LITTLE ELONGATION COMPLEX (LEC) AND SUPER ELONGATION COMPLEX (SEC) AS REGULATORS OF TDP-43-ASSOCIATED NEURODEGENERATION

Chia-Yu Chung

A DISSERTATION

in

Cell and Molecular Biology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2018

Supervisor of Dissertation

_____

Nancy M. Bonini, PhD
Florence R.C. Murray Professor of Biology; Investigator of the Howard Hughes Medical Institute

Graduate Group Chairperson

_____

Daniel Kessler, PhD
Associate Professor of Cell and Developmental Biology

Dissertation Committee

Thomas A. Jongens, PhD
Associate Professor of Genetics

Shelley L. Berger, Ph.D.
Daniel S. Och University Professor

Alice S. Chen-Plotkin, MD
Parker Family Associate Professor of Neurology

Kenneth S. Zaret, PhD
Joseph Leidy Professor

Zhaolan (Joe) Zhou, PhD
Associate Professor of Genetics

LITTLE ELONGATION COMPLEX (LEC) AND SUPER ELONGATION COMPLEX (SEC) AS REGULATORS OF TDP-43-ASSOCIATED NEURODEGENERATION

ACKNOWLEDGMENTS

Six years ago, I came here to study from Taiwan with excitement and anxiety, and it has been a wonderful journey during which I have learned so much. I am truly grateful to my advisor, Dr. Nancy Bonini, for the guidance, support and brilliant advice throughout my thesis work. The solid training in Nancy's lab has equipped me with strong experimental techniques as well as logical and critical thinking as an independent scientist. Nancy has been a great inspiration and an encouraging mentor. In addition, I am surrounded by exceptional colleagues, and I thank them for all of their help and for all of the fun we have had in the last few years. I would especially like to thank Leeanne, who guided me during my rotation and has helped me a lot in many ways; Jason, Amit and Ashley, who collaborated with me on the research; and Inny, who has been kindly shared experiences and tips in different topics with me. I would also like to thank my thesis committee members, Dr. Tom Jongens, Dr. Shelley Berger, Dr. Alice Chen-Plotkin, Dr. Ken Zaret and Dr. Joe Zhou for insightful and valuable comments. I would like to thank Tom for being my committee chair and all of the help, to thank Joe for the advice and help in various aspects since I started my PhD, and to thank Alice for the collaboration on extending our finding in fly disease models to human patient samples. I also want to thank all other collaborators for the help on this work.

I would like to express my special appreciation to my family, who has always made me feel being loved and supported. Words cannot express how grateful I am to my parents for the love they have for me. Although they would prefer I could be nearer to them, they are fully supportive of my choice. Finally, I would like to express my deep appreciation to my husband, Ting-Yu, with whom I share all of my happiness and frustration. Although we are on opposite sides of the earth, he has always been my strongest support.

ABSTRACT

LITTLE ELONGATION COMPLEX (LEC) AND SUPER ELONGATION COMPLEX (SEC) AS REGULATORS OF TDP-43-ASSOCIATED NEURODEGENERATION

Chia-Yu Chung

Nancy M. Bonini

TDP-43 aggregation is the pathological hallmark of amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration with ubiquitinated inclusions (FTLD-TDP). To define pathways important in TDP-43 proteinopathy, a genetic screen for modifiers of TDP-43-mediated eye degeneration was conducted in the fruit fly (*Drosophila melanogaster*). This approach and the following studies in the fly identified transcriptional elongation factor Ell as a potent dose-dependent modifier. Ell is a shared component of little elongation complex (LEC) and super elongation complex (SEC). Our studies show that downregulation of additional components in LEC or SEC partially but consistently ameliorates TDP-43-associated eye degeneration, implicating both Ell-associated complexes in TDP-43 toxicity. LEC regulates transcription of RNA Polymerase II (Pol II)-transcribed small nuclear RNAs (snRNAs), while SEC mainly regulates transcription of select inducible genes harboring paused Pol II including stress-induced loci. We hypothesized that, in the disease state, TDP-43 increases the activity of LEC and SEC, leading to enhanced expression of select target genes that contribute to neurodegeneration. Our results support this hypothesis by showing that select targets of LEC and SEC become upregulated in fly heads expressing TDP-43 compared to controls. Among upregulated targets, U12 snRNA and a stress-induced long non-coding RNA (lncRNA) *Hsrω* were shown to functionally contribute to TDP-43-induced

degeneration in *Drosophila*. The increase in U12 snRNA leads to elevated activity of the U12-dependent spliceosome, assessed by examining U12-dependent splicing events. Among the known U12 targets, the splicing of genes CG15735, CG16941 and CG11839 were shown to be upregulated selectively by TDP-43. In addition, the novel target of TDP-43 *Hsrω* was identified by polytene chromosome association of TDP-43 and SEC components Lilli and Ell in the fly. We extended the findings of *Hsrω* to show that the human orthologue Satellite III repeat (Sat III) is elevated in both a human cellular disease model of TDP-43 toxicity, and FTLD-TDP patient tissue. Furthermore, TDP-43 was shown to interact with human ELL2 by co-immunoprecipitation from human HEK293 cells. These findings support a model whereby TDP-43 promotes the LEC and SEC activities through an interaction with ELL2, leading to abnormal activation of LEC and SEC target genes that contribute to degeneration. These studies reveal the critical roles of Ell and Ell-associated complexes in TDP-43 toxicity, identify important downstream targets, and provide potential therapeutic strategies to combat TDP-43-associated neurodegeneration.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1 : INTRODUCTION

THE ROLES OF ELL-CONTAINING TRANSCRIPTIONAL ELONGATION COMPLEXES IN TDP-43-ASSOCIATED NEURODEGENERATION

ALS/FTD and TDP-43

Amyotrophic lateral sclerosis (ALS; OMIM no. 105400) and frontotemporal dementia (FTD; OMIM no. 600274) are two neurological disorders with distinct clinical features yet share common pathogenic mechanisms [1-4]. ALS is the most common motor neuron disease, resulting from loss of upper and lower motor neurons in the motor cortex, the brainstem and spinal cord, which leads to fatal paralysis. FTD is the second most common cause of early onset dementia, caused by progressive neuron death in the frontal and temporal lobes. This results in changes in behavior, personality and/or language. Increasingly, evidence indicates that ALS and FTD may be within the same disease spectrum. First, about 15% of ALS patients develop the symptoms of FTD and vice versa [5-7]. Second, genetic mutations in shared genes link these two diseases [1,2,8]. Among the genes bearing causal mutations for both diseases, TAR DNA-binding protein 43 (TDP-43) further connects these two diseases. In 97% of ALS and 45% of FTD (FTLD-TDP/FTLD-U) patients, TDP-43 was identified to be the major component of aggregates in the central nervous system [9,10]. In the neurons and glia containing TDP-43 inclusions, physiological diffused nuclear TDP-43 is cleared and cytoplasmic aggregates of full length and truncated TDP-43 are observed with abnormal hyperphosphorylation and ubiquitination [9,10].

TDP-43 is a DNA and RNA binding protein with the molecular weight of 43 kDa, which was discovered and named for its ability to bind transactive response (TAR) element of human immunodeficiency virus type 1 (HIV-1) [11]. The total length of TDP-43 is 414 amino acids. It contains a N-terminal domain (NTD) including a nuclear localization signal, two RNA recognition motifs (RRMs) with a nuclear export signal within the second RRM domain and a C-terminal domain (CTD). The CTD containing QN-rich

residues is implicated in the aggregation propensity of TDP-43 and is where most of ALS/FTD-associated mutations are located [12-14]. TDP-43 is involved in various cellular processes, including splicing regulation, microRNA processing, RNA stability, gene expression, RNA translocation between nucleus and cytoplasm, stress granule regulation, cell proliferation and apoptosis [13,15]. Research in different disease models shows that both upregulation and depletion of TDP-43 leads to neuron loss, indicating that a balanced level of TDP-43 is critical in the nervous system [13,16,17].

Drosophila as a powerful genetic model system

To understand the mechanisms of TDP-43 protein pathology, animal models have been established. Each disease model has its strength, revealing disease mechanisms from different angles. *Drosophila melanogaster* has a short lifespan, conserved genes and pathways with humans as well as powerful genetics, providing an extraordinary system for unbiased genetic screens to identify novel pathways. The life cycle of each generation of the fruit fly is about 10 days; therefore, large numbers of progeny for research can be obtained in a short period of time. Moreover, although the genome size of *Drosophila* (1.2×10^8 base pairs) is much smaller than the human genome size (3.3×10^9 base pairs), around 70% of known human disease genes have fly orthologs in the fly based on sequence analysis [18,19]. This homology provides an opportunity for elucidation of mechanisms that may be involved in human disease. In fact, the smaller genome size of the fly reduces the complexity of gene redundancy of mammalian systems, and allows quicker and easier interpretation for loss of function studies.

The advantages of using *Drosophila* as a model organism also include a comprehensive and versatile toolbox for genetic manipulation. The fruit fly has been used as a model organism for more than 100 years, and tools have been developed to manipulate the expression of fly genes [20,21]. For example, the *GAL4-UAS* bipartite system is widely used for gene manipulation, allowing spatial and temporal expression of exogenous genetic constructs [22,23]. The system is composed of two parts. In one construct, a tissue specific promoter drives the expression of the yeast transcription factor GAL4; in the second construct, a desired transgene is combined with an upstream activation sequence (UAS), which can be activated by the GAL4. Therefore, the transgene will be selectively expressed in the tissue of interest. Additional regulation and modification of GAL4 can be used to control the temporal expression. For instance, $Gal80^{ts}$, a temperature sensitive inhibitor of the GAL4 protein, is used to turn on the expression system at higher temperature (>30°C), at which the inhibition by $Gal80^{ts}$ is repressed [24]. Another method is by using a ligand-inducible GAL4 chimera, GAL4-progesterone receptor fusion (GeneSwitch), which is activated in the presence of the drug mifepristone (RU486) [25].

Over time, the fly community has generated comprehensive collections of mutants and transgenic fly lines. The Bloomington *Drosophila* Stock center at Indiana University (BDSC) held 67,634 stocks at the end of 2017, including a genome-wide RNAi collection. There are also other *Drosophila* stock centers, such as the Vienna *Drosophila* Research Center (VDRC), Kyoto Stock Center (DGRC), Fly Stocks of National Institute of Genetics (NIG-FLY) and Zurich ORFome Project (FlyORF). Together, these stock centers provide an accessible resource allowing manipulation of almost every gene. The extensive resources and toolbox make *Drosophila* an excellent model organism for

genetic screen. Such screens have highlighted various factors in fly disease models of neurodegenerative disorders [26,27].

Fly models for TDP-43 toxicity

The TDP-43 pathologies, which includes nuclear clearance and cytoplasmic accumulation of TDP-43, suggests the possibility that degeneration may be caused by loss-of-function in the nucleus and/or gain-of-function in the cytoplasm [9,10]. Different fly diseases models have been used to elucidate mechanisms in TDP-43-associated neurodegeneration, including depletion and overexpression of *Drosophila* TDP-43 (TBPH) [28-31], and expression of different forms of human TDP-43 [32-38]. Downregulation of TBPH in the fly leads to neuronal defect and shortened lifespan [28-31], indicating that normal function of TBPH is important for neuronal function and supporting the loss-of-function model. Fly TBPH and human TDP-43 share conserved domains, and sequence analysis shows 59% amino acid identity between them [39]. Both proteins bind to UG repeat sequences [39]. Furthermore, the phenotype caused by TBPH depletion can be rescued by human TDP-43 [28], indicating functional conservation between fly and human TDP-43. Ectopic expression of TBPH or human TDP-43 also recapitulates features of human patients, such as neuron loss, motor disability and decreased lifespan [30-35,37,38]. Furthermore, at the molecular level, expression of fly and human TDP-43 can lead to cytoplasmic TDP-43 aggregation, and ubiquitinated and phosphorylated forms of TDP-43 aggregates, mimicking human molecular pathologies and potentially modeling the gain-of-function toxicity [31,33,36-38]. Therefore, dissecting the pathways involved in neurodegeneration in these different disease models may reveal mechanistic insights of TDP-43 proteinopathy in human disease.

In order to elucidate novel and important factors involved in TDP-43-associated neurodegeneration, studies in our lab using the *Drosophila* eye as a platform for genetic screening. We identified *Drosophila* eleven-nineteen lysine-rich leukemia (Ell; also known as Su(Tpl)) as a potential modifier of TDP-43-mediated degeneration. The *Drosophila* eye provides a platform for high throughput genetic screening since it is a nonessential organ for survival. The fly compound eye is a simple and delicate nervous system composed of thousands of neuronal and nonneuronal cells [40]. The basic unit is ommatidium, which contains 8 photoreceptor neurons and surrounded by 4 lens-secreting cone cells and 2 primary pigment cells. In between each ommatidium, there are 12 accessory cells. Because the structure of the eye is highly ordered, degeneration and cell death can be easily observed as morphological phenotypes, such as rough eye, small eye or change of pigmentation. Expression of the human TDP-43 in fly eye leads to disruption of eye integrity [37,38], and studies in our lab found that downregulation of Ell strongly suppresses the degeneration caused by TDP-43 toxicity.

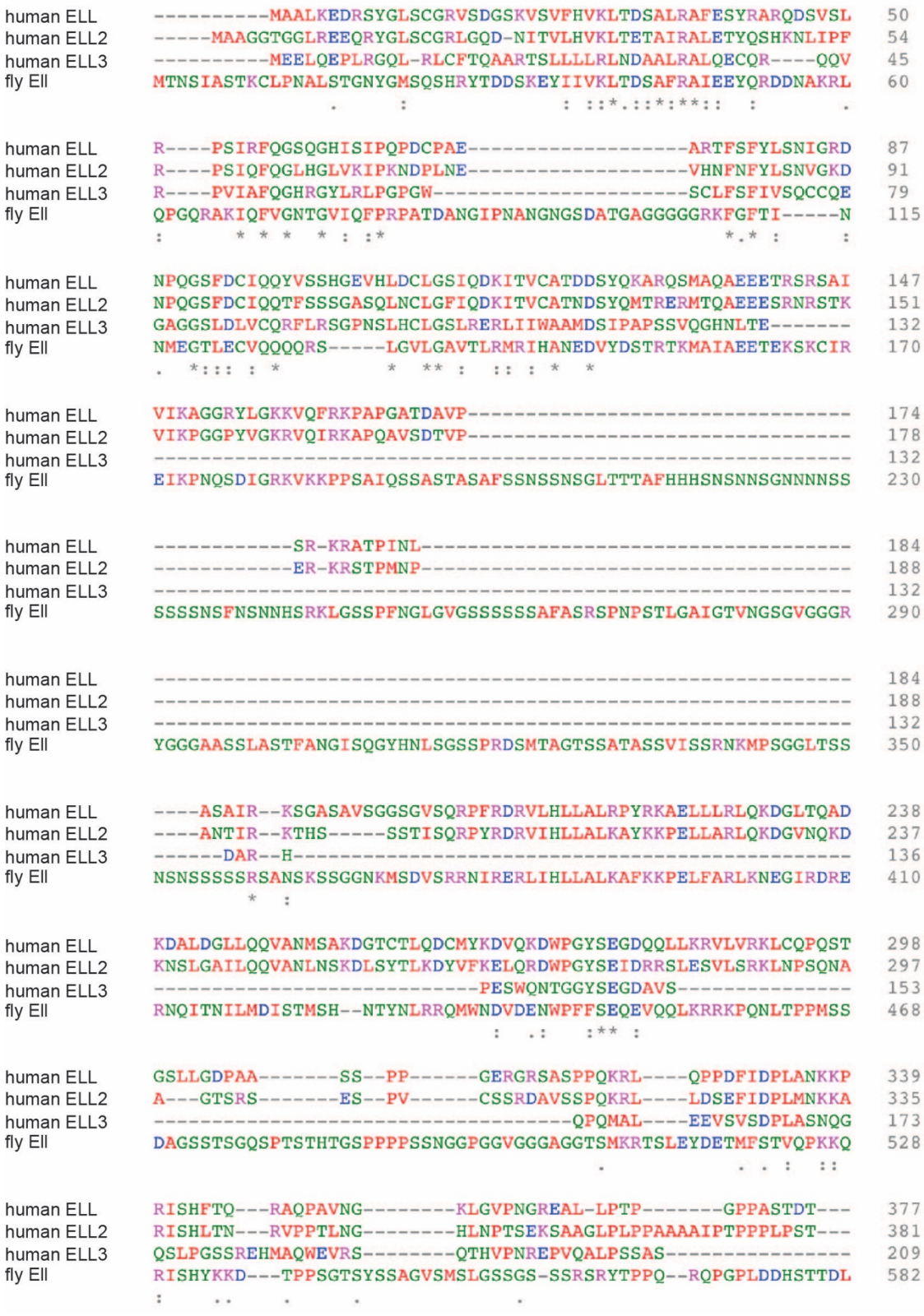
Transcription elongation factor: Ell

RNA Polymerase II (Pol II) is a 12-subunit enzyme that synthesizes protein-coding mRNAs and non-coding RNAs (ncRNAs). The process is tightly regulated at different stages of transcription, including initiation, pausing, elongation and termination. More and more attention has been brought toward the understanding of pausing regulation. After forming the preinitiation complex, Pol II can pause and accumulate in the proximal region (30-60 nucleotides downstream of the transcription start site) [41-43]. The release of paused Pol II into productive elongation requires the kinase, positive transcription elongation factor b (P-TEFb). P-TEFb is composed of the cyclin-dependent kinase 9

(CDK9) and cyclin T1 (CYCT1) or CYCT2. It can phosphorylate negative elongation factor (NELF), 5, 6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) sensitivity-inducing factor (DSIF) and Pol II on the carboxy-terminal domain (CTD). NELF and DSIF are factors working together to set up paused Pol II at the promoter-proximal region. Phosphorylation of these 3 factors leads to dissociation of NELF from Pol II and productive transcription elongation.

Indicated by large scale genomic studies, the release of paused Pol II into elongation is an important rate-limiting step of gene expression, potentially for all active genes [44]. Depending on the methods and cell type, 40%-70% of active genes show accumulation of Pol II at the promoter proximal region. However, studies using drug treatments inhibiting P-TEFb and genome-wide sequencing methods show that prevention of Pol II from releasing halts almost all transcription. These data suggest that P-TEFb dependent pausing regulation may be a highly variable rate-limiting regulatory step for all active genes. The importance of pausing regulation has been highlighted by research investigating its role on transcription subjected to regulatory control and stimuli, such as differentiation signals and heat stress [41-43]. Transcriptional pausing regulation allows rapid induction of transcription as well as enabling fine-tuning of gene activation. The paused Pol II is transcriptionally-engaged and awaiting signaling events, such as developmental cues, physiological signals and stress response. Upon stimulation, poised Pol II can activate gene expression rapidly and serve as a transcription checkpoint for precise and synchronized gene activation upon stimulation.

Several factors regulating transcriptional elongation have been identified by in vitro biochemical assays; these factors include ELL [45]. ELL was isolated from rat liver nuclei extract as a factor that promotes the catalytic rate of transcription elongation by repressing transient pausing of Pol II [46]. Human ELL is a protein containing 621 amino acids, which is also capable of simulating transcription elongation like the rat ELL protein [46]. Following the purification and investigation of ELL, two additional ELL-related proteins, ELL2 and ELL3, were also characterized in mammalian cells [47,48]. Through sequence comparison, *ELL2* and *ELL3* were identified and cloned for in vitro characterization. The *ELL2* open reading frame (ORF) encodes a 640-amino acid protein [48], while *ELL3* ORF encodes a protein containing 397-amino acids, which is testis specific based on Northern blot analysis in different tissues [47]. The alignment of all three proteins by CLUSTAL O (1.2.4) multiple sequence comparison is shown in Figure 1-1. Analysis by Clustal 12.1 shows that ELL and ELL2 share 49.43% identity; ELL and ELL3 share 32.9% identity; ELL2 and ELL3 share 31.97% identity (Table 1-1). Biochemically, all three ELL family proteins can increase the catalytical rate of Pol II transcription elongation [46-48]. In vivo studies indicate that the function of ELL3 is distinct from the other two. By using mouse embryonic stem cells, ELL3 was shown to occupy enhancer regions for priming of gene activation for proper differentiation [49].


human ELL ----------MAALKEDRSYGLSCGRVSDGSKVSVFHVKLTDSALRAFESYRARQDSVSL 50
human ELL2 -----MAAGGTGGLREEQRYGLSCGRLGQD-NITVLHVKLTETAIRALETYQSHKNLIPF 54
human ELL3 ----------MEELQEPLRGQL-RLCFTQAARTSLLLLRLNDAALRALQECQR----QQV 45
fly Ell MTNSIASTKCLPNALSTGNYGMSQSHRYTDDSKEYIIVKLTDSAFRAIEEYQRDDNAKRL 60
human ELL R----PSIRFQGSQGHISIPQPDCPAE------------------ARTFSFYLSNIGRD 87
human ELL2 R----PSIQFQGLHGLVKIPKNDPLNE------------------VHNFNFYLSNVGKD 91
human ELL3 R----PVIAFQGHRGYLRLPGPGW---------------------SCLFSFIVSQCCQE 79
fly Ell QPGQRAKIQFVGNTGVIQFPRPATDANGIPNANGNGSDATGAGGGGGRKFGFTI-----N 115
human ELL NPQGSFDCIQQYVSSHGEVHLDCLGSIQDKITVCATDDSYQKARQSMAQAEEETRSRSAI 147
human ELL2 NPQGSFDCIQQTFSSSGASQLNCLGFIQDKITVCATNDSYQMTRERMTQAEEESRNRSTK 151
human ELL3 GAGGSLDLVCQRFLRSGPNSLHCLGSLRERLIIWAAMDSIPAPSSVQGHNLTE------- 132
fly Ell NMEGTLECVQQQQRS-----LGVLGAVTLRMRIHANEDVYDSTRTKMAIAEETEKSKCIR 170
human ELL VIKAGGRYLGKKVQFRKPAPGATDAVP---------------------------------- 174
human ELL2 VIKPGGPYVGKRVQIRKAPQAVSDTVP---------------------------------- 178
human ELL3 -- 132
fly Ell EIKPNQSDIGRKVKKPPSAIQSSASTASAFSSNSSNSGLTTTAFHHHSNSNNSGNNNNSS 230
human ELL -----------SR-KRATPINL---------------------------------- 184
human ELL2 -----------ER-KRSTPMNP---------------------------------- 188
human ELL3 -- 132
fly Ell SSSSNSFNSNNHSRKLGSSPFNGLGVGSSSSSSAFASRSPNPSTLGAIGTVNGSGVGGGR 290
human ELL -- 184
human ELL2 -- 188
human ELL3 -- 132
fly Ell YGGGAASSLASTFANGISQGYHNLSGSSPRDSMTAGTSSATASSVISSRNKMPSGGLTSS 350
human ELL ----ASAIR--KSGASAVSGGSGVSQRPFRDRVLHLLALRPYRKAELLLRLQKDGLTQAD 238
human ELL2 ----ANTIR--KTHS-----SSTISQRPYRDRVIHLLALKAYKKPELLARLQKDGVNQKD 237
human ELL3 ------DAR--H-- 136
fly Ell NSNSSSSSRSANSKSSGGNKMSDVSRRNIRERLIHLLALKAFKKPELFARLKNEGIRDRE 410
human ELL KDALDGLLQQVANMSAKDGTCTLQDCMYKDVQKDWPGYSEGDQQLLKRVLVRKLCQPQST 298
human ELL2 KNSLGAILQQVANLNSKDLSYTLKDYVFKELQRDWPGYSEIDRRSLESVLSRKLNPSQNA 297
human ELL3 ----------------------------PESWQNTGGYSEGDAVS---------- 153
fly Ell RNQITNILMDISTMSH--NTYNLRRQMWNDVDENWPFFSEQEVQQLKRRKPQNLTPPMSS 468
human ELL GSLLGDPAA-------SS--PP------GERGRSASPPQKRL----QPPDFIDPLANKKP 339
human ELL2 A---GTSRS-------ES--PV------CSSRDAVSSPQKRL----LDSEFIDPLMNKKA 335
human ELL3 ----------------------------------QPQMAL----EEVSVSDPLASNQG 173
fly Ell DAGSSTSGQSPTSTHTGSPPPPSSNGGPGGVGGGAGGTSMKRTSLEYDETMFSTVQPKKQ 528
human ELL RISHFTQ---RAQPAVNG--------KLGVPNGREAL-LPTP-------GPPASTDT--- 377
human ELL2 RISHLTN---RVPPTLNG--------HLNPTSEKSAAGLPLPPAAAAIPTPPPLPST--- 381
human ELL3 QSLPGSSREHMAQWEVRS--------QTHVPNREPVQALPSSAS---------------- 209
fly Ell RISHYKKD---TPPSGTSYSSAGVSMSLGSSGS-SSRSRYTPPQ--RQPGPLDDHSTTDL 582

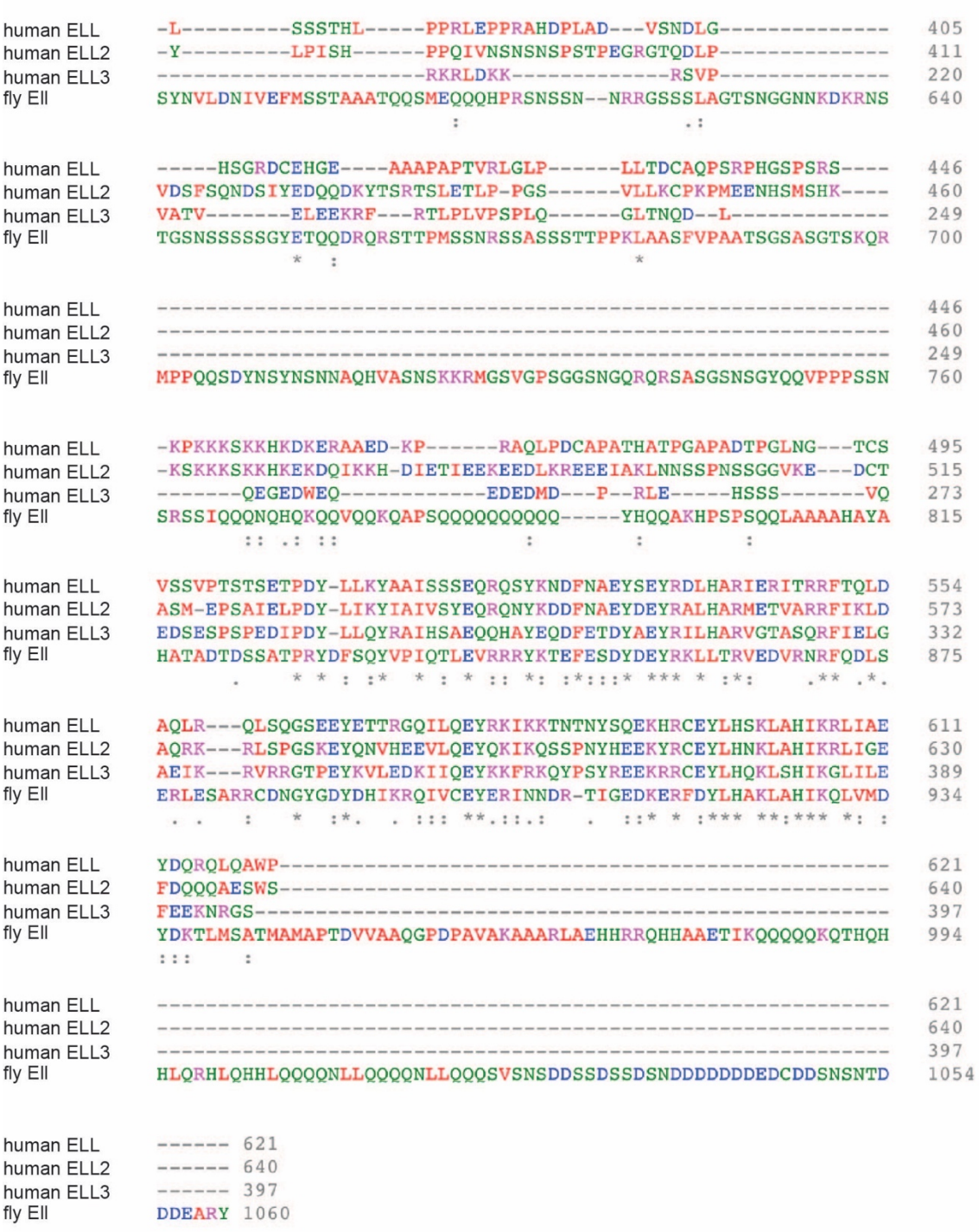


Figure 1-1. Alignment of human ELL, ELL2, ELL3 and *Drosophila* Ell
Comparison of protein sequences of ELL, ELL2, ELL3 and fly Ell by CLUSTAL O (1.2.4) multiple sequence comparison.

Table 1-1. Percentage identity matrix created by Clustal 12.1

	Human ELL	Human ELL2	Human ELL3	Fly Ell
Human ELL	100.00	49.43	32.90	30.38
Human ELL2	49.43	100.00	31.97	28.73
Human ELL3	32.90	31.97	100.00	20.73
Fly Ell	30.38	28.73	20.73	100.00

The ELL family proteins are highly conserved in mammal and nonmammal species [50]. In *Drosophila*, Ell is the orthologue with a conserved function in promoting Pol II elongation. Ell was cloned and shown to increase catalytic rate of Pol II elongation in vitro [51]. Although the neighbor-joining phylogenic tree indicates that fly Ell is equally divergent from human ELL family proteins (Figure 1-2), according to sequence alignment, fly Ell is more similar to human ELL (30.38% identity) and ELL2 (28.73% identity) than ELL3 (20.73% identity) (Figure 1-1 and Table. 1-1). Ell was also shown to regulate heat shock gene induction upon heat stress in vivo. By chromosome immunostaining in the fly salivary gland, Ell was found to co-localize with active elongating Pol II under heat stress treatment at heat shock loci, which have been used as an established model for investigating the regulation of paused Pol II in *Drosophila* [51,52].

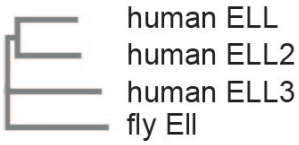


Figure 1-2. A phylogenic tree for human ELL family proteins and fly Ell
A neighbor-joining phylogenic tree shows that human ELL and ELL2 are less divergent.

