Identification and Characterization of Novel snRNA Processing Regulators in *C. elegans*

A Thesis Submitted to the College of

Graduate and Postdoctoral Studies

In Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy

In the Department of Veterinary Biomedical Sciences

Western College of Veterinary Medicine

University of Saskatchewan

Saskatoon

By Brandon M. Waddell

Permission to use and disclaimer statement

In presenting this thesis/dissertation in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for scrutiny. I further agree that permission for copying this thesis/dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis/dissertation work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis/dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use, which may be made of any material in my thesis/dissertation.

The company/corporation/brand name and website were exclusively created to meet the thesis and/or exhibition requirements for the degree of Postgraduate studies at the University of Saskatchewan. Reference in this thesis/dissertation to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement, recommendation, or favoring by the University of Saskatchewan. The views and opinions of the author expressed herein do not state or reflect those of the University of Saskatchewan and shall not be used for advertising or product endorsement purposes.

Requests for permission to copy or to make other uses of materials in this thesis/dissertation in whole or part should be addressed to:

Head of the Department of Veterinary Biomedical Sciences

Western College of Veterinary Medicine

University of Saskatchewan, S7N 5B4, Canada.

OR

Dean College of Graduate and Postdoctoral Studies

University of Saskatchewan

116 Thorvaldson Building, 110 Science Place

Saskatoon, Saskatchewan S7N 5C9 Canada

Abstract

Splicing of messenger RNA is a hallmark of eukaryotic cells. Small nuclear ribonucleic acid (snRNA) molecules are critical for precise splicing as they recognize and target the spliceosome to specific sequences of precursor mRNA that require splicing. The Integrator complex has been implicated in the processing and maturation of these snRNA molecules. Perturbations to splicing can have deleterious consequences, leading to various human diseases including cancer. Using the model organism *C. elegans* I identified several novel snRNA processing regulators. In Chapter 2, using a reverse genetic screen I found that the Argonaute, CSR-1, plays a pivotal role in snRNA processing. Loss of CSR-1 caused increased levels of snRNA misprocessing that was dependent on the catalytic activity of the CSR-1b isoform. RNA-seq studies revealed the transcriptome is altered similarly during both CSR-1 or INTS-4 knockdown, indicating these proteins affect similar biological pathways. Additionally, members of the nuclear pore complex NPP-1, NPP-3, and NPP-6 were also confirmed to cause snRNA misprocessing through qPCR analysis. In Chapter 3, I isolated additional snRNA processing regulators through a forward genetic screen and determined to which degree they cause misprocessing. Single nucleotide polymorphism (SNP) mapping identified mutations in *rde-11, rde-1, sago-2,* and *rsd-2*, all components of the RNAi pathway, causing snRNA misprocessing. Complementation analysis confirmed the identified mutations were causing the misprocessing defect. Finally, in Chapter 4, I examined the effects these novel regulators have on maintaining a normal lifespan. The catalytic activity of CSR-1b was found to be essential in promoting a healthy lifespan. Depletion of any of the RNAi components identified in Chapter 2 caused decreases in average lifespan, as did the majority of nuclear pore proteins. Overall these results highlight the importance of transcriptome regulation and integrity in maintaining a healthy lifespan.

Acknowledgments

I would like to extend my sincere thanks and gratitude to my supervisor Dr. Michael Wu for his unconditional support, encouragement and help that he provided throughout my research program. I would also like to extend my sincere thanks to my graduate chairs during my graduate study which included Drs. Daniel McPhee, Suraj Unniappan, Maud Ferrari, and my committee members Drs. Tony Ruzzini, Adelaine Leung, and Troy Harkness for their constant support and valuable input and advice during my Ph.D. research studies. I wish to express my gratitude and thanks to all my fellow lab members. Finally, I would like to thank my parents and family for their unending support throughout my degree.

I also extend my sincere thanks to Cindy Pollard (VBMS Graduate Program Coordinator), Cheryl Hack (Administrative Assistant), and Stacey Trainor for their help and support for the successful completion of my program of study. The generous funding for this research was supported by the Discovery Grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada to M.W. Infrastructure support was provided by the John Evans Leaders Fund from the Canada Foundation for Innovation, and an Establishment grant from the Saskatchewan Health Research Foundation (SHRF) to MW. The facilities and research infrastructure provided by the Department of Veterinary Biomedical Sciences and the University of Saskatchewan are deeply appreciated.

Permission to reproduce

A substantial part (data Chapter 2) of this thesis has been published in a peer-reviewed journal. All published parts of the thesis are distributed under the terms of the Creative Commons CC BY license, which permits unrestricted use, distribution, and reproduction in any medium provided the original work is properly cited. The copyright rules of the Public Library of Science (PLoS Genetics; Chapter 2), allows reproduction of published work in the author's thesis. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Table of Contents

List of Figures

List of Abbreviations

22G-RNA: 22 nt long small interfering RNA

26G-RNA: 26 nt long small interfering RNA

ADH: *csr-1* catalytic inactive mutant with an ADH motif in place of the DDH motif

AGO: Argonaute

AID: auxin-inducible degradation

AS: alternative splicing

C. elegans: *Caenorhabditis elegans*

cDNA: complementary deoxyribonucleic acid

CSR-1: Chromosome-Segregation and RNAi deficient

Co-IP: co-immunoprecipitation

CRISPR: clustered regularly interspaced short palindromic repeats

CTCF: corrected total cell fluorescence

CTDP1: Carboxy-Termina Domain Phosphatase 1

DICE1: Deleted In Cancer 1

DNA: deoxyribonucleic acid

dsRNA: double-stranded ribonucleic acid

eGFP: enhanced GFP

EMS: ethylmethane sulfonate

ETOH: ethanol

EV: empty vector

FDR: false discovery rate

GCFC: GC-rich sequence DNA-binding FaCtor homolog

GFP: green fluorescent protein

GOI: gene of interest

GRO-seq: global run on sequencing

HDR: homology-dependent repair

IAA: indole-3-acetic acid

IPTG: isopropyl β-D-thiogalactopyranoside

KO: knock out

L1-L4: larval stages 1-4

mRNA: messenger ribonucleic acid

MosSCI: Mos1-mediated Single Copy Insertion

MUT: Mutator

NGM: nematode growth media

NPC: nuclear pore complex

NPP: nucleoporin protein

nt: nucleotide

NTR2: NineTeen complex Related protein 2

PAM: protospacer adjacent motif

PCR: polymerase chain reaction

piRNA: Piwi-interacting RNA

Piwi: P-element Induced WImpy

Pol: polymerase

qPCR: quantitative polymerase chain reaction

RAPToR: real-age prediction using transcriptome staging

RDE: RNAi DEfective

RdRP: ribonucleic acid-dependent polymerase

RFP: red fluorescent protein

Ribo-seq: ribosome sequencing

RISC: RNA-induced silencing complex

RNA Pol II: ribonucleic acid polymerase II

RNA: ribonucleic acid

RNAi: ribonucleic acid interference

RNA-seq: RNA sequencing

RRF: RNA-dependent RNA polymerase Family

RSD: RNAi Spreading Defective

SAGO: Synthetic secondary siRNA-deficient ArGOnaute

SCF: Skp, Cullin, F-box containing complex

SEC: self excising cassette

sgRNA: synthetic guide ribonucleic acid

SIN: slicing inactive

siRNA: small interfering ribonucleic acid

SLBP: stem-loop binding protein

SM: survival of motor neurons

SNP: single nucleotide polymorphism

snRNA: small nuclear ribonucleic acid

snRNP: small nuclear ribonucleoprotein

SOSS: sensor of single-stranded DNA

SS: Splice site

TIR1: transcription inhibitor response 1

UV: ultraviolet

WT: wild type

Chapter 1

General Introduction

1.1 *C. elegans* history

The free-living nematode *Caenorhabditis elegans* has long provided researchers with an attractive model system to study an array of biological processes. While Emile Maupas first described them in the late 1800s, it was their introduction to Sydney Brenner's lab in 1965 that researchers began to understand their potential ("History of research on *C. elegans* and other free-living nematodes as model organisms" 2017). Dr. Brenner's focus at the time was to understand the genetic requirements of neuronal development and function (Riddle *et al.* 1997). He was interested in a model organism with a simple reproductive system capable of generating many offspring yet was small enough to facilitate the observation of large populations (Ankeny 2001).

C. elegans encompasses all these criteria, and is often simply referred to as the 'worm'. These worms are small (1.0 to 1.5 mm in length) and have an exceedingly fast developmental period of 2 days followed by a 3-5 day reproductive cycle where they can self-reproduce ~ 300 genetically identical progeny from a single worm (Riddle *et al.* 1997). Past the reproduction period, wildtype *C. elegans* can live for another 2-3 weeks before natural death; the relatively short lifespan of these animals from birth to death has also led to the eventual use of this model organism in aging studies. A graphical description of the lifecycle can be seen in **Figure 1.1.** Populations of worms can be maintained on agar-based media seeded with *Escherichia coli*, or cryopreserved frozen at extremely low temperatures (-80°C) for long-term storage (Riddle *et al.* 1997).

The majority of *C. elegans* in a wildtype population are hermaphroditic (> 99.9%), although males do arise from chromosome non-disjunction at a rate of 0.1% (Brenner 1974). This allows for straightforward maintenance of homozygous populations but still affords a simple method of transferring genetic material between different backgrounds through mating. The complete sequence of the *C. elegans* genome was obtained in the late 1990s, which has greatly facilitated cloning strategies and comparative genomics. The genome is relatively small, approximately 100 Mb, and each worm has less than 1000 total cells (*C. elegans* Sequencing Consortium 1998). Combined with the eutely of individual worms and the complete mapping for cell lineage, it is an ideal organism to study developmental biology due to its invariant and predictable cell fate (Kimble and Nüsslein-Volhard 2022). Another useful advantage of *C. elegans* is its almost complete transparency, making it an ideal candidate for microscopy-based studies (Kaletta and Hengartner 2006).

The usefulness of *C. elegans* as a model system has been underscored by the insights into biology they have provided. The presence of the RNAi system and its effects on gene regulation were first described in the worm (Fire *et al.* 1998). Apoptosis, programmed cell death, was also initially characterized in worms (Conradt *et al.* 2016). As a final example, the cloning and expression of the green fluorescent protein (GFP) was pioneered in *C. elegans* (Zimmer 2009). These discoveries all led to Nobel prizes for their contributions to the research community. A striking observation between *C. elegans* and humans is the relative similarity between their genomes. While the human genome is estimated to contain 3.2 billion base pairs and *C. elegans* has only 100 million base pairs, ~80% of human genes have a known homolog in worms (Brown 2002; Kim *et al.* 2018). As such *C. elegans* has also provided an attractive

platform to model human diseases such as Alzheimer's and Parkinson's disease (Apfeld and Alper 2018).

Researchers have taken advantage of the nematode's simplicity and developed an ensemble of genetic tools to manipulate the worm. An essential tool is the production of transgenic strains. Worms can be injected with DNA into their reproductive gonad, which is inherited by and expressed in the resulting progeny (Nance and Frøkjær-Jensen 2019). With this tool, researchers can add fluorescent tags to proteins of interest, express mutant alleles, or overexpress genes. Transgenic injections also served as the original method of RNAi introduction, though this was streamlined by the discovery that RNAi in *C. elegans* can also be accomplished by feeding (Fire *et al.* 1998). This advancement has led to the construction of feeding RNAi libraries that can target genes of *C. elegans* in a high-throughput manner (Kamath and Ahringer 2003).

Compared to *C. elegans, Drosophila melanogaster* (fruit flies) and *Saccharomyces cerevisiae* (yeast) are also prominent model organisms. Much like worms, the fly and yeast genomes share considerable homology with the human genome, allowing for the extrapolation of their biological processes into higher-order organisms (Jennings 2011; Vanderwaeren *et al.* 2022). Yeast is relatively easy to culture under laboratory conditions and can reproduce daughter cells in as little as 90 minutes (Vanderwaeren *et al.* 2022). This makes them an excellent candidate for high-throughput screens. However, the lack of differentiated tissues and behavioural phenotypes can exclude yeast as an appropriate model for some areas of biology. Fruit flies undergo similar development programs to humans and contain several types of differentiated tissues, however, their 12-day life cycle and larger body size are not ideal for high-throughput screens (Jennings 2011). Unlike yeast and worms, flies require mating to produce

subsequent generations and need continual culturing as stocks that can not be frozen for cryopreservation (Jennings 2011). Worms combine the benefits of both flies and yeast as they have a short life cycle, contain differentiated tissues, produce behavioural phenotypes, and are amenable to long-term storage with the added benefit of transparency for microscopy studies.

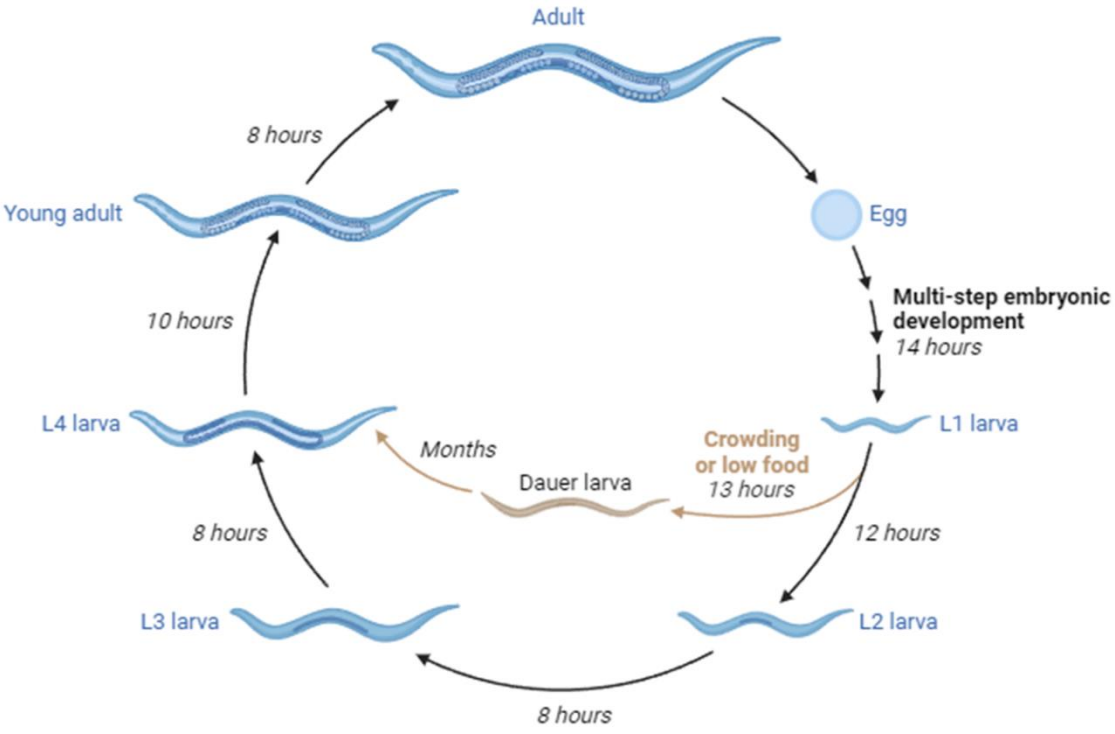


Figure 1.1: Graphical description of the *C. elegans* lifecycle at 20 °C. Eggs are laid onto the media by the mother and hatch into their first larval stage, L1. A growth period followed by cuticle molting signifies the transition into the second larval stage, L2, which is repeated until the final molt after the L4 stage. This produces adult worms capable of reproduction. During times of stress, low food, or overpopulation, *C. elegans* enters a stress-resistant developmental phase termed dauer, which involves metabolic reduction and encourages movement towards a more favourable environment. This stage can last for many months, and upon introduction to a food

source worms restart their developmental program at the L4 stage. Figure generated in BioRender©.

1.2 *C. elegans* Genetics

1.2.1 Classical Genetics

The first screen employed in *C. elegans* was a forward mutagenesis screen performed by Dr. Brenner; using the mutagen ethylmethane sulfonate (EMS), wild-type worms were mutagenized, and the resulting progeny screened for a suite of developmental or movement defects (Brenner 1974). EMS is a chemical that causes alkylation of guanine into 0^6 methyl guanine which interferes with the natural base pairing to cytosine, resulting in A/T transitions (Chen *et al.* 2023). In the *C. elegans* genome, EMS produces a null mutation to any given locus at a rate of 1 null per 2,000 copies of the gene screened (Jorgensen and Mango 2002). In this initial landmark study, Dr. Brenner was able to isolate over 600 individual mutations that produced locomotion defects as well as other physical characteristics such as underdeveloped body size and blistering of the cuticle (Brenner 1974).

Stemming from this initial screen eventually led to the idea of modifier screens. Rather than mutagenizing wildtype worms, previously generated mutants were mutagenized a second time. The resulting progeny were then screened for reduction of the mutant phenotype, also known as a suppressor screen, or for their enhancement of the mutant phenotype, an enhancer screen (Jorgensen and Mango 2002). While these experiments were powerful, it was an exorbitant amount of work to manually screen thousands of progeny for a rare phenotype. The selection screen was developed to alleviate this issue. One of the first selection screens performed was based on using drug resistance as a phenotype; in these assays, most/all wildtype

animals would die after drug exposure while mutations of interest would confer survivability of the affected progeny (Jorgensen and Mango 2002). These screens can be performed on a large-scale using *C. elegans*, often analyzing millions of individuals (Jorgensen & Mango, 2002). The ease of *C. elegans* culture offers the advantage of reaching mutagenic saturation, that is the ability to screen enough animals to cover the probability that each gene within a genome would be mutated at least once. The screen of a large number of initial worms also leads to a second major advantage to increase the probability of isolating partial loss of function or gain of function alleles (Jorgensen and Mango 2002).

With the discovery of RNAi came the idea of the reverse genetic screen. The direct effects of a genetic knockdown can be observed without the need to obtain a homozygous mutant (Apfeld and Alper 2018). While powerful on its own, the initial introduction of RNA triggers was extremely labour-intensive. The dsRNA first had to be produced and purified *ex vivo*, then directly injected into the tissue of interest (Grishok 2005). This issue was partially alleviated when it was found that injection into the intestine allowed for the induction of RNAi in other tissues. The process was streamlined when researchers found that simple ingestion of dsRNA was sufficient to cause gene knockdown. This led to the development of bacteria engineered to express dsRNA, allowing for the uptake of silencing RNA during feeding (Kamath *et al.* 2001). Subsequent development of this process produced libraries of bacteria that individually express RNAi triggers against all known genes in *C. elegans* (Kamath and Ahringer 2003). In the present day, the entire genome can be knocked down in as little as 12 weeks, providing a relatively fast and robust method to thoroughly investigate gene function on a specific phenotype. A major advantage of reverse RNAi screens over forward EMS screens is that

phenotypes associated with the knockdown of viable genes can be observed, this is particularly useful when studying essential genes where null mutations would result in lethality.

1.2.2 Contemporary genetics

In the past decade, Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) Cas-9 genome editing has become a powerful tool in many research fields. Initially used by prokaryotes as a viral defence mechanism, CRISPR-Cas9 technology has since been employed by researchers to alter DNA sequences in a site-specific manner (Jinek *et al.* 2012). During regular operation, the Cas-9 enzyme recognizes a protospacer adjacent motif (PAM), binds DNA, and uses a segment of "guide RNA" to direct double-stranded breaks in a target DNA sequence (Barrangou and Doudna 2016). This allows prokaryotic cells to destroy invading DNA-based pathogens. By manipulating the guide RNA, the Cas-9 enzyme can be directed to cleave DNA at any desired sequence. A graphical representation of CRISPR can be seen in **Figure 1.2**.

While this can efficiently cleave a target DNA sequence, there is no editing of the genome. By introducing a repair template of the desired sequence content with homology arms that flank the region of cleavage, cells can use homology-dependent repair (HDR) to repair the double-stranded break and incorporate the desired DNA sequence (Paix *et al.* 2017). While *C. elegans* does not harbour native Cas-9 enzyme, it can be supplied exogenously due to the transgenic system to express foreign DNA. This can be accomplished through the injection of a plasmid containing the Cas-9 sequence driven by a worm-specific promoter such as *eft-3p* that is ubiquitously expressed in somatic cells (Dickinson *et al.* 2013). The caveat here is that these plasmids are taken up as multi-copy extrachromosomal arrays and must act in the germline to produce edited progeny in the next generation. The germline of *C. elegans* is efficient at silencing over-expressed or multicopy arrays, which reduces the effectiveness of Cas-9

expression via multicopy arrays and the efficiency of genome editing (Schwartz *et al.* 2021). The inefficiency has been improved in the last few years by generating transgenic worms that constitutively express the Cas-9 protein in the germline through single copy insertion via Mos1-mediated Single Copy Insertion (MosSCI) (Schwartz *et al.* 2021).

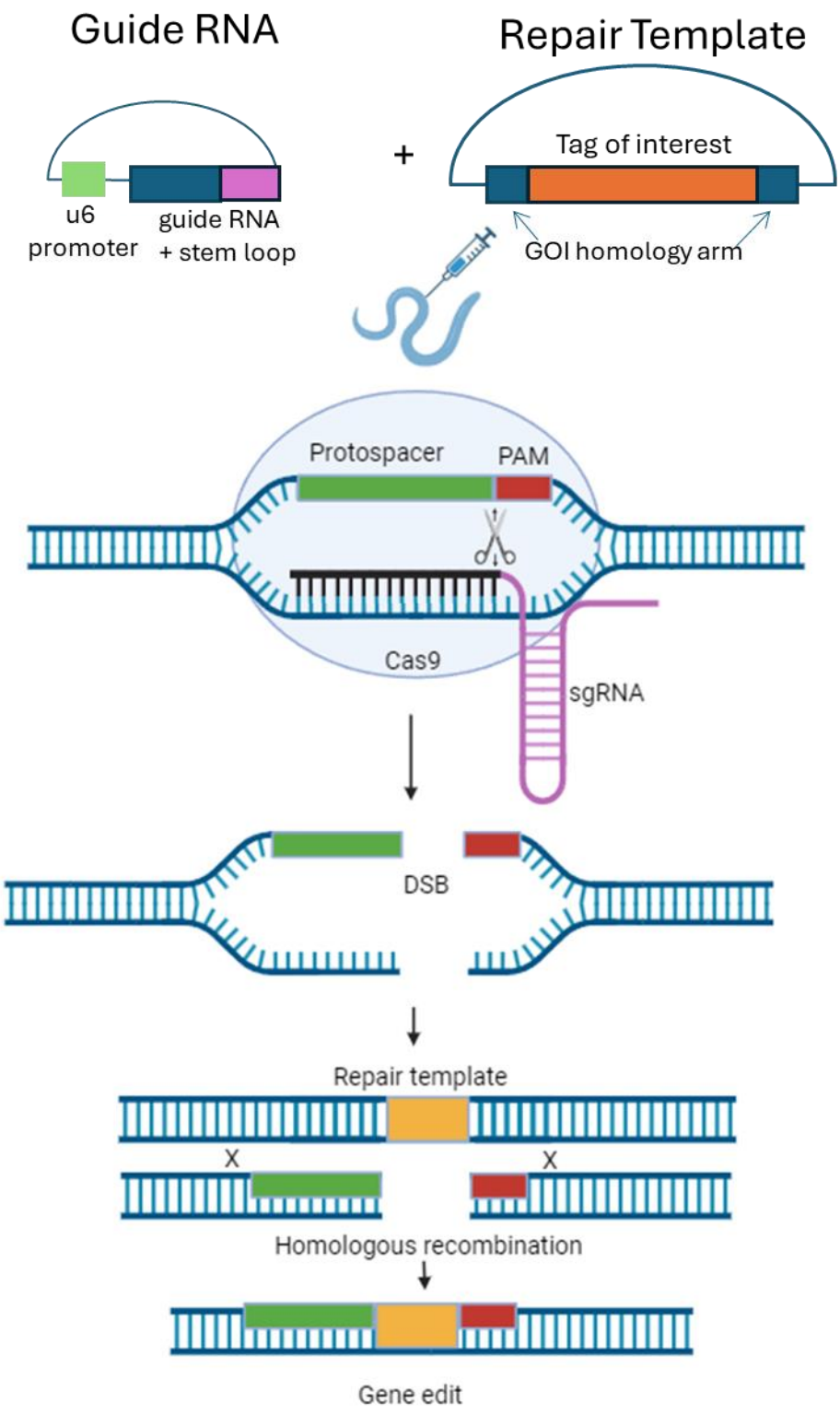


Figure 1.2: CRISPR editing mechanism. Preparations of the guide RNA and repair template are injected into reproductive worms. Expression of the injected constructs allows the guide RNA to

direct the cleavage site of Cas9 into the gene of interest, creating a double-stranded DNA break. Homologous recombination then uses the supplied repair template to ligate the sections of DNA back together in a sequence-dependent manner, incorporating the edit of interest. Figure generated in BioRender©.

1.2.3 Auxin inducible degradation

The development of the CRISPR/Cas-9 gene editing toolbox in *C. elegans* has opened up a wide range of possible edits that can be applied to the genome. One type of edit that is of interest to this thesis is the endogenous insertion of an auxin-inducible degradation element. In plants, the degradation of specific transcriptional repressors is mediated by the auxin hormone family (Nishimura *et al.* 2009). The binding of auxin to the F-box transcription inhibitor response 1 (TIR1) protein causes interaction with an SCF (Skp1, Cullin, and F-box) E3 ubiquitin ligase complex that recruits an E2 ubiquitin-conjugating enzyme to polyubiquitylate the bound repressor (Nishimura *et al.* 2009). The ubiquitinylated repressor is then degraded by the proteasome. This system has been utilized by researchers to provide robust and efficient knockdown of proteins of interest.

Many eukaryotes share the SCF pathway but lack the auxin response. By fusing an auxin-inducible degron (AID) from these auxin-responsive transcriptional repressors and providing cells with the TIR1 protein, researchers can simulate the auxin degradation pathway in virtually any biological system (Nishimura *et al.* 2009). The degradation is quick (~30 minutes), reversible, and conditional, allowing researchers to study the loss of protein function in a variety of scenarios (Ashley *et al.* 2021). This system has been recently adapted for *C. elegans* through

integrating a single-copy of the TIR1 transgene under the control of various promoters to facilitate tissue-specific protein knockdown (Zhang *et al.* 2015; Ashley *et al.* 2021). This system is supported by CRISPR/Cas-9 that allows researchers to tag their gene of interest with an AID epitope within the gene reading frame to facilitate conditional knockdown of the translated protein in a rapid and reversible manner. A schematic of the AID system can be seen in **Figure 1.3.** In this thesis, I will introduce the use of this system to construct transgenic *C. elegans* that permits rapid degradation of the Integrator complex.

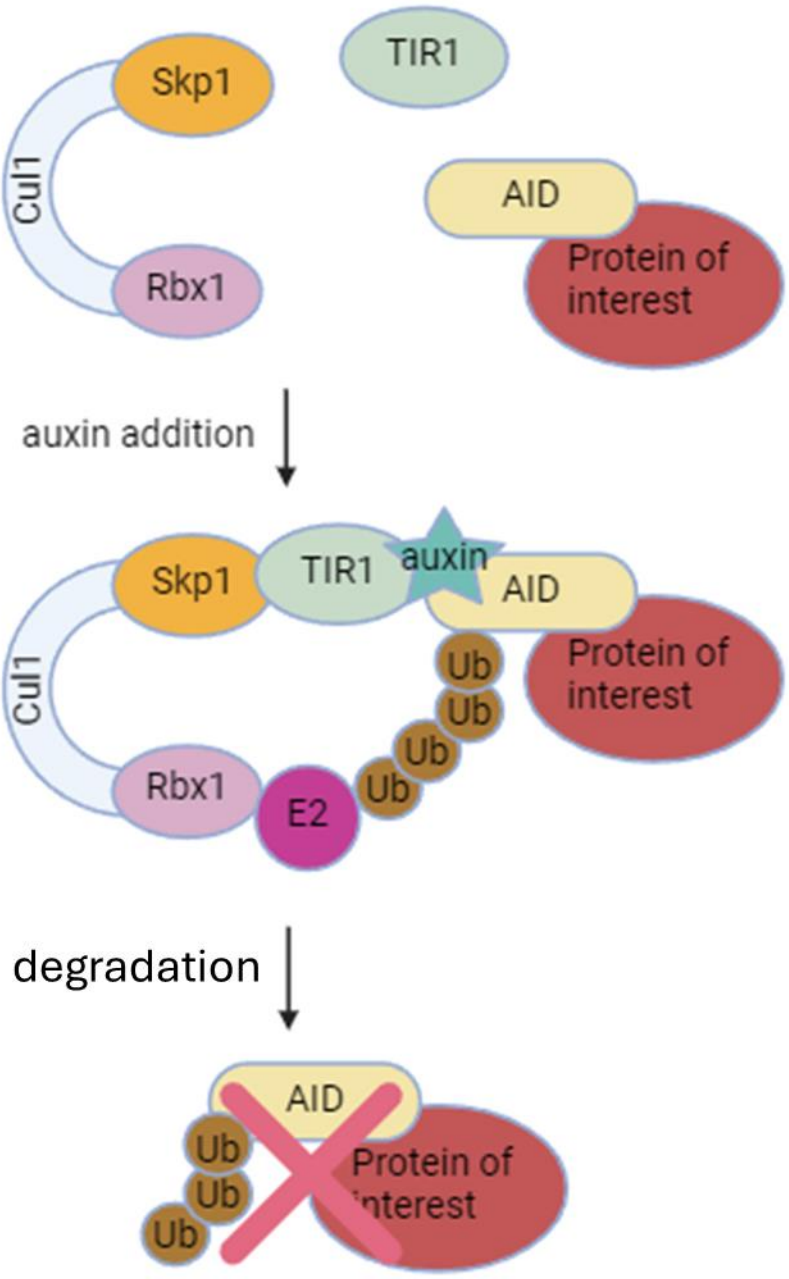


Figure 3: AID mechanism. The protein of interest is tagged with an AID motif, which can recognize and bind auxin. The binding of auxin to the AID motif is recognized by the TIR1 protein, which then recruits the conserved SCF ligase complex to the protein of interest.

Polyubiquitylation of the protein of interest then triggers its degradation by the proteasome. Figure generated in BioRender©.

1.3 RNAi silencing pathways and mechanism

1.3.1. Key RNAi regulatory proteins

A puzzling conundrum of RNAi was the robust response to limiting numbers of dsRNA triggers. Subsequent studies identified a mechanism of dsRNA signal amplification that could explain this observation. An RNA-dependent RNA polymerase (pol) (RdRP), RRF-3, was found to act in concert with the riboendonuclease DCR-1 and other cofactors to produce multiple short (26 nt) fragments of dsRNA from endogenous and exogenous sources of dsRNA (Simmer *et al.* 2002). These 26 nt species are termed 26G-RNA because of their length and their preference to being with a guanine. Argonaute proteins can then use these smaller fragments of dsRNA to target specific genes for silencing. This targeting event causes a cascade that requires another RdRP, RRF-1, to produce more short (22 nt) dsRNA moieties, characterized by their preference to begin with a guanine called 22G-RNA, that bind to additional worm-specific Argonaute proteins (WAGO) to further attenuate gene expression (Sijen *et al.* 2001). A catalytically restricted Argonaute protein, RDE-1, also functions in the production of these short (22 nt) dsRNA molecules. Involved primarily in passenger strand removal, RDE-1 can also bind several DCR-1 products to generate the 22 nt dsRNA fragment destined for WAGO binding and attenuation of expression (Tabara *et al.* 1999). A complex containing RDE-10/11 acts cooperatively with the RSD-2 protein to further amplify these 22 nt dsRNA signals (Yang *et al.* 2012). Yet another class of proteins, the Mutator proteins (MUT), also function along with RRF-

1 to produce the 22 nt dsRNA molecules to promote silencing in the germline (Zhang *et al.* 2011). These mechanisms act together to amplify the initial dsRNA trigger, be it an endogenous or exogenous source, to produce widespread gene silencing. A simplified graphical description of RNAi can be found in **Figure 1.4**.

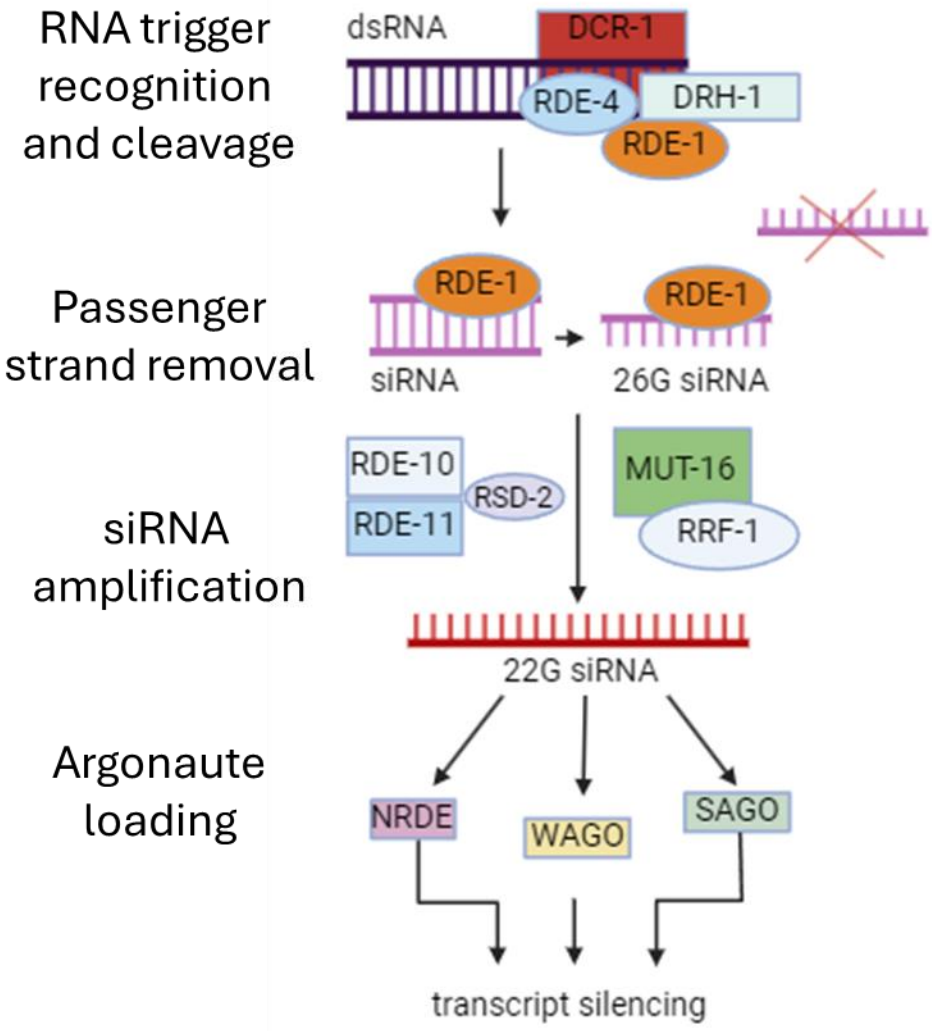


Figure 1.4: RNAi mechanism in *C. elegans*. Trigger dsRNA is recognized by DCR-1, which then cleaves the trigger dsRNA into smaller fragments which serve as a template for 26G siRNA synthesis. The double-stranded 26G siRNAs are recognized by RDE-1, which catalyzes the

removal of the passenger strand of RNA and shuttles the single targeting strand towards siRNA amplification complexes. The RDE-10/11 complex uses these 26G siRNAs to target mRNA destined for silencing, producing 22G siRNA as well as participating in mRNA transcript degradation. Similarly, Mutator foci also use these 26G siRNAs to target mRNA for 22G siRNA production via the RdRP RRF-1, although Mutator foci have not been observed in direct degradation of target mRNA. These 22G siRNAs can then bind their effector AGOs to further elicit transcript silencing. Figure generated in BioRender©.

1.3.2 Transgene silencing

Many studies in *C. elegans* make use of transgenic arrays to assess the function of various genes. The injected DNA sequences form repetitive arrays once inside the germline of the organism and become inherited in the next generation of progeny as an extrachromosomal array. These transgenes can be effectively expressed in the soma of the receiving progeny, however, the germline possesses a surveillance mechanism that is accurate and efficient at silencing these repetitive sequences (Kelly and Fire 1998). This process is thought to involve components of the RNAi machinery, as some mutants defective for RNAi are also defective in germline silencing (Schaner 2006). During meiosis, unpaired DNA elements are targeted by components of the RNAi pathway to receive histone 3 methylation of lysine 9, a chromatin mark associated with repressed transcription. So evidence exists to indicate that RNAi is involved with transcriptional gene silencing as well as post-transcriptional gene silencing in *C. elegans*. Silencing of a random transgene through feeding RNAi vectors was found to be dependent on the Argonaute protein RDE-1, the dsRNA binding protein RDE-4, the Dicer homolog DCR-1, the RdRP RRF-1, another Argonaute protein ALG-1 and a homolog of heterochromatin protein 1

HPL-2 (Grishok *et al.* 2005). Several auxiliary factors involved in the RNAi response have also been shown to participate in this somatic silencing, notably MUT-16 and RSD-6.

1.3.3 CSR-1 in *C. elegans*

Chromosome-Segregation and RNAi deficient (CSR-1) is an Argonaute protein that was originally identified for its role in chromosome segregation and involved in the expression of germline-specific genes (Claycomb *et al.* 2009). Contrary to other Argonaute proteins that possess cleavage activity, it was discovered that CSR-1 acts to promote germline gene expression by protecting targeted transcripts from degradation (Wedeles *et al.* 2013). CSR-1 has also been implicated in promoting transgene expression where it was found that loss of CSR-1 can lead to the silencing of transgenes (Fischer *et al.* 2013). In this mechanism, siRNAs direct CSR-1 to target transcripts, where they are protected from the actions of the Argonaute NRDE-3. NRDE-3 directs trimethylation of H3K9, a marker of silent chromatin, whereas CSR-1 may direct trimethylation of H3K4 associated with active chromatin (Fischer *et al.* 2013).

Aside from these functions, CSR-1 has also been shown to positively regulate the cleavage of replication-dependent histone mRNA, which are the only known transcripts that are processed via cleavage in an analogous manner to snRNAs (Avgousti *et al.* 2012). The histone downstream element and U7 snRNA normally involved in processing these mRNAs are absent in *C. elegans*, creating a need for new components to manage this processing function. Some of the siRNAs bound by CSR-1 are complementary to histone transcripts near its 3' cleavage site, and it has been proposed that loss of CSR-1 leads to misprocessing of these mRNAs (Avgousti *et al.* 2012). Whether the cleavage activity of CSR-1 is required for this event is still unknown. Together, these studies suggest that CSR-1 is a central player in regulating various aspects of

small RNA biology. In this thesis, we add to the growing list of CSR-1 functions by implicating a role for this protein in the regulation of germline small nuclear RNA processing.

1.3.4 P granules

Germ granules are found in many phyla across the animal kingdom (Price *et al.* 2021). In *Drosophila,* mice, and *C. elegans*, they have been called nuage, chromatoid bodies, and P granules, respectively. They are membrane-less organelles composed of a variety of proteins and RNA that are associated with the nuclear periphery in germ cells (Seydoux 2018). Germ granules serve as hot spots for post-transcriptional gene regulation and as such play a pivotal role in organismal development (Sundby *et al.* 2021). In *C. elegans*, P granules are present throughout the life cycle of the worm, and the loss of these granules in developing larvae results in sterile adults (Sundby *et al.* 2021). During the majority of its life cycle, P granules associate closely with the nuclear periphery. However, during oocyte maturation, these granules diffuse into the cytoplasm via a poorly understood mechanism before re-localizing to the germ line precursor cells.

A major function of the P granule is the maintenance of the germline transcriptome (Price *et al.* 2021). This is accomplished through a pool of small endogenous RNA molecules and associated Argonaute proteins that can distinguish germline transcripts from those of foreign origin. piRNA interacts with the PIWI-related Argonaute PRG-1 to silence deleterious foreign transcripts whereas actual germline transcripts are recognized by a different subset of small RNA molecules that associate with CSR-1, which protects these authentic transcripts (Sundby *et al.* 2021). In this way foreign sequences and incorrectly expressed somatic genes can be eliminated before their expression in the germ cell, allowing for the proper development of the germ cell and successful generation of progeny.

