Modeling splicing-related autosomal dominant retinitis pigmentosa in *Caenorhabditis elegans*

Karinna Denise Rubio Peña

TESI DOCTORAL UPF - 2017

Director de la Tesi

Dr. Julián Cerón Madrigal

Modeling human diseases in *C. elegans*. Genes, disease and therapies programe-IDIBELL

DEPARTAMENTO DE CIENCIAS EXPERIMENTALES
Y DE LA SALUD
UNIVERSIDAD POMPEU FABRA

A mis padres,a mi familia

y a los que ya no estan

ABSTRACT

Mutations in some ubiquitously expressed splicing factors genes have been linked to the autosomal dominant form of a rare genetic disease called Retinitis Pigmentosa (adRP). RP is characterized by a progressive visual degeneration produced by the apoptosis of photoreceptors.

Taking advantage of RNA mediated-interference (RNAi) and RNA-sequencing we started to build a *C. elegans* model where to study the disease. We found two important similarities between s-adRP and the RNAi phenotypes observed in *C. elegans*: (i) there is a cell-type-specific apoptosis and (ii) it seems to be associated with transcriptionally active tissues.

We have stated a working model to investigate the mechanisms that triggers apoptosis in these s-adRP retinal cells. This model involves inefficient splicing, reduced transcriptional efficiency, and presence of R-loops as source of replicative stress and genomic instability.
Additionally, we are using CRISPR/Cas9 to introduce specific s-adRP mutations in the *C. elegans* genome to constitute a platform where to screen for genetic or drug modifiers of the disease.

RESUM

Mutacions en gens que codifiquen per factors de *splicing* expressats de forma ubiqua s'han associat a la forma autosòmica dominant d'una malaltia genètica minoritària anomenada Retinitis Pigmentosa (adRP). La RP es caracteritza per una degeneració visual progressiva causada per l'apoptosi de fotoreceptors.

Utilitzant les tècniques d'ARN d'interferència (RNAi) i seqüenciació d'ARN vàrem iniciar un model de la malaltia en *C. elegans*. Trobàrem dues semblances importants entre la s-adRP i els fenotips de RNAi en *C. elegans*: (i) hem observat que l'apoptosi és específica d'un tipus cel·lular i (ii) aquesta apoptosi sembla estar associada a teixits amb alta activitat transcripcional.

Hem establert un model on poder investigar els mecanismes que desencadenen l'apoptosi en les cèl·lules de la retina afectades per s-adRP. Aquest model comprèn un *splicing* ineficient, una reducció en l'activitat transcripcional i la presència de R-loops com a font d'estrès replicatiu i inestabilitat genòmica. A més, hem utilitzat la tècnica de CRISPR/Cas9 per introduir mutacions específiques de s-adRP en el genoma de *C. elegans* amb la finalitat de poder utilitzar-ho com a plataforma per identificar gens o fàrmacs que modifiquin el curs de malaltia.

PREFACE

PREFACE

Retinitis pigmentosa (RP) is a rare disease characterized by the progressive loss of photoreceptor cells. With a worldwide incidence of approximately 1 in 3500 individuals it is also the most common type of inherited retinal degeneration. Since a subtype of autosomal dominant RP (adRP) was related to splicing factor proteins less than two decades ago (Chakarova et al., 2002; McKie et al., 2001; E N Vithana et al., 2001), outstanding efforts has been taken to study the nature of this disease. In fact, typing 'retinitis pigmentosa' and 'splicing factors' in Pubmed gives us more than 200 publications regarding mutations, studies in cellular and animal models, and/or characterization of protein structures and molecular functions of the related splicing factors. This is extraordinary taking into consideration that retinitis pigmentosa is a rare disease and that only a subtype of RP is related to splicing. Unfortunately, despite the amount of information available, the mechanisms behind the role of ubiquitous proteins in this retinal-specific disease remain elusive. Even more, no effective therapies have been proposed to cure or at least delay the progression of the disease.

Meanwhile, in the last few decades the nematode *Caenorhabditis elegans* has been gaining prominence in the biomedical research field. This is not surprising, since *C. elegans* harbors a set of traits such as a short life cycle, easy genetics, conserved mechanisms and pathways, easy manipulation, among others that will be discussed in detail later. Such traits, makes the worm an appealing organism in which to model and study human diseases. The fact that most of the important signaling pathways and cellular mechanisms have remained essentially

unchanged during evolution, has allowed the worm to be used to model neurodegenerative diseases like Parkinson or Alzheimer, muscular atrophies, metabolic disorders and cancer.

Being retinitis pigmentosa a hereditary disease without cure and with still obscure mechanisms driven the pathology, it was appealing to try to model it in *C. elegans*. Even more when adRP is associated with evolutionary conserved essential genes, such as those encoding splicing factors. The fact that the splicing factors related to adRP are essential, has hampered the development of animal models to study this disorder because the complete depletion of these genes would be lethal. Fortunately, the partial inactivation of a desired gene is pretty straightforward in the worm by RNA mediated-interference (RNAi). Furthermore, direct genome editing to produce partial loss-of-function mutations is also possible using the CRISPR/Cas9.

Altogether, we spotted the potential that *C. elegans* had to be a model for splicing related adRP (s-adRP). Our work led to a publication in RNA journal in 2015. In that manuscript we established *C. elegans* as a model where to investigate s-adRP and proposed a working model in which an inefficient splicing, reduced transcriptional efficiency, R-loops accumulation, replicative stress and genomic instability in highly metabolically active tissues, may converge to trigger a fatal apoptosis.

The recent incorporation of CRISPR/Cas9 to our model has opened new possibilities since, we can now, introduce specific s-adRP mutations into the genome of the worm providing a personalized model. Thus, not only studies regarding the identification of mechanisms of the disease can be pursue, but also high-throughput screenings to identify genetic or drug modifiers towards a personalized therapy can be easily accomplished.

Although *C. elegans* s-adRP model might not perfectly represent the pathophysiology of such a complex disease, the observations we make at the molecular and cellular levels would provide new insights into the mechanisms, regulation and development of s-adRP.

"And thus we find ourselves in a surprising position: As incredible as it seems, future research on flies and worms will quite often provide the shortest and most efficient path to curing human disease."

Bruce Alberts-Editor-in-Chief of Science, 2010

KEY WORDS

C. elegans

Retinitis pigmentosa

Splicing factors

Apoptosis

Intron retention

RNA-seq

CRISPR

Replication stress

R-loops

Genomic instability

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

I.1. Pre-mRNA splicing

Right after the discovery of the DNA structure and the genetic code, everyone assumed that genes were continuous and that the corresponding messenger RNA (mRNA) was directly translated into a protein. They were not wrong assuming this for prokaryotes, but these notions of gene expression did not seem to fit observations in eukaryotic organisms. It was not until 1977, that the recognition of the existence of a discontinuous mRNA was possible. The comparison of the sequence length between the cytoplasmic mRNA and its corresponding DNA sequence in the nucleus lead to the discovery that most eukaryotic genes are transcribed as precursors of mRNA (pre-mRNAs) which contain non-coding (introns) and coding (exons) sequences (Berget, Moore, & Sharp, 1977; Sharp, 2005). These observations led to the discovery of the existence of split genes and pre-mRNA splicing.

Pre-mRNA splicing is the process by which introns are accurately removed and exons are ligated together to produce a mature mRNA via two transesterification reactions (reviewed in (Grainger, 2005; Korir, Roberts, Ramesar, & Seoighe, 2014; Will & Lührmann, 2011) (Fig. I. 1). The mature mRNA consists of the sequentially connected exons that are exported to the cytoplasm to act as template for protein synthesis by the translation machinery (Wahl, Will, & Lührmann, 2009).

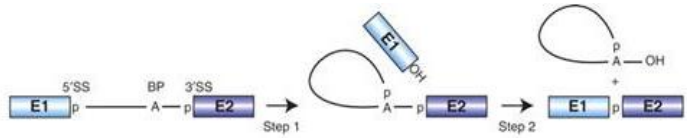

Fig. I. 1 Schematic representation of the two-step mechanism of pre-mRNA splicing.
Exons are represented by boxes and solid lines represent the intron. Letter A indicates the branch site adenosine and the phosphate groups at the 5' and 3' splice sites are indicated by (p), which are conserved in the splicing products. (*Modified from Will and Luhrmann, 2011*).

Most pre-mRNA exons are constitutive; they are always included in the mature mRNA. However, sometimes splicing can occur in alternative patterns generating multiple mRNA from the same genomic sequence. Alternative splicing (AS) is a major contributor to protein diversity in metazoan, as these alternatively spliced mRNAs give rise to protein isoforms with different chemical and biological activities (Black, 2003; Nilsen & Graveley, 2010). At the same time, we can find different types of alternative splicing mechanisms that include: exon skipping, alternative splice site usage or intron retention (Fig. I. 2.).

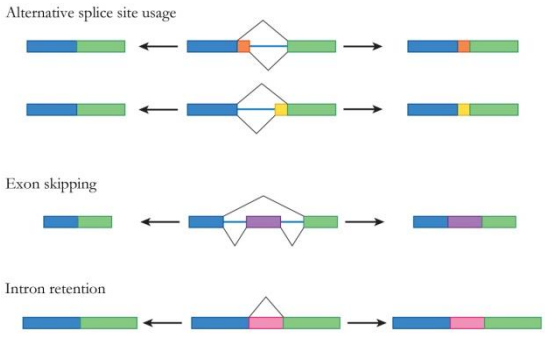

Fig. I. 2. Types of alternative splicing.
There are four basic types of alternative splicing: alternative splice site usage (5′ss and 3´ss), exon skipping and intron retention. (*Adapted from* Nilsen & Graveley, 2010).

Alternative splicing can also act as a post-transcriptional regulatory mechanism of gene expression. Many AS events introduce a premature termination codon (PTC) in one of the alternative isoforms. Barberan-Soler *et al*, showed that these AS transcripts containing PTCs are degraded by the Nonsense Mediated Decay pathway (NMD) (Barberan-Soler, Lambert, & Zahler, 2009). The NMD acts as a quality control of translation in eukaryotes and it had already been associated with transcripts that contain PTCs as consequence of nonsense mutations in the DNA sequence for decades (Losson & Lacroute, 1979).

From a biochemical point of view, pre-mRNA splicing seems to be a simple process that consists of a couple of transesterification reactions. However, an accurate recognition of the introns is necessary to assure a correct spliceosome assembly and consequently a correct splicing. For this to be possible, there exist a complex interplay of several *cis*- and *trans*-acting factors. *Cis*-acting signals comprise conserved sequences at the 5´ splice site (ss), 3´ss and the branch site (BS), which is located 20-50bp upstream the 3´ss. In higher eukaryotes, the BS is followed by a polypyrimidine tract (PPT). Additionally, splicing regulation relies on *cis*-acting regulatory sequences that can act as exonic and intronic splicing enhancers (ESEs and ISEs) or silencers (ESSs and ISSs)(Faustino, Cooper, & Andre, 2003; Lim & Burge, 2001; H. X. Liu, Zhang, & Krainer, 1998; Sun & Chasin, 2000; Wang & Cooper, 2007).

To ensure a proper removal of introns, a stepwise assembly of the spliceosome, guided by these elements, is needed. Once the spliceosome is assembled, a series of re-arrangements within the

spliceosome complex leads to its activation and to the two consequent transesterification reactions needed to accomplish intron removal and exon ligation (Corvelo, Hallegger, Smith, & Eyras, 2010; Poulos, Batra, Charizanis, & Swanson, 2011; Will & Lührmann, 2011).

I.1.1. The spliceosome

The spliceosome is one of the most complex macromolecules in the cell. Both, its conformation and composition are highly dynamic, with around 200 proteins going in and out of the complex depending on the splicing step that is taking place (Galej, Hoang, Nguyen, Newman, & Nagai, 2014). The two primary functions of splicing are performed by the spliceosome: (i) recognition of intron/exon boundaries, (ii) catalysis of the reactions that remove introns and join exons.

The core of the spliceosome is composed of five small nuclear ribonucleoproteins (snRNPs) named U1, U2, U5, and U4/U6 duplex for the U2-type spliceosome. Each snRNP consists of a small nuclear RNA (snRNA), except U4/U6 that has two, different snRNP-specific proteins, and a common set of Sm proteins. These snRNPs are remodeled and interact with different proteins to facilitate the formation of the RNA-RNA, RNA-protein and protein-protein interactions needed in each step of splicing (Will & Lührmann, 2006, 2011).

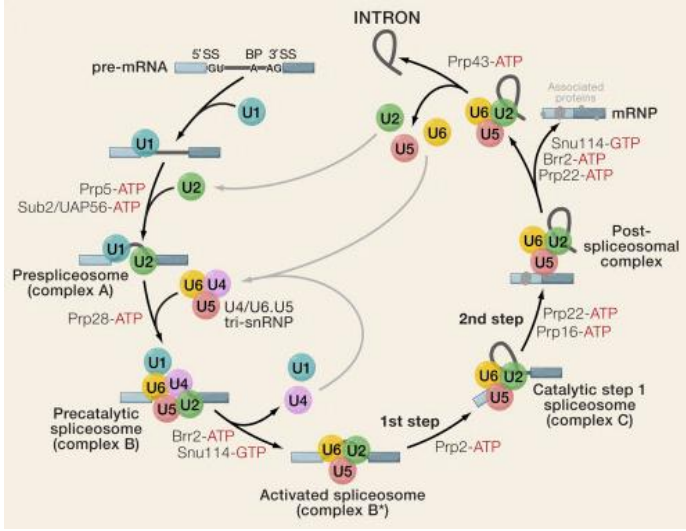

Fig. I. 3. Canonical assembly and disassembly pathway of the spliceosome. The ordered interactions of the snRNPs are shown (indicated by circles). For simplicity only few spliceosome proteins are shown in this scheme. The various spliceosomal complexes are named according to the metazoan nomenclature. Exon and intron sequences are indicated by boxes and lines, respectively. (*Extracted from* Wahl et al., 2009).

I.1.2. Spliceosome assembly and activation

Spliceosome assembly occurs by the ordered interaction of the spliceosomal snRNPs and other splicing factors. An overview of the spliceosome assembly and activation would be as follows: An intron flanked by two exons is recognized by the pre-spliceosome, which is composed by the U1 and U2 snRNPs complexes. Next, the U4/U6·U5 tri-snRNP remodels the spliceosome to later exclude U1 and U4 snRNPs, and keep U2, U5, and U6 snRNPs to form the activated spliceosome. Then, two catalytic steps occur to remove the intron and join the flanking exons (Fig. I.3.).

Behind this simple description, there is a very complex set of interactions between the snRNPs and different splicing factors. The biochemical and structural characterization of several of the spliceosome assembly and catalysis intermediates showed that transitions between assembly and catalysis stages of the spliceosome require deep compositional and conformational remodeling of the spliceosome's RNA and protein interaction networks (Wahl et al., 2009; Will & Lührmann, 2011). A deeper description of these interactions and rearrangements follows.

Intron recognition

The spliceosome assembles on its pre-mRNA substrate in a stepwise and carefully orchestrated manner. The assembly begins with U1 snRNP recognizing and base pairing to the 5'ss. At the same time, the branchpoint binding protein (BBP) is attaching to the branch site (BS). Then two other proteins, U2 auxiliary factors 35 and 65 (U2AF35 and U2AF65), bind to the 3'ss and PPT, respectively. This conformation represents the earliest splicing complex called the **E (early) complex**.

The primary function of U1 is to recruit the other spliceosome components to the substrate pre-mRNA. Thereafter, the U2 snRNP removes the BBP and binds to the BS, forming the pre-spliceosomal **complex A** in an ATP-dependent manner (Fig I.3.).

Spliceosome activation

Formation of the pre-catalytic **complex B** requires the ATP-dependent addition of the preformed U4/U6·U5 tri-snRNP to

complex A and numerous non-snRNP protein factors from the NineTeen Complex (NTC). At this point, all the necessary components for the splicing reaction are present; however, Complex B catalytic core remains inactive. A major rearrangement of its RNA network and its structure resulting in the release of U1 and U4 from the complex are required for its activation and the formation the activated spliceosome. This major remodeling event is catalyzed by DExD/H box helicases (SNRNP-200/Brr2 and DHX38), and regulated by PRP8 and SNU114, leading to the formation of the catalytically competent **complex B*** where only U5, U6 and U2 snRNPs remain.

Splicing catalysis

In complex B*, the first transesterification step of splicing takes place (Fig. I.3.). Here, a conserved branch point adenosine in the intron attacks the phosphodiester bond at the 5´-ss, leaving the 5´exon with a free 3´-hydroxyl group (Fig.I.1.). This leads to the formation of **C complex**, which catalyzes the second step. In the second step, the free 3´-hydroxyl group left by the previous reaction attacks the 3´-ss, leading to exon ligation and the elimination of the intron as a lariat (Fig.I.1.) (Fabrizio et al., 2009; Galej et al., 2014; Mozaffari-Jovin et al., 2014; Sanford & Caceres, 2004; Will & Lührmann, 2011).
After splicing is concluded, the whole machinery disassembles and its components are recycled for the next round of pre-mRNA maturation (Wahl et al., 2009).

I.1.3. Pre-mRNA splicing in human disease

Pre-mRNA splicing is a central step in gene expression. In fact, most human genes contain introns and require splicing to become mature transcripts, and not only that but 70% of them also express multiple mRNA by alternative splicing (Johnson et al. 2003; Wang & Cooper 2007). This requires a very complex regulation of the splicing process; an error that adds or removes even one nucleotide (nt) during splicing reaction would disrupt the open reading frame of a mRNA, however, intron-exon boundaries are recognized with extreme accuracy.

In average, a human gene contains 8 exons of 145 nt approximately. These exons are separated from each other by introns of up to hundreds or thousands of nucleotides of length, meaning that the introns represent around 90% of the transcription unit (Lander et al. 2001; Faustino et al. 2003; Wang & Cooper 2007). As mentioned before, within these intronic and exonic sequences several *cis*-acting regulatory elements, a part from classical splice sites, are found (ESEs, ISEs, ESSs and ISSs). Additionally, pseudo splice-sites are also present and the spliceosome must distinguish them from the authentic splice sites with the help of these regulatory elements (Faustino et al. 2003).

Being pre-mRNA splicing so essential for a correct gene expression and regulation in the cell, there is no surprise that mutations or any other alteration that disturbs this process can eventually lead to a bad functioning of the cell and finally to disease. For a long time, mutation analyses were primarily or exclusively performed at the genomic DNA level, and the effect of a mutation on the encoded

mRNA and protein was predicted from the primary sequence alone, leaving out of the equation mutations that could affect gene expression at transcriptional and post-transcriptional level.

In 1992, exonic and intronic splicing regulatory elements were not known yet, then, a study based uniquely in classical splice sites mutations published that only 15% of point mutations related with human disease disrupted splicing (Krawczak, Reiss, & Cooper, 1992; López-Bigas, Audit, Ouzounis, Parra, & Guigó, 2005). From this point of view, silent mutations in the *cis*-acting exonic elements (ESEs and ESSs) or in the intronic regions (ISEs and ISSs) that would not affect translation of the encoded transcript but could affect splicing regulation were probably underestimated. Additionally, now we know that mutations that affect *trans*-acting splicing elements, for example mutation in spliceosome proteins, are also implicated in various diseases (Faustino et al., 2003; Philips & Cooper, 2000; Scotti & Swanson, 2016; Wang & Cooper, 2007).

Splicing mutations, thus, may play a more important role than previously thought in human hereditary disease. It could be that some or many of the point mutations reported as silent mutations are in fact affecting the splicing pattern of the gene (López-Bigas et al., 2005).

Alterations in splicing can cause disease directly, modify the severity of the disease phenotype or be linked with disease susceptibility. It is important to make a distinction between the mechanisms causing altered splicing. In one hand, mutations in *cis*-acting elements will have a direct impact on the expression of only one gene, whereas

effects in *trans*-acting elements will have the potential to affect the expression of multiple genes (Faustino et al., 2003; Wang & Cooper, 2007). Independently of if it is a *cis*- or *trans*- mutation, these splicing-related mutations can either produce unnatural transcripts, or an aberrant regulation of splicing modifying the use of natural alternatively splice sites.

I.1.3.1. *Cis*-acting splicing mutations

Disruption of constitutive splice sites

It is now clear that a large but still unknown fraction of mutations affect splicing by disrupting components of the splicing code or by creating cryptic splice sites. Most mutations affecting splicing are single nucleotide substitutions located within intronic or exonic sequences of the classical splice sites. These mutations can cause several different outcomes that include: complete exon skipping, intron retention, activation of a pseudo splice site or generation of a new splice site within an exon or intron (cryptic splice site). Usually, these splicing alterations lead to the introduction of a premature termination codon (PTC) into the transcript which typically results in the degradation of the molecule by the NMD and as consequence the loss of function of the mutated allele. Mutations that cause pathogenic splicing abnormalities were identified in ESEs of breast cancer susceptibility genes, BRCA1 and BRCA2 (Fackenthal, Cartegni, Krainer, & Olopade, 2002; H. X. Liu, Cartegni, Zhang, & Krainer, 2001).

Alternative splicing deregulation

This type of mutations does not generate an aberrant splice; instead, these modifications promote a shift in the ratio of natural protein isoforms modifying normal alternative splicing patterns.

An example of how the modification of the AS pattern by a *cis-acting* mutation can lead to disease is the familial isolated growth hormone deficiency type II (IGHD II). This disorder is inherited in a dominant manner and affects the ratio of the alternative isoforms of the single growth hormone gene (GH-1). All IGHD II mutations cause increase alternative splicing of exon 3 as a consequence of the disruption of ISE, ESE or the 5´ss. (Binder, Brown, & Parks, 1996; Cogan et al., 1997; Moseley, Mullis, Prince, & Phillips, 2002).

I.1.3.2. *Trans*-acting splicing mutations

There are several genetic diseases in which a mutation disrupts either the constitutive components of the spliceosome or auxiliary factors that regulate alternative splicing. Spliceosome components are essential, so null mutations in these genes are generally lethal. It would be expected that mutations that cause a dysfunction of the basal splicing machinery would be lethal regardless of the cell type; however, mutations that disrupt the assembly or function of spliceosomal snRNPs are responsible for two human diseases that affect specific neuronal cell types, spinal muscular atrophy (SMA) and retinitis pigmentosa (RP).