The heat shock response is highly conserved for the protection of cells from a heat stress-induced proteotoxic environment [53,54]. Upon heat shock and other stresses, heat shock genes are rapidly and strongly induced. Heat shock proteins encoded by these genes are molecular chaperones that maintain proteostasis and promote cell survival. Research on heat shock genes in *Drosophila* is one of the early studies that indicates the existence of a transcriptional elongation control mechanism. In *Drosophila* cells, Pol

II was found to be transcriptionally engaged and form a nascent RNA chain of about 25 nucleotides at one of the major heat shock genes, HSP70, prior to heat shock treatment [55]. The paused Pol II can be released to synthesize the full-length *HSP70* transcription unit upon heat stress treatment [56]. In the fly, staining of the salivary gland polytene chromosomes is often used for study of heat shock response. In the salivary gland of the 3rd instar larvae, the chromosomes go through multiple rounds of replication without cell division, resulting in giant polytene chromosomes. This provides a unique opportunity to study the relationship between chromatin structure and transcription. Upon heat shock, stress genes are actively expressed and causes heat shock "puffs" on the polytene chromosomes, such that the major heat shock loci and encoded genes have been well characterized and identified [57-63]. Regulations of the transcription elongation of heat shock genes is conserved in mammals [64,65].

Ell is involved in SEC and LEC

Ell has been reported to be part of two elongation complexes: super elongation complex (SEC) and little elongation complex (LEC) [66], and these complexes are highly conserved from Drosophila to human (Table 1-2 and Figure 1-2). *ELL* was first identified and named as a gene that fuses to mixed-lineage leukemia (*MLL*) in acute myeloid leukemia [50]. *MLL* is located at chromosome 11 band q23, where aberrant translocation occurs frequently in hematologic malignancies. In these cases, the abnormal chromosome translocation leads to an in-frame fusion of the N terminus of MLL to the C-terminal part of various proteins with very little sequence similarity. Frequent fusion partners of MLL include ELL, AFF1, AF9, ENL and AF10 [50,67]. To better understand the relationship among these factors, the Shilatifard group generated cell lines expressing major MLL-

chimera complexes and purified them for mass spectrometry analysis. This approach led to the isolation of the SEC [68]. The composition of SEC was also investigated and reported using tagged normal proteins [69]. The isolation of SEC was also reported by the Zhou and the Kiernan groups through purification of transcription factors associated with Tat protein of human immunodeficiency virus type 1 (HIV-1) virus, which hijacks host transcription machinery for promoting transcription elongation for viral gene expression [70,71].

Table 1-2. Conserved genes involved in SEC and LEC from *Drosophila* to human

Super Elongation Complex (SEC)			
Drosophila melanogaster		*Homo sapiens*	
Gene	CG#	Gene	NCBI Gene ID
Su(Tpl)/Ell	CG32217	ELL/ELL1	8178
		ELL2	22936
		ELL3	80237
Eaf	CG11166	EAF1	85403
		EAF2	55840
lilli	CG8817	AFF1	4299
		AFF2	2334
		AFF3	3899
		AFF4	27125
ear	CG4913	MLLT1/ENL	4298
		MLLT3/AF9	4300
Cdk9	CG5179	CDK9	1025
CycT	CG6292	CCNT1	904
Little Elongation Complex (LEC)			
Drosophila melanogaster		*Homo sapiens*	
Gene	CG#	Gene	NCBI Gene ID
Su(Tpl)/Ell	CG32217	ELL/ELL1	8178
Eaf	CG11166	EAF1	85403
		EAF2	55840
Ice1	CG13550	KIAA0947/ICE1	23379
Ice2	CG10825	NARG2/ICE2	79664

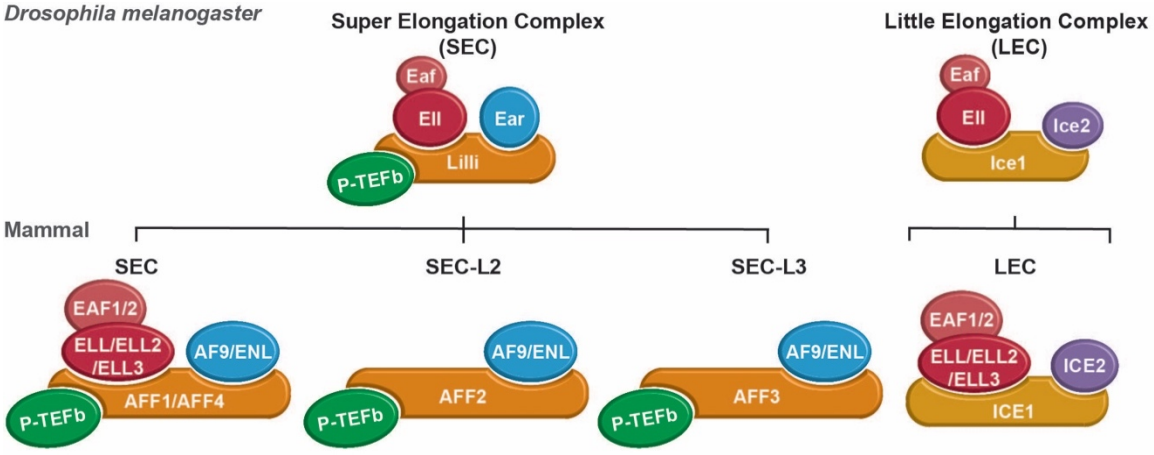


Figure 1-3. Schematics of SEC and LEC in fly and mammals
SEC and LEC are highly conserved from *Drosophila* to mammals.

The components of SEC include ELL family proteins ELL, ELL2 and ELL3, AF4/FMR2 family member 1 (AFF1; also known as AF4), AFF4, eleven-nineteen leukemia (ENL) and ALL1-fused gene from chromosome 9 (AF9), and P-TEFb [66]. AFF4 is the scaffold protein that directly interacts with P-TEFb, ENL/AF9 and ELL family proteins and is essential for SEC assembly [68]. The conservation of SEC in *Drosophila* was reported by a study that used a flag-tagged C terminus of Ell in fly S2 cells. Isolation by flag-affinity chromatography and analysis by mass spectrometry identified the fly SEC, which contains Ell (human ELL, ELL2 and ELL3-related), Lilli (human AFF1/4-related), Ear (human ENL/AF9-related) and P-TEFb [69,72] (Figure 1-3). P-TEFb is the kinase responsible for releasing paused Pol II. Containing P-TEFb, SEC has been shown to activate transcription through pausing regulation for genes involved in immediate responses, including developmental genes and stress-induced genes [68,73] (Figure 1-4).

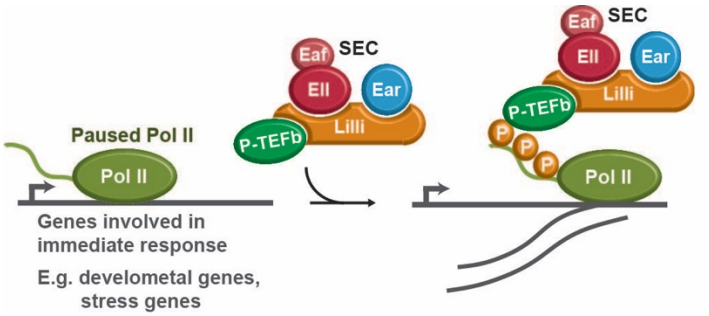


Figure 1-4. The function of SEC SEC activates transcription, mostly through pausing regulation, for genes involved in immediate responses

Identified targets of SEC include developmental genes. One study that identified the SEC complex through MLL-chimera also showed that in leukemia cell lines, SEC component AFF4 is responsible for the expression of developmental genes HOXA9 and HOXA10; these genes are often upregulated in cells bearing an MLL translocation [68]. This study suggests that in leukemia, the fusion of MLL and SEC components leads to premature release of paused Pol II and increased levels of developmental genes. The role of SEC in regulating transcription elongation on developmental genes was also demonstrated in mouse embryonic cells in response to retinoic acid (RA) and human HCT-116 cells in response to serum [73]. These results show that SEC is responsible for the rapid induction of a small subset of genes upon exposure to developmental signals. These genes have paused Pol II prior to stimulation, except one gene, *Cyp26a1*, whose induction requires SEC in the absence of paused Pol II, indicating there might be other mechanisms for rapid induction of transcription that involves SEC.

Another group of identified SEC targets are heat shock genes. Given that heat shock gene regulation is a well-established model for pausing regulation and that SEC is involved in the regulation of transcriptional pausing, the regulation of heat shock genes by SEC was tested. Chromosome immunostaining in *Drosophila* salivary gland showed

that the major component of SEC, Lilli, colocalizes with Ell and elongating Pol II on the chromosomes [68]. After heat shock treatment, Lilli co-localizes with Ell on the heat shock loci together with elongating Pol II, suggesting that SEC releases paused Pol II on heat shock genes [68]. Studies using chromatin immunoprecipitation and quantitative PCR (ChIP-qPCR) and qPCR confirmed that the scaffold protein of SEC, AFF4, is also bound to HSP70 and is required for its induction upon heat stress [68].

In mammals, besides the SEC complex, there are 2 related complexes: SEC-like 2 (SEC-L2) and SEC-like 3 (SEC-L3) ([66,72] and Figure 1-3). In mammals, except AFF1 and AFF4 which are involved in SEC, the AFF family also includes AFF2 (also known as FMR2) and AFF3 (also known as LAF4). These proteins share conserved N-terminal and C-terminal regions. SEC-L2 and SEC-L3, containing AFF2 and AFF3 respectively, were biochemically purified from human cells. Similar to SEC, SEC-like complexes also contain the kinase P-TEFb and ENL or AF9, but ELL family proteins were not identified as components involved. Genome wide studies showed that SEC, SEC-L2 and SEC-L3 regulate different sets of genes [72]. In addition, although all 3 complexes can phosphorylate Pol II CTD in vitro, only SEC is responsible for the induction of Hsp70 under heat shock in vivo [72]. These different complexes in mammals are predicted to expand the regulatory ability of cells in response to different cellular signals.

Ell is also identified to be involved in LEC (Figure 1-5). In the study identifying components that interact with *Drosophila* Ell, two proteins Ice1 and Ice2 were isolated as well as the components of SEC (Ice stands for “interacts with the C terminus of Ell”) [69]. Ice1 and Ice2 interact with Ell and Eaf, but not other components of SEC. This led to the identification of LEC, which is composed of Ell, Eaf, Ice1 and Ice2. Ell is implicated in

enhancing transcription elongation of target genes of LEC [74]. The major transcriptional targets of LEC are Pol II-transcribed small nuclear RNAs (snRNAs) (Figure 1-5). ChIP-seq studies using Ell and Ice antibodies identified snRNA genes as sites of LEC and Pol II enrichment [69]. RNA sequencing (RNA-seq) with Ell and Ice1 depletion also show that Pol II-transcribed snRNA genes are the most altered class of genes in LEC downregulated cells [69].

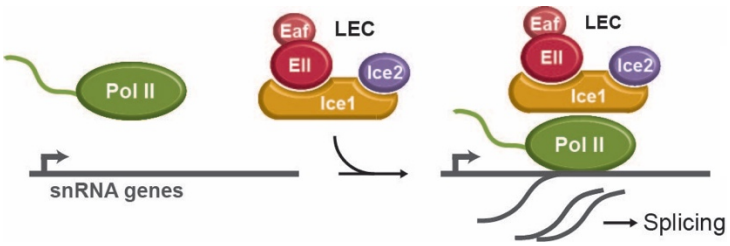


Figure 1-5. The function of LEC LEC is responsible for the transcription of snRNAs.

snRNAs are short non-coding RNAs that form small nuclear ribonucleoprotein particles (snRNPs). snRNPs further assemble with numerous proteins into the spliceosomes, functioning in pre-mRNA splicing [75]. Splicing is an essential step of gene expression, which can also increase the number of unique proteins from a single pre-mRNA species. Splicing is catalyzed by the spliceosomes of which there are two types: the U2-dependent spliceosome (also known as the major spliceosome), which is responsible for more than 99% of the splicing events, and the U12-dependent spliceosome (also known as the minor spliceosome), which is less abundant. The U2 spliceosome includes snRNPs: U1, U2, U4, U5 and U6. The U12 spliceosome is composed of U11, U12, U4atac, U5 and U6atac snRNPs [76,77] (Table 1-3). In addition to the snRNAs involved in the minor or major spliceosomes, there is another snRNA U7, which is involved in the maturation of replication-dependent histone mRNA [78]. U1, U2, U4, U4atac, U5, U7, U11

and U12 are transcribed by Pol II, whereas U6 and U6atac are generated by RNA polymerase I (Pol I).

Table 1-3. Major vs Minor spliceosome snRNAs

Spliceosome	U2-dependent/major	U12-dependent/minor
snRNAs	U1	U11
	U2	U12
	U5	U5
	U4	U4atac
	U6	U6atac

The composition and function of LEC is highly conserved from *Drosophila* to mammals. The orthologues of Ice1 and Ice2 in human are ICE1 and ICE2, respectively [69]. Based on biochemical purification, human LEC is composed of ELL family proteins, EAF1/EAF2, ICE1 and ICE2 [69]. In human HCT-116 cells, a protein ZC3H8 was also co-purified with LEC [74]. ICE1 is the scaffold protein which is critical for LEC assembly, proper localization of LEC in subnuclear bodies as well as downstream target expression [74]. ChIP-seq using HCT-116 cells shows that LEC subunits occupy Pol II-transcribed snRNA genes, and consistently RNA-seq study found that downregulation of ICE1 leads to a global reduction of snRNA expression [74]. Pol II ChIP-seq analysis, comparing cells with ICE1 depletion to control cells, further shows that depletion of ICE1 decreases Pol II occupancy on LEC-bound snRNA genes [74].

Implications of the roles of LEC and SEC in diseases

The disrupted function of SEC has been directly associated with human disorders including immunodeficiency, cancer and developmental syndromes [79-82]. SEC is also critical for HIV-1 replication [70,71]. In HIV-1 gene expression, elongation of paused Pol II near the transcription start site is a major rate-limiting step. The Tat protein encoded by HIV-1 is an RNA binding protein, which binds to Tat response element (TAR). Together,

they recruit SEC to the 5' end of all viral transcripts to produce full length products. Another disease associated with disrupted SEC activity is leukemia [68]. As discussed previously, many components of SEC have been found as translocation partners with MLL in disease. In leukemia cells carrying an MLL translocation, depletion of the essential component of SEC AFF4 results in loss of *HOXA9* expression, which is a key target of MLL fusion proteins. The involvement of SEC in MLL fusion proteins suggests that the regulation of transcription elongation may be disrupted and contribute to MLL-rearranged leukemias. Germline gain-of -function mutations in AFF4 were identified to be the cause of the developmental disorder CHOPS syndrome [83]. The name CHOPS is an abbreviation for features of the disease, including cognitive impairment, coarse facies, heart defects, obesity, pulmonary involvement, short stature and skeletal dysplasia. ChIP-seq analysis suggests that in the disease, transcriptional elongation is disturbed by altered binding of AFF4 and cohesin on chromosomes [83].

Although neither LEC or SEC has been implicated in neurodegenerative disorders, misregulations of targets of LEC and SEC have been reported in neurodegeneration. As introduced and discussed before, the targets of LEC are Pol II-transcribed snRNAs, which are involved in splicing. Global splicing alterations have been observed by genome-wide analysis in both FTD and ALS patients [84,85]. Furthermore, in ALS patient brain tissue and human cell disease models, the levels of several snRNAs are reported to be altered, with both higher and lower levels [86,87]. Onodera's group showed that the levels of U1, U2, U4atac and U6atac were significantly decreased in SH-SY5Y neuroblastoma cells treated with TDP-43 siRNA compared to control cells [86]. They also showed that U12 levels are decreased in ALS affected tissues, including spinal cord, motor cortex and thalamus, but not in unaffected tissues, including cerebellum, kidney

and muscle. However, Yamanaka's group reported that U4, U5 and U6 are upregulated in SH-SY5Y cells with TDP-43 downregulation compared to control cells, and that several snRNAs including U1, U2, U4, U5 U11 and U12 are upregulated in spinal cord tissue of ALS patients [87]. These studies indicate that in TDP-43-associated diseases snRNA levels and the spliceosome functions are disrupted, although there are discrepancies on the changes observed.

The minor spliceosome was discovered about a decade later than the discovery of the major spliceosome and has been studied less; however, more attention has been brought to it because of its involvement in diseases including neurodevelopmental and neurodegeneration disorders [88,89]. A point mutation in RNU12 encoding the U12 snRNA has recently been reported to be associated with early onset cerebellar ataxia (EOCA). EOCA is characterized by progressive cerebellar ataxia, brisk tendon reflexes and sensory loss. The mutation (84C>T) is located on one of the stem loop structures and is predicted to weaken the stem-closing base pairing and potentially affect protein interactions and spliceosome assembly [90]. Other known diseases associated with defects in minor spliceosome components include Microcephalic Osteodysplastic Primordial Dwarfism type I/Taybi-Linder Syndrome (MOPD1/TALS), Myelodysplastic Syndrome (MDS), Roifman syndrome (RFMN), Lowry Wood Syndrome (LWS) and Isolated Growth Hormone Deficiency (IGHD) [89,91-95]. Three of these diseases (MOPD1/TALS, RFMN and LWS), share clinical features of abnormal brain development related to cephalon-skeletal dysplasia, intrauterine and postnatal growth retardation and microcephaly.

The minor spliceosome is also associated with neurodegeneration. In addition to ALS, another neurodegenerative disease caused by death of motor neurons spinal muscular atrophy (SMA), has also been linked to dysfunction of the minor spliceosome. SMA results from mutations in the *SMN* gene, which leads to decreased levels of the SMN protein. SMN is essential for assembly of Sm-class snRNPs, including U1, U2, U4, U4atac, U5, U7, U11 and U12 snRNPs [96]. Studies using different SMA models have shown preferential decrease of minor snRNP levels [97-100], implying that the alteration of splicing regulated by the minor spliceosome may contribute to the SMA phenotype.

Heat shock genes are one group of identified and characterized targets of SEC. As neurodegenerative diseases are associated with the accumulation of misfolded proteins, the important role of protein folding machinery maintaining proteostasis has been indicated in neurodegeneration through different studies [101,102]. Among these, misregulation of the heat shock response has been reported to be associated with TDP-43 proteinopathy [103-105]. By using a human cell disease model of TDP-43 toxicity, Hsiang-Yu el al. show that Hsp90 inhibitor 17-AAG, which inhibits Hsp90 activity but activates HSF-1, Hsp27 and Hsp70, prevent TDP-43 aggregation caused by reactive oxygen species (ROS) [104]. In another study, Han-Jou el al. demonstrated in a TDP-43 cellular model that promoting the heat shock response by expression of the master regulator HSF1 leads to a decrease of insoluble TDP-43 aggregates and improves cell survival [105]. In recent work by Berson et al. using fruit fly and human cell models, the induction of several heat shock genes was found to be reduced in these disease models, and upregulation of Hsc4 and Hsp68 suppressed eye degeneration caused by TDP-43 toxicity [103]. These studies implicate a potential therapeutic role of the heat shock response in TDP-43-mediated neurodegeneration.

In *Drosophila*, among the heat shock genes, which typically encode proteins, one gene encodes a long non-coding RNA (lncRNA): heat shock response RNA omega (*Hsrω*) [106]. The *Hsrω* locus is composed of 2 exons, which are separated by an intron and followed by tandem repeats of 280 bp units. Nuclear and cytoplasmic lncRNAs named *Hsrω-n* and *Hsrω-c* are produced from this locus. *Hsrω* is induced dramatically by heat shock and is required for normal development and for survival after heat stress [106,107]. *Hsrω-n* together with various proteins forms specific nuclear structures named omega speckles [106], which are thought to serve as dynamic storage sites for RNA-processing and associated proteins. *Hsrω* is also involved in other pathways potentially functioning in maintenance of cellular homeostasis [106]. In human cells, the functional orthologue of *Hsrω* is the Satellite III repeat RNA (Sat III) [108]. Sat III is encoded by the Sat III repeat sequences mainly located in the pericentrimeric region of human Chromosome 9 [109]. The size of Sat III repeat RNA has been reported to be 2-5 kb, with the basic repeat unit of Sat III being a pentamer sequence GGAAT, which can be interrupted by the terminator sequence CAAC(C/A)CGAGT [110]. Like *Hsrω*, Sat III is induced by heat stress and forms a comparable nuclear structure called the nuclear stress body (nSB) at the site of synthesis [109]. The roles of these heat stress-induced non-coding RNAs in neurodegenerative diseases remain to be explored.

The new findings of the thesis adding to the field

The main findings of my thesis work include: 1) the identification of Ell and Ell-containing transcription elongation complexes LEC and SEC as modifiers of TDP-43 toxicity in the fly. 2) demonstration that the levels of several LEC target snRNAs are upregulated upon

TDP-43 expression including U12, which is functionally involved in TDP-43 toxicity. 3) that splicing of select U12-dependent spliceosome targets are elevated in TDP-43 fly disease model. 4) that TDP-43 co-localizes with SEC specific component Lilli on a chromosome locus encoding a stress-induced lncRNA *Hsrω*. 5) that *Hsrω* is identified as abnormally elevated and functionally important to TDP-43-mediated degeneration. 6) that the human orthologue of *Hsrω* ,Sat III, is increased in human cells with TDP-43 expression and in FTLD-TDP-43 patient tissue. 7) that TDP-43 physically interacts with one of the human orthologues of Ell, ELL2.

These findings suggest a model in which TDP-43 promotes the activities of LEC and SEC through interaction with ELL2, leading to the upregulation of specific targets of LEC and SEC. This upregulation results in misregulation of splicing and cell homeostasis and contributes to TDP-43-associated neurodegeneration (Figure 1-6). These findings reveal important roles of LEC and SEC in TDP-43 toxicity and link transcription elongation complexes to neurodegeneration for the first time, identify important downstream targets and provide potential therapeutic insights for TDP-43-mediated neurodegenerative diseases.

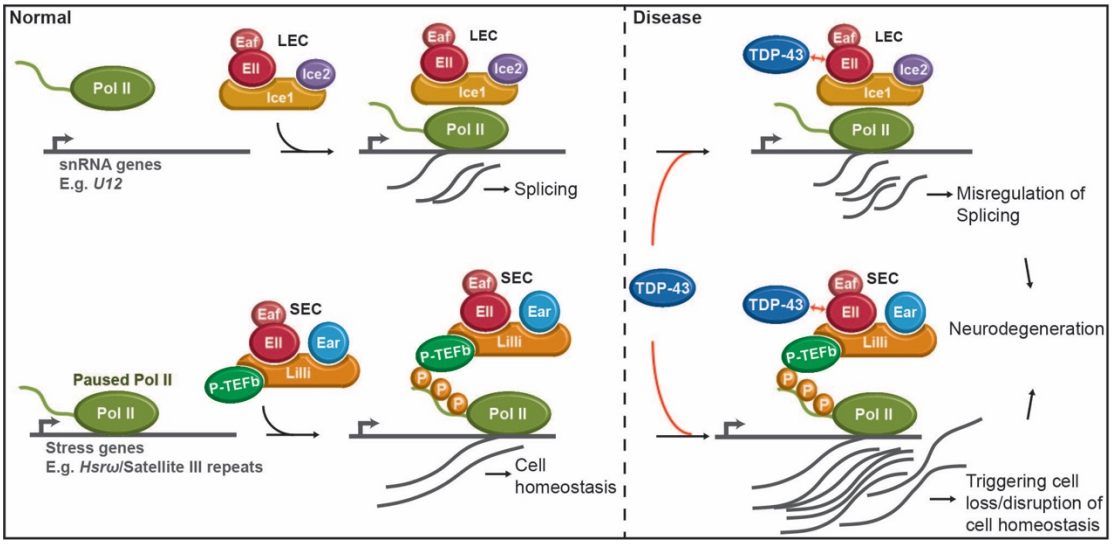


Figure 1-6. The model for the roles of LEC and SEC in TDP-43-mediated neurodegeneration

CHAPTER 2: MANUSCRIPT

ABERRANT ACTIVATION OF NON-CODING RNA TARGETS OF TRANSCRIPTIONAL ELONGATION COMPLEXES CONTRIBUTES TO TDP-43 TOXICITY

Chia-Yu Chung, Amit Berson, Jason R. Kennerdell, Ashley Sartoris, Travis Unger, Sílvia Porta, Hyung-Jun Kim, Edwin R. Smith, Ali Shilatifard, Vivianna Van Deerlin, Virginia M-Y Lee, Alice Chen-Plotkin, and Nancy M. Bonini

This chapter is a manuscript in press at *Nature Communications*.

The contributions of CYC are: 1) verification of the suppression effects of SEC components by detailed characterizations of the suppression effect and by various control experiments in the fly. 2) raising and testing the hypothesis that both SEC and LEC are involved in TDP-43 toxicity. 3) demonstration of the involvement of LEC in TDP-43-mediated eye degeneration 4) demonstration of Ell as a dose-dependent modifier by showing that upregulation of Ell enhances TDP-43-caused eye degeneration. 5) characterization of Ell as a modifier in the fly nervous system. 6) identification of important downstream targets of LEC and SEC, and demonstration of the functional involvement of these targets in fly disease models. 6) extension of the findings from fly disease models to a human cell disease model and FTLD-TDP patient tissues by developing qPCR primers and assessing the human repetitive RNA Sat III in RNA samples collected by collaborators. 7) determination of the RNA and protein levels of human ELL and ELL2 using samples from a human cell disease model and/or human patient tissues (RNA samples were collected by collaborators) 8) demonstration of the interaction of TDP-43 and human ELL2 using human HEK293 cells. 9) all figure panels except Fig 3c,d and Supplementary Data 1. 10) preparation of the manuscript under the supervision of Dr. Bonini.

Abstract

TDP-43 is the major disease protein associated with amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration with ubiquitinated inclusions (FTLD-TDP). Here we identify the transcriptional elongation factor Ell—a shared component of little elongation complex (LEC) and super elongation complex (SEC)—as a strong modifier of TDP-43-mediated neurodegeneration. Our data indicate select targets of LEC and SEC become up-regulated in the fly ALS/FTLD-TDP model. Among them, U12 snRNA and a stress-induced long non-coding RNA *Hsrω*, functionally contribute to TDP-43-mediated degeneration. We extend the findings of *Hsrω*, which we identify as a chromosomal target of TDP-43, to show that the human orthologue Sat III is elevated in a human cellular disease model and FTLD-TDP patient tissue. We further demonstrate an interaction between TDP-43 and human ELL2 by co-immunoprecipitation from human cells. These findings reveal important roles of Ell-complexes LEC and SEC in TDP-43-associated toxicity, providing potential therapeutic insight for TDP-43-associated neurodegeneration.

Introduction

Amyotrophic lateral sclerosis (ALS; OMIM no. 105400) is the most common motor neuron disease, resulting from loss of motor neurons in the motor cortex, brainstem and spinal cord, whereas frontotemporal dementia (FTD; OMIM no. 600274) is characterized by progressive changes in behavior, personality, and/or language due to gradual deterioration of the frontal and temporal lobes. Despite differences in primary sites of neurodegeneration, these two diseases share neuropathological and genetic commonalities as well as clinical overlap [5,6]. TAR DNA binding protein 43 (TDP-43) is the major component of inclusion bodies in most ALS and in half of FTD known as FTLD-TDP [9,10]. Both depletion and upregulation of TDP-43 cause neuronal loss, indicating that TDP-43 levels are critical in the brain [13,16]. TDP-43 is an RNA-binding protein with many functions in RNA regulation and metabolism and is one of a number of RNA binding proteins associated with ALS and FTD [13,111,112]. This leads to an RNA dysregulation-centered view of TDP-43 disease mechanisms. TDP-43 however also binds DNA and has been shown to regulate transcription[11,113,114]. The role of transcriptional dysfunction in TDP-43 proteinopathies remains largely unexplored.

Drosophila melanogaster has played an important role in elucidating roles of many genes in human neurological disease [27,115,116]. In *Drosophila,* expression of human TDP-43 leads to neuronal degeneration, shorten lifespan and climbing defects, recapitulating fundamental disease features and serving as a platform to provide insight into underlying pathways and therapeutic targets [37,38]. To define novel disease mechanisms associated with TDP-43, we used *Drosophila* to screen for modifiers of toxicity. Our unbiased genetic screen uncovered the transcription elongation factor Ell (also known as Su(Tpl)) as a novel and robust modulator. Ell is present in two complexes: little elongation

complex (LEC) and super elongation complex (SEC) [68,69,72]. In *Drosophila*, LEC, containing Ell, Eaf, Ice1 and Ice2, regulates the initiation and elongation of Pol II-transcribed small nuclear RNA (snRNA) genes [69,74]. SEC is composed of Ell, Eaf, Ear, Lilli and P-TEFb [66,69]. The kinase P-TEFb phosphorylates Pol II, leading to the release of paused Pol II into productive elongation. The composition and functions of both complexes are highly conserved in mammals [68,72,74]. Neither complex has been implicated in neurodegenerative disorders.

snRNAs together with a range of proteins form small nuclear ribonucleoproteins (snRNPs), which are essential components of the splicing machinery of the spliceosome. In both ALS and FTD patients, splicing changes have been observed by transcriptome and microarray analyses, implicating aberrant splicing as an important disease mechanism [84,117]. In ALS and human cell disease models with TDP-43 depletion, disruption of several snRNA levels has been reported, both higher [87] and lower levels [86], but the detailed mechanism and whether these alterations are functionally important in disease is unclear. In normal physiology, SEC activity has been shown to be specific to select conditions and genes: developmental genes under differentiation signals [73], and heat shock genes upon stress [68,72]. The critical role of heat shock genes in neurodegenerative diseases has been revealed by studies on molecular chaperone and other proteins, which maintain proteostasis [101,102]. Among the heat shock genes, one distinct from others encodes no protein product but rather a long non-coding RNA (lncRNA), named Heat shock RNA omega (*Hsrω*). *Hsrω* is functionally required for normal development and the heat shock response in *Drosophila*, and is also involved in multiple other pathways potentially contributing to cellular homeostasis [106,107]. In humans,

Satellite III repeat RNA (Sat III) is the functional orthologue of *Hsrω* [108] and its role in neurodegenerative disease is unknown.

Here, we implicate Ell-associated transcriptional elongation complexes LEC and SEC as misregulated upon TDP-43 toxicity. We identify key non-coding RNA targets, U12 and *Hsrω*, as abnormally activated and functionally contributing to TDP-43-induced degeneration. Extension of these data to disease tissue and demonstration of the interaction between TDP-43 and one of the human orthologue of Ell, ELL2, implicates misregulation of human Ell orthologues as a contributor to TDP-43-associated pathologies.