1.4 RNA Splicing

1.4.1 Mechanisms of RNA splicing

The central dogma of biology follows that DNA is transcribed into RNA which is translated into protein. Most transcripts of the human genome require the removal of introns to produce a mature transcript (Will and Lührmann 2011). This function is performed by the spliceosome, a multiprotein complex that also contains unique RNA molecules (Will and Lührmann 2011). Two different spliceosome complexes exist, the major spliceosome is responsible for the bulk of splicing events and the minor spliceosome acts only on a select few transcripts. Both the major and minor spliceosomes are multiprotein complexes; however, they differ in their RNA content. The major spliceosome contains U1, U2, U5, and U4/6 small nuclear RNA (snRNA) particles while the minor spliceosome contains U11, U12, U15, and U4atac/U6atac (Wilkinson *et al.* 2020). In *C. elegans,* the function of the major spliceosome is evolutionarily conserved, however, the minor spliceosome appears to be lost through evolution (Bartschat and Samuelsson 2010).

The sequences in the pre-mRNA are critical to proper splicing, which include the 5`-splice site (SS), the 3`SS, and the branch point that serves to direct the removal of the intronic sequence and the ligation of the exonic sequences (Burge *et al.* 1998). The spliceosome assembly begins at the 5` SS where the U1 small nuclear riboprotein (snRNP) binds to the 5` SS, followed by U2 snRNP binding to the branch site. The remaining snRNPs U4/6 along with U5 associate with the complex, where changes in the RNA-RNA and protein-RNA interactions produce an active spliceosome (Brow 2002). Following catalytic activation by the RNA helicase Prp2, the spliceosome is now poised to remove introns and ligate exons together in a 2-step reaction (Will and Lührmann 2011). A key component that facilitates RNA binding of the spliceosome is

snRNA which is transcribed by RNA pol II. However, before incorporation into the spliceosome, the snRNA molecule must be modified by cleavage of its 3' tail upstream a motif termed the 3' box (Ezzeddine *et al.* 2010). During transcription of the snRNA, a protein complex called Integrator associates with RNA pol II and cleaves the newly forming RNA molecule to form mature snRNA (Baillat *et al.* 2005). The snRNA molecule then interacts with an snRNA-specific export protein called PHAX in the nucleus for export into the cytoplasm before further assembly with its protein counterparts (Ohno *et al.* 2000). In the cytoplasm, association with survival of motor neurons (SM) constitutes the mature snRNP molecule, which in turn signals for the hypermethylation of the 5`-guanosine cap that directs the snRNP back to the nucleus to form the active spliceosome (Kiss 2004).

The core function of splicing for intron removal and exon joining produces a mature mRNA transcript ready for translation. In some instances, certain exons are retained or excluded during splicing to generate unique transcripts in a process termed alternative splicing. Many forms of alternative splicing can occur to produce transcript variants with mutually excluded exons, intron retention, exon exclusion, or alternative 5' and 3' SS (Nilsen and Graveley 2010). The choices made for which splicing form is used are determined by the binding of protein factors to enhancer and silencer sequences proximal to the transcript initiation site in the genome. By creating new combinations of exons within a single gene, genomic complexity can be greatly increased with minimal addition to the genome (Nilsen and Graveley 2010).

1.5 The Integrator complex

The Integrator complex is conserved across metazoan and was discovered in 2005 as the elusive protein machinery that catalyzes the 3' cleavage and subsequent maturation of snRNA transcript post-transcription (Baillat *et al.* 2005). Loss of function to the Integrator complex has

been shown to lead to aberrant RNA splicing, presumably caused by the defect in snRNA processing which is required to form a functional spliceosome (Gómez-Orte *et al.* 2019). In most organisms, mutations to the Integrator subunits cause lethality, highlighting the biological importance of this complex. The Integrator complex is a large molecular weight protein complex with at least 14 subunits identified to date that associate into modules that play different roles in various biological processes, which are described below (Wagner *et al.* 2023).

1.5.1 Integrator Core Module

The Integrator complex is a large macromolecular complex initially described to be comprised of 11 individual subunits. Subsequent proteomic and biochemical analyses have identified 3 additional Integrator subunits (Wagner *et al.* 2023). These subunits have been found to associate into separate modules that come together to make up the functional Integrator complex. The core of the complex consists of a backbone and shoulder module. INTS1, the largest Integrator subunit, along with INTS2 and 7 make up the backbone of the complex. The shoulder module comprises an INTS5 and 8 heterodimer and interacts with the C terminal repeats of INTS1 and 2 (Wagner et al., 2023). Although it has yet to be visualized in a crystallographic structure, yeast 2 hybrid studies indicate that INTS12 also interacts with the backbone and shoulder modules to function as a scaffold for the additional Integrator subunits (Chen *et al.* 2013).

1.5.2 Integrator Cleavage Module

The Integrator cleavage module is responsible for the endonuclease activity of the complex that includes INTS9 and INTS11 which are paralogues of the cleavage and polyadenylation specificity factors CPSF100 and CPSF73, respectively (Albrecht *et al.* 2018). These Integrator subunits form a heterodimer and interact with INTS4 through their C terminal

domains, which can then tether the cleavage module to the backbone. Compared to subunits of the core module, the knockdown of any one of the cleavage module components results in the largest induction of snRNA misprocessing, underscoring the importance of this module in Integrator function (Albrecht *et al.* 2018).

1.5.3 Integrator Phosphatase Module

Aside from the canonical snRNA processing functions of Integrator, more recent studies have revealed that Integrator also regulates the attenuation of promoter proximally paused genes (Baillat and Wagner 2015). This function has been attributed to components of the phosphatase module of Integrator which include INTS6 that interacts with protein phosphatase 2 A and C (PP2A and PP2C) (Stein *et al.* 2022). While INTS6 seems critical for PP2C recruitment, INTS8 of the shoulder module interacts with PP2A (Huang *et al.* 2020). Unlike most PP2A complexes, the Integrator-PP2A complex is missing the canonical B regulatory subunit (Shi 2009; Wagner *et al.* 2023). It has been proposed that Integrator fulfills this regulatory role and targets Integrator-PP2A to nascent transcripts. By targeting PP2A to transcriptionally paused genes, the C terminal domain of RNA pol II can be dephosphorylated by PP2A. This prevents the binding of positive elongation factors and allows for the cleavage module to release the incomplete transcript from the pol thereby terminating transcription (Huang *et al.* 2020). Although it has yet to be seen in any structural studies, INTS6 has the potential to bind INTS3. To date, INTS3 has not been described as a *bona fide* component of the phosphatase module and instead has reported functions in the sensor of single-stranded DNA (SOSS) complex (Welsh and Gardini 2023).

1.5.4 Integrator Auxiliary Module

The auxiliary module is the least characterized module of the Integrator complex, but recent studies have begun to shed light on its importance. INTS13 and 14 come together in a

heterodimer and form a stable complex with INTS10 (Sabath *et al.* 2020). This complex shares homology with the Ku7-Ku80 proteins involved in double-stranded break repair (Walker *et al.* 2001). Unlike Ku70 and Ku80, which have an affinity for DNA, the auxiliary module prefers RNA hairpin structures and interacts with the cleavage module through INTS13, where it is hypothesized to target Integrator to snRNA targets (Sabath *et al.* 2020). Interestingly, knockdown experiments suggest that the auxiliary module is not required for this function but is thought to increase the efficiency of Integrator-mediated snRNA cleavage, as mutations to the auxiliary module only mildly affect snRNA processing (Wagner *et al.* 2023).

While the C terminal domain of INTS13 is responsible for the association with the cleavage module required for efficient snRNA cleavage, INTS13 also seems to play a role in spermatogenesis in flies. In mutant flies, mutation to this region of INTS13 causes a perturbation in dynein-dynactin recruitment to the nuclear envelope resulting in sterile offspring (Sabath *et al.* 2020). Finally, this module also has proposed roles in the release of paused transcripts where mutations to INTS14 caused increased rates of elongation when compared to mutation of the catalytic subunit INTS11, suggesting that the auxiliary module plays a larger role in transcript termination than the cleavage module (Sabath *et al.* 2020). While it has yet to be visualized in crystallographic structures, studies indicate that INTS10 and INTS15 interact with one another.

1.5.5 Integrator and human diseases

In recent years, direct links between mutation to subunits of the Integrator complex and human disease have slowly emerged. An initial study examining patients suffering from severe neurodevelopmental syndromes revealed mutations to two distinct Integrator subunits, 1 and 8 (Oegema *et al.* 2017). Patients with mutations in INTS1 and INTS8 exhibit neurodevelopmental defects characterized by cognitive delay, motor impairment, and congenital face abnormalities

(Oegema *et al.* 2017). Cultured fibroblast cells derived from patients with mutations to INTS1 and INTS8 both show an increase in misprocessing of snRNA transcripts and splicing defects. These observations are consistent with the role Integrator has on RNA splicing via its function in snRNA processing (Oegema *et al.* 2017). In a separate study, biallelic mutations to INTS1 have been observed to cause rare neurodegenerative disease that was characterized by growth defects, intellectual disability, and cataracts (Zhang *et al.* 2020). Mechanistically, it was determined that INTS1 associates with CTDP1 (Carboxy-Termina Domain Phosphatase 1) which when mutated on its own results in congenital cataracts-facial dysmorphism-neuropathy that is symptomatically similar to patients with INTS1 mutation alone. CTDP1 primarily regulates RNA pol II to regulate transcript elongation; this suggests that the cellular defects associated with INTS1 may be linked to the RNA pol II pause release function of the Integrator.

Most recently, biallelic mutations to INTS11, which is part of the catalytic core, were reported in patients with neurodevelopmental syndromes as a result of impaired neurogenesis and/or neurodegenerative processes (Tepe *et al.* 2023). Interestingly, when these mutated amino acids in INTS11 were introduced into the *Drosophila* INTS11, it led to a significant reduction in lifespan. This implicates a role for INTS11 mutation to cause accelerated aging in addition to neurodevelopmental defects (Tepe *et al.* 2023). Mutations in INTS13 have been shown to disrupt ciliogenesis in humans, resulting in oral-facial-digital syndromes (Mascibroda *et al.* 2022).

Next to neurodevelopmental defects, mutations to Integrator have also been associated with cancer. INTS6 was initially identified as Deleted In Cancer 1 (DICE1) as it was frequently downregulated in lung carcinomas (Wieland *et al.* 1999). INTS3 has been found to be overexpressed in hepatocellular carcinoma patients, as has INTS14 in lung cancer tissues. Microarray and whole exome studies have implicated INTS8 and INTS2 mutations in gastric

cancer patients (Federico *et al.* 2017). However, it should be noted that most studies have not shown a direct interaction with Integrator mutation in the progression of tumour progression. Instead, perturbations to Integrator function in critical biological processes are thought to drive the system towards cancer development (Federico *et al.* 2017).

1.6 Rationale

snRNA transcripts are a key component of the eukaryotic RNA spliceosome machinery. RNA splicing is an integral part of gene regulation and defects in this process have been linked to premature aging and a host of chronic human diseases. Although the Integrator complex was recently identified as a principal regulator of snRNA maturation, how this protein complex is regulated has not been explored. To address this knowledge gap, my thesis set out to answer the core question of *how is the snRNA processing function of the Integrator complex regulated?* By identifying new regulators of snRNA processing, we can construct pathways that serve as the upstream regulators of the spliceosome, which could provide insights into pathologies associated with RNA splicing defects. In addition, given the recent emergence of reports directly linking Integrator mutation to neurodevelopmental syndromes and shortened lifespan, identifying regulators of the Integrator may reveal additional genes important for maintaining normal organismal development and function.

1.6.1 Hypothesis

I hypothesize that novel genes required for snRNA processing can be discovered through the use of genetic screening in the *C. elegans* model, which will uncover new regulators of the Integrator complex that are functionally important in maintaining normal organismal homeostasis.

1.6.2 Specific thesis objectives

1. Identify novel genes that can affect the snRNA processing function of the Integrator complex via genetic screens.
2. Characterize the mechanistic pathway through which novel regulators affect snRNA processing through the Integrator.
3. Determine the requirements of the Integrator complex and its novel regulators on aging.

TRANSITION

The following chapter focuses on Objectives 1 and 2 of my hypothesis, identification and characterization of novel genes that can affect the snRNA processing function of the Integrator complex via the use of a reverse genetic RNAi screen.

Publication: This chapter has been published in the journal *PLoS Genetics*.

Waddell BM, Wu CW (2024) A role for the *C. elegans* Argonaute protein CSR-1 in small nuclear RNA 3' processing. *PLoS Genet* 20(5): e1011284. https://doi.org/10.1371/journal.pgen.1011284

Contributions: BMW planned and conducted the experiments outlined in this study with input and technical assistance from CWW. The manuscript was drafted and edited by both BMW and CWW. Original ideas and funding were provided by CWW.

Chapter 2

A role for the *C. elegans* Argonaute protein CSR-1 in small nuclear RNA 3' processing

2.1 Abstract

The Integrator is a multi-subunit protein complex that catalyzes the maturation of snRNA transcripts via 3' cleavage, a step required for snRNA incorporation with snRNP for spliceosome biogenesis. Here we developed a GFP-based *in vivo* snRNA misprocessing reporter as a readout of Integrator function and performed a genome-wide RNAi screen for Integrator regulators. We found that loss of the Argonaute encoding *csr-1* gene resulted in widespread 3' misprocessing of snRNA transcripts that is accompanied by a significant increase in alternative splicing. Loss of the *csr-1* gene down-regulates the germline expression of Integrator subunits 4 and 6 and is accompanied by a reduced protein translation efficiency of multiple Integrator catalytic and non-catalytic subunits. Through isoform and motif mutant analysis, we determined that CSR-1's effect on snRNA processing is dependent on its catalytic slicer activity but does not involve the CSR-1a isoform. Moreover, mRNA-sequencing revealed high similarity in the transcriptome profile between *csr-1* and Integrator subunit knockdown via RNAi. Together, our findings reveal CSR-1 as a new regulator of the Integrator complex and implicate a novel role of this Argonaute protein in snRNA 3' processing.

2.2 Introduction

Eukaryotic RNA splicing is catalyzed by the spliceosome that removes noncoding intron segments from pre-mRNA transcripts to produce a mature mRNA for protein translation (Wilkinson *et al.* 2020). A core component of the spliceosome is the uridylate-rich small nuclear RNA (snRNA) molecules U1, U2, U4, U5, and U6 that are incorporated within small nuclear ribonucleoprotein (snRNP) complexes that serve to facilitate splice site recognition for intron removal (Karijolich and Yu 2010; Wilkinson *et al.* 2020). The biosynthesis of snRNA transcripts begins with transcription by RNA pol II to yield a pre-snRNA transcript with an extended 3' precursor (Eliceiri and Sayavedra 1976; Wieben *et al.* 1985). Post transcription, the pre-snRNA transcripts are processed and cleaved by the Integrator complex at the 3' end to yield mature snRNA transcripts that are then incorporated with snRNP towards spliceosome biogenesis (Baillat *et al.* 2005; Chen and Wagner 2010). The Integrator is a metazoan conserved protein complex that is composed of at least 15 distinct subunits in humans and was discovered in 2005 as the elusive molecular machinery for snRNA 3' processing (Baillat *et al.* 2005; Azuma *et al.* 2023). The Integrator catalytic module includes subunits 4, 9, and 11 that are directly involved in snRNA cleavage, while the functions of non-catalytic subunits are not well-defined (Albrecht *et al.* 2018; Mascibroda *et al.* 2022). Knockdown of both catalytic and non-catalytic subunits of the Integrator leads to RNA pol II termination failure and 3' misprocessing of the snRNA transcripts. This results in transcriptional read-through errors that can lead to the aberrant polyadenylation of snRNA, or the synthesis of a long chimeric RNA that is composed of the snRNA transcript and its unprocessed 3' end tethered to its downstream mRNA gene (Skaar *et al.* 2015; Gómez-Orte *et al.* 2019).

Beyond snRNA 3' processing, recent evidence has elucidated a broader role for the Integrator in contributing to transcriptional homeostasis; these functions include the 3' processing of non-coding Piwi-interacting RNAs as well as cleavage of nascent mRNAs at RNA pol II paused sites to facilitate either gene transcription activation or repression (Gardini *et al.* 2014; Beltran *et al.* 2021; Welsh and Gardini 2023). Phenotypically, human mutations to the Integrator complex have been linked to severe neurodevelopmental syndrome and developmental ciliopathies resulting in oral-facial digital syndromes (Oegema *et al.* 2017; Mascibroda *et al.* 2022; Tepe *et al.* 2023). In model organisms, the knockdown of Integrator causes developmental arrest and results in a shortened lifespan of *C. elegans*, and depletion of Integrator in mice results in cortical neuron migration defects leading to neurological disorders (van den Berg *et al.* 2017; Gómez-Orte *et al.* 2019; Wu *et al.* 2019). To date, while the core functions of the Integrator are well developed, regulators of the Integrator complex itself are less understood. Identifying mechanisms of Integrator regulation is of interest given its diverse influence on the transcriptome and its emergence in various human diseases.

In this study, we utilize the genetic model *C. elegans* to develop a GFP-based *in vivo* snRNA misprocessing reporter as a readout for Integrator malfunction and performed a genome-wide RNAi screen to identify potential Integrator regulators. We identified a novel role for the *csr-1* gene encoding an essential Argonaute protein as a regulator of the Integrator complex in snRNA processing (Claycomb *et al.* 2009). CSR-1 is well characterized for its core role in the protection of germline gene expression, this is achieved through a tethered interaction with target transcripts that is mediated by interfacing with small RNAs (22G-RNAs) that are antisense to the targeted gene (Wedeles *et al.* 2013; Cecere *et al.* 2014). More recently, a role for CSR-1 in the cleavage of maternal mRNAs in the embryo to facilitate clearance and removal has also been

demonstrated (Quarato *et al.* 2021). Deletion of *csr-1* results in sterility that is accompanied by loss of P-granule formation, defects in chromosome segregation, and misexpression of replication-dependent histone proteins (Claycomb *et al.* 2009; Avgousti *et al.* 2012). Here, we show that loss of *csr-1* results in an aberrant increase in snRNA 3' misprocessing and the alternative splicing of ~400 transcripts across the transcriptome. Mechanistically, our results show that loss of *csr-1* down-regulates translation of Integrator subunit proteins in the germline that is supported by Ribo-Seq analysis, indicating a reduced translation efficiency of Integrator subunits functioning in both the catalytic and non-catalytic domains. Together, this study provides new insights into *csr-1* as a regulator of the Integrator complex that can influence snRNA 3' processing in *C. elegans*.

2.3 Materials and Methods

2.3.1 *C. elegans* strains

All *C. elegans* strains were cultured at 20°C using standard methods unless noted otherwise (Brenner 1974). The following strains were used: N2 Bristol wildtype, MWU3 *cwwIs1*[*C47F8.9p::C47F8.9::GFP*; *myo-2p::tdTomato*], USC1258 *csr-1a(cmp135)*, WM182 *csr-1(tm892) IV; nT1*[*unc-?(n754) let-?*]*(IV;V)*, WM194 *csr-1(tm892) IV*; *neIs20* [*pie-1::GFP::csr-1 + unc-119(+)*], OD923 *ltSi240*[*csr-1p::csr-1(re-encoded) + Cbr-unc-119(+)*] *II* ; *unc-119(ed3) III*, OD925 *ltSi242* [*csr-1p::csr-1(re-encoded; D606A, D681A: isoform b numbering) + Cbr-unc-119(+)*] *II*; *unc-119(ed3) III*, MWU193 *cwwSi1*[*ints-4::mKATE2::AID*::3xFLAG*]; *wrdSi23* [*eft-3p::TIR1::F2A::mTagBFP2::AID*::NLS::tbb-2 3'UTR*] *(I:-5.32)*, MWU200 *cwwSi1*[*ints-4::mKATE2::AID*::3xFLAG*]; *wrdSi23*[*eft-3p::TIR1::F2A::mTagBFP2::AID*::NLS::tbb-2 3'UTR*] *(I:-5.32)*; *csr-1(tm892) IV/nT1* [*unc-?(n754) let-?*] *(IV;V),* EG9882 F53A2.9(*oxTi1127* [*mex-5p::Cas9(+smu-2 introns)::tbb-2 3'UTR + hsp-16.41p::Cre::tbb-2 3'UTR + myo-2p::2xNLS::cyOFP::let-858 3'UTR + lox2272*] *III*.), JCP341 *jcpSi10*[*ints-6p::ints-6::3xFLAG::eGFP::ints-6 3'UTR + unc-119(+)*] *II*, MWU233 *jcpSi10; csr-1(tm892) IV; nT1*[*unc-?(n754) let-?*]*(IV;V)*. The MWU3 strain was generated by cloning an 883 bp fragment containing the U2 snRNA (C47F8.9 gene) promoter, transcript, and 3' downstream sequence that was then fused in frame to the GFP fluorescent protein (Hobert 2002). This construct was microinjected into *C. elegans* and stably integrated into the genome via U.V. irradiation. The stably integrated snRNA misprocessing strain was outcrossed 4 times before use.

2.3.2 Genome-wide RNAi screen and RNAi experiments

RNAi screen was performed using a protocol previously described in detail (Chomyshen *et al.* 2022). Briefly, synchronized L1 MWU3 larvae obtained from hypochlorite treatment were grown in liquid nematode growth media (NGM) and fed with dsRNA producing *HT115(DE3)* bacteria for 3 days, followed by manual screening for snRNA misprocessing reporter GFP activation using an Olympus SZX61 stereomicroscope. The MRC genomic RNAi feeding library (Geneservice, Cambridge, UK) and the ORFeome RNAi feeding library (Open Biosystems, Huntsville, AL) were used totaling approximately 19,000 clones screened. Clones that activated the snRNA misprocessing reporter from the primary screen were rescreened three additional times using solid NGM agar plates for confirmation. NGM agar RNAi plates were prepared with 50 μg mL^{-1} carbenicillin and 100 μg mL^{-1} of isopropyl β-D-thiogalactopyranoside (IPTG) and seeded with *HT115(DE3) E. coli* expressing the corresponding target dsRNA clone or expressing the pPD129.36(L4440) plasmid that serve as the empty vector (EV) control.

2.3.3 Microscopy and fluorescent analysis

For the snRNA misprocessing reporter, worms were synchronized at the L1 stage and fed with the corresponding RNAi for 72 hours followed by imaging using a Zeiss Axioskop 50 microscope fitted with a Retiga R3 camera. Worms were mounted on a glass slide containing a 2% agar pad and immobilized in a 2% sodium azide solution dissolved in the M9 buffer. To image INTS-4 tagged with mKate2 and INTS-6 tagged with eGFP *in vivo*, the worms were synchronized and grown to the L4 stage and immobilized with 0.65% sodium azide on microscope slides containing a 2% agarose pad and imaged using the Delta Vision (GE) deconvolution system. To image INTS-4::mKate2 in the dissected germline, L4 stage worms were paralyzed with M9 buffer containing 10 mM of levamisole hydrochloride in the cavity of a

concave microscope slide followed by dissection with two 25 gauge needles just below the pharynx to extrude the germline follow by imaging with the Delta Vision (GE) deconvolution system. Fluorescence intensity was determined by the measure function in ImageJ and used to calculate CTCF as defined by [integrated density – (area of selection *x* mean background fluorescence)]; each value was divided by the median value of the wildtype control to determine relative CTCF. Background fluorescence was determined for each image by measuring the signal intensity in an area of the image where fluorescence was absent. Grayscale images taken were colourized in ImageJ using the merge channel function.

2.3.4 RNA extraction, qPCR and RNA sequencing

RNA was extracted by using the Purelink RNA mini kit (ThermoFisher, 12183020) with worm lysis accomplished with a QSonica Q55 sonicator. For each condition, N = 4 biological replicates were prepared with each replicate containing approximately 500 – 1000 worms. For RNA extraction of the *csr-1(tm892)* strain, 100 non-Unc worms were manually picked for each replicate from a synchronized population for RNA extraction. RNA extracted for qPCR analysis was first treated with DNAseI (ThermoFisher, EN0521) followed by cDNA synthesis with the Invitrogen Mutiscribe™ reverse transcriptase system (ThermoFisher, 4311235) using an Applied Biosystems ProFlex Thermocycler. A QuantStudio 3 system was used to perform qPCR with the PowerUp™ SYBR™ Green Master Mix (ThermoFisher, A25741). Relative gene expression was normalized to the housekeeping gene *cdc-42*. Primers used for qPCR are as follows:

Gene name	Forward primer (5'-3')	Reverse primer (5'-3')
numr-1	AGACGTCACTGTTTTGGTGGA	CCGAATCCTCCAGTTGGACC
cdc-42	CTTCTGAGTATGTGCCGACAGTCT	GGCTCGCCACCGATCAT
U2 total	TCTTCGGCTTATTAGCTAAGATCA	CCGAGTCTTCCCTAGGTTCC
U2 misprocessed	CTTCCCAAGGGTCGTCCTGG	GCGGGGAAATGTGTTAAATAGGAC
U4 total	CGAGGTGCGTTTATTGCTGG	GCGTATGCCACCCATGTTTC
U4 misprocessed	CCCCTGAAACATGGGTGGCATACG	ACGCGATTACGGTGTAATGCCAGA

For whole-transcriptome RNA sequencing, RNA was extracted using the same methods described above with the exception that each sample contained ~3,000 worms and that 3 biological replicates were prepared and sequenced for each condition. The RNA samples were sent to Novogene (Sacramento, CA) on dry ice followed by cDNA library construction with the oligo(dT) enrichment method for mRNA sequencing. Sequence annotation and data analysis were performed by Novogene. Mapping of sequence reads to the WBcel235 genome was performed with HISAT2 (v2.0.5), gene quantifications were performed with FeatureCounts (V1.5.0-p3), DESeq2 (v1.20.0) was used to analyze differentially expressed genes, and alternative splicing was analyzed by rMATS v3.2.5 (Anders and Huber 2010; Shen *et al.* 2014; Kim *et al.* 2019). For age estimation, normalized counts of genes from EV and *csr-1* (RNAi) fed worms were analyzed using the RAPToR R package (V1.2) with the wormRef database (V0.5) of Cel_YA_1 as a reference (Bulteau and Francesconi 2022).

2.3.5 CRISPR/Cas9 genome editing and auxin experiments

To insert mKate2::AID*::3xflag in the C-terminus of the *ints-4* locus, we followed the SEC-based protocol described in (Dickinson *et al.* 2015; Ashley *et al.* 2021). Briefly, 5' and 3' homology arm flanking the C-terminal insertion site was amplified from N2 wildtype genomic DNA using the Q5® High-Fidelity DNA Polymerase (NEB, M0491L), followed by HiFi DNA assembly (NEB, E5520S) with the pJW1589 plasmid that was digested with AvrII and SpeI to assembly the repair template. The *ints-4* sgRNA was inserted into plasmid pDD162 via the Q5 Site-Directed Mutagenesis Kit (NEB, E0554S). The sgRNA plasmid and repair template were Sanger sequenced to confirm correct construct assembly followed by microinjection into EG9882 strain with integrated Cas9 activity (10 ng μl^{-1} repair template, 50 ng μl^{-1} sgRNA

plasmid) (Schwartz *et al.* 2021). Primers used for sgRNA insertion into pDD162 and for generation of 5' and 3' homology arms are as follows:

Construct	Forward primer 5'-3'	Reverse primer 5'-3'
ints-4 sgRNA	AAAAAGGTGATATTTATCAA GTTTTAGAGCTAGAAATAGCAAGT	CAAGACATCTCGCAATAGG
ints-4 5' homology arm	**P1:** ACGTTGTAAAACGACGGC CAGTCGCCGGCAATTT GTTCGGTGAGGCTCGG	**P2:** CATCGATGCTCCTGAGGC TCCCGATGCTCCACGGGA ATGTGGATGAATACGAAC
ints-4 3' homology arm	**P3:** CGTGATTACAAGGATGACGA TGACAAGAGATGATAAATATCACCT TTTTTTTTCGTCAT	**P4:** GGAAACAGCTATGACCA TGTTATCGATTTCGCG AAATTGATTGGGGCCTG

Worms homozygous for the roller phenotype were outcrossed with N2 wildtype to remove the integrated Cas9 background followed by heat shock to remove the SEC cassette. Three additional rounds of outcross with the N2 wildtype background were performed to remove any potential non-specific edits. Worms expressing the mKate2::AID*::3xflag insertion to the *ints-4* locus were crossed with JDW225 to introduce the TIR1 protein expressed by the somatic *eft-3* promoter.

For auxin experiments, a 400 mM stock of indole-3-acetic acid 98% (IAA, referred to as auxin) (MilliporeSigma, I3750-25G-A) dissolved in 100% ethanol (ETOH) was used to produce NGM agar plates with a final concentration of 1 mM of auxin. NGM agar plate containing 0.25% ETOH only was used as the corresponding control. Worms were exposed to auxin either at the L1 stage or the L4 stage as described in the figure caption.

2.3.6 Developmental and lifespan analysis

For the developmental assay, synchronized L1 MWU193 worms were grown on NGM agar plates containing 0.25% ETOH or 1 mM auxin and imaged with an Olympus SZX61

stereomicroscope fitted with a Retiga R3 camera after 48 hours to determine the body length. For the lifespan assay, synchronized L1 wildtype or MWU193 worms were grown on NGM agar plates containing 0.25% ETOH or 1 mM auxin at 25°C. The first day of adulthood is marked as 1 day old and adult worms were separated from their progeny via manual picking. Worms were scored every 1-2 days for death via gentle prodding with a sterilized metal pick. Worms were considered dead if they did not respond to the prodding and were censored if they exhibited protruding vulva or gonads. Lifespan assays were performed at 25°C as we observed a high rate of censorship due to vulva protrusion when the lifespan assay was initially performed at 20°C in the MWU193 strain.

2.3.7 Statistical analysis

The GraphPad Prism software (V7.04) was used to generate graphical data and perform statistical analysis. Student's t-test was used when comparing two groups, one-way ANOVA with Dunnett's test was used for comparison of more than two groups, two-way ANOVA with Holm-Sidak's multiple comparisons was used for assessment of two factors with multiple groups, the F-test was used to calculate statistical significance to linear regression data, lifespan data were analyzed using the log-rank test via OASIS2 (https://sbi.postech.ac.kr/oasis2) (Han *et al.* 2016), and the false discovery rate (FDR) correction was applied to determine the statistical significance of RNA-sequencing data.

2.4 Results

2.4.1 Genome-wide screening of snRNA 3' processing factors

The Integrator complex serves as the principle regulator of snRNA processing in eukaryotes that catalyzes 3' post-transcriptional cleavage required for snRNA maturation (Baillat *et al.* 2005). Disruption of the Integrator complex has been shown to impair *C. elegans* development and can mimic a transcriptome profile similar to cadmium exposure (Gómez-Orte *et al.* 2019; Wu *et al.* 2019). To identify novel regulators of the Integrator or snRNA processing, we developed a visual biomarker of snRNA misprocessing in *C. elegans* by adapting the strategy previously employed in the *Drosophila* S2 cells (Ezzeddine *et al.* 2010). We chose to design the snRNA misprocessing reporter using the C47F8.9 transcript encoding the U2 snRNA as we previously showed that the knockdown of Integrator subunits by RNAi results in the misprocessing and increased aberrant polyadenylation of this transcript (Wu *et al.* 2019). A PCR-amplified genomic fragment of C47F8.9 containing the promoter, transcript, and a potential 3' motif for cleavage recognition was cloned in frame with GFP (**Figure 2.4.1a**). *C. elegans* lack a conserved 3' box sequence 9-19 nucleotides downstream of the coding region that is found in other metazoans serving as a cleavage signal for the Integrator (Thomas *et al.* 1990; Chen and Wagner 2010). As such, we cloned approximately 75 base pairs downstream of the C47F8.9 transcript which contains a potential 3' motif that is conserved across U2 snRNA transcripts. Under normal conditions, the GFP signal is absent as snRNA transcripts are cleaved by the Integrator complex resulting in the loss of GFP transcript; however, RNAi knockdown of *ints-4* encoding a catalytic subunit of the Integrator results in transcriptional read-through that strongly activates GFP expression (**Figure 2.4.1b**).

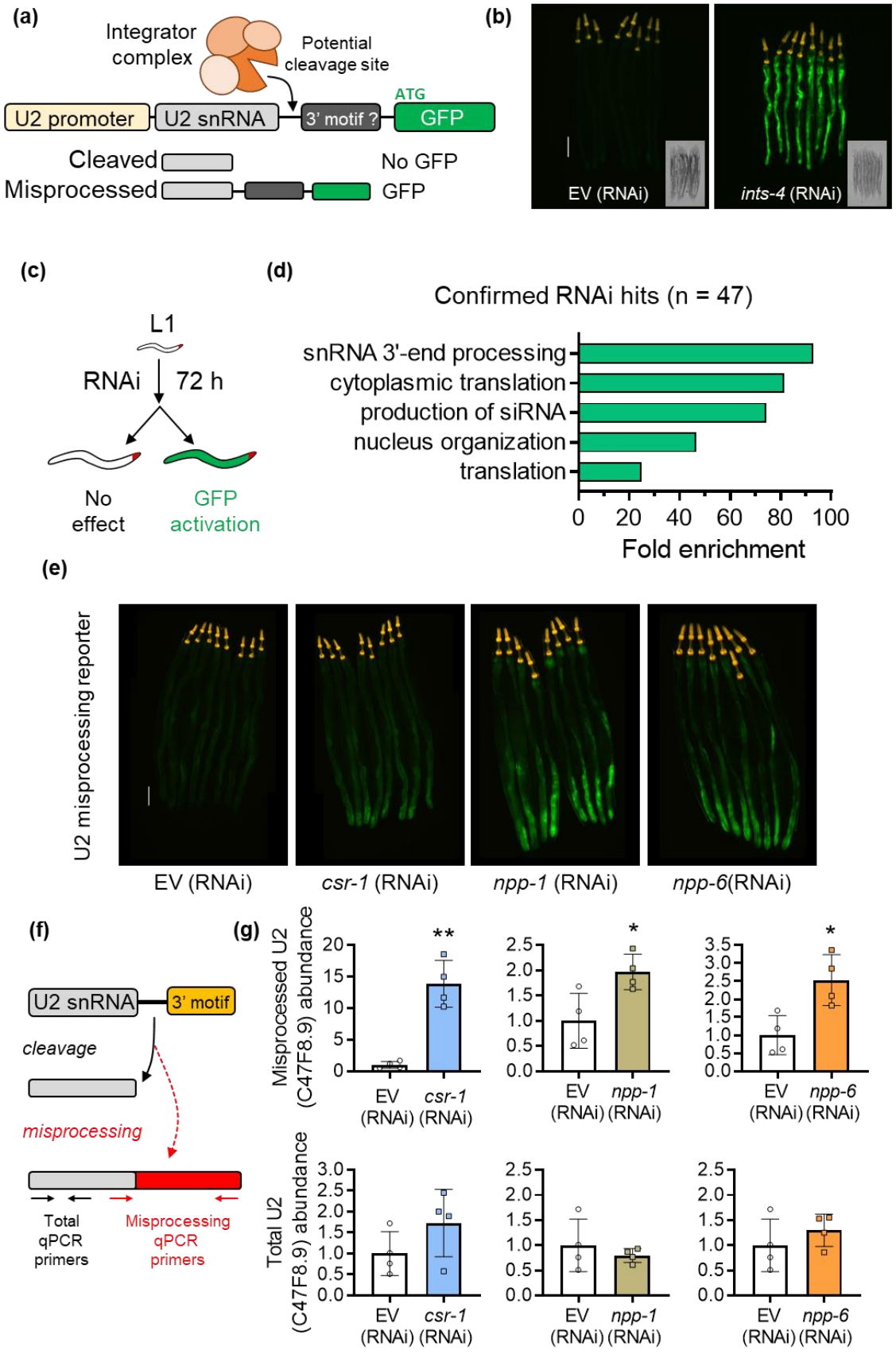

(a)
Integrator complex
Potential cleavage site
ATG
U2 promoter
U2 snRNA
3' motif ?
GFP
Cleaved
No GFP
Misprocessed
GFP
(b)
EV (RNAi)
ints-4 (RNAi)
(c)
L1
RNAi
72 h
No effect
GFP activation
(d)
Confirmed RNAi hits (n = 47)
snRNA 3'-end processing
cytoplasmic translation
production of siRNA
nucleus organization
translation
0
20
40
60
80
100
Fold enrichment
(e)
U2 misprocessing reporter
EV (RNAi)
csr-1 (RNAi)
npp-1 (RNAi)
npp-6(RNAi)
(f)
U2 snRNA
3' motif
cleavage
misprocessing
Total qPCR primers
Misprocessing qPCR primers
(g)
Misprocessed U2 (C47F8.9) abundance
Total U2 (C47F8.9) abundance
EV (RNAi)
csr-1 (RNAi)
npp-1 (RNAi)
npp-6 (RNAi)

Figure 2.4.1. Genome-wide screening of snRNA 3' processing factors. **a)** Schematic of U2 snRNA misprocessing reporter. An 883bp fragment containing the U2 snRNA (C47F8.9 gene), transcript, and 3' sequence was cloned and fused to GFP and stably integrated into the *C. elegans* genome to create an *in vivo* snRNA misprocessing reporter. **b)** Representative fluorescent micrograph and brightfield image showing basal expression of the snRNA misprocessing reporter and GFP activation after Integrator disruption via *ints-4* (RNAi). The yellow signal near the pharynx represents the *myo-3p*::tdTomato co-injection makers. **c)** Outline of the genome-wide RNAi screen to identify gene knockdowns that activate the snRNA misprocessing reporter. **d)** Enrichment analysis of 47 genes that when knocked down activate the snRNA misprocessing reporter. **e)** Representative fluorescent micrograph of the snRNA misprocessing reporter fed with EV, *csr-1*, *npp-1*, or *npp-6* RNAi. The scale bar is 100 μm**. f)** Primer design to measure total and misprocessed transcripts of the U2 snRNA. **g)** Relative levels of misprocessed and total U2 snRNA in worms fed with EV, *csr-1*, *npp-1*, *npp-6* as determined via qPCR. *P<0.05, **P<0.01 as determined by student t-test.

2.4.2 Genome-wide RNAi screen hit list

Next, we performed a genome-wide RNAi screen to identify genes that when knocked down via RNAi result in the activation of the snRNA misprocessing reporter. We screened ~19,000 genes and verified 47 genes that when silenced via RNAi result in GFP activation (**Figure 2.4.1c**). These include genes encoding additional subunits of the Integrator, those involved in nuclear membrane and transport, and regulators of siRNA processing machinery (**Figure 2.4.1d**, **Figure 2.4.2**). We focused on 3 genes that showed the strongest GFP activation that did not encode a known subunit of the Integrator complex, which were *csr-1*, *npp-1* (Nuclear Pore complex Protein), and *npp-6* (**Figure 2.4.1e**). To verify that the increase in GFP reporter fluorescence reflects the misprocessing of the endogenous snRNA transcript, we designed a pair of primers that measure the total and misprocessed levels of C47F8.9 (**Figure 2.4.1f**). We then knocked down *csr-1*, *npp-1*, and *npp-6* in N2 wildtype worms via RNAi and found that the knockdown of all three genes resulted in a significant increase in misprocessed levels of C47F8.9 without affecting the total transcript levels. To confirm that the processing of other snRNA transcripts was also regulated by these three genes, we employed the same qPCR strategy and found that the knockdown of *csr-1* and *npp-6* led to increased misprocessing of the U4 snRNA transcript K03B8.10 (**Figure 2.4.3a**). Overall, the results here identified several novel regulators of snRNA processing and verified a role for *csr-1* and *npp-6* as a requirement of the 3' processing of U2 and U4 snRNA transcripts.