SMA is an autosomal recessive disorder caused in most of the cases by the loss of the telomeric copy of the survivor of motor neuron gene (SMN1) (Wirth, 2000). This disease is characterized by the

progressive loss of lower motor neurons in the anterior horn of the spinal cord, accompanied by wasting of associated muscles and, ultimately, paralysis. SMN1 is an ubiquitously expressed protein that belongs to a complex required for the cytoplasm assembly and biogenesis of the core snRNPs but its specific function is still unknown (Briese, Esmaeili, & Sattelle, 2005).

RP is a retinal degenerative disease associated with more than 50 genes. One subtype of the disease has recently been related to mutations in splicing factor genes that belong to the U4·U6/U5 tri-snRNP complex. In the next chapter, retinitis pigmentosa etiology and its relation to mutations in ubiquitously expressed splicing factors will be treated in detail.

Trans-acting splicing mutations and cancer

It is clear that *cis*-acting mutations that affect the splicing of oncogenes, tumour suppressors and other cancer-relevant genes can have causal roles in cancer initiation and progression. However, most cancer-associated splicing changes are not associated with nucleotide changes in the affected genes implying that dramatic alterations in the *trans*-acting splicing environment are more likely to be the cause (Nissim-Rafinia & Kerem, 2002; Philips & Cooper, 2000). As a example, a recent report on cancer genetics showed that mutations in splicing factor genes were found in myelodysplastic syndrome (MDS), and in some other hematopoietic disorders and solid tumors (Yoshida & Ogawa, 2014).

Another example are the mutations on the splicing factor 3B subunit 1 (SF3B1). This protein belongs to the U2 snRNP complex and

mediates its binding to the branchpoint sequence, thus it may have an important role in splice site recognition. SF3B1 is mutated at significant rates in both haematological and solid cancers, including MDS, chronic lymphocytic leukemia (CLL), uveal melanoma and breast cancer, being the most commonly mutated spliceosomal gene in cancer (Dvinge, Kim, Abdel-Wahab, & Bradley, 2016).

I.1.4. Co-transcriptional splicing and disease

Co-transcriptional splicing has been demonstrated in a number of different systems and has been shown to play roles in coordinating both constitutive and alternative splicing. The nature of co-transcriptional splicing suggests that changes in transcription can dramatically affect splicing and viceversa.

A key player in coordinating transcription with splicing is the RNA polymerase itself. RNA polymerase II (RNA Pol II) is distinguished from the other eukaryotic RNA polymerases by the presence of a C-terminal domain (CTD), which is regulated by phosphorylation. Early studies investigating the role of the CTD in pre-mRNA splicing proposed that the CTD interacts directly with RNA splicing proteins to recruit them to the nascent transcript (Mortillaro et al., 1996). It has been demonstrated that truncation or mutation of the CTD of RNA Pol II leads to changes in the splicing process affecting the interaction of splicing proteins with the pre-mRNA (David, Boyne, Millhouse, & Manley, 2011; Hirose, Tacke, & Manley, 1999; McCracken et al., 1997). Proper CTD phosphorylation is also required for an efficient splicing of pre-mRNA substrates, as

demonstrated *in vitro* comparing splicing efficiency between T7 RNA polymerase and RNA Pol II transcribed RNAs (Das et al., 2006; Ghosh & Garcia-Blanco, 2000).

There is a growing body of evidence indicating that the relationship between transcription and splicing could work both ways, namely, that splicing and splicing factors could also influence transcription (Merkhofer, Hu, & Johnson, 2014). Indeed, it has been observed that interactions between U snRNPs and transcription elongation factors stimulated RNA polymerase II elongation (Fong & Zhou, 2001). Par contrary, *in vivo* depletion of either of SR proteins decreases nascent RNA production, with dramatic effects on transcription elongation seen upon SRSF2 depletion (Lin, Coutinho-Mansfield, Wang, Pandit, & Fu, 2008).

Since most of the pre-mRNA splicing in humans occurs co-transcriptionally, it is expected that defects in co-transcriptional splicing would have severe implications in human development and disease. Consistently, genes that are highly co-transcriptionally alternatively spliced in the fetal brain have been implicated in critical neuro developmental processes (Ameur et al., 2011).

Further analysis to better understand which genes are co-transcriptionally spliced and the mechanisms of co-transcriptional splicing will help us to identify genes whose dysregulation leads to disease or defects in development. Newer high-throughput technologies such as RNA-seq that have increased our knowledge of co-transcriptional splicing combined with traditional methods such as

classical genetics and biochemistry will lead to a better understanding of the mechanisms and spectrum of cotranscriptional splicing.

I.2. Retinitis Pigmentosa

Retinitis pigmentosa (RP) is a rare hereditary disease that is characterized by the progressive degeneration of the retina. Typical symptoms include night blindness followed by decreasing visual field, leading to tunnel vision and in later stages, complete blindness (Fig. I.4.) (Ferrari et al., 2011; Hartong, Berson, & Dryja, 2006).

Fig. I. 4. Tunnel Vision.
During the progression of retinitis pigmentosa, peripheral vision is affected due to the dead of rod photoreceptors leading to a condition known as tunnel vision (right panel).

These conditions are consequence of the gradual loss of the photoreceptors, the neurons responsible for transforming the light stimuli into an electric signal. There are two types of photoreceptors: (i) the **rods**, which mediate achromatic vision in starlight or moonlight; and (ii) the **cones**, which are important for color vision and fine acuity in daylight (Hartong et al., 2006; A. M. Rose & Bhattacharya, 2016). Despite the fact that perception in typical daytime light levels is dominated by cone-mediated vision, the total number of rods in the human retina (91 million) far exceeds the number of cones (roughly 4.5 million). As a result, the density of rods is much greater than cones throughout most of the retina (Purves et al., 2001).

The photoreceptors are highly specialized and compartmentalized neurons with four regions, each one with specific functions: the synaptic region, a cell body, an inner segment (IS) and an outer segment (OS) (Fig. I.5.).

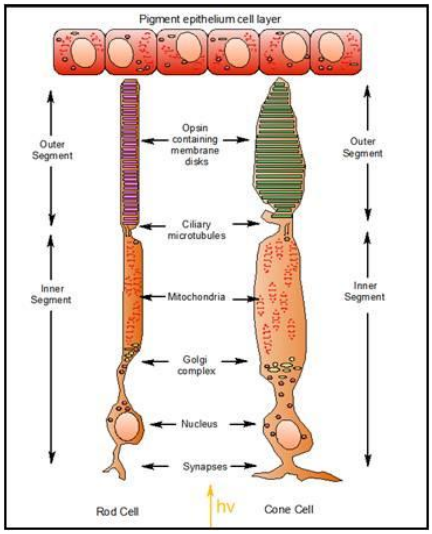

Fig. I. 5. Schematic representation of photoreceptor cells.
Rod photoreceptor cell representation on the left and cone on the right. Photoreceptors structure consists of four parts: (i) an outer segment where membrane disks are found, (ii) a cilium that connects the outer and inner segments, (iii) the inner segment that contains organelles, the cellular body where the nucleus and the synaptic body is localized, and (iv) the axon terminal. The retinal pigment epithelium (RPE) is in direct contact with photoreceptor´s outer segment (Hamel, 2006).

Phototransduction takes place in the photoreceptors´ outer segment that contains a stack of membrane discs which holds a great amount of visual transduction proteins, particularly rhodopsin, and cytoskeleton proteins. Daily, discs of the apex of the rod outer segment are phagocytosed by the retinal pigment epithelium (RPE). To compensate this phenomenon a boost of disc synthesis at the base of the outer segment is required. The molecular processes involved in biosynthesis, energy metabolism and membrane trafficking takes place within the IS of the photoreceptor cell, which is connected to the OS by the so called connecting cilium (Ferrari et al., 2011). This means that photoreceptors are under intense activity of mRNA and

protein synthesis, as well as an important protein trafficking from the rod´s IS, through the connecting cilium, to the rod´s OS. This cellular activity generates important energy consumption, requiring high content of mitochondria and oxygen, and mechanisms to protect the cell against the oxidative stress (Hamel, 2006).

In RP, the thickness of the outer nuclear layer suffers a reduction due to apoptosis of the photoreceptors, as well as lesions and/or retinal pigment deposits in the fundus (Fig. I. 6) (Ferrari et al., 2011; Hartong et al., 2006).

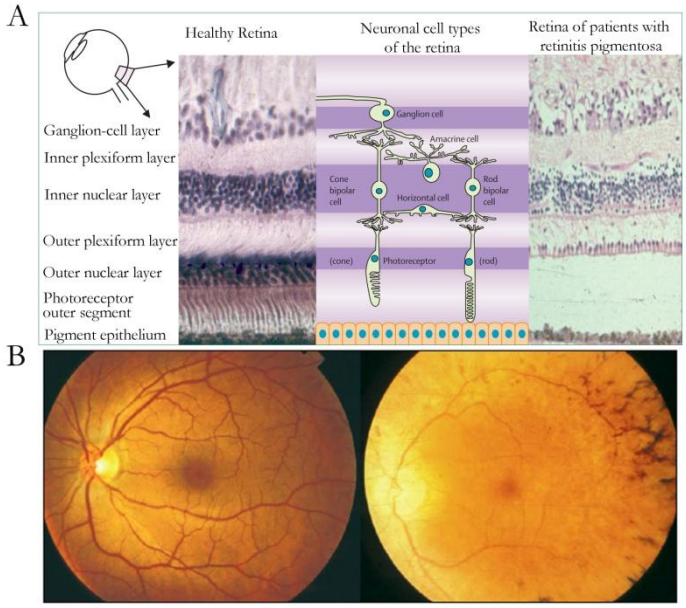

Fig. I. 6. Retinitis pigmentosa development of the disease.
(A) Histological appearance of healthy human retina (left) and retina of a patient with retinitis pigmentosa at a mid-stage of disease (right). Reduced thickness of the outer nuclear layer is evident in retinitis pigmentosa histological cut. (B) Fundi of a healthy individual (left) and a patient with retinitis pigmentosa (right). In the image of the diseased eye, optic-disc pallor, attenuated retinal arterioles, and peripheral intraretinal pigment deposits in a bone-spicule configuration are seen (Hartong et al., 2006)

I.2.1. RP is a hereditary disease

RP shows great genetic and phenotypic heterogeneity and it affects approximately one out of 3000-5000 individuals. Around 50 genes have been associated with the nonsyndromic form of retinitis pigmentosa. It can be inherited as an autosomal-dominant, autosomal-recessive, or X-linked trait (Hartong et al., 2006). The second most common inheritance pattern of RP is the autosomal dominant form (adRP) accounting for 30-40% of the cases (Daiger, Sullivan, & Bowne, 2013; A. M. Rose & Bhattacharya, 2016).

Over the last years, 25 genes have been associated with adRP, most of them are involved specifically in aspects of retinal biochemistry or structure; but 8 of them are spliceosome core components (PRPF3, PRPF4, PRPF6, PRPF8, PRPF31, SNRNP200/Brr2) or other splicing factors (RP9 and DHX38). These are all ubiquitously expressed and essential genes, however, when certain mutations appear they cause disease only in the retina while no adverse phenotypes in non-retinal tissues are known (Hartong et al., 2006; Korir et al., 2014; Ruzickova & Stanek, 2016; Tanackovic, Ransijn, Ayuso, et al., 2011; Wang & Cooper, 2007).

I.2.2. Retinitis pigmentosa and splicing factor mutations

Splicing-related genes linked with adRP (s-adRP) encode proteins present in the U4/U6·U5 tri-snRNP complex, which is a main actor in the process of splicing. Among these proteins, PRPF3, PRPF4, and PRPF31 are U4 snRNP components, whereas SNRNP200/BRR2, PRPF6, and PRPF8 are part of the U5 snRNP (Maita et al. 2005; Liu

et al. 2006). For the sake of simplicity, we can place the s-adRP proteins within the process of splicing as follows: An intron flanked by two exons is recognized by the pre-spliceosome (U1 and U2 snRNPs complexes). Next, the U4/U6·U5 tri-snRNP remodels the spliceosome to later exclude U1 and U4 snRNPs, and keep U2, U5, and U6 snRNPs to form the activated spliceosome. Then, two catalytic steps occur to remove the intron and join the flanking exons (Fig.I.7.).

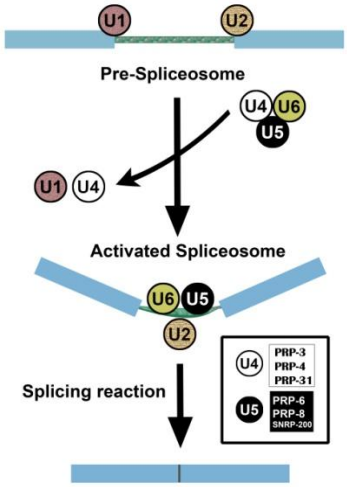

Fig. I. 7. Simplified scheme of the role of s-adRP genes in the splicing process. The six *C. elegans* s-adRP proteins are part of the tri-snRNP U4/U6•U5 complex. After the formation of the activated spliceosome, U5 but not U4 s-adRP proteins are required for subsequent splicing steps involving transesterification reactions. Exons are linked by thinner rectangles that represent an intron. White box: PRP-3, PRP-4 and PRP-31 are U4-specific, while PRP-6, PRP-8 and SNRP-200 are U5-specific.

Thus, a distinct implication in the splicing process of U4 and U5 snRNP becomes evident. While some of the U5 snRNP proteins remain in the activated complex, the U4 snRNP proteins leave the complex and are not required for the catalytic process itself. Still, as common function for all splicing-related adRP genes, their corresponding protein act on the pre-spliceosome assembly. Even when these mutations have been associated with adRP, the

mechanisms that drive the processes leading to the disease are still unknown.

All PRP splicing factors related to adRP are evolutionary conserved (Fig.I.8.) and not only that but the splicing process is essentially identical in *C. elegans* and in vertebrates (Riddle, 1997). This allows us to study the effect of partial loss-of-function of these genes in the process and unravel what mechanisms are responsible of the apoptotic outcome.

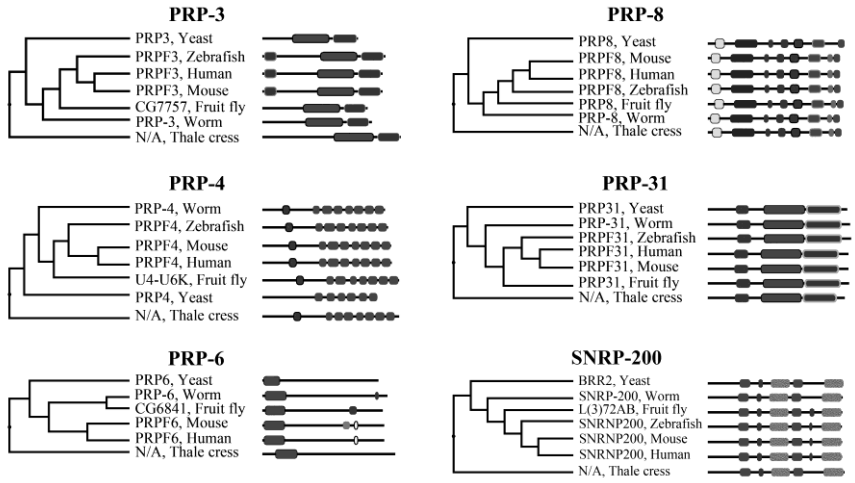

Fig. I. 8. Conservation of s-adRP genes from yeast to human.
Phylogenetic tree of s-adRP genes and schematic representation of protein domains that remain conserved across evolution. Tree and protein domain information retrieved from TreeFam (Schreiber, Patricio, Muffato, Pignatelli, & Bateman, 2014).

I.3. *C. elegans*: a Top Model organism

The nematode, *Rhabditis elegans*, was first identified in the soil by Emile Maupas over a century ago (Blaxter, 2011). Some years after, subsequent phylogenic studies renamed the species *Caenorhabditis elegans* or more commonly known as *C. elegans* (*Caeno* meaning recent; *rhabditis* meaning rod; *elegans* meaning nice). However, it was not until 1974, that Sydney Brenner recognized the potential of this ´nice´ worm as a model organism that could be used to unravel the genetic basics of development in metazoan systems (Brenner, 1974).

In the last four decades, the use of *C. elegans* as a model organism has grown exponentially. Today, it is used to study a large variety of biological processes including apoptosis, cell signaling, cell cycle, cell polarity, gene regulation, metabolism, ageing and sex determination (Kaletta & Hengartner, 2006). The reason of this widespread popularity is that *C. elegans* holds the desirable characteristics to be useful in research being easy to manage and maintain, having a short life cycle, simple genetics, among others. Additionally, its anatomy has been studied in detail and its invariant cell linage is fully known (J. E. Sulston & Horvitz, 1977). Its genome was fully sequenced and became available in 1998 (*C. elegans* Sequencing Consortium, 1998). It was the first multicellular organism sequenced. After that, the genetic and molecular tools that became available turned the worm into one of the most powerful models for research.

Furthermore, the computational tools and the existence of high throughput studies has fostered the development of open-access databases that act as a repository of genomic and curation, holding

data regarding allelic variations, mutant phenotypes, expression patterns during development and proteomics. The central online platform for *C. elegans* information is WormBase (www.wormbase.org).

Many studies of basic biology and biomedicine have revealed a surprisingly strong conservation of molecular and cellular pathways between the worm and mammals. One example was the identification of genes controlling programmed cell death in *C. elegans*, related genes with similar functions were lately identified in humans (Yuan, Shaham, Ledoux, Ellis, & Horvitz, 1993). For such discovery Sydney Brenner, H. Robert Horvitz and John E. Sulston were jointly awarded with The Nobel Prize in Physiology or Medicine in 2002.
Many genes and whole molecular pathways involved in human diseases are being studied in the worm. Some examples are Alzheimer (Levitan & Greenwald, 1995; Sundaram & Greenwald, 1993), Diabetes (Ogg et al., 1997) and depression (Ranganathan, Sawin, Trent, & Horvitz, 2001). This is not surprising since it has been estimated that 40% of genes that are associated with human disease have homologues in *C. elegans* (Culetto & Sattelle, 2000).

I.3.1. *C. elegans* biology

C. elegans is a small nematode of about 1mm in length. Wild strains can be found worldwide in the soil in temperate and humid environments feeding on bacteria present in decomposing plant material (Frézal et al., 2015). In the laboratory, the wild type strain N2, isolated in Bristol (England) is maintained in agar plates seeded with *Escherichia coli* as food source, and kept between 15 and 25°C.

Among the features that make *C. elegans* such a powerful tool for research these are the more relevant:

- Short life cycle

C. elegans life cycle begins with an embryonic stage that lasts around 12 hours. After hatching, its development consists in four main larval stages (L1 to L4), each of them ending with a molt, followed by the adult reproductive stage (Fig. I.9). The whole process is temperature dependent and can be completed in three days at 20°C. In response to unfavorable environmental conditions, late L1 larvae can enter dauer stage and arrest development up to several months (Fig.I.9). Its short generation time and the amount of progeny per worm (300 embryos per self-fertilizing hermaphrodite) enable a large-scale production of animals per day.

- Easy maintenance

C. elegans is easy and cheap to maintain. Techniques for freezing and thawing of living worms are available, which enables long term storage. Since *C. elegans* is non-parasitic, its handling does not represents a risk for the researcher (Riddle, 1997).

- Dual mode of reproduction

C. elegans is a diploid animal with two sexes. The most frequent sex under normal conditions is the hermaphrodite (XX), while males (XO) only appear in a low frequency (0.1%). Hermaphrodites produce sperm and oocytes and can self-fertilize, however, when males are present in the population they mate. The cross-fertilization not only promotes the refreshing of genomes but also allows the

combination of genetic modifications to generate new strains that would be useful to uncover unknown genetic interactions.

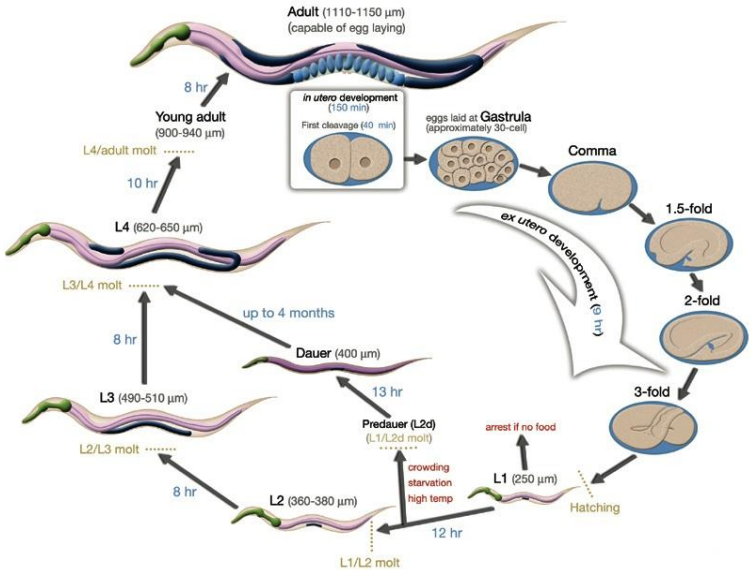

Fig. I. 9. Life cycle of a *C. elegans* hermaphrodite at 22°C.
Temperature-dependent *C. elegans* life cycle takes approximately 3 days at 22°C. 0 minutes corresponds to fertilization. Numbers along the arrows indicate duration of each stage. Size of the worm is expressed in micrometers (µm). Picture from: http://www.wormatlas.org

- Invariable cellular linage

It is a sophisticated multicellular animal with differentiated tissues and organs (muscle, hypodermis, intestine, glands, reproductive and nervous system). The developmental fate of every somatic cell (959 in hermaphrodites and 1031 in males) has been characterized at single cell resolution (J. E. Sulston & Horvitz, 1977; J. E. Sulston, Schierenberg, White, & Thomson, 1983; J. Sulston, 1988). These cell lineages are invariant among individuals, which facilitate the detection of mutations affecting cell division and developmental processes.

- Transparent body

Worm´s body is transparent allowing visualization of internal structures at the cellular and subcellular level by direct visualization with differential interference contrast (DIC). Additionally, fluorescent reporters can also be detected.

- Generation of mutant strains

The fact that *C. elegans* has a short lifecycle and self-fertilization, facilitates the isolation and propagation of new mutants. A large collection of mutant strains are available for general use in the worm community (*Caenorhabditis* Genetic Center, CGC, and National Bioresource Project, NBP). However, powerful systems like MosSCI (Frøkjær-Jensen et al., 2008) and CRISPR/Cas9 (Frøkjær-Jensen, 2013) are also available to generate single copy reporters and mutations *´a la carte´* in one gene of interest.