Results

TDP-43 toxicity is mitigated by modulation of *Ell*

Flies expressing human TDP-43 in the eye show retinal degeneration [37,38]. We employed these animals to screen for genes that enhanced or suppressed the degeneration of TDP-43. In the screen, 2933 fly lines with different genetic modifications were tested. Among them, 12 fly lines enhanced TDP-43 toxicity and 23 fly lines showed suppression, indicating a limited number of genes can modify TDP-43 toxicity. From this screen, we found that knockdown of the gene *Ell* strongly mitigated TDP-43-induced deterioration of the external eye and internal retina (Fig. 1a, b and Supplementary Fig. 1a). Furthermore, up-regulation of *Ell* enhanced TDP-43 toxicity with more severe external eye and internal retinal degeneration (Fig. 1a, b and Supplementary Fig. 1a). These data indicated that *Ell* is a potent dose-dependent modifier of TDP-43, with *Ell* knockdown or up-regulation on its own having no effect on eye morphology

(Supplementary Fig. 1b). A RNAi control line against Luciferase did not mitigate TDP-43-mediated eye degeneration (Fig. 1a, b).

Ell protein is a shared component of two transcriptional elongation complexes: LEC and SEC (Fig. 1c). In order to investigate whether the suppression due to *Ell* knockdown was through one or the other of these complexes, we downregulated additional components of LEC and SEC in the presence of TDP-43. Reduction of SEC components *ear* and *lilli* partially suppressed the external and internal retinal deterioration conferred by TDP-43, as did reduction of LEC component *Ice1* (Fig. 1d, e and Supplementary Fig. 1a). The rescue effect of any of these components was not as strong upon depletion of the shared component *Ell*, suggesting that both LEC and SEC contribute to TDP-43 toxicity. Depletion of any of the components on their own had no effect on eye integrity (Supplementary Fig. 1b).

The levels of the TDP-43 protein were assessed by western immunoblot. These data indicated that enhancement by *Ell* up-regulation or suppression by lowered levels of *Ell* or other components of LEC or SEC was not through modulating the levels of TDP-43 protein (Supplementary Fig. 1c). We further confirmed that the *GAL4-UAS* expression system was not impacted by any of these components by examining the levels of a control protein β-galactosidase (β-gal) (Supplementary Fig. 1d). Additional fly lines, including a genetic mutation of *Ell*, and two RNAi lines of *Ice1* with proper controls were tested and showed consistent suppression effect on TDP-43-caused eye degeneration (Supplementary Fig. 1e).

To extend these studies to the nervous system generally, we expressed TDP-43 in all neurons in the adult animal using a conditional drug-inducible driver line. With expression in the adult fly brain induced by RU486, TDP-43 animals consistently show an age-associated decline in climbing ability, indicative of neural dysfunction (Fig. 1f, Supplementary Fig. 1f and [38]). Knockdown of *Ell*, although showing no effect on its own, restored climbing ability to normal without affecting TDP-43 protein levels (Fig. 1f, g and Supplementary Fig. 1g). We also assessed the suppression effect of *Ell* on TDP-43 toxicity in the nervous system by lifespan assays using the drug-inducible neuronal driver. Knockdown of *Ell* on its own caused mild but statistically significant extension of lifespan, and downregulation of *Ell* caused a mild suppression of TDP-43-shorten lifespan and a shift of the early stage of the lifespan curve (Supplementary Fig. 1h). Lifespan assays assessing *Ell* knockdown effect on TDP-43 toxicity were also tested by using a ubiquitous drug-inducible driver, showing consistent results that *Ell* downregulation has some effect to mitigate TDP-43 toxicity (Supplementary Fig. 1I). These data indicate that increased activity of both Ell-complexes LEC and SEC may contribute to TDP-43-mediated degeneration in the nervous system.

LEC snRNA target U12 contributes to TDP-43 toxicity

To determine whether the activity and function of LEC and SEC are promoted by TDP-43, we assessed the downstream targets in the fly disease model. Targets of LEC are Pol II-transcribed snRNAs [69,74]. We therefore used Northern blot analysis to assess the levels of snRNAs U1, U2, U4, U4atac, U5, U7, U11 and U12 in fly heads, with or without added TDP-43 driven by a ubiquitous drug-inducible driver. TDP-43 expression significantly increased levels of U1, U4, U7 and U12 (Fig. 2a and Supplementary Fig. 2b). Importantly, knockdown of *Ell* restored the levels of the elevated snRNAs back to

normal without affecting TDP-43 expression (Fig. 2a and Supplementary Fig. 2a, b). These data suggest that misregulation of snRNA components may be a consequence of disrupted *Ell* function in the animals and contribute to TDP-43 toxicity.

To test the functional role of the snRNAs in TDP-43-induced degeneration, we determined whether downregulation of the elevated snRNAs could mitigate toxicity. We examined fly lines predicted to downregulate various snRNAs (U1, U4, U7, U12). There was little effect of fly lines directed to U1 and U4 (there are multiple copies of these snRNAs in the genome making interference challenging), and U7 depletion suppressed the external eye but did not suppress the internal deterioration. Although we cannot exclude the potential importance of U1, U4 and U7, we focused on U12. *U12* knockdown partially, but consistently mitigated TDP-43-associated retinal disruption (Fig. 2b, c and Supplementary Fig. 2c), indicating that misregulation of U12 levels is functionally important to TDP-43 toxicity. We confirmed that U12 downregulation had no effect on TDP-43 protein, and that reduction of U12 on its own did not affect the eye (Supplementary Fig. 2d, e). U12 is a component of the minor spliceosome (U12-type spliceosome), indicating that targets of the minor spliceosome may be misregulated by TDP-43.

There are 18 genes containing a U12-type intron in *Drosophila* [100]. Total RNA from fly heads was prepared for reverse transcription quantitative PCR (RT-qPCR) with qPCR primer sets spanning the U12-regulated introns to assess the levels of spliced gene products (Fig. 2d). Among the 18 genes, six (*CG16941, CG11839, CG33108, CG11328, CG15735* and *CG3294*) were up-regulated with expression of TDP-43, with the other 12 genes not affected (Fig. 2e; Supplementary Fig. 2f, g), indicating increased U12 levels

leads to elevation of specific minor spliceosome targets. Knockdown of *Ell* corrected the levels back to normal (Fig. 2e; Supplementary Fig. 2g), consistent with the ability of *Ell* downregulation to normalize the levels of U12 and protect from TDP-43 toxicity. To further determine whether the elevation resulted from transcription or splicing, we assessed the levels of the unspliced transcripts. The unspliced products of 3 genes, *CG33108*, *CG11328* and *CG3294* were also increased upon TDP-43 expression, indicating a change in total transcript levels (Supplementary Fig. 2g). The unspliced RNA levels of the 3 genes *CG15735*, *CG16941* and *CG11839* were not altered significantly by the presence of TDP-43, indicating that the splicing of these U12-regulated introns was increased (Fig. 2e). These data indicate that the U12-type spliceosome is abnormally activated in the fly disease model to cause disrupted regulation of selected downstream targets.

TDP-43 and Lilli co-localize at the *Hsrω* lncRNA locus

Ell is also a component of SEC and SEC activates transcriptional elongation, which is critical for genes involved in developmental signaling pathways and the stress response [68,72,73]. The SEC scaffold protein Lilli binds to and regulates specific targets on the chromosomes in *Drosophila* and human cells [68,72,73]. Recently, genome-wide chromatin immunoprecipitation (ChIP-seq) shows that *Drosophila* TDP-43, TBPH, associates with chromatin [114]. We therefore considered that TDP-43 may co-localize to the same genes as Lilli. To explore this, we expressed TDP-43 tagged with YFP in *Drosophila* salivary gland and assessed Lilli and TDP-43 localization on the polytene chromosomes by immunostaining. Under ambient temperature, we observed that TDP-43 binds to polytene chromosomes with specificity (Fig. 3a). TDP-43 and Lilli each localized to ~115 sites (Fig. 3a). Most of these were euchromatin regions where active transcription

usually occurs. Detailed analysis of the co-localization pattern indicated that TDP-43 and Lilli consistently overlapped at 16 sites across chromosomes X, 2R, 3R and 3L (Fig. 3a and Supplementary Table 1). Although it is difficult to precisely identify the genes with co-localization of TDP-43 and Lilli due to the resolution of polytene bands, these data indicate partial overlap of TDP-43 and SEC targets in normal conditions; these genes may be important in toxicity.

Upon heat stress, SEC is recruited to major heat shock loci on the polytene chromosomes and, further, has been shown to regulate the expression of a major heat shock protein Hsp70 [68]. Dysfunction of the heat shock response, which functions to maintain proteostasis, is associated with TDP-43 toxicity [103-105]. We therefore probed whether TDP-43 co-localized to these targets of SEC. As noted [68], with heat shock Lilli localizes to the major molecular chaperones at polytene chromosome sites 63B, 67B, 95D, 87A and 87C and a unique heat shock locus, 93D that encodes the lncRNA *Hsrω* (Fig. 3b). When we examined co-localization with TDP-43, we noted some co-localization at 63B (encoding *Hsp83*) and 67B (encoding small heat shock genes) (Supplementary Table 2). Moreover, we observed robust and consistent co-localization at 93D in all polytene chromosome spreads examined (Fig. 3b and Supplementary Table 2); this was a site at which TDP-43 and Lilli also co-localized without added stress (Supplementary Table 1), suggesting that this stress-induced lncRNA may be a common target of TDP-43 and Lilli. We also assessed the co-localization of TDP-43 and Ell. Consistent with the results of immunostaining with Lilli, Ell also co-localized with TDP-43 at the 93D locus encoding *Hsrω* (Supplementary Fig. 3a). Precision nuclear run-on sequencing (PRO-seq) has shown that pol II is paused on *Hsrω* [118], and chromatin

staining shows that SEC components, Ell and Lilli, co-localize with elongating Pol II at the 93D locus after heat stress [68], both indicating that *Hsrω* is a target of SEC.

Given the limited resolution of polytene chromosome staining, we analyzed the published ChIP-seq data for TBPH [114] and Lilli [69] to better define genes bound by both factors. This analysis showed that TBPH binds to 383 genes and Lilli binds to 4162 genes (Fig. 3c and Supplementary Data 1). Among them, 328 genes were bound by both factors, which comprises ~86% of the genes bound by TBPH and ~8% of the genes bound by Lilli (Fig. 3c). Consistent with the polytene data, *Hsrω* is a target of both TBPH and Lilli (Fig. 3d). We therefore considered that *Hsrω* might be a target of SEC that becomes misregulated by TDP-43.

Hsrω functionally modulates TDP-43 toxicity

The 93D locus encodes nuclear and cytoplasmic lncRNAs referred to as *Hsrω-n* and *Hsrω-c*, respectively, both of which are induced by stress. Upon heat stress, *Hsrω-n* transcripts are up-regulated at the 93D site [106], with a balanced level of *Hsrω* being critical for the survival and recovery of flies following heat stress [107]. Given the striking co-localization of TDP-43 and Lilli at the *Hsrω* locus, we considered that dysregulation of *Hsrω* may occur with TDP-43. RNA from fly heads in the presence or absence of added TDP-43 expression was extracted for RT-qPCR and the levels of total *Hsrω* (*Hsrω-c* and *Hsrω-n*), and *Hsrω-n* were determined. Total *Hsrω* and *Hsrω-n* were increased ~2-fold upon TDP-43 expression (Fig. 4a). Importantly, this misregulation was reduced toward normal levels by *Ell* knockdown (Fig. 4a). These data indicate that *Hsrω* may be a functional target of SEC activity that contributes to TDP-43-associated degeneration.

To assess whether misregulation of *Hsrω* is functionally important to TDP-43 toxicity, we reduced the levels of *Hsrω* in animals expressing TDP-43. Reduction of *Hsrω* by placing animals expressing TDP-43 *in trans* to an *Hsrω* null mutation partially mitigated TDP-43-associated external and internal retinal deterioration (Fig. 4b, c). This was not associated with changed levels of TDP-43 protein, and loss of one copy of *Hsrω* on its own had no effect (Supplementary Fig. 4a, b). We also confirm the suppression effect of *Hsrω* downregulation on TDP-43-mediated eye degeneration by a RNAi line with proper controls (Supplementary Fig. 4c). The effect of lacking one copy of *Hsrω* on climbing defect caused by TDP-43 was unable to be determined because reduction of *Hsrω* on its own caused a decrease in climbing ability (data not shown). Taken together, these findings on the co-localization of TDP-43 and Lilli at the *Hsrω* gene locus, on elevated levels of *Hsrω* RNAs in TDP-43-expressing animals, and on functional mitigation of TDP-43-induced degeneration by modulation of *Hsrω* levels suggest that the stress-induced lncRNA *Hsrω* is a functional target of TDP-43 that contributes to TDP-43-mediated degeneration.

Sat III is upregulated in a cellular model and in patient tissue

These data on *Hsrω* raised the possibility that the functional orthologue of *Hsrω* in humans, the stress-induced Satellite III repeat RNA (Sat III) [108], may become misregulated in disease and contribute to toxicity. Although not sharing precise sequence similarity, Hsrω and Sat III transcripts share functional features and regulatory mechanisms: they are repeats and non-coding in nature; they are induced and accumulate at the site of synthesis upon stress, and associate with heterogeneous nuclear ribonucleoproteins (hnRNPs) and other RNA-processing factors; they are both Pol II-dependent transcripts [106,108,109]. Moreover, although Sat III is essential for cell

survival after heat shock, upregulation of Sat III promotes cell death and acute senescence in various cell models [119,120]. We therefore investigated whether Sat III transcripts became misregulated by TDP-43.

To assess Sat III levels in mammalian cells, we used human embryonic kidney 293 (HEK293) cells expressing doxycycline-induced GFP-tagged TDP-43 (GFP-TDP-43) or GFP alone. RNA was prepared following 6d of induction, and the levels of Sat III transcripts were examined by RT-qPCR. We compared Sat III levels in cells expressing GFP-TDP-43 to control cells expressing GFP. TDP-43 expression led to a ~4-fold increase in the steady state level of Sat III (Fig. 5a). The Sat III primers were validated by heat shock induction of Sat III followed by RT-qPCR analysis (Supplementary Fig. 5a). Thus, the human counterpart of *Hsrω* became misregulated upon aberrant TDP-43 expression in cells.

These findings raised the possibility that Sat III may become dysregulated in human disease. To address this, we used RNA prepared from frontal cortex of 13 FTD patients (six sporadic or with unknown family history and seven patients with a family history of FTD and/or a disease-associated mutation) and six clinically normal controls (Fig. 5b). All of the FTD patients had TDP-43 pathology (FTLD-TDP) [121,122]. In accordance with our studies in the fly and human cells, the results showed that the levels of Sat III transcripts were significantly higher in the frontal cortex of FTD patients compared to controls (Fig. 5b). Case numbers and details are indicated in Table 1. Taken together with the functional data from the fly, misregulation of the stress-induced lncRNA Sat III may be a contributor to TDP-43-associated disease.

TDP-43 and ELL2 proteins interact

To further dissect the relationship between Ell and TDP-43 and define additional mechanisms by which Ell may contribute to TDP-43-mediated neurodegeneration, we assessed whether Ell levels were upregulated in disease models; upregulation of Ell could lead to increase levels of downstream targets and in this manner contribute to TDP-43-mediated neurodegeneration. RNA levels of fly *Ell* were examined by RT-qPCR using RNA extracted from fly heads in the presence or absence of added TDP-43. These data showed that *Ell* is elevated ~50% upon TDP-43 expression (Fig. 6a). We next examined human ELL proteins in the HEK293 cell model. In mammals, there are three proteins in the ELL family: ELL, ELL2 and ELL3 [66], with ELL3 being the most distinct based on sequence (~50% identity) and enriched in testis [47]. All three ELL proteins can be pulled down with factors in LEC and SEC [68,69]. We examined the RNA levels of *ELL* and *ELL2* in cells expressing GFP-TDP-43 or GFP control by RT-qPCR. The results are consistent with the fly data, showing that TDP-43 expression led to a 50% increase of *ELL* and *ELL2* levels (Fig. 6b). However, when we assessed protein levels by western immunoblot, the results indicate that the levels of ELL and ELL2 were not changed (Supplementary Fig. 6a). We also examined *ELL* and *ELL2* by RT-qPCR using RNA prepared from frontal cortex of the 13 FTD patients with TDP-43 pathology; expression levels of *ELL* and *ELL2* were not changed significantly compared to controls (Supplementary Fig. 6a). These results, together with the data showing that upregulation of *Ell* on its own does not cause eye degeneration (Fig. 1a and Supplementary Fig. 1b), indicate that mechanisms beyond simply a global increase in levels of *Ell* may be involved in TDP-43-associated toxicity.

The polytene chromosome immunostaining showed that TDP-43 co-localized with Ell and Lilli at the 93D locus (Fig. 3a and Supplementary Fig. 3a), indicating that interactions between the proteins might occur. We thus assessed whether endogenous TDP-43 interacts with ELL or ELL2 in nuclear extracts of HEK293 cells using co-immunoprecipitation (co-IP). TDP-43 was immunoprecipitated by an anti-TDP-43 antibody (Fig. 6c), and western blots then probed for presence of co-immunoprecipitating ELL or ELL2. ELL was not detected; however, ELL2 was co-immunoprecipitated together with TDP-43 (Fig. 6c). This was specific because ELL2 was not co-immunoprecipitated with IgG control. These data indicate that TDP-43 interacts with ELL2, suggesting that this interaction may contribute to dysfunction of ELL-associated complexes upon aberrant TDP-43 function.

Discussion

Misregulation of several transcriptional elongation factors has been reported in diseases, including viral pathogenesis and cancer [79,123], yet these factors have not been implicated in neurodegenerative disorders. Here we report that Ell and Ell-containing transcriptional elongation complexes LEC and SEC are novel modifiers of TDP-43 toxicity. The levels of several LEC target snRNAs including U12 are upregulated by TDP-43 proteinopathy. Our data suggest that the U12-type spliceosome is abnormally activated and contributes to toxicity. Through chromatin immunostaining, we identify a key target of SEC, a stress-induced lncRNA *Hsrω*, as aberrantly elevated and functionally important to TDP-43-mediated degeneration. In addition, we show that the levels of the human orthologue of *Hsrω*, Sat III, is increased in a human cell model and in FTLD-TDP frontal cortex tissue, indicating that Sat III dysfunction may contribute to TDP-43-associated disease. Finally,

we demonstrate TDP-43 interacts with the one of the human orthologues of Ell, ELL2, indicating that the aberrant elevation of LEC and SEC activity in disease may be promoted through the interactions between TDP-43 and ELL2 (Fig. 7). These findings highlight the critical roles of LEC and SEC in TDP-43-mediated pathologies, and highlight that approaches to normalize the activity of human orthologues of the shared component Ell may be of therapeutic benefit.

Our data suggest a model whereby SEC and LEC contribute to TDP-43-mediated degeneration in parallel (Fig. 7). We find that that depletion of the shared component *Ell* led to nearly full suppression of TDP-43-mediated eye degeneration, whereas, by contrast, downregulation of LEC- and SEC-specific components or downstream targets U12 and *Hsrω* showed only partial suppression, despite the downregulation effect for the genes being robust. In our study, we identified important targets that are functionally involved in TDP-43-mediated degeneration, but we do not exclude the possibility that there may be other targets regulated by LEC and SEC that are important in the disease. Furthermore, we show that up-regulation of the shared component *Ell* enhances TDP-43 degeneration, and the abnormally elevated levels of LEC and SEC targets caused by TDP-43 toxicity can be rescued by *Ell* depletion. Our study demonstrates that TDP-43-mediated degeneration can be alleviated by decreasing the activities of LEC and SEC, which is predicted to lead to reduction of elongating Pol II on downstream key targets and thus lower their levels.

As an RNA binding protein, TDP-43 has been demonstrated to regulate targets through direct binding [13,111,124]. Our studies propose another layer of regulation: we hypothesize that TDP-43 affects transcription through misregulation of LEC and SEC activities,

contributing to degeneration. Furthermore, the data showing that TDP-43 interacts with ELL2 (Fig. 6c) suggests a mechanism which TDP-43 may promote the activities of LEC and SEC at target genes, through interactions of the TDP-43 with ELL2 (Fig. 7).

Our data highlight elements of specificity in TDP-43-associated toxicity. First, we show that only selected targets of LEC and SEC are elevated in expression in the fly upon TDP-43 expression. The selectivity of targets can be defined by accessory factors associated with the LEC and SEC: in human cells, a mediator subunit MED26 has been shown to interact through EAF1 and EAF2 and promote the expression of select target genes, including a subset of snRNAs and Hsp70 [125,126]. Other factors like PAF1 and integrator, interacting with SEC [127,128], might also contribute to target specificity.

Our data showing that TDP-43 binds to *Drosophila* polytene chromosomes with selectivity (Fig. 3a, b), and that TDP-43 interacts with ELL2 (Fig. 6c), suggest specificity could also be defined by TDP-43. In HIV infection and leukemia, the sequence specific DNA binding activators, HIV-1 Tat protein and MLL respectively, direct SEC to specific target genes for their abnormal expression [68,70,71]. Thus, it is possible that TDP-43 serves as a sequence specific DNA or RNA binding factor and therefore defines target specificity in associated diseases. A second layer of specificity is that, upon increased U12 snRNA levels, among 18 target genes with a U12-type intron, only 6 become misregulated with TDP-43. The detailed mechanisms by which specificity is established is an intriguing question that remains to be addressed.

In models of the motor neuron disease spinal muscular atrophy (SMA), expression levels and splicing of several genes containing a U12-type intron are decreased [100]. Our data

show that the activity of the minor spliceosome is affected in an opposite direction in TDP-43 expressing animals, highlighting the importance of a balanced level of the minor spliceosome in maintaining normal motor neuron function. Products of 6 minor spliceosome-regulated genes were increased with TDP-43, with 3 genes potentially regulated through transcription and 3 genes likely through misregulation of splicing (see Fig. 2e and Supplementary Fig. 2f, g); among these, expression of 3 genes (*CG33108*, *CG16941* and *CG11839*) are decreased in a fly model of SMA [100]. This finding indicates misregulation of common target genes in different motor neuron disease models. Misregulation of orthologues of some of these genes have been reported in mammalian models of ALS and FTD or in human patients: the splicing of a U12-type intron in *C19orf54*, the human orthologue of *CG33108*, is affected in an ALS transgenic mouse model expressing human FUS, which is encoded by a gene that, when mutated, can cause ALS and rare cases of FTD [129]. Another gene, *SF3A1*, the human orthologue of *CG16941*, shows altered poly-A usage in ALS patients bearing the *C9orf72* GGGGCC hexanucleotide repeat expansion [84], the most common known genetic cause of ALS and FTD [84,130,131]. *SF3A1* is also implicated as a novel risk factor in FTD by gene co-expression network analysis [132]. Our findings, together with others, indicate that these genes may be critical in TDP-43-associated disorders.

We report that the stress-induced repetitive RNA Sat III is increased in TDP-43-associated disease, which suggest that the elevation may contribute to neuronal loss and degeneration. We used both HEK293 cell disease model and human patient samples to verify findings in different systems. Stress genes maintain proteostasis and promote cell survival. However, the stress response can be diverse depending on the circumstances [133,134]. The induction of Sat III by TDP-43 is significant although mild (~4-

fold; see Fig.5), leading us to consider that expression of TDP-43 may be akin to a chronic stress that triggers distinct pathways from those induced by an acute heat shock response (whereby Sat III is induced many thousand-fold (see Supplementary Fig. 4)). In support of this idea, although Sat III is required for cell survival after heat shock, forced expression of Sat III triggers cell death and rapid cellular senescence [119,120]. Elevated Sat III transcripts have been noted in senescent cells and fibroblasts from patients with the premature aging disease Hutchinson-Gilford progeria syndrome (HPGS) [135,136]. Furthermore, normalization of elevated Sat III largely rescues the mitotic dysregulation and senescence phenotype of SIRT6-depleted cells [120], indicating that elevated Sat III contributes to age-related cellular abnormalities. These studies support our findings and model that elevated Sat III is detrimental and may contribute to TDP-43-associated neuronal dysfunction. The influence of ELL proteins on the induced Sat III levels by TDP-43 is an interesting question that remains to be further investigated.

Given our data in *Drosophila*, the shared component *Ell*, whose depletion rescued the elevated levels of LEC and SEC targets, may be an effective therapeutic target. Importantly, our data also showed that knockdown of *Ell* on its own did not cause deleterious effects like degeneration or compromised mobility, further supporting that *Ell* may be a promising therapeutic candidate (see Supplementary Fig. 1b, f). Domain analysis of human orthologues of Ell, ELL and ELL2, identified the N-terminal 150 amino acids as critical for elongation function [48,137]. Given that the ELL family is also reported to display additional activities, such as serving as an E3 ubiquitin ligase for c-Myc degradation [138] and inhibiting P53 function [138], a strategy to apply molecules targeting domains specific for elongation activity, or even targeting the interaction between ELL proteins and the assembly components of LEC and SEC, ICE1 and AFF4 (human

orthologue of Lilli), respectively, may be promising. Recent research resolving the structure of AFF4 and ELL2 binding interface reveal a cavity that is a potential binding site for small molecules to interrupt SEC activity [139].

The therapeutic potential of ELL family proteins may not be limited to TDP-43-associated toxicity. Our findings indicate that knockdown of *Ell* protects against not only TDP-43 but also GGGGCC hexanucleotide repeat toxicity (Supplementary Fig. 9), which is also a major disease locus for ALS and FTD. These data suggest that shared targets of these two toxic insults and of ELL family proteins might be central to degeneration associated with these mechanisms.

Methods

Drosophila melanogaster

Flies were raised at 25°C. Transgenic flies with TDP-43-YFP and recombinant fly lines gmr>TDP-43 and elavGS>TDP-43 are described [37,38]. The recombinant fly line daGS>TDP-43 was generated with the fly line *daughterless-GeneSwitch* (*dsGS*), a generous gift from Dr. Veronique Monnier [140]. *Ice1* RNAi fly line was generated with two constructs (SH09112.N and SH09113.N) provided by DRSC/TRiP center. Fly lines used are listed in Supplementary Table 3.

Drosophila external eye and internal eye imaging

Adult female flies (1-2d) were used. For external eyes, flies were anesthetized with ether and imaged. For retinal tissue imaging, fly heads were embedded in paraffin blocks and sectioned for images of endogenous autofluorescence. Images of retina sections were

analyzed by ImageJ to measure the thickness of retina tissue in the middle for quantification.

RU486 induction in adult fly

Adult male flies (0-1d) were collected and aged in vials containing fly media added with or without RU486 as indicated (4 mg/ml or 8 mg/ml, 100 µl) to induce gene and RNAi expression at 25°C for indicated days. 100% ethanol (EtOH) was used as vehicle.

Adult fly climbing assay

Adult male flies were collected for RU486 induction. Assays were conducted between 10-11 am. Plastic vials (height=9.5 cm) were used for the assay. Flies were transferred into plastic vials at the density of 25 flies per vial 20 min before the assay. For the assay, flies were tapped to the bottom of the vial, and the number of flies climbing above 5.5 cm was counted after 15 sec (3 repeats). For each trial, a cohort of 100 flies was assessed for each genotype over 21 days.

Fly lifespan assay

Male flies were collected 0-1 days after eclosion and transferred into vials containing fly media with or without RU486 at a density of 20 flies per vial. Flies were transferred to fresh media every other day and the numbers of dead flies were scored. For each group, 200 flies were used in each experiment.

Western blotting

Adult male fly heads (1-2d unless otherwise noted) were homogenized in Laemmli sample buffer (Biorad) with βME, boiled and centrifuged to remove debris. NuPAGE 4-

12% Bis-Tris gel (Thermo Fisher) was used to run the samples. Proteins were transferred to nitrocellulose membrane by the iBlot blotting system (ThermoFisher). For human cell samples, cells (~$9.6*10^5$) were lysed in RIPA buffer (Cell signaling; #9806) supplemented with PMSF and proteinase inhibitor protease inhibitor cocktail (Roche; #1183670001). Extracts were sonicated for total 6 min using using QSonica (Newtown, CT) water bath sonicator (amplitude = 100, 30 s on, 30 s off). After centrifugation the concentrations of the protein samples were measured by Pierce BCA kit (Thermo; #23225). Same amount of proteins was prepared for each sample with NuPAGE LDS sample buffer (Invitrogen; #NP0007), boiled at 95°C, 5min and run on NuPAGE 4-12% Bis-Tris gel (Thermo Fisher). Proteins were transferred to PVDF membrane by the XCell SureLock system (Invitrogen). Primary antibodies used were anti-TDP-43 rabbit polyclonal antibody (1:5000; Proteintech; #10782-2-AP), anti-β-gal mouse monoclonal antibody (1:2000; Promega; #Z378A), anti-α-tubulin rabbit polyclonal antibody conjugated with HRP (1:1000; Cell signaling; #9099), anti-ELL rabbit polyclonal antibody (1:1000; Proteintech; #51044-1-AP), anti-ELL2 rabbit polyclonal antibody (1:1000; Bethyl; #A302-505A), anti-ELL2 mouse monoclonal for co-IP immunoblots (1:1000; Santa Cruz; #sc-515276) and anti-GAPDH mouse monoclonal (1: 10,000; AbD Seotec; #4699-9555). The secondary antibodies used were goat anti-rabbit IgG-HRP (1:5000; Milipore, #AP307P) and goat anti-mouse (1:5000; Jackson ImmunoResearch; #115-035-146). All blocking and antibody incubations were done in 5% milk in PBS overnight (O/N) at 4°C for primary and 1 hr at room temperature (RT) for secondary. Signals were developed by ECL plus (ThermoFisher) or ECL prime (GE healthcare) western blotting reagents. The images were scanned by Fujifilm LAS-3000 imager (Fujifilm) or Amersham Imager 600 (GE healthcare), and quantification was performed using ImageJ. The uncropped blots in the main figures are provided in Supplementary Fig. 7.