Gene name	Wormbase ID	Gene name	Wormbase ID
npp-6	WBGene00003792	*rps-7*	WBGene00004476
mut-16	WBGene00003508	*inf-1*	WBGene00002083
ints-8	WBGene00013021	*T25B9.9*	WBGene00012015
dcr-1	WBGene00000939	*rde-4*	WBGene00004326
W03D2.3	WBGene00020980	*pabp-2*	WBGene00003904
npp-1	WBGene00003787	*abce-1*	WBGene00012714
npp-8	WBGene00003794	*rpl-36*	WBGene00004450
T01C3.11	WBGene00011325	*rps-23*	WBGene00004492
ints-4	WBGene00012234	*rpl-23*	WBGene00004435
csr-1	WBGene00017641	*rps-3*	WBGene00004472
rps-5	WBGene00004474	*alg-4*	WBGene00006449
ints-7	WBGene00008361	*rpl-6*	WBGene00004417
ints-5	WBGene00013075	*rps-11*	WBGene00004480
rde-12	WBGene00010280	*rps-18*	WBGene00004487
rpl-26	WBGene00004440	*ubl-1*	WBGene00006725
rpl-33	WBGene00004447	*rpl-26*	WBGene00004440
rps-30	WBGene00004499	*rpl-30*	WBGene00004444
rla-0	WBGene00004408	*rpl-33*	WBGene00004447
rpl-36.a	WBGene00004454	*rpl-22*	WBGene00004434
rpl-11.1	WBGene00004422	*inf-1*	WBGene00002083
rps-22	WBGene00004491	*T25B9.9*	WBGene00012015
rps-18	WBGene00004487	*sars-1*	WBGene00005663
rps-11	WBGene00004480	*abce-1*	WBGene00012714
		pabp-2	WBGene00003904

Figure 2.4.2. List of genes and their corresponding Wormbase ID identified in the genome-wide RNAi screen that when knocked down results in the GFP activation of the U2 snRNA misprocessing reporter.

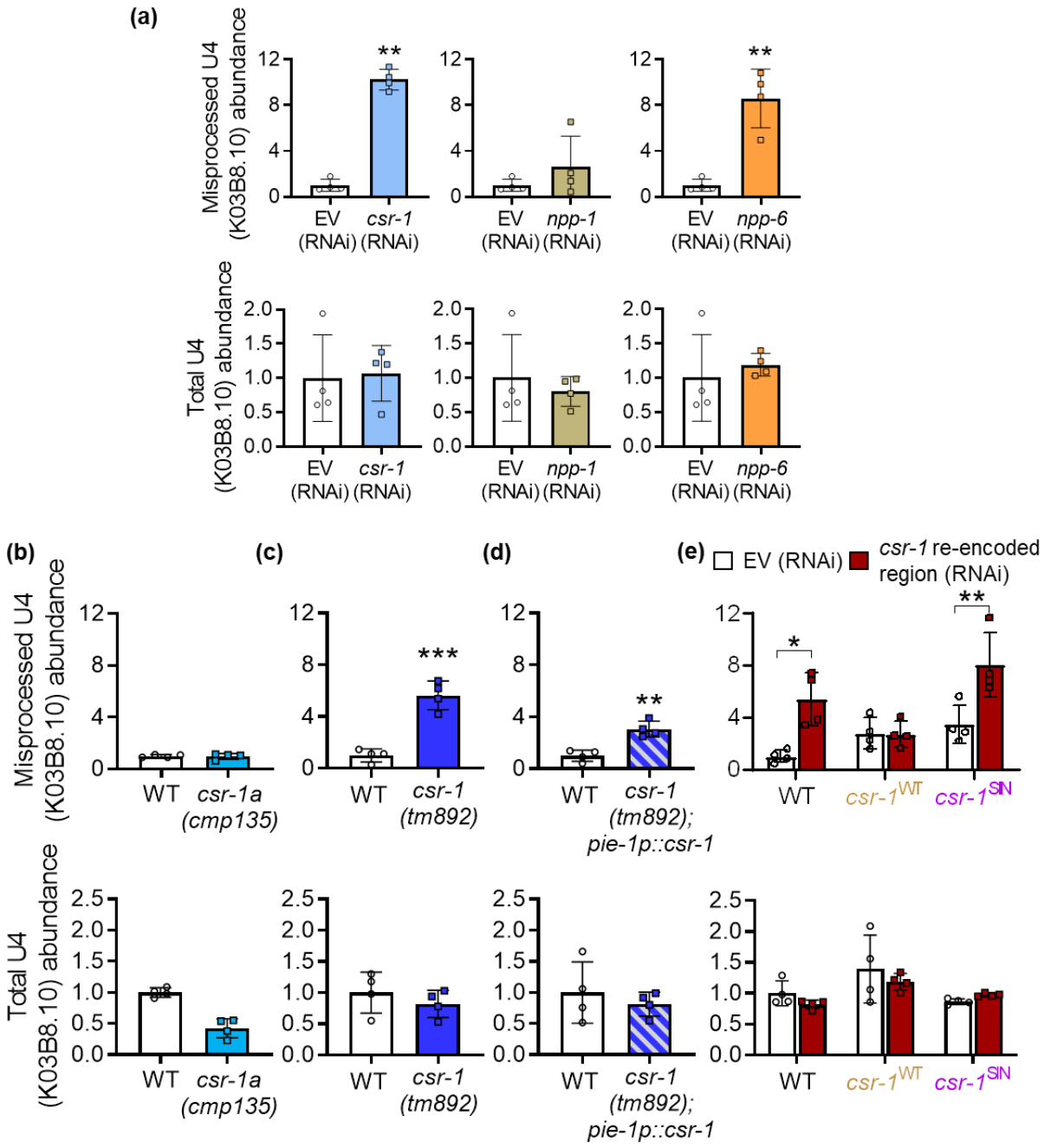


Figure 2.4.3. Regulation of the U4 snRNA transcript by genes identified from the RNAi screen. **a)** Relative levels of misprocessed and total U4 snRNA in worms fed with EV, *csr-1*, *npp-1*, *npp-6* RNAi as determined via qPCR. Relative levels of misprocessed and total U4 snRNA in N2 wildtype (WT) and **b)** *csr-1a(cmp135)* mutant, **c)** *csr-1(tm892)* mutant, and **d)** *csr-1(tm892)* mutant with germline *csr-1* rescue. **e)** Effects of inserting a single copy of RNAi resistant

wildtype *csr-1*^{WT} or slicing-inactive (SIN) *csr-1*^{SIN} on levels of misprocessed and total U4 snRNA after feeding with *csr-1* dsRNA targeting the re-encoded region. All bar graphs indicate mean ± standard error, *P<0.05 **P<0.01, ***P<0.001 as determined by student t-test in **a** to **d**, and by two-way ANOVA in **e**.

2.4.3 Isoform analysis of *csr-1* function in snRNA processing

We decided to focus on characterizing *csr-1* as the endogenous levels of snRNA misprocessing were the greatest compared to the *npp-6*. Since *csr-1* RNAi is predicted to knockdown both *csr-1a* and *csr-1b* isoforms given that the dsRNA targets a shared coding region, we first utilized a *csr-1a* null mutant (*cmp135*) that removes 20 bp of the coding sequence in exon 1 that is exclusively expressed by the *csr-1a* isoform (**Figure 2.4.4a**). Compared to wildtype, the *csr-1a(cmp135)* mutants do not show increased levels of misprocessed U2 or U4 snRNA transcripts (**Figure 2.4.3b, Figure 2.4.4b**). We next tested the *csr-1(tm892)* mutant that contains a 400 bp deletion to both *csr-1* isoforms and found that compared to wildtype, *csr-1(tm892)* mutants showed a significantly increased level of misprocessed U2 and U4 snRNA, with the level of misprocessing comparable to those observed via *csr-1* RNAi (**Figure 2.4.3c, Figure 2.4.4c**). We then examined a hypomorphic allele where *csr-1* is partially rescued in the germline and found that the levels of misprocessed U2 and U4 snRNA were also significantly elevated, albeit not to the same degree as the *csr-1(tm892)* mutant (**Figure 2.4.3d**, **Figure 2.4.4d**). Together, these results suggest that the *csr-1b* isoform, but not the *csr-1a* isoform, is required for snRNA processing. However, given that *csr-1(tm892)* deletes both isoforms, we cannot rule out the possibility that snRNA misprocessing observed in this mutant is caused by the loss of both isoforms.

The CSR-1 protein encodes a Piwi domain that is catalytically active with RNA slicing activity (Aoki *et al.* 2007). To determine if the endo-nucleolytic cleavage or "slicing" activity of CSR-1 is required for snRNA misprocessing, we utilized two worm strains that express either a single copy insertion of wildtype (*csr-1*^{WT}) or slicing inactive (*csr-1*^{SIN}; D606A, D681A mutations) variant of *csr-1b* (Gerson-Gurwitz *et al.* 2016). The two variants of the single copy *csr-1b* also contain a re-encoded region in exon 6 that renders resistance to a *csr-1* RNAi targeting 420 bp within exon 6 that is only effective against the endogenous *csr-1* gene (**Figure 2.4.4a**) (Gerson-Gurwitz *et al.* 2016). Wildtype worms fed with dsRNA targeting the *csr-1* re-encoded region resulted in a significant increase in U2 and U4 snRNA misprocessing (**Figure 2.4.3e**, **Figure 2.4.4e**,). In the *csr-1*^{WT} strain, RNAi against the re-encoded region did not cause an increase in U2 or U4 snRNA misprocessing, suggesting that single copy addition of RNAi resistant wildtype *csr-1b* is sufficient to compensate for the knockdown of endogenous *csr-1* (**Figure 2.4.3e**, **Figure 2.4.4e**). In contrast, RNAi against the re-encoded region in worms expressing *csr-1*^{SIN} resulted in the misprocessing of U2 and U4 snRNA to a similar extent observed in the wildtype strain (**Figure 2.4.3e**, **Figure 2.4.4e**). This suggested that the single copy addition of a slicing inactive *csr-1b* variant does not rescue snRNA misprocessing caused by the knockdown of the endogenous *csr-1* gene (**Figure 2.4.3e**, **Figure 2.4.4e**). Overall, the results here support the requirement of the CSR-1's enzymatic slicing activity for snRNA processing.

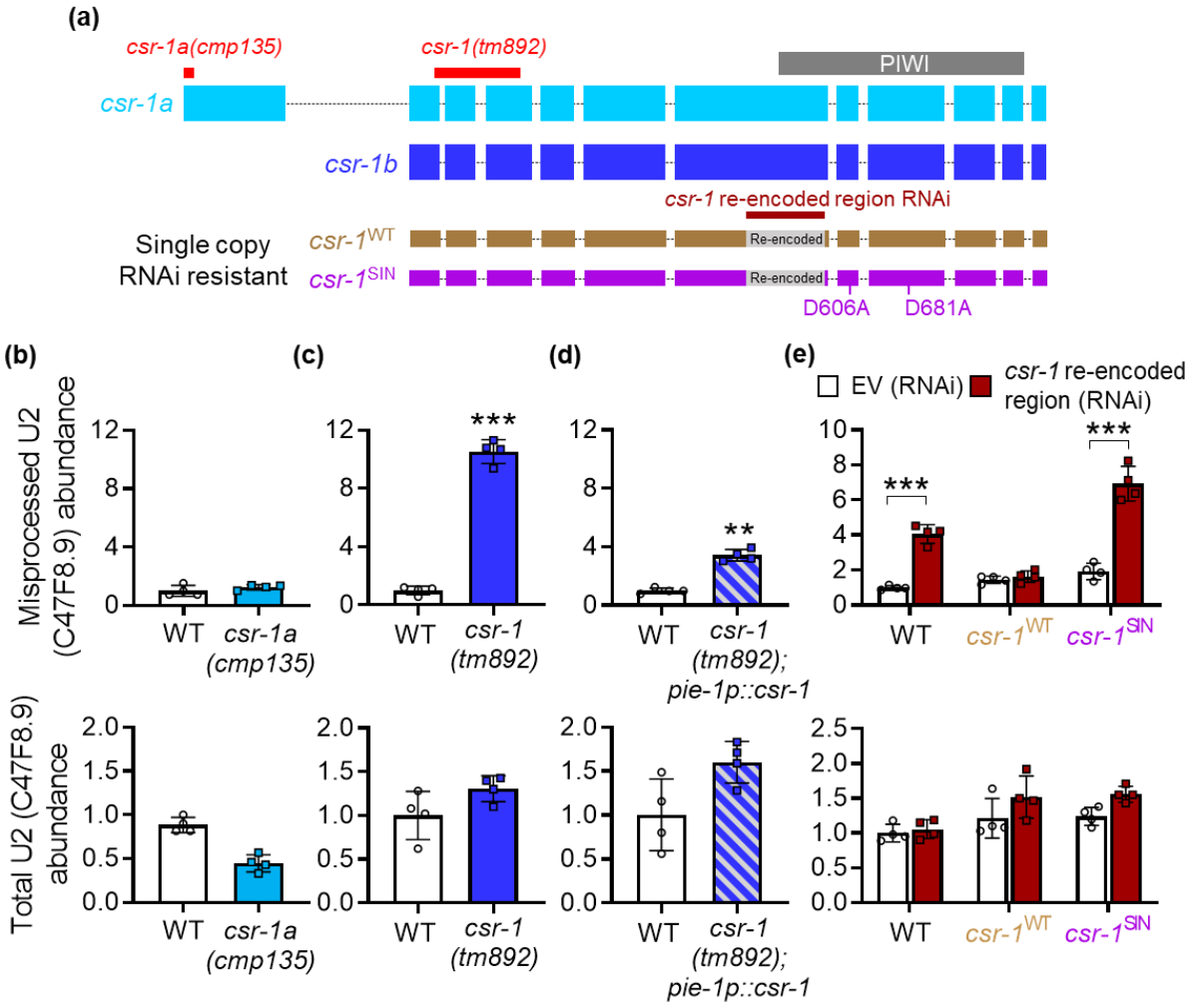


Figure 2.4.4. Isoform analysis of *csr-1* function in snRNA processing. **a)** Diagrammatic illustration of two CSR-1 isoforms, their range of nucleotide deletion in loss of function mutants, and sites of modification to RNAi-resistant single copy insertion strains. Relative levels of misprocessed and total U2 snRNA in N2 wildtype (WT) and **b)** *csr-1a(cmp135)* mutant, **c)** *csr-1(tm892)* mutant, and **d)** *csr-1(tm892)* mutant with germline *csr-1* rescue. **e)** Effects of inserting a single copy of RNAi resistant wildtype *csr-1*^{WT} or slicing-inactive (SIN) *csr-1*^{SIN} on levels of misprocessed and total U2 snRNA after feeding with *csr-1* dsRNA targeting the re-encoded region. All bar graphs indicate mean ± standard error, **P<0.01 and ***P<0.001 as determined by student's t-test in **c** and **d**, and by two-way ANOVA in **e**.

2.4.4 The transcriptome effect of *csr-1* knockdown resembles Integrator disruption

To investigate the role *csr-1* has on the transcriptome, we knocked down *csr-1* using RNAi and performed mRNA-sequencing of oligo(dT) enriched transcripts. An increased accumulation of snRNA transcripts beyond their 3' end is observed in the *csr-1* knocked-down worms indicating transcriptional read-through as evidence for 3' misprocessing (**Figure 2.4.5a**). We found that expression for 38 snRNA transcripts can be detected in the EV control sample, and knockdown of *csr-1* resulted in a 2-fold increase in 27/38 of these transcripts (**Figure 2.4.6a**). Furthermore, 18 snRNA transcripts were detected in the *csr-1* RNAi knocked-down worms that were not detected in the EV control-fed worms. (**Figure 2.4.6b**). Given that snRNA transcripts are only polyadenylated as a consequence of misprocessing that results in transcriptional read-through (Skaar *et al.* 2015; Wu *et al.* 2019), an increased expression of snRNA abundance in oligo(dT) enriched samples also reflects an increase in snRNA misprocessing. To account for any potential developmental differences caused by *csr-1* (RNAi) in the transcriptomic data, we performed real-age prediction using transcriptome staging (RAPToR) (Bulteau and Francesconi 2022). We found that EV and *csr-1* (RNAi) samples have a predicted age of 71.82 ± 0.15 and 71.74 ± 0.16 hours, respectively, which suggest that gene expression variance between the two conditions were unlikely to be caused by potential developmental differences (**Figure 2.4.6c**). However, it should be noted that loss of *csr-1* has been shown to delay the onset of oocyte production which can lead to potential alternation of germline gene expression (Singh *et al.* 2021).

A recent study has shown that the transcriptional read-through effect of Integrator malfunction results in the up-regulation of genes that are located downstream of the misprocessed snRNA (Gómez-Orte *et al.* 2019). We compared the fold change of genes that are

located directly downstream of each snRNA transcript after Integrator subunit-4 (*ints-4*) RNAi knockdown to the fold change observed for the same genes after *csr-1* RNAi knockdown and found that the expression of snRNA downstream genes is highly correlated (R = 0.626) between *ints-4* and *csr-1* knockdown (**Figure 2.4.5b**). We then performed the same analysis using RNA-seq data of the *csr-1(tm892)* mutant recently reported by Singh et al. and found a similar result where expressions of snRNA downstream genes were significantly correlated between *csr-1(tm892)* mutants and *ints-4* (RNAi) (**Figure 2.4.6d**) (Singh *et al.* 2021). This indicated that knockdown or loss of *csr-1* also results in similar up-regulation of snRNA downstream genes observed after *ints-4* depletion as a consequence of transcriptional read-through.

Given that *csr-1* knockdown resulted in snRNA misprocessing, we next determined the effects it has on alternative splicing. Analysis of the RNA sequencing data showed that 6,808 alternative splicing events were detected in *csr-1* knocked-down worms that were not present in EV controls, with 391 of these events found to be statistically significant and composed of primarily exon skipping events (**Figure 2.4.5c**). Enrichment analysis of these 391 significant events reveals clustering to a wide range of cellular processes including calcium transport, muscle functions, and response to stimulus (**Figure 2.4.5d**). However, these enriched processes of alternatively spliced transcripts were largely distinct from the enriched cellular processes of genes that were up or down-regulated by *csr-1* RNAi (**Figure 2.4.5e-f**). Of the transcripts that were alternatively spliced by *csr-1* (RNAi), 60% (232/391) were differentially expressed at the mRNA levels (139 up-regulated, 93 down-regulated), which is ~2-fold higher than the global effect *csr-1* (RNAi) has on the whole transcriptome where 31% of the genes were differentially regulated (FDR<0.05, no fold change cut off). While this suggests that transcripts alternatively

spliced in response to *csr-1* depletion have increased rates of differential expression, the functional significance of these alterations remains to be experimentally determined.

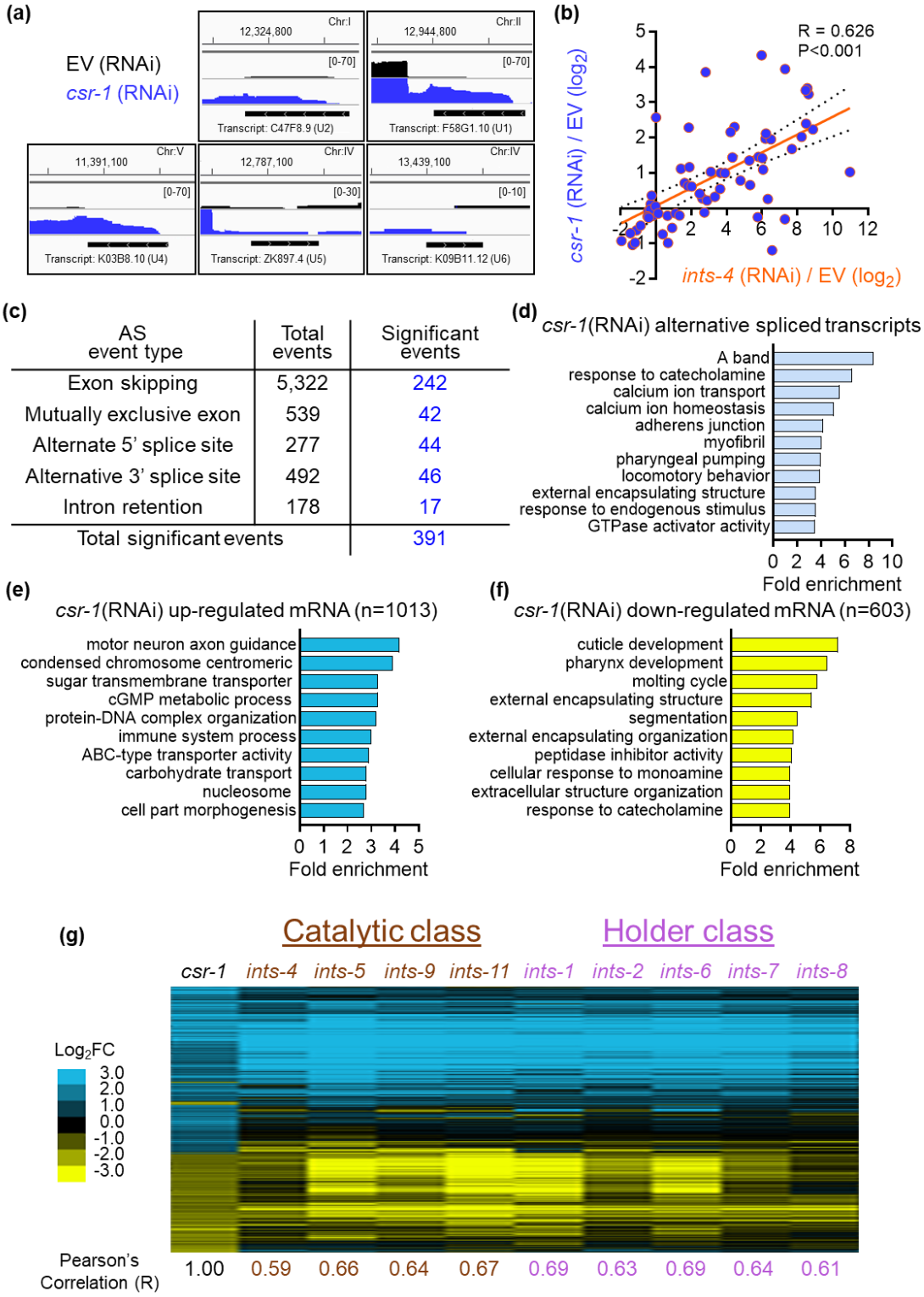


Figure 2.4.5. The transcriptome effect of *csr-1* knockdown resembles Integrator disruption. **a)** RNA sequencing reads in EV or *csr-1* (RNAi) fed worms aligned to the *C. elegans* genome in the region of different snRNA genes visualized using the Integrated Genomics Viewer. **b)** Linear regression analysis on $\log_2$ fold change of snRNA downstream genes between *ints-4* (RNAi) and *csr-1* (RNAi), ***P<0.001 as determined by the F-test. N = 72 snRNA downstream genes are plotted. **c)** Number of transcripts alternatively spliced by *csr-1* (RNAi) compared to EV and **d)** Gene Ontology (GO) terms enriched by transcripts that undergo significant alternative splicing events. Enrichment analysis of GO terms by mRNA transcripts that were **e)** up-regulated or **f)** down-regulated by >2-fold in *csr-1* (RNAi) compared to EV. **g)** Clustered heat map of $\log_2$ fold change in gene expression changes caused by *csr-1* (RNAi) / EV compared to RNAi knockdown of different Integrator complex subunit genes / EV that function within the catalytic or holder class. Pearson's correlation (R) value for each gene knockdown compared to *csr-1* (RNAi) is shown below. RNA-sequencing expression data for Integrator subunit knockdown were retrieved from (Gómez-Orte *et al.* 2019) for analysis. N = 1353 genes up or down-regulated by >2-fold after *csr-1*(RNAi) are plotted.

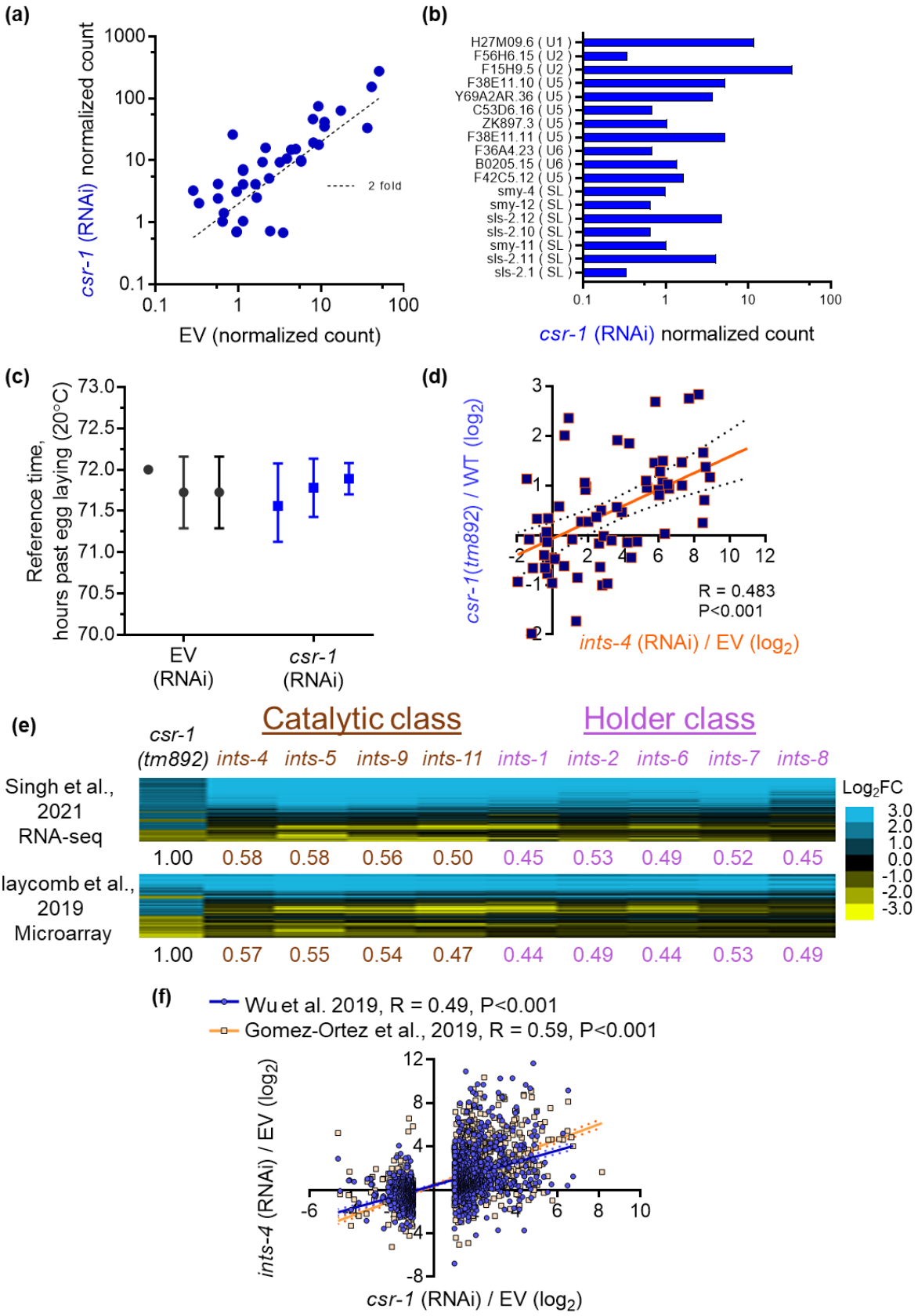

(a)
csr-1 (RNAi) normalized count
EV (normalized count)
2 fold
(b)
H27M09.6 (U1)
F56H6.15 (U2)
F15H9.5 (U2)
F38E11.10 (U5)
Y69A2AR.36 (U5)
C53D6.16 (U5)
ZK897.3 (U5)
F38E11.11 (U5)
F36A4.23 (U6)
B0205.15 (U6)
F42C5.12 (U5)
smy-4 (SL)
smy-12 (SL)
sls-2.12 (SL)
sls-2.10 (SL)
smy-11 (SL)
sls-2.11 (SL)
sls-2.1 (SL)
csr-1 (RNAi) normalized count
(c)
Reference time, hours past egg laying (20°C)
EV (RNAi)
csr-1 (RNAi)
(d)
csr-1(tm892) / WT (log2)
R = 0.483
P<0.001
ints-4 (RNAi) / EV (log2)
(e)
Catalytic class
Holder class
csr-1 (tm892)
ints-4
ints-5
ints-9
ints-11
ints-1
ints-2
ints-6
ints-7
ints-8
Singh et al., 2021 RNA-seq
1.00 0.58 0.58 0.56 0.50 0.45 0.53 0.49 0.52 0.45
Claycomb et al., 2019 Microarray
1.00 0.57 0.55 0.54 0.47 0.44 0.49 0.44 0.53 0.49
Log2FC
(f)
Wu et al. 2019, R = 0.49, P<0.001
Gomez-Ortez et al., 2019, R = 0.59, P<0.001
ints-4 (RNAi) / EV (log2)
csr-1 (RNAi) / EV (log2)

Figure 2.4.6. Depletion of *csr-1* leads to a widespread increase in snRNA abundance. **a**) Normalized count of snRNA transcripts in poly-A-tailed mRNA sequenced RNA samples in worms fed with EV as compared to *csr-1* RNAi. The dotted line indicates 2 fold increase in expression. **b)** Normalized count of snRNA transcripts in worms fed with *csr-1* RNAi that was not detected in worms fed with EV. **c)** Age estimation of EV and *csr-1* RNAi-fed worm RNA samples based on the expression profile of 25,902 genes detected via RNA-sequencing as determined via RAPToR (Bulteau and Francesconi 2022). **d)** Linear regression analysis on $\log_2$ fold change of snRNA downstream genes between *ints-4* (RNAi) and *csr-1(tm892)* using data from (Singh *et al.* 2021), N = 71 snRNA downstream genes are plotted. **e)** Clustered heat map of $\log_2$ fold change in gene expression changes caused by *csr-1* (*tm892*) / N2 wildtype (WT) compared to RNAi knockdown of different Integrator complex subunit genes / EV that function within the catalytic or holder class. Pearson's correlation (R) value for each gene knockdown compared to *csr-1(tm892)* is shown below. N = 2133 and N = 3576 of genes up or down-regulated by >2-fold in *csr-1*(*tm892*) relative to wildtype are plotted for the Singh et al dataset (Singh *et al.* 2021) and Claycomb et al dataset (Claycomb *et al.* 2009) respectively. **f)** Linear regression analysis of genes with >2-fold change in *csr-1* (RNAi) compared to the expression of the corresponding genes in *ints-4* (RNAi) obtained from two independent studies (Gómez-Orte *et al.* 2019; Wu *et al.* 2019). N= 1353 genes plotted. For **d** and **f**, statistical significance was determined by the F-test.

2.4.5 Depletion of *csr-1* leads to a widespread increase in snRNA abundance

To further compare the transcriptome profile between *csr-1* and the Integrator, we generated a clustered heat map of expression changes for genes that were up or down-regulated by *csr-1* RNAi by > 2-fold in comparison to expression changes for these same genes after RNAi knockdown of genes encoding the catalytic and holder class of the Integrator subunit (Gómez-Orte *et al.* 2019). We observed a striking similarity across the 9 different Integrator subunit genes which when knocked down by RNAi exhibit a correlation value of 0.59 – 0.69 in the expression of genes that were differentially regulated by *csr-1* RNAi by > 2-fold (**Figure 2.4.5g**). A similar correlation in expression patterns was also observed between Integrator subunit knockdown and the previously published *csr-1(tm892)* mutant transcriptome obtained from RNA-seq and microarray studies (**Figure 2.4.6e**) (Claycomb *et al.* 2009; Singh *et al.* 2021). We did not compare the gene expression changes to the auxiliary class (*ints-3*, *ints-10*, *ints-12*, and *ints-13*) as RNAi knockdown of these subunits resulted in minimal changes in gene expression compared to the EV control (Gómez-Orte *et al.* 2019). Given that we are comparing transcriptomic data between studies that employed different growth condition that likely introduces batch variation in the RNA samples, we next compared the differentially regulated gene profile of *csr-1* knockdown to another study that measured the effects of *ints-4* RNAi on the transcriptome (Wu *et al.* 2019). Via linear regression analysis, we observed that the gene expression changes caused by *csr-1* knockdown were significantly correlated to those induced by *ints-4* RNAi from two independent studies and that the correlation values between the two studies were comparable at R = 0.49 and R = 0.59 (**Figure 2.4.6f**). Together, the RNA-sequencing data presented here show that knockdown of *csr-1* by RNAi results in the aberrant expression of snRNA transcripts, an

increase in alternatively spliced transcripts, and exhibits a high similarity to the transcriptome profile induced by knockdown of the Integrator complex.

2.4.6 Auxin degradation of INTS-4 on U4 snRNA processing

Given that recent reports indicated that *csr-1b* functions primarily in the germline (Nguyen and Phillips 2021; Charlesworth *et al.* 2021), we next tested for potential epistatic interactions with *ints-4* in snRNA processing via the auxin degradation system. We chose to study *ints-4* as this gene functions within the catalytic subunit of the Integrator complex and that knockdown of *ints-4* via RNAi generated the strongest developmental defect and shortened lifespan compared to the other subunits (Gómez-Orte *et al.* 2019; Wu *et al.* 2019). Via CRISPR, we endogenously tagged the C-terminal end of the *ints-4* locus with mKate2::AID*::3xflag (referred to now as INTS-4::degron) to permit visualization of the INTS-4 protein via a fluorescent tag, and to enable auxin-mediated degradation of Integrator function by the co-expression of the plant TIR1 F-box protein (**Figure 2.4.7a**). We introduced the TIR1 protein under the control of the somatically expressed *eft-3* promoter and found that exposure to 1 mM auxin starting at the L1 stage for 48 hours caused a significant increase in misprocessed U2 and U4 snRNA **(Figure 2.4.7b**, **Figure 2.4.8**). Total levels of U2 snRNA were also elevated by 4.8-fold after INTS-4 degradation, but are considerably lower than the 111-fold increase in misprocessed U2 snRNA. We also confirmed that the metal-responsive *numr-1* gene we previously found to be up-regulated in response to *ints-4* RNAi was also significantly up-regulated. To further validate that the auxin-inducible degradation of INTS-4 is phenotypically similar to previous effects as characterized by *ints-4* RNAi, we measured the effects of auxin exposure on worm development and aging. Exposure to auxin starting at the L1 stage caused a significant reduction in body length after 48 hours and a decreased lifespan of the INTS-

4::degron strain in comparison to ethanol control (**Figure 2.4.7c-d**). These observations are consistent with the previously reported larval arrest and shortened lifespan phenotypes caused by *ints-4* RNAi (Gómez-Orte *et al.* 2019; Wu *et al.* 2019).

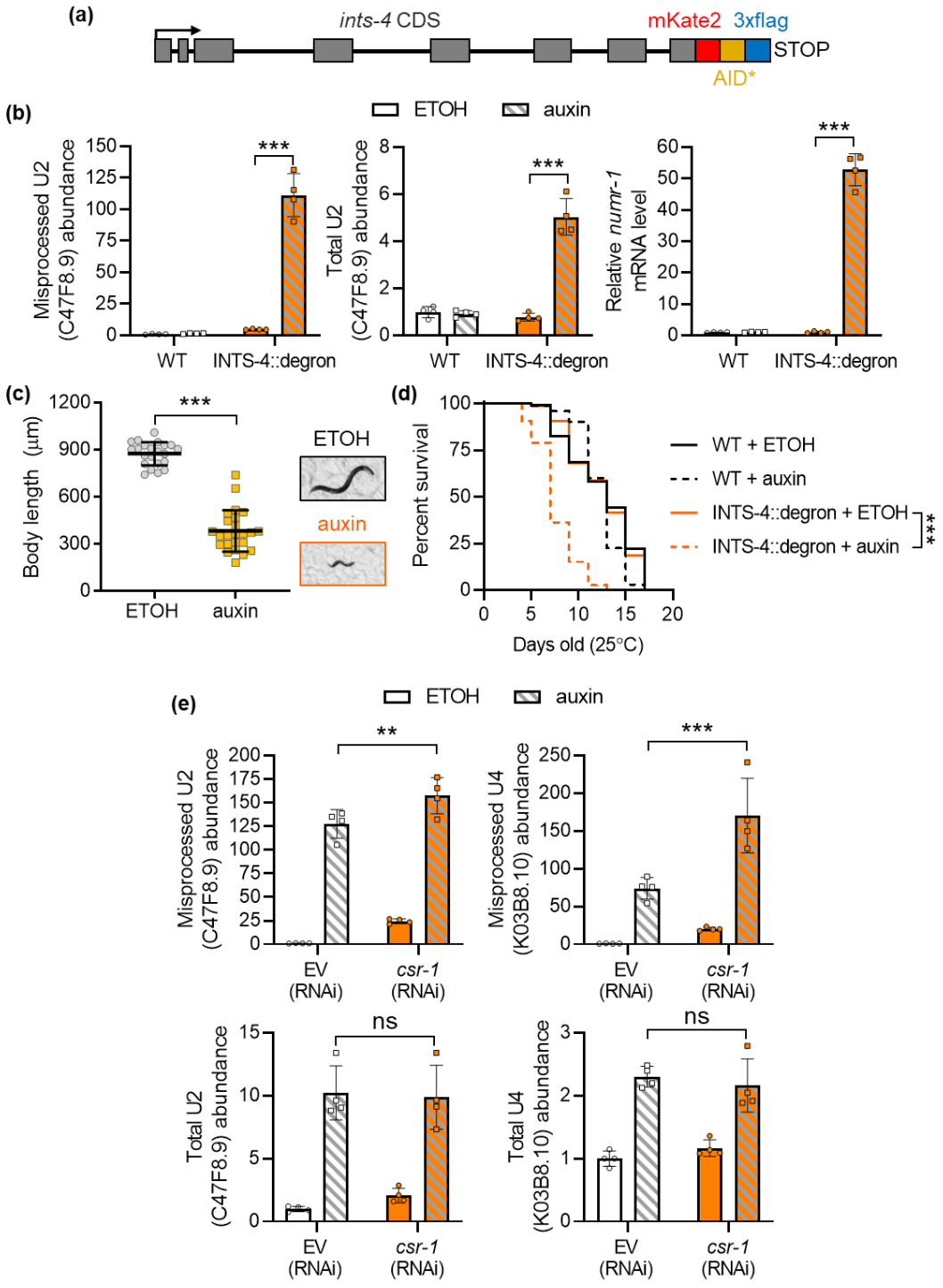


Figure 2.4.7. Genetic interaction between *csr-1* and *ints-4* in snRNA processing. **a)** Schematic illustration of the INTS-4::degron strain generated by CRISPR insertion of mKate2::AID*::3xflag in the C-terminal of the *ints-4* locus. **b)** Relative misprocessed U2, total U2, and *numr-1* transcript levels in N2 wildtype (WT) and INTS-4::degron strain expressing *eft-3p*::TIR1 treated with ethanol (ETOH) or 1 mM of auxin at the L1 stage. **c)** Effects of auxin treatment beginning at the L1 stage to INTS-4::degron; *eft-3p*::TIR1 strain on development after 48 hours compared to ETOH control. N=135-172 worms scored for each condition. ***P<0.001 as determined by the student's t-test. **d)** Lifespan of N2 wildtype and INTS-4::degron strain treated with ETOH or auxin from the L1 stage measured at 25 °C. N = 68 – 139 worms scored for each condition, a summary of lifespan data is presented in Table S3. ***P<0.001 as determined by the log-rank test. **e)** Relative levels of misprocessed and total U2 and U4 snRNA in INTS-4::degron worm strains expressing *eft-3p*::TIR1 fed with EV or *csr-1* (RNAi) at the L1 stage and treated with ETOH or 1 mM auxin for 24 hours beginning at the L4 stage. All bar graphs indicate mean ± standard error. For qPCR data in **b** and **e**,**P<0.01, ***P<0.001 as determined by two-way ANOVA.

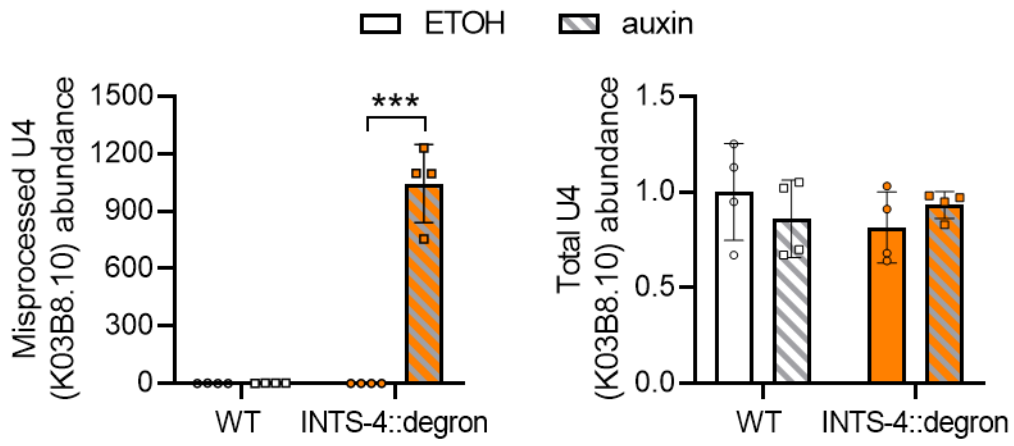


Figure 2.4.8. Auxin degradation of INTS-4 on U4 snRNA processing. Effects of ETOH or 1 mM auxin exposure beginning at the L1 stage on the levels of misprocessed and total U4 snRNA in N2 wildtype (WT) or INTS-4::degron strain. All bar graphs indicate mean ± standard error; ***<0.001 as determined by two-way ANOVA.