- Gene silencing by RNA-mediated interference (RNAi)

In *C. elegans*, the function of a specific gene can be disrupted by RNAi. RNAi allows the study of phenotypic effects of the knock down of a gene by the sequence-specific degradation of the target mRNA. This process is usually systemic and transgenerational and can be delivered to the worm by different means. The most common form is by feeding the worms with transformed bacteria expressing the double stranded RNA (dsRNA) of interest; these bacteria clones can be found in two libraries that together cover around 94% of the whole genome (Kamath et al., 2003; J. F. Rual et al., 2004). Alternatively, worms can also be soaked in or injected with a solution containing the dsRNA (Fig. I.10.).

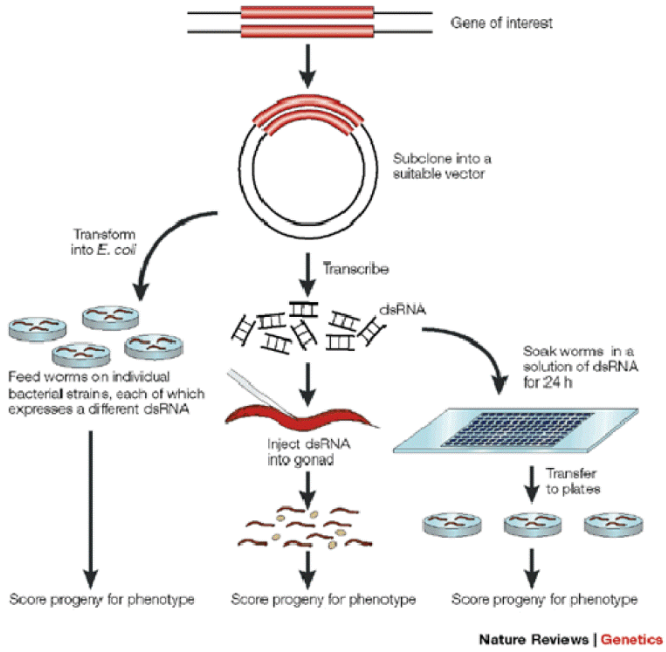

Fig. I. 10. Different protocols to inactivate gene expression by RNAi in *C. elegans.*

From left to right: RNAi by feeding, by microinjection and by soaking (*Taken from* Kim, 2001).

I.4. Silencing gene expression by RNA-mediated interference (RNAi)

RNAi mediated silencing of gene expression is one of the most widely used tools in cell biology nowadays. In *C. elegans*, dsRNA induces a homology-dependent and highly effective decrease in the activity of the corresponding homologous gene, without any evident effect in other genes. Actually, this mechanism was first identified in the worm by Andy Fire, Craig Mello and their colleagues (Fire et al., 1998) and they were awarded with the Nobel Prize in Physiology or Medicine in 2006 for this discovery.

The principle of the technique derives from an evolutionary conserved mechanism of defense against dsRNA from viruses. Briefly, when the dsRNA enters the cell, it is recognized and cleaved by the enzymatic Dicer-RDE complex. The resulting small RNAs are between 21 and 23 bp and are known as small interfering RNAs (siRNAs). Then, the siRNA attaches to an Argonaute protein where the sense strand is degraded while the antisense remains attached and forms, along with other proteins, the RNA-induced silencing complex (RISC). RISC is the complex that recognizes the complementary endogenous mRNA and finally degrades it in small fragments allowing one to mimic a partial loss-of-function mutation in the corresponding gene (Corsi, 2006; Grishok, 2005).

As mentioned before, there are three methods by which dsRNA can be delivered to the worm: injecting or soaking the animals with *in vitro* transcribed dsRNA or, feeding animal with *E. coli* that holds a plasmid that expresses the desired dsRNA (Fig.I.10.).

I.4.1. RNAi by feeding

The two RNAi libraries available for *C. elegans* cover almost the whole genome (Kamath et al., 2003; J. F. Rual et al., 2004). These RNAi libraries consist on plasmids that contain a 500-3000 bp insert of the target mRNA. This inserts have been obtained using different substrates; while Ahringer´s library was obtained by PCR amplification of genomic DNA (Kamath & Ahringer, 2003), Vidal´s library was obtained by cloning full length open reading frames (ORF) from cDNA sequences (J. F. Rual et al., 2004). These plasmids are cloned into the ampicillin resistant L4440 vector, which contains two inverted T7 promoters for bidirectional transcription. This vector is then transformed into the RNase III-deficient *E. coli* strain HT115, which can produce T7 polymerase from an IPTG-inducible promoter (Corsi, 2006).

I.4.2. RNAi by microinjection

In this case, an *in vitro* synthesized dsRNA is introduced directly into the developing gonad of the adult hermaphrodite. Then, the offspring can be scored for mutant phenotypes. The efficiency of this RNAi is stronger than with the other delivery methods.

I.5. Genomic engineering in *C. elegans*

C. elegans was the first multicellular organism to have its whole genome sequenced (*C. elegans* Sequencing Consortium, 1998). Having this information available allowed the development of new approaches to study gene function such as trangenesis (Chalfie, 1994) or directed genome editing techniques like CRISPR/Cas9 (Friedland et al., 2013; Frøkjær-Jensen, 2013) or MosSCI method (Frøkjær-Jensen et al., 2008)

I.5.1. Transgenesis

The introduction of exogenous DNA into *C. elegans* to generate a transgenic strain can be done microinjecting the germ line of self-fertilizing hermaphrodites or by particle bombardment.

C. elegans transgenes obtained following microinjection assemble into multicopy extrachromosomal arrays that behave as artificial chromosomes, as they are efficiently replicated and transmitted to progeny (Kadandale, Chatterjee, & Singson, 2009). To make the transformation more stable, extrachromosomal arrays can be integrated following treatment of a transgene strain with ionizing radiation or chemical mutagenesis.

Microparticle bombardment is usually done on *unc-119* mutant animals (uncoordinated worms that are not capable of moving) that are ´shot´ with DNA-coated beads. By this approach, extrachromosomal and integrated arrays can be obtained (Praitis & Maduro, 2011).

The majority of the markers used to identify transformed worms are easily distinguishable under a scope; most of them contain the rescue of a non-lethal mutation, like *unc-119*, or are strong fluorescent markers.

I.5.2. CRISPR/Cas9: Editing the genome at will

In 2012, a new mechanism was presented to the scientific community with very high expectations on the potential to exploit it for RNA-programmable genome editing: the CRISPR/Cas9 system (Jinek et al., 2012). It was initially identified as a mechanism of immune defense against viruses and plasmids in bacteria and *archaea* (Mojica, Díez-Villaseñor, García-Martínez, & Almendros, 2009), but quickly turned into the most powerful genome editing tool of all times. Just one year after the first CRISPR/Cas9 publication, the system had already been applied and optimized to almost every model organism and even food crops like rice or wheat. *C. elegans* was not an exception and several papers have been published to optimize genome editing by CRISPR/Cas9 in this model (Dickinson & Goldstein, 2016; Paix, Folkmann, Rasoloson, & Seydoux, 2015; Waaijers & Boxem, 2014).

Clustered, regularly interspaced, short palindromic repeats (CRISPR) and CRISPR-associated (Cas) systems are adaptive mechanisms evolved by bacteria and *archaea* to repel invading viruses and plasmids (Mojica et al., 2009). These systems incorporate foreign DNA sequences into host to generate short CRISPR RNAs (crRNAs) that direct sequence-specific cleavage of homologous target double-stranded DNA by Cas endonucleases (Jinek et al., 2012).

Briefly, the system consists of (i) the CRISPR associated endonuclease (Cas9) and (ii) two noncoding RNAs (crRNA and the trans-activating tracrRNA). The specificity of the target is acquired by the recognition of a 20 nucleotides spacer region in the crRNA, complementary to the target DNA sequence. Next to this complementary sequence a 3-nt motif (NGG) known as protospacer adjacent motif (PAM) is necessary to produce the DSB (Gasiunas, Barrangou, Horvath, & Siksnys, 2012; Jinek et al., 2012). Since in CRISPR/Cas9, the only requirement is to localize the PAM sequence in the target region, a wide range of sites can be chosen providing a good access to the entire *C. elegans* genome for engineering. Conveniently, the system was improved by fusing the crRNA to the tracrRNA to form a single guide RNA (sgRNA) that was proven to be sufficient to direct Cas9 to a specific site and generate DSBs (Friedland et al., 2013).

In the *C. elegans* germline, the expression of Cas9/sgRNA system produces DSB that can be repaired by two methods: (i) non-homologous end joining (NHEJ), an error-prone process that can create insertions, deletions or mutations at the cut site, and (ii) homology-directed repair (HDR), a precise mechanism that repairs the break using a homologous donor molecule (Fig.I.11).

Thus, to generate specific mutations a donor molecule that carries the desired edit flanked by homologous sequences called "homology arms" should be incorporated. Then, these edits could be integrated as part of the repair process (Paix et al., 2015; Waaijers & Boxem, 2014). Using this strategy it is possible to generate deletions of different lengths, endogenous tagging, single modifications or insertions anywhere in the genome.

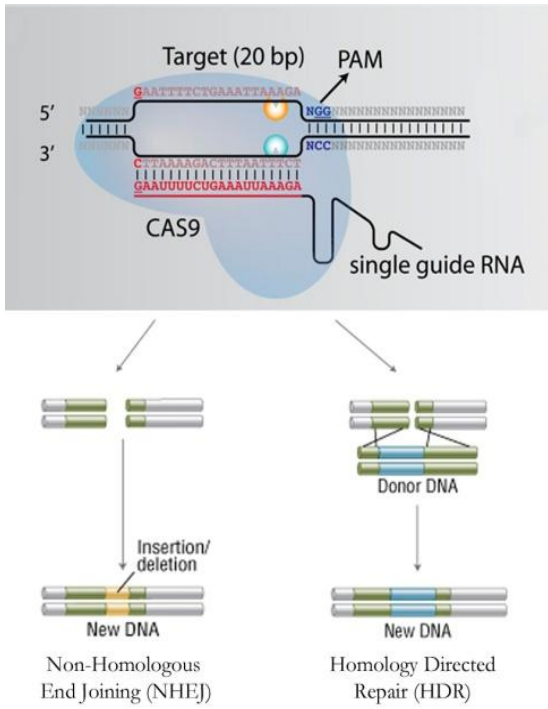

Fig. I. 11. Cas9/sgRNA in complex with a target site.

Cas9 generates a double strand break. The PAM sequence in the genome is in blue, the 20 nucleotide recognition sequence in red, and sgRNA sequence is indicated. Cas9 nuclease site specifically cleaves double-stranded DNA (dsDNA) activating double-strand break repair machinery. In the absence of a homologous repair template non-homologous end joining (NHEJ) can result in indels disrupting the target sequence. Alternatively, precise mutations and knock-ins can be made by providing a homologous repair template and exploiting the homology directed repair (HDR) pathway. (*Modified from* Frøkjær-Jensen, 2013).

I.5. *C. elegans* seam cells

The hypodermis in *C. elegans* embryos is the outer monolayer of 78 epithelial cells that secrete the components of the cuticle (J. E. Sulston et al., 1983). One of its functions is to establish the basic body architecture of the worm.

At the beginning of the embryonic development the larger hypodermal cells (65 cells) are arranged in six parallel rows: two central rows that shape the dorsal hypodermis, two lateral rows that will form the seam cells and two external rows that will become the ventral hypodermis (P cells) (Fig.I.12A). The remaining small hypodermal cells belong to the head and tail of the worms. Shortly after, cells of the central row migrate to interlace forming a single row of dorsal cells.

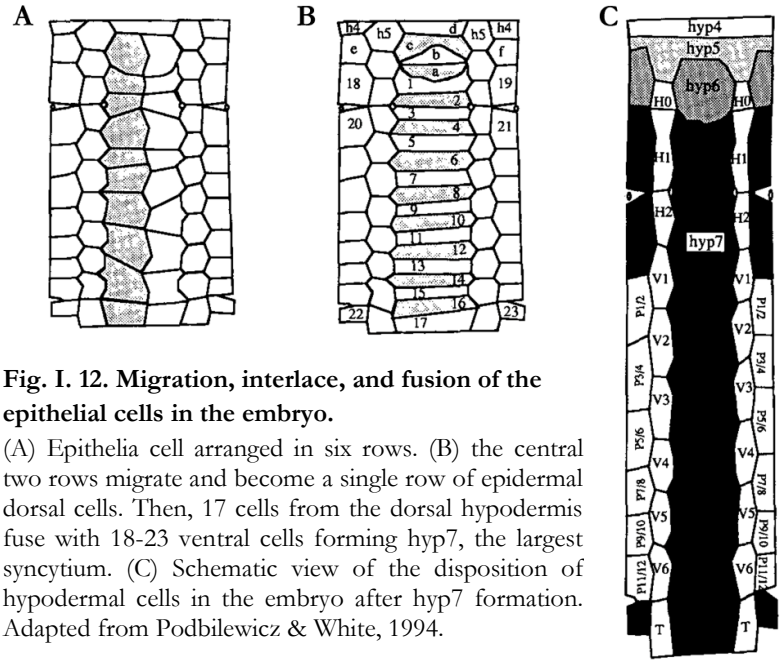

Fig. I. 12. Migration, interlace, and fusion of the epithelial cells in the embryo.
(A) Epithelia cell arranged in six rows. (B) the central two rows migrate and become a single row of epidermal dorsal cells. Then, 17 cells from the dorsal hypodermis fuse with 18-23 ventral cells forming hyp7, the largest syncytium. (C) Schematic view of the disposition of hypodermal cells in the embryo after hyp7 formation. Adapted from Podbilewicz & White, 1994.

Then the larger cell in the hypodermis, known as hyp7, is formed by the fusion of 23 cells, most of them, cells in this central dorsal row (1-23 in Fig.I.12.B). This 23-nucleate hyp7 syncytium covers most of the dorsal area of the embryo and larval stage 1 (L1). Syncytia are multinucleate cells formed by cell-to-cell fusions and one third of all the nuclei of an adult worm reside in these multinucleate compartments. During post embryonic development additional 110 cells, product of divisions of the lateral seam cells and ventral cells, join hyp7 forming the largest syncytium in the adult body.

At the beginning of L1, each lateral row of the larva holds 10 seam cells (H0-H2, V1-V6 and T, Fig.12C). These cells are embedded in the hypodermal syncytium hyp7 and are also in close contact with ventral epithelial cells, the P cells. Most of these seam cells, but H0, act as stem cells and continue dividing through development. The posterior daughter cell remains a seam cell while the anterior daughter cell differentiates, with the exception of H1 where it happens the other way around. After division of H1, V1-V4 and V6, the anterior cell differentiates and fuses with the epidermal syncytium hyp7, while H2, V5 and T become either a neuron or a neural support cell (Podbilewicz & White, 1994; J. E. Sulston & Horvitz, 1977; Wildwater, Sander, de Vreede, & van den Heuvel, 2011).

At larval stage 2 (L2), a proliferative division occurs and the number of seam cells in each lateral row of the worm is increased to 16 cells (Fig.I.13). Then, the 16 row pattern is maintained and divisions are carried out until larval stage L4, so that resulting anterior cells keep joining hyp7.

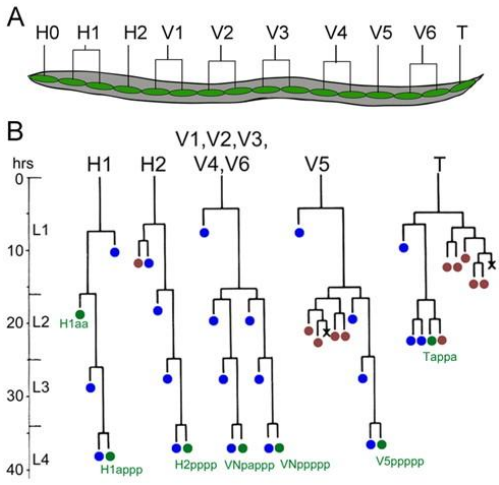

Fig. I. 13. Seam cells organization.
(A) From L2 and on, 16 Seam cells (green) are aligned in a linear row on each lateral side of *C. elegans* larvae.
(B) Postembryonic seam cell lineages. y-axis indicates time (hours) of development. Circle color indicates fate: green, seam cells; blue, hyp7 fusion fate; red, neuronal fate; cross, apoptosis. V1-V4 and V6 undergo a proliferative division in early L2.

Thus, hypodermal seam cells play a stellar role in *C. elegans´* body constitution, not only because they are part of the hyp7 syncytium, but also because they participate in the secretion of the cuticle components.

I.5.1. The cuticle

The cuticle of the worm is an exoskeleton composed of collagen-like molecules. It is secreted by underlying epithelial cells, known as the hypodermis and seam cells. The cuticle provides a relatively impermeable barrier between the worm and its environment, it determines the shape of the worm, having an appropriate balance between flexibility and rigidity to allow proper worm movement (AP. Page, 2007; Johnstone, 1994).

During development the worm synthesizes 5 cuticles, one within the egg prior to hatching which constitutes the cuticle of L1, and the

remaining four during the molt at the end of each of the four larva1 stages (Fig.I.14). Molting is a temporally reiterated process driven in a cyclical fashion and it is synchronized with hypodermal seam cells division. They seem to play a central role in the process of cuticle synthesis as it has being observed that these cells swallow and shows signs of synthetic activity during this period (Frand, Russel, & Ruvkun, 2005; Johnstone, 1994; Singh & Sulston, 1978).

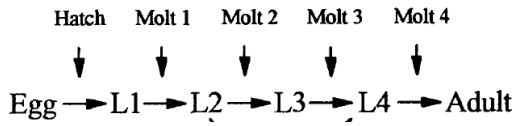

Fig. I. 14. Molting points during worm´s life cycle.
Five cuticles are synthesized during development. The first one prior hatching and then, one after each developmental stage Adapted from Johnstone, 1994.

Cuticle formation in each larval stage represents a high molecular activity process as it requires the synthesis of a great amount of proteins each time. This has been observed in pulse-labeling and pulse-chase-labeling experiments in which synthesis of cuticle components was detected at high levels during the molting periods and at much reduced rates during the inter-molt periods (Cox, Kusch, DeNevi, & Edgar, 1981). Additionally, the abundance of collagen mRNA during molting cycle and peaks prior to synthesis has also been detected (Cox & Hirsh, 1985). Thus the synthesis of the cuticle is not only spatially restricted to the hypodermis, but occurs in a temporally regulated fashion with the majority of cuticular proteins being synthesized in a short period of time immediately prior to, and during, the deposition of each cuticle (Johnstone, 1994).

I.6. Apoptosis in *C. elegans*

C. elegans cell lineages are invariant, and so is the number of cells that undergo apoptosis during development. Apoptotic cell death is under genetic control and the genes implicated in this molecular program are conserved throughout evolution (Ellis & Horvitz, 1986; Metzstein, Stanfield, & Horvitz, 1998). Deregulation of this process is associated with several diseases so a better understanding of it is of great interest.

Programmed cell death in *C. elegans* occurs during two stages of its life and in two different types of tissues: during embryonic and postembryonic development of the soma (developmental cell death) (J. E. Sulston et al., 1983), and in the gonad of adult hermaphrodites (germ cell death) (Gumienny, Lambie, Hartwieg, Horvitz, & Hengartner, 1999; J. Sulston, 1988; White, 1988).

I.6.1. Developmental cell death

During development, around 10% (131 cells) out of the 1090 cells that are generated to form an adult hermaphrodite undergo apoptosis in an invariant pattern. First, 113 cells die in the first steps of embryonic development (J. E. Sulston et al., 1983). Then, during post-embryonic development 18 neuronal cells are removed at L2 stage, being the last somatic cells that undergo apoptosis.

Somatic cell death in *C. elegans* is regulated by four genes: three death-promoting genes, *egl-1*, *ced-3*, and *ced-4*; and the anti-apoptotic gene *ced-9*, which protects cells from undergoing programmed cell death during *C. elegans* development (Hengartner, Ellis, & Horvitz, 1992). *ced-3*, *ced-4*, *egl-1*,and *ced-9* appear to act in a simple genetic pathway in which *egl-*

1 acts upstream of *ced-9* to induce cell death, *ced-9* acts upstream of *ced-4* to inhibit cell death, and *ced-4* acts upstream of *ced-3* to kill cells (Fig.I. 15) (Barbara Conradt & Horvitz, 1998; Hengartner et al., 1992; Shaham & Horvitz, 1996)

These are all conserved apoptotic regulators being *ced-3* a protease of the caspase family (Yuan et al., 1993) and *ced-4* an adaptor protein similar to mammalian apoptotic protease-activating factor-1 (APAF1) (Yuan & Horvitz, 1992). *ced-9* and *egl-1* encode members of the Bcl-2 family, being *ced-9* an anti-apoptotic gene, and EGL-1 a pro-apoptotic BH3-only-domain protein (Barbara Conradt & Horvitz, 1998; Hengartner & Horvitz, 1994).

egl-1 is the key activator of the activation phase of apoptotic cell death and it is regulated at the transcriptional level. Thus, in cells that are programmed to die *egl-1* activity is high, while in cell that are programmed to life, *egl-1* remains absent. High *egl-1* activity inhibits *ced-9* activity, resulting in the activation of *ced-4* and *ced-3* and the induction of the execution phase of apoptotic cell death (Horvitz, 2003).

I.6.2. Physiological germ line apoptosis

The soma and germ line use a common apoptotic execution machinery. However, these two types of tissues use different regulatory mechanisms to control the activation of this machinery. The majority of germ cells generated in the hermaphrodite germ line are eliminated by apoptotic cell death. This physiological germ cell

death can be suppressed by *ced-9* and is dependent on *ced-4* and *ced-3*, but in an *egl-1* independent manner (Gumienny et al., 1999). What triggers this physiological germ cell death remains to be determined, but recent evidence suggests that the transcriptional regulation of components of the central apoptosis machinery, namely *ced-9*, *ced-4* and *ced-3*, might play an important role (Nehme & Conradt, 2009).

