May of 2003 she graduated Valedictorian from Mexia High School. In August of 2003 she enrolled at The University of Texas at Arlington in Arlington, Texas, and graduated from that institution *Magna cum Laude* with a Bachelors of Science degree in microbiology in August of 2006. In August of 2007 she enrolled in the Graduate School of Biomedical Sciences at The University of Texas Health Science Center in Houston, Texas.