2.4.7 Genetic interaction between *csr-1* and *ints-4* in snRNA processing

To test for a potential epistatic relationship between *csr-1* and *ints-4*, we depleted *csr-1* via RNAi and INTS-4 via auxin to determine the effects of single or double knockdown on snRNA processing. Depletion of *csr-1* (via RNAi) or INTS-4 (via auxin) alone both significantly increased misprocessed U2 and U4 snRNA, with INTS-4 depletion expectedly showing a greater effect compared to *csr-1* depletion (**Figure 2.4.7e**). Interestingly, double knockdown of *csr-1* and INTS-4 resulted in a further increase in U2 misprocessing and U4 misprocessing relative to INTS-4 depletion alone (**Figure 2.4.7e**). This could suggest that the knockdown of *csr-1* causes snRNA misprocessing that is additive to the effects of Integrator disruption via depleting the catalytic subunit INTS-4 alone. However, it should be noted that since we did not use null alleles

in this assay, potential residual functions of CSR-1 and INTS-4 may still be present in these double knockdowns.

2.4.8 Loss of *csr-1* alters Integrator subunit expression

To explore the relationship between *csr-1* and the Integrator complex, we first examined 22G-RNA molecules bound to immunoprecipitated CSR-1a or CSR-1b proteins as previously reported by Charlesworth et al. (Charlesworth *et al.* 2021). We found that CSR-1b showed significant enrichment towards 22G-RNA that are antisense to 11 out of 12 transcripts encoding subunits of the Integrator complex (**Figure 2.4.9a**). In contrast, no enrichment of 22G-RNA antisense to Integrator subunits was found for CSR-1a, consistent with our data that loss of *csr-1a* does not influence snRNA processing. Interestingly, immunoprecipitated CSR-1b also shows an increase in enrichment towards 22G-RNA that are antisense to numerous snRNA transcripts, suggesting that CSR-1b may also directly interact with snRNA molecules (**Figure 2.4.10a**). It remains to be determined the functional significance of this interaction as sequence analysis of 22G-RNAs targeting snRNA shows binding to the 5' coding region instead of the potential 3' cleavage region which was observed for histone transcripts targeted by CSR-1 (Avgousti *et al.* 2012).

It has been demonstrated that CSR-1 functions in a protective role of germline expression through a tethered interaction with 22G-RNAs that are antisense to the targeted transcript (Wedeles *et al.* 2013). To explore the possibility that CSR-1 regulates the expression of Integrator subunit genes, we analyzed whole-transcriptome sequencing data between the CSR-1 KO (knockout, *tm892*) mutant and wildtype previously reported by Singh et al. (Singh *et al.* 2021). In CSR-1 KO, total RNA sequencing revealed that the majority of 22G-RNAs targeting Integrator subunits show reduced expression (**Figure 2.4.10b**), this is consistent with the report

that CSR-1 is required for 22G-RNA biogenesis (Singh *et al.* 2021). Analysis of mRNA-seq and GRO-seq data of CSR-1 KO mutant revealed a lack of consistent patterns in the mRNA levels and nascent transcription of genes encoding Integrator subunits (**Figure 2.4.9b-c**). The minimal changes to steady-state mRNA levels of Integrator subunit genes were also observed in our *csr-1* (RNAi) mRNA sequencing results and the previously published *csr-1(tm892)* microarray data (**Figure 2.4.10c**) (Claycomb *et al.* 2009). In the Ribo-seq data, however, a decrease in transcript abundance at active ribosomes was observed for 8 out of 12 Integrator subunits in the CSR-1 KO mutant, suggesting that loss of *csr-1* decreases the translation efficiency of multiple Integrator subunit proteins (**Figure 2.4.9d**).

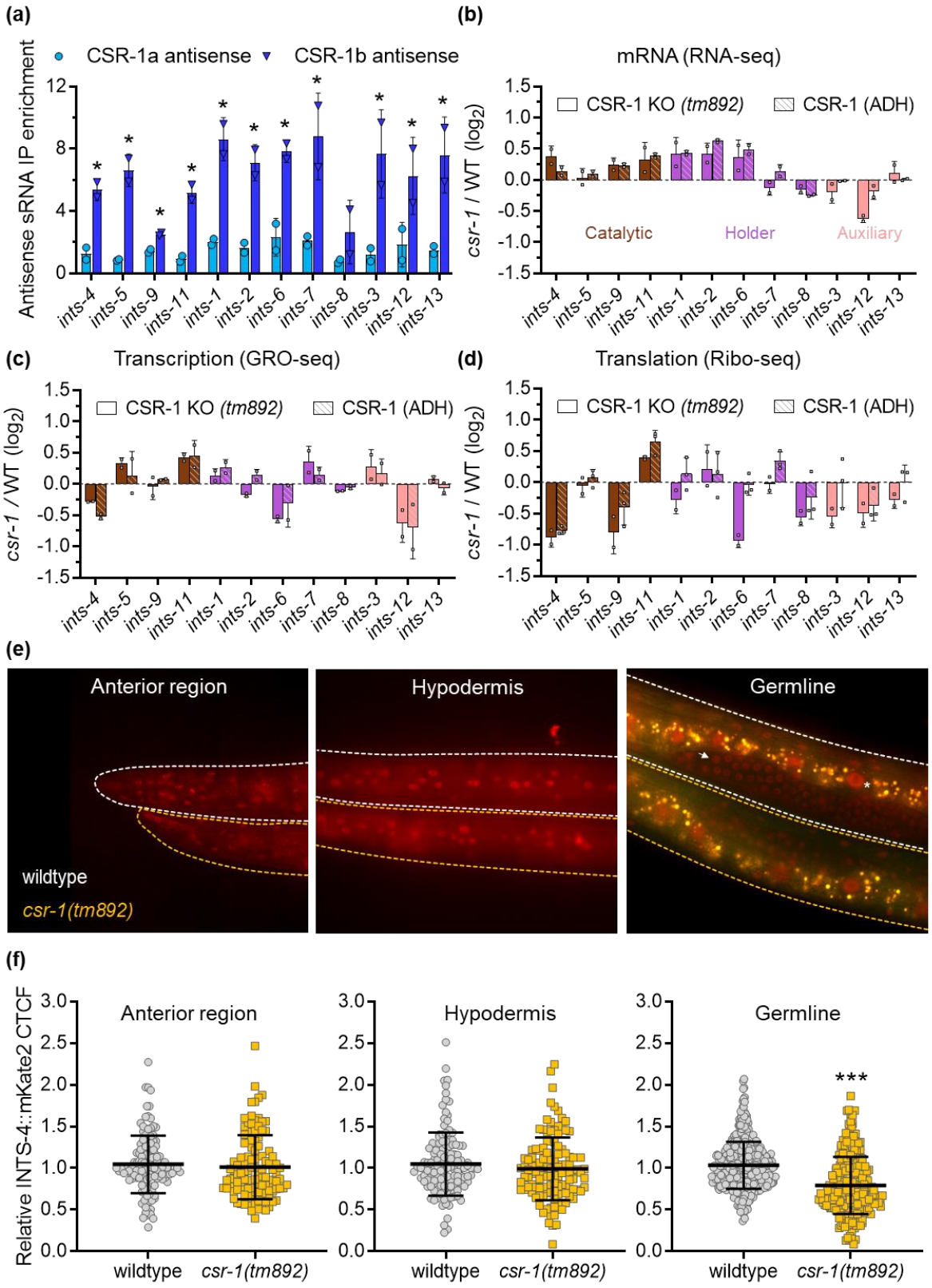

(a)
CSR-1a antisense
CSR-1b antisense
Antisense sRNA IP enrichment
ints-4
ints-5
ints-9
ints-11
ints-1
ints-2
ints-6
ints-7
ints-8
ints-3
ints-12
ints-13
(b)
mRNA (RNA-seq)
CSR-1 KO (tm892)
CSR-1 (ADH)
csr-1 / WT (log2)
Catalytic
Holder
Auxiliary
(c)
Transcription (GRO-seq)
(d)
Translation (Ribo-seq)
(e)
Anterior region
Hypodermis
Germline
wildtype
csr-1(tm892)
(f)
Relative INTS-4::mKate2 CTCF
***

Figure 2.4.9. Loss of *csr-1* alters Integrator subunit expression. **a)** Enrichment of antisense 22G-RNAs complementary to genes encoding Integrator subunits bound to CSR-1a or CSR-1b as determined by small RNA-sequencing data (N = 2 biological replicates) obtained from (Charlesworth *et al.* 2021). Log_2 fold change of CSR-1 KO *(tm892)* or CSR-1 (ADH) compared to N2 wildtype (WT) for **b)** mRNA expression as determined via RNA-seq, **c)** nascent transcription as determined by GRO-seq, and **d)** mRNA translation efficiency as determined by Ribo-seq (N = 2 biological replicates). Data from **c-d** were analyzed from (Singh *et al.* 2021). All bar graphs indicate mean ± standard error, **e)** Representative *in* vivo fluorescent micrograph and **f)** corrected total cell fluorescence (CTCF) quantification of INTS-4::mKate2 expression in the anterior, hypodermis region, and germline in the wildtype or *csr-1(tm892)* mutant background. Two independent imaging trials were performed with N = 112 – 128 nuclei quantified for the anterior region, N = 120 – 132 nuclei quantified for the hypodermis, and N = 254 – 454 nuclei quantified for the germline. RFP in the image represents INTS-4::mKate2 fluorescence, while the yellow is an overlay of GFP and RFP channels which represent auto-fluorescent intestinal granules. ***P<0.05 as determined by student's t-test. The germline composite image was created by merging images taken from the RFP and GFP filters to illustrate the non-specific intestinal gut auto-fluorescent signal represented in yellow. INTS-4 signal in the germline is marked by a white arrow and in the intestine is marked by an asterisk.

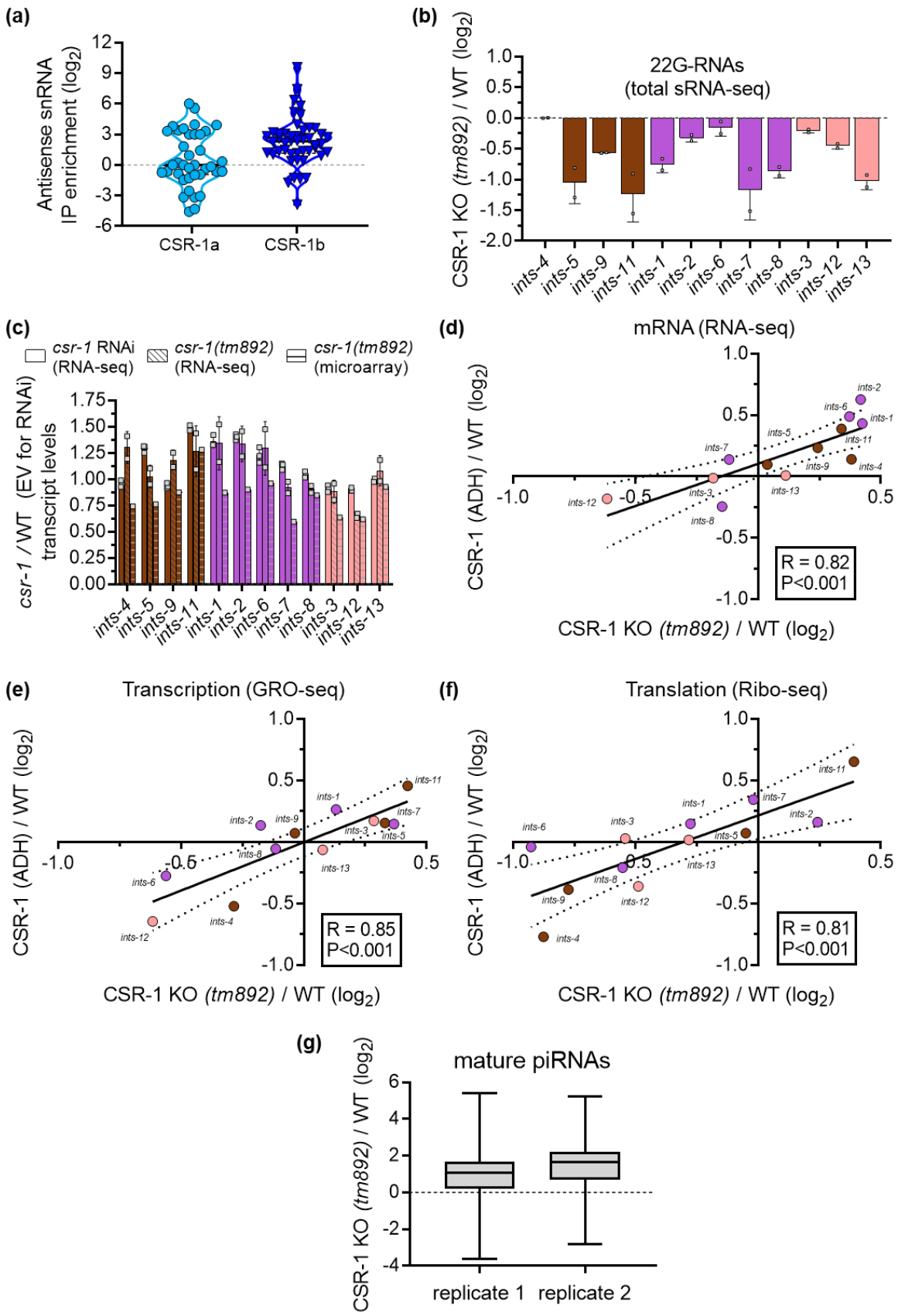

(a)
Antisense snRNA IP enrichment (log2)
CSR-1a
CSR-1b
(b)
22G-RNAs (total sRNA-seq)
CSR-1 KO (tm892) / WT (log2)
(c)
csr-1 RNAi (RNA-seq)
csr-1(tm892) (RNA-seq)
csr-1(tm892) (microarray)
csr-1 / WT (EV for RNAi) transcript levels
(d)
mRNA (RNA-seq)
CSR-1 (ADH) / WT (log2)
CSR-1 KO (tm892) / WT (log2)
R = 0.82
P<0.001
(e)
Transcription (GRO-seq)
R = 0.85
P<0.001
(f)
Translation (Ribo-seq)
R = 0.81
P<0.001
(g)
mature piRNAs
replicate 1
replicate 2

Figure 2.4.10. Requirement of *csr-1* for expression of 22G-RNAs targeting the Integrator. **a)** Enrichment of antisense 22G-RNAs complementary to snRNA transcripts bound to CSR-1a or CSR-1b as determined by small RNA-sequencing data (N = 2 biological replicates) obtained from (Charlesworth *et al.* 2021). **b)** Expression of 22G-RNAs targeting subunits of the Integrator complex in *csr-1(tm892)* mutants relative to N2 wildtype (WT) as determined by total sRNA-sequencing. **c)** mRNA expression level of Integrator subunit genes in *csr-1* RNAi or *csr-1(tm892)* mutant relative to control as determined via mRNA-sequencing and microarray using data from (Claycomb *et al.* 2009; Singh *et al.* 2021). Linear regression analysis on $\log_2$ fold change of Integrator subunit genes between CSR-1 KO (*tm892)* and CSR-1 (ADH) in **d)** mRNA-seq, **e)** GRO-seq, and **f)** Ribo-seq. $P<0.001$ as determined by the F-test. N = 12 Integrator subunit genes plotted for **d-f**. **g)** Global expression of mature piRNA levels in CSR-1 KO mutant / N2 wildtype. Data from **b-g** are obtained and analyzed from (Singh *et al.* 2021) with N=2 samples per biological replicate.

As we showed that the catalytic activity of CSR-1 was required for snRNA processing, we next examined the sequencing dataset of the CSR-1 catalytic dead mutant (termed CSR-1 ADH) reported from the same study (Singh *et al.* 2021). We found that for all three sequencing methods, the effects between CSR-1 KO and CSR-1 (ADH) mutants on Integrator subunit expressions were highly consistent (**Figure 2.4.9b-d**), as supported by a correlation value of 0.82, 0.85, and 0.81 between CSR-1 KO and CSR-1 (ADH) for RNA-seq, Gro-seq, and Ribo-seq, respectively (**Figure 2.4.10d-f**). Together, these data show that loss of CSR-1 expression or its catalytic activity has a robust influence on the translation of select Integrator subunit proteins.

A role for the Integrator in piRNA processing was recently demonstrated which showed that RNAi knockdown of *ints-11* caused a global decrease in piRNA expression (Beltran *et al.* 2021). Given that our data suggest that *csr-1* is required for Integrator gene expression, we examined the effects of CSR-1 KO on piRNA expression. In contrast, we observed that CSR-1 KO mutants exhibit an overall increase in the global expression of mature piRNA transcripts relative to the wild type (**Figure 2.4.10g**). The requirement of other Integrator subunits beyond *ints-11* in piRNA processing has not yet been explored, and curiously, CSR-1 KO mutants show an increase in *ints-11* abundance in Ribo-seq compared to a general trend of decrease in translation efficiency of other subunits (**Figure 2.4.9d**). As such, it remains to be determined whether the global increase in piRNA expression in CSR-1 KO mutants is influenced by its effects on *ints-11* expression, or via other potential pleiotropic functions that CSR-1 may exert in the germline that can affect piRNA expression.

2.4.10 Loss of *csr-1* affects INTS-4::mKate2 expression

To confirm the loss of *csr-1* affects Integrator protein expression, we introduced the INTS-4::degron strain that is also tagged to the mKate2 fluorescent protein into the *csr-1(tm892)*

mutant background. Fluorescent microscopy indicates that INTS-4 is broadly expressed in many tissues of the worm, and is most obvious in the anterior region, germline, and hypodermis tissues (**Figure 2.4.11a**). Treatment with auxin abolishes the mKate2 signal across all tissues, indicating that the fluorescence is associated with INTS-4 (**Figure 2.4.11a**). Next, we compared the expression of INTS-4 in the wild type and *csr-1(tm892)* mutant worm across the 3 regions where mKate2 can be reliably quantified. INTS-4 signal is also observed in the intestine but was not quantified due to the high levels of auto-fluorescence from gut granules observed in the intestinal tract. We found that *csr-1(tm892)* mutants showed normal expression of INTS-4 in the anterior region and hypodermal cell nuclei, but demonstrated a significant reduction in fluorescence within the germline (**Figure 2.4.9e, f**). To rule out the potential contribution of gut auto-fluorescence to germline INTS-4 quantification due to the proximity of these tissues, we dissected the germline from wildtype and *csr-1(tm892)* mutant and measured INTS-4 fluorescence. In the dissected germlines, we observed a similar decrease of INTS-4 fluorescence in the *csr-1(tm892)* mutant compared to the wildtype and the relative fold change was consistent with the *in vivo* quantification (**Figure 2.4.11b**). A caveat to this analysis is that we did not use Western blot to measure protein expression which may provide a more robust quantification of the precise INTS-4 expression difference between wildtype and the *csr-1(tm892)* mutant.

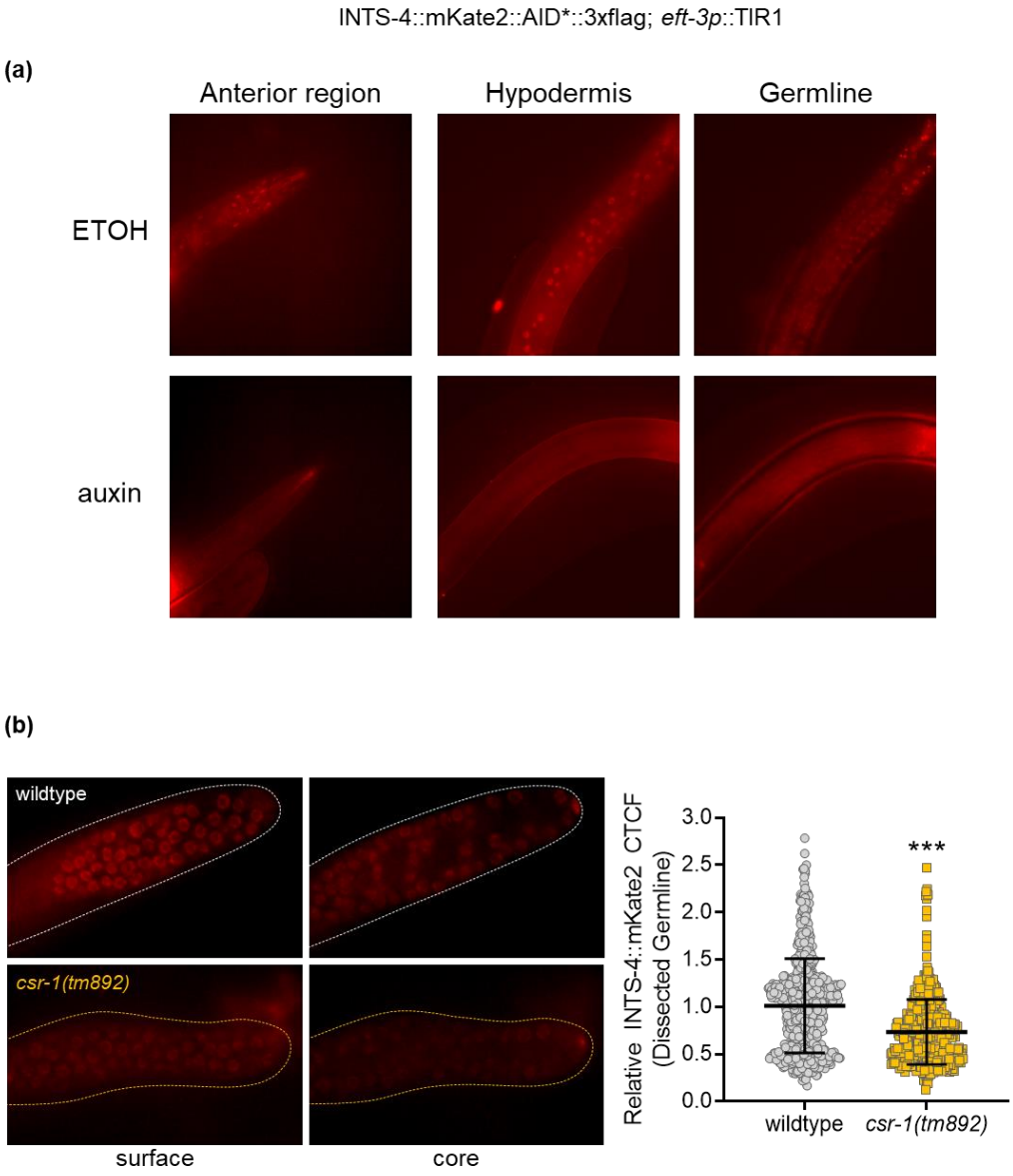


Figure 2.4.11. INTS-4::mKate2 expression. **a)** Representative fluorescent micrograph showing expression of endogenous INTS-4::mKate2 in an L4 worm at the anterior region, hypodermis, and germline after treatment with ETOH or 1 mM auxin. **b)** Representative fluorescent micrograph and CTCF quantification of INTS-4::mKate2 in the distal tip of dissected wildtype and *csr-1(tm892)* mutant germline. Surface nuclei are located on the outer perimeter of the germline while core nuclei are buried inside. Two independent imaging trials were performed

with 612 nuclei quantified for wildtype and 478 nuclei quantified for *csr-1(tm892).* The bar graph indicates mean ± standard error, ***P<0.05 as determined by student's t-test.

2.4.11 Co-contribution of Integrator subunits 4 and 6 to snRNA processing

To explore if loss of *csr-1* affects the expression of other Integrator subunits, we chose to measure the expression of INTS-6 given that *ints-6* also showed a decrease in Ribo-seq abundance in the *csr-1(tm892)* mutant (**Figure 2.4.9d**). Using a strain of worm expressing INTS-6::eGFP under its native promoter, we observed a significant decrease of *in vivo* INTS-6::eGFP fluorescence in the germline of the *csr-1(tm892)* mutant compared to the wildtype (**Figure 2.4.12a**). To determine if the additive effects of snRNA misprocessing observed when INTS-4 and *csr-1* are simultaneously depleted may be caused by the decrease in translation efficiency to multiple Integrator subunits in the *csr-1* mutant, we knocked down *ints-6* via RNAi and depleted INTS-4 via auxin. We found that co-depletion of Integrator subunits 4 and 6 exacerbated the degree of U2 snRNA misprocessing, total U2 expression, and *numr-1* activation compared to depletion of the individual subunit alone (**Figure 2.4.12b**). The effect of multiple Integrator subunit knockdown on expression changes is of synergistic nature, and may be related to the structural dependency of Integrator subunits on each other where disruption of a single Integrator subunit may compromise the activity and function of the associating subunits. For example, it is demonstrated that INTS-4 serves as an anchor for the catalytic module that contains subunits 9 and 11 while INTS-6 interacts and stabilizes with subunits 2, 5, and 8 to form the Integrator backbone core (Wagner *et al.* 2023). Together, these data suggest that *csr-1* contributes to snRNA processing by regulating the translation of multiple Integrator subunits.

Overall, the results in this study show that CSR-1b is bound to 22G-RNA antisense to Integrator subunit genes, and that loss of *csr-1* negatively impacts the translation efficiency of

Integrator subunit transcripts on active ribosomes that can disrupt the post-transcriptional cleavage of snRNA.

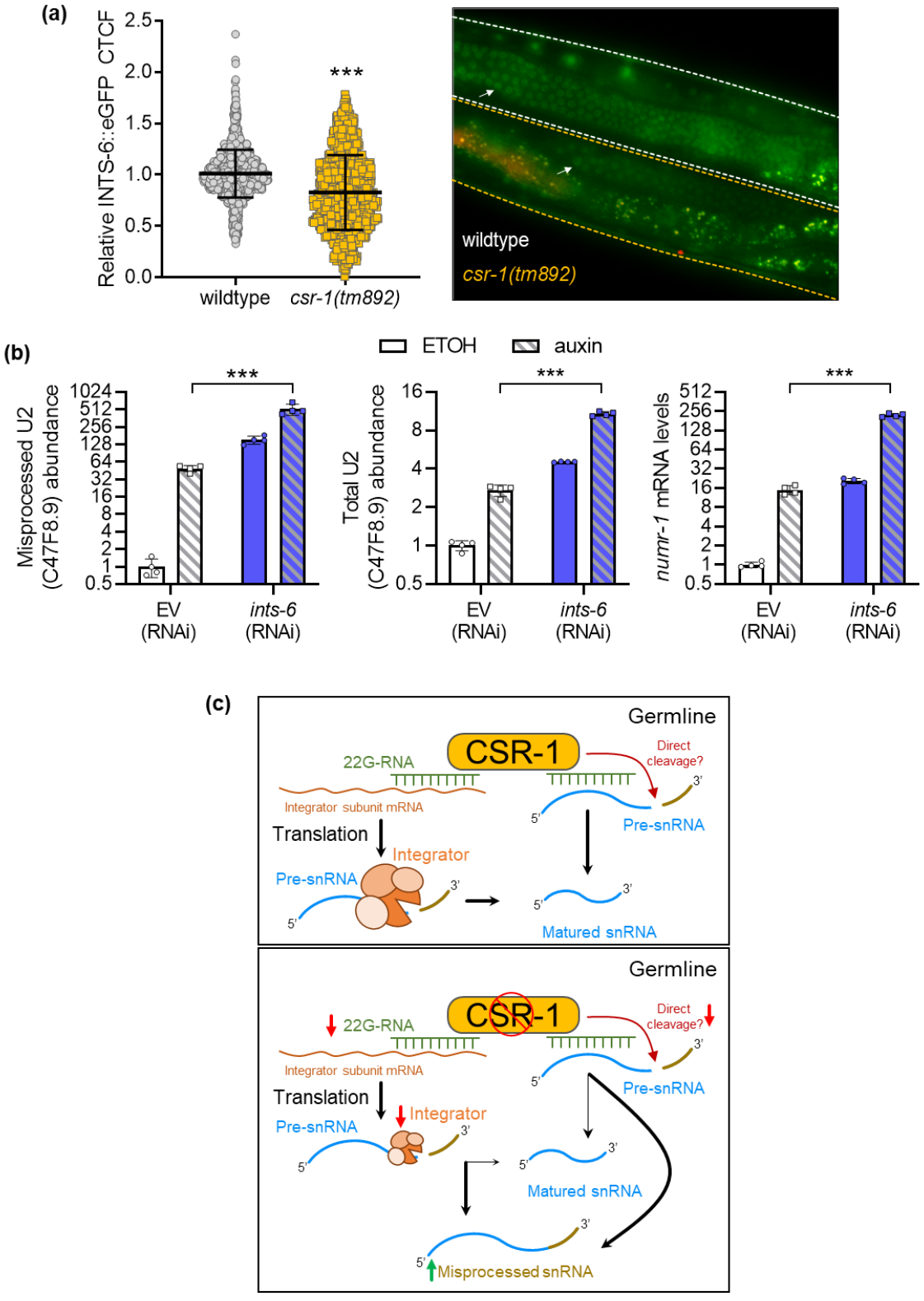

(a)
Relative INTS-6::eGFP CTCF
2.5
2.0
1.5
1.0
0.5
0.0
***
wildtype
csr-1(tm892)
wildtype
csr-1(tm892)
(b)
ETOH
auxin
Misprocessed U2 (C47F8.9) abundance
1024
512
256
128
64
32
16
8
4
2
1
0.5
***
EV (RNAi)
ints-6 (RNAi)
Total U2 (C47F8.9) abundance
16
8
4
2
1
0.5
***
EV (RNAi)
ints-6 (RNAi)
numr-1 mRNA levels
512
256
128
64
32
16
8
4
2
1
0.5
***
EV (RNAi)
ints-6 (RNAi)
(c)
Germline
CSR-1
22G-RNA
Direct cleavage?
Integrator subunit mRNA
Translation
Integrator
Pre-snRNA
Matured snRNA
Germline
CSR-1
22G-RNA
Direct cleavage?
Integrator subunit mRNA
Translation
Integrator
Pre-snRNA
Matured snRNA
Misprocessed snRNA

Figure 2.4.12. Co-contribution of Integrator subunits 4 and 6 to snRNA processing. **a)** Representative fluorescent *in vivo* micrograph and CTCF quantification of INTS-6::eGFP expression in the germline of wildtype or *csr-1(tm892)* mutant worms. Green represents INTS-6::eGFP fluorescence, while the bright yellow foci result from an overlay of the GFP and RFP channels that delimit auto-fluorescent intestinal granules. The bar graph indicates mean ± standard error with N = 1743 and N = 979 nuclei quantified for wildtype and *csr-1(tm892)* respectively, ***P<0.05 as determined by student's t-test. **b)** Relative levels of misprocessed U2 snRNA, total U2 snRNA, and *numr-1* in INTS-4::degron worm strains expressing *eft-3p*::TIR1 fed with EV or *ints-6* (RNAi) at the L1 stage and treated with ETOH or 1 mM auxin for 24 hours beginning at the L4 stage. All bar graphs indicate mean ± standard error, ***P<0.001 as determined by two-way ANOVA. **c)** Proposed mechanism of snRNA processing regulation by CSR-1. CSR-1 in the germline binds to 22G-RNA targeting transcripts encoding subunits of the Integrator complex and this interaction supports the translation of Integrator proteins required for snRNA processing. CSR-1 also binds to 22G-RNA targeting snRNA transcripts suggesting a possible alternative mechanism via direct cleavage. Loss of CSR-1 function or catalytic activity decreases the abundance of 22G-RNA targeting the Integrator and reduces the translation efficiency of Integrator subunits leading to increased accumulation of misprocessed snRNA transcripts.

2.5 Discussion

The Integrator is a metazoan-specific multi-protein complex that was initially discovered as the elusive termination machinery that facilitates the 3' processing of U-rich snRNAs (Baillat *et al.* 2005). Human mutations to the Integrator complex are characterized by increased levels of misprocessed U-rich snRNA transcripts that are accompanied by disruptions to gene expression and RNA processing (Oegema *et al.* 2017). While recent studies have expanded on the core functions of the Integrator beyond snRNA processing including cleavage of nascent mRNAs during RNA pol II pause-release (Gardini *et al.* 2014), additional factors that regulate the Integrator complex or influence snRNA 3' processing remain underexplored. In this study through the use of an *in vivo* snRNA misprocessing reporter in the *C. elegans* system, we identified several genes, including the Argonaute encoding *csr-1*, that when knocked down result in the misprocessing of snRNA transcripts. We propose that *csr-1* is required for the germline expression of Integrator subunit proteins, and that loss of *csr-1* contributes to snRNA misprocessing by altering the abundance of Integrator complex subunits (**Figure 2.4.12c**). Additionally, given that CSR-1 also binds to 22G-RNA targeting snRNA transcripts, it is possible that CSR-1 can directly cleave snRNA molecules given its catalytic slicing activity, a function that has been proposed for the 3' processing of histone transcripts in *C. elegans* (Aoki *et al.* 2007; Avgousti *et al.* 2012; Gerson-Gurwitz *et al.* 2016; Singh *et al.* 2021).

2.5.1 CSR-1 isoforms and slicing activity in snRNA 3' processing

The CSR-1 Argonaute protein is composed of two isoforms, CSR-1a and CSR-b, which have been recently shown to exhibit distinct expression patterns and interact with diverse 22G-RNAs to regulate unique downstream targets (Nguyen and Phillips 2021; Charlesworth *et al.*

2021). CSR-1a expression is more readily detected during adulthood in the spermatogenesis region of the germline and somatic tissues including the intestine, whereas CSR-1b is constitutively detected throughout all developmental stages but strictly expressed in the germline (Nguyen and Phillips 2021; Charlesworth *et al.* 2021). While both CSR-1 isoforms are involved in fertility regulation, loss of *csr-1b* results in complete sterility and loss of *csr-1a* contributes to loss of sperm-based fertility in a transgenerational manner (Nguyen and Phillips 2021; Charlesworth *et al.* 2021). Our mutant analysis shows that snRNA transcripts are not affected in the *csr-1a* mutant but are misprocessed in the *csr-1(tm892)* mutant that contains deletion to both isoforms. This suggests that loss of function to *csr-1b*, but not *csr-1a*, impairs snRNA processing. However, given that a mutant with loss of function to only *csr-1b* is not possible, it is also conceivable that the snRNA misprocessing observed in the *csr-1(tm892)* mutant is caused by the loss of both isoforms. An isoform-specific role of CSR-1 is supported by the co-IP experiments showing that only CSR-1b, but not CSR-1a, binds to endo-siRNAs that are antisense to Integrator subunit genes (Charlesworth *et al.* 2021). Furthermore, we found that the introduction of an RNAi-resistant slicing-inactive (SIN) variant of *csr-1b* was not able to rescue snRNA misprocessing induced by RNAi knockdown of endogenous *csr-1*. Given that the slicing activity has been reported to be essential for CSR-1 in catalyzing the biogenesis of 22G-RNA antisense to its germline targets, this data proposes a model where CSR-1b contributes to snRNA processing by targeting Integrator subunit transcripts via endo-siRNA interaction to governor germline expression (Gerson-Gurwitz *et al.* 2016; Singh *et al.* 2021). A role for CSR-1 as a requirement in maintaining the expression of Integrator subunits is supported by the evidence that genes that are differentially regulated after *csr-1* knockdown are also similarly affected upon Integrator subunit knockdown (**Figures 3g, S2e**). In the future, it will be of interest to be

determined whether these gene expression changes are linked to snRNA misprocessing which affects downstream RNA splicing, or are a consequence of disruption to Integrator's role in nascent mRNA cleavage during RNA pol II pause-release that affects transcriptional efficiency (Stein *et al.* 2022).

2.5.2 CSR-1 in histone and snRNA processing

Akin to snRNA transcripts, replication-dependent histone mRNA are also unique in that they are not polyadenylated after transcription, but rather undergo 3' post-transcriptional cleavage for maturation (Chen and Wagner 2010; Skaar *et al.* 2015). While the 3' processing of histone mRNAs is primarily carried out by the stem-loop binding protein (SLBP) complex, depletion of the Integrator has also been shown to cause aberrant histone polyadenylation, suggesting that the Integrator controls 3' termination of diverse classes of transcripts beyond snRNA (Marzluff *et al.* 2008; Skaar *et al.* 2015; Wu *et al.* 2019). In this study, we identified a role for *csr-1* in snRNA processing, which is intriguing given that *csr-1* has also been previously demonstrated to process 3' cleavage of histone transcripts in *C. elegans* (Avgousti *et al.* 2012). In the histone mechanism, it is proposed that CSR-1 binds to antisense endo-siRNA that are complementary to sequences adjacent to the stem-loop region to facilitate 3' processing (Avgousti *et al.* 2012). Direct cleavage activity of CSR-1 on histone transcripts has not been demonstrated, but it was shown that CSR-1 binds to histone mRNA and knockdown of *csr-1* results in an increased abundance of misprocessed histones that is accompanied by a decrease in histone protein expression (Avgousti *et al.* 2012). While CSR-1 also binds to endo-siRNA that are antisense to snRNA transcripts, the majority of these 22G-RNA are complementary to the 5' coding region of snRNA transcripts instead of the potential 3' cleavage region downstream of the coding sequence, and the significance of this interaction has not been elucidated (Nguyen and

Phillips 2021). Rather, our data suggests that CSR-1 may indirectly influence snRNA processing by regulating the abundance of Integrator proteins; however, we do not rule out the possibility that CSR-1 may also facilitate direct post-transcriptional 3' processing in a similar manner that is proposed for histone transcripts (**Figure 6c**). We show in this study that loss of *csr-1* influences INTS-4 and INTS-6 protein abundance in the germline, and this is consistent with the known role of CSR-1 in licensing the translation of germline-expressed protein-encoding genes in *C. elegans* (Wedeles *et al.* 2013). While our study only directly analyzed INTS-4 and INTS-6 expression, Ribo-seq analysis has revealed that loss of *csr-1* results in the decreased translation efficiency of multiple Integrator subunits that function within the catalytic (*ints-4, ints-9*), holder (*ints-1, ints-6, ints-8*), and auxiliary (*ints-3, ints-12, ints-13*) modules (Singh *et al.* 2021). This decrease in expression of multiple Integrator subunits could also explain why simultaneous knockdown of *csr-1* and INTS-4 caused an additive effect on snRNA misprocessing compared to depletion of INTS-4 alone (**Figure 2.4.7e**), as knockdown of *csr-1* decreased the translation of multiple Integrator subunits, each of which may contribute to overlapping or distinct cleavage activity on snRNA 3' processing (Gómez-Orte *et al.* 2019). This is supported by our data showing that co-depletion of Integrator subunits 4 and 6 exacerbates snRNA misprocessing compared to single subunit disruption (**Figure 2.4.12b**). Overall, this evidence illustrates the broad regulatory role of CSR-1 Argonaute protein in the post-transcriptional processing of histone and snRNA transcript maturation, potentially via a direct and indirect mechanism respectively.

2.5.3 snRNA regulators beyond *csr-1*

Our genome-wide RNAi screen identified a total of 47 genes that are required for snRNA processing as determined by the *in vivo* snRNA misprocessing reporter, 43 of which were not

direct subunits of the Integrator and function in processes such as nuclear organization and siRNA biogenesis. Next to *csr-1*, we verified that the knockdown of *npp-1* and *npp-6* encoding nuclear pore protein that facilitates nucleocytoplasmic transport resulted in the misprocessing of endogenous U2 and U4 snRNA transcripts (**Figure 2.4.1g**) (Galy *et al.* 2003). A direct role for nuclear pore proteins in snRNA processing has yet to be demonstrated, however, it has been shown that snRNA transcripts are briefly exported from the nucleus to the cytoplasm for modification before re-entry into the nucleus for incorporation with snRNP for spliceosome biogenesis (Takata *et al.* 2012). It is possible that the depletion of *npp* genes interferes with this step in nucleocytoplasmic shuttling and disrupts the post-transcriptional maturation of snRNA transcripts. Alternatively, the knockdown of *C. elegans npp* genes has also been shown to disrupt the germline P granule formation, which is the primary subcellular location where CSR-1 is enriched to facilitate 22G-RNA loading for antisense gene targeting (Claycomb *et al.* 2009; Voronina and Seydoux 2010). As such, the knockdown of *npp* may contribute to snRNA misprocessing indirectly by interfering with CSR-1 function through the disruption of P granule integrity and formation.