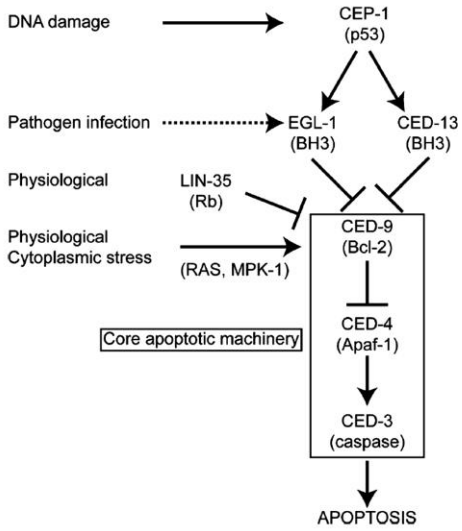

Fig. I. 15. *C. elegans'* apoptotic pathway.
Core apoptotic machinery required for physiological apoptosis is represented. In the germ line, apoptosis can also be induced by external factors like pathogen infection, stress or DNA damage (*Adapted from Gartner et al, 2008*).

I.6.3. DNA damage induced apoptosis

C. elegans´ germ cells are capable to undergo apoptosis in response to bacterial infections or genotoxic agents (Gartner, Boag, & Blackwell, 2008). In contrast to constitutive germ cell death, pathogen and DNA damage-induced germ cell death is not only dependent on *ced-4* and

ced-3 but also on *egl-1*. The control of transcription of *egl-1* upon DNA damage is carried on by CEP-1, the *C. elegans* ortholog of mammalian p53 damage (Schumacher, Hofmann, Boulton, & Gartner, 2001). Upstream of this central death machinery, DNA damage response (DDR) components lie and act either as sensors of the lesion or as transducers of the initial signal detected (Stergiou & Hengartner, 2004).

Studies regarding the regulation of DDR activation in *C. elegans* have shown that upon induction of DNA damage by some genotoxic agents (IR, cisplatin or etoposide), somatic cells do not activate DDR and that most DDR factors are not expressed in *C. elegans* soma. Furthermore, they conclude that worm somatic cells do not show any apparent response to DNA-damaging agents and do not undergo apoptosis. However, the also observed that somatic cells were capable to repair DNA damage in a DDR independent manner, so they do not rule out the existence of additional, yet uncharacterized, signal pathways in somatic cells (Vermezovic, Stergiou, Hengartner, & d'Adda di Fagagna, 2012).

I.7. R-loops: a double-edged sword structure

R-loops are three-stranded nucleic acid structures composed of an RNA-DNA hybrid and a displaced ssDNA (Hamperl & Cimprich, 2014). They are thought to form co-transcriptionally when nascent messenger RNA hybridizes with the DNA template. These structures were first identified in bacteria, and are found in organisms from yeast to mammals where they seem to participate in regulation of several cellular processes (Chan et al., 2014; Drolet, Bi, & Liu, 1994; El Hage, Webb, Kerr, Tollervey, & Andujar, 2014).

I.7.1. R-loops as a regulatory structure

Recent experiments suggest that R-loops have a role in regulation of gene expression. It has been shown that they are usually formed at vertebrate CpG-islands at promoter regions generating stable structures. Then, R-loops can act as a positive regulator of gene expression as it protects the DNA from *de novo* DNA methylations, an epigenetic mark associated with transcriptional silencing (Ginno, Lott, Christensen, Korf, & Chédin, 2012). However, they are also found in organisms as *C. elegans* that does not have CpG island or methylation (Deaton & Bird, 2011). R-loops can also act as regulators of transcription termination when they are found at the termination regions of particular genes. There, R-loops recruit the RNA/DNA helicase senataxin (SETX) which finally resolves the R-loop structures ending transcription (Skourti-Stathaki, Proudfoot, & Gromak, 2011).

I.7.2. R-loops and DNA damage

R-loops have also been associated with induction of DNA damage and genomic instability (Aguilera & García-Muse, 2012; Castellano-Pozo, García-Muse, & Aguilera, 2012). Thus, their formation and resolution are processes that must be regulated to maintain the equilibrium that prevents R-loop formation to become a threat.

Defects in factors that resolve or prevent R-loops can lead to DNA damage and genomic instability. For example, double strand breaks (DSB) can be produced by the transcription-coupled nucleotide excision repair pathway (TC-NER) when R-loops are generated during transcription as consequence of missing RNA-processing factors or topoisomerase I (Sollier et al., 2014).

Other endogenous mechanisms can act over R-loops regions as a mistake and produce DNA damage. This is the case of the activation-induced cytidine deaminase (AID) which initiates class switch recombination (CSR) and somatic hypermutation (SHM) in B lymphocytes. AID can act over ssDNA in R-loops structures and make them susceptible to excision repair mechanisms that will finally produce a DNA lesion (Basu et al., 2011). Additionally, R-loops can represent a physical barrier to replication forks generating DSB by collision with the replication machinery (Gan et al., 2011).

Further work is needed to understand the respective contributions of the replication fork, AID, and TC-NER factors to R-loop-induced DSB formation and genome instability. Deciphering when and how each of these processes act, will provide a better understanding of the

mechanisms underlying R-loop-induced genome instability and their potential contributions to human disease.

I.7.3. Mechanisms of protection against R-loops

As mentioned above, an excess of R-loops can be harmful to the cell. Therefore, several mechanisms are known to resolve R-loops or prevent their formation (Hamperl & Cimprich, 2014). One of them is RNase H, which resolves R-loops specifically degrading the RNA moiety in RNA–DNA hybrids (Wahba, Amon, Koshland, & Vuica-Ross, 2011). They can also be resolved by the RNA/DNA helicase activity of proteins like SETX or aquarius (AQR). Depletion of human SETX and mutations in its yeast homolog, Sen1, have shown accumulation of R-loops (Mischo et al., 2011; Skourti-Stathaki et al., 2011). An accumulation of R-loops was also recently detected in the nucleus of cells depleted of the putative RNA/DNA helicase AQR, however, where they accumulate in the genome is not known (Sollier et al., 2014). Techniques such as DNA:RNA immunoprecipitation sequencing (DRIP-seq) are helping to identify these R-loops harboring sequences.

During transcription, R-loop formation may also be prevented by the action of topoisomerase I, which resolves negative torsional stress behind the RNA Pol II (Tuduri et al., 2009). Additionally, RNA-processing and RNA-export factors can also avoid R-loop formation by binding to the nascent RNA as it emerges from RNA Pol II making it unavailable to interact with the DNA (Huertas & Aguilera, 2003; X. Li, Niu, & Manley, 2007).

OBJECTIVES

I. **Establish the nematode *Caenorhabditis elegans* as a model for splicing-related autosomal dominant retinitis pigmentosa (adRP).**

II. **Identify the mechanisms behind the cell-type-specific apoptosis caused by mutations in essential splicing factor genes related to adRP.**

III. **Utilization of CRISPR/Cas9 system to introduce s-adRP mutations into the genome of *C. elegans.***

RESULTS

R.1. Splicing factors involved in adRP are essential for the development and fertility of *C. elegans*

The essential function of s-adRP genes has been a major obstacle in the use of animal models to study retinitis pigmentosa (RP). In any multicellular organism the complete inactivation of s-adRP genes leads to lethality. Therefore it is not surprising that mutations in splicing-related genes associated with adRP are either partial-loss-of-function mutations, mostly missense, or deletion alleles that cause haploinsufficiency.

C. elegans allows studying the consequences of a partial decrease in the normal amount of a desired transcript/protein by using RNA-mediated-interference (RNAi) by feeding. In this way, we can mimic in worms the partial-loss-of-function of s-adRP genes in RP patients. RNAi clones corresponding to 5 of the 8 s-adRP genes were obtained, and validated by sequencing, from the two existing *C. elegans* RNAi libraries (Kamath et al., 2003; J.-F. Rual et al., 2004). In the case of *snrp-200*, the RNAi clone was generated in our laboratory by Gateway cloning (Fig. R.1A).

To assess how efficient was the inactivation of the targeted genes under our RNAi conditions, we used quantitative PCR to quantify the levels of mRNA after RNAi treatment and found that more than 50% of the targeted transcripts were depleted (Fig. R.1B).

There are deletion mutants available for *prp-31* and *prp-8* that were generated by the *C. elegans* Deletion Mutant Consortium (2012) that

would be good tools to study haploinsufficiency in worms. *prp-31(gk1094)* allele consists of a 5 bp insertion/1953 bp deletion and *prp-8(gk3511*) is a 1823 bp deletion (Fig. R.1A). These are not viable mutations when they are in homozygosis (larval arrest (Lva)), but still valuable to evaluate if heterozygous animals are haploinsufficient.

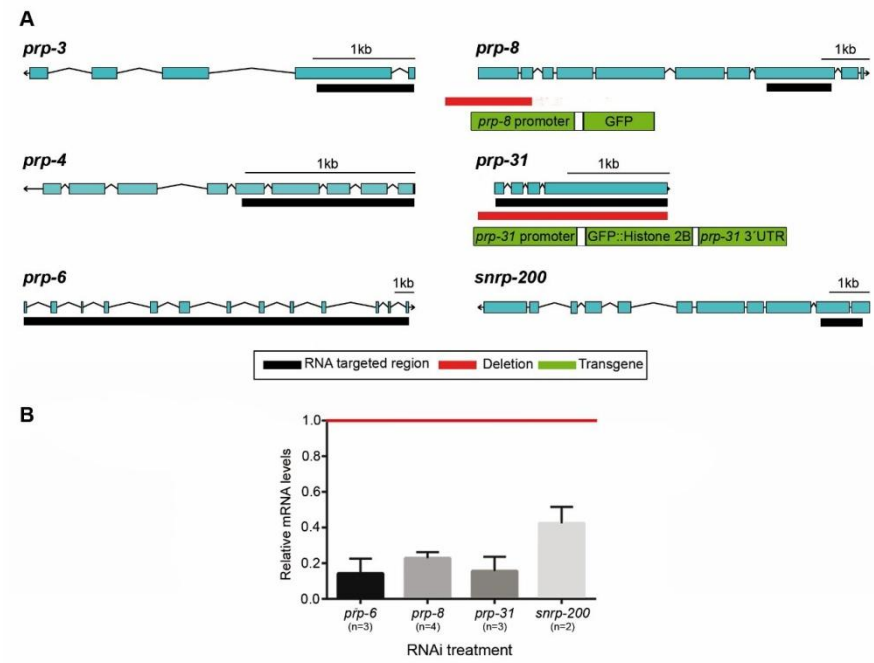

Fig. R. 1. *C. elegans* tool kit 1.0 to study adRP.
(A) Scheme of s-adRP genes in *C. elegans* including the regions that are targeted by RNAi clones (black bars), the deleted fragment in the *prp-31(gk1094)* and *prp-8(gk3511)* alleles (red bars) and the elements of the *prp-8* and *prp-31* reporters generated in this study (green rectangles). (B) Quantification of *prp* genes expression levels after RNAi inactivation. mRNA levels of *prp* genes are represented relative to their expression in *gfp(RNAi)* control animals (arbitrary value of 1, indicated with a red line). Transcript levels are normalized against *tbb-2* levels in each case. RNA for analysis was obtained from up to four biological replicates (n). RNA from samples used for RNA-Seq analyses were included. Error bars represent standard error of the mean.

R.1.1. RNAi of s-adRP splicing factors reveals a correlation between the function of the gene and severity of the phenotype

To study the phenotypes of worms with a partial depletion of s-adRP genes, we used RNAi by feeding. Wild type N2 worms were synchronized following the sodium hypochloride treatment (Porta-de-la-Riva, Fontrodona, Villanueva, & Cerón, 2012) and L1 animals were grown at 20°C and fed with RNAi clones for *prp-3*, *prp-4*, *prp-6*, *prp-8*, *prp-31* and *snrp-200*, and *gfp* dsRNA as control. The partial reduction of the activity of these genes produced phenotypes that allowed us to separate them in two different phenotypic clusters.

We found that RNAi of genes belonging to the U5 snRNP complex (*prp-6*, *prp-8* and *snrp-200*) produced larval arrest, on the other hand, animals treated with RNAi of genes belonging to the U4 snRNP complex (*prp-3*, *prp-4* and *prp-31*) could develop and reach the adult stage but they were sterile (Table R. 1, Fig. R. 2).

	RNAi by feeding	**RNAi by injection**
U4 protein coding genes		
prp-3	Sterility	Embryonic lethality
prp-4	Sterility	Embryonic lethality
prp-31	Sterility	Embryonic lethality
U5 protein coding genes		
prp-6	Larval arrest	Embryonic lethality[a]
prp-8	Larval arrest	Embryonic lethality[a]
snrp-200	Larval arrest	Embryonic lethality[a]

[a]Information obtained from the *C. elegans* PhenoBank (Sönnichsen et al. 2005)

Table R. 1. Phenotypes observed after RNAi of s-adRP genes by two different methods.

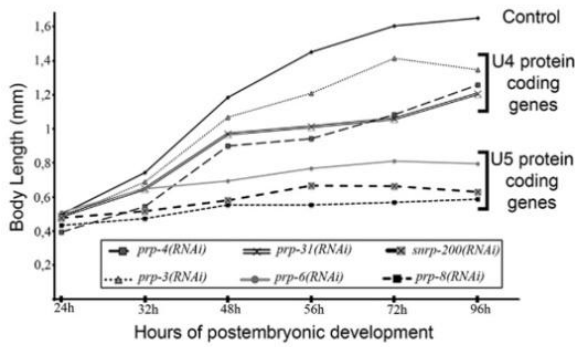

Fig. R. 2 RNAi experiments classify s-adRP genes in two phenoclusters.
Developmental growth in animals treated with RNAi for s-adRP genes. Body length of wild type worms fed with the indicated RNAi clone, starting from synchronized L1 and grown at 20°C. Mean body length of 50 worms per RNAi condition was scored at the indicated time points. Animal length was measured using the Image J software. Control worms were fed with *E. coli* containing the L4440 plasmid (control).

As mentioned in the introduction, all six proteins are components of the U4/U6·U5 tri-snRNP complex and participate in the assembly of the spliceosome, but only PRP-6, PRP-8 and SNRP-200, the U5 snRNP proteins, remain in the activated spliceosome (Fig. I. 7). Thus, the bigger presence of U5 snRNP proteins within the splicing process could be the reason for the stronger phenotype we observe upon RNAi.

RNAi by injection produces a strong gene inhibition from worm to worm as its efficiency is higher than other RNAi delivery methods. A previous large-scale RNAi by injection study included all six s-adRP genes and showed a common embryonic lethality (Emb) phenotype (Sönnichsen et al., 2005). We wanted to reproduce the experiment and to do this, we synthesized and microinjected dsRNA for *prp-3*, *prp-4* and *prp-31* and we were able to confirm the Emb phenotype.

All the observed phenotypes confirmed that normal levels of each of these s-adRP genes are essential for the viability of the animal.

R.1.2. Deletion alleles in *prp-8* and *prp-31* do not cause haploinsufficiency in germline or soma development

For some genes, deletion of one functional copy from a diploid genome causes a phenotype or a disease. These genes are called haploinsufficient because a single copy of them is insufficient to maintain normal functions (Dang, Kassahn, Marcos, & Ragan, 2008). We investigated haploinsufficiency in *C. elegans* by using deletion alleles for *prp-31* and *prp-8*. We used strains in which one copy of *prp-31(gk1094)* or *prp-8(gk3511)* allele is balanced with the translocation hT2 [*bli-4(e937) let-?(q782)* qIs48] (I;III) or with classic genetic markers (*unc-57(ad592) dpy(e61)I* for *prp-31*) to study their progeny. We found that these deletions did not produce haploinsufficiency in terms of fertility or embryonic development (Fig. R. 3A).

We further explored other signs of haploinsufficiency in worms carrying the *prp-31* deletion by investigating different penetrance among individuals. We monitored the progeny of *prp-31(gk1094)* heterozygous individuals and observed no interindividual variability in terms of brood size (Fig. R. 3B). Moreover, the mendelian proportions in the offspring of heterozygous animals were normal, with ≈1/4 of arrested larvae (*prp-31(gk1094)* homozygous) (Fig. R. 3C).

In conclusion, *prp-31(gk1094)* and *prp-8(gk3511)* are haplosufficient in terms of somatic and germline development, although we cannot discard haploinsufficiency in any other specific cellular process. Strains carrying these mutations could be useful to screen for enhancers, and therefore for potential modifiers of the disease.

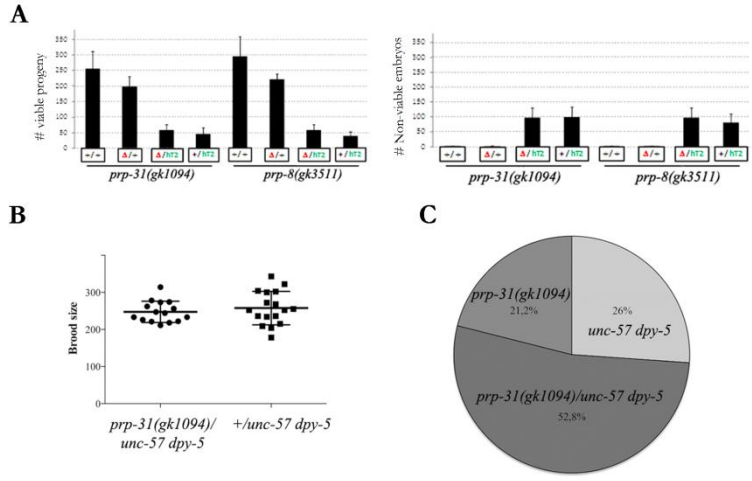

Fig. R. 3. *prp-31(gk1094)* and *prp-8(gk3511)* heterozygous animals are happlosufficient in terms of fertility and embryonic development.
(A) Progeny and embryonic lethality of at least 8 worms of the indicated genotypes was scored at 15ºC. Heterozygous worms for both alleles (without the balancer) show a slight reduction in the number of viable progeny. Additionally, the hT2 balancer by itself is deleterious for the survival producing fewer progeny and embryonic lethality. So the use of the alleles without the balancer is recommended for careful examination of the mutant phenotype. Bars represent the standard deviation. (B) Brood size of heterozygous *prp-31(gk1094)* worms balanced with the two genetic markers *unc-57(ad592)* and *dpy-5(e61)* compared to control worms. Progeny of 15 (dots) and 17 (squares) worms was scored at 25 °C for each genotype. Bars represent the standard deviation. (C) Distribution of genotypic proportions among the progeny of heterozygous *prp-31(gk1094); unc-57(ad592) dpy-5(e61)* animals. Proportions represent the average of the progeny of five worms counted at 25 °C. Similar Mendelian inheritance was observed at 15 °C.

R.2. Generation of transgenic strains

Being *C. elegans* a transparent worm it is possible to visualize the expression of fluorescent proteins under the control of regulatory regions of a gene of interest. We studied the expression of s-adRP genes using transgenic reporter strains with Green Fluorescent Protein (GFP).

We used a transcriptional reporter strain for *prp-8* that was available with the extrachromosomal transgene sEx12486[rCesC50C3.6::GFP + pCeh361]. In this strain the *prp-8* promoter (1052bp; Promoterome Database WorfDB http://worfdb.dfci.harvard.edu/promoteromedb/) controls the transcription of GFP. We studied GFP transcription in these worms and observed a ubiquitous expression (Fig. R. 4A) (Hunt-Newbury et al., 2007).

This transgene, although informative, did not contain all the regulatory regions of the gene and moreover was not integrated in the worm´s genome resulting in mosaicism in its expression. Thus, we made several attempts to express this protein tagged with GFP, but we were not able to obtain transgenic animals by these means, suggesting that ectopic expression of these proteins could be toxic for animal development.

We also generated a transcriptional reporter containing the promoter region of *prp-31* in frame with the nuclear histone 2B::GFP sequence and the 3'UTR of *prp-31*. Animals were transformed by gene bombardment or microinjection and we obtained several strains carrying this reporter either integrated in the genome or as an extrachromosomal array. All these strains broadly expressed the transgenes along the animal providing a useful genetic background to

screen for other genes or small molecules capable of regulating the expression levels of *prp-31* (Fig. R. 4B).

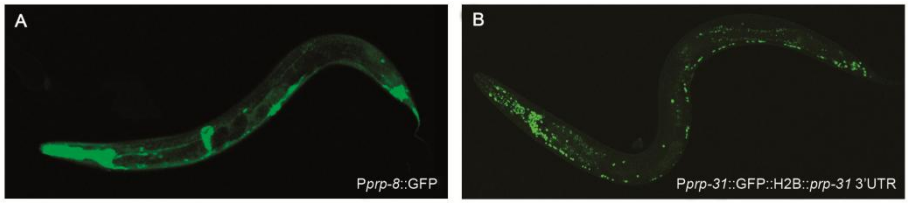

Fig. R. 4. Transgenic reporter strains of *prp-8* and *prp-31*.
(A) Representative confocal image showing a transgenic worm expressing the transgene *sEx12486*, which consists of GFP under the control of a promoter region for *prp-8*. (B) Representative confocal image showing a transgenic worm expressing the transgene *cerEx79*, which consists of the fusion protein GFP::H2B under the control of *prp-31* promoter and 3'UTR.

R.3. Transcriptomic analysis of worms deficient for s-adRP genes

To study the molecular consequences of a partial inactivation of s-adRP genes, we analyzed the transcriptomes of animals treated with RNAi against *prp-6*, *prp-8* and *prp-31*. L1 animals were fed with the corresponding RNAi clone and harvested for RNA isolation after 24 hours at 20°C, before developmental alterations due to RNAi became perceptible.

R.3.1. s-adRP RNAi produce low intron retention

A major concern when dealing with a deficiency in constitutive splicing during development related genes is the failure of the splicing process *per se* that would be evident in the accumulation of introns or aberrant pre-mRNA molecules. In order to estimate the global levels of intron retention, we quantified the proportion of reads (read length 73 bp) mapping in intronic sequences in experimental and control samples. To ensure that our analysis was detecting intron retention and not stabilization of excised introns, we scored the proportion of reads mapping within introns and reads mapping in exon-intron junctions (Fig. R. 5).

Mild intron retention levels were observed in developing animals treated with RNAi against these three s-adRP genes. The strongest levels of intron retention were detected in *prp-8(RNAi)* samples. More precisely, 6.49% and 4.07% of the reads mapped in exon-intron junctions in *prp-8(RNAi)* and *gfp(RNAi)* animals respectively. These results indicate a basal level of unspliced transcripts in control animals, suggesting that the splicing machinery is not 100% efficient in the

excision of introns in normal conditions. We corroborated the increase of intron retention events in *prp-8(RNAi)* animals by semiquantitative RT-PCR (Fig. R. 6A) and by observing the amount of exon-intron border and intronic reads of some pre-mRNAs in a Genome Browser (Fig R. 6BC) (Rubio-Peña et al., 2015).