Small RNA Northern blotting

Total RNAs were extracted from fly heads using Trizol reagent (ThermoFisher; #15596026), following the manufacturer's protocol. RNA quality was checked by gel, and 0.6-3 µg of total RNA was loaded in 15% TBE-urea gel (ThermoFisher; #EC68852BOX). RNAs were then transferred onto nylon membrane (GE HealthCare; #RPN303B) and cross-linked by UV. The membrane was then prehybridized by UltraHyb Oligo Hybridization Buffer (ThermoFisher; #AM8663) and then hybridized with P^{32}-labeled probes overnight at 50°C. To make the probes, DNA oligos were annealed to obtain the template for RNA probes, which were synthesized by in vitro transcription using MAXIscript T7 kit (ThermoFisher; #AM1312), supplemented with P^{32}-α-UTP. Signals were detected by GE Amersham Typhoon 9410 Imager and analyzed by ImageJ for quantification. Probes used are listed in Supplementary Table 4. The same blots were used to probe 2 different snRNAs: either U1 and U11; U2 and U5; U4 and U7 or U4atac and U12. The uncropped scans of the blots are shown in Supplementary Fig. 8.

Drosophila polytene immunohistochemistry

To obtain larvae in the approximately same developmental stage, fly food was mixed with 0.05% bromophenol blue, and light blue wandering larvae were selected for experiments. Salivary glands were dissected in PBST (0.05% Tween20) and fixed with 45% acetic acid and 2% paraformaldehyde for 1 min and transferred into a drop of 45% acetic acid on a sigma-coated coverslip. The salivary glands were then squashed onto Fisherbrand Superfrost Plus Microscope slides (Fisher Scientific), and liquid nitrogen was used to freeze the slides. Slides were blocked at RT for 1 hr, and incubated with

anti-GFP mouse monoclonal antibody (1:500; Takara; #632380), anti-Lilli rabbit polyclonal antibody (1:500; a generous gift from Dr. Ali Shilatifard) and anti-Ell rabbit polyclonal antibody (1:500; a generous gift from Dr. Ali Shilatifard) overnight at 4°C. The secondary antibodies were goat anti-mouse IgG Alexa Fluor488 (1:500; ThermoFisher; #A-11001) and goat anti-rabbit IgG Alexa Fluor488 (1:500; ThermoFisher; #A-11037). All blocking and antibody incubations were done in 3% BSA in TBST (0.05% Tween20). DAPI (ThermoFisher; #D3571) was used for nucleic acid staining. The spreads were mounted using Dako fluorescence mounting medium (Dako; #S3023) and imaged on a Leica DM6000 CS confocal microscope. For heat shock treatment, larvae were incubated for 5 min at 37°C.

Reverse transcription quantitative PCR

Total RNA from larvae, whole fly, fly heads, human cultured cells or human tissue was extracted using Trizol reagent (ThermoFisher; #15596026), following the manufacturer's protocol. DNA was removed by Turbo DNA-free kit (ThermoFisher; #AM1907) for fly samples, RNeasy micro kit (QIAGEN; #74004) for human cultured cells or DNase I (ThermoFisher; #18068015) for human tissue samples. RNA quality was checked by gel or Agilent 2100 Bioanalyzer system (Agilent). Reverse transcription was performed with random hexamers using High-Capacity cDNA Reverse Transcription kit (ThermoFisher; #4368814) for samples from fly tissues and human cultured cells or SuperScript III Reverse Transcriptase (ThermoFisher; #18080093) for human tissue samples. Fast SYBR green master mix (ThermoFisher; #4385614) was used for qPCR performed by ViiA 7 Real-Time PCR system (ThermoFisher). Primers used for qPCR are listed in Supplementary Table 5, 6. Sequences of primers for *CG6323*, *CG8408*, *CG16941*,

CG11839, *CG7892*, *CG13431*, *CG33108*, *CG16941* U12int and *CG11839* U12int are as previously described [100].

ChIP-Seq data analysis

Raw fastq data from GSM345570 and GSM345568 (input control and Lilli ChIP-seq, respectively, were used to call Lilli peaks [69]. Raw fastq data from GSM2224492 and GSM2224493 (input control replicates), GSM2224501 GSM2224502, GSM2224503, and GSM2224504 (TBPH ChIP-seq) were used to call narrow TBPH peaks [114]. Narrow peaks were called using the mosaics R package [141]. Read trimming, alignment and peak calling steps were performed in R using a published pipeline [142]. Software packages used were mosaics v2.12.0, Rbowtie v1.14.0, dada2 v1.2.2, quasR v 1.14.0, TxDb.Dmelanogaster.UCSC.dm6.ensGene v3.3.0, GenomicRanges v1.26.4, GenomicFeatures v1.26.4, BSgenome v1.42.0, biomaRt v2.30.0, AnnotationDbi v1.36.2, deeptools v3.1.0, python v2.7.10, Integrative Genomics Viewer v2.3.93, and R v3.3.2. Mosaics peaks were called using the default parameters, with the exceptions of analysis type IO, false discovery rate (FDR)=0.05. Only regions with a called mosaics narrow peak in all biological replicates, when compared to both controls, were counted as a peak. If these peaks lay within 250 bp of a transcription start site, the corresponding gene was considered to have a peak.

Human cell culture

The doxycycline-inducible (TetON) HEK293 cell lines, TetOn-GFP (clone#9.3) and TetOn-GFP-TDP-43 (clone#12.5) generated using a subclone of HEK293 cells (QBI-293) are generous gifts from Dr. Virginia Lee's laboratory. Cells were grown in

Dulbecco's modified Eagle's medium (DMEM) with L-glutamine, glucose and sodium pyruvate (Corning; #MT10013CV), supplemented with 10% Tet system approved FBS, 1% penicillin/streptomycin. Media was further added 400 µg/ml G418 and 1 µg/ml puromycin. Cells were cultured at 37°C and 5% CO_2 and routinely sub-cultured at 1:10 ratio every 7 days. To induce expression, 1000 ng/ml doxycycline was added for 6 d. For heat shock treatment, cells were subjected to heat stress at 46°C followed by 6 hr recovery. Cells were checked for appropriate GFP-tagged protein expression.

Human frontal cortical tissue

Human postmortem brain samples were obtained from the University of Pennsylvania Center for Neurodegenerative Disease Brain Bank. All relevant ethical regulations were compiled, and informed consent from next of kin was obtained for all cases. These comprised samples from clinically normal individuals (n = 6), as well as individuals with FTLD-TDP (n = 13). The region sampled was midfrontal cortex (BA9), and all disease cases were previously reported and confirmed to have TDP-43 pathology [121,122]. DNA was extracted from all cases and screened for mutations in the two most common genes associated with FTLD-TDP, *GRN* and *C9orf72*. Briefly, the coding regions of *GRN* were bi-directionally sequenced by Sanger sequencing using flanking primers to each exon as previously described [143]. Sequence analysis was done with Mutation Surveyor (SoftGenetics, State College, PA). Analysis for hexanucleotide repeat expansions in *C9orf72* was performed using a modified repeat-primed PCR method [144]. Analysis of the valosin containing protein gene (*VCP)* analysis was performed by targeted Sanger sequencing of the relevant exon in a case with known family history of a *VCP* mutation [145].

Co-IP

Endogenous Co-IP was performed by using the nuclear extract of HEK293 TetOn-GFP (clone#9.3) cells without doxycycline induction. Cells (~$2.6*10^7$) were resuspended in hypotonic solution (20 mM Tris-HCl (pH 7.5), 20 mM NaCl, 5 mM $MgCl_2$) supplemented with protease inhibitor cocktail (Roche; #1183670001) and homogenized by Dounce homogenizer. After centrifugation (3,000 g, 4°C, 15 min), the pellet containing nucleus was resuspended in Pierce IP lysis buffer (Thermo; #87787) supplemented with protease inhibitor cocktail (Roche; #1183670001). After centrifugation (15,000 g, 4°C, 10 min), input sample was saved, and the rest of the lysate were divided to incubate at 4°C, overnight with Dynabeads Protein G (Invitrogen; #1004D) prepared with 5 μg of anti-TDP-43 mouse monoclonal antibody- mAb 5028[146] (a generous gift from Dr. Virginia Lee's lab) or same amount of mouse IgG control (Santa Cruz; #sc-2025). The next day, the beads were washed with lysis buffer 3 times, with the 3rd time rotating the tubes for 5 min in 4°C. Elute the proteins by NuPAGE LDS sample buffer (Invitrogen; #NP0007) by boiling at 95°C, 5min. The following western blotting were performed as described above. Experiments were repeated to confirm the results.

Statistical analysis

Graphs are represented as mean±standard deviation (SD). The statistics used are indicated in each figure legend. Comparison between two groups were calculated using the two-tailed unpaired Student's *t* test. If data sets show significant variance according to variance F-test analysis, two-tailed unpaired Student's *t* test with Welch's correction was used as indicated in the figure legend. Shapiro-Wilk normality test was used for

testing normal distribution for all data sets except data sets normalized to 1. If the data are not normally distributed, two-tailed unpaired Mann-Whitney Test (nonparametric test) was used as indicated in the figure legend. The differences among three groups were calculated using ANOVA followed by Tukey's multiple comparison test. Brown-Forsythe test and Shapiro-Wilk normality test were used to test variance differences and normality respectively for data analyzed by One-way ANOVA. Differences with *P*-values <0.05 were considered statistically significant. The number of sample size and biological replicates is indicated in the methods section or figure legend. No statistical method was used to predetermine sample sizes. No sample was excluded from the analysis. No method of randomization was used. The investigators were not blinded to allocation during experiments.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Code availability

The code used for analysis is available upon request.

Acknowledgments

We thank Dr. Maya Capelson and Terra Kuhn for helpful input, and Dr. Joe Zhou and members of the Bonini laboratory for critical reading of the manuscript. We thank former Bonini laboratory members William Motley and Ross Weber for early studies that helped launch this research. This work was supported by the NIH (R01-NS078283, N35-

NS097275 to NMB; AG-017586 to VVD), a Glenn Award for Research in the Biological Mechanisms of Aging (to NMB).

Author contributions

Conceptualization: C.Y.C., N.M.B.; Methodology: C.Y.C., A.B., J.K., A.Sartoris, T.U., V.V.D., E.R.S., A.Shilatifard, A.C.; Investigation: C.Y.C., J.K., H.J.K.; Resources: A.B., A.Sartoris, T.U., S.P., V.M.L., E.R.S., A.Shilatifard, A.C.; Supervision: N.M.B.; Writing: C.Y.C., N.M.B.

Competing interests

The authors declare no competing interests.

Figures

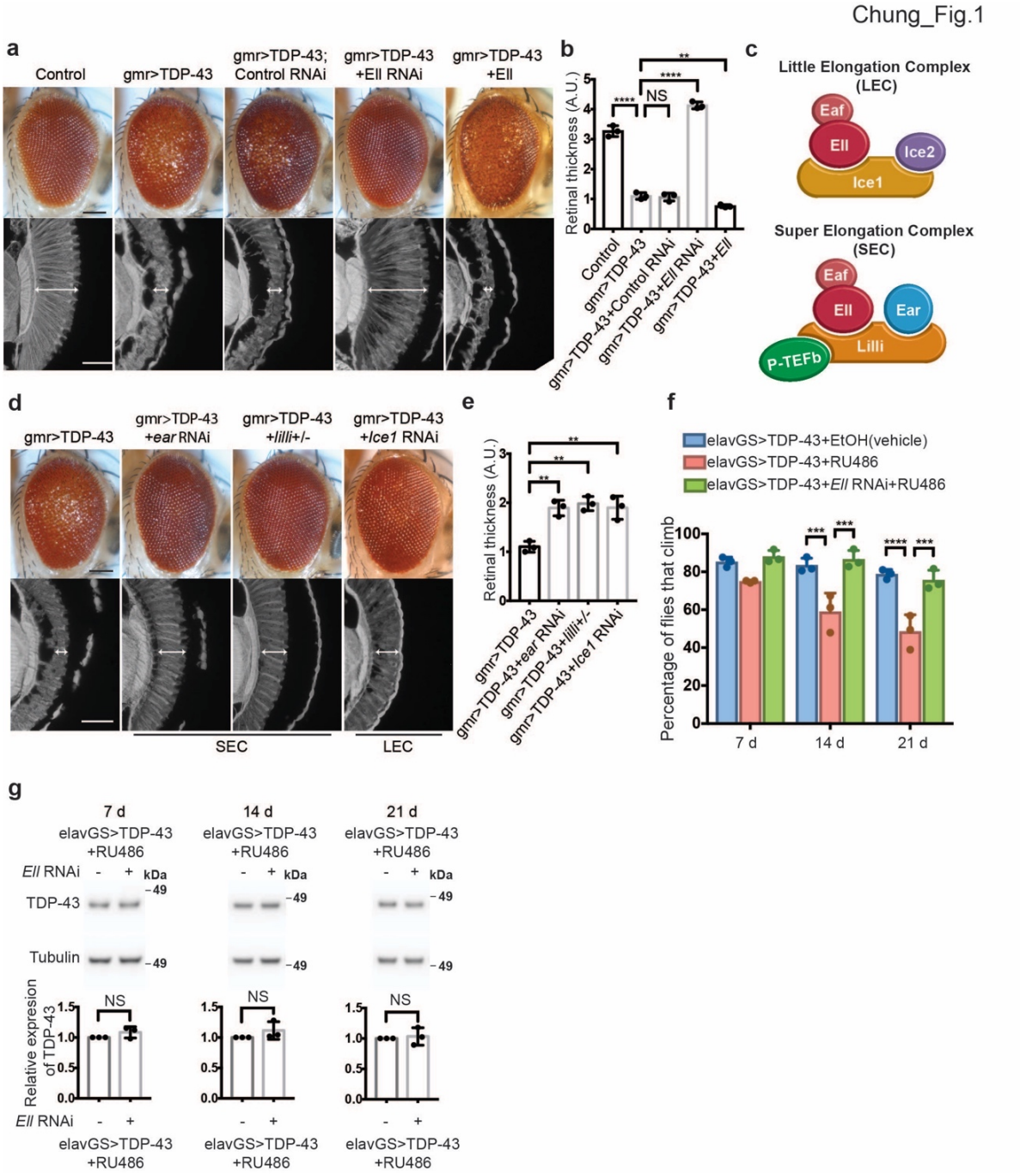


Figure 1. Components of LEC and SEC modulate TDP-43 toxicity
(**a**) TDP-43 expression causes eye degeneration that is suppressed by *Ell* RNAi and enhanced by *Ell* upregulation. Scale bars: external eye images (top), 100 μm; internal retinal section (bottom), 5 μm. Genotypes: Control is *gmr-GAL4(YH3)/+*. gmr>TDP-43 is

UAS-TDP-43/+; gmr-GAL4(YH3)/+. gmr>TDP-43+Control RNAi is *UAS-TDP-43/+; gmr-GAL4(YH3)/UAS-Control.RNAi*^{JF01355}. gmr>TDP-43+*Ell* RNAi is *UAS-TDP-43/+; gmr-GAL4(YH3)/UAS-Ell.RNAi*^{HMS00277}. gmr>TDP-43+*Ell* is *UAS-TDP-43/+; gmr-GAL4(YH3)/UAS-Ell*$^{P\{EP\}G4098}$.

(**b**) Quantification of retina thickness related to Fig. 1a. Three flies of each genotypes were measured (n=3). Bars represent mean (SD). (**) *P* <0.01, (****) *P*<0.0001(Two-tailed unpaired Student's *t* test). Genotypes are the same as indicated in Fig. 1a. (A.U.) arbitrary units.

(**c**) Schematic of LEC and SEC.

(**d**) Downregulation of SEC components (*ear* and *lilli*), or LEC component (*Ice1*) suppresses TDP-43 toxicity. Scale bars: external eye (top), 100 µm; internal retina section (bottom), 5 µm. Genotypes: gmr>TDP-43 is *UAS-TDP-43/+; gmr-GAL4(YH3)/+*. gmr>TDP-43+*ear* RNAi is *UAS-TDP-43/+; gmr-GAL4(YH3)/UAS-ear.RNAi*^{HMS00107}. gmr>TDP-43+*lilli+/-* is *UAS-TDP-43/lilli*$^{17\text{-}2}$*; gmr-GAL4(YH3)/+*. gmr>TDP-43+*Ice1* RNAi is *UAS-TDP-43/UAS-Ice1.RNAi* (SH09112.N from DRSC/TRiP)*; gmr-GAL4(YH3)/+*.

(**e**) Quantification of retina thickness related to Fig 1d. Three flies of each genotypes were measured (n=3). Bars represent mean (SD). (**) *P* <0.01 (Two-tailed unpaired Student's *t* test). Genotypes are the same as indicated in Fig. 1d. (A.U.) arbitrary units.

(**f**) TDP-43 expression in the adult neurons by *elavGS* causes climbing defects. Knockdown of *Ell* restores climbing ability. RU486 (8 mg/ml) was used to induce the expression of TDP-43 and *Ell* RNAi. EtOH was used as vehicle. Total of 100 flies were measured 3 times for each genotype at different time points. Bars represent mean (SD). (***) *P* <0.001, (****) *P*<0.0001 (Two-way ANOVA followed by Tukey's multiple comparison test). Significant differences are only indicated within the same time point. Genotypes: elavGS>TDP-43 is *elavGS-GAL4, UAS-TDP-43/+*. elavGS>TDP-43+*Ell* RNAi is *elavGS-GAL4, UAS-TDP-43/UAS-Ell.RNAi*^{HMS00277}.

(**g**) TDP-43 protein levels in heads from 7, 14 and 21d flies are not altered by *Ell* RNAi. *n*=3 biological replicates. Tubulin was used as internal control. Bars represent mean (SD). (NS) not significant (Two-tailed unpaired Student *t* test with Welch's correction). Genotypes and RU486 treatment are as in Fig. 1f.

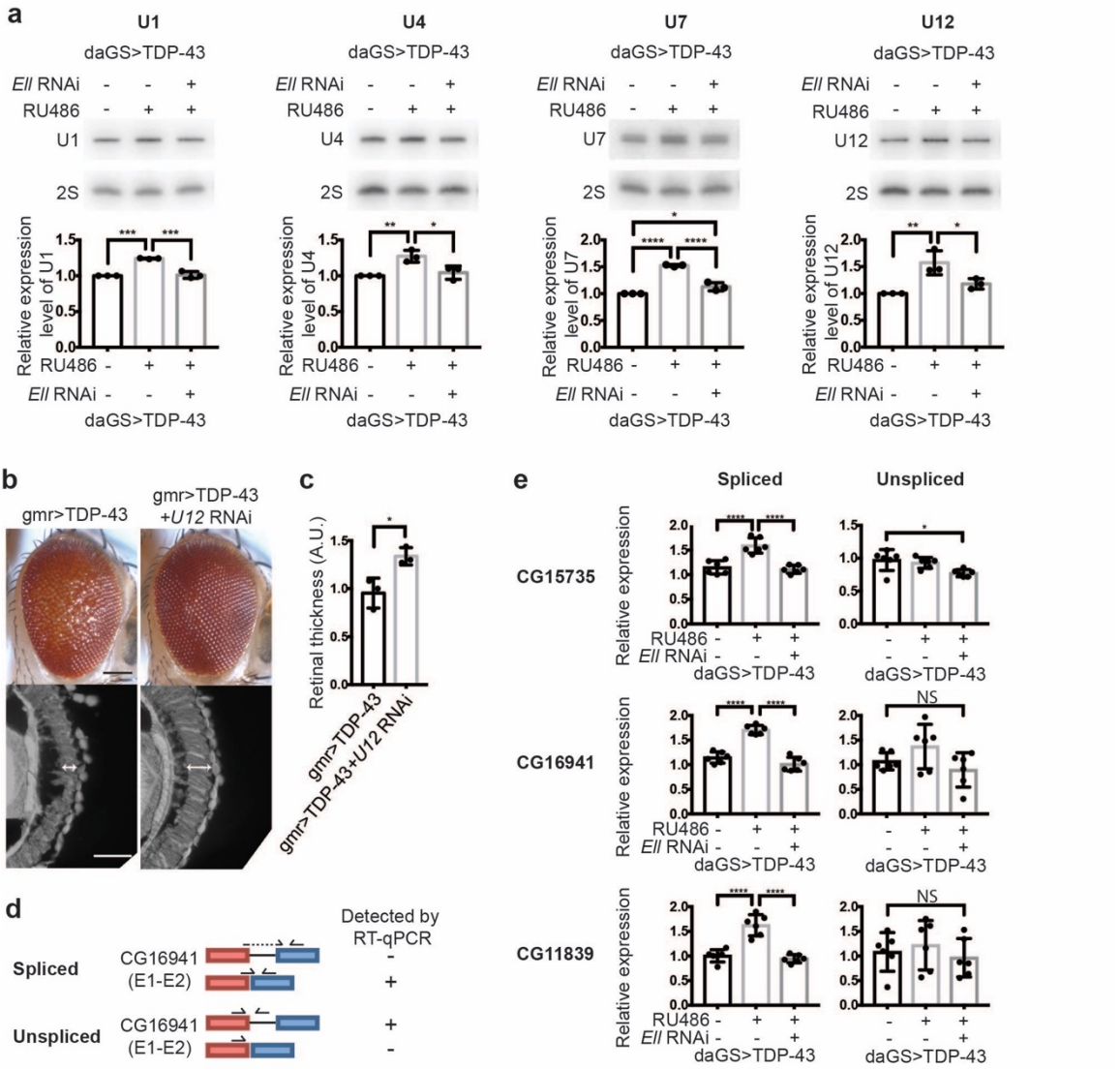


Figure 2. U12 snRNA is upregulated and a functional target of TDP-43
(**a**) Northern blot analysis shows that the level of snRNAs U1, U4, U7 and U12 are upregulated by TDP-43 expression driven by a ubiquitous drug-inducible promoter, *daGS*, in heads. *Ell* RNAi corrects the increased levels of snRNAs. RU486 (4 mg/ml) was used to induce the expression of TDP-43 and *Ell* RNAi for 8d. 2S rRNA was the loading control. *n*=3 biological replicates. Bars represent mean (SD). (*) *P* <0.05, (**) *P* <0.01, (***) *P* <0.001, (****) *P*<0.0001 (One-way ANOVA followed by Tukey's multiple comparison test). Genotypes: daGS>TDP-43 is *daGS-GAL4/+; UAS-TDP-43/+*. daGS>TDP-43+Ell RNAi is *daGS-GAL4/+; UAS-TDP-43/UAS-Ell.RNAi*^{HMS00277}.
(**b**) Knockdown of *U12* suppresses eye degeneration caused by TDP-43 toxicity. Scale bars: external eye (top), 100 µm; internal retina (bottom), 5 µm. Genotypes: gmr>TDP-43 is *UAS-TDP-43/+; gmr-GAL4(YH3)/+*. gmr>TDP-43+*U12* RNAi is *UAS-TDP-43/ UAS-snRNA:U12:73B.RNAi*^{HMC03841}*; gmr-GAL4(YH3)/+*.

(**c**) Quantification of retina thickness related to Fig 2b. Three flies of each genotypes were measured (n=3). Bars represent mean (SD). (*) P <0.05 (Two-tailed unpaired Student's *t* test). Genotypes are the same as indicated in Fig. 2b. (A.U.) arbitrary units.
(**d**) Schematic illustrates detection of spliced or unspliced products by RT-qPCR for a U12-type intron.
(**e**) RT-qPCR analysis shows that spliced products of *CG15735*, *CG16941*, *CG11839* are up-regulated by TDP-43 expression driven by *daGS* in fly heads, and downregulation of *Ell* rescues the increased levels, while the unspliced products are not changed significantly upon TDP-43 expression. RU486 (4 mg/ml) was used to induce expression for 8d. mRNA levels were normalized to RpL32 mRNA. n=6 biological replicates. Bars represent mean (SD). (*) P <0.05, (****) P<0.0001 (One-way ANOVA followed by Tukey's multiple comparison test). Genotypes are as indicated in Fig. 2a.

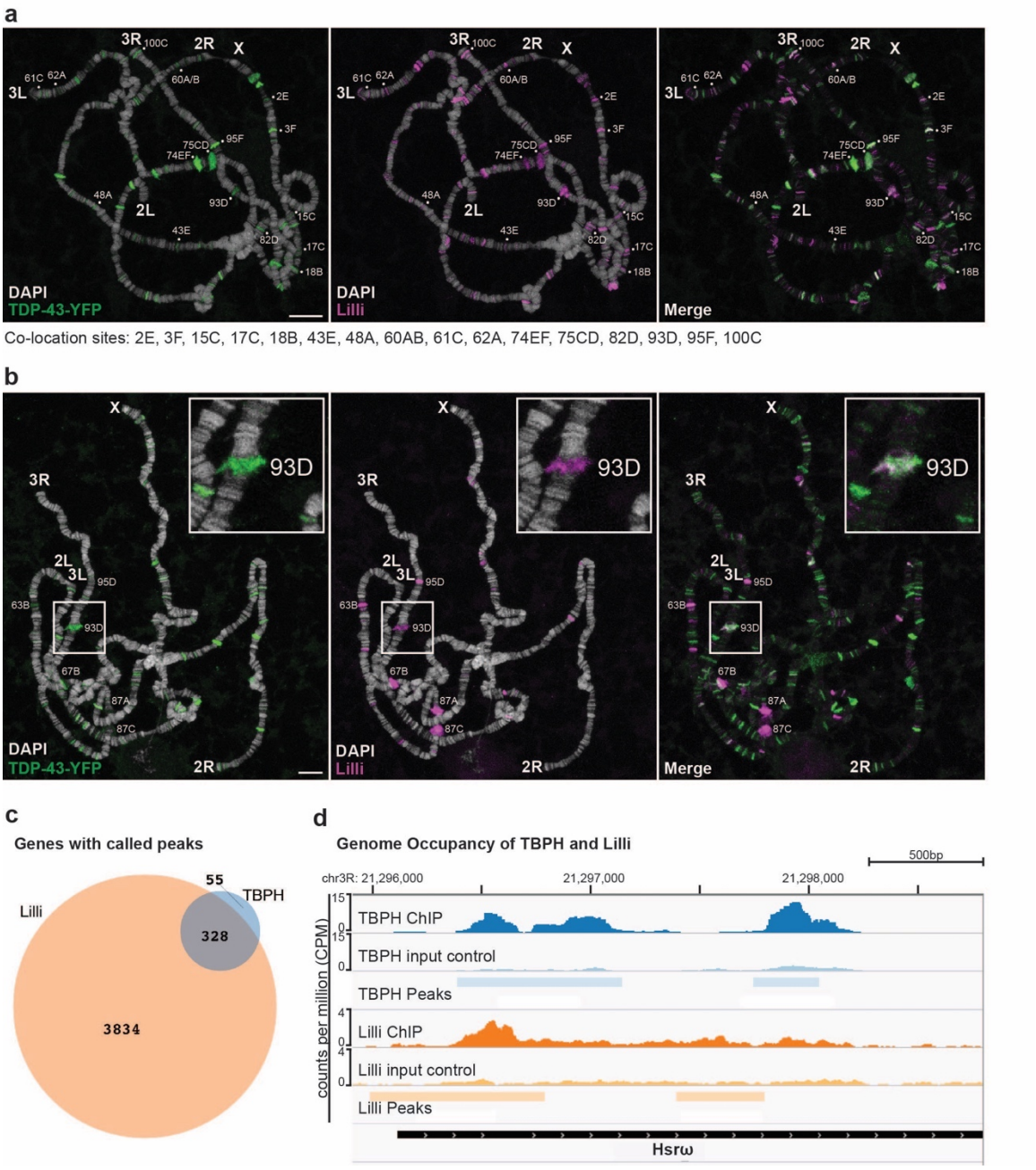


Figure 3. TDP-43 and Lilli co-localize at the 93D locus on polytene chromosomes
(**a**) Immunostainings of DAPI (white), TDP-43-YFP (green) and Lilli (magenta) show partial overlap of TDP-43-YFP and Lilli on chromosomes from salivary glands expressing TDP-43-YFP. Sixteen sites of consistent co-localization are indicated (bottom, see also Supplementary Table 1).
(**b**) Upon heat shock, Lilli localizes to major heat shock loci and co-localizes with TDP-43-YFP at the 93D locus (enlarged at the upper right corner). See also Supplementary Table 2.

For (**a**) and (**b**), scale bars: 10 μm. Genotype: *sgs3-GAL4/UAS-TDP-43-YFP*.
(**c**) Venn diagram illustrates the numbers of genes bound by TBPH and Lilli.
(**d**) A ChIP-seq profile of TBPH (top, blue) and Lilli (bottom, orange) shows overlap at the *Hsrω* transcription start site. Bar intervals indicate called peaks in TBPH (light blue, top) and Lilli (light orange, bottom). The TBPH profile is one of four biological replicates that show called peaks at *Hsrω* transcription start site.

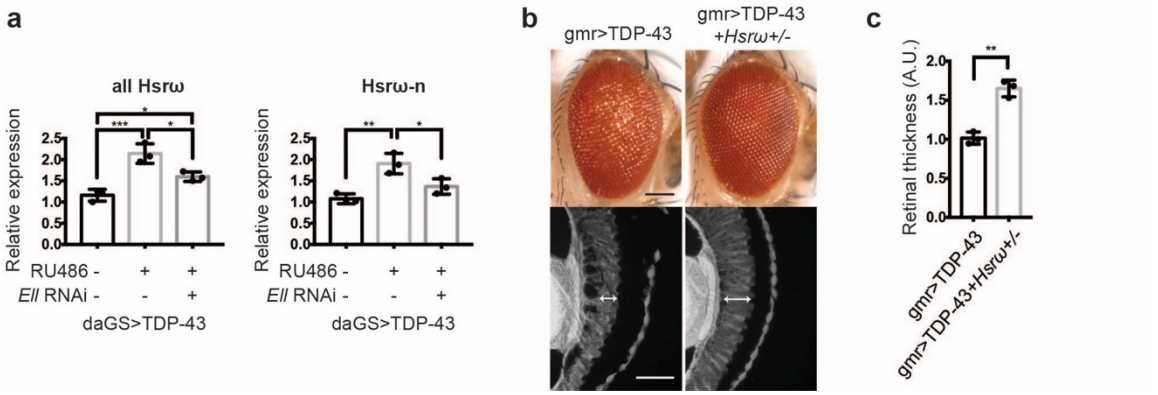


Figure 4. Hsrω is elevated and contributes to TDP-43 toxicity
(**a**) RT-qPCR analysis of fly heads shows that expression of TDP-43 driven by a drug-inducible promoter *daGS* leads to an increase of total *Hsrω* and *Hsrω-n*. Knockdown of *Ell* reduces the elevated levels of total *Hsrω* and *Hsrω-n*. RU486 (4 mg/ml) was used to induce expression for 4d. Relative RNA levels were normalized to *Pgk*, *RpL32* and *RpS20* mRNAs (geometric mean). *n*=3 biological replicates. Bars represent mean (SD). (*) *P* <0.05, (**) *P* <0.01, (***) *P* <0.001 (One-way ANOVA followed by Tukey's multiple comparison test). Genotype: daGS>TDP-43 is *daGS-GAL4/+; UAS-TDP-43/+*. daGS>TDP-43+*Ell* RNAi is *daGS-GAL4/+; UAS-TDP-43/UAS-Ell.RNAi*^{HMS00277}.
(**b**) Images of external eyes (top) and internal retina (bottom) show that loss of one copy of *Hsrω* ameliorates TDP-43-caused eye degeneration. Scale bars: external eye, 100 µm; internal retina section, 5 µm. Genotypes: gmr>TDP-43 is *UAS-TDP-43/+; gmr-GAL4(YH3)/+*. gmr>TDP-43+*Hsrω+/-* is *UAS-TDP-43/+; gmr-GAL4(YH3)/Hsrω*^{66}.
(**c**) Quantification of retina thickness related to Fig 4b. Three flies of each genotypes were measured (n=3). Bars represent mean (SD). (**) *P* <0.01 (Two-tailed unpaired Student's *t* test). Genotypes are the same as indicated in Fig. 4b. (A.U.) arbitrary units.