Our RNAi screen also identified genes that encode various components of the RNAi machinery involved in the 26G-RNA pathway including *mut-16* (MUTator), *dcr-1* (DiCer Related), *rde-4* (RNAi DEfective), *alg-4* (Argonaute Like Gene) (Zhang *et al.* 2011; Youngman and Claycomb 2014). While we did not verify the degree of endogenous snRNA misprocessing caused by knockdown of these genes via qPCR, the activation of the misprocessing reporter may suggest that the defect in endogenous siRNA pathways is linked to snRNA transcript misprocessing. As such, it appears that additional and potentially more complex mechanisms of

snRNA regulation via the 26G-RNA pathway beyond the CSR-1 mediated 22G-RNA mechanism may exist that require future investigations.

2.6 Conclusion

Overall, we demonstrate in this study a positive role for the *csr-1* gene encoding the only essential Argonaute protein in *C. elegans* as a regulator of snRNA processing, through a mechanism where *csr-1* is required for the translation and expression of Integrator subunit genes within the germline. Beyond *csr-1*, the genome-wide RNAi screen presented in this study has also identified several yet to be characterized regulators of snRNA processing including those encoding nuclear protein complex as well as members of the endogenous siRNA pathway. Given the recent expansion of a wide-ranging role for the Integrator complex in gene expression control beyond snRNA processing, and its emerging implication in human diseases (Oegema *et al.* 2017; Mascibroda *et al.* 2022; Tepe *et al.* 2023), it will ultimately be of interest to determine whether these novel regulators may also influence snRNA-independent functions of the Integrator in contributing to transcriptome stability.

TRANSITION

The following chapter also addresses Objective 1 of my hypothesis, the identification of novel genes that can affect the snRNA processing function of the Integrator complex via genetic screens. A forward mutagenesis screen indicated that several components of the RNAi pathway participate in the regulation of snRNA processing.

Publication: Identification of RNAi regulatory genes in snRNA processing of *C. elegans* (to be submitted).

Contributions: BMW conducted all of the experimental study. CWW performed the bioinformatics to map the mutation loci. A manuscript will be drafted by BMW and edited by CWW (Cheng-Wei-Wu).

Chapter 3

Identification of non-lethal *C. elegans* mutants with snRNA processing defect

3.1 Abstract

Splicing is a hallmark of eukaryotic cells. The spliceosome catalyzes the removal of introns from pre-mRNA to promote their maturation into translatable transcripts. The non-coding snRNA molecules are a critical component of RNA splicing and require the function of the Integrator to catalyze its 3' cleavage for the maturation before their incorporation into the spliceosome. Here we utilized a GFP-based *in vivo* snRNA misprocessing reporter as a readout of the Integrator function and performed an EMS chemical mutagenesis screen for snRNA processing regulators. We found that mutation to several components of the RNAi pathway induced the activation of the GFP-based reporter to signal snRNA misprocessing. Among the mutants identified, mutations to *rde-11* or *mut-16* resulted in the highest levels of misprocessing. Perturbations to *rde-1* or *rsd-2* produce moderate and mild levels of misprocessing respectively. Finally, *sago-2* or *rrf-1* mutations cause a mild level of misprocessing. These proteins function at various levels of the RNAi pathway to elicit proper gene expression. While many of these proteins are worm-specific and have no known homologs, it opens the possibility of snRNA processing regulation through the RNAi pathway in other systems.

3.2 Introduction

The spliceosome is a conserved eukaryotic protein complex that functions to remove noncoding introns from pre-mRNA transcripts to generate mature mRNA for protein translation, a function called splicing (Wilkinson *et al.* 2020). Critical components of this complex include the uridylate-rich small nuclear RNAs (snRNA) U1, U2, U4, U5, and U6 that complex with small nuclear ribonucleoproteins (snRNP) complexes, where they direct splice site recognition and intron removal based on mRNA sequences (Karijolich and Yu 2010; Wilkinson *et al.* 2020). Produced from RNA pol II these snRNA transcripts encode extended 3' sequences that are recognized and post-transcriptionally cleaved by the Integrator complex without the use of poly A sequences (Eliceiri and Sayavedra 1976; Wieben *et al.* 1985; Proudfoot 2011; Baillat and Wagner 2015). Integrator complex was initially discovered to contain 12 novel subunits, but recent work has revealed additional regulatory subunits to suggest that this complex contains at least 15 distinct subunits, with the majority of the subunits conserved across metazoans (Baillat *et al.* 2005; Wagner *et al.* 2023). The well-defined catalytic activity of snRNA processing is conferred by subunits 4, 9, and 11, while the functions of other Integrator components are less understood (Albrecht *et al.* 2018; Mascibroda *et al.* 2022).

Recently the Integrator complex has been implicated in additional cellular processes, such as the 3' end processing of Piwi-interacting RNAs (piRNA) and the pause/release of stalled RNA pol II during transcription (Gardini *et al.* 2014; Beltran *et al.* 2021; Welsh and Gardini 2023). At the transcription level, impairment of Integrator subunits results in the loss of 3' end processing and transcriptional read-through into downstream genes (Skaar *et al.* 2015; Gómez-Orte *et al.* 2019). In the context of physiology, perturbations to Integrator function have been linked to human diseases including neurodevelopmental disorders, ciliopathies, and primary

tumours (Gardini *et al.* 2014; Federico *et al.* 2017; Oegema *et al.* 2017; Mascibroda *et al.* 2022). In model systems, studies of Integrator depletion in *C. elegans* and *Drosophila* result in developmental arrest and shortened life span (Wu *et al.* 2019; Tepe *et al.* 2023), while loss of Integrator function mouse models show defects in neuronal migration (van den Berg *et al.* 2017). Substantial effort has gone into understanding the functions of the Integrator, yet how this complex is regulated remains poorly understood.

In most organisms, loss of Integrator function results in lethality as this complex controls various essential cellular processes (Wagner *et al.* 2023). This is supported by our recent genome-wide RNAi screen that identified several genetic regulators of Integrator mediated snRNA processing, with all genes identified as being essential for *C. elegans* development or reproduction (Waddell and Wu 2024). In this study, we utilized the same GFP-based *in vivo* snRNA misprocessing reporter as a readout for Integrator malfunction and performed an EMS mutagenesis screen to identify potentially viable mutants with an Integrator defect. Here we isolate 6 non-essential genes involved in snRNA processing regulation, all of which function within the RNAi pathway in *C. elegans*. The mutants isolated show constitutive activation of the snRNA misprocessing reporter, and function at various junctions of the RNAi pathway including siRNA production (*rde-11*), mutator foci aggregation (*mut-16*), Argonaute protein (*rde-1*), and those function in endogenous RNAi amplification and spreading (*rsd-2*, *sago-2*, *rrf-1*).

The *rde-11* (RNAi DEfective 11) gene encodes a protein critical for small interfering RNA (siRNA) production during RNA interference (RNAi) and experiences among the highest degree of snRNA misprocessing (Zhang *et al.* 2012). The *mut-16* (MUTator 16) gene encodes a disordered worm-specific protein that is indispensable for Mutator foci aggregation, a major compartment of siRNA production, and also experiences a high degree of misprocessing

(Phillips *et al.* 2012). The *rde-1* gene encodes a primary Argonaute (AGO) protein involved in the targeting of primary transcripts for siRNA production and experiences a moderate level of misprocessing (Grishok *et al.* 2005). The *rsd-2* (RNAi Spreading Defective 2) gene encodes an accessory protein that interacts with RDE-11 to produce siRNA, having mild effects on snRNA processing (Sakaguchi *et al.* 2014). The *sago-2* (Synthetic secondary siRNA-deficient ArGOnaute mutant) gene encodes a worm-specific AGO protein that uses siRNA to target transcripts for silencing and has mild effects on snRNA processing (Yigit *et al.* 2006). Finally, the *rrf-1* (RNA-dependent RNA polymerase Family 1) gene encodes an RNA-dependent RNA polymerase (RdRP) that catalyzes the production of siRNA from target mRNA and also has a mild effect on snRNA processing (Grishok *et al.* 2005). Given that all of the genes identified in our EMS screen have previously been shown to cause a defective RNAi response when mutated, this suggests that a functional RNAi pathway is also required for snRNA processing in *C. elegans.*

3.3 Materials and Methods

3.3.1 *C. elegans* Strains

All *C. elegans* strains were cultured at 20°C using standard methods unless noted otherwise (Brenner 1974). The following strains were used: N2 Bristol wildtype, MWU3 *cwwIs1* [*C47F8.9p::C47F8.9::GFP*; *myo-2p::tdTomato*], MWU44 [*rde-1(cww1);cwwIs1*], MWU46 [*rde-11(cww2);cwwIs1*], MWU48 [*mut-16(cww3);cwwIs1*], MWU56 [*rde-1(cww4);cwwIs1*], MWU58 [*mut-16(cww5);cwwIs1*], MWU60 [*mut-16(cww6);cwwIs1*], MWU62 [*rrf-1(cww7);cwwIs1*], MWU64 [*sago-2(cww8);cwwIs1*], MWU66 [*rde-1(cww9);cwwIs1*], MWU68 [*mut-16(cww10);cwwIs1*], MWU70 [*rsd-2(cww11);cwwIs1*],

GR1823 [*mut-16(mg461) I*], MAH23 [*rrf-1(pk1417) I*], NL3307 [*rsd-2(pk3307)*], WM154 [*sago-2(tm894) I*], VS27 [*rde-11(hj37) IV*], WM27 [*rde-1(ne219) V*], and CB4856.

3.3.2 Forward genetic screen and SNP mapping

A large population of gravid worms were age synchronized and grown for 2 days at 20℃ until L4 stage. Worms were then collected in 5 mL of M9 buffer in a 15 mL conical tube with 50 uL of Luria broth added to prevent worms from sticking to the plastic. Stock EMS was added to a final concentration of 50 mM, and the conical tube was wrapped in aluminum foil and set to rock for 4 hours at room temperature. Worms were washed 4 times with M9 by gravity settling before being transferred to a 10 cm Na22 plate seeded with Na22 bacteria for a 2-hour recovery period. Worms were collected in M9 buffer following recovery and ~50 worms were plated onto each of 30 10-cm seeded Na22 plates. Worms were allowed to lay eggs overnight at 16 ℃. Before hatching of the F1 generation, P0 adults were removed from all plates by gently washing with M9 buffer and discarded leaving behind the F1 generation. The F1 was grown for 3 days at 20℃ to adulthood and allowed to lay eggs overnight. The F1 adults were removed from all plates by gentle washing with M9 leaving behind the F2 generation. The F2 was allowed to grow for 2 or 3 days to adulthood while avoiding starvation conditions before screening for GFP activation. Single GFP-positive worms were selected to their own plate and followed to find true breeding populations.

After mutagenesis, newly isolated mutants were backcrossed 4 times to the parent strain to remove unwanted mutations. SNP mapping was used to identify the induced mutations. Male populations of the isolated mutants were generated and crossed with the Hawaiian strain CB4856. Successfully mated F1s were identified by the presence of the RFP co-injection marker associated with the snRNA misprocessing biomarker. F1 progeny were individually picked to

their own plates and allowed to self-fertilize. The resulting F2 progeny were also individually picked to their own plates and allowed to self-fertilize. The F3 populations of each individual mutant were pooled amongst themselves and genomic DNA was extracted with a PureLink® Genomic DNA Mini Kit. Sequencing libraries were generated from genomic DNA using a Nextera Flex® Library Prep Kit (Illumina, Cat# 20018704 and 20027213) and sequenced on a NextSeq 550 using a Mid Output kit (Illumina, Cat#20024904). Reads were extracted and trimmed using the Illumina Generate FASTQ BaseSpace pipeline (version 1.37.0). Sequencing data was analyzed as described by (Doitsidou *et al.* 2010).

3.3.3 Microscopy

Worms were mounted on a glass slide containing a 2% agar pad and immobilized in a 2% sodium azide solution dissolved in the M9 buffer. For the snRNA misprocessing reporter, worms were synchronized at the L1 stage and fed with empty vector RNAi for 72 hours followed by imaging using a Zeiss Axioskop 50 microscope fitted with a Retiga R3 camera. For complementation imaging, mated F1 progeny were allowed to develop for 72 hours before selecting approximately 100 progeny from each mating pair into a single well of a 96-well black wall plate containing 100 uL of M9 with 2% sodium azide. Imaging was performed on an Agilent BioTek Cytation 5 Cell Imaging Multi-Mode Reader at 4x magnification.

3.3.4 RNA extraction and qPCR

RNA was extracted by using the Purelink RNA mini kit (ThermoFisher, 12183020) with worm lysis accomplished with a QSonica Q55 sonicator. For each condition, N = 4 biological replicates were prepared with each replicate containing approximately 500 worms. RNA extracted for qPCR analysis was first treated with DNAseI (ThermoFisher, EN0521) followed by cDNA synthesis with the Invitrogen Mutiscribe^{TM} reverse transcriptase system (ThermoFisher,

4311235) using an Applied Biosystems ProFlex Thermocycler. A QuantStudio 3 system was used to perform qPCR with the PowerUp™ SYBR™ Green Master Mix (ThermoFisher, A25741). Relative gene expression was normalized to the housekeeping gene *cdc-42*. Primers used for qPCR were as previously described (Waddell and Wu 2024).

3.3.5 Mating and complementation

Mating plates were prepared using NGM agar spotted with 25 uL of OP50 bacterial culture grown at 37˚C for 18 hours and allowed to dry overnight. Male populations of the EMS mutants were generated by heat-shocking L4 worms at 30˚C for 6 hours, then backcrossing to their strain of origin to produce robust male populations. For each mating plate, ten males carrying the EMS-induced mutations that show constitutive activation of the GFP reporter are crossed with five L4 hermaphrodites carrying a mutation to the same gene but at different loci. After 3 days, F1 hermaphrodites expressing the GFP reporter (also identified by an RFP co-injection marker) are imaged.

3.3.6 Statistical analysis

The GraphPad Prism software (V7.04) was used to generate graphical data and perform statistical analysis. Two-way ANOVA with multiple comparisons was used for comparison of two factors with multiple groups.

3.4 Results

3.4.1 Forward genetic screen concept and workflow

In eukaryotes, the Integrator complex plays a key role in the 3’ post-transcriptional cleavage of snRNA molecules required for their proper maturation (Baillat et al., 2005). *C.*

elegans development and transcriptomic profile are dependent upon this complex, as perturbations to the Integrator complex result in reduced growth and an altered transcriptome(Gómez-Orte *et al.* 2019; Wu *et al.* 2019). To identify novel regulators of the Integrator or snRNA processing, we previously developed a visual biomarker of snRNA misprocessing in *C. elegans* by adapting the strategy previously employed in the *Drosophila* S2 cells (Ezzeddine *et al.* 2010; Waddell and Wu 2024). The C47F8.9 transcript encoding the U2 snRNA was chosen as the target of the misprocessing reporter given that we previously showed that disruption of the Integrator results in the strong and aberrant polyadenylation of this transcript (Wu *et al.* 2019). We chose to design the snRNA misprocessing reporter using the C47F8.9 transcript encoding the U2 snRNA as it was previously shown that the knockdown of Integrator subunits by RNAi results in the misprocessing and increased aberrant polyadenylation of this transcript. A genomic fragment of C47F8.9 expressing the promoter, transcript, and a potential 3' motif for cleavage recognition was amplified by PCR and cloned in frame with GFP using the fusion PCR approach (Hobert 2002). (**Figure 2.4.1 a**, **Chapter 2)**. Ordinarily, Integrator cleaves these transcripts preventing the expression of GFP. However RNAi knockdown of *ints-4*, a catalytic Integrator subunit, impairs the cleavage function inducing transcriptional read-through activating GFP expression (**Figure 2.4.1 b**, **Chapter 2)**.

Next, we performed a forward genetic screen using EMS searching for viable mutations that cause activation of the misprocessing reporter (**Figure 3.4.1 a-b**) Following mutagenesis, 20 novel mutants were isolated that all showed biomarker activation (**Figure 3.4.1c**). Using qPCR, misprocessed levels of C47F8.9 were compared between the wildtype (MWU3) and the EMS-generated mutants (**Figure 3.4.1d**). We also designed qPCR primers to measure the levels of misprocessed U4 snRNA from the K03B8.10 transcript, which generally showed similar patterns

of snRNA misprocessing compared to C47F8.9 (**Figure 3.4.1d**). Based on the degree of U2 and U4 misprocessing, we selected 11 mutants to perform SNP mapping followed by WGS.

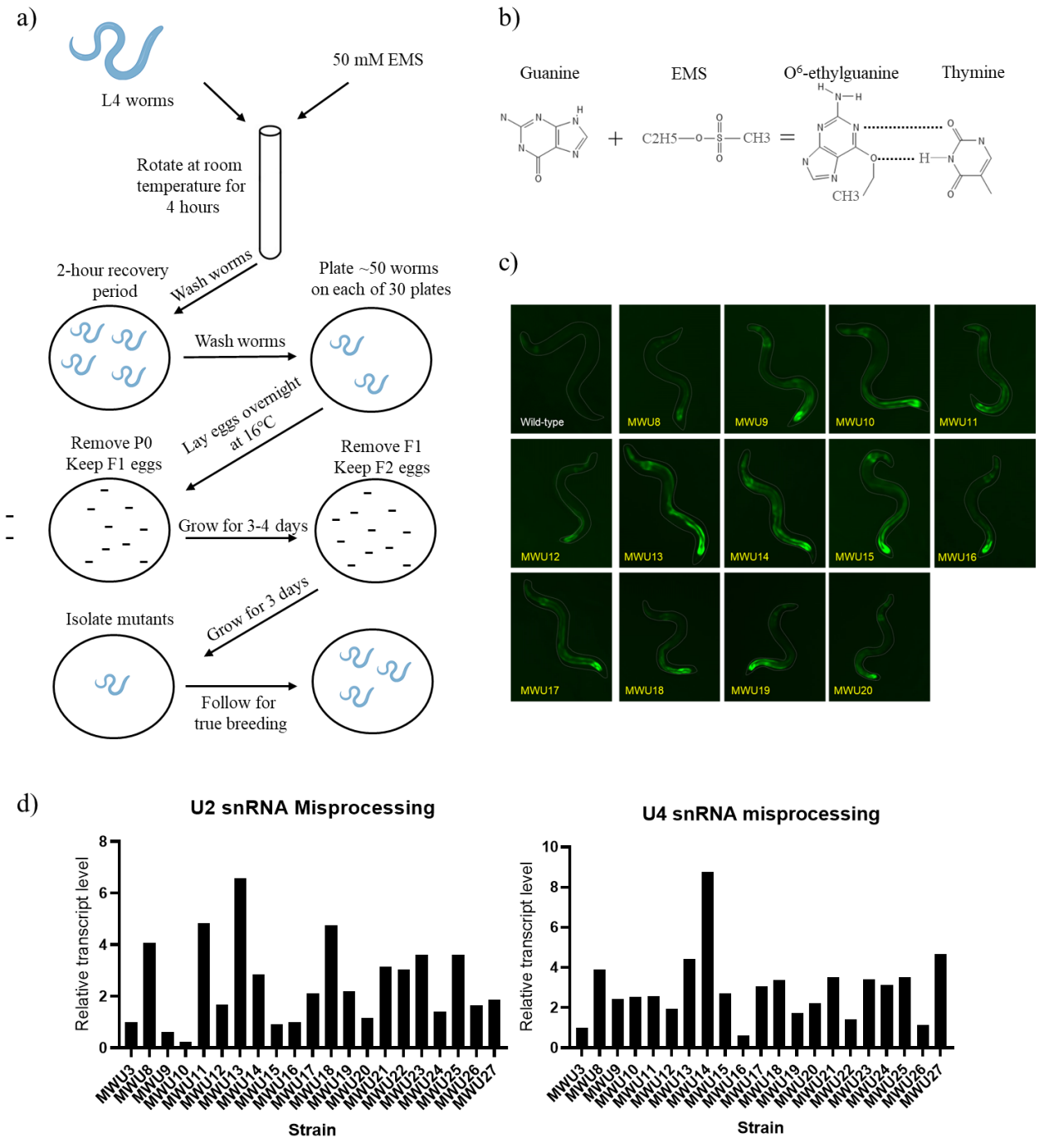


Figure 3.4.1: Forward genetic screen concept and workflow. a) Outline of the EMS mutagenesis protocol. Image generated using BioRender©. b) EMS mutagenic mechanism. Alkylation of guanine by EMS to produce O^6-ethylgaunine induces improper base pairing with thymine resulting in G/C to A/T or A/T to G/C transition mutations. **c)** Isolated EMS mutants showing snRNA misprocessing GFP reporter activation. Images were only taken with the GFP filter and the RFP co-injection marker is not shown. 13 of the 20 isolated mutants are shown. **d)** Confirmation of mutant snRNA misprocessing via qPCR. Mutant and control populations were age-matched via hypochlorite treatment and grown on RNAi media seeded with empty vector bacteria for 72 hours before RNA extraction and cDNA synthesis. Only a single replicate of ~500 worms was performed per condition.

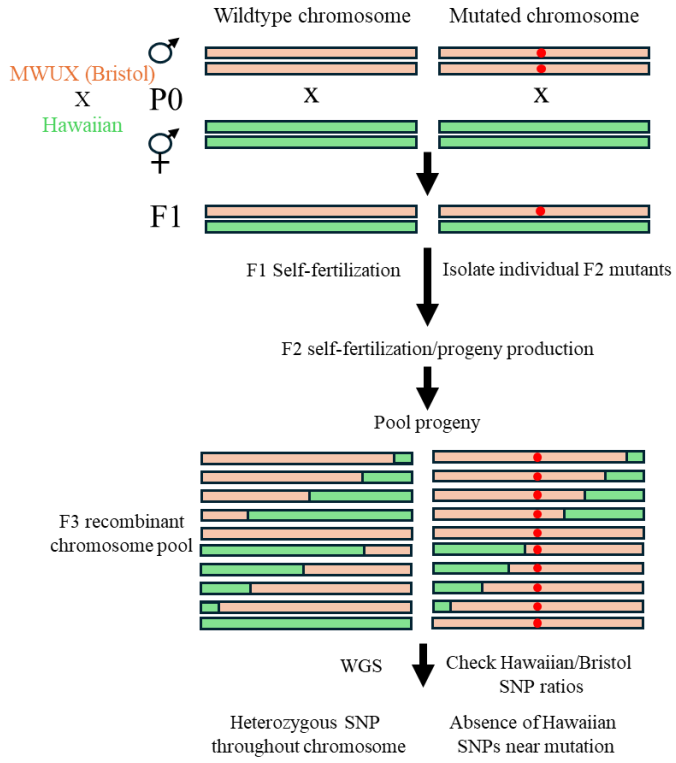
3.4.2 Mutant identification through SNP mapping

To maximize the number of mutants subject to sequencing, we chose to use a streamlined one-step whole genome sequencing and SNP mapping strategy initially developed by (Doitsidou *et al.* 2010), depicted pictorially in **Figure 3.4.2 a-b**. Sequencing space on the flow cell permitted 11 of the 20 EMS mutants to be sequenced and allowed for approximately 20x coverage of each mutant. After SNP mating, the identities of the isolated mutations can be seen in **Figure 3.4.2 c**. A total of 6 separate genes were identified, many of them in duplicate, indicating the screen was approaching saturation. Interestingly, all isolated mutations are components of the RNAi pathway in *C. elegans*. We isolated 3 mutations within the *rde-1* gene that include a premature stop codon, mutation to splice site acceptor, and a nonsynonymous mutation. RDE-1 is a primary Argonaute protein involved in the recognition of primary siRNA produced by Dicer cleavage of dsRNA to form the RNA-induced silencing complex (RISC) in *C. elegans* (Tabara *et al.* 1999). Loss of function mutations to *rde-1* render worms insensitive to

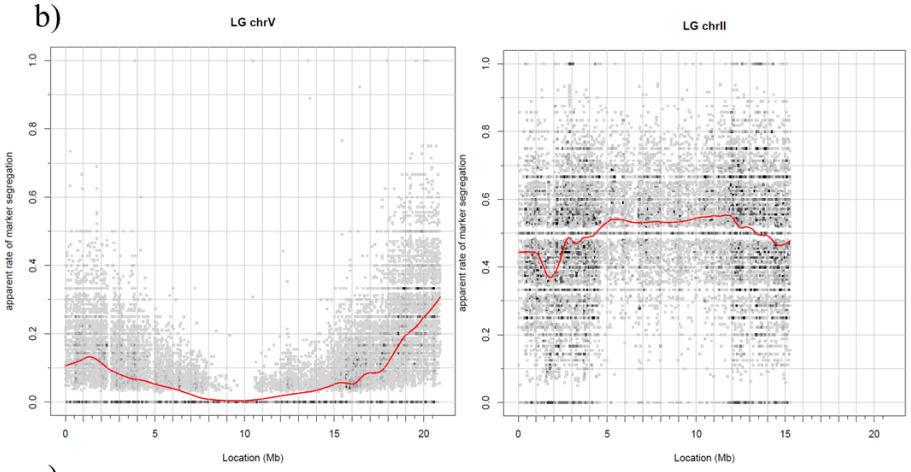
RNAi effects, underscoring its importance in facilitating dsRNA-mediated gene silencing (Tabara *et al.* 1999). Downstream of RDE-1, we also isolated worms with mutations within the secondary siRNA amplification pathway which included non-synonymous or amino acid altering mutations to *rde-11*, *rrf-1*, and *sago-2*, along with premature stop mutation to *rsd-2*. In our *rde-11* mutant, cysteine 241 has been substituted for a tyrosine residue. While this residue resides in a disordered region of the protein, the loss of cysteine removes a thiol group from the overall structure which may participate in dimerization with the RDE-11 complex. The mutation also sits just downstream of the zinc finger domain. The larger size of tyrosine and its aromatic side chain could elicit steric hindrance upon substrate binding or recognition. In our *sago-2* mutant, asparagine was substituted for lysine and this mutation lies outside of a functional domain but is part of an alpha-helical structure. It is possible that this substitution of a charged residue inside the helix could induce conformational changes that hinder protein function.

A complex involving *rde-11* and *rsd-2* is required for the accumulation of primary siRNAs generated by RDE-1 to bridge the steps towards secondary siRNA accumulation that is required to mediate RNAi at low dsRNA concentrations (Zhang *et al.* 2012). Upon recognition by RdRP, the primary siRNA are then amplified by *rrf-1* to produce secondary siRNA, which binds to SAGO-2 secondary Argonaute proteins to exert systemic target inhibition (Zhang *et al.* 2012). We also isolated 4 separate mutations to the *mut-16* gene that include premature stop codon or frameshift mutation. MUT-16 is a worm-specific protein that is a scaffolding protein involved in Mutator foci aggregation that is proposed to function with RRF-1 to facilitate siRNA amplification for gene silencing (Phillips *et al.* 2012). Taken together, these results reveal an unexpected but consistent involvement of the RNAi pathways as a requirement for snRNA processing in *C. elegans.*

a)



b)



c)

Strain	Allele	Chromosome	Gene	Mutation type
MWU44	*cww1*	V	*rde-1*	stop gained (9991213, C to T)
MWU46	*cww2*	IV	*rde-11*	non_synonymous coding (13106331, G to A)
MWU48	*cww3*	I	*mut-16*	stop gained (10084919, C to T)
MWU56	*cww4*	V	*rde-1*	splice acceptor mutation (9989470, C to T)
MWU58	*cww5*	I	*mut-16*	stop gained (10081651, G to A)
MWU60	*cww6*	I	*mut-16*	deletion --> frame shift (10080714, TT to T)
MWU62	*cww7*	I	*rrf-1*	non_synonymous coding (7646247, C to T)
MWU64	*cww8*	I	*sago-2*	non_synonymous coding (618566 C to G)
MWU66	*cww9*	V	*rde-1*	non_synonymous coding (9988753, C to T)
MWU68	*cww10*	I	*mut-16*	stop gained (10083990, C to T)
MWU70	*cww11*	IV	*rsd-2*	stop gained (13529481, C to T)

Figure 3.4.2: Mutant identification through SNP mapping. a) Schematic of the workflow involved in SNP mapping. **b)** Example result of WGS identified Bristol/Hawaiian SNP ratios in strain MWU44. Chromosome V has a low frequency of Hawaiian SNP loci near the region of mutation while chromosome II where no mutations occur has almost equal ratios of SNPs along its length. **c)** Table depicting the mutations identified through SNP mapping.

3.4.3 Validation of isolated mutants for snRNA misprocessing

Of the 6 mutants uncovered, 5 were selected for further analysis (**Figure 3.3.3**). While it was a potentially interesting hit, we found that the mutant carrying the *rrf-1* mutation had exceedingly low levels of GFP biomarker activation after several rounds of outcross. In order to more accurately assess the changes to snRNA processing in these mutants, the snRNA misprocessing reporter was removed through outcrossing with the N2 wildtype, and offspring with only the homozygous mutation was isolated. This is because the biomarker could essentially be considered another locus for snRNA production given that it is comprised of a full promoter, transcript, and 3' UTR that could influence total levels of snRNA molecules and their transcription efficiencies. Upon removing the snRNA misprocessing reporter, analysis with qPCR showed that increases in misprocessed U2 and U4 were still detected. Total snRNA remained near wild-type levels, indicating that post-transcriptional snRNA processing was affected and not total transcription. Aside from *sago-2* and *rsd-2*, the degree of change to U2 and U4 misprocessing in the other 3 mutants was relatively similar to those during *csr-1* knockdown as determined in Chapter 2. qPCR measurements for the 5 mutants can be found in **Figure 3.4.4**.

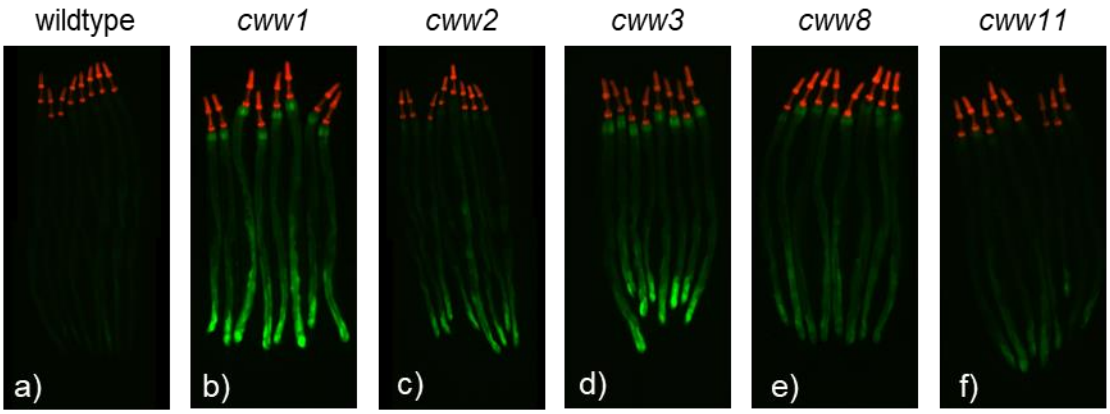


Figure 3.4.3: Hits identified in the genetic screen. EMS-generated mutants showing snRNA misprocessing biomarker activation. The affected gene is indicated under each image as well as the mutation type. 24 young adults were imaged for each strain, 8 worms are represented in each panel. Images were taken on a Ziess Axioskop 50 microscope at 10x magnification.

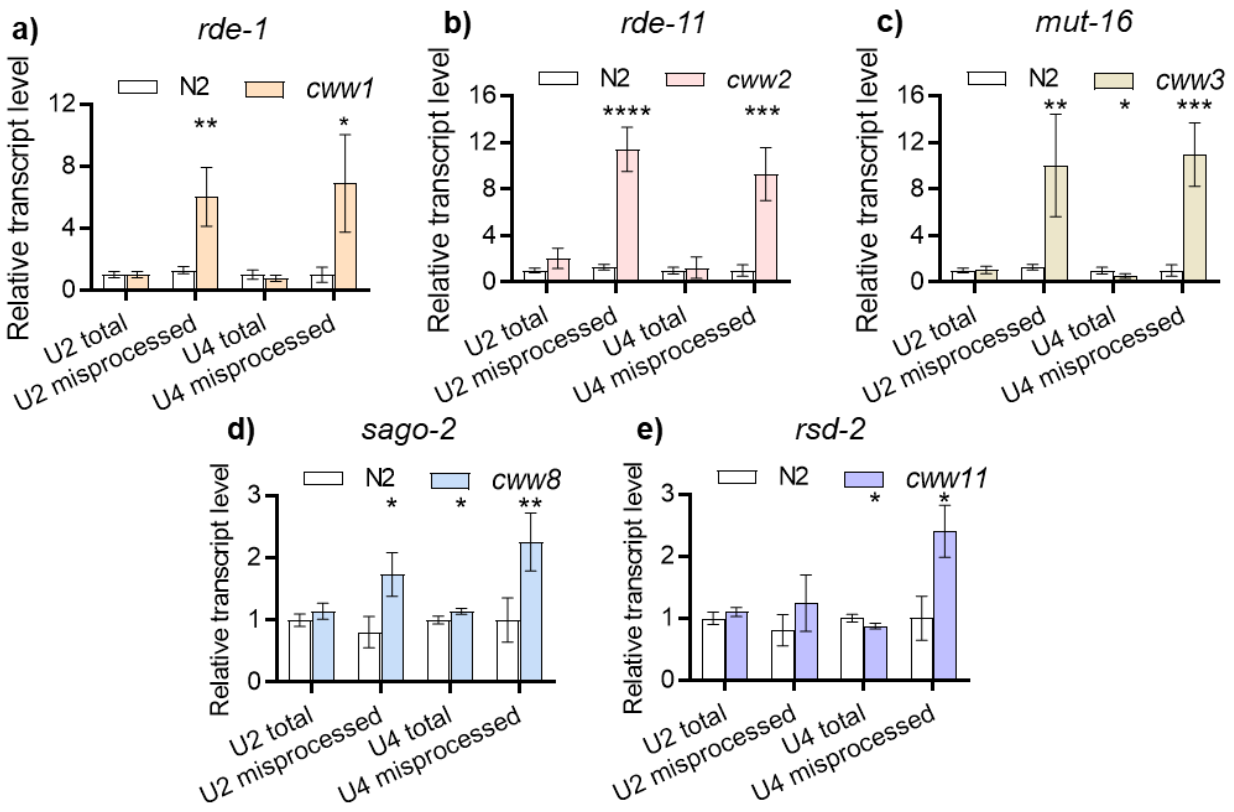


Figure 3.4.4: qPCR confirmation of snRNA misprocessing. Populations of wildtype control and EMS-generated mutants were aged matched via hypochlorite treatment and grown on RNAi media seeded with empty vector RNAi control bacteria for 72 hours. Animals were collected, RNA extracted, and cDNA synthesized before subjecting the samples to qPCR analysis. 4 replicates per strain were performed using approximately 400 animals per extraction and average values were reported. Mutations examined in each panel are **a)** *rde-1*, **b)** *rde-11*, **c)** *mut-16*, **d)** *sago-2*, and **e)** *rsd-2*. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001 as determined by student's t-test compared to N2.

3.4.4 Complementation with verified RNAi mutants

To confirm our mapping strategy accurately identified our mutations, we chose to use a complementation approach with existing mutant worm strains isolated by the *C. elegans*

community that carry characterized loss of function recessive mutations to the same target gene but at a different locus. The classic methods of mutant confirmation often use PCR amplified wildtype alleles of the gene in question to rescue the mutant phenotype. This requires cloning of the promoter region as well as introns, exons, and the 3' untranslated regions. Most of the genes identified in our study have relatively long sequences such as *rsd-2* that is >20 kb which increases the complexity in amplicon generation required for rescue. Given that for all of our 5 genes of interest there exist well-characterized recessive loss of functions alleles, we chose to utilize the complementation strategy. For 4 out of 5 strains, we were able to successfully carry out the mating assay and found that our isolated mutations to *rde-1*, *sago-2*, *mut-16*, and *rsd-2* failed to complement with existing loss of function mutations of these genes. These are illustrated by the constitutive activation of the GFP misprocessing reporter in the F1 offspring of the cross where the worm carries 1 copy of each of the mutated alleles (**Figure 3.4.5**). It should be noted that the F1 offspring also expresses the GFP reporter in the heterozygote, which could explain the reduced fluorescent compared to its homozygous form as illustrated in **Figure 3.4.3**. Interestingly, the *rde-11* mutant we isolated failed to mate with both their complement strain as well as with themselves after several attempts, suggesting that the mutation we isolated may cause infertility in males. Therefore, it was omitted from the complementation assay.

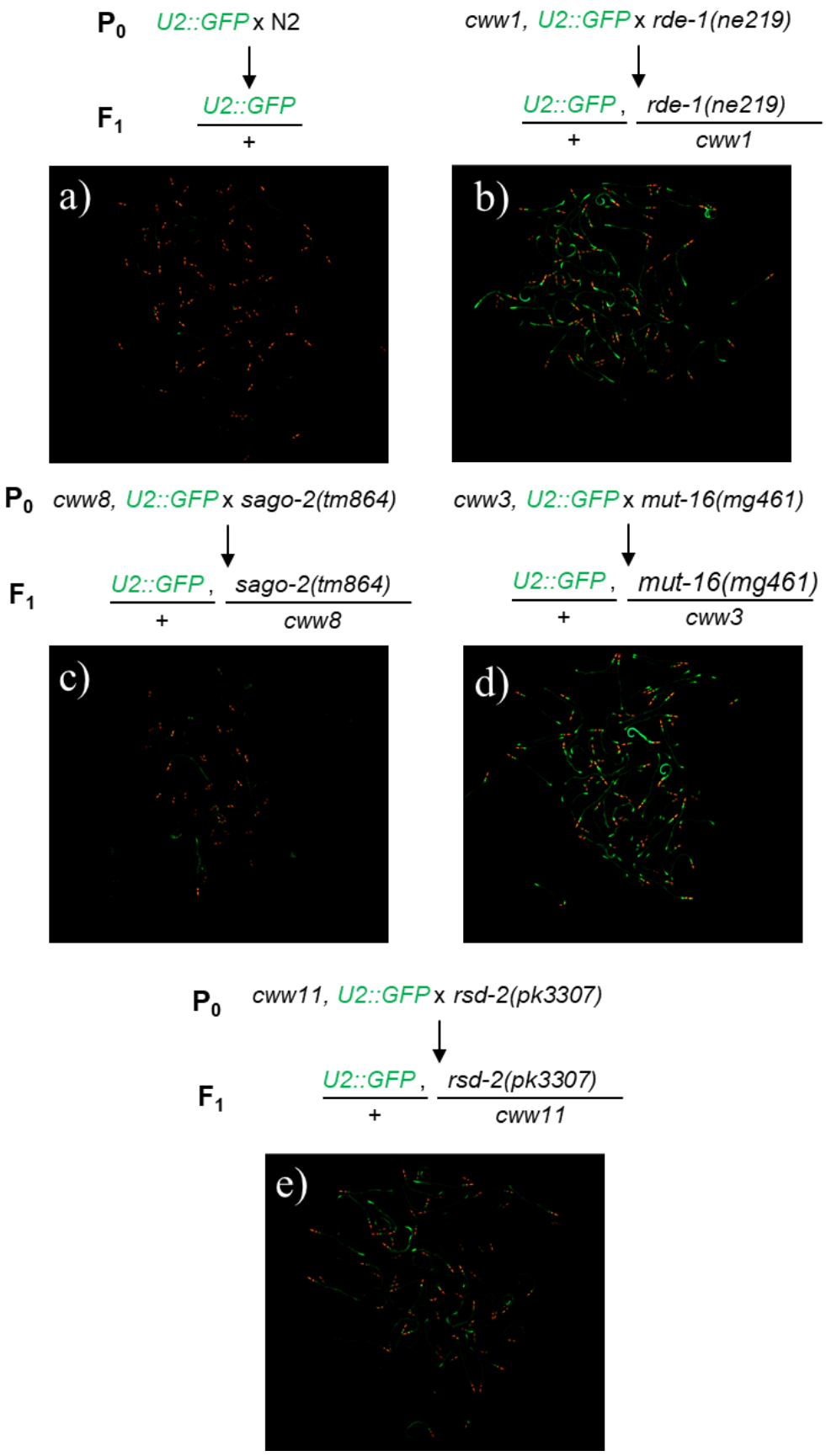

P0 U2::GFP x N2
F1 U2::GFP / +
a)
P0 cww1, U2::GFP x rde-1(ne219)
F1 U2::GFP, rde-1(ne219) / + cww1
b)
P0 cww8, U2::GFP x sago-2(tm864)
F1 U2::GFP, sago-2(tm864) / + cww8
c)
P0 cww3, U2::GFP x mut-16(mg461)
F1 U2::GFP, mut-16(mg461) / + cww3
d)
P0 cww11, U2::GFP x rsd-2(pk3307)
F1 U2::GFP, rsd-2(pk3307) / + cww11
e)

Figure 3.4.5: mutant identification confirmation through complementation. Male populations for the EMS mutant strains were generated and individually mated to their respective complement strain. Control animals with no mutations were mated to wild-type N2. The F1 generation was identified by the presence of the RFP co-injection marker associated with the misprocessing biomarker. Approximately 100 mated F1 progeny were imaged per complementation. Images taken on an Agilent BioTek Cytation 5 Cell Imaging Multi-Mode Reader at 4x magnification.