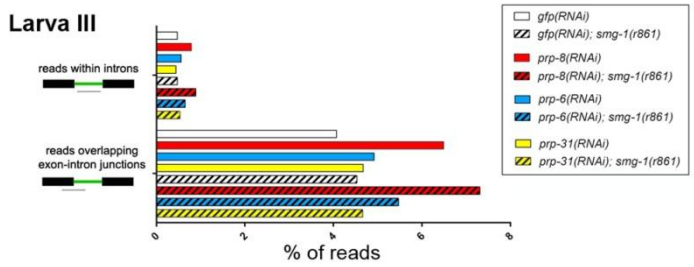

Fig. R. 5. Intron retention caused by RNAi of s-adRP genes in wild type and NMD mutants developing worms.
Proportion of reads (73 bp) mapping within intron or in the exon-intron border in L3 N2 and *smg-1* mutants upon the indicated RNAi treatment.

By considering only RNA-seq reads that mapped in annotated introns, we identified 69 statistically significant (*p-value*≤ 0,01) common intron retention events in the whole genome of L3 animals treated with RNAi against s-adRP genes (Supplemental table S.1 in Rubio-Peña et al, 2015). We did not detect any common characteristic of the retained introns in terms of length of the gene (ranging from 114 to 4506 bp) or position within the gene (Rubio-Peña et al., 2015).

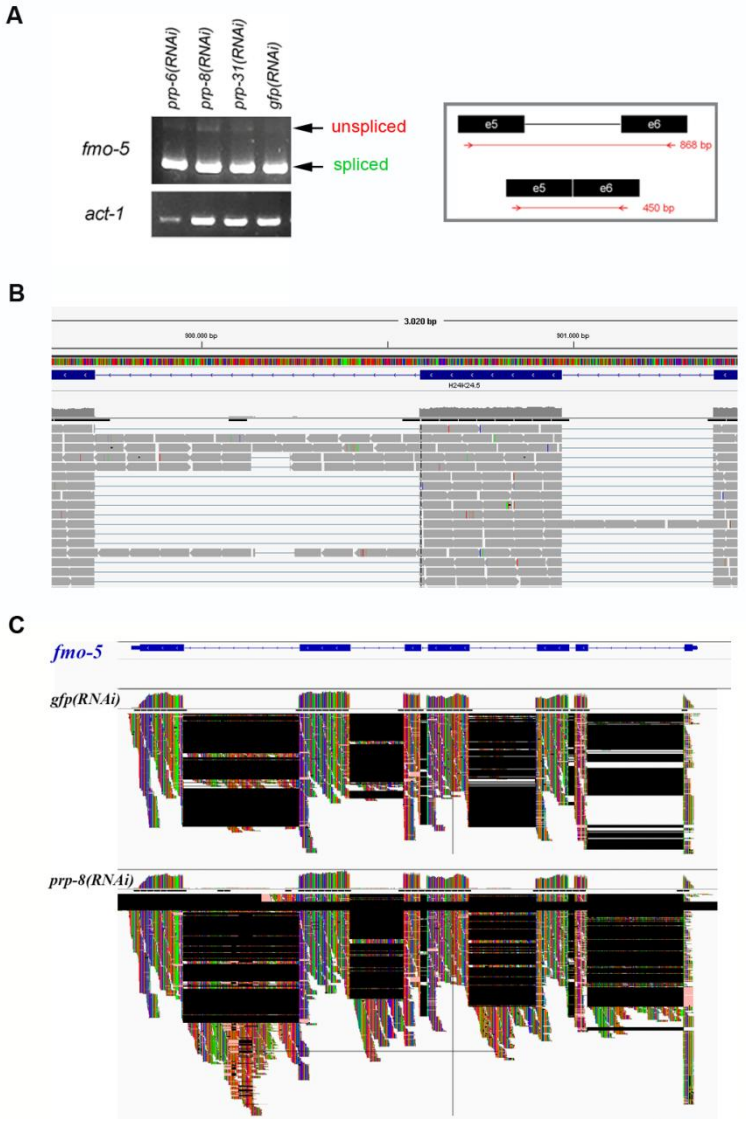

Fig. R. 6. Intron retention in *fmo-5* mRNA.

(A) Semiquantitative RT-PCR was used to amplify the intron indicated in the box. We detected a slight accumulation of unspliced transcripts after RNAi of s-adRP genes. (B) Screen capture of the IGV genome browser at H24K24.5 gene (*fmo-5*) showing RNA-Seq reads corresponding to the *prp-6(RNAi)* sample. Grey boxes represent the 73 bp reads. Blue lines link reads mapping to two contiguous exons. The picture shows reads mapping in exon-intron borders. (C) View of the IGV genome browser at H24K24.5 gene (*fmo-5*) showing RNA-Seq reads corresponding to *gfp(RNAi)* and *prp-8(RNAi)* samples. Black regions indicate absence of reads. Both samples display similar *fmo-5* expression levels, but the *prp-8(RNAi)* sample presents more reads in intronic regions.

R.3.2. Nonsense-Mediated-Decay does not mask higher intron retention.

The Nonsense-Mediated-Decay (NMD) machinery prevents the expression of truncated proteins by degrading transcripts with premature termination codons (PTCs) (Chang, Imam, & Wilkinson, 2007; Hodgkin, Papp, Pulak, Ambros, & Anderson, 1989). Previous transcriptomic studies estimated that 20% of *C. elegans* genes produce transcripts that are degraded by the NMD pathway and many of those transcripts arise from splicing errors as retention of introns or wrong splice site selection (Arun K Ramani et al., 2009). Therefore, in our RNAi experiments the NMD system could be masking a stronger effect on intron retention.

To address this concern we performed the same transcriptomic studies in the NMD defective strain *smg-1(r861)*. Although the proportion of reads on introns was slightly higher in all *smg-1(r861)* samples, results were comparable with those obtained in wild type animals and we did not observe a more dramatic effect on intron retention in NMD mutants upon RNAi of s-adRP genes (Fig. R. 5).

R.3.3. The observed intron retention could be related to transcriptional activity

During development, cells have higher transcriptional activity than postmitotic cells. We wanted to know if intron retention levels would be reduced in an adult animal where transcription in somatic tissues is lower. To do this, we needed to get rid of the transcriptional noise that a very active germline would provide. Thus, we used an abnormal germ line proliferation (*glp*) mutant, *glp-4(bn2)*, which is completely devoid of germline at the restrictive temperature of 25°C.

The transcriptome of 5-days-old germline-less animals was analyzed. Interestingly, we did not observe the mild intron retention associated with *prp-8(RNAi)*, probably because of the much more reduced levels of transcriptional activity in adult animals compared to animals during development (Fig. R. 7).

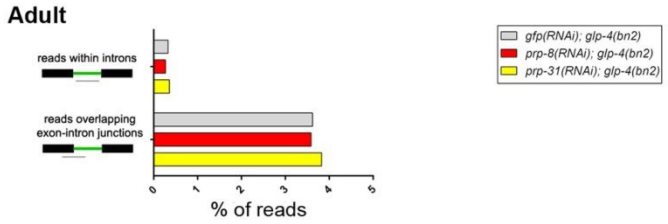

Fig. R. 7. Intron Retention caused by RNAi of s-adRP genes in germline-less adult worms.
Proportion of reads (73bp) mapping within intro or introns-exon borders in adult *glp-4* mutants, upon the indicated RNAi treatment.

R.3.4. The DNA damage response gene *atl-1* and *egl-1*, a cell death-related gene, are upregulated upon RNAi of s-adRP genes

Using RNA-sequencing data we analyzed the differential expression of genes in *prp-6(RNAi)*, *prp-8(RNAi)* and *prp-31(RNAi)* animals compared with those exposed to control RNAi. This bioinformatic analysis brought to light a list of genes misregulated with distinct *p-values* (Rubio-Peña et al., 2015).

There are different strategies to take benefit of this transcriptomic dataset. To set up a cut-off, we followed a candidate-gene approach since among the genes with the lowest *p-value* ($\leq$ 0,0001) we found the effector of apoptosis *egl-1*, which encodes a conserved BH3-only

domain protein (Barbara Conradt & Horvitz, 1998; Nehme & Conradt, 2009).

Among the top upregulated genes a primary sensor of DNA damage in humans also called our attention (Fig. R. 8A). It was *atl-1*, human homolog of ATR. ATR is highly conserved in all eukaryotes and functions as a cell-cycle checkpoint nuclear kinase in response to DNA damage and DNA replication arrest (Garcia-Muse & Boulton, 2005; Kastan & Bartek, 2004; Suetomi, Mereiter, Mori, Takanami, & Higashitani, 2013).

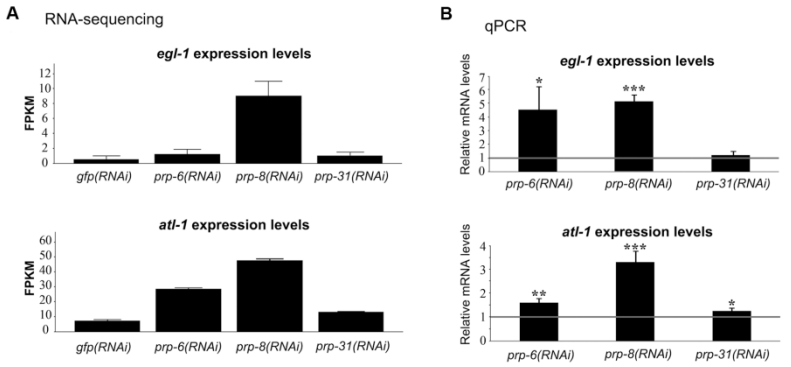

Fig. R. 8. Upregulation of *atl-1* and *egl-1* after RNAi of s-adRP genes follows a gradient from *prp-8(RNAi)* to *prp-31(RNAi)*.

(A) RNA-Seq data of *atl-1* and *egl-1* after RNAi in wild type worms. FPKM represents the Fragments per Kilobase of transcript per Million mapped reads. Bars indicate the average confidence intervals values provided by Cufflinks (Trapnell et al., 2012) for each gene. (B) Validation of the RNA-Seq data by qPCR. mRNA levels of *atl-1* and *egl-1* upon RNAi of some s-adRP genes are represented relative to their expression in *gfp(RNAi)* control animals (arbitrary value of 1, indicated with a grey line). qPCR expression data was normalized to transcript levels of *tbb-2*. Three separate experiments were analyzed. Error bars represent the standard deviation. Student's t-test for independent samples was used to analyze the statistical significance: one, two, and three asterisks indicate $p<0.05$, $p<0.01$ and $p<0.001$, respectively.

To validate the upregulation of both genes, total RNA from a synchronized population of s-adRP*(RNAi)* and *gfp(RNAi)* treated worms grown at 20°C was isolated 24 hours post L1. From these samples cDNA was synthesized and analyzed by qPCR (Fig. R. 8B). qPCR showed up to *5-fold* up-regulation of *egl-1* and *3-fold* upregulation of *atl-1* in *prp-8(RNAi)* worms. Considering that all three RNAi clones (*prp-6*, *prp-8* and *prp-31*) are similarly effective (Fig. R. 1B), we conclude that the upregulation of *atl-1* and *egl-1* follows a gradient as follows: *prp-8* > *prp-6* > *prp-31*.

R.4. RNAi of s-adRP genes induces the expression of the pro-apoptotic factor *egl-1* in a cell type-specific manner.

In *C. elegans*, the genetic pathway for the execution of developmental programmed cell death is well established. Programmed cell death can occur either during embryonic and postembryonic development in the soma (Sulston & Horvitz, 1977; Sulston, Schierenberg, White, & Thomson, 1983) or in the gonad of adult hermaphrodites (J. Sulston, 1988; Tina L. Gumienny, Eric Lambie, Erika Hartwieg, 1999; White, 1988). Somatic cell linage in *C. elegans* is invariant, therefore, it is not only known the number of cells that are destined to undergo programmed cell death but also which cell will die and when (Sternberg & Horvitz, 1984; J. E. Sulston & Horvitz, 1977; J. E. Sulston et al., 1983). In a hermaphrodite wild type worm 113 somatic cells undergo programmed cell death during embryonic and 18 during postembryonic development (Sulston and Horvitz 1977; Sulston et al. 1983).

R.4.1. *egl-1* is ectopically expressed in somatic cells in s-adRP RNAi animals

As mentioned above, one of the genes that were found upregulated in the transcriptomic analysis of s-adRP partially depleted worms was the pro-apoptotic factor *egl-1*. In order to see if the over-expression of *egl-1* detected in the transcriptomic analysis was global or restricted to a specific cell-type, we used an *egl-1* transcriptional reporter (*egl-*

*1*p::2xNLS::GFP) and exposed the worms to either *prp-8(RNAi)* or control RNAi.

During post embryonic development, only 18 cells are expected to express *egl-1* in a specific period of time during larval stage L2 (Fig. R. 9A). However, we observed that *prp-8(RNAi)* animals displayed GFP expression in somatic cells that are not expected to suffer apoptosis during normal development (J. E. Sulston & Horvitz, 1977). This *egl-1* ectopic expression was particularly evident in hypodermal seam cells, which form a row from the head to the tail of the worm (Fig. R. 9B).

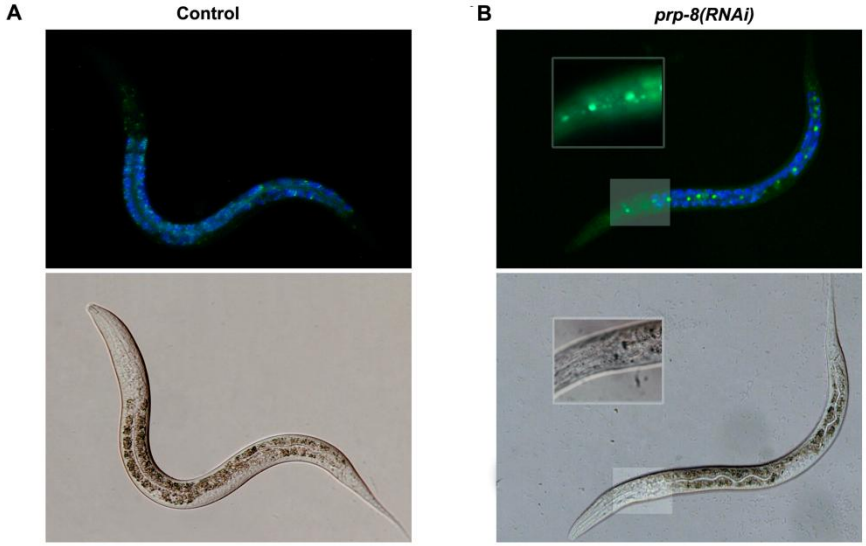

Fig. R. 9. *egl-1*::GFP expression is ectopically induced in *prp-8(RNAi)* animals.

Expression of the transgene *opIs56* [P*egl-1::2xNLS::GFP*] after 24 hours at 20°C in (A) control RNAi (empty vector) and (B) *prp-8*(*RNAi)*. Size of the animals and the germline development stage indicate that in our experimental conditions *prp-8(RNAi)* worms develop to a similar stage as N2 during the first 24 hours. *prp-8(RNAi)* worms ectopically express GFP in additional cells, including hypodermal seam cells (magnified area). Blue fluorescence is shown to label areas with autofluorescence. Images displayed are representative of three different experimental replicates.

R.4.2. *egl-1* ectopic expression is observed in hypodermal seam cells upon *prp-8(RNAi)* treatment.

Looking for an additional confirmation that hypodermal cells, specifically the seam cells, are the ones mostly affected by the partial inactivation of *prp-8* by RNAi, we used a reporter strain for *dlg-1 [dlg-1::RFP + unc-119(+)]*. DLG-1 is localized in adherent junctions of intestinal, epithelial and hypodermal seam cells (Firestein & Rongo, 2001).

We combined this strain with the GFP reporter of *egl-1*, so we could be able to distinguish the expression pattern of both genes in the same animal. The double reporter worms were synchronized and treated with *prp-8(RNAi)* following the same conditions as previous experiments. After 24 hours of treatment worms were recovered and observed under fluorescent microscopy. As we expected, *egl-1* expressing cells were all surrounded by DLG-1::RFP signal, indicating that the cells expressing *egl-1* ectopically are seam cells (Fig. R. 10).

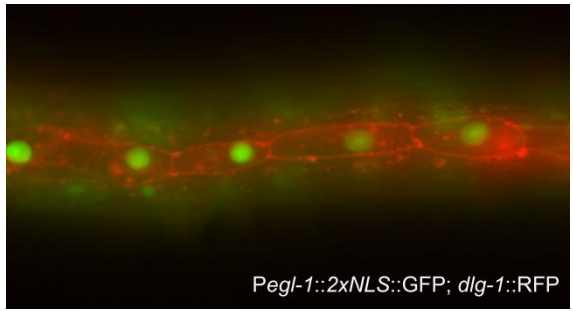

Fig. R. 10. *egl-1* ectopic expression in seam cells of *prp-8(RNAi)* animals. EGL-1::GFP; DLG-1::RFP double reporter upon *prp-8(RNAi)* treatment. Representative image was taken under fluorescent microscopy 24 hours post L1 at 20°C. *egl-1* ectopic expression is observed in the nuclei of the seam cells. Seam cell junctions are evident by the RFP signal expressed under the promoter of *dlg-1*.

R.4.3. *egl-1* ectopic expression leads to the formation of apoptotic corpses.

We wondered if this ectopic expression of *egl-1* was actually leading to apoptosis in these cells. In *C. elegans*, cells undergoing programmed cell death change their morphology and refractivity and can be observed in living animals using differential interference contrast (DIC) microscopy (B. Conradt, Wu, & Xue, 2016; Robertson & Thomson, 1982). In *C. elegans*, the transition from life to death of cells culminates in the appearance of highly refractile button-like discs, which are rapidly recognized and removed by neighboring cells in a process that takes around 101 minutes (Robertson & Thomson, 1982).

The clearance of dying cells in *C. elegans* is carried out by neighboring cells that engulf them for a posterior degradation. Some *cell death abnormal (ced)* genes, for example *ced-6*, are involved in this engulfment step. In mutants of these genes, the pattern and kinetics of cell death are normal, but cell clearance is impaired, resulting in the accumulation of 'persistent cell corpses', which can remain for many hours, or even days (Lettre & Hengartner, 2006) facilitating their visualization.

We took advantage of the mutant strain *ced-6(n2095)*, which is defective in the clearance of apoptotic cells, to look for apoptotic cell corpses in animals treated with *prp-8(RNAi)*. As expected, we detected apoptotic corpses in somatic cells (probably seam cells) of animals treated with *prp-8(RNAi)* that were not present in *gfp(RNAi)* control worms (Fig. R. 11A,B).

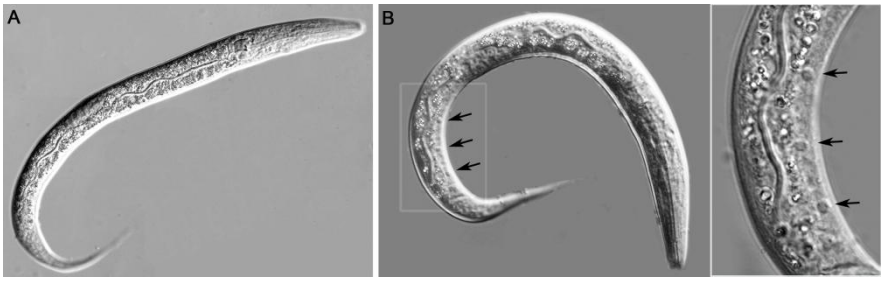

Fig. R. 11. Animals treated with *prp-8(RNAi)* display additional apoptotic cell corpses.
The *ced-6* mutant larvae treated with (A) control RNAi and (B) with *prp-8(RNAi)*. Black arrows indicate apoptotic cells (button-like refractile corpses) that were found in *prp-8(RNAi)* worms but not in control animals. The right panel shows a magnified image of the area highlighted with a white box.

R.4.4. *egl-1* ectopic expression is partially dependent of *cep-1*

Developmental apoptotic pathways have been described in somatic and germ cells; however, DNA damage induced apoptosis has only been described in germ cells. As discussed in the introduction, the apoptotic pathway in the germ line is well known, and we can distinguish three different, genetically separable apoptotic pathways (Gartner et al., 2008): (i) Physiological germ cell death, (ii) DNA damage and (iii) pathogen induced germ cell death (Aballay & Ausubel, 2001; Gartner, Milstein, Ahmed, Hodgkin, & Hengartner, 2000; Gumienny et al., 1999). In the germ line, unlike what happens during physiological germ cell death, DNA-damage-induced apoptosis occurs in an *egl-1* dependent manner (Nehme & Conradt, 2009) and it requires the participation of the transcription factor *cep-1*, the p53 homolog (Hofmann et al., 2002).

We wondered if *egl-1* dependent apoptosis in somatic cells upon RNAi treatment with adRP genes also needed the intervention of *cep-1*. To answer this question we combined the strain carrying *cep-1(gk138)* mutation, which blocks the induction of *egl-1* expression and is defective in the activation of apoptosis upon DNA damage in the germ line (Hofmann et al., 2002; Morthorst & Olsen, 2013), and the *egl-1*p::2xNLS::GFP reporter. This strain was exposed to *prp-8(RNAi)* and grown at 20°C. EGL-1::GFP expression was visualized after 24 hours post L1 and the number of cells per worm presenting ectopic expression of *egl-1* was estimated for each strain. The number of cells presenting ectopic expression of EGL-1::GFP was significantly higher in worms with a functional *cep-1*; however, when *cep-1* is mutated we can still observe ectopic expression in somatic cells suggesting that transcription of *egl-1* is, at least, partially dependent of *cep-1* (Fig. R. 12).

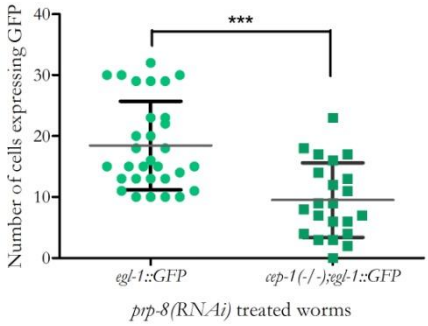

Fig. R. 12. *egl-1* ectopic expression upon *prp-8(RNAi)* treatment is partially dependent on *cep-1*.