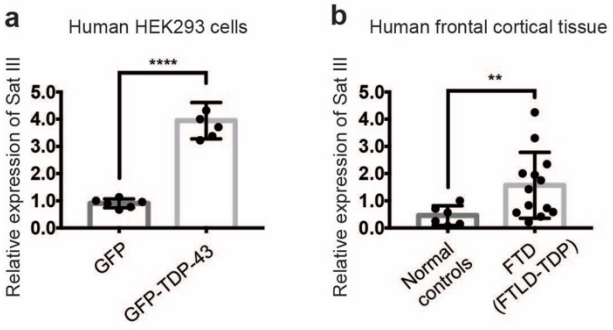


Figure 5. Sat III is upregulated in a human cell model and FTD patient samples
(**a**) RT-qPCR analysis shows that the levels of Sat III are increased in HEK293 cells expressing GFP-TDP-43 after 6d induction. Relative RNA levels were normalized to *GAPDH* and *ACTB* mRNAs (geometric mean). n=6 biological replicates. Bars represent mean (SD). (****) *P<0.0001* (Two-tailed unpaired Student's *t* test with Welch's correction).
(**b**) RT-qPCR analysis shows that the levels of Sat III are increased significantly in frontal cortex of FTD patients compared to normal frontal cortex controls. Relative RNA levels were normalized to *GAPDH* and *ACTB* mRNAs (geometric mean). All FTD patients had TDP-43 pathology (FTLD-TDP). Case numbers and details are as indicated in Table 1. Bars represent mean (SD). (**) *P<0.01* (Two-tailed unpaired Student's *t* test with Welch's correction).

Chung_Fig.6

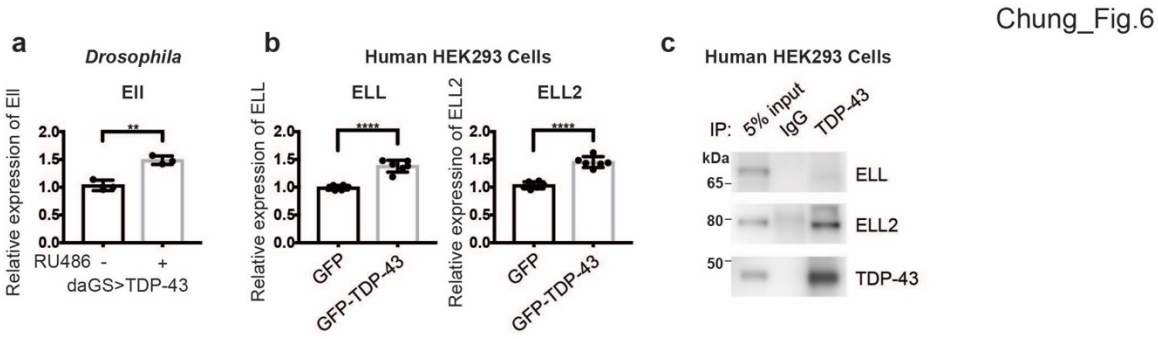


Figure 6. TDP-43 interacts with ELL2 in human cells

(**a**) RT-qPCR analysis of fly heads shows that expression of TDP-43 driven by *daGS* leads to an increase of *Ell*. RU486 (4 mg/ml) was used to induce expression for 4d. Relative RNA levels were normalized to *Pgk*, *RpL32* and *RpS20* mRNAs (geometric mean). *n*=3 biological replicates. Bars represent mean (SD). (**) *P* <0.01 (Two-tailed unpaired Student's *t* test). Genotypes are the same as indicated in Fig. 4a.

(**b**) RT-qPCR analysis shows that the levels of *ELL* and *ELL2* are increased in HEK293 cells expressing GFP-TDP-43 compared to cells expressing GFP after 6d induction. Relative RNA levels were normalized to *GAPDH* and *ACTB* mRNAs (geometric mean). n=6 biological replicates. Bars represent mean (SD). (****) *P<0.0001* (Two-tailed unpaired Student's *t* test).

(**c**) IP with anti-TDP-43 antibody or mouse IgG as negative control followed by immunoblotting studies with anti-ELL, ELL2 or TDP-43 antibody show that TDP-43 interacts with ELL2. The co-IP assays were repeated independently four times and show consistent results (a different ELL2 antibody was used for detection for two repeats of the experiments).

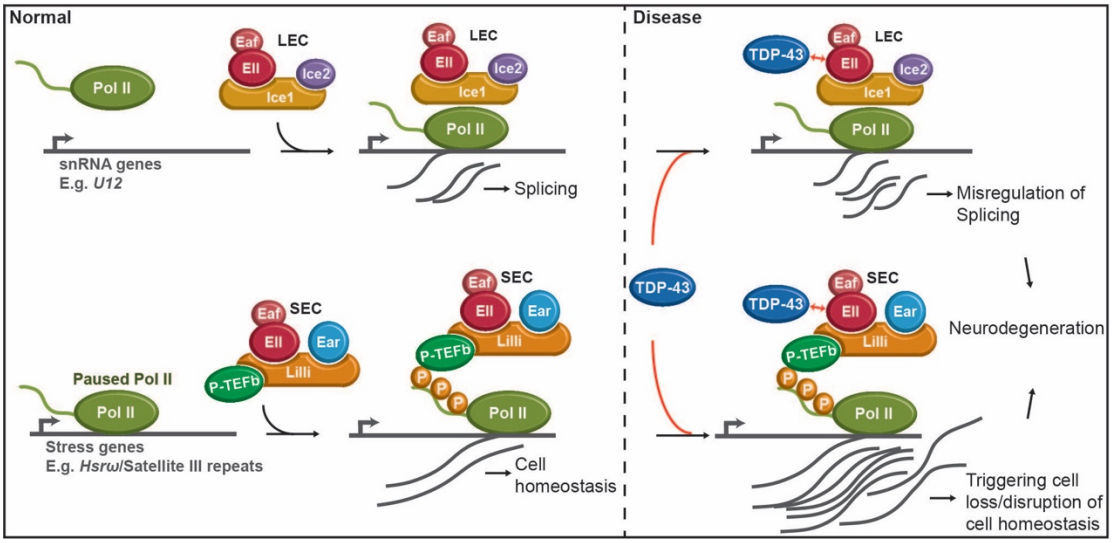


Figure 7. TDP-43 promotes the levels of the targets of LEC and SEC contributing to neurodegeneration

A model for TDP-43 toxicity associated with LEC and SEC. LEC regulates the transcription of snRNAs, forming spliceosomes, and SEC regulates stress response genes, maintaining cell homeostasis. In the disease state, TDP-43 promotes the levels of LEC and SEC targets through interactions with Ell in fly/ELL2 in human cells, including U12 and *Hsrω*/Sat III, leading to misregulation of splicing, cell loss and disruption of cell homeostasis and contributes to neurodegeneration.

Tables

Table 1. Human brain samples

		Number
Normal controls	No mutation identified	6
FTD (FTLD-TDP)	No mutation identified	6
	GRN mutation	5
	C9Orf72 mutation	1
	VCP mutation	1

Supplementary Information

Supplementary Table 1. Quantification of co-localization of TDP-43-YFP and Lilli on 16 sites in 4 polytenes, related to Figure 3.

Chromosome	~Cytological bands	Polytene 1	Polytene 2	Polytene 3	Polytene 4
X	2E	+	+	+	+
	3F	+	?	+	+
	15C	+	+	+	+
	17C	+	+	+	+
	18B	+	+	+	+
2L					
2R	43E	+	+	+	+
	48A	+	+	+	+
	60A/B	+	+	+	+
3L	61C	+	+	+	+
	62A	+	+	?	+
	74 EF	+ (weak)	+	+	+
	75CD	+ (weak)	+	+ (weak)	+
3R	82D	+	+	+	+
	93D	+	+	+	+
	95F	+	+	+	+
	100C	+	+	+	+

(+) co-localization, (?) unable to define due to pattern of chromosome spreading.

Supplementary Table 2. Quantification of co-localization of TDP-43-YFP and Lilli on 6 major heat shock loci in 15 polytenes, related to Figure 3.

Major heat shock loci	63B	67B	93D	95D	87A	87C
Polytene 1	+ (weak)	+ (weak)	+	-	-	-
Polytene 2	-	+	+	-	-	-
Polytene 3	+ (weak)	-	+	-	-	-
Polytene 4	+ (very week)	+ (weak)	+	-	-	-
Polytene 5	+ (weak)	-	+	-	-	-
Polytene 6	+ (weak)	-	+	-	-	-
Polytene 7	+ (weak)	-	+	-	-	-
Polytene 8	-	-	+ (weak)	-	-	-
Polytene 9	-	-	+	-	-	-
Polytene 10	+	-	+	-	-	-
Polytene 11	+ (very weak)	-	+	-	-	-
Polytene 12	+ (very weak)	-	+	-	-	-
Polytene 13	-	-	+	-	-	-
Polytene 14	-	-	+	-	-	-
Polytene 15	-	-	+	-	-	-

(+) co-localization, (-) no co-localization.

Supplementary Table 3. Fly lines.

Fly lines	Insertion chromosome(s)	Reference/Source
w^{1118}	1	Bloomington
UAS-TDP-43-YFP	3	[37]
UAS-TDP-43	2	[37]
UAS-TDP-43	3	[37]
elavGS-GAL4	3	Bloomington
daGS-GAL4	2	[140]
da-GAL4	3	Bloomington
gmr-GAL4(YH3)	3	[37]
sgs3-GAL4	3	Bloomington
UAS-LacZ	2	Bloomington
$lilli^{17-2}$	2	Bloomington
$Hsr\omega^{66}$	3	Bloomington
UAS-$Ell^{P\{EP\}G4098}$	3	Bloomington
UAS-Ell.RNAi^{HMS00277}	3	Bloomington
$Ell^{[S-192]}$	3	Bloomington
UAS-ear.RNAi^{HMS00107}	3	Bloomington
UAS-Ice1.RNAi (SH09112.N from DRSC/TRiP)	2	This paper
UAS-Ice1.RNAi (SH09113.N from DRSC/TRiP)	3	This paper
UAS-Ice1.RNAi (13550R-2 from NIG)	2	NIG
UAS-snRNA:U12:73B.RNAi^{HMC03841}	2	Bloomington
UAS- Hsrω.RNAi^{HMC05093}	2	Bloomington
UAS-Control.RNAi^{JF01355}	3	Bloomington
UAS-$(G_4C_2)_{48}$	2	[147]
UAS-mCherry RNAi	2	Bloomington

Bloomington is to the Bloomington Drosophila Stock Center, https://bdsc.indiana.edu/

Supplementary Table 4. Probes for small RNA Northern blot.

Name	Sequence
U1	5′-GATAATACGACTCACTATAGGGAGA-3′ 5’-AAAAAACTGAGTTGACCTCTGCGATTATTCCTCTCCCTATAGTGAGTCGTATTATC-3′
U2	5′-GATAATACGACTCACTATAGGGAGA-3′ 5’-AAAAAATCGGCCTTATGGCTAAGATCAAATCTCCCTATAGTGAGTCGTATTATC-3′
U4	5′-GATAATACGACTCACTATAGGGAGA-3′ AAAAAAGAAAACTTTAACCAATACCCCGCCTCTCCCTATAGTGAGTCGTATTATC-3′
U4atac	5′-GATAATACGACTCACTATAGGGAGA-3′ 5’-AAAAAATCAATGAACGTCTAGTGAGGACATTTCTCCCTATAGTGAGTCGTATTATC-3′
U5	5′-GATAATACGACTCACTATAGGGAGA-3′ 5’-AAAAAATCTGGTTTCTCTTCAATTGTCGAATTCTCCCTATAGTGAGTCGTATTATC-3′
U7	5′-GATAATACGACTCACTATAGGGAGA-3′ 5’-AAAAAACTCTTTGAAATTTGTCTTGGTGGGATCTCCCTATAGTGAGTCGTATTATC-3′
U11	5′-GATAATACGACTCACTATAGGGAGA-3′ 5’-AAAAAAGTTTCCGATCACGAAACTCAAGTGTCTCCCTATAGTGAGTCGTATTATC-3′
U12	5′-GATAATACGACTCACTATAGGGAGA-3′ 5’-AAAAAAAATGAGTAAGGAAAACCAATCAGCCTCTCCCTATAGTGAGTCGTATTATC-3′
2S	5′-GATAATACGACTCACTATAGGGAGA-3′ 5’-TGCTTGACTACATATGGTTGAGGGTTGTATCTCCCTATAGTGAGTCGTATTATC-3′

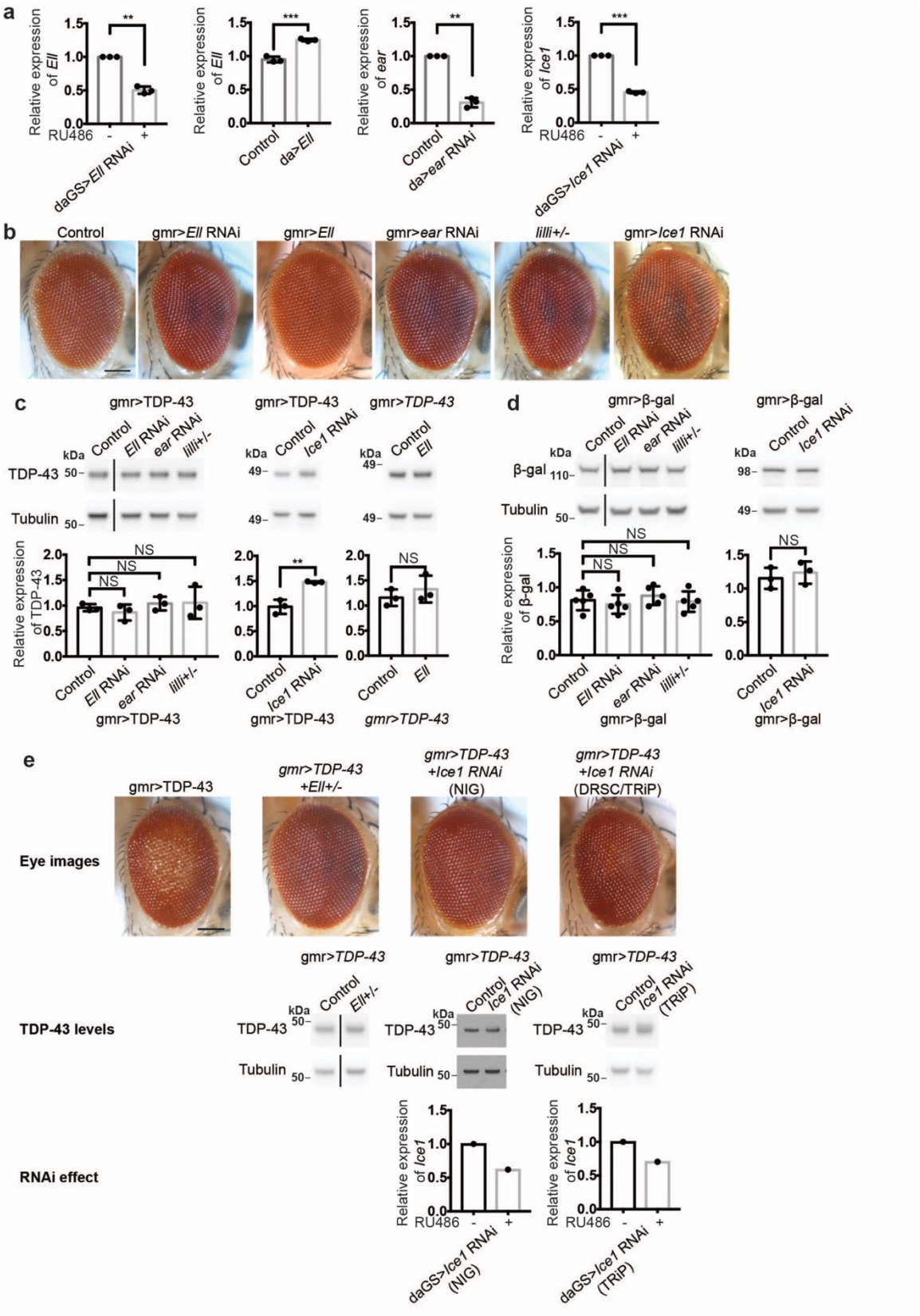
Two DNA oligonucleotides were annealed to obtain the template for each RNA probe.

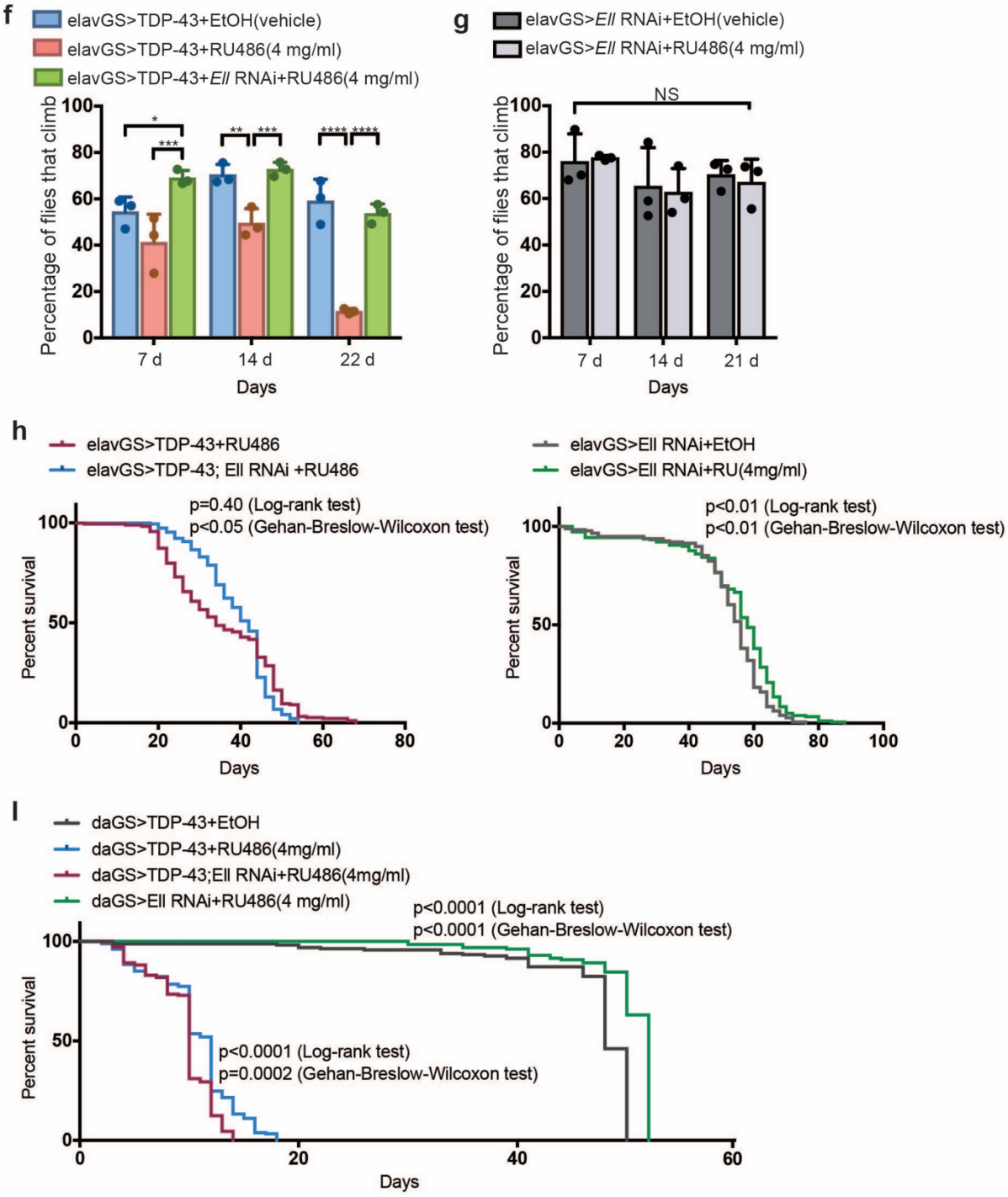
Supplementary Table 5. qPCR primers for *Drosophila* genes.

Name	Forward Sequences (5'-3')	Reverse Sequences (5'-3')
Ell	AAAACAAACTCACATACAACCAAAAA	GCACTCGTTGTGGAAGTAGACA
Ear	TTCCGAAACCCAAACGAGT	CGTCTCGGTTGCGAAAGTA
Ice1	ACCACTTCCAGCCCTATGAG	GCTTCCCCACCAGGTTCT
U12	TTTGCAAGGGCACAGGTC	GCTAGCCGGACGCAAAGTAG
CG6323	TGCTCTTCTTCTACATGATCATTCTGTT	GCACTTGTTTGCGCAAATCC
CG8408	GTCGCCTACGTATTCCTGCAAACA	CCAGCGCCGAACAGAAACATATGA
CG16941	GCCGGAAGTTAGAAATATCGTTGAC	ACTTTGGATTGCCCAGCTCA
CG11839	CACCATTGGCAAAGTATGACTCCTCG	CCACCTTATTCTGCTTAGCCTGATCG
CG7892	CTTTCAAGAAATCTATGTGATAACGGAG	GGATCTGGTACAGGAACACCTT
CG13431	ATGGCCCAAATCCTTCTGGGATGA	CCAAACGTGCGAGTGCGTGATATT
CG33108	GCCTACTAGTGCCCTACGATG	TGCGTGTCCACCAGATAACCTA
CG7736	CCTCGAGGACACTATCAGCA	CCTTCATCTGCTTGACCTCATC
CG17912	CCTCCAAACCCTGGTGCT	GGGTCCGGTGTACAACTTCTT
CG32705	CGCCCATCGTAGCCATTATC	CATGTGGAAGCCGGTCAGT
CG11328	CGTTGAGCGATTTCTTCTCC	TCGAACCCGAGTAAATTTGG
CG17228	TGACGGCATGGCTCCTAC	GGGATAGCGCACCCAGAA
CG15081	AGAAGCCGAGGCTGCTAA	TGCGTGCAATGCTTTGAG
CG11984	GCGAGCCGAATCTAGTCACA	GATGCGACGCACTCCAC
CG4894	ATTCGCATTGTGGAGTGG	GTAAACAGCCAAGGCAATACAG
CG18177	GGCAAGGACGGATTCACCT	CATCGACGCGTAGTGCTTG
CG15735	GGCCTTCGATCACAACACA	GTTGCCGTTGCACTCCTT
CG3294	GAACTGCATGCCAAGAGG	GGTGTACTCCAGTTCCGTGT
CG16941 U12int	ATGGCAACTTTAGACGCGGAA	TTGTTTTCTTTTTGGCAAAGGATGC
CG11839 U12int	GTTGTTTACACTTTCTAGCCGGTGT	GCCCGAGGAGTCATACGTATAGTTA
CG33108 U12int	GCATGTAGGTCCTTTGGACTG	CAGTGTGCTGTTCACCGTTTAC
CG11328 U12int	TAACTCCCACCGCACATGA	GGCCACCACACCAGTAAGTT
CG15735 U12int	GGCCTTCGATCACAACACA	AGAAAGTGCTCGCTTTCACC
CG3294 U12int	GCGGTACGTGAACTGCAT	CAGTCGTTACCTTTGGGACA
Hsrω-all	TATCTAATGTCCGGGGTCGT	CACAATCCGCACAATCAATC
Hsrω-n	ATAGTCCCTCGGAGGAAAGG	GCGCTCACAGGAGATCAA
βTub56D	CATCCAAGCTGGTCAGTG	GCCA TGCTCATCGGAGAT
Pgk	ATCACCAGCAACCAGAGAATTG	TGCCAGGGTGTACTTGATGTT
RpS20	CCGCATCACCCTGACATCC	TGGTGATGCGAAGGGTCTTG
RpL32	CATCCGCCCAGCATACAG	CCATTTGTGCGACAGCTTAG

Supplementary Table 6. qPCR primers for human genes.

Name	Forward Sequences (5'-3')	Reverse Sequences (5'-3')
Sat III	GTGCAATCGAATGGAATCG	CCATTCCTGTACTCGGGTTG
GAPDH	CGAGATCCCTCCAAAATCAA	TTCACACCCATGACGAACAT
ACTB	TTCCTGGGCATGGAGTC	CAGGTCTTTGCGGATGTC
ELL	GTCGGAGACGCCTGACTACT	TACTCGGCATTGAAGTCGTTC
ELL2	TGGGAGCAATTCTGCAAC	ATCCAGGCCAGTCTCTTTGA

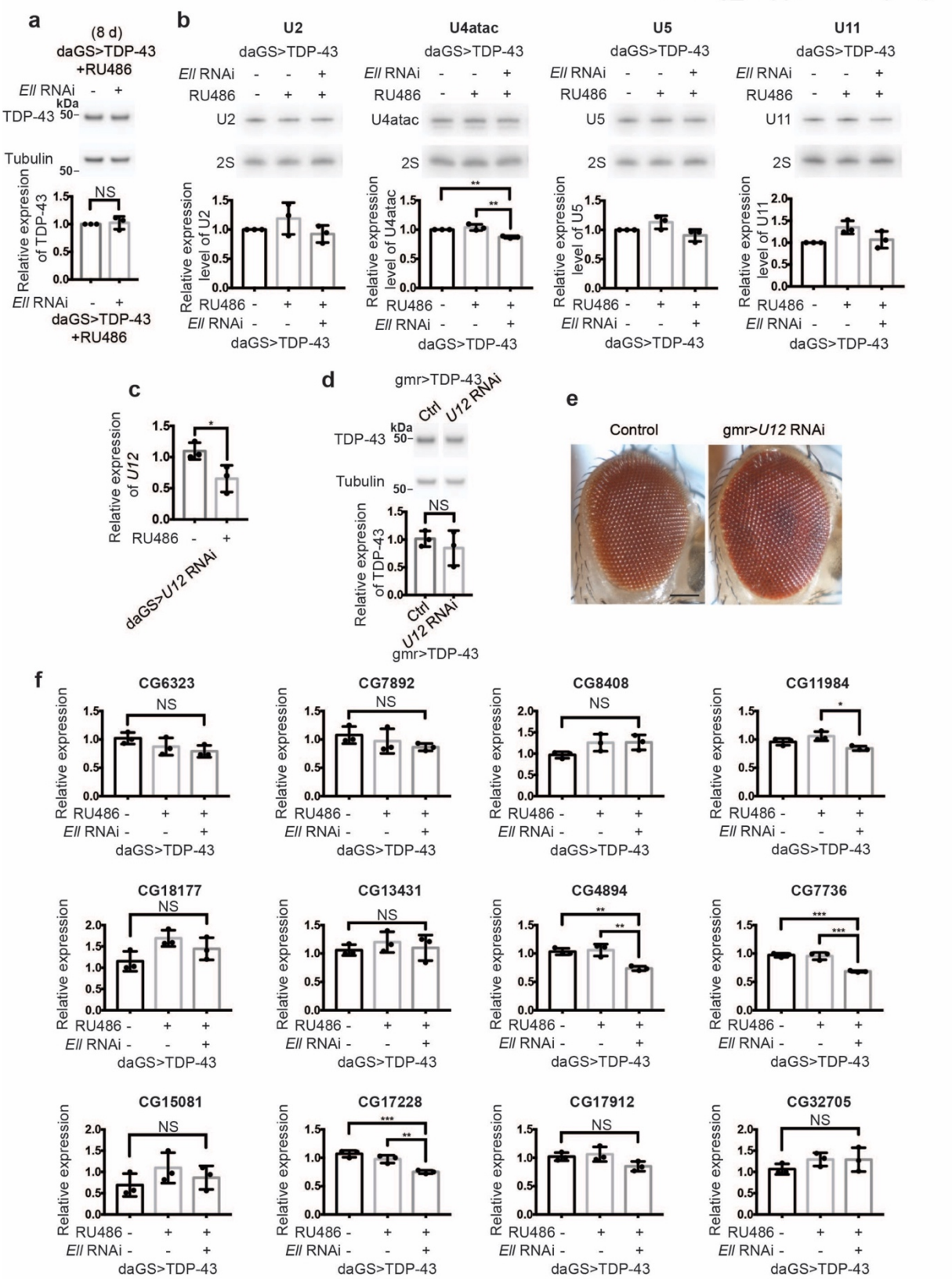

a
Relative expression of Ell
RU486
daGS>Ell RNAi
Control
da>Ell
Relative expression of ear
da>ear RNAi
Relative expression of Ice1
daGS>Ice1 RNAi
b
Control
gmr>Ell RNAi
gmr>Ell
gmr>ear RNAi
lilli+/-
gmr>Ice1 RNAi
c
gmr>TDP-43
TDP-43
Tubulin
kDa
Relative expression of TDP-43
NS
gmr>TDP-43
gmr>TDP-43
gmr>TDP-43
d
gmr>β-gal
β-gal
Relative expression of β-gal
gmr>β-gal
gmr>β-gal
e
Eye images
gmr>TDP-43
gmr>TDP-43 +Ell+/-
gmr>TDP-43 +Ice1 RNAi (NIG)
gmr>TDP-43 +Ice1 RNAi (DRSC/TRiP)
TDP-43 levels
RNAi effect
Relative expression of Ice1
daGS>Ice1 RNAi (NIG)
daGS>Ice1 RNAi (TRiP)