3.4.6 qPCR of complementation parents for snRNA misprocessing

To further illustrate that these mutations cause snRNA misprocessing, we performed qPCR on the previously isolated mutant alleles that were used for the complementation assay which were *rde-1(ne219), sago-2(tm864)*, *mut-16(mg461)*, and *rsd-2(pk3307).* As shown in **Figure 3.4.6a-e**, these alleles also cause varying degrees of snRNA misprocessing. Interestingly, significant changes in snRNA misprocessing were mostly observed in the U4 transcript rather than the U2 transcript. We also performed qPCR analysis with known mutations to *rrf-1* and *rde-11* that were not included in the complementation assay due to low fluorescence and mating infertility respectively and found that both of these mutation alleles resulted in a significant increase in U2 and U4 misprocessing. This variation in endogenous snRNA misprocessing may be caused by different degrees of mutation penetrance on gene function in separate loci of a given gene. The *rsd-2* allele isolated from our EMS screen formed a premature stop codon resulting in a loss of function while *rsd-2(pk3307)* carries a substitution to arginine 587, a large positively charged residue, to glycine. This mutation occurs between two β-sheets in *rsd-2(pk3307)* and may only partially impact protein function resulting in a lower degree of snRNA misprocessing. Our *rde-1* allele also produces a premature stop codon resulting in a loss of function. This is compared to the complement allele *rde-1(ne219)* that carries a substitution of glutamate 413, a negatively charged residue, with a positively charged lysine. This mutation occurs within a β-sheet of the PAZ domain and could affect the recognition of RNA substrates.

Our *mut-16* allele produces a premature stop codon. The complement allele includes a 451 bp deletion ~500 bp upstream of the start site. This changes the promoter and not the coding sequence, which may only decrease expression and allow low levels of transcription. This is reflected in **Figure 3.4.6 f**, as the complement allele did not experience any significant

misprocessing. Our *rde-11* allele contains a substitution of cysteine 241 with tyrosine and occurs just downstream of the zinc finger domain. The removal of the cysteine residue could impact dimerization with other components, and its placement near the zinc finger site could affect substrate recognition. The complement allele produces a premature stop codon resulting in a loss of function. Finally, our *sago-2* allele contains a substitution of asparagine, an uncharged residue, with a positively charged lysine. This mutation occurs outside of a functional domain but inside an α-helix. The complement allele contains a splice site acceptor variant that may only also have minor effects on protein function. As such, both our allele and the complement allele may not generate a complete loss of function, and this is reflected by the low levels of misprocessing experienced during both conditions (**Figure 3.4.4 d** and **Figure 3.4.6 d**).

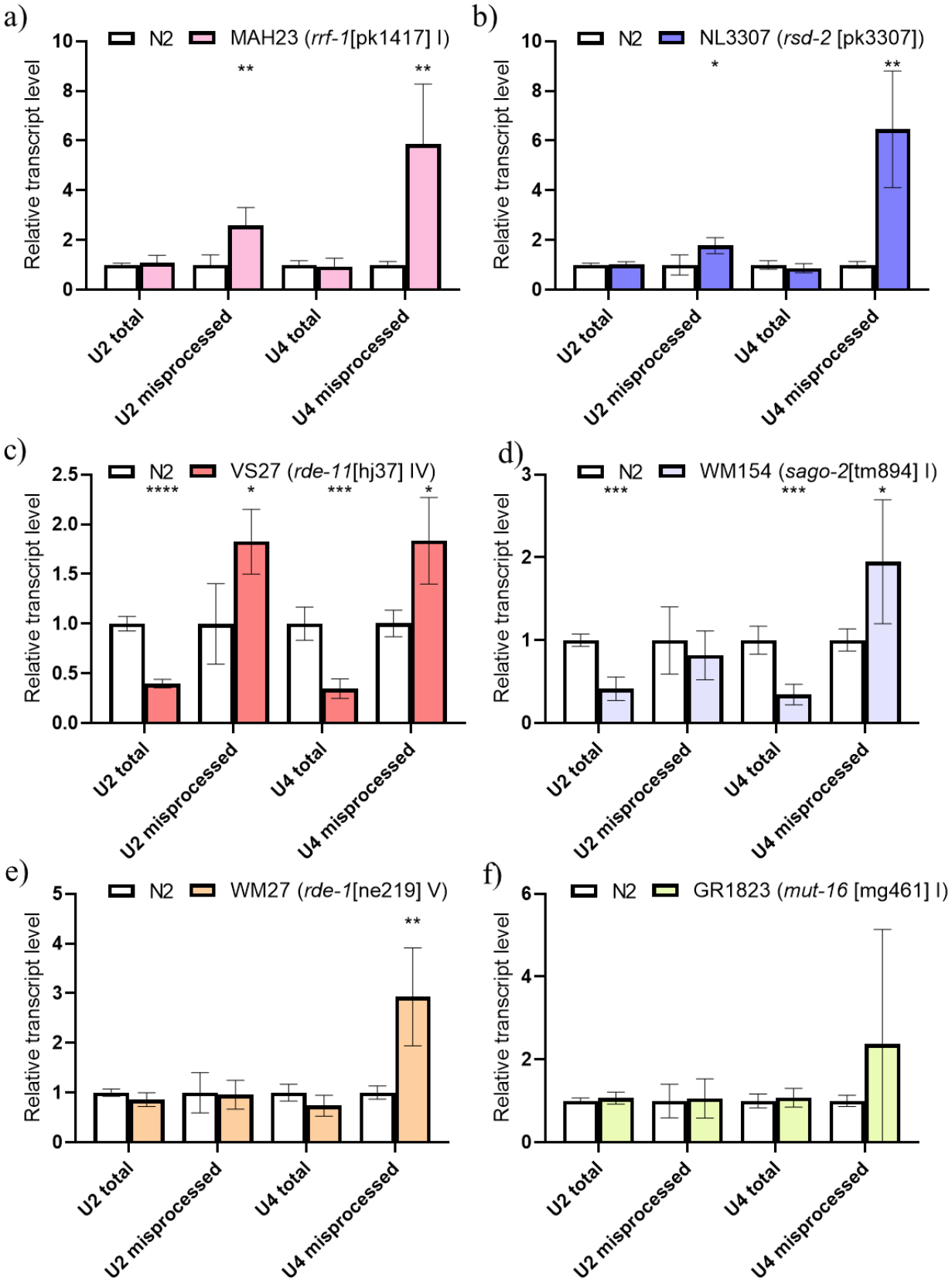

a)
N2
MAH23 (rrf-1[pk1417] I)
Relative transcript level
U2 total
U2 misprocessed
U4 total
U4 misprocessed
b)
N2
NL3307 (rsd-2 [pk3307])
c)
N2
VS27 (rde-11[hj37] IV)
d)
N2
WM154 (sago-2[tm894] I)
e)
N2
WM27 (rde-1[ne219] V)
f)
N2
GR1823 (mut-16 [mg461] I)

Figure 3.4.6: qPCR of complementation parents for snRNA misprocessing. qPCR results examining the differences between total and misprocessed levels of U2 and U4 snRNA between N2 and **a)** MAH23 (*rrf-1*[*pk1417*] I), **b)** NL3307 (*rsd-2*[*pk3307*]), **c)** VS27 (*rde-1*[*hj37*] IV), **d)** WM154 (*sago-2*[*tm894*] I), **e)** WM27 (*rde-1*[*ne219*] V), and **f)** GR1823 (*mut-16*[*mg461*] I). Populations of each respective strain were age-matched via hypochlorite treatment and grown on RNAi media seeded with empty vector bacteria for 72 hours before RNA extraction and cDNA synthesis. RNA from a total of 4 replicates per strain were extracted with each containing approximately 400 worms. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001 as determined by student's t-test compared to N2 wildtype.

3.5 Discussion

The regulation of gene expression via post-transcriptional splicing is paramount to transcriptome stability and snRNA transcripts sit at a crux in this process (Zahler 2012). Once transcribed, snRNA transcripts must undergo 3' cleavage for maturation and subsequent incorporation into the spliceosome. Here we have identified several new gene regulators of snRNA 3' processing aside from the well-described Integrator complex (Baillat *et al.* 2005). This study identified several key components of the RNAi pathway that when mutated to confer a loss of function cause increased snRNA misprocessing. While it is reasonable to speculate that defects in the RNAi machinery may affect Integrator expression and function, mutation or knockdown to Integrator subunits critical for snRNA processing results in severe developmental defects that were not recapitulated in the mutants isolated in this study. In fact, no discernable phenotype was observed for any of our mutants with respect to development. However, the degree of snRNA misprocessing triggered by mutations to the RNAi pathway is modest compared to direct Integrator depletion (Waddell and Wu 2024), which could be a contributing factor in the lack of a clear developmental defect.

A possible mechanism through which defects in the RNAi pathway could cause snRNA misprocessing would be a direct effect on Integrator gene expression. As the endogenous RNAi pathway functions to influence protein expression, defects in this mechanism could result in altered levels of Integrator expression. This could leave separate modules of Integrator with enough function to perform their accessory roles in other processes but impair its snRNA processivity. Alternatively, since many of the proteins identified here have RNA binding properties, there is also a potential role for direct interaction with the snRNA transcripts. For example, the Argonaute encoding *rde-1* is essential for both somatic and germline RNAi and

binds directly to 26G dsRNA *dcr-1* products to catalyze the removal of the passenger strand to produce ssRNA that targets mRNA for further degradation (Tabara *et al.* 1999; Grishok 2005). In addition to RDE-1, RRF-1 and SAGO-2 are both RNA-binding proteins that can bind to ssRNA and snRNA transcripts contain regions of complementarity that can pair with each other and form potential dsRNA (Thomas *et al.* 1990; Yigit *et al.* 2006). As such these factors could lead to potential direct interactions with RNAi-binding proteins to interact directly with snRNA transcripts to aid in its 3' processing. Furthermore, as snRNAs need to be exported from the nucleus for incorporation into the spliceosome, these proteins could also bind and facilitate their transport or recruit additional cofactors involved in this process.

This EMS screen also identified several mutated alleles of *mut-16*, which serve as an essential scaffolding component of perinuclear Mutator foci compartments that serve as sites of 22G siRNA amplification (Phillips *et al.* 2012). These foci are involved in secondary siRNA amplification and are consistent with a role for the RNAi pathway in snRNA processing. Loss of function to *mut-16* has also been shown to deplete CSR-1 class siRNAs by 90%, which we have previously shown to be required for the germline expression of Integrator subunits (Zhang *et al.* 2011; Waddell and Wu 2024). As such, snRNA misprocessing caused by mutations to *mut-16* observed in this study could be influenced by its effect on CSR-1 siRNA expression.

The EMS screen employed in this study identified several non-essential loss of function mutations that caused snRNA misprocessing, which is in contrast with our previous RNAi screen study using the same misprocessing reporter that identified almost exclusively essential genes that when knocked down by RNAi-induced snRNA misprocessing (Waddell and Wu 2024). We identified 6 non-essential gene mutations using the EMS screen compared to almost 50 essential genes in the RNAi screen (**Figure 2.4.2 Chapter 2**), underscoring the importance of snRNA

processing as an essential process for organismal development. There is little overlap between the two genetic screens with only *mut-16* identified in both experiments. This could be due to the fact that the EMS screen exclusively identified genes functioning in the RNAi pathway, for which these genes would be refractory to RNAi and cannot be reliably depleted using dsRNA feeding that would be dependent on a functional RNAi pathway.

A potential limitation of this study is the number of snRNA transcripts we measured using qPCR to determine misprocessing. In *C. elegans* 72 snRNA loci have been identified to date, including 19 separate transcripts encoding the U2 snRNA (Thomas *et al.* 1990; Gómez-Orte *et al.* 2019). These transcripts are dispersed across chromosomes in varying locations and may well be regulated independently based on the gene loci. This may be a source of bias if the C47F8.9 (U2) or K03B8.10 (U4) transcripts we measured are not regulated in the same fashion as the other snRNA transcripts. A possible alternative is to perform RNA sequencing to capture a more complete description of global snRNA processing defects. As many of the genes identified in this study participate in the RNAi response, it would be beneficial to examine any epistatic effects present between these components. This could help pinpoint critical steps within the RNAi pathway that influence snRNA processing.

3.6 Conclusion

In this study we identified several components of the RNAi pathway as novel regulators of snRNA processing in *C. elegans*. While the mechanism of this regulation is still unknown, we have provided speculation into multiple avenues of how this pathway may be regulated. As the genes identified in this study function at different levels of the RNAi response, it will be interesting to tease this pathway apart and pinpoint the steps critical for snRNA processing. As the roles of the Integrator complex have grown to encompass a wide range of transcriptome

control beyond snRNA processing and is emerging as a target of various human diseases, it will be beneficial to determine if these novel regulators within the RNAi pathway also affect other aspects of Integrator function beyond snRNA maturation (Oegema *et al.* 2017; Krall *et al.* 2019; Azuma *et al.* 2023; Tepe *et al.* 2023).

TRANSITION

The following chapter focuses on Objective 3 of my hypothesis, the requirements of Integrator regulators on aging. My findings conclude that mutations to several regulators of Integrator mediated snRNA processing exert deleterious effects on overall lifespan. One gene has been extensively characterized in terms of isoform and catalytic requirements.

Publication: Involvement of snRNA processing regulatory genes in aging (to be submitted)

Contributions: BMW conducted all of the study. Manuscript was drafted by BMW and edited by CWW (Cheng-Wei-Wu).

Chapter 4

Involvement of snRNA processing regulatory genes in aging

4.1 Abstract

Aging is a natural progression that affects all living organisms. While the consequences of aging have been increasingly understood in the context of disease progression, there is much debate about its cause. Recently, defects in RNA splicing have been linked directly to aging, implicating the importance of this post-transcriptional process in maintaining organismal homeostasis. Here we use the *C. elegans* model to investigate the involvement of snRNA processing regulators in aging. We found that depletion of the snRNA principal regulator Integrator complex via auxin-induced degradation of catalytic subunits-4 during early larval development drastically reduced lifespan. However, when depletion was performed during adulthood, lifespan was unaffected. We then characterized a number of genes we previously identified to be required for snRNA processing to determine their role in aging. We found that mutations to *csr-1* that cause snRNA misprocessing reduced *C. elegans* lifespan and was dependent on its catalytic activity. Interestingly, loss of *csr-1a* did not affect lifespan, suggesting a potential isoform-specific role. In addition, knockdown of several genes encoding nuclear pore proteins (*npp-1*, *3*, and *6*) impaired snRNA processing and also led to significant lifespan reduction. Together, this study has identified several snRNA processing regulators as essential genes in maintaining normal lifespan in *C. elegans.*

4.2 Introduction

The emergence of introns in the genome defines an important evolutionary step between prokaryotes and eukaryotes (Olthof et al., 2022). In vertebrates, this increase in intron content may have contributed to organismal complexity (Olthof et al., 2022). An indication of this increased complexity is highlighted by the presence of alternatively spliced transcripts that can drastically affect protein structure, function, and expression, (Chen & Manley, 2009). This alternative splicing (AS) allows an organism to effectively expand its proteome without proportionally increasing its genomic content. AS has been critically implicated in various cellular processes including cell cycle progression and cell lineage specification and differentiation (Baralle & Giudice, 2017; Olthof et al., 2022).

Aging has been characterized as the time-dependent decline of tissue and organ function (Li et al., 2017). More than 90% of human genes undergo AS, and as such AS may play an important role in aging (Li et al., 2017). This is an idea supported by the multitude of diseases that appear when AS is altered. Recent estimates suggest that 60% of all disease-causing mutations disrupt splicing (Wang & Cooper, 2007). Perturbations to cis and trans elements of splicing can both lead to disease. Mutations to core spliceosome components have been linked with Cerebro-Costo-Mandibular Syndrome characterized by facial feature abnormalities, rib/chest cavity developmental deficiencies, and difficulty breathing/eating (Jiang & Chen, 2021; Wang & Cooper, 2007). Myelodysplastic syndromes have also been associated with alterations to trans-splicing factors (Jiang & Chen, 2021). Alterations to cis-acting sequences have been linked to diseases such as cystic fibrosis, where even synonymous nucleotide substitutions have been implicated to impact splicing (Wang & Cooper, 2007). Changes to AS can also lead to

incorrect ratios of protein isoforms, inducing tauopathies such as Alzheimer's, Parkinson's, and Huntington's disease (Jiang & Chen, 2021).

The Integrator complex was initially described as a potent regulator and processing factor for proper snRNA transcription (Baillat et al., 2005). Recently, new roles for the Integrator have been described that extend beyond snRNA processing. The release of paused RNA Pol II during early transcription to attenuate non-productive transcription has been attributed to the Integrator complex (Elrod et al., 2019). Much like splicing, this process contributes to overall transcriptome integrity. While the genetic screens employed in this study focused on factors that promote snRNA transcription, it is not unreasonable to conclude that these factors could also influence other Integrator functions. As the catalytic activity of Integrator has been implicated in this role of RNA Pol II pause/release and is also required for snRNA processing, there may be an overlap between factors that control both of these processes (Baillat et al., 2005; Elrod et al., 2019).

In this study, we examined the requirement of Integrator in aging in a tissue and temporal manner through the use of the auxin degradation system. We also characterized several regulators of Integrator function on aging identified previously in our genome-wide RNAi screen (**Figure 2.4.2 Chapter 2**). In Chapter 2 we found that knockdown or depletion of Integrator causes widespread snRNA misprocessing during both larval development and adulthood. Here we report that depletion of Integrator during only larval development has a deleterious effect on lifespan and is dispensable for lifespan maintenance in adulthood. Knockdown of Integrator snRNA processing regulators also causes decreases in lifespan. Overall, this study implicates the importance of maintaining snRNA processing in a normal lifespan.

4.3 Materials and Methods

4.3.1 *C. elegans* strains

All *C. elegans* strains were cultured at 20°C using standard methods unless noted otherwise (Brenner 1974). The following strains were used: N2 Bristol wild type, MWU193 (*cwwSi1[ints-4::mKATE2::AID*::3xFLAG]; wrdSi23 [eft-3p::TIR1::F2A::mTagBFP2::AID*::NLS::tbb-2 3'UTR] (I:-5.32)*), MWU212 (*cwwSi1[ints-4::mKATE2::AID*::3xFLAG]; wrdSi3 [sun-1p::TIR1::F2A::mTagBFP2::AID*::NLS::tbb-2 3'UTR] (II:0.77)*), WM182 (*csr-1(tm892) IV/nT1 [unci-?(n754) let-?] (IV;V)*), USC1258 (*csr-1a(cmp135) IV*), OD923 (*ltSi240 [csr-1p::csr-1(re-encoded) + Cbr-unc-119(+)] II*), and OD925 (*ltSi242 [csr-1p::csr-1(re-encoded; D606A, D681A: isoform b numbering) + Cbr-unc-119(+)] II*).

4.3.2 Life span analysis

Worms were aged matched via hypochlorite treatment and approximately 180 worms plated across 3 plates. RNAi media and empty vector control bacteria were used for all assays unless otherwise noted. Worms were manually picked every day to fresh plates for the first 5 days, then moved to fresh plates every other day until day 10. Worms were scored as dead once they failed to respond to mechanical stimulus. Animals with vulval protrusions or that desiccated on the sidewalls were censored out. 3 trials were performed per condition. Life span assays were conducted at 20 ℃ or 25 ℃ as indicated in the figure legends. For auxin exposure assays, 1 mM of auxin was used in all experiments. Bacteria on auxin media was seeded as a 2x concentrated solution as auxin interferes with bacterial growth. Auxin stock solution was made in 100%

ethanol and filter sterilized, made fresh for each batch of media. Auxin exposure began at L1 or L4 larval stage depending on the assay. Ethanol was used as the vehicle control in control media.

4.3.3 Statistical analysis

GraphPad Prism (V 8.4.3) was used to generate the data graphs. Mean lifespan was calculated for each curve and the Log-rank tests were performed using the OASIS 2 software to determine statistical significance between groups.

4.4 Results

4.4.1 Integrator effects on lifespan

Disruptions to the metazoan Integrator complex result in embryonic lethality (Hata and Nakayama 2007). Therefore to study this complex we employed 2 separate knockdown strategies to examine the effects of Integrator depletion on lifespan. Initially, the catalytic *ints-4* subunit was targeted for degradation via RNAi during early larval development. As can be seen in **Figure 4.4.1 a**, RNAi depletion of *ints-4* resulted in a severe decrease in total lifespan. N2 worms were observed to live on average 20.92 days, whereas worms depleted of *ints-4* only survived an average of 5.77 days. Aside from RNAi, the auxin-inducible degradation (AID) system was also used to deplete INTS-4 protein from worms. The AID sequence from the IAA17 plant protein is fused to the protein of interest and requires the plant Transport Inhibitor Response 1 (TIR1) protein for auxin recongition (Nishimura *et al.* 2009). Exposure to auxin then causes TIR1 to bind and promote ubiquitylation of the AID-tagged protein and its subsequent degradation (Ashley *et al.* 2021). This system is attractive as it allows for expression of TIR in different tissues through the use of specific promoters to induce spatial target degradation that

can be achieved in as little as 30 minutes of auxin exposure and in a reversible manner (Zhang *et al.* 2015; Ashley *et al.* 2021).

We found that the AID strain generated in this study experienced a high degree of vulval protrusions during the lifespan assay when performed at 20°C, resulting in a large number of censored subjects. To alleviate this issue, lifespan analysis of all AID strains was performed at 25 ℃ as vulval protrusions are less prominent at this temperature (Leiser *et al.* 2016). Using the AID system, we found that INTS-4 depletion resulted in similar decreases to overall lifespan as can be seen in **Figure 4.4.1 b**. Worms expressing AID tagged *ints-4* had a considerable decrease in their average lifespan when exposed to auxin (7.57 days) versus the ethanol control (13.43 days). N2 worms exposed to the ethanol control or auxin had similar average lifespans, 13.15 days and 13.4 days respectively (**Figure 4.4.1c**).

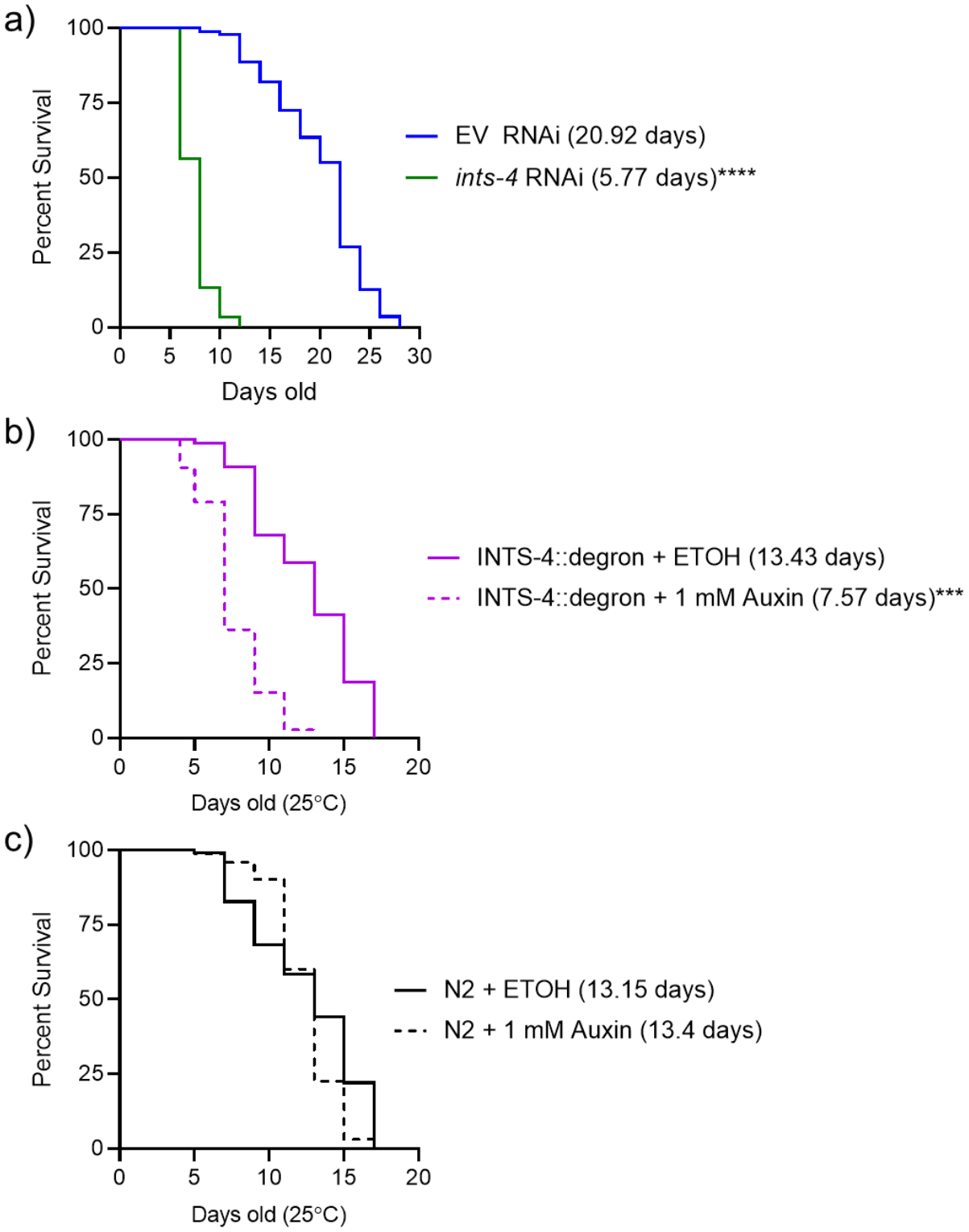

a)
Percent Survival
100
75
50
25
0
0 5 10 15 20 25 30
Days old
EV RNAi (20.92 days)
ints-4 RNAi (5.77 days)****
b)
Percent Survival
100
75
50
25
0
0 5 10 15 20
Days old (25°C)
INTS-4::degron + ETOH (13.43 days)
INTS-4::degron + 1 mM Auxin (7.57 days)***
c)
Percent Survival
100
75
50
25
0
0 5 10 15 20
Days old (25°C)
N2 + ETOH (13.15 days)
N2 + 1 mM Auxin (13.4 days)

Figure 4.4.1: Integrator effects on lifespan. a) Lifespan effects of Integrator knockdown via RNAi. b) Lifespan effects of auxin-induced Integrator depletion in the INTS-4::degron strain. c) Effects of 1 mM auxin exposure on the N2 wildtype. *P<0.05, ****P<0.0001 as determined by the log-rank test compared to N2 (panel a) or ETOH control (panel b). All worms were age-matched via hypochlorite treatment and RNAi or auxin exposure beginning during L1 larval stage. All animals were fed empty vector RNAi control bacteria unless otherwise noted. All auxin concentrations were 1 mM and average lifespan are indicated in parentheses.

To assay for critical development time points that require Integrator, we then used the AID system to knockdown *ints-4* during different age stages. We also examined the germline-specific requirements of Integrator on average lifespan given that we previously identified CSR-1 as a germline specific regulator of Integrator function (Waddell and Wu 2024). N2 worms experienced little change to their average lifespan during auxin exposure. At 25 ℃ N2 worms lived for 12.49 days on average, while L1 and L4 auxin exposure had average lifespans of 12.41 days and 12.06 days respectively and can be seen in **Figure 4.4.2 a**. While the degree of change was small for L4 exposure, it was still found to be significant. First, we examined the effect of somatic Integrator knockdown at L1 and L4 in worms expressing the INTS-4::degron tag in addition to TIR driven by the *eft-3* promoter. Worms treated with ethanol lived for an average of 12.52 days while L1 and L4 exposed worms lived an average of 7.57 days and 13.12 days respectively (**Figure 4.4.2b**). This indicated that depletion of INTS-4 during development shortened lifespan but had no effect when the protein was degraded once the worms reached adulthood. Next, we examined the germline requirements of Integrator at different developmental time points using the germline-specific *sun-1* promoter to drive TIR1 expression

in the germline. Depletion of germline INTS-4 resulted in negligible changes to lifespan, as can be seen in **Figure 4.4.2 c**. ETOH control worms lived for an average of 12.86 days while L1 and L4 auxin-exposed worms lived for 13.44 days and 12.68 days respectively. Taken together these results indicate that depletion of Integrator in somatic cells during early development is deleterious to lifespan.

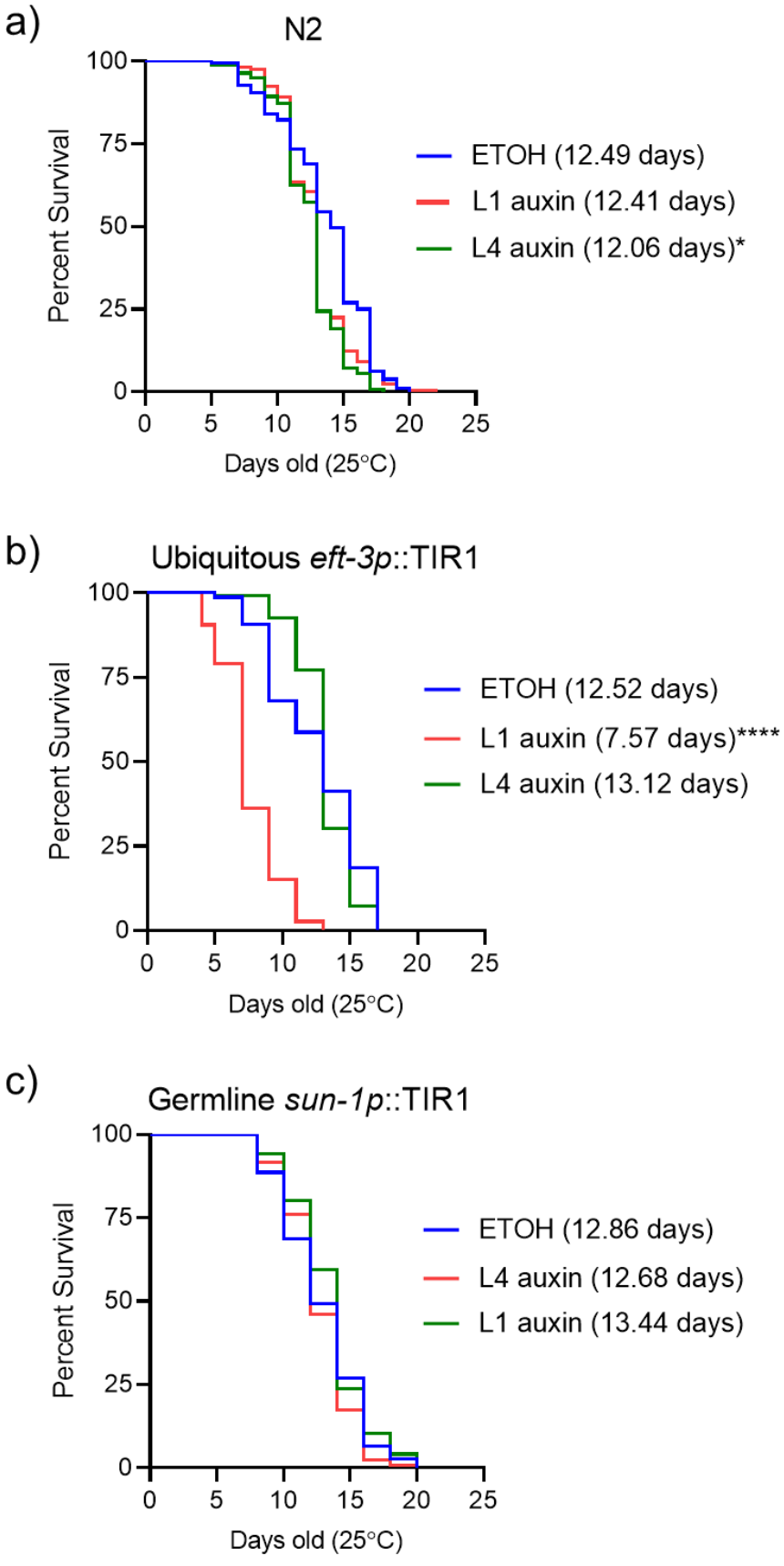

a)
N2
Percent Survival
Days old (25°C)
ETOH (12.49 days)
L1 auxin (12.41 days)
L4 auxin (12.06 days)*
b)
Ubiquitous eft-3p::TIR1
Percent Survival
Days old (25°C)
ETOH (12.52 days)
L1 auxin (7.57 days)****
L4 auxin (13.12 days)
c)
Germline sun-1p::TIR1
Percent Survival
Days old (25°C)
ETOH (12.86 days)
L4 auxin (12.68 days)
L1 auxin (13.44 days)

Figure 4.4.2: Tissue and age-specific effects of Integrator depletion. a) Effects of delayed auxin exposure on N2 worms. b) Effects of delayed auxin exposure on worms expressing AID tagged *ints-4* with TIR1 expressed under the control of the *eft-3* promoter to enable somatic depletion. c) Effects of delayed auxin exposure on worms expressing AID tagged *ints-4* with TIR1 under the control of the *sun-1* promoter to enable germline-specific depletion. *P<0.05, ****P<0.0001 as determined by the log-rank test compared to the ETOH control. All worms were age-matched via hypochlorite treatment and RNAi beginning during L1 larval stage. All animals were fed empty vector RNAi control bacteria. All auxin concentrations were 1 mM. Average lifespan indicated in parentheses. Auxin exposure lifespan assays were performed at 25 °C.

4.4.2 *csr-1* effects on lifespan

Through a genome-wide RNAi screen using a snRNA misprocessing reporter, we discovered that disruption to *csr-1* encoding an Argonaute protein leads to widespread snRNA misprocessing (**Figure 2.4.1 g Chapter 2)**. We then sought to examine the effects of *csr-1* depletion on lifespan. Much like Integrator depletion, null mutants for *csr-1* experience a significant decrease in lifespan. As can be seen in **Figure 4.4.3 a** N2 worms had an average lifespan of 20.92 days while the *csr-1(tm892)* mutant had an average lifespan was 13.4 days. As *csr-1(tm892)* causes deletion to both isoform A and B, we analyzed the effects of another mutant that specifically deletes isoform A. Comparison of N2 worms with a *csr-1a(cmp135)* null mutant revealed little difference between their average lifespans as can be seen in **Figure 4.4.3 b**. N2 worms lived for 17.89 days on average while the *csr-1a* mutant lived for an average of 16.95 days. While this suggests *csr-1b* isoform is likely to be responsible for the decreased lifespan in

the *csr-1(tm892*) mutant, it is also possible that the decrease observed is contributed by the simultaneous loss to both isoforms.

To further examine the effects of the CSR-1b catalytic PIWI domain on lifespan, we utilized a pair of transgenic strains of worm harbouring a single copy insertion of a genetically re-encoded *csr-1b* locus that is wildtype (CSR-1^{WT}) or contains mutation to the catalytic domains (CSR-1^{SIN}). This single copy inserted *csr-1b* also contains a re-encoded region that is resistant to *csr-1* RNAi to permit RNAi-mediated knockdown of endogenous *csr-1* without affecting the single copy transgene (Gerson-Gurwitz *et al.* 2016; Waddell and Wu 2024). In this way, RNAi targeting the re-encoded region can be used to deplete the endogenous *csr-1a* and *csr-1b* transcripts while leaving the single copy transgene unaffected.

RNAi knockdown of endogenous *csr-1* caused a 10% decrease in the *csr-*1^{WT} lifespan compared to EV (EV lifespan of 17.7 days vs *csr-1* RNAi lifespan of 15.94 days), suggesting that single copy insertion of *csr-1* may not be sufficient to fully overcome depletion of endogenous *csr-1* (**Figure 4.4.3c**). Interestingly, RNAi knockdown of *csr-1* in the *csr-*1^{SIN} strain caused a 26% decrease in lifespan compared to EV (EV lifespan of 18.69 days vs *csr-1* RNAi lifespan of 13.94 days) (**Figure 4.4.3c**). The enhanced reduction of lifespan after endogenous *csr-1* knockdown in the *csr-*1^{SIN} strain indicates that single copy rescue with a catalytically dead *csr-1* is not sufficient in restoring lifespan to the effects observed for *csr-*1^{WT}.

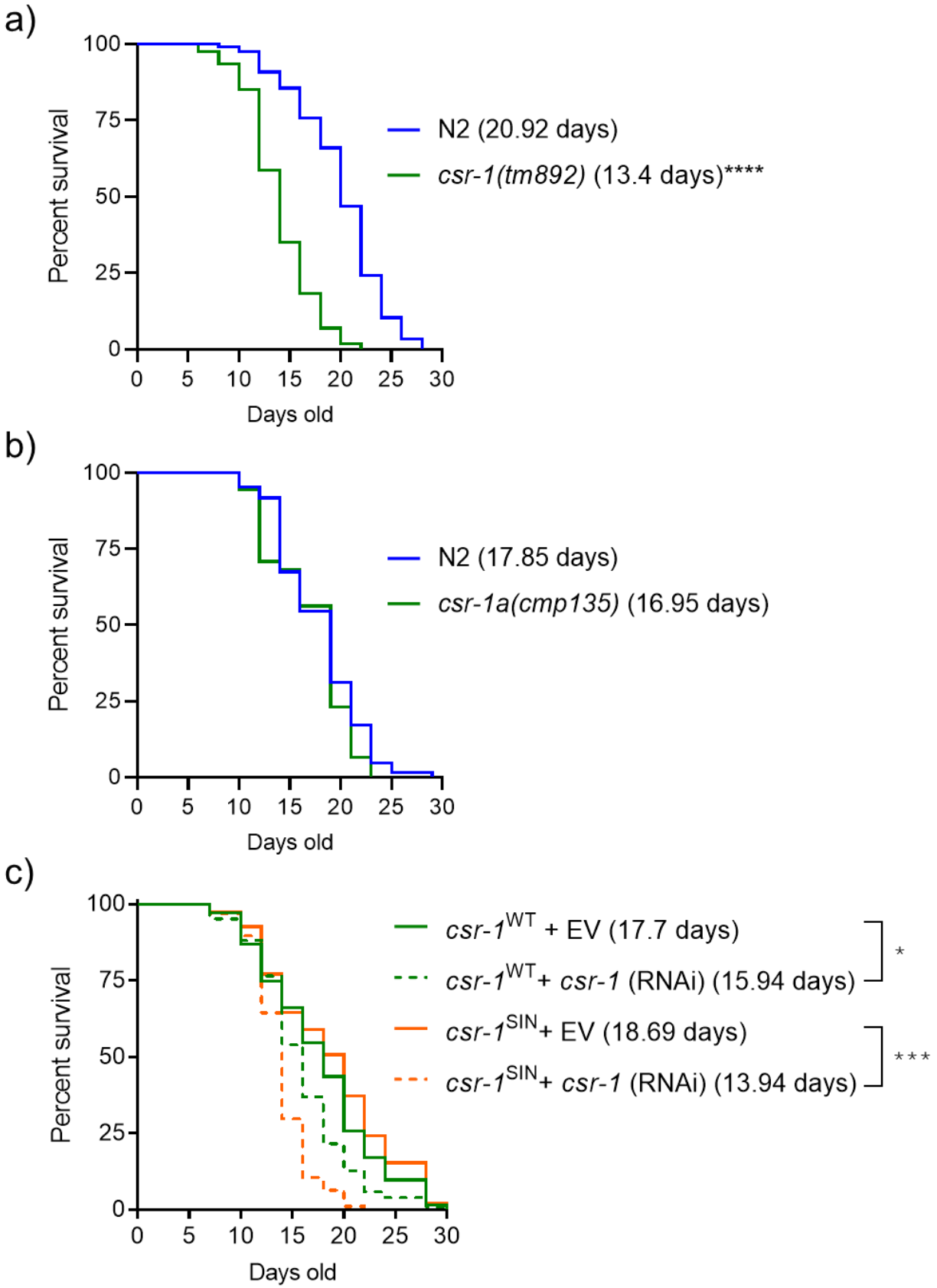

a)
Percent survival
100
75
50
25
0
0 5 10 15 20 25 30
Days old
N2 (20.92 days)
csr-1(tm892) (13.4 days)****
b)
Percent survival
100
75
50
25
0
0 5 10 15 20 25 30
Days old
N2 (17.85 days)
csr-1a(cmp135) (16.95 days)
c)
Percent survival
100
75
50
25
0
0 5 10 15 20 25 30
Days old
csr-1WT + EV (17.7 days)
csr-1WT+ csr-1 (RNAi) (15.94 days)
csr-1SIN+ EV (18.69 days)
csr-1SIN+ csr-1 (RNAi) (13.94 days)
*
***

Figure 4.4.3: *csr-1* effects on lifespan. All worms were age-matched via hypochlorite treatment and RNAi exposure beginning during L1 larval stage. All animals were fed empty vector RNAi control bacteria unless otherwise noted. a) Comparison of lifespan between N2 wildtype and a null *csr-1* mutant absent for both isoforms (WM182; *csr-1(tm892)*). b) Comparison of lifespan between N2 wildtype and a *csr-1* isoform A mutant (USC1258; *csr-1a(cmp135)*). c) Requirements of the catalytic PIWI domain of *csr-1* isoform B on lifespan. Average lifespan is indicated in parentheses. *P<0.05, ****P<0.0001 compared to N2 wildtype in a or to the respective EV fed condition in panel c as determined by the log-rank test.