Quantification of cells showing ectopic expression of *egl-1* in wild type and *cep-1(gk138)* background in worms treated with *prp-8(RNAi)* for 24-30 hours at 20°C. Each dot represents the amount of GFP-positive cells in one worm. Number of GFP-positive cells in wild type worms is significantly higher than in the *cep-1* mutant background upon *prp-8(RNAi)* (p *value*<0.0001). Error bars (black) represent standard deviation and grey line represents the mean. For each condition the cells of at least 20 worms were counted.

R.5. Tissue-specific RNAi of s-adRP genes

To further investigate this cell-type-specific induced apoptosis, we performed tissue-specific RNAi. We used *rde-1* mutant strains that are resistant to RNAi but also carry a wild type copy of *rde-1* expressed under the control of distinct tissue-specific promoters. Since many tissue-specific promoters are available for *C. elegans*, this system can be applicable to different tissues or selected cells (Qadota et al., 2007).

R.5.1. Tissue specific *prp-8(RNAi)* causes larval arrest when RNAi is active in hypodermal cells.

We used strains that ectopically express *rde-1* either in muscle or hypodermal cells only. Synchronized L1 worms were grown at 20°C and fed with *prp-8(RNAi)* or empty vector clones. After 48 hours, phenotypes were observed under the scope and body length was measured using ImageJ software.

Strikingly, we observed the characteristic growth arrest of *prp-8(RNAi)* when the RNAi machinery was functional in hypodermal cells, but we did not detect such phenotype when the *prp-8(RNAi)* was effective in muscle cells (Fig. R. 13A,B). Interestingly, these hypodermal cells (including seam cells) synthesize the proteins that form the cuticle and, similarly to retinal cells, require high transcriptional activity (AP. Page, 2007).

Knowing that *prp-8(RNAi)* produced a tissue-specific effect, we wanted to know if this effect correlated with *atl-1* and *egl-1* increased gene expression. We isolated RNA from wild type and hypodermal- and muscle-specific RNAi worms and performed a qPCR to estimate gene expression levels of both genes. Interestingly, we observed that

the tissue-specific up-regulation of *atl-1* and *egl-1* occurred when *prp-8(RNAi)* was efficient in the hypodermis but did not happen when the RNAi worked in muscle cells only (Fig. R. 13C).

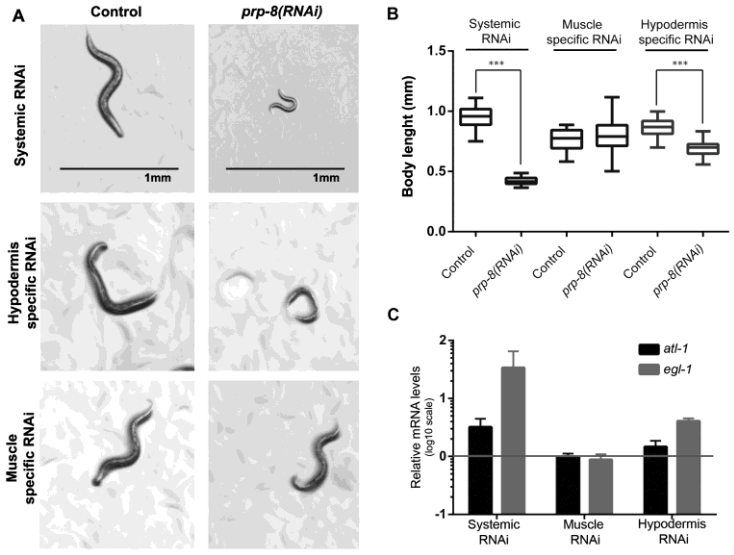

Fig. R. 13 *prp-8(RNAi)* animals display tissue-specific phenotype.
(A) Larval arrest phenotype was observed in wild-type and hypodermis-specific RNAi animals *(rde-1(ne219) V; KzIs9)*, while no obvious phenotype was observed in muscle-specific RNAi worms *(rde-1(ne219) V; KzIs20)*. Representative images were taken under the microscope after 48 h at 20°C. Scale bar, 1 mm. (B) Body length measure of wild-type, and hypodermis- and muscle-specific RNAi animals fed with *prp-8(RNAi)* and control (empty vector) clones. Animals synchronized at L1 stage were treated with the corresponding RNAi clone and grown at 20°C. More than 25 worms were measured after 48 h of treatment in each condition. Animal length is showed in millimeters and was scored using the ImageJ software. Statistical significance was calculated using Student's t-test for independent samples. Three asterisks indicate statistical significance with p-value<0.001. Whiskers were plotted by Tukey's test. (C) Up-regulation of *atl-1* and *egl-1* after *prp-8(RNAi)* is tissue specific. qPCR results for *atl-1* and *egl-1* expression in wild-type and hypodermis- and muscle-specific RNAi animals represented in a bar graph. mRNA levels of these genes after *prp-8(RNAi)* are relative to their expression in control animals. Results obtained from three independent biological replicates. mRNA transcript levels of *atl-1* and *egl-1* are normalized against *tbb-2* levels and represented in a *log10* scale. Error bars represent standard deviation.

R.5.2. *prp-8(RNAi)* worms present less number of seam cells.

Additionally, we used a GFP reporter under the control of a seam cell specific promoter (SCMp::GFP) to see how was the pattern of seam cells upon *prp-8(RNAi)*. After L2 larval stage (~24h post L1), we expect to see 16 seam cells in each row along the worm´s body. At mid-L4 (~40h post L1), the final divisions had been taken place and, as a last step in the hypodermis development, the cell junctions between the seam cells in each row disappear forming two continuous lateral syncytia (with 16 nuclei each) (Podbilewicz & White, 1994).

We observed the number of seam cells in *prp-8(RNAi)* and control worms at different time points. At 48h post L1, the number of seam cells in *prp-8(RNAi)* treated animals was reduced with respect to controls (Fig. R. 14A,B). However, at 24h post L1, we can observe the 16 cells in each row showing that a normal number of seam cells are present at the beginning of the development, but they start dying as development progresses (Fig. R. 14C). This confirms that seam cells are actually being affected in a drastic manner by the partial inactivation of *prp-8*.

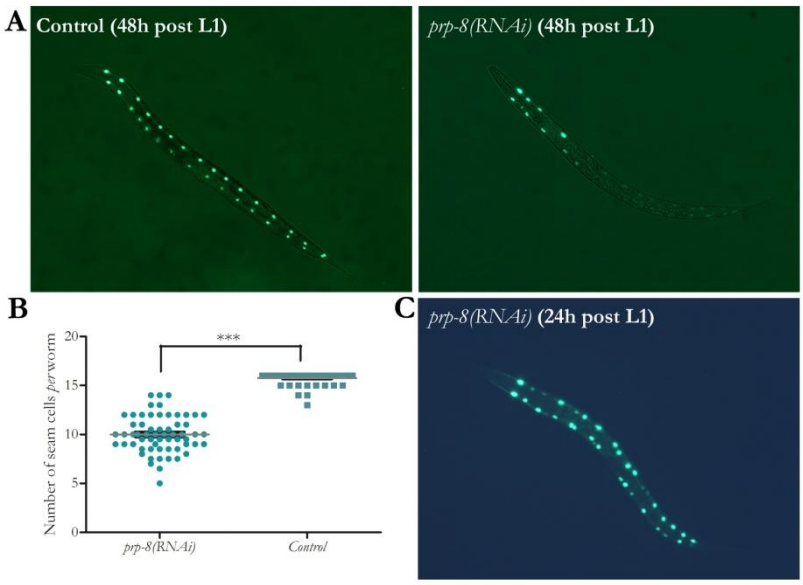

Fig. R. 14. The number of seam cells decreases upon *prp-8* RNAi.
(A). Representative images of the reporter strain that carries the transgene *wIs51*[SCMp::GFP + *unc-119(+)],* which expresses GFP in hypodermal seam cells, upon control and *prp-8(RNAi)* conditions. Animals were studied under fluorescence microscopy 48 hours post L1 stage at 20°C. (B). Number of GFP-positive cells in control or *prp-8(RNAi)* worms is represented in a bar graph. The statistical test applied is a t-student test and asterisks represent p-value<0.001. Error bars represent standard deviation. The data is obtained from three different biological replicates. For each independent experiment 30 worms were scored. (C) Representative image of *prp-8(RNAi)* treated worm at the first steps of the worm´s development is shown with the purpose of demonstrating that a normal number of seam cells is present at the beginning.

R.6. Generation of avatar worms carrying s-adRP mutations

New direct genome editing techniques are now available and they are changing the way we do science. As mentioned in the introduction, CRISPR/Cas9 technique allows us to modify the genome at will and we are using it to introduce the same missense mutations found in retinitis pigmentosa patients into the *C. elegans* genome.

At least nineteen different mutations in PRPF8 related to adRP have been identified to date including missense changes, premature stop codons and deletions, eleven of which correspond to missense mutations (Table 2) (Boon et al., 2007; Martínez-Gimeno et al., 2003; Ruzickova & Stanek, 2016; Towns et al., 2010). These mutations are clustered in the C-terminal Jab1/MPN domain, a conserved region which interacts with SNRNP200 and regulates its helicase activity (Boon et al., 2007; Mozaffari-Jovin et al., 2013; Pena, Liu, Bujnicki, Lührmann, & Wahl, 2007; Ruzickova & Stanek, 2016) (Fig. R. 15).

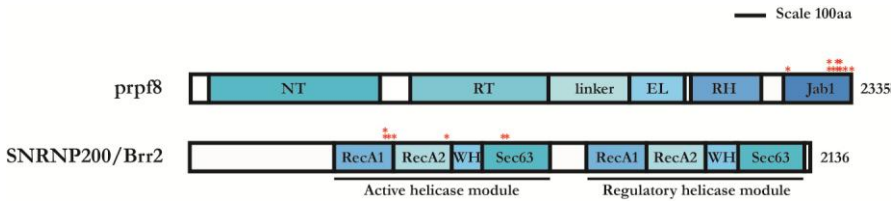

Fig. R. 15. Domain structure and RP mutations of prpf8 and SNRNP200/Brr2.

Scale representation of prpf8 and SNRNP200/Brr2 protein domains. The positions of mutations linked to adRP are marked by red asterisks. Protein domains are indicated in the scheme. Abbreviations: NT- N-terminal domain; RT- reverse transcriptase -like domain; EL - endonuclease-like domain; RH - RNase H-like domain; WH- Winged Helix (*Adapted from* Ruzickova & Stanek 2016).

For SNRNP200 7 mutations have been identified in adRP patients (Table 2). Two of them (Ser1087Leu and Arg1090Leu) are in the Sec63-like domain of the enzymatically active N-terminal helicase module. These mutations affect U4/U6 unwinding and reduce splicing fidelity by promoting the usage of cryptic sites. However, they can be normally incorporated in snRNPs and do not affect their formation. The other 5 (Arg681Cys, Arg681His, Val683Leu, Tyr689Cys and Gln885Glu) have been recently found in the RecA-like domains RecA1 and RecA2 of the same helicase module. These two domains constitute the "core" helicase domains and provide the motor associated with helicase activity (Fig. R.15) (Benaglio et al., 2011; Johnson & Jackson, 2013; T. Liu et al., 2012; Ruzickova & Stanek, 2016).

Exon	Nucleotide Mutation	Aminoacid Sustitution
	prpf8	
38	c.6353C>T	Ser2118Phe
42	c.6901C>A	Pro2301Thr
42	c.6901C>T	Pro2301Ser
42	c.6912C>G	Phe2304Leu
42	c.6926A>C	His2309Pro
42	c.6926A>G	His2309Arg
42	c.6926G>A	Arg2310Lys
42	c.6928A>G	Arg2310Gly
42	c.6930G>C	Arg2310Ser
42	c.6942C>A	Phe2314Leu
42	c.7000T>A	Tyr2334Asn
42	c.7006T>C	Term2336Arg
	SNRNP200/Brr2	
16	c.2041C>T	Arg681Cys
16	c.2042G>A	Arg681His
16	c.2047G>T	Val683Leu
16	c.2066A>G	Tyr689Cys
20	c.2653C>G	Gln885Glu
25	c.3260C>T	Ser1087Leu
25	c.3269G>T	Arg1090Leu

Table R. 2. List of missense mutations and amino acid substitutions of prpf8 and SNRNP200/Brr2 linked to adRP.
(*Adapted from* Ruzickova & Stanek, 2016)

R.6.1. Generation of s-adRP worm mutants by CRISPR

Among the 11 PRPF8 missense mutations related to adRP, we decided to target the specific mutation p.R2310G that affects one Spanish family (Martínez-Gimeno et al., 2003; Towns et al., 2010). Based on their position, the residue changes in PRP8 related to adRP are divided into three types, each likely associated with different functions in Prp8 (Mozaffari-Jovin et al., 2014). This mutation belongs to the type II which maps to the proximal region of the Jab1 C-terminal tail. Based on its position, this residue is expected to interact with the Sec63 unit of the active helicases module of SNRNP200.

As SNRNP200 mutations are located in two different regions of its active helicase module, we decided to target at least one mutation in each region so we can see if mutations targeting different motifs have different outcomes. For practical reasons regarding efficiency of CRISPR/Cas9 genome editing, mainly proximity to the PAM sequence, we chose p.V683/676L (Human/Worm position) and p.S1087/1080L.

As described in the introduction, to generate specific mutations using CRISPR we need, along with the single guide (sgRNA=crRNA+ tracrRNA) and the Cas9 protein, a repair template which carries the nucleotide changes that will produce the amino acid substitution. Then, we rely on homology directed repair (HDR) to get our mutation into the genome.

However, this is not the only way the DNA can be repaired after it is cleaved by the Cas9 protein. In the absence of a repair template, the DNA is repaired by the error-prone Non-Homologous End-Joining

(NHEJ) pathway that usually produces small indels (Ran et al., 2013). So it is not surprising that even in the presence of a repair template, NHEJ repair can also be triggered, producing small indels.

To identify in which worms Cas9 protein is active and HDR took place we co-inject, along with our target gene sgRNA and repair template, a sgRNA and repair template for *dpy-10* that works as a selection marker that would point us in the direction of the plates where Cas9 was cutting efficiently. *dpy-10* sgRNA produces dumpy (Dpy) animals, which can be easily identified as short and fat worms, where Cas9 protein is active. Additionally, if the repair template is incorporated by HDR, worms display a roller (Rol) phenotype (for details see *Material and Methods*). The plates where Dpy and/or Rol worms are found are known as "jackpot plates".

The strategy to introduce this mutation into the worm´s genome follows the one reviewed by Paix and collaborators, and relies in the use of purified Cas9 and commercial crRNA, tracrRNA and ssODNs (Paix et al., 2015). An overview of the CRISPR design and mutations obtained in *prp-8* and *snrp-200* are shown in figure R. 16 and R. 17, respectively.

R.6.1.1. *prp-8* CRISPR mutations

To generate the R2310/2303G mutation, we used a single-stranded oligonucleotide (ssODN) to serve as the repair template. This repair template includes several silent mutations, used to prevent re-cutting of the Cas9 and facilitate PCR screening, and c.6907C>G and c.6909G>C nucleotide substitutions that will change the arginine for the glycine in the translated protein (Fig R. 16).

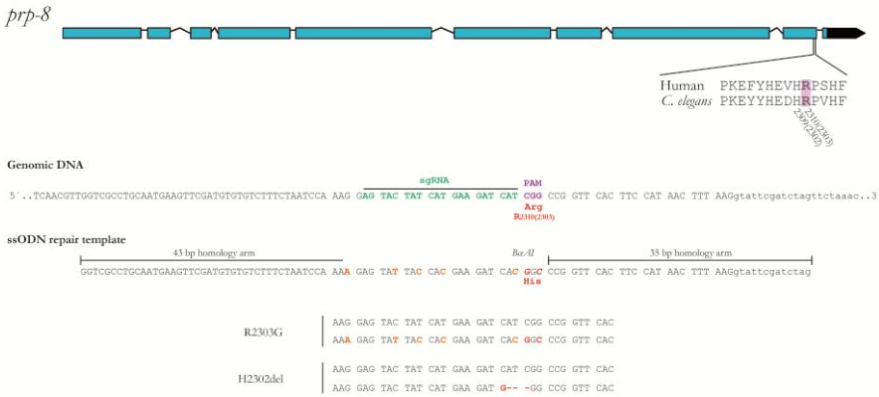

Fig. R. 16. Schematic representation of the CRISPR design to introduce *prp-8* mutations into *C. elegans* genome

Silent nucleotide substitutions (orange) were included to prevent re-cutting of Cas9 once the repair template has been incorporated. Nucleotide changes incorporated to produce the mutation are indicated in red. *prp-8* mutations: R2310 human residue corresponds to R2303 in *C. elegans* that is located in exon 9. A restriction site for *BceAI* was also incorporated to facilitate identification of the mutant worms. Two different mutant strains were obtained: One that incorporated the repair template generating the R2310/2303G substitution, and one small indel.

13 worms were injected with CRISPR/Cas9 mix as described in material and methods. From these injected worms, 5 jackpot plates were obtained and one of them contained mutated *prp-8* worms. 28 worms were scored through PCR using primers specifically designed to detect worms with the silent mutations plus the CGG>GGC codon change. We were able to identify at least 10 different mutant lines by PCR. They were sequenced for confirmation and nine of them had the desired mutation.

Additionally, sequencing of one of the samples showed that as a result of NHEJ, an indel that consisted of a 4nt deletion/1nt insertion was formed. This small indel results in the absence of H2309/2302, a residue that is also mutated in adRP patients (Fig.R 16; Table.R.2).

The nucleotide substitutions that we introduced in both mutants allow molecular identification by PCR using regular primers. Primers are specifically designed to distinguish R2310G mutants from indel mutants and from wild type worms (Fig. R. 18 A, B).

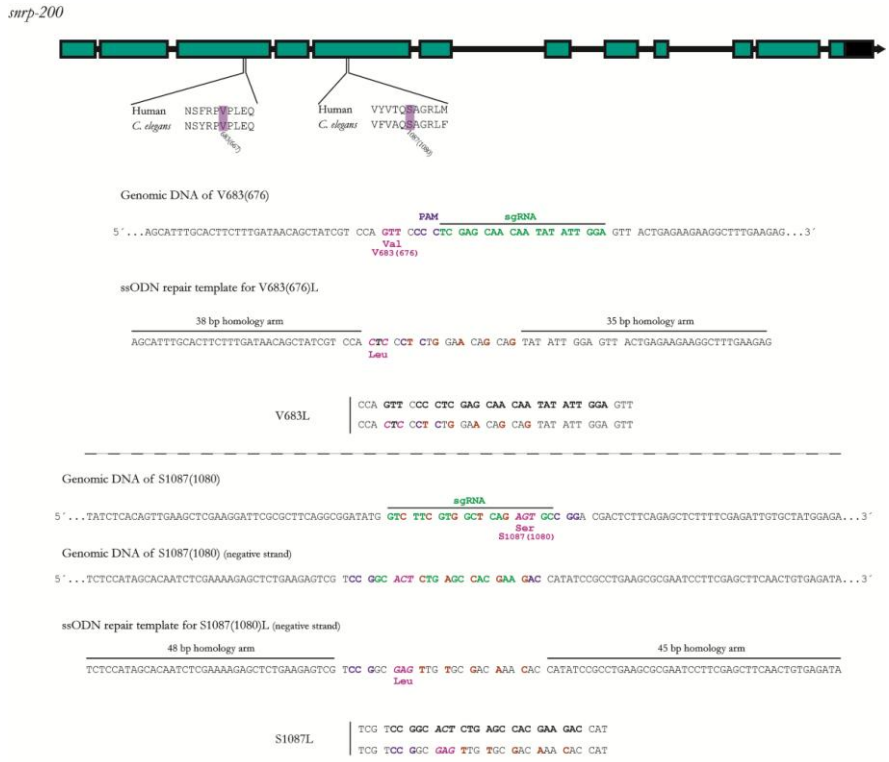

Fig. R. 17. Schematic representation of the CRISPR design to introduce *snrp-200* mutations into *C. elegans* genome.
Silent nucleotide substitutions (orange) were included to prevent re-cutting of Cas9 once the repair template has been incorporated. Nucleotide changes incorporated to produce the mutation are indicated in red. *snrp-200* mutations: V683 amino acid corresponds to V676 in *C. elegans* located in exon 3. S1087 human residue corresponds to S1080 in the worm and is located in exon 5.

Fig. R. 18. Genotype of *prp-8* mutant alleles *cer14* and *cer22*.
Mutants are genotyped using a normal PCR. (A) Scheme of genotyping strategy, one common forward primer is used for all reactions, while specific primers are used to distinguish between each genotype. (B) PCR amplification products from heterozygous, homozygous for the mutations and wild type worms. Bands size are between 680-682bp, Hyperladder 50bp is used, Agarose gel 1,5%.

R.6.1.2. *snrp-200* novel mutations

Two ssODNs were designed to introduce the V676L and S1080L mutations into the worm. V676L change is generated by the substitution of nucleotides c.2026G>C and c.2028T>C. S1080L is generated by the whole replacement of one codon in the position c3238AGT>CTC. In average, 20 worms were injected for each mutation. From V676L injected animals we obtained 3 jackpot plates, and we found the mutation in all of them. 7 jackpot plates were obtain from S1080L injected worms, and 6 of them had the mutation. In this case, mutants are also easily distinguished from wild type worms using PCR with specific primers for each mutation and silent adjacent mutations (Fig. R. 19).

Fig. R. 19. Genotype of *snrp-200* mutant alleles *cer23* and *cer24*.
Mutants are genotyped using a standard PCR. (A) Scheme of genotyping strategy, one common forward primer is used for all reactions, while specific primers are used to distinguish between each genotype from each other.

R.6.2. Characterization of CRISPR *prp-8* and *snrp-200* mutants

As detailed above, using CRISPR/Cas9 we were able to generate four mutations. In *prp-8,* R2310/2303G substitution is caused by mutant allele *cer22* whereas the indel H2309/2302del was named *cer14*. Regarding *snrp-200* alleles, *cer23* and *cer24* cause V683/676L and S1087L/S1080L substitutions respectively (Table R.3).

prp-8(cer22), the allele that produces the p.R2310G substitution, is viable in homozygosis and does not exhibit obvious phenotypes. Worms grown at 15°C or 25°C look healthy in general. However, during a developmental assay we observed a low proportion of Pvl (2 out of 25 worms) at 20°C, and we also could observe at 25°C a low proportion of adult lethality (Adl) mostly because of larvae that hatch into the maternal body (Bag) or worms that explode through the vulva

(Rup) (2 out of 25). Regarding H2302del, *prp-8(cer14)*, heterozygous mutants are healthy and do not exhibit obvious phenotypes.