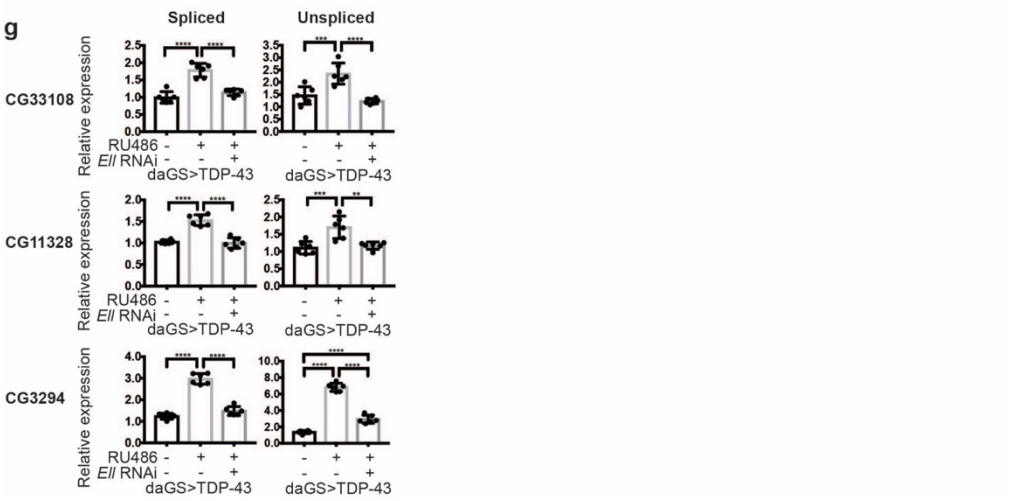

f
elavGS>TDP-43+EtOH(vehicle)
elavGS>TDP-43+RU486(4 mg/ml)
elavGS>TDP-43+Ell RNAi+RU486(4 mg/ml)
Percentage of flies that climb
100
80
60
40
20
0
*
***
**
***
****
****
7 d
14 d
22 d
Days
g
elavGS>Ell RNAi+EtOH(vehicle)
elavGS>Ell RNAi+RU486(4 mg/ml)
NS
Percentage of flies that climb
7 d
14 d
21 d
Days
h
elavGS>TDP-43+RU486
elavGS>TDP-43; Ell RNAi +RU486
p=0.40 (Log-rank test)
p<0.05 (Gehan-Breslow-Wilcoxon test)
Percent survival
Days
elavGS>Ell RNAi+EtOH
elavGS>Ell RNAi+RU(4mg/ml)
p<0.01 (Log-rank test)
p<0.01 (Gehan-Breslow-Wilcoxon test)
Percent survival
Days
I
daGS>TDP-43+EtOH
daGS>TDP-43+RU486(4mg/ml)
daGS>TDP-43;Ell RNAi+RU486(4mg/ml)
daGS>Ell RNAi+RU486(4 mg/ml)
p<0.0001 (Log-rank test)
p<0.0001 (Gehan-Breslow-Wilcoxon test)
p<0.0001 (Log-rank test)
p=0.0002 (Gehan-Breslow-Wilcoxon test)
Percent survival
Days

Supplementary Figure 1. Downregulation of components in SEC and LEC suppresses TDP-43-caused degeneration

(**a**) RT-qPCR analysis verifies that *Ell*, *ear* and *Ice1* are knocked down by RNAi and *Ell* is elevated by the upregulation fly line. RU486 (4 mg/ml) was used to induce the expression of RNAi driven by *daGS* promoter. Whole flies were used for accessing knockdown effects of *Ell* RNAi and *Ice1* RNAi (8d RU486 induction), and upregulation effect of Ell (1d). Larvae were collect for evaluate *ear* RNAi effect. Relative mRNA levels were normalized to βTub56D mRNA. *n*=3 biological replicates. Bars represent mean (SD). (**) *P* <0.01, (***) *P* <0.001 (Two-tailed unpaired Student's *t* test with Welch's correction for RNAi effect; Two-tailed unpaired Student's *t* test for upregulation effect of *Ell*). Genotype: daGS>*Ell* RNAi is *daGS-GAL4/+; UAS-Ell.RNAi*^{HMS00277}/+, Control for Ell is *da-GAL4/+*, da>*Ell* is *da-GAL4/UAS- Ell*$^{P\{EP\}G4098}$, Control for ear is *da-GAL4/mCherry RNAi*, da>*ear* RNAi is *da-GAL4/UAS-ear.RNAi*^{HMS00107}, daGS>*Ice1* RNAi is *daGS-GAL4/UAS-Ice1.RNAi* (SH09112.N from DRSC/TRiP).
(**b**) Reduction of components involved in SEC and/or LEC on their own does not affect eye integrity. Scale bars: 100 μm. Genotype: Control is *gmr-GAL4(YH3)/+*, gmr>*Ell* RNAi is *gmr-GAL4(YH3)/UAS-Ell.RNAi*^{HMS00277}, gmr>*Ell* is *gmr-GAL4(YH3)/UAS-Ell*^{EP}G4098, gmr>*ear* RNAi is *gmr-GAL4(YH3)/UAS-ear.RNAi*^{HMS00107}, *lilli*+/- is *lilli*$^{17-2}$*/+; gmr-GAL4(YH3)/+* and gmr>*Ice1* RNAi is *UAS-Ice1.RNAi* (SH09112.N from DRSC/TRiP)*/+; gmr-GAL4(YH3)/+*.
(**c**) The levels of TDP-43 are unaltered by *Ell* RNAi, *ear* RNAi, loss a genomic copy of *lilli* or upregulation of *Ell*. The levels of TDP-43 are elevated by Ice1 RNAi. *n*=3 biological replicates. Bars represent mean (SD). (NS) not significant, (**) *P* <0.01 (Two-tailed unpaired Student's *t* test). Genotypes of the flies are the same as indicated in Fig. 1a, d.
(**d**) Western blot analysis of β-gal shows that the GAL4/UAS system is not altered by downregulation of components in SEC and/or LEC. *n*=3 biological replicates. Bars represent mean (SD). (NS) not significant (Two-tailed unpaired Student's *t* test). Genotypes: gmr>β-gal; Control is *gmr-GAL4, UAS-β-gal/+,* gmr>β-gal+*Ell* RNAi is *gmr-GAL4, UAS-β-gal/+; UAS-Ell.RNAi*^{HMS00277}*/+,* gmr> β-gal+*ear* RNAi is *gmr-GAL4, UAS-β-gal/+; UAS-ear.RNAi*^{HMS00107}*/+,* gmr> β-gal+*lilli*+/- is *gmr-GAL4, UAS-β-gal/lilli*$^{17-2}$*,* gmr> β-gal+*Ice1* RNAi is *gmr-GAL4, UAS-β-gal/UAS-Ice1.RNAi* (SH09112.N from DRSC/TRiP).
(**e**) Depletion of Ell and Ice1 by additional fly lines show the suppression effect on TDP-43-mediated eye degeneration. The levels of TDP-43 and RNAi effect were confirmed by western blotting and RT-qPCR. For RNAi effect confirmation, RU486 (4 mg/ml) was used to induce the expression of RNAi driven by *daGS* promoter for 8 days. Genotypes: gmr>TDP-43 is *UAS-TDP-43/+; gmr-GAL4(YH3)/+*. gmr>TDP-43+*Ell*+/- is *UAS-TDP-43/+; gmr-GAL4(YH3)/Ell*$^{[S-192]}$. gmr>TDP-43+*Ice1 RNAi(NIG)* is *UAS-TDP-43/UAS-Ice1.RNAi* (13550R-2 from NIG)*; gmr-GAL4(YH3)/+*. gmr>TDP-43+*Ice1* RNAi(TRiP) is *UAS-TDP-43/+; gmr-GAL4(YH3)/UAS-Ice1.RNAi* (SH09113.N from DRSC/TRiP). daGS>*Ice1* RNAi(NIG) is *daGS-GAL4/UAS-Ice1.RNAi* (13550R-2 from NIG) and daGS>*Ice1* RNAi(TRiP) is *daGS-GAL4/+; UAS-Ice1.RNAi* (SH09113.N from DRSC/TRiP)*/+.*
(**f**) TDP-43 expression in the adult neurons by *elavGS* causes climbing defects, which can be rescued by knockdown of *Ell*. Genotypes are as indicated in Fig. 1f.
(**g**) Knockdown of *Ell* on its own in neurons causes no climbing defects. Genotypes: elavGS>*Ell* RNAi is *elavGS-GAL4, UAS-Ell.RNAi*^{HMS00277}*/+.*

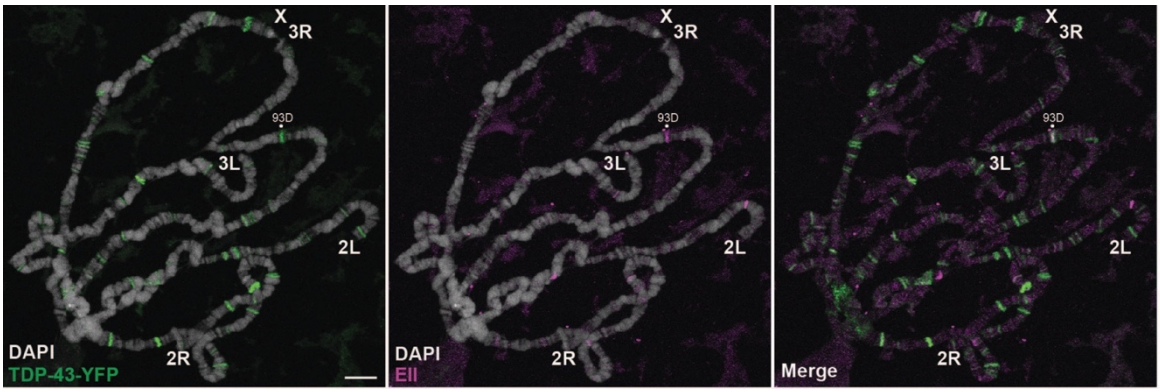
For (**f**) and (**g**), RU486 (4 mg/ml) was used to induce the expression of TDP-43 and *Ell* RNAi. EtOH was used as vehicle. 100 flies were measured 3 times for each genotype at different time points. Bars represent mean (SD). (*) P <0.05, (**) P <0.01, (***) P <0.001, (****) P<0.0001, (NS) not significant (Two-way ANOVA followed by Tukey's multiple comparison test). Significant differences are only indicated within the same time point. (**h**) (**I**) Lifespan analysis of flies expressing TDP-43 and/or *Ell* RNAi by neuronal driver *elavGS* (**h**) or ubiquitous driver *daGS* (**I**) show that Ell depletion mildly prolongs the lifespan on its own and mildly extends the shortened lifespan mediated by TDP-43 expression. Genotypes are as indicated in Fig. 1f and Fig. 2a. n=200 flies per group. P values are indicated in the figure (Log-rank test and Gehan-Breslow-Wilcoxon test).


a
(8 d)
daGS>TDP-43
+RU486
Ell RNAi
kDa
TDP-43
50–
Tubulin
50–
NS
Relative expression of TDP-43
daGS>TDP-43
+RU486
b
U2
U4atac
U5
U11
daGS>TDP-43
Ell RNAi
RU486
2S
Relative expression level of U2
Relative expression level of U4atac
Relative expression level of U5
Relative expression level of U11
c
Relative expression of U12
daGS>U12 RNAi
d
gmr>TDP-43
Ctrl
U12 RNAi
e
Control
gmr>U12 RNAi
f
CG6323
CG7892
CG8408
CG11984
CG18177
CG13431
CG4894
CG7736
CG15081
CG17228
CG17912
CG32705
Relative expression
RU486
Ell RNAi
daGS>TDP-43



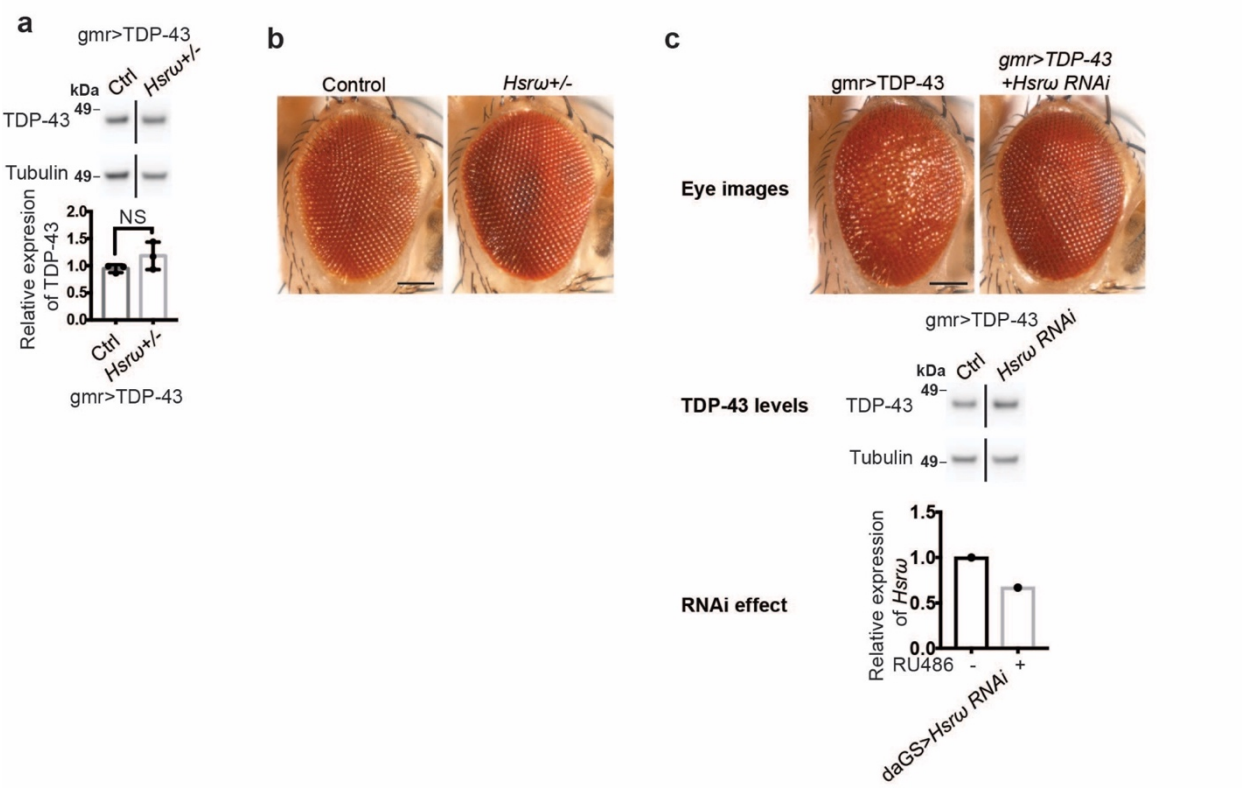
Supplementary Figure 2. Expression of TDP-43 affects the levels of selected snRNAs and U12 intron-containing genes in fly heads

(**a**) Western blot analysis shows that downregulation of *Ell* does not alter TDP-43 protein levels driven by ubiquitous drug-inducible diver *daGS* in fly heads. Genotypes of the flies and RU486 treatment are the same as indicated in Fig. 2a, e.
(**b**) Northern blot analysis of Pol II-transcribed snRNAs shows that U2, U4atac, U5 and U11 are not significantly affected by TDP-43 expression in fly heads. 2S rRNA was the internal control. Genotypes of the flies and RU486 treatment are the same as indicated in Fig. 2a, e.
(**c**) RT-qPCR analysis verifies that *U12* is knocked down by RNAi. RU486 (4 mg/ml) was used to induce the expression of RNAi for 8d, and whole flies were used. Relative mRNA levels were normalized to RpL32 mRNA. Genotype: daGS>*U12* RNAi is *daGS-GAL4/UAS-snRNA:U12:73B.RNAi*^{HMC03841}.
(**d**) Western blot analysis confirms that the levels of TDP-43 protein are not altered by *U12* RNAi. Genotypes of the flies are the same as indicated in Fig. 2b.
(**e**) Knockdown of *U12* on its own does not affect eye integrity. Scale bars: 100 μm. Genotype: Control is *gmr-GAL4(YH3)/+* and gmr>*U12* RNAi is *UAS-snRNA:U12:73B.RNAi*^{HMC03841}*/+; gmr-GAL4(YH3)/+*.
(**f**) RT-qPCR analysis of U12 intron-containing genes shows that 12 out of 18 genes are not altered by TDP-43 expression in fly heads. Relative mRNA levels were normalized to RpL32 mRNA. Genotypes of the flies and RU486 treatment are the same as indicated in Fig. 2a, e.
For (**a**-**d**) and (**f**), *n*=3 biological replicates. Bars represent mean (SD). (*) P <0.05, (**) P <0.01, (***) P <0.001, (NS) not significant (Two-tailed unpaired Student *t* test for comparisons between 2 groups; One-way ANOVA followed by Tukey's multiple comparison test for comparisons among 3 groups).
(**g**) RT-qPCR analysis of 18 genes regulated by U12-dependent spliceosome shows that both the spliced and unspliced products of *CG33108*, *CG11328*, *CG3294* are up-regulated by TDP-43 expression driven by *daGS* in fly heads, and downregulation of *Ell* rescues the increased levels. RU486 (4 mg/ml) was used to induce expression for 8d. mRNA levels were normalized to RpL32 mRNA. Statistical analysis and genotypes are as indicated in Fig. 2a,e.



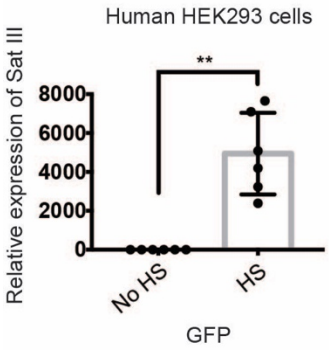
Supplementary Figure 3. TDP-43 and Ell co-localize at the 93D locus on polytene chromosomes

Immunostainings of DAPI (white), TDP-43-YFP (green) and Ell (magenta) show partial overlap and co-localization of TDP-43-YFP and Ell on *Hsrω* locus in salivary glands expressing TDP-43-YFP. Scale bars: 10 µm. Genotype is as indicated in Fig. 3a, b.



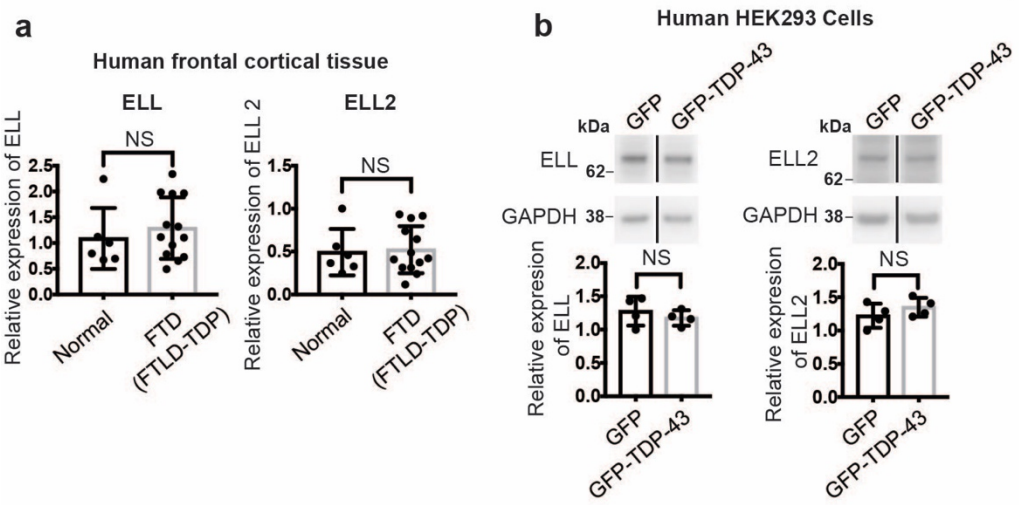
Supplementary Figure 4. Controls for the specificity of Hsrω on TDP-43 toxicity

(**a**) Western blot analysis confirms that the levels of TDP-43 are not altered by downregulation of *Hsrω*. *n*=3 biological replicates. Bars represent mean (SD). (NS) not significant (Two-tailed unpaired Student *t* test). Genotypes are as indicated in Fig. 4b. (**b**) Loss of one copy of *Hsrω* does not disrupt eye integrity. Scale bars: 100 μm. Genotype: Control is *gmr-GAL4(YH3)/+* and *Hsrω+/-* is *gmr-GAL4(YH3)/Hsrω*^{66}. (**c**) Depletion of *Hsrω* by an RNAi fly line suppresses the eye degeneration caused by TDP-43 expression. The levels of TDP-43 and RNAi effect were confirmed by western blotting and RT-qPCR. For RNAi effect confirmation, RU486 (4 mg/ml) was used to induce the expression of RNAi driven by *daGS* promoter for 8 days, and whole flies were used. Genotypes: gmr>TDP-43 is *UAS-TDP-43/+; gmr-GAL4(YH3)/+*. gmr>TDP-43+*Hsrω* RNAi is *UAS-TDP-43/UAS-Hsrω.RNAi*^{HMC05093}*; gmr-GAL4(YH3)/+.* daGS>*Hsrω* RNAi is *daGS-GAL4/UAS-Hsrω.RNAi*^{HMC05093}.



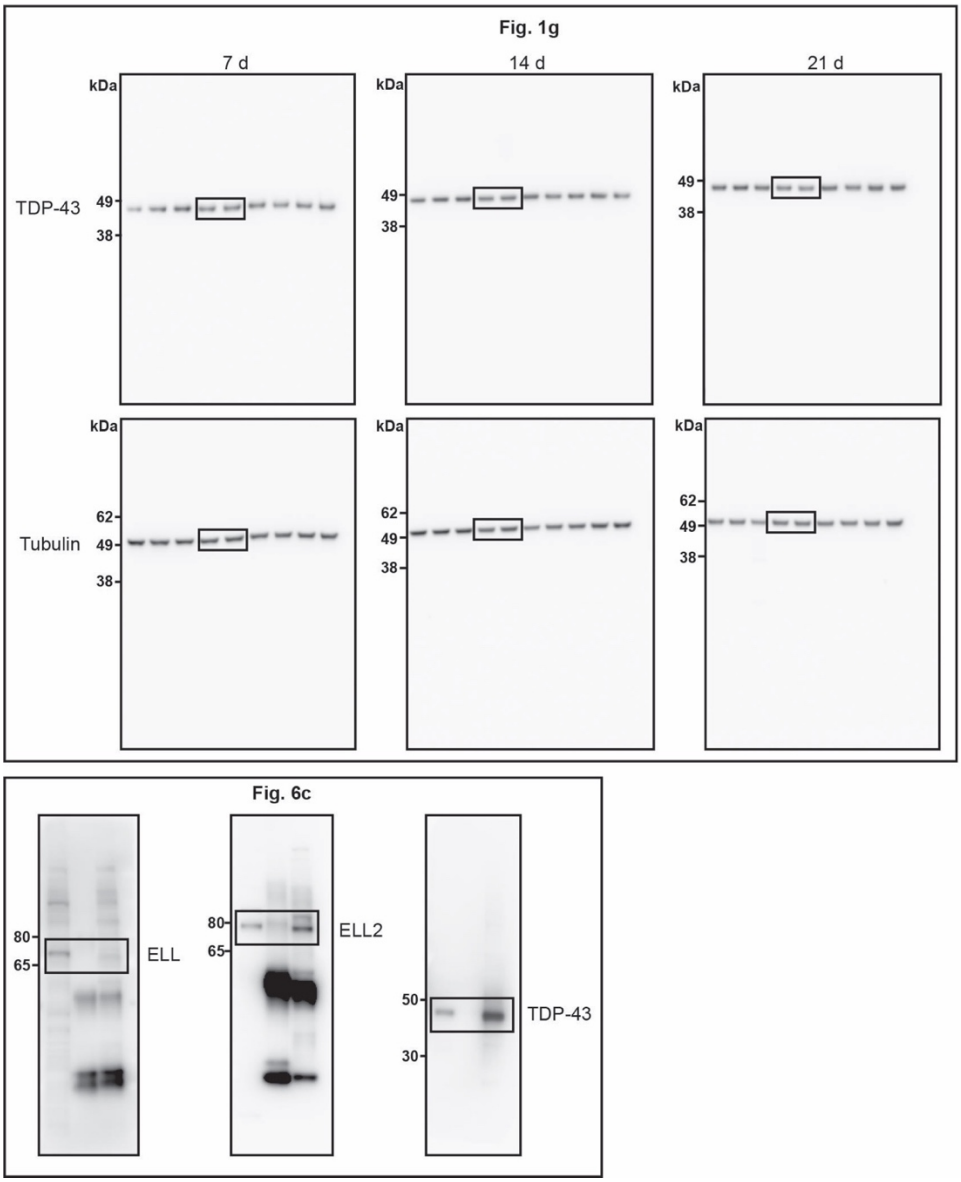
Supplementary Figure 5. Sat III qPCR primers detect a large-fold induction by heat stress

RT-qPCR analysis shows that the levels of Sat III are dramatically induced by heat stress in HEK293 cells expressing GFP after 6d induction. Relative RNA levels were normalized to *GAPDH* and *ACTB* mRNAs (geometric mean). n=6 replicates. Bars represent mean (SD). (**) *P<0.01* (Two-tailed unpaired Student's *t* test with Welch's correction).

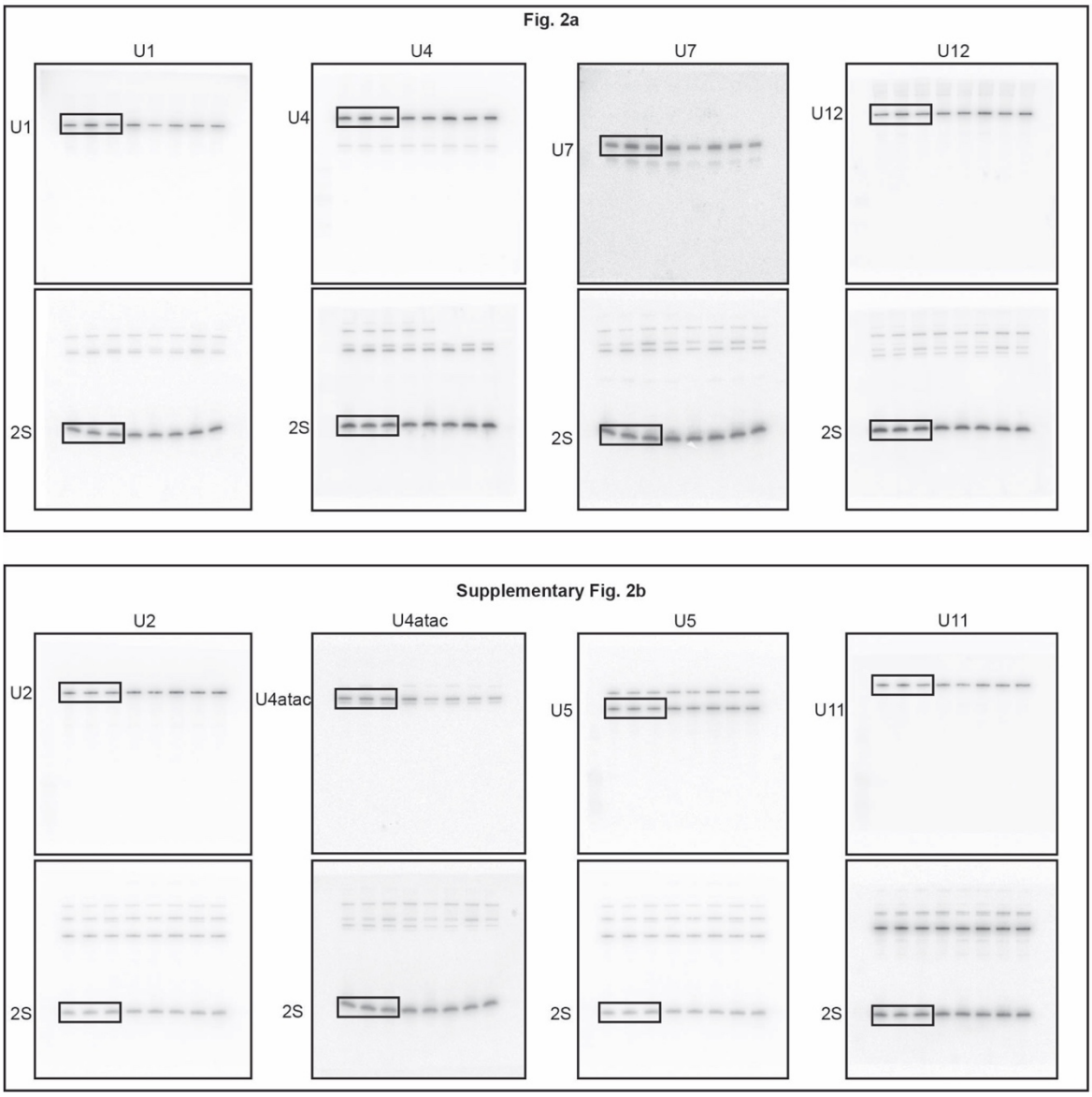


Supplementary Figure 6. RNA levels and protein levels of ELL and ELL2 are not altered in patient samples and a HEK293 cell disease model respectively

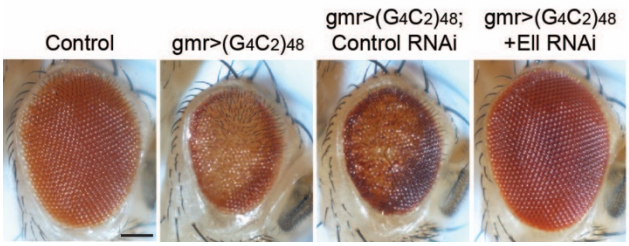
(**a**) RT-qPCR analysis shows that the levels of *ELL* and *ELL2* are increased significantly in frontal cortex of FTLD-TDP compared to normal frontal cortex controls. Bars represent mean (SD). (NS) not significant (Two-tailed unpaired Student's *t* test and Two-tailed unpaired Mann-Whitney test for ELL; Two-tailed unpaired Student's *t* test for ELL2). Sample details are as indicated in Fig. 6b.
(**b**) Western blot analysis of ELL and ELL2 show that protein levels of ELL and ELL2 are not altered in HEK293 cells expressing GFP-TDP-43 compared to cells expressing GFP after 6d induction. GAPDH was used as internal control. *n*=4 biological replicates. Bars represent mean (SD). (NS) not significant (Two-tailed unpaired Student *t* test).