4.4.3 Nucleoporin protein's effects on lifespan

Amongst the group of genes required for snRNA processing in *C. elegans* are those encoding multiple components of the nuclear pore complex protein (*npp*) (**Figure 2.4.2 Chapter 2**). We next sought to examine whether knockdown of *npp* genes that cause snRNA misprocessing influences lifespan. As can be seen in **Figure 4.4.4**, RNAi knockdown of 3 out of 4 *npp* encoding genes caused a decrease in lifespan. Depletion of *npp-1* or *npp-3* resulted in a modest decrease in average lifespan to 12.54 and 12.24 days respectively compared to 18.41 days in the EV-fed control group. NPP-1 and NPP-3 proteins function in proximity within the nuclear pore complex with NPP-1 localized to the central channel through which cargo is transported into and out of the nucleus while NPP-3 is constrained within the inner ring that supports the central channel (Cohen-Fix and Askjaer 2017). Loss of *npp-6* resulted in the highest degree of lifespan decrease to 9.4 days on average compared to the 18.41 of EV control-fed worms. NPP-6 is part of the ring structure that is exposed to both cytoplasmic and nucleoplasmic compartments of the cell (Cohen-Fix and Askjaer 2017). NPP-6 encodes the mammalian Nup-

160 that has a critical function in the assembly of nuclear pore complex and its depletion causes reduced expression of other nucleoporins and decreased nuclear pore complex density (Walther *et al.* 2003). Interestingly, we found that RNAi depletion of *npp-4* had negligible effects on lifespan. This result is confounding as NPP-4 functions in the central channel alongside NPP-1 and that it has been suggested that it functions with NPP-3 and NPP-4 in spindle orientation during cell division (Schetter *et al.* 2006; Cohen-Fix and Askjaer 2017).

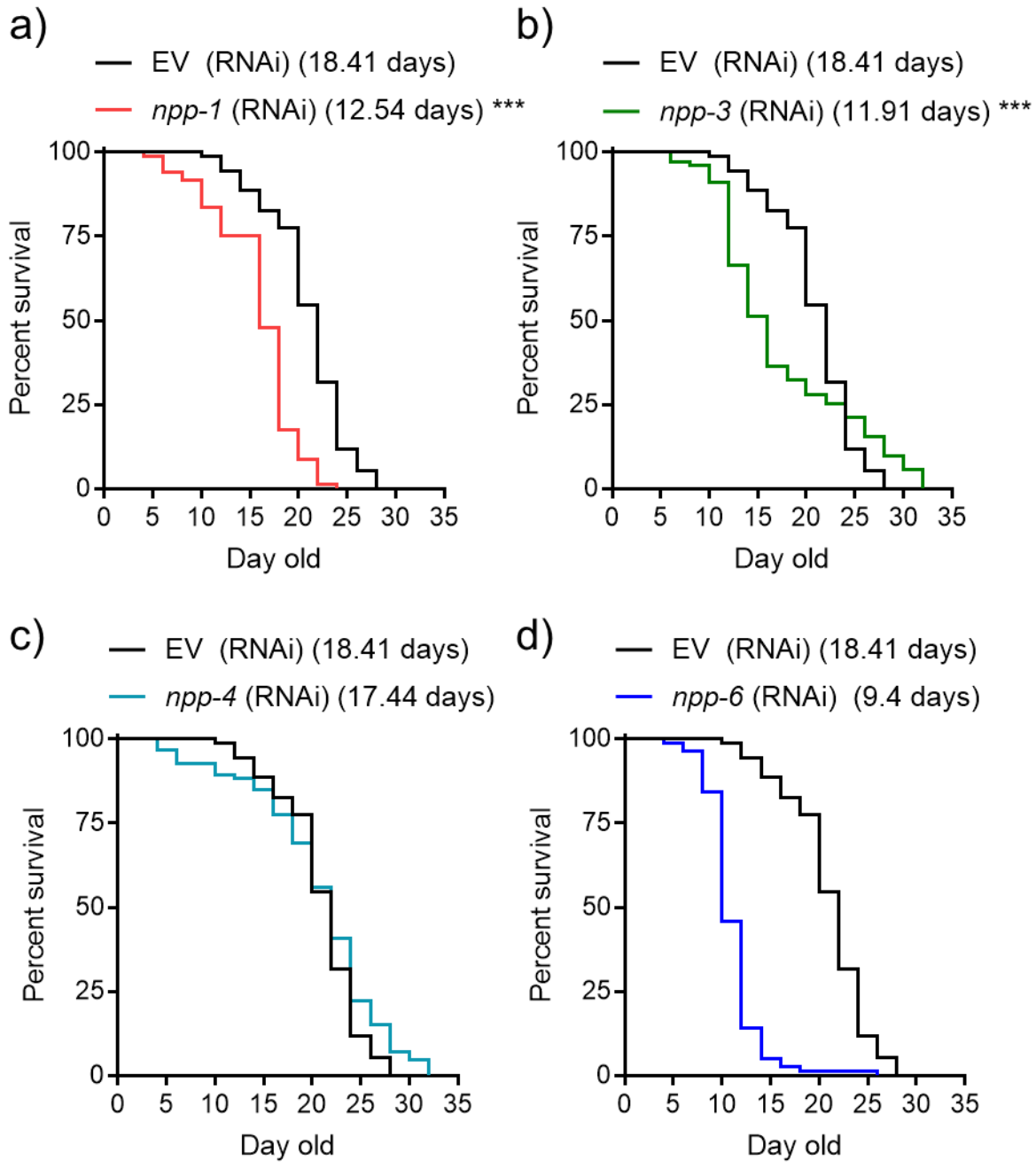


Figure 4.4.4: Nucleoporin proteins' effects on lifespan. All worms were age-matched via hypochlorite treatment and RNAi exposure beginning during L1 larval stage. N2 worms were fed either empty vector control or RNAi targeting a specific nucleoporin gene. Average lifespan is indicated in parentheses. ****P<0.0001 as determined by the log-rank test compared to N2.

4.5 Discussion

The results of this study suggest that the spatial and temporal requirements of Integrator function are critical during early development and in somatic cells. Germline depletion of INTS-4 resulted in negligible changes to lifespan regardless of the timing of INTS-4 depletion. Meanwhile, somatic depletion of INTS-4 resulted in lifespan reduction only when the protein is depleted during early larval development. These results are interesting as the germline undergoes extensive changes during development and would presumably require the function of the Integrator in RNA splicing and transcriptome control to assert these changes. Recent evidence has indicated at least one protein in the *C. elegans* germline is capable of functioning as an independent pre-mRNA splicing factor. The newly characterized *mog-7* gene contains NineTeen complex Related protein 2 (Ntr2) and GC-rich sequence DNA-binding FaCtor homolog (GCFC) domains, which have been shown to control splicing in *S. cerevisiae* (Cao *et al.* 2022). This could serve as an alternative mechanism to protect germline transcripts in the absence of canonical splicing pathways. Though it is possible that depletion of Integrator in the germline may affect other aspects of physiology such as reproduction and embryogenesis instead of aging.

In somatic cells, we found that depletion of INTS-4 only shortened lifespan if it occurred during larval development. This may be attributed to the fact that miotic division of somatic cells occurs during early embryonic and larval stages of development and cells become irreversibly arrested in a postmitotic state once they reach adulthood (Kipreos and van den Heuvel 2019). In this context, Integrator function may only be critical in actively dividing cells and dispensable once cells have fully differentiated. This may reflect why loss of Integrator during the larval stage led to shortened lifespan while having minimal effects when the depletion occurs during adulthood. While our results show a requirement for Integrator activity in the soma but not the

germline, it is possible the somatic defects we observe result from a culmination of Integrator failure in multiple tissues. As the *eft-3* promoter used to drive somatic TIR1 expression to promote INTS-4 degradation is constitutively active in all somatic tissues, small defects from individual tissues may compound into larger whole-organism effects. Future studies of using tissue-specific promoters for other individual tissue types could help fill these gaps in knowledge.

A key regulator of the Integrator complex we recently discovered was the Argonaute protein CSR-1 which functions in the RNAi pathway (Waddell and Wu 2024). The endogenous RNAi pathways exert control over many cellular processes and stabilize the genome in the germline (Grishok 2005). The CSR-1 dependent RNAi pathway promotes transcriptional reprogramming during the oocyte to embryo transition (Fassnacht *et al.* 2018). Removal of individual components of these pathways can have deleterious outcomes in terms of accurate and timely expression of the transcriptome. Loss of DCR-1, the sole Dicer homolog of *C. elegans*, leads to embryonic lethality and sterility (Grishok *et al.* 2001). Like-wise, loss of the catalytic Argonaute CSR-1 is also attributed to embryonic lethality and sterility (Claycomb *et al.* 2009). Here, we demonstrate that loss of function mutation to *csr-1* decreases in lifespan that requires its catalytic Piwi domain. Interestingly, a mutant that specifically deletes the *csr-1a* isoform did not influence lifespan, with this isoform recently found to be predominantly involved in spermatogenesis (Nguyen and Phillips 2021; Charlesworth *et al.* 2021). The general function of CSR-1 has been suggested to play a protective rather than silencing role in the licensing and expression of germline-regulated genes (Wedeles *et al.* 2013). As such, loss of *csr-1* function could affect lifespan through dysregulation of germline gene expression that negatively affects germ cell development. As evidence mounts for germline to soma communication in the

progression of aging, maintenance of *C. elegans* germline cells by CSR-1 could preserve the signal transduction between these tissues.

Our study also identified several members of the nuclear pore complex proteins as regulators of lifespan. These are large macromolecular complexes that associate with and span the nuclear membrane, where they control the cellular material that enters and exits the nucleus (Pemberton and Paschal 2005). Few disease-causing mutations have been observed to affect the nuclear pore complex, underscoring its critical roles in biological processes (Pemberton and Paschal 2005). A paramount function of the NPC is the shuttling of RNA pol II from the cytoplasm into the nucleus (Forget *et al.* 2013). The inability of pol to enter the nucleus would be disastrous to the cell, as transcription would come to a standstill. Of equal importance is the import of the spliceosome into the nucleus and the absence of intron removal from pre-mRNA would be equally detrimental as translated products would have severely altered functions (Wilkinson *et al.* 2020). With respect to snRNA processing, snRNA transcripts destined for spliceosome incorporation are first exported from the nucleus to interact with small nuclear ribonucleoproteins before re-entry in the nucleus to assemble into a functional spliceosome (Takata *et al.* 2012; Becker *et al.* 2019).

We observed that loss of *npp-6* resulted in the largest change in average lifespan and this protein forms part of the ring structure that is exposed to both the cytoplasm and nucleoplasm. The absence of NPP-6 could impact the integrity of the nuclear pore complex as a whole and its association with the nuclear membrane. It could also impact the binding of additional nuclear pore complex cofactors involved in import/export signal recognition as this ring contacts the nuclear basket and cytoplasmic recognition domains (Cohen-Fix and Askjaer 2017). This could lead to widespread disruption of NPC localization and function which would hinder many

cellular processes. Depletion of *npp-1* and *npp-3* reduced overall lifespan, though the effects were not as pronounced as *npp-6.* The NPP-1 protein functions in the central channel through which cargo passes through, and loss of this protein could alter the kinetics of cargo movement through the NPC as well as the selectivity of the nuclear pore complex. The NPP-3 protein is found in the inner ring that makes contacts with virtually all the subdomains of the NPC, and loss of this protein could alter the size of the nuclear pore allowing for unregulated transport across the nuclear membrane due to increased pore size or reduced transport if the pore size is diminished. Interestingly, it has been observed that nuclei purified from old nematodes show an increase in permeability to 70 kDa dextrans, suggesting that the nuclear pore complex function may degrade as a function of age (D'Angelo *et al.* 2009). As such, it appears that gene knockdown of *npp* genes can shorten lifespan, and that the normal function of nuclear pore complex in controlling nuclear-cytoplasmic transport also degrades as the organism ages.

4.6 Conclusion

In this study, we examined how depletion of genes required for snRNA processing in *C. elegans* affects lifespan. The majority of genes assessed in this study cause a decrease in average lifespan, suggesting snRNA processing is important in maintaining a healthy lifespan. However in all of the genes we have identified, including the Integrator, they function in various critical biological processes in addition to snRNA processing. As such, our study does not exclude the overall effects of impaired nuclear transport or altered functions of endogenous RNAi regulation on aging, nor does it rule out the effects of changes to Integrator's pause/release function may have aging independent of its role in snRNA processing. Here, we have initiated a study examining how snRNA regulators can contribute to normal aging, and future studies addressing

these pathways in conjunction with impaired snRNA processing will be essential in determining their mechanism of action.

Chapter 5

General Discussion

The overarching goal of my thesis was to identify novel regulators of snRNA processing in *C. elegans*. Prior to the beginning of my study the Integrator complex was identified as the sole cleavage factor to terminate snRNA transcription and produce mature snRNA transcripts (Baillat *et al.* 2005). While much progress has been made to deduce the various functionalities of this complex, less work has been done examining how this complex is regulated or its impacts on development and aging in a whole-organism context. This chapter will summarize and provide a general discussion of the overall outcomes of my research and its contributions to the field of transcriptional regulation in *C. elegans*. This chapter will also discuss the major limitations of my research as well as the future directions of my study.

5.1 CSR-1 as a novel regulator of snRNA processing in *C. elegans*

Given the growing body of evidence describing the functions of the Integrator complex and a lack of knowledge governing the regulation of these processes, my study focused on identifying novel components that can regulate specific functions of the Integrator, namely its 3` snRNA processing. In Chapter 2 I focused on identifying novel regulators of snRNA processing through the use of a genome-wide RNAi screen. I found that the *csr-1* gene, which encodes an essential AGO protein in *C. elegans* that participates in the RNAi response and germline maintenance, to be critical for proper snRNA processing. Using qPCR analysis of a *csr-1* mutant as well as RNAi knockdown of *csr-1*, I found that levels of misprocessed U2 and U4 snRNA was significantly increase (**Figure 2.4.3 a and b**). Through isoform analysis, I further determined that CSR-1 mediated snRNA regulation is likely to involve the B isoform and requires its catalytic

PIWI activity domain (**Figure 2.4.3 b**). To better contextualize the importance of CSR-1 in whole animal physiology, I performed lifespan analysis and found that loss of *csr-1* results in decreases in average lifespan and this decrease is also partially dependent on the catalytic activity of *csr-1b* isoform (**Figure 4.4.1 a and b, Figure 4.4.3 c**). To obtain a broader view of the effect of *csr-1* on the transcriptome, I used RNA sequencing and found that the transcriptomic changes in response to *csr-1* knockdown was highly similar to those observed when subunits of the Integrator were knocked down (**Figure 2.4.5 b**). This suggests the same biological processes are perturbed in each condition. Mechanistically, I used Ribo-seq data published in a previous study to uncover that loss of *csr-1* decreased the translation efficiency of multiple Integrator subunits (**Figure 2.4.9 d**) (Singh *et al.* 2021). I was able to verify this effect by using CRISPR/Cas9 to add the mKate2 fluorophore to the *ints-4* gene to study its endogenous localization patterns. This fluorescent analysis of INTS-4::mKate2 in a *csr-1* mutant revealed that loss of *csr-1* decreases INTS-4::mKate2 expression in the germline (**Figure 2.4.12 a**). Taken together, my results in Chapter 2 have identified CSR-1b as a novel regulator Integrator expression in the germline to promote snRNA processing.

CSR-1 in *C. elegans* is unique in several aspects. Firstly, *csr-1* is the only AGO in *C. elegans* that is required for embryonic viability and fertility (Claycomb *et al.* 2009). Second, the majority of CSR-1 targets in the germline are not silenced but instead protected from silencing to license their expression (Wedeles *et al.* 2013). The second point is interesting as CSR-1 is one of the few *C. elegans* AGOs that exhibit a catalytically active Piwi domain (Youngman and Claycomb 2014). A conserved function of AGOs in many systems is to down-regulate target mRNA via the endonuclease activity of their Piwi domain, which acts to ‘slice’ target sequences to undergo endolytic cleavage (Hutvagner and Simard 2008). The implications of CSR-1’s

slicing function are slowly being elucidated. Studies have shown that the slicing activity of CSR-1 is required to accurately load the *C. elegans* CENP-A homolog HCP-3 onto worm chromosomes to direct kinetochore formation and promote proper chromosome segregation (Wong *et al.* 2024). In addition, CSR-1's slicing activity has been found to promote microtubule dynamics during embryonic divisions and is required for the 3` processing of histone mRNA (Avgousti *et al.* 2012; Gerson-Gurwitz *et al.* 2016). The slicing activity of CSR-1 has also been implicated in balancing the load of maternal mRNAs incorporated into developing oocytes to ensure accurate embryonic divisions based on the density of 22G RNA bound to CSR-1 (Gerson-Gurwitz *et al.* 2016).

As my results have shown that the loss of *csr-1* results in decreased translational efficiency of several Integrator subunits and decreased protein expression of INTS-4 in the germline, the simplest explanation for this interaction is that *csr-1* controls the expression of the Integrator complex. Indeed, analysis of small RNA sequencing data from immunoprecipitated CSR-1 has found that CSR-1 is bound to many 22G-RNA that are antisense to genes encoding subunits of the Integrator complex (Charlesworth *et al.* 2021; Waddell and Wu 2024). However, if *csr-1* is a master regulator of Integrator expression, its depletion should cause the most robust increase in snRNA misprocessing. My results show that the depletion of *csr-1* alone results in a lesser degree of snRNA misprocessing than the depletion of *ints-4* directly (**Figure 2.4.7 e**). This may be due to the fact that we found that loss of *csr-1* primarily causes INTS-4 depletion in only the germline and mot in other somatic tissues. I also found that 22G RNAs antisense to snRNA transcripts are bound to CSR-1 targeting the complex to the 5` ends of snRNA transcripts (Waddell and Wu 2024). While this does not directly support a role for *csr-1* in the direct cleavage of snRNA transcripts, as the cleavage motif occurs at the 3` box sequence, it does open

up the possibility of their regulation by CSR-1 by alternative mechanisms in addition to influencing the expression of Integrator subunits. Perhaps binding of the CSR-1 complex during snRNA transcription promotes the cleavage function of Integrator. In this way depletion of CSR-1 alone would impair snRNA processing independently of Integrator expression, and the loss of both CSR-1 and INTS-4 would cooperate to produce a more robust misprocessing defect. A more complicated mechanism would be the global dysregulation of multiple proteins that converge to produce snRNA processing defects, although this would require extensive characterization to determine exactly which factors are involved.

My study includes several limitations. Firstly, my study identified a host of factors that can influence snRNA processing yet only *csr-1* was examined in detail. These additional factors could share regulatory mechanisms with *csr-1* that could help confirm the action of *csr-1* on snRNA processing. On the other hand, they could also represent many unexplored avenues of potential snRNA regulation that could function independently of *csr-1*. Transcript expression is subject to many forms of regulation, and this study only scratches the surface of a single possibility. Another caveat of my study is the lack of Western Blot analysis. While the fluorescent analysis used here serves as a starting point to determine the effects on Integrator expression, the exact degree of these changes will require protein immunoblotting for a more accurate determination. A possible future direction would be to examine potential sex-specific effects of this regulation as this study only examined snRNA processing in hermaphrodite worms. As males do exist in the general population, albeit only accounting for 0.1%, this study did not examine any sex-based effects of snRNA processing. In the past 6 years, the Integrator's newly described function in the pause/release of stalled RNA pol II has added a degree of complexity when studying individual Integrator functions that we did not take into account in

this work. As this pause/release affects a large number of transcripts across all developmental stages, the overall transcriptome could be perturbed independently of altered snRNA processing. This could affect the Integrator itself, decreasing the availability of functional Integrator for snRNA processing, the expression of snRNA transcripts themselves, or the expression of yet to be identified factors/regulators of snRNA processing. This global alteration of the transcriptome could also be the driving force behind the lifespan effects seen in this study. Elucidation of the subunits responsible for this pause/release function will be beneficial in examining the independent roles of snRNA processing on development and aging.

5.2 Identification of additional snRNA processing regulators

Building upon my results in Chapter 2, the overall objective in Chapter 3 was to identify additional regulators of snRNA processing through using a complementary genetic screen. Here I employed a forward EMS mutagenesis screen to isolate viable mutants with snRNA processing defects. An additional six regulators of snRNA processing were identified as components of the RNAi response in *C. elegans*. Interestingly the mutations I isolated with the most robust misprocessing defect affected components involved in secondary siRNA amplification, *mut-16* and *rde-11* (**Figure 3.4.4**). This is striking as mutation to either component results in a reduced RNAi response and not complete refraction to RNAi (Phillips *et al.* 2012; Yang *et al.* 2012). This could be due to the loss of interactions with other components of the RNAi machinery when these proteins are absent/nonfunctional. MUT-16 is a required scaffolding component of Mutator foci, and its loss prevents their aggregation. This in turn could alter the localization of RRF-1, the RdRP responsible for the amplification of siRNA targeting somatic transcripts, decreasing the availability of siRNA for AGO (Phillips *et al.* 2012). This could also affect the loading of mutator-associated AGOs themselves, although they have yet to be observed in these foci. In this

way mutation to *mut-16* could alter the location and/or function of multiple RNAi components by destabilizing the Mutator foci whose effects descend on perturbations to snRNA processing.

RDE-11 functions similarly to MUT-16, but it is not associated with a defined subcellular compartment (Yang *et al.* 2012). Aside from siRNA amplification, the RDE-10/11 complex also participates in the degradation of mRNA targets and the loss of this complex would impact transcript silencing by decreasing RDE-10/11 activity and decreasing the production of amplified siRNAs. RSD-2, a known interacting component of RDE-10/11 identified as an snRNA processing regulator in this study, also participates in siRNA amplification and systemic RNAi spreading (Tijsterman *et al.* 2004; Yang *et al.* 2012). In this way loss of RDE-10/11 could also perturb RSD-2 function and exacerbate the effects seen in the *rsd-2* mutants. Much like MUT-16, the RDE-10/11 is poised to affect other RNAi components which could result in global transcript dysregulation that ultimately impedes snRNA processing.

I also isolated 3 independent mutations in *rde-1* that encodes a primary Argonaute protein involved in the maturation of primary siRNA triggers (Steiner *et al.* 2009). Loss of *rde-1* results in a severely reduced response to RNAi in all tissues, indicating a reduced function of the RNAi pathway (Tabara *et al.* 1999). Interestingly, RDE-1 and CSR-1 are the only two Argonaute proteins that have a catalytically active RNAse H domain exhibiting slicer activity. Though unlike CSR-1, the slicer activity of RDE-1 is involved in passenger-strand RNAi turnover prior to the incorporation of the guide strand into the RISC and is not involved in target mRNA cleavage (Steiner *et al.* 2009). As such, it is likely that loss of *rde-1* likely contributes to snRNA misprocessing via disruption of the RNAi pathway, rather than via potential loss of direct cleavage activity of the snRNA. Finally, I also identified the secondary Argonaute *sago-2* as a novel regulator of snRNA processing. Along with *rsd-2*, the mutation to *sago-2* produced the

least pronounced misprocessing defects. SAGO-2 is a worm-specific AGO that functions in the intestine of *C. elegans*, where it binds to secondarily amplified siRNA to promote target silencing (Yigit *et al.* 2006; Seroussi *et al.* 2023). As *C. elegans* contains at least 27 distinct AGO members, it is not unfeasible that redundant secondary AGOS exist that could partially rescue the loss of SAGO-2 function. A separate scenario could be that the number of transcripts regulated by SAGO-2 is relatively small or not as critical for snRNA processing, which could explain the low levels of misprocessing I observed in this mutated strain.

There are several limitations that I have identified for this study. Firstly, there has been evidence of RNAi components playing roles in transgene silencing which include the RDE-10/11 complex that participates in somatic transgene silencing (Grishok *et al.* 2005; Yang *et al.* 2012). As my study employed a fluorescent *in vivo* transgene as an initial indicator of snRNA misprocessing, this could lead to a high false positive rate where the transgene is incorrectly activated without defects in snRNA processing. In this scenario, loss of RNAi components leads to de-silencing that can lead to artificially high levels of reporter activation that does not match changes to endogenous gene expression. Varying degrees of transgene desilencing have been observed in mutants of *mut-16, rde-1, and rrf-1,* which were amongst genes I identified in this mutagenic screen (Grishok *et al.* 2005). However, through the use of qPCR I measured the levels of snRNA misprocessing to confirm changes occurred to the endogenous transcript. However, this may explain why the intensity of the GFP reporter activation appears to be disproportionally strong compared to the levels of endogenous snRNA processing in mutants such as *rde-1*, which may be a result of enhanced transgene de-silencing that increases GFP fluorescence.

Another limitation of my study was the complementation strategy used to confirm the identities of the isolated mutations. This method was chosen over classical wildtype DNA rescue

as many of the amplicons required for this procedure were large and proved difficult to accurately amplify in my experience. The complementation strategy was not applicable to all mutations identified in my study, due to issues in mating or transgene expression, and as a result not all alleles were subject to confirmation. An example of this was an inability to successfully induce the mating of *rde-11* males for the complementary assay. My study only went as far as to characterize the levels of snRNA misprocessing experienced by the novel mutants. Additional experiments examining Integrator subunit expression could provide insight into the mechanism of this regulation. As for future directions, given that all of the mutated genes identified are members of the RNAi response, epistatic experiments in creating double or triple mutants of these genes will provide critical information to understand how the RNAi pathway as a whole influences snRNA processing.

5.3 snRNA regulators are essential genes for proper lifespan maintenance

Expanding upon my work in Chapters 2 and 3, I characterized the effects of snRNA processing defects on organismal lifespan. The most informative observation I made was the requirements of Integrator function on aging. Using the AID degron system, I found that Integrator depletion in somatic tissues during early development resulted in severe developmental arrest and a drastically decreased lifespan. Most interestingly, I found that this effect is largely dispensable in adulthood as INTS-4 depletion in adult worms had no effect on lifespan. As far as I am aware, this is the first description for differential requirements of the Integrator at different developmental time points. As there are large transcriptional changes that accompany tissue differentiation during development, tight control over splicing or transcriptional control likely requires a functional Integrator. However as animals develop into adults, many of the somatic tissues have reached their potential and become post mitotic and the

need for such tight splicing/transcriptome regulation may not be as critical in an aging context (Korta and Hubbard 2010).

The next focus of my lifespan analysis centered around the Argonaute encoding *csr-1* gene. Fluorescent analysis of INTS-4::mKate2 in a *csr-1* mutant indicated that the loss of *csr-1* resulted in reduced expression of INTS-4::mKate2 in the germline. This *csr-1* mutant also experienced snRNA processing defects and a reduction in average lifespan, which were dependent on the catalytic activity of the b isoform. Interestingly, loss of *csr-1* causes embryonic lethality but does not affect larval growth and development. These results indicate that while *csr-1* promotes snRNA processing, it may influence development and aging separately from the Integrator.

I also characterized the effects on lifespan of some of the additional snRNA processing regulators identified in this study. I found that several members of the nuclear pore complex to be involved in snRNA processing regulation. Depletion of NPP-1 or NPP-6 causes significant snRNA misprocessing defects as well as decreases to average lifespan. Additionally, depletion of NPP-3 and NPP-4 were found to cause decreases in average lifespan but their effects on misprocessing were only evaluated based on the snRNA misprocessing reporter rather than via qPCR analysis. Nucleoporins manage the transport of cellular components across the nuclear membrane, so it is not surprising that disrupting the nuclear pore has deleterious consequences. By altering the size of the pore, the interactions inside of the transport channel, or disturbing receptors on either the nucleoplasmic or cytoplasmic faces would disrupt virtually all cellular processes. Molecules such as snRNA require transport from the nucleus to the cytoplasm for incorporation into snRNPs, and then back into the nucleus to form the spliceosome (Will and Lührmann 2011). If either step of their transport is disrupted, RNA splicing would be negatively

affected due to the inappropriate mis-location of the snRNA transcript. Additionally, if any component of the Integrator is unable to translocate into the nucleus, snRNA processing itself would also be negatively affected.

While the results of my study provide new information on the effects of snRNA processing on aging, it has various areas for improvement. My study only examined two time points of Integrator depletion, beginning during early development at the L1 stage and during the transition into adulthood at the L4 stage. With the auxin degron system, it is possible to include additional time points that encompass the L2 and L3 larval stages which could pinpoint the exact timing of Integrator requirement during development. Additionally, given that a collection of *C. elegans* strain that drives TIR1 expression in a tissue-specific manner was recently developed, there exists the possibility to test Integrator depletion in individual somatic tissues (*e.g.* muscle, neurons, hypodermis, etc.) and the germline. Lastly, given that the auxin effect is reversible, it will also be of interest to examine whether recovery from L1 auxin exposure could reinitiate the developmental cycle to reverse its effect on reduced body growth and shortened lifespan.

Lastly, my study only examined snRNA processing effects in the hermaphrodite, which has a distinct germline with distinct signaling pathways compared to the male gonad. Therefore, any sex-specific effects have gone unnoticed. It would be of interest to see if the spatial and temporal effects of Integrator degradation have the same effect on both sexes of *C. elegans.*

5.4 Overall significance and contributions

The Integrator is quickly gaining appreciation as a regulator of the transcriptome. From its first description related to snRNA processing in 2005, focus on Integrator function has led to the attribution of several modes of transcriptional control, in addition to module-specific sub-

functions and the identification of novel subunits (Pfleiderer and Galej 2021; Welsh and Gardini 2023). A major function of Integrator that has been recently identified is the attenuation of paused RNA pol II at protein-coding genes (Elrod *et al.* 2019). Integrator is recruited to stalled RNA pol II through interactions with the negative elongation factor (Stadelmayer *et al.* 2014). This interaction allows INTS-8 of the phosphatase module to recruit PP2A to RNA pol II and dephosphorylate its C terminal domain, preventing the association of positive elongation factors (Fianu *et al.* 2024). Then INTS-11 of the cleavage module can cleave the nascent transcript and release it from the pol (Elrod *et al.* 2019). A newly described crystal structure indicates a novel module composed of INTS-10, 13, 14, and 15 can then disrupt the DNA clamp binding domain of RNA pol II, dissociating the pol from the DNA strand (Fianu *et al.* 2024).

This mechanism proposes a model where RNA pol II is actively engaged with a large number of transcripts, but the actions of Integrator prevent transcript elongation and promote the "recycling" of the pol onto productive genes. In this way, Integrator can control the transcriptome independently of its contributions to splicing. As such, it is possible that many of the phenotypes we observed with Integrator depletion including developmental defect and shortened lifespan could be attributed to Integrator's control over pol II pause-release, rather than RNA splicing control vis snRNA processing. The endonuclease function of INTS-11 is required for efficient RNA pol II mRNA processing, and as such might share regulatory mechanisms with snRNA processing (Stadelmayer *et al.* 2014). However, studies have indicated that only INTS-1, 4, 9, and 11 are the only subunits that are directly required for snRNA processing (Ezzeddine *et al.* 2010). The requirements of the phosphatase module along with the newly characterized "scorpion" module in RNA pol II pause/release opens up the possibility of differential regulation of the individual sub-modules to affect separate aspects of Integrator functions.

In humans, Integrator mutations are often associated with neurodevelopmental deficiencies (Oegema *et al.* 2017; Tepe *et al.* 2023). Mutations to either *INTS-1* or *INTS-8* have been found to significantly affect RNA splicing patterns in patient-derived fibroblast and are thought to lead to brain abnormalities and neurodevelopmental syndromes (Oegema *et al.* 2017). Similarly, mutations to *INTS-11* have also been reported in separate cases of neurodevelopmental syndromes (Tepe *et al.* 2023). While the exact mechanism of INTS-11 disruption has yet to be identified, studies indicate that there is a general transcript dysregulation in INTS-11 deficient cells (Elrod *et al.* 2019). Mutations to *INTS-11* that impair its function have recently been reported to cause decreases in lifespan in *Drosophila* (Tepe *et al.* 2023), which is in agreement with the results of my study. While it has yet to be identified as a driver of cancer, INTS-7 and 8 have been observed to be highly mutated in several different types of malignancies and the transcription of *INTS-7, 8*, and *13* have also been found to be altered in many cancer types (Federico *et al.* 2017).

The implications of heterozygous Integrator mutations in cancer and neurodegenerative diseases highlight the importance of this complex in promoting healthy aging and development, two issues components of organismal fitness I address in my study. My data shows that loss of the Integrator causes developmental defects and shortened lifespan, highlighting Integrator's importance in regulating biological processes. The simple nature of model organisms like *C. elegans* has allowed researchers to characterize these conserved biological processes, and as such provide researchers an excellent opportunity to study the complex functions required for life. The similarities between these model systems and humans allows us to study disease states and extrapolate their evolutionarily conserved mechanism of action into high-ordered organisms. This is also one of the first studies to place a tissue-specific requirement on Integrator function,

opening the possibility of additional uncharacterized tissue-specific functions which may provide insights toward different pathogenic states associated with Integrator human mutations. This study also serves as an initial insight into the regulations of Integrator function, which currently have not been well studied. In my work here, we have identified several potential regulators of Integrator, including the Argonaute protein CSR-1, along with other components of the RNAi pathway and the nuclear transport system. Further work towards understanding mechanisms of Integrator regulation would prove beneficial toward uncovering Integrator related disease mechanisms and potentially provide sources of therapeutic targets to treat these diseases.

5.5 Conclusion

The main goal of my Ph.D. thesis was to identify novel regulators of integrator-mediated snRNA processing in *C. elegans*. My hypothesis was these novel regulators would prove to be essential in maintaining normal organismal function. My studies were equally focused on genetic screens to identify these novel regulators, characterization of their mechanisms of action, and the degree to which they impact aging. Overall my results are supportive of my hypothesis; the genes I identified all impacted snRNA processing and their loss resulted in decreased lifespan for the most part. Utilizing the genetic amenability of *C. elegans* I was able to perform an in-depth characterization of the requirements of *csr-1* on snRNA processing, providing for the first time evidence of snRNA processing regulation through the control Integrator expression. Expanding upon our knowledge of the Integrator, I was able to narrow down the timing and placement of Integrator function in the context of aging and development. This spatial-temporal specificity of Integrator subunit-4 in development and aging prompts the rationale for future studies to examine the specificity of other Integrator subunits in contributing to organismal physiology. Of particular interest will be the role Integrator plays in neurodevelopment, as it has been implicated

in an array of neurodevelopment diseases which are becoming more prevalent in our aging population.