Type of mutation	Allele	Description	Phenotype
prp-8			
point mutation	*cer22*	R2310/2303G	No obvious phenotype
indel	*cer14*	H2309/2302del	pSte, pLva, pEmb
snrp-200			
point mutation	*cer23*	V683/676L	pLva, pEmb, pSte, pMlt
point mutation	*cer24*	S1087/1080L	No obvious phenotype

Table R. 3. *prp-8* and *snrp-200* mutations generated by CRISPR.
Phenotypes abbreviations: Ste: sterility, Lva: larval arrest, Emb: embryonic lethality, Mlt: molting defect. Variable levels of penetrance is indicated by the letter p before the phenotype.

On the other hand, *cer14* homozygotes, although viable at 15°C, only about 8% of them look healthy and are capable of reproduce. Most predominant phenotypes are sterility (Ste; 31%) and embryonic lethality (Emb; 48%). Sterile worms do not seem to be capable of self-fertilizing as deduced by the absence of fertilized oocytes in the uterus cavity during adult stage; however the cause of this sterility should be further study. Additionally, they display several phenotypes with at low penetrance including: protruding vulva (Pvl), lethality (Let), developmental delay and larval arrest (Lva) (Table 4).

Phenotype	Number of worms displaying the phenotype (%)	
Viable/Non Ste	66	(7,88)
Ste	261	(31,15)
Emb	401	(47,85)
Others (Pvl, Lva, Let)	110	(13,13)
Total	**838**	**(100)**

Table R. 4. Phenotypes observed in *prp-8(cer14)* homozygous mutants.
Phenotypes abbreviations: Ste: sterility, Emb: embryonic lethality, Pvl: protruding vulva, Lva: larval arrest, Let: lethality,

snrp-200(cer23) is the allele that produces the V683L/V676L residue change. This mutant is viable in homozygosis between 15 and 25°C but it displays a variety of phenotypes that includes Lva, Emb, Pvl, Ste and molting defects (Mlt); indeed, the penetrance of the phenotypes is clearly increased when the worms are maintained at 25°C.

Finally, *snrp-200(cer24)*, responsible for S1087/1080L amino acid change, is also a viable mutant in homozygosis. However, unlike *cer23* allele, *cer24* homozygous worms do not display any obvious phenotypes and look apparently healthy at temperatures from 15 to 25°C.

R.6.3. All s-adRP CRISPR mutants display developmental delay

We wanted to know the impact of these mutations in *C. elegans* development. We compared the developmental timing of wild type N2 worms with *prp-8* and *snrp-200* mutants. The transition from the first larvae stage (L1) to the L1 stage of the next generation is a highly synchronized cycle that takes 3 and 2,5 days in wild type strains grown

at 20°C or 25°C, respectively. For this experiment, L1 synchronized worms were singled out and placed at 20°C or 25°C. Every 24 h their developmental stage was estimated by anatomical structures and size. All mutants showed developmental delay, being *prp-8(cer22)* the less affected strain (Fig. R.20. A, B).

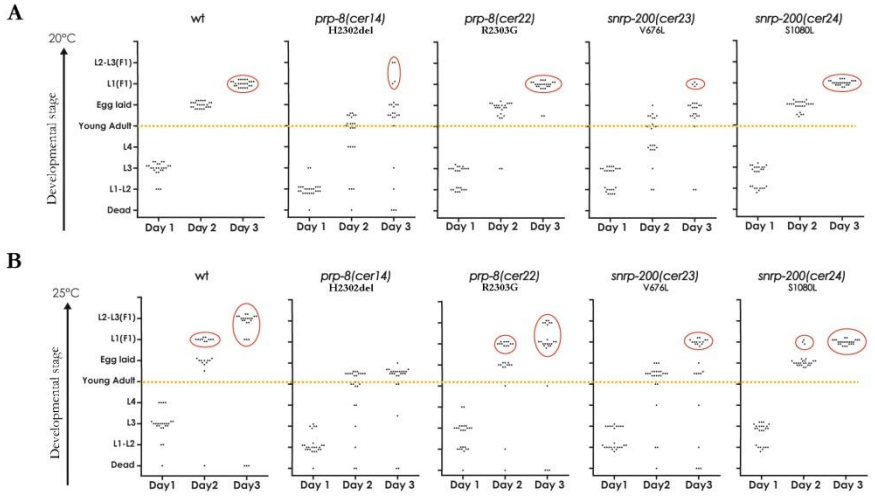

Fig. R. 20. s-adRP mutants are viable but present delayed development.

prp-8 and *snrp-200* mutations provoked delay in development from L1 larval stage into next generation (F1) L1 stage. Developmental progress of singled mutant and wild type L1 larvae was monitored every 24h for 3 days at 20 °C and 25 °C. Every dot represents a worm. Worms above orange line reached adult stage. Worms inside red circle were able to lay viable progeny. In each case, the development of 25 worms was followed.

R.6.5. *prp-8* and *snrp-200* CRISPR mutations effect in *C. elegans* fertility

Two of our mutants, *prp-8(cer22)* and *snrp-200(cer24)* did not present an obvious phenotype, thus we wondered if these mutations had any effect in *C. elegans* fertility. We studied brood sizes of homozygous mutants *prp-8(cer22), snrp-200(cer23)* and *snrp-200(cer24)* at 25°C. *prp-8(cer14)* was studied independently due to its severe Ste phenotype. In

any case, synchronized L4 worms growing at the indicated temperature were single out and placed in individual plates. They were transferred to new plates every 8-12 hours to facilitate the counting of F1 larvae until they started to lay oocytes or died.

Neither *prp-8(cer22)* nor *snrp-200(cer24)* homozygous mutants presented significative reduce brood size (Rbs) compared to wild type worms at 25°C. On the other hand, *snrp-200(cer23)* showed a significant Rbs compared to wild type worms at this temperature (Fig. R. 21).

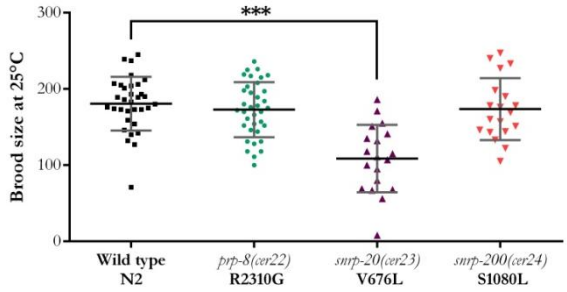

Fig. R. 21. Brood size estimation of s-adRP mutants at 25°C.
Worms were grown at 25°C until reaching L4 stage, singled out and incubated at the corresponding temperature (n≥20). Until the end of their reproductive phase, the animals were transferred to new plates every 12-16 hours and the number of larvae (F1) determined. *cer22* and *cer24* did not show significant reduced brood size (Rbs) compared to wild type. However, brood size of *cer23* homozygous animals was significantly different from wild type's (p *value*<0.0001). Error bars indicate the standard deviation (grey) and black line indicates the Mean.

As mentioned before, *cer14* H2309/2302del homozygotes although viable, they showed a high percentage of Ste animals at 15°C. For this reason we decided to study their brood size independently. For these mutants, brood sizes were estimated at 15 and 25°C. These mutants also presented a high percentage of Emb so dead embryos were not considered and only larvae were counted for the final brood size estimation (Table R.4.).

A significant Rbs was observed in *cer14* indel homozygotes at 15°C, this phenotype could be at least in part, explained by the high proportion of Emb these worms produce at this temperature (Fig. R.22.). Additionally, more than 30% of the *cer14* homozygous mutants are completely Ste at 15°C, so these Ste worms were not considered in the experiment. Most of the worms were Ste if they grown at 25°C since larval stage 1, the ones that were able to lay some progeny do not produce more than 14 larvae, in this particular study at least, and all of the F1 animals turned out to be either Ste or Lva.

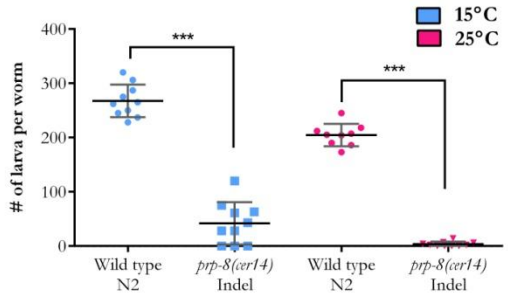

Fig. R. 22. Brood size estimation for *prp-8(cer14)* mutants at 15 and 25°C
Worms were grown at 15 or 25°C until reaching L4 stage, singled out and incubated at the corresponding temperature. Each dot represents an animal. Until the end of their reproductive phase, the animals were transferred to new plates every 8-12 hours and the number of larvae (F1) determined. Brood size of mutant worms were significantly different from wild type's (p *value*<0.0001) at both temperatures. Error bars indicate the standard deviation (grey) and black line indicates the mean.

In summary, to mantain the cer14 homozygous mutant, the strain needs to be kept at 15°C, and still the fertile animals produce less than half of the progeny of a wild type worm.

R.6.6. *atl-1* and *egl-1* are also upregulated in s-adRP mutants

We wondered if our CRISPR s-adRP mutants also present upregulation in the DNA damage sensor *atl-1*/ATR and the pro-apoptotic gene *egl-1*. Most of the phenotypes observed in our mutants are temperature dependent and show a stronger effect in the worm at 25°C. For this reason, synchronized L1 wild type and s-adRP CRISPR mutants were grown at 25°C for 48h. Then, they were recovered and total RNA was isolated from each sample and used as template for cDNA synthesis.

atl-1 upregulation was present in all samples. An increase of 5-, 8-, 12- and 48-fold was observed for *prp-8(cer14)*, *snrp-200(cer24)*, *prp-8(cer22)* and *snrp-200(cer23)*, respectively. *egl-1* expression was also upregulated in these worms with an increase of 64-, 245-,415- and over 1000 fold in *snrp-200(cer24)*, *prp-8(cer22)*, *prp-8(cer14)* and *snrp-200(cer23)* (Fig. R. 23). The increased expression of both genes correlates with the severity of the phenotypes we have observed so far, showing that the biggest differences in gene expression are observed in *snrp-200(cer23)* and *prp-8(cer14)*.

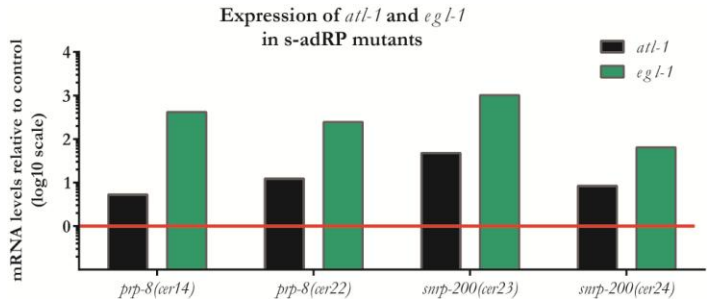

Fig. R. 23. Quantification of mRNA levels of *atl-1* and *egl-1* in s-adRP novel mutants.

All four mutations produce an increase in *atl-1* and *egl-1* expression in worms grown at 25°C for 48 h. mRNA levels of *atl-1* and *egl-1* in mutant worms are relative to their expression in wild type worms. qPCR expression data was normalized to transcript levels of *tbb-2* (red line) and represented in a *log10* scale.

R.7. Indications of replicative stress and R-loops in s-adRP deficient animals

R.7.1. *atl-1*/ATR is upregulated after *prp-8(RNAi)* in a DNA damage independent manner.

atl-1/ATR and *atm-1*/ATM are primary DNA damage sensors in the DNA damage response (DDR) pathway, an important step in the maintenance of the genomic integrity. Even when ATM and ATR can share elements of the signaling pathway, they basically respond to different types of DNA damage. While ATM mainly senses DNA double strand breaks (DSBs), ATR responds to other DNA lesions such as those produced by replication fork stalling or UV damage (Garcia-Muse & Boulton, 2005).

We wanted to uncover what triggers *atl-1*/ATR upregulation in s-adRP genes deficient animals. We targeted this question from different approaches focusing in the presence of either DNA damage or replication stress. Either the reason of *atl-1*/ATR increased expression, it would be a cause of genomic instability that, if is not fully solved by the ATR checkpoint, could lead the cell to an apoptotic fate.

R.7.1.1. *prp-8(RNAi)* produces sensitivity to DNA damage induced by UV-C

We used the RPA-1::YFP reporter, which forms foci upon DNA damage, to study the presence of DNA damage in *prp-8(RNAi)* animals. RPA, RPA-1 in *C. elegans,* is a sensor of DNA damage and is required for ATR activation. When there is a single-strand DNA

(ssDNA) region originated by replication interference, RPA recruits ATR to the ssDNA (Garcia-Muse & Boulton, 2005). The subsequent activation of ATR downstream checkpoint signals can lead to cell cycle arrest or apoptosis. RPA-1::YFP is expressed in all worm cells, but the distribution of this tagged protein is diffuse and does not form visible *foci* (Vermezovic et al., 2012). Formation of R-loops (DNA:RNA hybrid structures) is also capable of leaving a ssDNA exposed, what represents a threat to genomic stability as this ssDNA is left exposed to genotoxic agents.

To study the dynamics of RPA-1 in s-adRP deficient worms, we synchronized L1 RPA-1::YFP animals and treated them either with *prp-8(RNAi)* or control RNAi for 24 hours. Then, *foci* formation was assessed using confocal microscopy. We did not observe RPA-1::YFP foci formation after *prp-8(RNAi)* in normal growth conditions. Then we decided to evaluate RPA-1::YFP expression pattern upon exposure to a genotoxic agent such as UV-C radiation. Animals treated with *prp-8(RNAi)* or control clones were exposed to 100 J/m^2 UV-C 24 h post L1. 24 hours later, we studied *foci* formation by confocal microscopy. We observed that the number of *foci* in *prp-8(RNAi)* treated worms was higher than in control worms, and that these *foci* were present in both germ and somatic cells (Fig. R. 24). This result suggests that a partial depletion of *prp-8* activity may cause an alteration in the DNA that is not reported by our RPA-1::YFP transgenic strain in normal conditions, but it is uncovered upon UV exposure. In other words, *prp-8(RNAi)* does not cause DNA damage but produces susceptibility to genotoxic agents.

This result is compatible with genomic instability and presence of R-loops, since these DNA:RNA structures also leave exposed a ssDNA that can be susceptible to genotoxic agents.

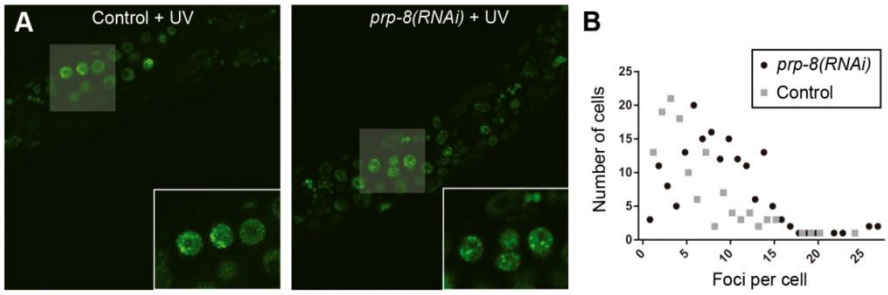

Fig. R. 24. Higher sensitivity to UV-induced DNA damage observed in *prp-8(RNAi)* treated animals.

After UV-C exposure, foci formation evidenced by the transgene *opIs263 [Prpa-1::rpa-1::YFP+unc-119(+)]* is more abundant in animals with partial inactivation of *prp-8* compared to control worms (fed with empty vector clone). (A) Representative confocal images of transgenic worms carrying the RPA-1::YFP transgene under control and *prp-8(RNAi)* conditions. *Foci* formation was observed in somatic and germline cells in both conditions. White squares show magnified images of germline cells with RPA-1::YFP foci. (B) Quantification of *foci* formation per cell represented in a dispersion graph. Each square (control) and dot (*prp-8(RNAi)*) represents the number of cells displaying the corresponding amount of *foci*. Only somatic and germline cells that displayed one or more *foci* were scored.

We wanted to further investigate the role of R-loops in the up-regulation of *atl-1* in s-adRP depleted worms. R-loops are considered a barrier to replication (Hamperl et al., 2016), so the presence of these structures would make the worms prompt to replicative stress. We addressed this point by using hydroxyurea (HU). Replication stress in the worms can be induced by exposing them to hydroxyurea which is an inhibitor of ribonucleotide reductase that leads to replication fork stalling (Garcia-Muse & Boulton, 2005).

We exposed *egl-1*p::2xNLS::GFP worms to HU and observed, similarly to the effect of *prp-8(RNAi)*, ectopic expression of *egl-1* and arrested development (Fig. R. 25). We think that this could be an indirect indication that the phenotypes we observe in *prp-8(RNAi)* treated worms could be caused, at least in part, by replication stress.

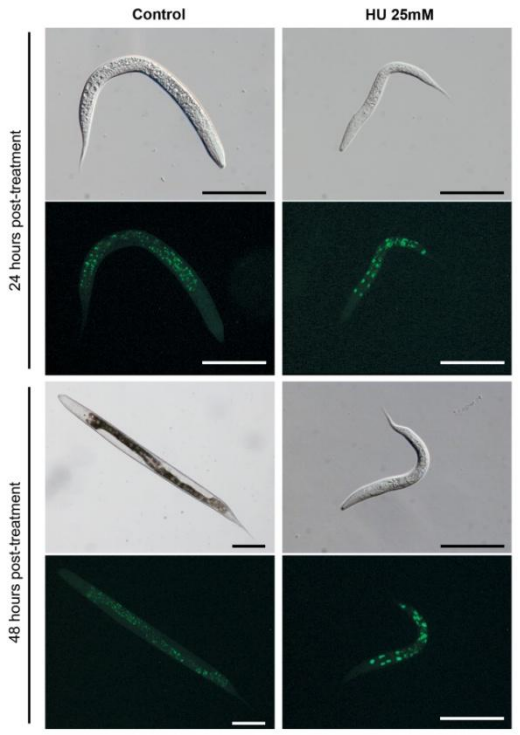

Fig. R. 25. Larval arrest and *egl-1*::GFP ectopic expression is observed upon hydroxyurea treatment.
Panels show Nomarski and fluorescence microscopy images of synchronized animals that were grown at 20°C either in NGM plates containing 25mM of hydroxyurea or normal NGM plates. Pictures were taken 24 and 48 hours post L1. Scale bar represents 100μm in all cases.

We also asked whether the expression of *atl-1* and *egl-1* was induced under the effect of UV-C and exposure to HU. For that purpose, we isolated RNA from N2 worms treated either with HU or exposed to UV-C 24 hours post L1, and perform qPCR. These two DNA insults were capable of increasing the expression levels of *atl-1* and *egl-1* (Fig. R. 26).

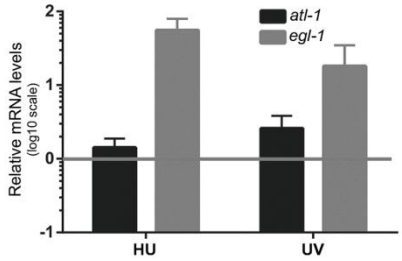

Fig. R. 26. Quantification of *atl-1* and *egl-1* upon exposure to DNA insults.
Both DNA insults, UV and hydroxyurea produce an increase in *atl-1* and *egl-1* expression. Quantification of expression levels of *atl-1* and *egl-1* in wild type animals treated either with UV (100J/m²) or hydroxyurea (25mM). Expression levels of these genes are normalized to *tbb-2* and represented relative to the ones in untreated worms.

R.7.3. s-adRP CRISPR mutant worms are hypersensitive to replicative stress

We wanted to know if our splicing CRISPR derived mutant worms were sensitive to replication fork stalling. To monitor this, we exposed worms to HU with the intention of exacerbating replication fork stalling. L1 larvae of wild type and mutant worms were exposed to different concentrations of HU (0-20mM) for 16 hours and allowed to recover for 3 days in absence of the inhibitor. Then we estimated the percentage of worms that could recover from the HU pulse and

reached the adult stage. *met-2, set-25* mutant was used as a control for replicative stress (Zeller et al., 2016).

The stronger effect, as expected, was seen in worms exposed to 20mM HU. Wild type worms recovered the 88.61% of the time, while recovery of mutants was observed in 80.13%, 60.54% and 57.59% of the time in *prp-8(cer22)*, *snrp-200(cer24)* and *snrp-200(cer23)* respectively. In this experiment, *prp-8(cer14)* mutant had to be excluded due to technical issues. One of the best indications that an animal has recovered and reached adulthood is the presence of embryos; since *cer14* allele produces a big percentage of Ste worms, it was too difficult to evaluate recovery in these worms (Fig. R. 27). These results indicate that our mutants are under replication stress.

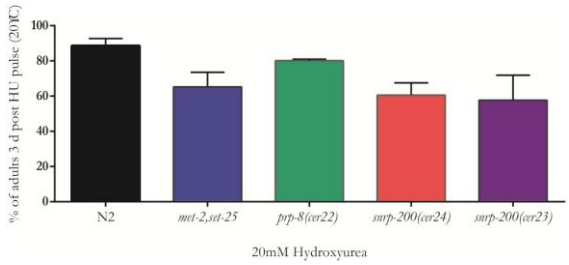

Fig. R. 27. s-adRP novel mutants are hydroxyurea sensitive.
Synchronized populations of wild type and s-adRP mutant L1 larvae were exposed to 20mM HU for 16 h, and then the worms that develop into adults after 3 d were quantified. Error bars represent standard deviation of the mean. Results represent observation of two biological replicates.

DISCUSSION

Discussion

The fact that splicing-related adRP (s-adRP) genes and the mechanisms of splicing in which they are implicated are very well conserved through evolution, allows the study of splicing-associated diseases in model organisms. Here, for first time, we propose the multicellular organism *C. elegans* as a model to study the molecular mechanisms that lead to tissue-specific apoptosis in s-adRP.

Similarly to humans, s-adRP depletion causes a tissue-specific effect in the worm. In humans, s-adRP mutations lead to apoptosis in photoreceptors in the retina, hypodermal seam cells of *C. elegans* also undergo apoptosis upon partial inactivation of s-adRP genes by RNAi (Fig.D.1).