Supplementary Figure 7. The uncropped scans of western blots in Fig. 1g and Fig. 6c



Supplementary Figure 8. The uncropped scans of small RNA Northern blots



Supplementary Figure 9. Depletion of *Ell* suppresses the eye degeneration caused by G_4C_2 expansion

Expression of GGGGCC expansion causes eye degeneration, which is suppressed by *Ell* RNAi. Scale bars: 100 μm. Genotypes: Control is *gmr-GAL4(YH3)/+*. gmr>$(G_4C_2)_{48}$ is *UAS-*$(G_4C_2)_{48}$*/+; gmr-GAL4(YH3)/+*. gmr>$(G_4C_2)_{48}$+Control RNAi is *UAS-*$(G_4C_2)_{48}$*/+; gmr-GAL4(YH3)/UAS-Control.RNAi*^{JF01355}. gmr>$(G_4C_2)_{48}$+*Ell* RNAi is *UAS-*$(G_4C_2)_{48}$*/+; gmr-GAL4(YH3)/UAS-Ell.RNAi*^{HMS00277}.

Supplementary Data

File Name: Supplementary Data 1

Description: The lists of genes bound by TBPH, Lilli or both, derived from [69,114]

(The file will be able to be downloaded from https://www.nature.com/ncomms/)

CHAPTER 3: CONCLUSION AND FUTURE DIRECTIONS

CONCLUSION OF THE FINDINGS AND OPEN QUESTIONS

Conclusions

Our work defined Ell, the shared subunit of LEC and SEC, as a strong dose-dependent modifier of TDP-43-mediated degeneration by using a powerful *Drosophila* disease model expressing human TDP-43 in the fly eye system (Figure 1a-c). The identification of Ell was through a fly genetic screen for modifiers of TDP-43-mediated eye degeneration. Other potential modifiers of TDP-43 toxicity indicated by the screen are listed in Table 3.1 (derived from Ross Weber's data). Depletion of the components specific for LEC or SEC partially suppressed the eye degeneration, although the suppression is not as strong as downregulation of the shared component Ell (Figure 1d,e). The ability to suppress TDP-43-mediated neurodegeneration was further demonstrated in the fly nervous system (Figure 1f). These findings reveal the important roles of Ell and Ell-containing complexes in TDP-43-mediated neurodegeneration, and implicate misregulation of LEC and SEC in mechanisms of disease.

We sought evidence indicating that the activities of LEC and SEC were upregulated. By Northern blot analysis, several LEC target snRNAs, including U1, U4, U7 and U12 snRNAs, were upregulated in fly heads with TDP-43 expression compared to controls (Figure 2-2a). Among these, depletion of U12 snRNA functionally ameliorated eye degeneration caused by TDP-43, indicating that the elevation of U12 contributes to TDP-43-mediated degeneration (Figure 2b,c). Further examination of the levels of 18 genes containing an intron regulated by the U12-dependent spliceosome showed that the spliced products of 3 genes, *CG15735, CG16941 and CG11839*, are upregulated in fly heads with expression of TDP-43, with the levels of the unspliced products unchanged (Figure 2e). These results suggest that the activity of LEC might be upregulated, leading

Table 3-1. Additional potential modifiers of TDP-43 toxicity

Gene Symbol	Gene (CG #)	Bloomington Stock #	Effect
E(bx)	CG32346	27471	Enhancer
pcm	CG3291	33263	Enhancer
SREBP	CG8522	27232	Suppressor
CG3542	CG3542	17291	Suppressor
CG7220	CG7220	29681	Suppressor
CG44837	CG44837	17283	Suppressor
CG16711	CG16711	27182	Suppressor
CG17187	CG17187	30066	Suppressor
Mnt; Parg	CG13316; CG2864	10086	Suppressor
CG5009	CG5009	[148] (P{EP}CG5009B227.2)	Suppressor
Smyd5	CG3353	30047	Suppressor
bin3	CG8276	22584	Suppressor
RpS15	CG8332	27125	Suppressor
Ac76E	CG7978	17485	Suppressor
Top3α	CG10123	17238	Suppressor
Vha68-2	CG3762	17243	Suppressor
Vamp7	CG1599	28488	Suppressor
ppk5; Rpn10	CG33289; CG7619	27212	Suppressor
CG14182	CG14182	27134	Suppressor
CG6923	CG6923	30077	Suppressor
Hph	CG44015	20142	Lethal
fus	CG8205	27457	Lethal
CG11110	CG11110	27466	Lethal
CG7530	CG7530	15762	Lethal
CG5789	CG5789	17083	Lethal
CG15771; lin-52	CG15771; CG15929	17004	Lethal
S6KL; CG6961	CG7001; CG6961	10144	Lethal
CG6294; CG6299	CG6294; CG6299	33456	Lethal
RpI12; tsl	CG13418; CG6705	32613	Lethal

(The table is derived from Ross Weber's data)

to downstream abnormally elevated activity of the U12-dependent spliceosome, contributing to TDP-43 toxicity.

A different approach was undertaken to define important targets of SEC. By *Drosophila* polytene chromosome immunostaining under ambient and heat shock treatment, TDP-43 and Lilli (the essential component of SEC) were discovered to colocalize on the chromosomal locus of a stress-induced lncRNA *Hsrω*, suggesting that *Hsrω* is a shared target of TDP-43 and SEC (Figure 3a,b). Analyses of published ChIP-seq databases of fly TBPH and Lilli [69,114] also showed that *Hsrω* is a gene bound by both factors (Figure 3c,d). The levels of *Hsrω* are increased in the fly disease model and *Hsrω* was shown to functionally contribute to TDP-43-mediated degeneration (Figure 4a-c). These data show that TDP-43 co-localizes with central SEC components on the chromosomes, identifying a novel non-coding lncRNA target of TDP-43 with functional importance to degeneration.

Our results suggest a model in which SEC and LEC contribute to TDP-43-meidated degeneration in parallel. By comparing the retinal thickness in the fly disease model, downregulation of the shared component Ell restored eye degeneration caused by TDP-43 and caused a 4-fold increase in retinal thickness that was comparable to normal retinal thickness. By contrast, depletion of LEC- and SEC-specific components led to only a ~2-fold increase in retinal thickness, and depletion of downstream targets U12 and *Hsrω* led to only ~1.5-fold increase in retinal thickness. Comparison of the retinal thickness is shown in Figure. 3.1.

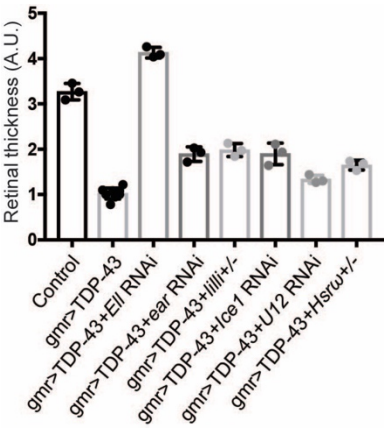


Figure 3-1. Comparison of suppression effects of downregulation of different factors Analysis of retinal thickness shows that depletion of the shared component Ell suppresses TDP-43-mediated eye degeneration more strongly than downregulation of LEC- or SEC-specific components, or downstream targets.

The relevance of this stress-induced lncRNA in human disease was further assessed using a human cell disease model and patient tissue. The levels of the human counterpart of *Hsrω*, Sat III, in HEK cells expressing GFP-tagged TDP-43 was upregulated ~4-fold compared to control (Figure 5a). Furthermore, the levels of Sat III were increased significantly in frontal cortical tissues from FTD patients with TDP-43 pathology (FTLD-TDP) compared to tissue from normal individuals (Figure 5b). Since upregulation of Sat III has been shown to promote cell death and acute senescence in different cell models [119,120], these results suggest that stress-induced non-coding RNA Sat III may be a contributor to TDP-43-mediated degeneration.

Given that ELL is the shared element of LEC and SEC, we sought to define how disruption of ELL function may be induced by TDP-43. We found that TDP-43 physically interacts one of the ELL orthologues in HEK293 cells by co-immunoprecipitation. We were able to test ELL and ELL2, and found evidence that TDP-43 interacts in lysates with human ELL2 (Figure 6c). Together with the immunostaining data showing that TDP-43 co-localizes with SEC on polytene chromosomes, these results indicate that a

mechanism contributing to the promotion of LEC and SEC activity may include interaction between TDP-43 and the proteins of the complexes.

These findings reveal novel roles of the transcription elongation factor Ell-containing complexes, highlight the importance of non-coding RNAs in disease, elucidate new mechanistic insight into neurodegenerative disorders and provide therapeutic targets for TDP-43-associated neurodegeneration.

Future directions

Targets of LEC and SEC

In the investigation of the targets of LEC and SEC, we used information from the literature to define potential important targets and used different assays to examine the role of these potential targets. Through these approaches, we identified important targets, including U12 snRNA and *Hsrω*, that we showed are functionally involved in TDP-43-mediated degeneration (Figure 2, 3 and 4). The relationship and interaction between U12 and *Hsrω* in suppressing TDP-43 toxicity can be further dissected. One interesting question is whether downregulation of one of these targets can affect the levels of the other. This can be addressed by assessing the levels of U12 and *Hsrω* in fly head tissue upon downregulation of U12 or *Hsrω*.

In addition to the identified targets U12 and *Hsrω* , there may be other targets regulated by LEC and SEC that are critical in disease. In addition, in previous studies, targets of LEC and SEC were identified and investigated in non-neuronal cells, including *Drosophila* S2 cells [69], mouse embryonic cells[73], human HCT-116 cells [72] and human

HEK cells [74]. It is therefore unknown the extent to which the targets of LEC and SEC are the same in the nervous system. Given the suppression effect of LEC and SEC downregulation on TDP-43-mediated neurodegeneration (Figure 1), identifying the protein-coding and non-coding RNA targets of LEC and SEC in neurons, as well as revealing the targets important in neurodegenerative situations may help us understand the roles of LEC and SEC in the nervous system and in TDP-43 toxicity more thoroughly.

To define the global targets of LEC and SEC in the nervous system and identify important genes in TDP-43 toxicity, I propose to use ChIP-seq complemented with RNA-seq to identify genes bound by different factors and detect the transcriptomic changes under different conditions. ChIP-seq can be conducted using antibodies against Ell (the shared component), Ice1 (specific to LEC) and Lilli (specific to SEC) and TDP-43, in fly brains/heads with or without TDP-43 expression. Targets of LEC are predicted to be those bound by Ell and Ice1, while targets of SEC are predicted to be those bound by Ell and Lilli. Comparisons of LEC or SEC targets in TDP-43 expressing animals to controls are predicted to elucidate whether LEC or SEC bind to different sets of genes in the fly disease model compared to the normal condition.

RNA-seq can be used to analyze and compare the transcriptome in fly heads or brains with the combination of TDP-43 expression, depletion of Ell (the shared component), Ice1 (specific to LEC) and Lilli (specific to SEC) (Table 3-2). By comparing Group 1 and 2, gene alterations caused by TDP-43 expression can be identified. This analysis would also improve our understanding of this fly disease model. The LEC-regulated genes in fly brain are predicted to be downregulated in both Group 3 and 5, compared to Group 1.

The SEC-regulated genes will be those downregulated in Group 3 and 7 compared to Group 1. Important targets of LEC in disease are predicted to be those upregulated in Group 2 and rescued in Group 4 and 6. Likewise, potential important targets of SEC are those upregulated in Group 2 and decreased back toward normal in Group 4 and 8. Defined target genes then can be further analyzed for confirmation and functional importance using fly and human cell disease models. Alternatively, to reduce the complexity and numbers of samples, RNA-seq could be done focusing only the combinations of TDP-43 expression and Ell downregulation. The involvement of Ice1 and Lilli could be assessed by ChIP-qPCR later once specific genes are identified. The best time point and conditions for sample collection would require determination by pilot studies using target genes identified in this thesis work, such as U12 and *Hsrω*.

Table 3-2. Experimental design for identification of global LEC and SEC targets

Group	1	2	3	4	5	6	7	8
TDP-43	-	+	-	+	-	+	-	
Ell RNAi	-	-	+	+	-	-	-	-
Ice1 RNAi	-	-	-	-	+	+	-	-
Lilli RNAi	-	-	-	-	-	-	+	+

Does Ell, U12 or *Hsrω* affect stress sensitivity?

TDP-43 expressing flies show hypersensitivity to heat stress, paraquat-induced oxidative stress and starvation [103]. In my thesis, we show that the heat shock-induced lncRNA *Hsrω* is aberrantly upregulated in fly heads with TDP-43 expression. These data suggest that the stress response may be misregulated in these animals. Environmental stress can be broadly defined as environmental changes that threaten or disrupt physiological homeostasis and reduce the performance or fitness of the organism [149,150]. The effect of stress on an organism depends on the type of stressor, the stress intensity and duration,

and different situations may trigger different stress responses [151]. Therefore, it is likely that TDP-43 expressing animals may react differently according to the specific stress situation [151,152].

A question we have is whether the suppression effect of depletion of Ell, U12 or *Hsrω* on TDP-43-mediated degeneration is through modulating the stress response. To test this hypothesis, one can start from assessing sensitivity of flies to a variety of stress situations with TDP-43 expression combined with downregulation of Ell, U12 and *Hsrω*. The prediction is that expression of TDP-43 may lead to altered sensitivity to different stressors and stress situations, and depletion of Ell, U12 and may rescue the affected stress sensitivity.

Is transcriptional pausing regulation affected?

Misregulation of transcriptional elongation has been associated with different diseases, including cancer, virus infection and developmental diseases [79,82,123]. Yet very little was known about its involvement in neurodegenerative diseases prior to this work. Our study suggests a model of abnormal activation of LEC and SEC in TDP-43-mediated neurodegeneration (Figure 2-7). This predicts that transcription pausing regulated by LEC and SEC is misregulated. This idea is supported by our data showing targets of LEC and SEC are upregulated in the disease models (Figure 2-2 and 2-4). However, to further strengthen and prove this hypothesis, direct evidence of dysfunction of transcriptional pausing regulation is required.

The regulation of transcription elongation in TDP-43-mediated degeneration can be assessed for specific targets or investigated globally by different assays in the disease models used in this study. The stress-induced lncRNA *Hsrω* regulated by SEC is defined as a novel target of TDP-43 (Figure 2-3 and 2-4), and this specific gene can be an interesting target for accessing pausing regulation. I propose to use ChIP-qPCR with antibody against total Pol II and qPCR primers to detect Pol II occupancy on *Hsrω* gene at either the transcriptional start site (~0), pausing region (~+30) or intergenic regions as a control. A second method, permanganate ($KMnO_4$) footprinting, can also be used to assess pausing regulation on *Hsrω*. Paused and transcriptionally engaged Pol II leads to single-stranded DNA regions, where potassium permanganate can oxidize thymines. The oxidized thymines can then be cleaved, and the positions of DNA breaks can be determined by ligation-mediated PCR (LM-PCR) using primers specific to *Hsrω*.

The genome-wide identification of pausing regulation in TDP-43 disease models can be assessed by precision nuclear run-on sequencing (PRO-seq) [153] by using 1) nuclei isolated from fly heads with or without TDP-43 expression 2) nuclei isolated from HEK cells with GFP-TDP-43 expression or GFP expression. In PRO-seq, isolated nuclei are treated with Sarkosyl, which prevents new incorporation of Pol II into the chromosomes, and used for identification of sites with transcriptional-engaged Pol II by extending nascent RNAs with biotin-labeled ribonucleotide triphosphate analogs (biotin-NTP). By comparing the Pol II active sites in the fly and human cells with TDP-43 expression to controls, any alterations to transcriptional pausing regulation on coding and non-coding genes can be revealed.

Ell as a universal therapeutic target in neurodegenerative diseases?

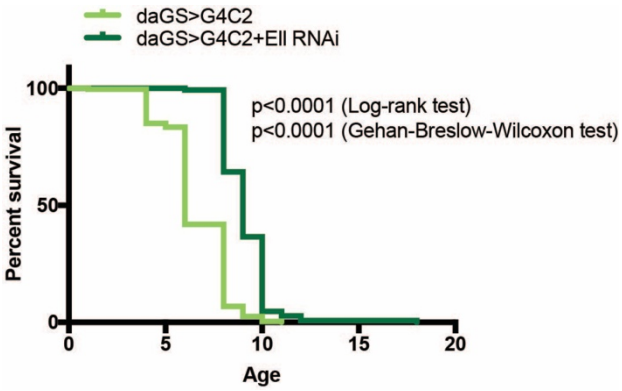


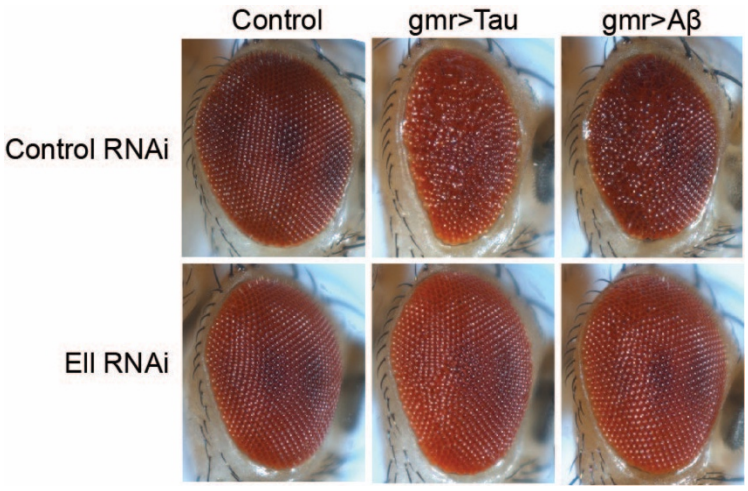
Figure 3-2. Ell depletion suppresses the toxicity of G_4C_2 expansion
Lifespan assay show that Ell depletion extends the lifespan of flies with G_4C_2 repeat expression.

In ALS and FTD, another important disease factor is the GGGGCC (G_4C_2) hexanucleotide repeat expansion. Over the last 7 years, a G_4C_2 hexanucleotide repeat expansion in the 5' region of *C9ORF72* has been identified as the most common genetic mutation in ALS and FTD [130,131]. Expression of transgenes bearing a G_4C_2 hexanucleotide repeat expansion in the fly causes neurodegeneration [147,154-156]. In our study, Ell was identified as a strong dose-dependent modifier of TDP-43 toxicity (Figure 2-1). Given that ALS and FTD are within the same disease spectrum and share molecular features, we tested the effect of Ell downregulation on fly eye degeneration caused by expression of G_4C_2 hexanucleotide repeat expansion. The results show that Ell depletion also markedly ameliorates toxicity of G_4C_2 hexanucleotide repeat expansion (Figure 2-S9). In additional preliminary results, the suppression effect of Ell downregulation on G_4C_2 hexanucleotide repeat expansion toxicity in a lifespan assay was also observed, by using the drug-inducible ubiquitous daGS promoter (Figure 3-2). Although a modest effect, the impact on lifespan is rigorous when compared with controls (Figure 3-2). These data suggest that shared targets of Ell might be critical to degeneration associated with these two disease factors.

Moreover, my data indicate that different neurodegenerative diseases beyond ALS and FTD might share similar pathways. Neurodegenerative diseases accord common

pathogenetic mechanisms involving aggregation of misfolded proteins, as with Tau and Amyloid-beta (Aβ) in Alzheimer's disease, α-Synuclein in Parkinson's disease, Huntingtin in Huntington and TAR DNA binding protein 43 (TDP-43) in ALS and FTLD-TDP [157,158]. Although the type of disease proteins and the cellular and regional distribution of aggregates are varied among the diseases, similar biological and genetic pathways may link these disorders. To assess the involvement of Ell in other neurodegenerative diseases associated with additional toxic proteins, the effect of Ell downregulation was examined in fly disease models expressing Tau and Aβ in the fly eye. Expression of Tau and Aβ driven by the eye specific driver gmr led to eye degeneration that was be suppressed by Ell depletion (Figure 3-3).

Figure 3-3. Depletion of Ell suppressed the toxicity of Tau and Aβ
External eye images show that expression of Tau and Aβ cause eye degeneration, which are ameliorated by Ell downregulation.



These preliminary data suggest that Ell might be a central factor among a range of different toxicities associated with different neurodegenerative proteins. To identify the common and distinct coding and noncoding targets of Ell in different disease models may provide effective therapeutic insight. Before further addressing the potential central

role of Ell in neurodegeneration, one should clarify whether the suppression effects of Ell are through decreasing the levels of the disease factors. The reason for this is that Ell might play different roles in mechanisms associated with different disease factors. For example, depletion of a transcriptional elongation factor Spt4 has been shown to decrease the transcripts of expanded G_4C_2 repeats [159]. Potentially, downregulation of Ell, which is also involved in the regulation of elongation, could also affect the production of G_4C_2 repeat transcripts and therefore suppress the toxicity associated with the repeat expansion.

Next, to identify the important targets of Ell in different disease models, I propose to test the toxicity of disease factors and the suppression effect of Ell carefully by the drug-inducible ubiquitous driver daGS using lifespan and other assays. RNA-seq can then be used to compare the transcriptomes in fly heads/brains with the combination of expression of different disease factors and Ell RNAi (Table 3-3). Transcriptome changes in different disease models can be identified by comparing Group 2, 5 or 7 to Group 1. Important Ell targets that are common in all disease models are predicted to be upregulated in Group 2, 5, and 7 and downregulated in Group 4, 6 and 8. Furthermore, critical Ell targets that are specific for each disease factor can also be identified by comparing changes among the different groups. To reduce complexity, it may be impactful to start from TDP-43 and G_4C_2 hexanucleotide repeat expansion, which are two distinct factors but important for the same disease spectrum of ALS/FTD.

Table 3-3. Experimental design for identification of common and distinct Ell targets

Group	1	2	3	4	5	6	7	8
TDP-43	-	+	-	+	-	-	-	-
G_4C_2	-	-	-	-	+	+	-	-
Tau or Aβ	-	-	-	-	-	-	+	+
Ell RNAi	-	-	+	+	-	+	-	+

Can Sat III affect TDP-43 LLPS or aggregation?

Our work identified the *Drosophila* stress-induced lncRNA *Hsrω*, which is upregulated in a fly disease model, as a novel modifier of TDP-43-mediated degeneration (Figure 2-3, 2-4). Furthermore, human orthologue Sat III repeats are abnormally elevated in a human cellular model and frontal cortical tissue from FTD patients with TDP-43 pathology (Figure 2-5). One question is how elevated Sat III may contribute to TDP-43-mediated toxicity. TDP-43, containing a low complexity domain, forms aggregates in neurons and glia of ALS and FTLD-TDP patients [9,10], and purified TDP-43 can undergo liquid liquid phase separation (LLPS) and form aggregates in vitro [160,161]. Many of the disease-associated TDP-43 mutations promote TDP-43 aggregation in disease models, suggesting formation of inclusions may be causal to the TDP-43-associated dysfunction[162,163].

Distinct from other stress genes, the Sat III RNAs form a transient subnuclear organelle, the nuclear stress body (nSB), which is associated with the site of Sat III DNA; the comparable structure in *Drosophila* is the omega speckle [106,108,109]. Our data show that TDP-43 localizes to the *Hsrω* locus (Figure 2-3), suggesting that TDP-43 may also bind the ω speckle. In fact, TDP-43 has been reported to localize to nSB under heat shock in HeLa cells [164], suggesting that Sat III may interact with TDP-43. Thus, it may be that highly repetitive Sat III RNAs increase the local concentration of various RNA-binding

proteins, including TDP-43, and promote LLPS to form nSB. These data lead us to hypothesize that there might exist a feedback loop between TDP-43 and Sat III: in disease, TDP-43 expression leads to the elevation of Sat III levels, which in turn contributes to TDP-43-mediated toxicity by promoting TDP-43 LLPS and/or aggregation, potentially in the nSB.

LLPS has been proposed to be a dynamic process that gives rise to membrane-less organelles [165,166], and RNAs have been shown to regulate LLPS [167-169]. As one test of this hypothesis, Sat III generated by in vitro transcription can be added to TDP-43, and LLPS and aggregation can be measured using different concentrations of RNA and protein to assess whether Sat III can alter TDP-43 droplet formation or aggregation. Sat III by itself can also be tested for LLPS. These in vitro assays can test the hypothesis and might provide more mechanistic insights for the findings from the in vivo studies.

Conclusion

Our studies have defined a critical role for ELL in toxicity of TDP-43. Importantly, the ELL regulated pathway may be broad, and the future directions proposed seek to define those pathways in greater detail, as well as define the specific mechanisms of the pathways revealed in this thesis with greater clarity.

BIBLIOGRAPHY

1 Bennion Callister, J. & Pickering-Brown, S. M. Pathogenesis/genetics of frontotemporal dementia and how it relates to ALS. *Exp Neurol* **262 Pt B**, 84-90, doi:10.1016/j.expneurol.2014.06.001 (2014).

2 Lattante, S., Ciura, S., Rouleau, G. A. & Kabashi, E. Defining the genetic connection linking amyotrophic lateral sclerosis (ALS) with frontotemporal dementia (FTD). *Trends Genet* **31**, 263-273, doi:10.1016/j.tig.2015.03.005 (2015).

3 Ling, S. C., Polymenidou, M. & Cleveland, D. W. Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron* **79**, 416-438, doi:10.1016/j.neuron.2013.07.033 (2013).

4 Weishaupt, J. H., Hyman, T. & Dikic, I. Common Molecular Pathways in Amyotrophic Lateral Sclerosis and Frontotemporal Dementia. *Trends Mol Med* **22**, 769-783, doi:10.1016/j.molmed.2016.07.005 (2016).

5 Wheaton, M. W. *et al.* Cognitive impairment in familial ALS. *Neurology* **69**, 1411-1417, doi:10.1212/01.wnl.0000277422.11236.2c (2007).

6 Ringholz, G. M. *et al.* Prevalence and patterns of cognitive impairment in sporadic ALS. *Neurology* **65**, 586-590, doi:10.1212/01.wnl.0000172911.39167.b6 (2005).

7 Burrell, J. R., Kiernan, M. C., Vucic, S. & Hodges, J. R. Motor neuron dysfunction in frontotemporal dementia. *Brain* **134**, 2582-2594, doi:10.1093/brain/awr195 (2011).

8 Fecto, F. & Siddique, T. Making connections: pathology and genetics link amyotrophic lateral sclerosis with frontotemporal lobe dementia. *J Mol Neurosci* **45**, 663-675, doi:10.1007/s12031-011-9637-9 (2011).

9 Arai, T. *et al.* TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem Biophys Res Commun* **351**, 602-611, doi:10.1016/j.bbrc.2006.10.093 (2006).

10 Neumann, M. *et al.* Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130-133, doi:10.1126/science.1134108 (2006).

11 Ou, S. H., Wu, F., Harrich, D., Garcia-Martinez, L. F. & Gaynor, R. B. Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. *J Virol* **69**, 3584-3596 (1995).

12 Chen-Plotkin, A. S., Lee, V. M. & Trojanowski, J. Q. TAR DNA-binding protein 43 in neurodegenerative disease. *Nat Rev Neurol* **6**, 211-220, doi:10.1038/nrneurol.2010.18 (2010).

13 Lee, E. B., Lee, V. M. & Trojanowski, J. Q. Gains or losses: molecular mechanisms of TDP43-mediated neurodegeneration. *Nat Rev Neurosci* **13**, 38-50, doi:10.1038/nrn3121 (2011).

14 Sun, Y. & Chakrabartty, A. Phase to Phase with TDP-43. *Biochemistry* **56**, 809-823, doi:10.1021/acs.biochem.6b01088 (2017).

15 Ratti, A. & Buratti, E. Physiological functions and pathobiology of TDP-43 and FUS/TLS proteins. *J Neurochem* **138 Suppl 1**, 95-111, doi:10.1111/jnc.13625 (2016).

16 Igaz, L. M. *et al.* Dysregulation of the ALS-associated gene TDP-43 leads to neuronal death and degeneration in mice. *J Clin Invest* **121**, 726-738, doi:10.1172/JCI44867 (2011).

17 Liu, Y. C., Chiang, P. M. & Tsai, K. J. Disease animal models of TDP-43 proteinopathy and their pre-clinical applications. *Int J Mol Sci* **14**, 20079-20111, doi:10.3390/ijms141020079 (2013).
18 Rubin, G. M. *et al.* Comparative genomics of the eukaryotes. *Science* **287**, 2204-2215 (2000).
19 Reiter, L. T., Potocki, L., Chien, S., Gribskov, M. & Bier, E. A systematic analysis of human disease-associated gene sequences in Drosophila melanogaster. *Genome Res* **11**, 1114-1125, doi:10.1101/gr.169101 (2001).
20 McGuire, S. E., Roman, G. & Davis, R. L. Gene expression systems in Drosophila: a synthesis of time and space. *Trends Genet* **20**, 384-391, doi:10.1016/j.tig.2004.06.012 (2004).
21 del Valle Rodriguez, A., Didiano, D. & Desplan, C. Power tools for gene expression and clonal analysis in Drosophila. *Nat Methods* **9**, 47-55, doi:10.1038/nmeth.1800 (2011).
22 Brand, A. H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401-415 (1993).
23 Duffy, J. B. GAL4 system in Drosophila: a fly geneticist's Swiss army knife. *Genesis* **34**, 1-15, doi:10.1002/gene.10150 (2002).
24 McGuire, S. E., Le, P. T., Osborn, A. J., Matsumoto, K. & Davis, R. L. Spatiotemporal rescue of memory dysfunction in Drosophila. *Science* **302**, 1765-1768, doi:10.1126/science.1089035 (2003).
25 Osterwalder, T., Yoon, K. S., White, B. H. & Keshishian, H. A conditional tissue-specific transgene expression system using inducible GAL4. *Proc Natl Acad Sci U S A* **98**, 12596-12601, doi:10.1073/pnas.221303298 (2001).
26 Lenz, S., Karsten, P., Schulz, J. B. & Voigt, A. Drosophila as a screening tool to study human neurodegenerative diseases. *J Neurochem* **127**, 453-460, doi:10.1111/jnc.12446 (2013).
27 McGurk, L., Berson, A. & Bonini, N. M. Drosophila as an In Vivo Model for Human Neurodegenerative Disease. *Genetics* **201**, 377-402, doi:10.1534/genetics.115.179457 (2015).
28 Feiguin, F. *et al.* Depletion of TDP-43 affects Drosophila motoneurons terminal synapsis and locomotive behavior. *FEBS Lett* **583**, 1586-1592, doi:10.1016/j.febslet.2009.04.019 (2009).
29 Lu, Y., Ferris, J. & Gao, F. B. Frontotemporal dementia and amyotrophic lateral sclerosis-associated disease protein TDP-43 promotes dendritic branching. *Mol Brain* **2**, 30, doi:10.1186/1756-6606-2-30 (2009).
30 Diaper, D. C. *et al.* Loss and gain of Drosophila TDP-43 impair synaptic efficacy and motor control leading to age-related neurodegeneration by loss-of-function phenotypes. *Hum Mol Genet* **22**, 1539-1557, doi:10.1093/hmg/ddt005 (2013).
31 Vanden Broeck, L. *et al.* TDP-43 loss-of-function causes neuronal loss due to defective steroid receptor-mediated gene program switching in Drosophila. *Cell Rep* **3**, 160-172, doi:10.1016/j.celrep.2012.12.014 (2013).
32 Hanson, K. A., Kim, S. H., Wassarman, D. A. & Tibbetts, R. S. Ubiquilin modifies TDP-43 toxicity in a Drosophila model of amyotrophic lateral sclerosis (ALS). *J Biol Chem* **285**, 11068-11072, doi:10.1074/jbc.C109.078527 (2010).
33 Li, Y. *et al.* A Drosophila model for TDP-43 proteinopathy. *Proc Natl Acad Sci U S A* **107**, 3169-3174, doi:10.1073/pnas.0913602107 (2010).