Appendix A. Lifespan data statistics

Figure 4.4.1 a	Mean Survival (Days)	P value vs EV (RNAi)	95% confidence interval	# of subjects scored	# of censors
Lifespan (Trial 1)					
N2 EV (RNAi)	18.41		17.33 ~ 19.48	53	6
N2 *ints-4* (RNAi)	6.91	0	6.52 ~ 7.31	69	10
Lifespan (Trial 2)					
N2 EV (RNAi)	20.03		19.03 ~ 21.03	100	28
N2 *ints-4* (RNAi)	7.45	0	7.18 ~ 7.73	121	0
Lifespan (Trial 3)					
N2 EV (RNAi)	20.92		20.06 ~ 21.78	91	13
N2 *ints-4* (RNAi)	5.77	0	5.44 ~ 6.10	103	0

Figure 4.4.1 b	Mean Survival (Days)	P value vs N2 ETOH	95% confidence interval	# of subjects scored	# of censors
Lifespan (Trial 1)					
N2 ETOH	12.49		11.75 ~ 13.24	111	20
N2 L1 auxin	12.41	0.2495	11.89 ~ 12.93	85	18
INTS-4::degron ETOH	12.52	1	11.76 ~ 13.28	85	10
INTS-4::degron L1 auxin	7.57	0	7.28 ~ 7.86	223	0
Lifespan (Trial 2)					
N2 ETOH	13.15		12.57 ~ 13.73	157	18
N2 L1 auxin	13.4	1	12.57 ~ 14.23	79	11
INTS-4::degron ETOH	13.43	1	12.65 ~ 14.20	98	9
INTS-4::degron L1 auxin	7.57	0	7.31 ~ 7.83	138	0
Lifespan (Trial 3)					
N2 ETOH	13.65		13.04 ~ 14.26	150	18
N2 L1 auxin	11.85	0.0003	11.30 ~ 12.40	135	32
INTS-4::degron ETOH	14.59	0.1155	13.93 ~ 15.25	190	22
INTS-4::degron L1 auxin	10.01	0	9.77 ~ 10.25	261	58

Figure 4.4.2 a	Mean Survival (Days)	P value vs N2 ETOH	95% confidence interval	# of subjects scored	# of censors
Lifespan (Trial 1)					
N2 ETOH	12.49		11.75 ~ 13.24	111	20
N2 L1 auxin	12.06	0.0998	11.89 ~ 12.93	85	18
N2 L4 auxin	12.41	0.0169	11.62 ~ 12.50	164	27
Lifespan (Trial 2)					
N2 ETOH	13.14		12.56 ~ 13.71	156	17
N2 L1 auxin	11.24	0.0001	11.84 ~ 12.69	56	11
N2 L4 auxin	12.27	8.60E-03	10.79 ~ 11.69	162	19
Lifespan (Trial 3)					
N2 ETOH	13.65		13.04 ~ 14.26	150	18
N2 L1 auxin	11.85	0.0001	11.30 ~ 12.40	135	32
N2 L4 auxin	13.13	0.4034	12.57 ~ 13.70	149	43

Figure 4.4.2 b	Mean Survival (Days)	P value vs sun1p::TIR1 ETOH	95% confidence interval	# of subjects scored	# of censors
Lifespan (Trial 1)					
sun1p::TIR1 ETOH	12.86		12.30 ~ 13.41	133	23
sun1p::TIR1 L1 auxin	13.44	0.5564	12.89 ~ 13.99	111	13
sun1p::TIR1 L4 auxin	12.68	0.5783	12.26 ~ 13.10	150	21
Lifespan (Trial 2)					
sun1p::TIR1 ETOH	13.58		13.00 ~ 14.16	144	18
sun1p::TIR1 L1 auxin	14.33	0.2687	13.82 ~ 14.85	156	30
sun1p::TIR1 L4 auxin	14.88	0.0062	14.28 ~ 15.47	184	35
Lifespan (Trial 3)					
sun1p::TIR1 ETOH	13.62		13.20 ~ 14.04	134	14
sun1p::TIR1 auxin	13.91	0.7716	13.54 ~ 14.29	160	16
sun1p::TIR1 auxin	14.17	0.1204	13.74 ~ 14.60	203	36

Figure 4.4.2 c	Mean Survival (Days)	P value vs eft-3p::TIR1 ETOH	95% confidence interval	# of subjects scored	# of censors
Lifespan (Trial 1)					
eft-3p::TIR1 ETOH	12.52		11.76 ~ 13.28	85	10
eft-3p::TIR1 L4 auxin	13.12	1	12.77 ~ 13.46	144	8
eft-3p::TIR1 L1 auxin	7.57	0	7.28 ~ 7.86	223	0
Lifespan (Trial 2)					
eft-3p::TIR1ETOH	13.43		12.65 ~ 14.20	98	9
eft-3p::TIR1 L4 auxin	11.96	0.0016	11.46 ~ 12.47	97	8
eft-3p::TIR1 L1 auxin	7.57	0	7.31 ~ 7.83	138	0
Lifespan (Trial 3)					
eft-3p::TIR1 ETOH	14.59		13.93 ~ 15.25	190	22
eft-3p::TIR1 L4 auxin	14.15	0.3772	13.69 ~ 14.60	225	48
eft-3p::TIR1 L1 auxin	10.01	0	9.77 ~ 10.25	261	58

Figure 4.4.3 a	Mean Survival (Days)	P value vs N2	95% confidence interval	# of subjects scored	# of censors
Lifespan (Trial 1)					
N2	18.41		17.33 ~ 19.48	53	6
tm892	14.44	6.30E-07	13.44 ~ 15.43	52	6
Lifespan (Trial 2)					
N2	20.03		19.03 ~ 21.03	100	28
tm892	14.16	0	13.45 ~ 14.87	78	8
Lifespan (Trial 3)					
N2	20.92		20.06 ~ 21.78	91	13
tm892	13.4	0	12.68 ~ 14.11	83	12

Figure 4.4.3 b	Mean Survival (Days)	P value vs N2	95% confidence interval	# of subjects scored	# of censors
Lifespan (Trial 1)					
N2	17.89		17.12 ~ 18.67	125	29
cmp135	16.95	0.1823	15.76 ~ 18.15	87	52
Lifespan (Trial 2)					
N2	18.43		17.89 ~ 18.98	177	23
cmp135	16.58	0.0000015	16.05 ~ 17.10	186	43
Lifespan (Trial 3)					
N2	17.06		16.36 ~ 17.75	157	7
cmp135	15.46	0.000024	14.90 ~ 16.01	152	15

Figure 4.4.3 c	Mean Survival (Days)	P value vs *csr-1*^{WT} + EV (RNAi)	95% confidence interval	# of subjects scored	# of censors
Lifespan (Trial 1)					
csr-1^{WT} + EV (RNAi)	16.3		14.42 ~ 18.18	37	16
csr-1^{WT} + *csr-1* (RNAi)	17.44	1	15.98 ~ 18.90	98	79
csr-1^{SIN} + EV (RNAi)	17.96	0.75	16.27 ~ 19.65	66	34
csr-1^{SIN} + *csr-1* (RNAi)	14.72	0.68	13.43 ~ 16.01	105	75
Lifespan (Trial 2)					
csr-1^{WT} + EV (RNAi)	19.24		18.37 ~ 20.11	174	36
csr-1^{WT} + *csr-1* (RNAi)	17.93	0.151	17.01 ~ 18.85	134	51
csr-1^{SIN} + EV (RNAi)	17.92	0.0451	17.18 ~ 18.67	139	30
csr-1^{SIN} + *csr-1* (RNAi)	16.19	0.0001	15.16 ~ 17.22	126	61
Lifespan (Trial 3)					
csr-1^{WT} + EV (RNAi)	17.59		16.69 ~ 18.49	153	16
csr-1^{WT} + *csr-1* (RNAi)	15.94	0.022	15.10 ~ 16.79	124	14
csr-1^{SIN} + EV (RNAi)	18.69	0.389	17.53 ~ 19.84	117	19
csr-1^{SIN} + *csr-1* (RNAi)	13.94	0	13.43 ~ 14.46	167	66

Figure 4.4.4 and 4.4.5	Mean Survival (Days)	P value vs N2 EV (RNAi)	95% confidence interval	# of subjects scored	# of censors
Lifespan (Trial 1)					
N2 EV (RNAi)	18.41		17.33 ~ 19.48	53	6
npp-1 (RNAi)	12.54	0	11.69 ~ 13.39	85	4
npp-3 (RNAi)	12.24	0	11.34 ~ 13.14	95	7
npp-4 (RNAi)	18.77	1	17.53 ~ 20.01	102	11
npp-6 (RNAi)	9.4	0	8.98 ~ 9.81	89	9
dcr-1 (RNAi)	14.31	0	13.52 ~ 15.10	41	1
mut-16 (RNAi)	17.93	1	16.55 ~ 19.30	52	2
Lifespan (Trial 2)					
N2 EV (RNAi)	20.03		19.03 ~ 21.03	100	28
npp-1 (RNAi)	15.22	0	14.17 ~ 16.28	107	18
npp-3 (RNAi)	18.63	1	16.90 ~ 20.35	119	28
npp-4 (RNAi)	20.46	1	19.35 ~ 21.58	118	17
npp-6 (RNAi)	13.32	0	11.69 ~ 14.95	114	33
dcr-1 (RNAi)	15.74	0	15.18 ~ 16.30	87	15
mut-16 (RNAi)	19.35	1	18.34 ~ 20.37	87	17
Lifespan (Trial 3)					
N2 EV (RNAi)	20.92		20.06 ~ 21.78	91	13
npp-1 (RNAi)	15.89	0	14.97 ~ 16.81	86	11
npp-3 (RNAi)	17.97	0.6232	16.42 ~ 19.53	106	20
npp-4 (RNAi)	20.76	1	19.43 ~ 22.09	97	9
npp-6 (RNAi)	11.04	0	10.43 ~ 11.66	87	7
dcr-1 (RNAi)	16.48	0	15.65 ~ 17.30	53	9
mut-16 (RNAi)	18.69	0.0256	17.73 ~ 19.64	86	9

References

1. Albrecht T. R., S. P. Shevtsov, Y. Wu, L. G. Mascibroda, N. J. Peart, *et al.*, 2018 Integrator subunit 4 is a 'Symplekin-like' scaffold that associates with INTS9/11 to form the Integrator cleavage module. Nucleic Acids Res. 46: 4241–4255. https://doi.org/10.1093/nar/gky100

2. Anders S., and W. Huber, 2010 Differential expression analysis for sequence count data. Genome Biol. 11: 1–12. https://doi.org/10.1186/GB-2010-11-10-R106

3. Ankeny R. A., 2001 The natural history of *Caenorhabditis elegans* research. Nat. Rev. Genet. 2: 474–479. https://doi.org/10.1038/35076538

4. Aoki K., H. Moriguchi, T. Yoshioka, K. Okawa, and H. Tabara, 2007 In vitro analyses of the production and activity of secondary small interfering RNAs in *C. elegans*. EMBO J. 26: 5007–5019. https://doi.org/10.1038/SJ.EMBOJ.7601910

5. Apfeld J., and S. Alper, 2018 What Can We Learn About Human Disease from the Nematode C. elegans?, pp. 53–75 in *Methods in molecular biology (Clifton, N.J.)*, NIH Public Access.

6. Ashley G. E., T. Duong, M. T. Levenson, M. A. Q. Q. Martinez, L. C. Johnson, *et al.*, 2021 An expanded auxin-inducible degron toolkit for *Caenorhabditis elegans*. Genetics 217: iyab006. https://doi.org/10.1093/GENETICS/IYAB006

7. Avgousti D. C., S. Palani, Y. Sherman, and A. Grishok, 2012 CSR-1 RNAi pathway positively regulates histone expression in *C. elegans*. EMBO J. 31: 3821–3832. https://doi.org/10.1038/EMBOJ.2012.216

8. Azuma N., T. Yokoi, T. Tanaka, E. Matsuzaka, Y. Saida, *et al.*, 2023 Integrator complex subunit 15 controls mRNA splicing and is critical for eye development. Hum. Mol. Genet. 32: 2032–2045. https://doi.org/10.1093/hmg/ddad034

9. Baillat D., M. A. Hakimi, A. M. Näär, A. Shilatifard, N. Cooch, *et al.*, 2005 Integrator, a

multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. Cell 123: 265–276. https://doi.org/10.1016/j.cell.2005.08.019

10. Baillat D., and E. J. Wagner, 2015 Integrator: Surprisingly diverse functions in gene expression. Trends Biochem. Sci. 40: 257–64. https://doi.org/10.1016/j.tibs.2015.03.005

11. Barrangou R., and J. A. Doudna, 2016 Applications of CRISPR technologies in research and beyond. Nat. Biotechnol. 34: 933–941. https://doi.org/10.1038/nbt.3659

12. Bartschat S., and T. Samuelsson, 2010 U12 type introns were lost at multiple occasions during evolution. BMC Genomics 11: 106. https://doi.org/10.1186/1471-2164-11-106

13. Becker D., A. G. Hirsch, L. Bender, T. Lingner, G. Salinas, *et al.*, 2019 Nuclear Pre-snRNA Export Is an Essential Quality Assurance Mechanism for Functional Spliceosomes. Cell Rep. 27: 3199-3214.e3. https://doi.org/10.1016/J.CELREP.2019.05.031

14. Beltran T., E. Pahita, S. Ghosh, B. Lenhard, and P. Sarkies, 2021 Integrator is recruited to promoter-proximally paused RNA Pol II to generate *Caenorhabditis elegans* piRNA precursors. EMBO J. 40. https://doi.org/10.15252/EMBJ.2020105564

15. Berg D. L. C. van den, R. Azzarelli, K. Oishi, B. Martynoga, N. Urbán, *et al.*, 2017 Nipbl Interacts with Zfp609 and the Integrator Complex to Regulate Cortical Neuron Migration. Neuron 93: 348–61. https://doi.org/10.1016/j.neuron.2016.11.047

16. Brenner S., 1974 The genetics of *Caenorhabditis elegans*. Genetics 77: 71–94.

17. Brow D. A., 2002 Allosteric Cascade of Spliceosome Activation. Annu. Rev. Genet. 36: 333–360. https://doi.org/10.1146/annurev.genet.36.043002.091635

18. Brown T., 2002 *Chapter 1 The Human Genome*. Wile-Liss.

19. Bulteau R., and M. Francesconi, 2022 Real age prediction from the transcriptome with

RAPToR. Nat. Methods 2022 198 19: 969–975. https://doi.org/10.1038/s41592-022-01540-0

20. Burge C., T. Tuschl, and P. A. Sharp, 1998 20 Splicing of Precursors to mRNAs by the Spliceosome. Cold Spring Harb. Monogr. Arch. 37: 525–560.

21. *C. elegans* Sequencing Consortium, 1998 Genome sequence of the nematode *C. elegans*: A platform for investigating biology. Science. 282(5369): 2012–8. https://doi.org/10.1126/science.282.5396.2012

22. Cao W., C. Tran, S. K. Archer, S. Gopal, and R. Pocock, 2022 Functional recovery of the germ line following splicing collapse. Cell Death Differ. 29: 772–787. https://doi.org/10.1038/s41418-021-00891-z

23. Cecere G., S. Hoersch, S. O'keeffe, R. Sachidanandam, and A. Grishok, 2014 Global effects of the CSR-1 RNA interference pathway on the transcriptional landscape. Nat Struc Mol Biol 21: 358–365. https://doi.org/10.1038/nsmb.2801

24. Charlesworth A. G., U. Seroussi, N. J. Lehrbach, M. S. Renaud, A. E. Sundby, *et al.*, 2021 Two isoforms of the essential *C. elegans* Argonaute CSR-1 differentially regulate sperm and oocyte fertility. Nucleic Acids Res. 49: 8836–8865. https://doi.org/10.1093/nar/gkab619

25. Chen J., and E. J. J. Wagner, 2010 snRNA 3′ end formation: the dawn of the Integrator complex. Biochem. Soc. Trans. 38: 1082–7. https://doi.org/10.1042/BST0381082

26. Chen J., B. Waltenspiel, W. D. Warren, and E. J. Wagner, 2013 Functional analysis of the integrator subunit 12 identifies a microdomain that mediates activation of the drosophila integrator complex. J. Biol. Chem. 288: 4867–77. https://doi.org/10.1074/jbc.M112.425892

27. Chen L., L. Duan, M. Sun, Z. Yang, H. Li, *et al.*, 2023 Current trends and insights on EMS mutagenesis application to studies on plant abiotic stress tolerance and development. Front. Plant Sci. 13. https://doi.org/10.3389/fpls.2022.1052569

28. Chomyshen S. C., H. Tabarraei, and C. W. Wu, 2022 Translational suppression via IFG-1/eIF4G inhibits stress-induced RNA alternative splicing in *Caenorhabditis elegans*. Genetics 221: iyac075. https://doi.org/10.1093/GENETICS/IYAC075

29. Claycomb J. M., P. J. Batista, K. M. Pang, W. Gu, J. J. Vasale, *et al.*, 2009 The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation. Cell 139: 123–134. https://doi.org/10.1016/J.CELL.2009.09.014

30. Cohen-Fix O., and P. Askjaer, 2017 Cell biology of the *Caenorhabditis elegan*s nucleus. Genetics 205. https://doi.org/10.1534/genetics.116.197160

31. Conradt B., Y.-C. Wu, and D. Xue, 2016 Programmed Cell Death During *Caenorhabditis elegans* Development. Genetics 203: 1533–1562. https://doi.org/10.1534/genetics.115.186247

32. D'Angelo M. A., M. Raices, S. H. Panowski, and M. W. Hetzer, 2009 Age-dependent deterioration of nuclear pore complexes causes a loss of nuclear integrity in postmitotic cells. Cell 136: 284–295. https://doi.org/10.1016/J.CELL.2008.11.037

33. Dickinson D. J., J. D. Ward, D. J. Reiner, and B. Goldstein, 2013 Engineering the *Caenorhabditis elegans* genome using Cas9-triggered homologous recombination. Nat. Methods 2013 1010 10: 1028–1034. https://doi.org/10.1038/nmeth.2641

34. Dickinson D. J., A. M. Pani, J. K. Heppert, C. D. Higgins, and B. Goldstein, 2015 Streamlined genome engineering with a self-excising drug selection cassette. Genetics 200: 1035–1049. https://doi.org/10.1534/GENETICS.115.178335/-/DC1

35. Doitsidou M., R. J. Poole, S. Sarin, H. Bigelow, and O. Hobert, 2010 *C. elegans* mutant identification with a one-step whole-genome-sequencing and SNP mapping strategy. PLoS One 5: e15435. https://doi.org/10.1371/journal.pone.0015435

36. Eliceiri G. L., and M. S. Sayavedra, 1976 Small RNAs in the nucleus and cytoplasm of HeLa

cells. Biochem. Biophys. Res. Commun. 72: 507–512. https://doi.org/10.1016/s0006-291x(76)80070-8

37. Elrod N. D., T. Henriques, K.-L. Huang, D. C. Tatomer, J. E. Wilusz, *et al.*, 2019 The Integrator Complex Attenuates Promoter-Proximal Transcription at Protein-Coding Genes. Mol. Cell 76: 738-752.e7. https://doi.org/10.1016/j.molcel.2019.10.034

38. Ezzeddine N., J. Chen, B. Waltenspiel, B. Burch, T. Albrecht, *et al.*, 2010 A Subset of Drosophila Integrator Proteins Is Essential for Efficient U7 snRNA and Spliceosomal snRNA 3'-End Formation. Mol. Cell. Biol. 31: 328–341. https://doi.org/10.1128/MCB.00943-10

39. Fassnacht C., C. Tocchini, P. Kumari, D. Gaidatzis, M. B. Stadler, *et al.*, 2018 The CSR-1 endogenous RNAi pathway ensures accurate transcriptional reprogramming during the oocyte-to-embryo transition i*n Caenorhabditis elegans*. PLOS Genet. 14: e1007252. https://doi.org/10.1371/journal.pgen.1007252

40. Federico A., M. Rienzo, C. Abbondanza, V. Costa, A. Ciccodicola, *et al.*, 2017 Pan-cancer mutational and transcriptional analysis of the integrator complex. Int. J. Mol. Sci. 18: E936. https://doi.org/10.3390/ijms18050936

41. Fianu I., M. Ochmann, J. L. Walshe, O. Dybkov, J. N. Cruz, *et al.*, 2024 Structural basis of Integrator-dependent RNA polymerase II termination. Nature 629: 219–227. https://doi.org/10.1038/S41586-024-07269-4

42. Fire A., S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, *et al.*, 1998 Potent and specific genetic interference by double-stranded RNA in caenorhabditis elegans. Nature 391. https://doi.org/10.1038/35888

43. Fischer S. E. J., Q. Pan, P. C. Breen, Y. Qi, Z. Shi, *et al.*, 2013 Multiple small RNA pathways regulate the silencing of repeated and foreign genes in *C. elegans*. Genes Dev. 27: 2678–2695. https://doi.org/10.1101/gad.233254.113

44. Forget D., A.-A. Lacombe, P. Cloutier, M. Lavallée-Adam, M. Blanchette, *et al.*, 2013 Nuclear import of RNA polymerase II is coupled with nucleocytoplasmic shuttling of the RNA polymerase II-associated protein 2. Nucleic Acids Res. 41: 6881–6891. https://doi.org/10.1093/nar/gkt455

45. Galy V., I. W. Mattaj, and P. Askjaer, 2003 *Caenorhabditis elegans* Nucleoporins Nup93 and Nup205 Determine the Limit of Nuclear Pore Complex Size Exclusion In Vivo. Mol. Biol. Cell 14: 5104–5115. https://doi.org/10.1091/mbc.E03-04

46. Gardini A., D. Baillat, M. Cesaroni, D. Hu, J. M. Marinis, *et al.*, 2014 Integrator regulates transcriptional initiation and pause release following activation. Mol. Cell 56: 128–39. https://doi.org/10.1016/j.molcel.2014.08.004

47. Gerson-Gurwitz A., S. Wang, S. Sathe, R. Green, G. W. Yeo, *et al.*, 2016 A Small RNA-Catalytic Argonaute Pathway Tunes Germline Transcript Levels to Ensure Embryonic Divisions. Cell 165: 396–409. https://doi.org/10.1016/j.cell.2016.02.040

48. Gómez-Orte E., B. Sáenz-Narciso, A. Zheleva, B. Ezcurra, M. de Toro, *et al.*, 2019 Disruption of the *Caenorhabditis elegans* Integrator complex triggers a non-conventional transcriptional mechanism beyond snRNA genes. PLoS Genet. 15: e1007981.

49. Grishok A., A. E. Pasquinelli, D. Conte, N. Li, S. Parrish, *et al.*, 2001 Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. Cell 106: 23–34. https://doi.org/10.1016/S0092-8674(01)00431-7

50. Grishok A., 2005 RNAi mechanisms in Caenorhabditis elegans. FEBS Lett. 579: 5932–5939.

51. Grishok A., J. L. Sinskey, and P. A. Sharp, 2005 Transcriptional silencing of a transgene by RNAi in the soma of *C. elegans.* Genes Dev. 19: 683–696. https://doi.org/10.1101/gad.1247705

52. Han S. K., D. Lee, S.-J. V. Lee, D. Kim, H. G. Son, *et al.*, 2016 OASIS 2: online application for survival analysis 2 with features for the analysis of maximal lifespan and healthspan in aging research. Oncotarget 7: 56147–56152. https://doi.org/10.18632/oncotarget.11269

53. Hata T., and M. Nakayama, 2007 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 - Mol. Cell Res. 1773: 1039–1051. https://doi.org/10.1016/j.bbamcr.2007.04.010

54. History of research on *C. elegans* and other free-living nematodes as model organisms, 2017 WormBook 1–84. https://doi.org/10.1895/wormbook.1.181.1

55. Hobert O., 2002 PCR Fusion-Based Approach to Create Reporter Gene Constructs for Expression Analysis in Transgenic *C. elegans*. Biotechniques 32: 728–730. https://doi.org/10.2144/02324BM01

56. Huang K.-L., D. Jee, C. B. Stein, N. D. Elrod, T. Henriques, *et al.*, 2020 Integrator Recruits Protein Phosphatase 2A to Prevent Pause Release and Facilitate Transcription Termination. Mol. Cell 80: 345-358.e9. https://doi.org/10.1016/j.molcel.2020.08.016

57. Hutvagner G., and M. J. Simard, 2008 Argonaute proteins: key players in RNA silencing. Nat. Rev. Mol. Cell Biol. 2008 91 9: 22–32. https://doi.org/10.1038/nrm2321

58. Jennings B. H., 2011 Drosophila – a versatile model in biology & medicine. Mater. Today 14: 190–195. https://doi.org/10.1016/S1369-7021(11)70113-4

59. Jinek M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, *et al.*, 2012 A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337: 816–821. https://doi.org/10.1126/SCIENCE.1225829

60. Jorgensen E. M., and S. E. Mango, 2002 The art and design of genetic screens: *Caenorhabditis elegans.* Nat. Rev. Genet. 5: 179–89. https://doi.org/10.1038/nrg794

61. Kaletta T., and M. O. Hengartner, 2006 Finding function in novel targets: *C. elegans* as a model organism. Nat. Rev. Drug Discov. 5: 387–98. https://doi.org/10.1038/nrd2031

62. Kamath R. S., M. Martinez-Campos, P. Zipperlen, A. G. Fraser, and J. Ahringer, 2001 Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *Caenorhabditis elegans*. Genome Biol. 2: 1–10. https://doi.org/10.1186/GB-2000-2-1-RESEARCH0002/FIGURES/4

63. Kamath R. S., and J. Ahringer, 2003 Genome-wide RNAi screening in *Caenorhabditis elegans*. Methods 30: 313–321. https://doi.org/10.1016/S1046-2023(03)00050-1

64. Karijolich J., and Y.-T. Yu, 2010 Spliceosomal snRNA modifications and their function. RNA Biol. 7: 192–204. https://doi.org/10.4161/rna.7.2.11207

65. Kelly W. G., and A. Fire, 1998 Chromatin silencing and the maintenance of a functional germline in *Caenorhabditis elegans*. Development 125: 2451–2456. https://doi.org/10.1242/dev.125.13.2451

66. Kim Y., Y. Park, J. Hwang, and K. Kwack, 2018 Comparative genomic analysis of the human and nematode *Caenorhabditis elegans* uncovers potential reproductive genes and disease associations in humans. Physiol. Genomics 50: 1002–1014. https://doi.org/10.1152/physiolgenomics.00063.2018

67. Kim D., J. M. Paggi, C. Park, C. Bennett, and S. L. Salzberg, 2019 Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. 37: 907–915. https://doi.org/10.1038/s41587-019-0201-4

68. Kimble J., and C. Nüsslein-Volhard, 2022 The great small organisms of developmental genetics: *Caenorhabditis elegans* and *Drosophila melanogaster*. Dev. Biol. 485: 93–122.

69. Kipreos E. T., and S. van den Heuvel, 2019 Developmental Control of the Cell Cycle: Insights from *Caenorhabditis elegans.* Genetics 211: 797–829.

https://doi.org/10.1534/genetics.118.301643

70. Kiss T., 2004 Biogenesis of small nuclear RNPs. J. Cell Sci. 117: 5949–5951. https://doi.org/10.1242/jcs.01487

71. Korta D. Z., and E. J. A. Hubbard, 2010 Soma-germline interactions that influence germline proliferation in *Caenorhabditis elegans*. Dev. Dyn. 239: 1449–1459. https://doi.org/10.1002/DVDY.22268

72. Krall M., S. Htun, R. E. Schnur, A. S. Brooks, L. Baker, *et al.*, 2019 Biallelic sequence variants in INTS1 in patients with developmental delays, cataracts, and craniofacial anomalies. Eur. J. Hum. Genet. 2019 274 27: 582–593. https://doi.org/10.1038/s41431-018-0298-9

73. Leiser S. F., G. Jafari, M. Primitivo, G. L. Sutphin, J. Dong, *et al.*, 2016 Age-associated vulval integrity is an important marker of nematode healthspan. Age (Omaha). 38: 419–431. https://doi.org/10.1007/s11357-016-9936-8

74. Marzluff W. F., E. J. Wagner, and R. J. Duronio, 2008 Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. Nat. Rev. Genet. 2008 911 9: 843–854. https://doi.org/10.1038/nrg2438

75. Mascibroda L. G., M. Shboul, N. D. Elrod, L. Colleaux, H. Hamamy, *et al.*, 2022 INTS13 variants causing a recessive developmental ciliopathy disrupt assembly of the Integrator complex. Nat. Commun. 13: 6054. https://doi.org/10.1038/s41467-022-33547-8

76. Nance J., and C. Frøkjær-Jensen, 2019 The *Caenorhabditis elegans* Transgenic Toolbox. Genetics 212: 959. https://doi.org/10.1534/GENETICS.119.301506

77. Nguyen D. A. H., and C. M. Phillips, 2021 Arginine methylation promotes siRNA-binding specificity for a spermatogenesis-specific isoform of the Argonaute protein CSR-1. Nat. Commun. 2021 121 12: 1–15. https://doi.org/10.1038/s41467-021-24526-6

78. Nilsen T. W., and B. R. Graveley, 2010 Expansion of the eukaryotic proteome by alternative splicing. Nature 463: 457–63. https://doi.org/10.1038/nature08909

79. Nishimura K., T. Fukagawa, H. Takisawa, T. Kakimoto, and M. Kanemaki, 2009 An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nat. Methods 6: 917–922. https://doi.org/10.1038/nmeth.1401

80. Oegema R., D. Baillat, R. Schot, L. M. van Unen, A. Brooks, *et al.*, 2017 Human mutations in integrator complex subunits link transcriptome integrity to brain development. PLoS Genet. 13: e1006809. https://doi.org/10.1371/journal.pgen.1006809

81. Ohno M., A. Segref, A. Bachi, M. Wilm, and I. W. Mattaj, 2000 PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. Cell 101: 187–198. https://doi.org/10.1016/S0092-8674(00)80829-6

82. Paix A., A. Folkmann, and G. Seydoux, 2017 Precision genome editing using CRISPR-Cas9 and linear repair templates in *C. elegans.* Methods 121–122: 86–93. https://doi.org/10.1016/j.ymeth.2017.03.023

83. Pemberton L. F., and B. M. Paschal, 2005 Mechanisms of Receptor-Mediated Nuclear Import and Nuclear Export. Traffic 6: 187–198. https://doi.org/10.1111/j.1600-0854.2005.00270.x

84. Pfleiderer M. M., and W. P. Galej, 2021 Emerging insights into the function and structure of the Integrator complex. Transcription 12: 251–265. https://doi.org/10.1080/21541264.2022.2047583/ASSET/BDC5E5CA-1FE4-4926-A5DE-7BD2DCBB8800/ASSETS/IMAGES/KTRN_A_2047583_F0003_OC.JPG

85. Phillips C. M., T. A. Montgomery, P. C. Breen, and G. Ruvkun, 2012 MUT-16 promotes formation of perinuclear Mutator foci required for RNA silencing in the *C. elegans* germline. Genes Dev. 26. https://doi.org/10.1101/gad.193904.112

86. Price I. F., H. L. Hertz, B. Pastore, J. Wagner, and W. Tang, 2021 Proximity labeling

identifies LOTUS domain proteins that promote the formation of perinuclear germ granules in C. elegans. Elife 10. https://doi.org/10.7554/eLife.72276

87. Proudfoot N. J., 2011 Ending the message: poly(A) signals then and now. Genes Dev. 25: 1770–1782. https://doi.org/10.1101/gad.17268411

88. Quarato P., M. Singh, E. Cornes, B. Li, L. Bourdon, *et al.*, 2021 Germline inherited small RNAs facilitate the clearance of untranslated maternal mRNAs in *C. elegans* embryos. Nat Commun 12: 1441. https://doi.org/10.1038/s41467-021-21691-6

89. Riddle D., T. Blumenthal, and B. Meyer, 1997 *Section 1 The Biological Model*. ColSpringHarborLaboratoryPress.

90. Sabath K., M. L. Stäubli, S. Marti, A. Leitner, M. Moes, *et al.*, 2020 INTS10–INTS13–INTS14 form a functional module of Integrator that binds nucleic acids and the cleavage module. Nat. Commun. 11: 3422. https://doi.org/10.1038/s41467-020-17232-2

91. Sakaguchi A., P. Sarkies, M. Simon, A.-L. Doebley, L. D. Goldstein, *et al.*, 2014 *Caenorhabditis elegans* RSD-2 and RSD-6 promote germ cell immortality by maintaining small interfering RNA populations. Proc. Natl. Acad. Sci. 111. https://doi.org/10.1073/pnas.1406131111

92. Schaner C., 2006 Germline chromatin. WormBook. https://doi.org/10.1895/wormbook.1.73.1

93. Schetter A., P. Askjaer, F. Piano, I. Mattaj, and K. Kemphues, 2006 Nucleoporins NPP-1, NPP-3, NPP-4, NPP-11 and NPP-13 are required for proper spindle orientation in *C. elegans.* Dev. Biol. 289: 360–371. https://doi.org/10.1016/j.ydbio.2005.10.038

94. Schwartz M. L., M. Wayne Davis, M. S. Rich, E. M. Jorgensen, M. W. Davis, *et al.*, 2021 High-efficiency CRISPR gene editing in *C. elegans* using Cas9 integrated into the genome. PLOS Genet. 17: e1009755. https://doi.org/10.1371/journal.pgen.1009755

95. Seroussi U., A. Lugowski, L. Wadi, R. X. Lao, A. R. Willis, *et al.*, 2023 A comprehensive survey of *C. elegans* argonaute proteins reveals organism-wide gene regulatory networks and functions. Elife 12: 83853. https://doi.org/10.7554/ELIFE.83853

96. Seydoux G., 2018 The P Granules of C. elegans: A Genetic Model for the Study of RNA–Protein Condensates. J. Mol. Biol. 430: 4702–4710. https://doi.org/10.1016/j.jmb.2018.08.007

97. Shen S., J. W. Park, Z. Lu, L. Lin, M. D. Henry, *et al.*, 2014 rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc. Natl. Acad. Sci. U. S. A. 111: E5593-601. https://doi.org/10.1073/pnas.1419161111

98. Shi Y., 2009 Serine/Threonine Phosphatases: Mechanism through Structure. Cell 139: 468–484. https://doi.org/10.1016/j.cell.2009.10.006

99. Sijen T., J. Fleenor, F. Simmer, K. L. Thijssen, S. Parrish, *et al.*, 2001 On the Role of RNA Amplification in dsRNA-Triggered Gene Silencing. Cell 107: 465–476. https://doi.org/10.1016/S0092-8674(01)00576-1

100. Simmer F., M. Tijsterman, S. Parrish, S. P. Koushika, M. L. Nonet, *et al.*, 2002 Loss of the putative RNA-directed RNA polymerase RRF-3 makes C. elegans hypersensitive to RNAi. Curr. Biol. 12: 1317–1319. https://doi.org/10.1016/S0960-9822(02)01041-2

101. Singh M., E. Cornes, B. Li, P. Quarato, L. Bourdon, *et al.*, 2021 Translation and codon usage regulate Argonaute slicer activity to trigger small RNA biogenesis. Nat Commun 12: 3492. https://doi.org/10.1038/s41467-021-23615-w

102. Skaar J. R., A. L. Ferris, X. Wu, A. Saraf, K. K. Khanna, *et al.*, 2015 The Integrator complex controls the termination of transcription at diverse classes of gene targets. Cell Res. 25: 288–305. https://doi.org/10.1038/cr.2015.19

103. Stadelmayer B., G. Micas, A. Gamot, P. Martin, N. Malirat, *et al.*, 2014 Integrator complex

regulates NELF-mediated RNA polymerase II pause/release and processivity at coding genes. Nat. Commun. 5. https://doi.org/10.1038/NCOMMS6531

104. Stein C. B., A. R. Field, C. A. Mimoso, C. C. Zhao, K. L. Huang, *et al.*, 2022 Integrator endonuclease drives promoter-proximal termination at all RNA polymerase II-transcribed loci. Mol. Cell 82: 4232-4245.e11. https://doi.org/10.1016/J.MOLCEL.2022.10.004

105. Steiner F. A., K. L. Okihara, S. W. Hoogstrate, T. Sijen, and R. F. Ketting, 2009 RDE-1 slicer activity is required only for passenger-strand cleavage during RNAi in Caenorhabditis elegans. Nat. Struct. Mol. Biol. 16: 207–211. https://doi.org/10.1038/nsmb.1541

106. Sundby A. E., R. I. Molnar, and J. M. Claycomb, 2021 Connecting the Dots: Linking *Caenorhabditis elegans* Small RNA Pathways and Germ Granules. Trends Cell Biol. 31: 387–401. https://doi.org/10.1016/j.tcb.2020.12.012

107. Tabara H., M. Sarkissian, W. G. Kelly, J. Fleenor, A. Grishok, *et al.*, 1999 The rde-1 gene, RNA interference, and transposon silencing in *C. elegans*. Cell 99: 123–32. https://doi.org/10.1016/S0092-8674(00)81644-X

108. Takata H., H. Nishijima, K. Maeshima, and K.-I. Shibahara, 2012 The integrator complex is required for integrity of Cajal bodies. J. Cell Sci. 125: 166–175. https://doi.org/10.1242/jcs.090837

109. Tepe B., E. L. Macke, M. Niceta, M. Weisz Hubshman, O. Kanca, *et al.*, 2023 Bi-allelic variants in INTS11 are associated with a complex neurological disorder. Am. J. Hum. Genet. 110: 774–789. https://doi.org/10.1016/j.ajhg.2023.03.012

110. Thomas J., K. Lea, E. Zucker-Aprison, and T. Blumenthal, 1990 The spliceosomal snRNAs of *Caenorhabditis elegans*. Nucleic Acids Res. 18: 2633–2642.

111. Tijsterman M., R. C. May, F. Simmer, K. L. Okihara, and R. H. A. Plasterk, 2004 Genes required for systemic RNA interference in *Caenorhabditis elegans*. Curr. Biol. 14: 111–116.

https://doi.org/10.1016/J.CUB.2003.12.029

112. Vanderwaeren L., R. Dok, K. Voordeckers, S. Nuyts, and K. J. Verstrepen, 2022 Saccharomyces cerevisiae as a Model System for Eukaryotic Cell Biology, from Cell Cycle Control to DNA Damage Response. Int. J. Mol. Sci. 23: 11665. https://doi.org/10.3390/ijms231911665

113. Voronina E., and G. Seydoux, 2010 The *C. elegans* homolog of nucleoporin Nup98 is required for the integrity and function of germline P granules. Development 137: 1441–1450. https://doi.org/10.1242/DEV.047654/-/DC1

114. Waddell B. M., and C. W. Wu, 2024 A role for the *C. elegans* Argonaute protein CSR-1 in small nuclear RNA 3' processing. PLoS Genet. 20. https://doi.org/10.1371/JOURNAL.PGEN.1011284

115. Wagner E. J., L. Tong, and K. Adelman, 2023 Integrator is a global promoter-proximal termination complex. Mol. Cell 83: 416–427. https://doi.org/10.1016/j.molcel.2022.11.012

116. Walker J. R., R. A. Corpina, and J. Goldberg, 2001 Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. Nature 412: 607–614. https://doi.org/10.1038/35088000

117. Walther T. C., A. Alves, H. Pickersgill, I. Loïodice, M. Hetzer, *et al.*, 2003 The conserved Nup107-160 complex is critical for nuclear pore complex assembly. Cell 113: 195–206. https://doi.org/10.1016/S0092-8674(03)00235-6

118. Wedeles C. J., M. Z. Wu, and J. M. Claycomb, 2013 Protection of Germline Gene Expression by the *C. elegans* Argonaute CSR-1. Dev. Cell 27: 664–671. https://doi.org/10.1016/J.DEVCEL.2013.11.016

119. Welsh S. A., and A. Gardini, 2023 Genomic regulation of transcription and RNA processing by the multitasking Integrator complex. Nat. Rev. Mol. Cell Biol. 24: 204–220.

https://doi.org/10.1038/s41580-022-00534-2

120. Wieben E. D., J. M. Nenninger, and T. Pederson, 1985 Ribonucleoprotein organization of eukaryotic RNA. J. Mol. Biol. 183: 69–78. https://doi.org/10.1016/0022-2836(85)90281-5

121. Wieland I., K. C. Arden, D. Michels, L. Klein-Hitpass, M. Böhm, *et al.*, 1999 Isolation of DICE1: A gene frequently affected by LOH and downregulated in lung carcinomas. Oncogene 18: 4530–7. https://doi.org/10.1038/sj.onc.1202806

122. Wilkinson M. E., C. Charenton, and K. Nagai, 2020 RNA Splicing by the Spliceosome. Annu Rev Biochem 89: 359–388. https://doi.org/10.1146/ANNUREV-BIOCHEM-091719-064225

123. Will C. L., and R. Lührmann, 2011 Spliceosome structure and function. Cold Spring Harb. Perspect. Biol. 3: pii: a003707. https://doi.org/10.1101/cshperspect.a003707

124. Wong C. Y. Y., H. N. Tsui, Y. Wang, and K. W. Y. Yuen, 2024 Argonaute CSR-1 restricts holocentromere protein CENP-A/HCP-3 localization. J. Cell Sci. 137. https://doi.org/10.1242/JCS.261895/VIDEO-2

125. Wu C.-W. W., K. Wimberly, A. Pietras, W. Dodd, M. B. Atlas, *et al.*, 2019 RNA processing errors triggered by cadmium and integrator complex disruption are signals for environmental stress. BMC Biol. 17: 56. https://doi.org/10.1186/s12915-019-0675-z

126. Yang H., Y. Zhang, J. Vallandingham, H. Li, L. Florens, *et al.*, 2012 The RDE-10/RDE-11 complex triggers RNAi-induced mRNA degradation by association with target mRNA in *C. elegans.* Genes Dev. 26: 846–856. https://doi.org/10.1101/gad.180679.111

127. Yigit E., P. J. Batista, Y. Bei, K. M. Pang, C.-C. G. C. G. Chen, *et al.*, 2006 Analysis of the *C. elegans* Argonaute Family Reveals that Distinct Argonautes Act Sequentially during RNAi. Cell 127: 747–757. https://doi.org/10.1016/j.cell.2006.09.033

128. Youngman E. M., and J. M. Claycomb, 2014 From early lessons to new frontiers: The worm as a treasure trove of small RNA biology. Front. Genet. 5: 416. https://doi.org/10.3389/FGENE.2014.00416/BIBTEX

129. Zahler A. M., 2012 Pre-mRNA splicing and its regulation in Caenorhabditis elegans. WormBook.

130. Zhang C., T. A. Montgomery, H. W. Gabel, S. E. J. Fischer, C. M. Phillips, *et al.*, 2011 mut-16 and other mutator class genes modulate 22G and 26G siRNA pathways in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. U. S. A. 108: 1201–1208. https://doi.org/10.1073/PNAS.1018695108/-/DCSUPPLEMENTAL

131. Zhang C., T. A. Montgomery, S. E. J. Fischer, S. M. D. A. Garcia, C. G. Riedel, *et al.*, 2012 The Caenorhabditis elegans RDE-10/RDE-11 complex regulates RNAi by promoting secondary siRNA amplification. Curr. Biol. 22. https://doi.org/10.1016/j.cub.2012.04.011

132. Zhang L., J. D. Ward, Z. Cheng, and A. F. Dernburg, 2015 The auxin-inducible degradation (AID) system enables versatile conditional protein depletion in *C. elegans.* Development 142: 4374–4384. https://doi.org/10.1242/dev.129635

133. Zhang X., Y. Wang, F. Yang, J. Tang, X. Xu, *et al.*, 2020 Biallelic INTS1 Mutations Cause a Rare Neurodevelopmental Disorder in Two Chinese Siblings. J. Mol. Neurosci. 70: 1–8. https://doi.org/10.1007/s12031-019-01393-x

134. Zimmer M., 2009 GFP: from jellyfish to the Nobel prize and beyond. Chem. Soc. Rev. 38: 2823. https://doi.org/10.1039/b904023d