34 Voigt, A. *et al.* TDP-43-mediated neuron loss in vivo requires RNA-binding activity. *PLoS One* **5**, e12247, doi:10.1371/journal.pone.0012247 (2010).
35 Estes, P. S. *et al.* Wild-type and A315T mutant TDP-43 exert differential neurotoxicity in a Drosophila model of ALS. *Hum Mol Genet* **20**, 2308-2321, doi:10.1093/hmg/ddr124 (2011).
36 Choksi, D. K. *et al.* TDP-43 Phosphorylation by casein kinase Iepsilon promotes oligomerization and enhances toxicity in vivo. *Hum Mol Genet* **23**, 1025-1035, doi:10.1093/hmg/ddt498 (2014).
37 Elden, A. C. *et al.* Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature* **466**, 1069-1075, doi:10.1038/nature09320 (2010).
38 Kim, H. J. *et al.* Therapeutic modulation of eIF2alpha phosphorylation rescues TDP-43 toxicity in amyotrophic lateral sclerosis disease models. *Nat Genet* **46**, 152-160, doi:10.1038/ng.2853 (2014).
39 Ayala, Y. M. *et al.* Human, Drosophila, and C.elegans TDP43: nucleic acid binding properties and splicing regulatory function. *J Mol Biol* **348**, 575-588, doi:10.1016/j.jmb.2005.02.038 (2005).
40 Kumar, J. P. Building an ommatidium one cell at a time. *Dev Dyn* **241**, 136-149, doi:10.1002/dvdy.23707 (2012).
41 Adelman, K. & Lis, J. T. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet* **13**, 720-731, doi:10.1038/nrg3293 (2012).
42 Chen, F. X., Smith, E. R. & Shilatifard, A. Born to run: control of transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol* **19**, 464-478, doi:10.1038/s41580-018-0010-5 (2018).
43 Kwak, H. & Lis, J. T. Control of transcriptional elongation. *Annu Rev Genet* **47**, 483-508, doi:10.1146/annurev-genet-110711-155440 (2013).
44 Jonkers, I. & Lis, J. T. Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16**, 167-177, doi:10.1038/nrm3953 (2015).
45 Shilatifard, A., Conaway, R. C. & Conaway, J. W. The RNA polymerase II elongation complex. *Annu Rev Biochem* **72**, 693-715, doi:10.1146/annurev.biochem.72.121801.161551 (2003).
46 Shilatifard, A., Lane, W. S., Jackson, K. W., Conaway, R. C. & Conaway, J. W. An RNA polymerase II elongation factor encoded by the human ELL gene. *Science* **271**, 1873-1876 (1996).
47 Miller, T., Williams, K., Johnstone, R. W. & Shilatifard, A. Identification, cloning, expression, and biochemical characterization of the testis-specific RNA polymerase II elongation factor ELL3. *J Biol Chem* **275**, 32052-32056, doi:10.1074/jbc.M005175200 (2000).
48 Shilatifard, A. *et al.* ELL2, a new member of an ELL family of RNA polymerase II elongation factors. *Proc Natl Acad Sci U S A* **94**, 3639-3643 (1997).
49 Lin, C., Garruss, A. S., Luo, Z., Guo, F. & Shilatifard, A. The RNA Pol II elongation factor Ell3 marks enhancers in ES cells and primes future gene activation. *Cell* **152**, 144-156, doi:10.1016/j.cell.2012.12.015 (2013).
50 Thirman, M. J., Levitan, D. A., Kobayashi, H., Simon, M. C. & Rowley, J. D. Cloning of ELL, a gene that fuses to MLL in a t(11;19)(q23;p13.1) in acute myeloid leukemia. *Proc Natl Acad Sci U S A* **91**, 12110-12114 (1994).

51 Gerber, M., Ma, J., Dean, K., Eissenberg, J. C. & Shilatifard, A. Drosophila ELL is associated with actively elongating RNA polymerase II on transcriptionally active sites in vivo. *EMBO J* **20**, 6104-6114, doi:10.1093/emboj/20.21.6104 (2001).
52 Smith, E. R., Winter, B., Eissenberg, J. C. & Shilatifard, A. Regulation of the transcriptional activity of poised RNA polymerase II by the elongation factor ELL. *Proc Natl Acad Sci U S A* **105**, 8575-8579, doi:10.1073/pnas.0804379105 (2008).
53 Lindquist, S. The heat-shock response. *Annu Rev Biochem* **55**, 1151-1191, doi:10.1146/annurev.bi.55.070186.005443 (1986).
54 Richter, K., Haslbeck, M. & Buchner, J. The heat shock response: life on the verge of death. *Mol Cell* **40**, 253-266, doi:10.1016/j.molcel.2010.10.006 (2010).
55 Rougvie, A. E. & Lis, J. T. The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged. *Cell* **54**, 795-804 (1988).
56 Rougvie, A. E. & Lis, J. T. Postinitiation transcriptional control in Drosophila melanogaster. *Mol Cell Biol* **10**, 6041-6045 (1990).
57 Artavanis-Tsakonas, S., Schedl, P., Mirault, M. E., Moran, L. & Lis, J. Genes for the 70,000 dalton heat shock protein in two cloned D. melanogaster DNA segments. *Cell* **17**, 9-18 (1979).
58 Craig, E. A., McCarthy, B. J. & Wadsworth, S. C. Sequence organization of two recombinant plasmids containing genes for the major heat shock-induced protein of D. melanogaster. *Cell* **16**, 575-588 (1979).
59 Holmgren, R., Livak, K., Morimoto, R., Freund, R. & Meselson, M. Studies of cloned sequences from four Drosophila heat shock loci. *Cell* **18**, 1359-1370 (1979).
60 Livak, K. J., Freund, R., Schweber, M., Wensink, P. C. & Meselson, M. Sequence organization and transcription at two heat shock loci in Drosophila. *Proc Natl Acad Sci U S A* **75**, 5613-5617 (1978).
61 Moran, L. *et al.* Physical map of two D. melanogaster DNA segments containing sequences coding for the 70,000 dalton heat shock protein. *Cell* **17**, 1-8 (1979).
62 Schedl, P. *et al.* Two hybrid plasmids with D. melanogaster DNA sequences complementary to mRNA coding for the major heat shock protein. *Cell* **14**, 921-929 (1978).
63 Walldorf, U. *et al.* Cloning of heat-shock locus 93D from Drosophila melanogaster. *EMBO J* **3**, 2499-2504 (1984).
64 Bunch, H. RNA polymerase II pausing and transcriptional regulation of the HSP70 expression. *Eur J Cell Biol* **96**, 739-745, doi:10.1016/j.ejcb.2017.09.003 (2017).
65 Mahat, D. B., Salamanca, H. H., Duarte, F. M., Danko, C. G. & Lis, J. T. Mammalian Heat Shock Response and Mechanisms Underlying Its Genome-wide Transcriptional Regulation. *Mol Cell* **62**, 63-78, doi:10.1016/j.molcel.2016.02.025 (2016).
66 Luo, Z., Lin, C. & Shilatifard, A. The super elongation complex (SEC) family in transcriptional control. *Nat Rev Mol Cell Biol* **13**, 543-547, doi:10.1038/nrm3417 (2012).
67 Rowley, J. D. The critical role of chromosome translocations in human leukemias. *Annu Rev Genet* **32**, 495-519, doi:10.1146/annurev.genet.32.1.495 (1998).
68 Lin, C. *et al.* AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. *Mol Cell* **37**, 429-437, doi:10.1016/j.molcel.2010.01.026 (2010).
69 Smith, E. R. *et al.* The little elongation complex regulates small nuclear RNA transcription. *Mol Cell* **44**, 954-965, doi:10.1016/j.molcel.2011.12.008 (2011).

70 He, N. *et al.* HIV-1 Tat and host AFF4 recruit two transcription elongation factors into a bifunctional complex for coordinated activation of HIV-1 transcription. *Mol Cell* **38**, 428-438, doi:10.1016/j.molcel.2010.04.013 (2010).
71 Sobhian, B. *et al.* HIV-1 Tat assembles a multifunctional transcription elongation complex and stably associates with the 7SK snRNP. *Mol Cell* **38**, 439-451, doi:10.1016/j.molcel.2010.04.012 (2010).
72 Luo, Z. *et al.* The super elongation complex family of RNA polymerase II elongation factors: gene target specificity and transcriptional output. *Mol Cell Biol* **32**, 2608-2617, doi:10.1128/MCB.00182-12 (2012).
73 Lin, C. *et al.* Dynamic transcriptional events in embryonic stem cells mediated by the super elongation complex (SEC). *Genes Dev* **25**, 1486-1498, doi:10.1101/gad.2059211 (2011).
74 Hu, D. *et al.* The little elongation complex functions at initiation and elongation phases of snRNA gene transcription. *Mol Cell* **51**, 493-505, doi:10.1016/j.molcel.2013.07.003 (2013).
75 Will, C. L. & Luhrmann, R. Spliceosome structure and function. *Cold Spring Harb Perspect Biol* **3**, doi:10.1101/cshperspect.a003707 (2011).
76 Patel, A. A. & Steitz, J. A. Splicing double: insights from the second spliceosome. *Nat Rev Mol Cell Biol* **4**, 960-970, doi:10.1038/nrm1259 (2003).
77 Wahl, M. C., Will, C. L. & Luhrmann, R. The spliceosome: design principles of a dynamic RNP machine. *Cell* **136**, 701-718, doi:10.1016/j.cell.2009.02.009 (2009).
78 Schumperli, D. & Pillai, R. S. The special Sm core structure of the U7 snRNP: far-reaching significance of a small nuclear ribonucleoprotein. *Cell Mol Life Sci* **61**, 2560-2570, doi:10.1007/s00018-004-4190-0 (2004).
79 Sharma, N. Regulation of RNA polymerase II-mediated transcriptional elongation: Implications in human disease. *IUBMB Life* **68**, 709-716, doi:10.1002/iub.1538 (2016).
80 Smith, E., Lin, C. & Shilatifard, A. The super elongation complex (SEC) and MLL in development and disease. *Genes Dev* **25**, 661-672, doi:10.1101/gad.2015411 (2011).
81 Conaway, J. W. & Conaway, R. C. Transcription elongation and human disease. *Annu Rev Biochem* **68**, 301-319, doi:10.1146/annurev.biochem.68.1.301 (1999).
82 Lee, T. I. & Young, R. A. Transcriptional regulation and its misregulation in disease. *Cell* **152**, 1237-1251, doi:10.1016/j.cell.2013.02.014 (2013).
83 Izumi, K. *et al.* Germline gain-of-function mutations in AFF4 cause a developmental syndrome functionally linking the super elongation complex and cohesin. *Nat Genet* **47**, 338-344, doi:10.1038/ng.3229 (2015).
84 Prudencio, M. *et al.* Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. *Nat Neurosci* **18**, 1175-1182, doi:10.1038/nn.4065 (2015).
85 Tollervey, J. R. *et al.* Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat Neurosci* **14**, 452-458, doi:10.1038/nn.2778 (2011).
86 Ishihara, T. *et al.* Decreased number of Gemini of coiled bodies and U12 snRNA level in amyotrophic lateral sclerosis. *Hum Mol Genet* **22**, 4136-4147, doi:10.1093/hmg/ddt262 (2013).
87 Tsuiji, H. *et al.* Spliceosome integrity is defective in the motor neuron diseases ALS and SMA. *EMBO Mol Med* **5**, 221-234, doi:10.1002/emmm.201202303 (2013).

88 Turunen, J. J., Niemela, E. H., Verma, B. & Frilander, M. J. The significant other: splicing by the minor spliceosome. *Wiley Interdiscip Rev RNA* **4**, 61-76, doi:10.1002/wrna.1141 (2013).
89 Verma, B., Akinyi, M. V., Norppa, A. J. & Frilander, M. J. Minor spliceosome and disease. *Semin Cell Dev Biol* **79**, 103-112, doi:10.1016/j.semcdb.2017.09.036 (2018).
90 Elsaid, M. F. *et al.* Mutation in noncoding RNA RNU12 causes early onset cerebellar ataxia. *Ann Neurol* **81**, 68-78, doi:10.1002/ana.24826 (2017).
91 Argente, J. *et al.* Defective minor spliceosome mRNA processing results in isolated familial growth hormone deficiency. *EMBO Mol Med* **6**, 299-306, doi:10.1002/emmm.201303573 (2014).
92 Madan, V. *et al.* Aberrant splicing of U12-type introns is the hallmark of ZRSR2 mutant myelodysplastic syndrome. *Nat Commun* **6**, 6042, doi:10.1038/ncomms7042 (2015).
93 He, H. *et al.* Mutations in U4atac snRNA, a component of the minor spliceosome, in the developmental disorder MOPD I. *Science* **332**, 238-240, doi:10.1126/science.1200587 (2011).
94 Edery, P. *et al.* Association of TALS developmental disorder with defect in minor splicing component U4atac snRNA. *Science* **332**, 240-243, doi:10.1126/science.1202205 (2011).
95 Merico, D. *et al.* Compound heterozygous mutations in the noncoding RNU4ATAC cause Roifman Syndrome by disrupting minor intron splicing. *Nat Commun* **6**, 8718, doi:10.1038/ncomms9718 (2015).
96 Li, D. K., Tisdale, S., Lotti, F. & Pellizzoni, L. SMN control of RNP assembly: from post-transcriptional gene regulation to motor neuron disease. *Semin Cell Dev Biol* **32**, 22-29, doi:10.1016/j.semcdb.2014.04.026 (2014).
97 Zhang, Z. *et al.* SMN deficiency causes tissue-specific perturbations in the repertoire of snRNAs and widespread defects in splicing. *Cell* **133**, 585-600, doi:10.1016/j.cell.2008.03.031 (2008).
98 Gabanella, F. *et al.* Ribonucleoprotein assembly defects correlate with spinal muscular atrophy severity and preferentially affect a subset of spliceosomal snRNPs. *PLoS One* **2**, e921, doi:10.1371/journal.pone.0000921 (2007).
99 Boulisfane, N. *et al.* Impaired minor tri-snRNP assembly generates differential splicing defects of U12-type introns in lymphoblasts derived from a type I SMA patient. *Hum Mol Genet* **20**, 641-648, doi:10.1093/hmg/ddq508 (2011).
100 Lotti, F. *et al.* An SMN-dependent U12 splicing event essential for motor circuit function. *Cell* **151**, 440-454, doi:10.1016/j.cell.2012.09.012 (2012).
101 Lindberg, I. *et al.* Chaperones in Neurodegeneration. *J Neurosci* **35**, 13853-13859, doi:10.1523/JNEUROSCI.2600-15.2015 (2015).
102 Neef, D. W., Jaeger, A. M. & Thiele, D. J. Heat shock transcription factor 1 as a therapeutic target in neurodegenerative diseases. *Nat Rev Drug Discov* **10**, 930-944, doi:10.1038/nrd3453 (2011).
103 Berson, A. *et al.* TDP-43 Promotes Neurodegeneration by Impairing Chromatin Remodeling. *Curr Biol*, doi:10.1016/j.cub.2017.10.024 (2017).
104 Chang, H. Y., Hou, S. C., Way, T. D., Wong, C. H. & Wang, I. F. Heat-shock protein dysregulation is associated with functional and pathological TDP-43 aggregation. *Nat Commun* **4**, 2757, doi:10.1038/ncomms3757 (2013).

105 Chen, H. J. *et al.* The heat shock response plays an important role in TDP-43 clearance: evidence for dysfunction in amyotrophic lateral sclerosis. *Brain* **139**, 1417-1432, doi:10.1093/brain/aww028 (2016).
106 Lakhotia, S. C. Forty years of the 93D puff of Drosophila melanogaster. *J Biosci* **36**, 399-423 (2011).
107 Lakhotia, S. C., Mallik, M., Singh, A. K. & Ray, M. The large noncoding hsromega-n transcripts are essential for thermotolerance and remobilization of hnRNPs, HP1 and RNA polymerase II during recovery from heat shock in Drosophila. *Chromosoma* **121**, 49-70, doi:10.1007/s00412-011-0341-x (2012).
108 Jolly, C. & Lakhotia, S. C. Human sat III and Drosophila hsr omega transcripts: a common paradigm for regulation of nuclear RNA processing in stressed cells. *Nucleic Acids Res* **34**, 5508-5514, doi:10.1093/nar/gkl711 (2006).
109 Biamonti, G. & Vourc'h, C. Nuclear stress bodies. *Cold Spring Harb Perspect Biol* **2**, a000695, doi:10.1101/cshperspect.a000695 (2010).
110 Rizzi, N. *et al.* Transcriptional activation of a constitutive heterochromatic domain of the human genome in response to heat shock. *Mol Biol Cell* **15**, 543-551, doi:10.1091/mbc.E03-07-0487 (2004).
111 Buratti, E. & Baralle, F. E. Multiple roles of TDP-43 in gene expression, splicing regulation, and human disease. *Front Biosci* **13**, 867-878 (2008).
112 Ito, D., Hatano, M. & Suzuki, N. RNA binding proteins and the pathological cascade in ALS/FTD neurodegeneration. *Sci Transl Med* **9**, doi:10.1126/scitranslmed.aah5436 (2017).
113 Lalmansingh, A. S., Urekar, C. J. & Reddi, P. P. TDP-43 is a transcriptional repressor: the testis-specific mouse acrv1 gene is a TDP-43 target in vivo. *J Biol Chem* **286**, 10970-10982, doi:10.1074/jbc.M110.166587 (2011).
114 Swain, A. *et al.* Drosophila TDP-43 RNA-Binding Protein Facilitates Association of Sister Chromatid Cohesion Proteins with Genes, Enhancers and Polycomb Response Elements. *PLoS Genet* **12**, e1006331, doi:10.1371/journal.pgen.1006331 (2016).
115 Jaiswal, M., Sandoval, H., Zhang, K., Bayat, V. & Bellen, H. J. Probing mechanisms that underlie human neurodegenerative diseases in Drosophila. *Annu Rev Genet* **46**, 371-396, doi:10.1146/annurev-genet-110711-155456 (2012).
116 Lu, B. & Vogel, H. Drosophila models of neurodegenerative diseases. *Annu Rev Pathol* **4**, 315-342, doi:10.1146/annurev.pathol.3.121806.151529 (2009).
117 Tollervey, J. R. *et al.* Analysis of alternative splicing associated with aging and neurodegeneration in the human brain. *Genome Res* **21**, 1572-1582, doi:10.1101/gr.122226.111 (2011).
118 Duarte, F. M. *et al.* Transcription factors GAF and HSF act at distinct regulatory steps to modulate stress-induced gene activation. *Genes Dev* **30**, 1731-1746, doi:10.1101/gad.284430.116 (2016).
119 Goenka, A. *et al.* Human satellite-III non-coding RNAs modulate heat-shock-induced transcriptional repression. *J Cell Sci* **129**, 3541-3552, doi:10.1242/jcs.189803 (2016).
120 Tasselli, L. *et al.* SIRT6 deacetylates H3K18ac at pericentric chromatin to prevent mitotic errors and cellular senescence. *Nat Struct Mol Biol* **23**, 434-440, doi:10.1038/nsmb.3202 (2016).

121 Chen-Plotkin, A. S. *et al.* Variations in the progranulin gene affect global gene expression in frontotemporal lobar degeneration. *Hum Mol Genet* **17**, 1349-1362, doi:10.1093/hmg/ddn023 (2008).
122 Chen-Plotkin, A. S. *et al.* TMEM106B, the risk gene for frontotemporal dementia, is regulated by the microRNA-132/212 cluster and affects progranulin pathways. *J Neurosci* **32**, 11213-11227, doi:10.1523/JNEUROSCI.0521-12.2012 (2012).
123 Smith, E. & Shilatifard, A. Transcriptional elongation checkpoint control in development and disease. *Genes Dev* **27**, 1079-1088, doi:10.1101/gad.215137.113 (2013).
124 Ling, J. P., Pletnikova, O., Troncoso, J. C. & Wong, P. C. TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. *Science* **349**, 650-655, doi:10.1126/science.aab0983 (2015).
125 Takahashi, H. *et al.* Human mediator subunit MED26 functions as a docking site for transcription elongation factors. *Cell* **146**, 92-104, doi:10.1016/j.cell.2011.06.005 (2011).
126 Takahashi, H. *et al.* MED26 regulates the transcription of snRNA genes through the recruitment of little elongation complex. *Nat Commun* **6**, 5941, doi:10.1038/ncomms6941 (2015).
127 Gardini, A. *et al.* Integrator regulates transcriptional initiation and pause release following activation. *Mol Cell* **56**, 128-139, doi:10.1016/j.molcel.2014.08.004 (2014).
128 He, N. *et al.* Human Polymerase-Associated Factor complex (PAFc) connects the Super Elongation Complex (SEC) to RNA polymerase II on chromatin. *Proc Natl Acad Sci U S A* **108**, E636-645, doi:10.1073/pnas.1107107108 (2011).
129 Mirra, A. *et al.* Functional interaction between FUS and SMN underlies SMA-like splicing changes in wild-type hFUS mice. *Sci Rep* **7**, 2033, doi:10.1038/s41598-017-02195-0 (2017).
130 DeJesus-Hernandez, M. *et al.* Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* **72**, 245-256, doi:10.1016/j.neuron.2011.09.011 (2011).
131 Renton, A. E. *et al.* A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **72**, 257-268, doi:10.1016/j.neuron.2011.09.010 (2011).
132 Ferrari, R. *et al.* Frontotemporal dementia: insights into the biological underpinnings of disease through gene co-expression network analysis. *Mol Neurodegener* **11**, 21, doi:10.1186/s13024-016-0085-4 (2016).
133 Beere, H. M. Death versus survival: functional interaction between the apoptotic and stress-inducible heat shock protein pathways. *J Clin Invest* **115**, 2633-2639, doi:10.1172/JCI26471 (2005).
134 Fulda, S., Gorman, A. M., Hori, O. & Samali, A. Cellular stress responses: cell survival and cell death. *Int J Cell Biol* **2010**, 214074, doi:10.1155/2010/214074 (2010).
135 Enukashvily, N. I., Donev, R., Waisertreiger, I. S. & Podgornaya, O. I. Human chromosome 1 satellite 3 DNA is decondensed, demethylated and transcribed in senescent cells and in A431 epithelial carcinoma cells. *Cytogenet Genome Res* **118**, 42-54, doi:10.1159/000106440 (2007).
136 Shumaker, D. K. *et al.* Mutant nuclear lamin A leads to progressive alterations of epigenetic control in premature aging. *Proc Natl Acad Sci U S A* **103**, 8703-8708, doi:10.1073/pnas.0602569103 (2006).

137 Shilatifard, A., Haque, D., Conaway, R. C. & Conaway, J. W. Structure and function of RNA polymerase II elongation factor ELL. Identification of two overlapping ELL functional domains that govern its interaction with polymerase and the ternary elongation complex. *J Biol Chem* **272**, 22355-22363 (1997).
138 Chen, Y. *et al.* ELL targets c-Myc for proteasomal degradation and suppresses tumour growth. *Nat Commun* **7**, 11057, doi:10.1038/ncomms11057 (2016).
139 Qi, S. *et al.* Structural basis for ELL2 and AFF4 activation of HIV-1 proviral transcription. *Nat Commun* **8**, 14076, doi:10.1038/ncomms14076 (2017).
140 Tricoire, H. *et al.* The steroid hormone receptor EcR finely modulates Drosophila lifespan during adulthood in a sex-specific manner. *Mech Ageing Dev* **130**, 547-552, doi:10.1016/j.mad.2009.05.004 (2009).
141 Kuan, P. F. *et al.* A Statistical Framework for the Analysis of ChIP-Seq Data. *J Am Stat Assoc* **106**, 891-903, doi:10.1198/jasa.2011.ap09706 (2011).
142 Park, S. J., Kim, J. H., Yoon, B. H. & Kim, S. Y. A ChIP-Seq Data Analysis Pipeline Based on Bioconductor Packages. *Genomics Inform* **15**, 11-18, doi:10.5808/GI.2017.15.1.11 (2017).
143 Yu, C. E. *et al.* The spectrum of mutations in progranulin: a collaborative study screening 545 cases of neurodegeneration. *Arch Neurol* **67**, 161-170, doi:10.1001/archneurol.2009.328 (2010).
144 Suh, E. *et al.* Semi-automated quantification of C9orf72 expansion size reveals inverse correlation between hexanucleotide repeat number and disease duration in frontotemporal degeneration. *Acta Neuropathol* **130**, 363-372, doi:10.1007/s00401-015-1445-9 (2015).
145 Watts, G. D. *et al.* Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. *Nat Genet* **36**, 377-381, doi:10.1038/ng1332 (2004).
146 Kwong, L. K. *et al.* Novel monoclonal antibodies to normal and pathologically altered human TDP-43 proteins. *Acta Neuropathol Commun* **2**, 33, doi:10.1186/2051-5960-2-33 (2014).
147 Burguete, A. S. *et al.* GGGGCC microsatellite RNA is neuritically localized, induces branching defects, and perturbs transport granule function. *Elife* **4**, e08881, doi:10.7554/eLife.08881 (2015).
148 Bilen, J. & Bonini, N. M. Genome-wide screen for modifiers of ataxin-3 neurodegeneration in Drosophila. *PLoS Genet* **3**, 1950-1964, doi:10.1371/journal.pgen.0030177 (2007).
149 Klepsatel, P., Galikova, M., Xu, Y. & Kuhnlein, R. P. Thermal stress depletes energy reserves in Drosophila. *Sci Rep* **6**, 33667, doi:10.1038/srep33667 (2016).
150 Schulte, P. M. What is environmental stress? Insights from fish living in a variable environment. *J Exp Biol* **217**, 23-34, doi:10.1242/jeb.089722 (2014).
151 Chown, S. L. & Terblanche, J. S. Physiological Diversity in Insects: Ecological and Evolutionary Contexts. *Adv In Insect Phys* **33**, 50-152, doi:10.1016/S0065-2806(06)33002-0 (2006).
152 Boonstra, R. Reality as the leading cause of stress: rethinking the impact of chronic stress in nature. *Funct Ecol* **27**, 11-23, doi:10.1111/1365-2435.12008 (2013).

153 Kwak, H., Fuda, N. J., Core, L. J. & Lis, J. T. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science* **339**, 950-953, doi:10.1126/science.1229386 (2013).
154 Freibaum, B. D. *et al.* GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature* **525**, 129-133, doi:10.1038/nature14974 (2015).
155 Xu, Z. *et al.* Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. *Proc Natl Acad Sci U S A* **110**, 7778-7783, doi:10.1073/pnas.1219643110 (2013).
156 Zhang, K. *et al.* The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature* **525**, 56-61, doi:10.1038/nature14973 (2015).
157 Ross, C. A. & Poirier, M. A. Protein aggregation and neurodegenerative disease. *Nat Med* **10 Suppl**, S10-17, doi:10.1038/nm1066 (2004).
158 Cardinale, A., Chiesa, R. & Sierks, M. Protein misfolding and neurodegenerative diseases. *Int J Cell Biol* **2014**, 217371, doi:10.1155/2014/217371 (2014).
159 Kramer, N. J. *et al.* Spt4 selectively regulates the expression of C9orf72 sense and antisense mutant transcripts. *Science* **353**, 708-712, doi:10.1126/science.aaf7791 (2016).
160 Conicella, A. E., Zerze, G. H., Mittal, J. & Fawzi, N. L. ALS Mutations Disrupt Phase Separation Mediated by alpha-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. *Structure* **24**, 1537-1549, doi:10.1016/j.str.2016.07.007 (2016).
161 Wang, A. *et al.* A single N-terminal phosphomimic disrupts TDP-43 polymerization, phase separation, and RNA splicing. *EMBO J* **37**, doi:10.15252/embj.201797452 (2018).
162 Buratti, E. Functional Significance of TDP-43 Mutations in Disease. *Adv Genet* **91**, 1-53, doi:10.1016/bs.adgen.2015.07.001 (2015).
163 Gendron, T. F., Rademakers, R. & Petrucelli, L. TARDBP mutation analysis in TDP-43 proteinopathies and deciphering the toxicity of mutant TDP-43. *J Alzheimers Dis* **33 Suppl 1**, S35-45, doi:10.3233/JAD-2012-129036 (2013).
164 Udan-Johns, M. *et al.* Prion-like nuclear aggregation of TDP-43 during heat shock is regulated by HSP40/70 chaperones. *Hum Mol Genet* **23**, 157-170, doi:10.1093/hmg/ddt408 (2014).
165 Courchaine, E. M., Lu, A. & Neugebauer, K. M. Droplet organelles? *EMBO J* **35**, 1603-1612, doi:10.15252/embj.201593517 (2016).
166 Hyman, A. A., Weber, C. A. & Julicher, F. Liquid-liquid phase separation in biology. *Annu Rev Cell Dev Biol* **30**, 39-58, doi:10.1146/annurev-cellbio-100913-013325 (2014).
167 Elbaum-Garfinkle, S. *et al.* The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc Natl Acad Sci U S A* **112**, 7189-7194, doi:10.1073/pnas.1504822112 (2015).
168 Lin, Y., Protter, D. S., Rosen, M. K. & Parker, R. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol Cell* **60**, 208-219, doi:10.1016/j.molcel.2015.08.018 (2015).
169 Zhang, H. *et al.* RNA Controls PolyQ Protein Phase Transitions. *Mol Cell* **60**, 220-230, doi:10.1016/j.molcel.2015.09.017 (2015).