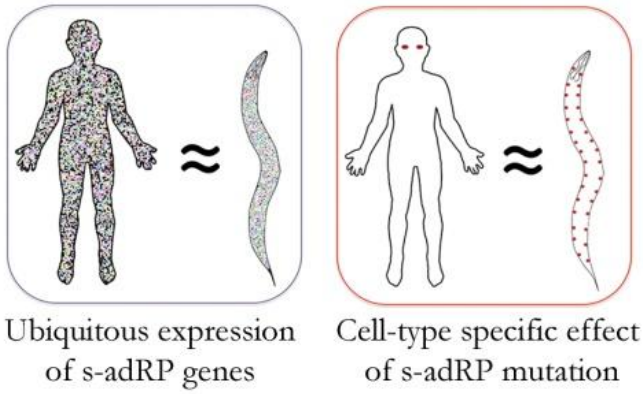

Fig. D. 1. *C. elegans* as a model for s-adRP.
s-adRP genes are expressed in all cell-types in both organisms. Partial inactivation of these genes causes tissue-specific defects in humans and worms.

D.1. The role of s-adRP proteins in splicing

s-adRP proteins are ubiquitous components of the spliceosome that belong to the U4 (PRPF3, PRPF4 and PRPF31) and U5 (PRPF6, PRPF8 and SNRNP200/BRR2) snRNP particles. Even when all of these s-adRP proteins have different functions within the splicing process, they all encode components of the tri-snRNP complex. This complex is implicated in the assembly and activation of the spliceosome, suggesting that mutations in these genes lead to a common effect within the splicing process that finally drives the cell to an apoptotic fate.

D.1.2. U4 s-adRP proteins: PRP3, PRP4 and PRP31

In the early spliceosome assembly, U4/U6 di-snRNP is required to form a tri-snRNP complex with U5 snRNP that will later be recruited to form the spliceosome core. In this process, PRP3 and PRP4 form a tri-complex with PPIH (Peptidyl-prolyl cis-trans isomerase H), which accelerates the folding of proteins (acts as a chaperone), and then binds U4/U6 snRNAs (S. Liu et al., 2015; Nottrott, Urlaub, & Lührmann, 2002). PRP3 is also known to stabilize the U4/U6·U5 tri-snRNP by interacting with PRP6 (S. Liu, Rauhut, Vornlocher, & Lührmann, 2006). When these two proteins do not associate properly there is a delayed spliceosome assembly and inefficient splicing (Tanackovic, Ransijn, Thibault, et al., 2011). The third adRP related U4 snRNP component to mention is PRP31, which is also involved in the assembly and stability of the complex. PRP31 also binds to the U5-specific protein PRP6 creating a bond which is essential for the formation of U4/U6·U5 tri-snRNP.

D.1.2. U5 s-adRP proteins: PRP6, PRP8 and SNRNP-200

The U5 PRP6 binds to four out of the other five tri-snRNP proteins involved in adRP, constituting the structural heart of the particle (Tanackovic, Ransijn, Ayuso, et al., 2011). Mutations in either these splicing factors, compromise the interaction of U4/U6 with U5 snRNP (S. Liu et al., 2006; Makarova, Makarov, Liu, Vornlocher, & Lührmann, 2002).

PRP8 and SNRNP200/Brr2 are the other two U5-specific proteins in s-adRP, that along with the GTPase Snu114, are known as key players both in spliceosome assembly and in formation of the catalytic core of splicing (Wahl et al., 2009). For the catalytic core to be activated, a rearrangement that unwinds U4/U6 and releases U4 from the complex is necessary. This unwinding is mediated by the helicase activity of SNRNP200/Brr2, which has also been implicated in the catalytic step of splicing and in spliceosome disassembly and recycling (Maeder, Kutach, & Guthrie, 2009; Zhao et al., 2009). Thus, SNRNP200 and PRP8 not only structurally support the assembly of the spliceosome but also are actively responsible of splicing activation and catalysis (Mozaffari-Jovin et al., 2014).

D.1.3. U4 versus U5 components

Overall, all s-adRP proteins are required to activate the spliceosome, but U5 proteins remain in the activated complex whereas U4 proteins do not. Interestingly, our analysis of RNAi phenotypes distinguishes two phenotypic clusters between s-adRP genes coding for U4 or U5 components, being U5 genes(RNAi) phenotype more severe than the other. The wider role of U5 snRNPs in splicing could explain these differences. Consistently, human genes coding for U5 snRNPs allow less harmful mutations than U4 snRNP coding genes.

For example, most PRP8 adRP mutations consist of a single amino acid change, caused by a point mutation in the last exon (exon 42) (Grainger, 2005; Towns et al., 2010), whereas genes coding U4 snRNP components such as PRP31 include null alleles that are able to be maintained in heterozygosis in adRP patients (Mordes et al., 2006; Rio Frio et al., 2008). Mutations in PRPF31 cause adRP with a variable penetrance that allows some carriers to stay asymptomatic. It has been shown that the haploinsufficiency of PRP31 can be compensated by higher expression from a wild type allele. Several genomic loci that enhance PRPF31 expression and prevent disease development have been identified (Rio Frio et al., 2008; Anna M. Rose et al., 2016; Eranga N Vithana et al., 2003). Our reporter strains carrying both the promoter and the 3' UTR of *prp-31* can be used as a system to identify genes or conditions that could influence the expression of *prp-31*. Additionally, although we did not observe haploinsufficiency in worms heterozygous for a deletion in *prp-31*, these animals can serve as a sensitive background to search for factors capable of influencing *prp-31* functions.

D.2. Effect of s-adRP mutations in splicing

Studies have been made to evaluate the effect of these mutations in spliceosome assembly. Using lymphoblast from patients with ten different mutations in PRPF3, PRPF31 and PRPF8 Tanackovic and collaborators, determinated that these cells were able to form spliceosomes *in vitro*, but the process had a reduced kinetics. They observed that the formation of complexes B and C were 3 times slower and much more inefficient than in control cells. Further experiments determined that the problem resides in the transition between complex A and B where the incorporation of the U4/U6·U5 tri-snRNP complex is needed; however, the formation of complex A, and consequently the previous steps of splicing, were not affected (Tanackovic, Ransijn, Thibault, et al., 2011).

In this same study, the effect of these mutations in pre-mRNA splicing was also evaluated. Just 5 out of 57 splicing units (8,8%) accumulated unspliced products.

This is consistent with the low intron retention events we found in the transcriptomic data obtained from *s-adRP(RNAi)* worms. This mild effect in splicing seems more likely to be due to a reduced splicing efficiency in terms of assembly and availability of the active complex, than to defects in intron recognition or alternative splicing.

Consistently with this, so far s-adRP has not been associated with mutations in splicing factors of the U1 or U2 snRNP complexes; however, mutations in splicing factors belonging to U2 snRNP have been related with other diseases. For example, somatic heterozygous

mutations in SF3B1 splicing factor have been associated with myelodysplastic syndrome (MDS). Almost all mutations are missense and usually occur at conserved positions. SF3B1 is a U2 snRNP–associated protein involved in branch point selection (Corrionero, Minana, & Valcarcel, 2011). Several studies suggest that SF3B1 mutations alter the splicing of transcripts involved in chromatin structure, DNA repair, and the DDR, thereby possibly providing an explanation for the accumulation of DNA damage in hematopoietic progenitor cells of MDS patients (Zhou et al., 2013).

Going back to the mild splicing defects detected in our s-adRP studies, such defects could be associated to high transcriptional activity since we found normal constitutive splicing in germline-less adults exposed to s-adRP*(RNAi)*. Metabolic requirements of retinal cells have been hypothesized as a possible cause of retinal pathology associated with mutations in pre-mRNA splicing factor genes (Gamundi et al., 2008). The retina is a human tissue with high transcriptional activity that is not present in a cellular culture of lymphoblasts or induced pluripotent stem cells (iPSCs). Hence, the effect of s-adRP mutations in these models can be masked by a basal metabolism. Having a model in which differentiated tissues, some of them with specific metabolic requirements like high transcriptional activity, can be useful to expose the effects that were unnoticed in other models.

Thus, splicing defects might not be the main cause of the retinal degeneration in s-adRP patients as the effects of these mutations on intron removal must be subtle. Indeed, the progression of the disease in a RP patient is slow and it takes several years until symptoms start to show up. A dramatic splicing defect in any cell-type, and even more

in cells with high transcriptional activity, would be expected to cause a more aggressive outcome. At any case, the molecular basis of the disease remains unknown so far.

PRPF8 adRP mutations

PRP8 is one of the largest nuclear proteins and it is a core subunit of the pre-catalytic and catalytic spliceosome. Mutations in PRPF8 related to s-adRP are located in the C-terminal region of the protein, mostly in its Jab1/MPN domain and the C-terminus tail. PRPF8 has also been related with diseases like MDS and others types of cancer in which splicing defects have been identified. However, mutations associated with these diseases and these missplicing events are located in other regions of PRP8 (Kurtovic-Kozaric et al., 2014). PRP8 accomplishes other functions within the spliceosome besides regulation of SNRNP200 function during assembly and activation. PRP8 also acts as a structural support for snRNAs that holds the pre-mRNA in the correct position, most likely through a RNA recognition motif (RRM) located at the center of the protein (Pena et al., 2007; Ruzickova & Stanek, 2016). Thus, these other functions of PRP8 may have implications in splicing fidelity, while the function of the C-terminal region of the protein might be restricted to regulation of SNRNP200 and therefore assembly and activation of the complex.

Uncovered s-adRP mutations

Since all s-adRP proteins function in assembly and activation steps of splicing, future studies should focus on spliceosome assembly rather than in other steps like intron recognition or removal (catalysis) to

unmask common mechanisms of action fueling the pathology. Interestingly, there are still adRP patients whose disease causing mutation has not been identified, other genes encoding proteins required for the tri-snRNP U4/U6·U5 formation, such as SNU114 (U5 snRNP) or SNU13 (U4 snRNP), are candidates to be screened for mutations responsible of adRP (de Sousa Dias et al., 2013).

D.3. Why is *C. elegans* a good model to study adRP?

As mentioned before, basic cellular processes like splicing and apoptosis are well conserved between the worm and humans. We propose the use of *C. elegans* as a model to investigate the cellular mechanisms causing s-adRP in which the final outcome is photoreceptors death by apoptosis in the retina. Such a proposal is strongly supported by the following arguments:

The partial loss-of-function caused by mutations in s-adRP genes can be easily mimicked in the worms using RNA mediated-interference (RNAi) (Longman, Johnstone, & Cáceres, 2000). Phenotypic and molecular characterization of s-adRP*(RNAi)* worms helped us to establish a model for adRP which so far has shown similarity with what is observed in other studies of s-adRP. Regarding splicing efficiency, RNA-seq analyses of s-adRP*(RNAi)* worms evidenced mild intron retention similar to what has been observed in other s-adRP models (Gamundi et al., 2008; Tanackovic, Ransijn, Thibault, et al., 2011).

More importantly, RP is a tissue-specific disease that affects only some cells in the retina while cells in other tissues remain unaffected. Similarly, our worm model display a tissue-specific apoptotic response inducing the ectopic expression of the pro-apoptotic gene *egl-1* and the onset of apoptotic cell corpses in developing hypodermal cells. The fact that *C. elegans* is a multicellular organism allows us to study tissue-specific responses to impairment of s-adRP genes in a relatively simple model. Having differentiated tissues allows to identify the effect of

partial loss-of-function mutations in specific-cell-types that may have special molecular requirements as high transcriptional activity or the ability to divide during post-embryonic development.

CRISPR

New genome editing techniques like CRISPR/Cas9 are available for *C. elegans*. Using this technique we have been able to generate four mutants, three of them carrying specific s-adRP mutations. Mutated s-adRP residues are usually conserved from yeast to humans, this could be an indication of an important role for the function of the protein (Pena et al., 2007). Therefore, being able to change a specific residue in the *C. elegans* genome gives us the opportunity to develop a more specific model.

Why the development of "personalized models" is so important?

Several mutations have been identified for each of the s-adRP genes, and the prognosis of the disease usually varies if the patient carries one allele or another. For example, prognosis and outcomes for PRPF8 RP patients differ with different mutations, with H2309P or H2309R having a worse prognosis than R2310K. This correlates with the observed difference in growth defect severity in yeast lines carrying the equivalent mutations (Towns et al., 2010). As mentioned before, PRPF8 adRP mutations are classified according to its position and presumed function within the C-terminal domain of the protein. Type I mutations like H2309R, are expected to destabilize the Jab1 domain´s structure, and it has been reported to abolish the interaction between PRP8 and SNRNP200 leading to defects in U5 snRNP maturation and reduced tri-snRNP levels in yeast. By contrast, type II residues are expected to mediate direct interactions between PRP8 and

SNRNP200, and mutations like R2310G or R2310K have only shown a reduced interaction between these proteins, resulting in a less amount of U4/U6·U5 tri-snRNP and reduced SNRNP200 activity (Mozaffari-Jovin et al., 2014).

This is also consistent with the fact that *cer22*, R2310/2303G, does not display an obvious phenotype in our *C. elegans* CRISPR mutants. It is possible that this type II mutation only slightly affects PRP-8-SNRP200 interactions, giving the splicing system in the worm the opportunity to overcome this defect. On the other hand, H2309/2302del does show stronger phenotypes which could be related with the complete destabilization of PRP-8 and SNRP200 interaction. However, we have to consider that H2309/2302del represents the total loss of the amino acid instead of just a substitution, a new CRISPR mutant should be generated targeting this specific residue to confirm this parallelism.

Mutations in SNRNP200/Brr2 also display different onset timing depending on the mutation, being S1087L less harmful than other residues (Zhao et al., 2009). It has not been described whether V683L produces a more severe onset than S1087L or if it is the other way around. However, S1087L is located in the Sec63 domain of the active N-Terminal module of SNRNP200, while V683L is located in the actual DExD/H box ATPase domain (which contains the RecA-like domains) (Fig.R.15). Even when Sec63 domain is required for the ATPase and unwinding activity of SNRNP200, it is the DExD/H box who catalyses the process, thus a severe phenotype in our CRISPR V683/676L mutants would be justified.

In summary, we believe that these CRISPR s-adRP mutants would help to develop "personalized" models to understand specific mechanisms of the disease and to screen for drugs and genetic modifiers.

Seam cells and retinal cells

Overall, so far the observations in our *C. elegans* model seems to be consistent with what other studies have been reporting and the worm rises like a good model for retinitis pigmentosa. However, there are several points to consider if we are trying to extrapolate our observations. It is clear that we cannot make a direct parity between the *C. elegans* hypodermis and a human retina, neither between a seam cell and a photoreceptor. Even when seam cells are ´special´ cells within the worm´s cellular lineages, it cannot be compared with the complexity a human photoreceptor harbors. Nevertheless, the basic nuclear processes and their gene expression machinery are basically the same, thus we expect to be able to extrapolate the observations we made at the molecular level.

Time is an important factor to take into account to understand this disease. As mentioned above, retinitis pigmentosa first symptoms appear generally at the adolescence when night blindness begins to be evident. Then, loss of visual field is gradual and takes several years (30-50 years), by contrast, the phenotypes we observe in *C. elegans* are detected within 2-5 days (~2 days for apoptosis in the seam cells and up to 5 days for sterility phenotypes). We think that replicative capacity of dividing cells, may act as an accelerator of the cell failure onset, thus we can observe these deleterious effect within a shorter

period of time. One thing that we also have to consider is that the life cycle timing of both organisms is very different and that the life time of a cell in *C. elegans* is much shorter that the life of a photoreceptor. Future studies can focus on the H0 seam cell, which is a seam cell that does not present postembryonic divisions and may help to compare effects of splicing mutations in differentiated cells *vs* dividing cells.

D.4. How do ubiquitously expressed and essential genes cause a cell-type-specific effect when inactivated?

Several hypotheses have been formulated to address this question and probably most of them are valid considering the complexity of the retina and the diversity of mutations leading to s-adRP. Taking into consideration the published data and our study, we find three compatible possibilities, which are: defects in splicing reactions, reduced transcriptional efficiency, and genome instability through the formation of R-loops.

D.4.1. Is adRP caused by splicing defects?

The most straightforward explanation is that splicing is defective in the retina, where cells present a high transcriptional activity, causing a reduction in the mRNA production and consequently a deficit in the amount of proteins needed for its adequate functioning.

As cited above, some studies, mostly in lymphoblastoid cells lines of s-adRP patients, showed that these s-adRP mutations can produce defects in splicing of some genes (Ivings et al., n.d.; Tanackovic, Ransijn, Thibault, et al., 2011). However, these defects in constitutive splicing are not general as they are usually present in a small proportion of the intronic units tested. Moreover, in a specific study where five out of 57 intronic units presented splicing defects, only three of those five intron retention events were common for all the PRPF mutations tested (Tanackovic, Ransijn, Ayuso, et al., 2011;

Tanackovic, Ransijn, Thibault, et al., 2011). Although a mild intron retention has been observed in s-adRP models (Tanackovic, Ransijn, Thibault, et al., 2011; Wilkie et al., 2008), a correlation between a slight intron retention in certain genes and retinal degeneration has not been clearly established yet (Mordes et al., 2006). Nevertheless, being the retina such a complex tissue, we have to consider that maybe just a couple of misspliced cell-type-specific genes can be enough to drive the cell to total failure, but so far this cause/consequence relation has not been proved.

Besides the analysis of the RNA-sequencing in which we observed a low intron retention in all our s-adRP genes depleted samples, the RT-PCR and intron retention events search in a genome browser (i.e., a *fmo-5* intron, Fig.R.6), helped to conclude that the amount of unspliced transcripts produced upon a partial depletion of s-adRP genes is low. Upon RNAi, expression levels of the s-adRP targeted genes were reduced in more than 50% (Fig.R.1), even then, the spliceosome machinery seems to have a decent performance. This can indicate a high buffering capacity of spliceosomal proteins to maintain their essential functions despite an important reduction of their activity. It may be possible that any transcript important for the viability of photoreceptors does not have such buffering capacity and a small reduction of mature mRNA can have catastrophic consequences. Again, this is pure speculation at the moment. By s-adRP RNAi we did not detect a prominent reduction in the abundance of any particular subset of transcripts. The transcriptomes of s-adRP CRISPR mutants that we generated may shed some light into this question.

Do s-adRP mutations cause alternative splicing defects?

The alteration of these adRP splicing components may affect alternative splicing (AS) as well, and thus the balance among isoforms. In fact, it has been shown that AS is affected in cell lines derived from patients carrying s-adRP mutations. However, the alterations in the levels of alternative transcripts were only observed in PRPF31 mutant cell lines and only three significant alternative splicing events were identified out of 96 tested, while PRPF8 and PRPF3 cell lines did not show any alternative splicing event. Thus, PRPF mutations do not seem to share a common effect in AS between each other (Tanackovic, Ransijn, Thibault, et al., 2011). On the contrary, a recent report based on siRNA in cell lines has uncovered a link between patterns of AS alterations and functionally related spliceosome components (Papasaikas, Tejedor, Vigevani, & Valcárcel, 2015). To explore the effect of our RNAi assays on AS we checked a panel of eight AS events occurring at L2 and L3 larval stages (A. K. Ramani et al., 2011) and did not find common AS alterations in s-adRP RNAi samples (Rubio-Peña et al, 2015).

In any case, it is difficult to justify *C. elegans'* phenotypes, as well as the retinal degeneration in humans, just by slight intron retention or by a modest unbalance of alternative transcripts, which argues in favor of a more complex mechanism. Nevertheless, the presence of an AS event specifically associated with one of the s-adRP mutations does not argue with a hypothesis where a common defective assembly and activation of the spliceosome would be the main cause of adRP. Actually, the specific additional effect of a mutation in constitutive or

alternative splicing would explain the different onsets observed in adRP patients carrying different mutations.

D.4.2. adRP mutations may affect transcriptional efficiency

It is well established that splicing and transcription are coordinated processes and that alterations in the spliceosome can affect the transcriptional machinery (Das et al., 2006; Fontrodona et al., 2013; Montecucco & Biamonti, 2013). As an example of such co-transcriptional splicing in a developing multicellular animal, our laboratory demonstrated that RSR-2 interacts with PRP-8 and the RNA polymerase II (RNA Pol II) in *C. elegans*, and that the impairment of RSR-2 functions results in a lower transcriptional efficiency (Fontrodona et al. 2013). Consistent with this, the mutant allele *prp-8(rr40)*, which reduces the levels of wild-type PRP-8, diminishes the production of highly expressed germline transcripts, but it does not affect the splicing reaction (Hebeisen et al. 2008). In transcriptionally active tissues the necessity of splicing factors is bigger, so a deficient spliceosome assembly may cause a reduction in transcriptional efficiency as well.

The retina is a highly metabolically active tissue. Large amounts of proteins are synthesized continually as the outer segments of photoreceptors have to be rapidly renewed (Ferrari et al., 2011). This tissue contains the highest volume of processed pre-mRNAs within the whole body, as measured by the amount of spliced genes that were common to 31 human tissues (Tanackovic, Ransijn, Thibault, et al., 2011).

Similarly, *C. elegans´* hypodermal seam cells are highly transcriptionally active cells as they have to produce the cuticle of the worm between each larval stage (Johnstone, 1994). RP mutations may compromise splicing efficiency; hence, a less efficient splicing resulting from a s-adRP mutation could be sufficient to fulfill metabolic requirements in other tissues but insufficient in the retina or the hypodermal seam cells.

D.4.3. s-adRP mutations can lead to genomic instability

We think that a reduction in the activity of s-adRP proteins could produce genomic instability that, in cells with high transcriptional activity, can contribute to programmed cell death.

R-loops form naturally during transcription by the hybridization of the nascent RNA with the template DNA strand leaving a single stranded DNA (ssDNA) exposed. These RNA:DNA structures are prevalent in mammalian genomes and are known to participate as regulatory elements in several molecular processes, however, their persistent formation can be a risky outcome with deleterious effects on genome integrity (Skourti-Stathaki & Proudfoot, 2014).

It has been shown that highly transcribed loci may be more prone to R-loop formation, as it was evidenced by the binding of the yeast´s R-loop suppressing factor Npl3 preferentially at the most highly transcribed genes (Santos-Pereira et al., 2013). Additionally, studies in which transcription was increased upon induction by the hormone estrogen (E2), which binds the estrogen receptor to promote

transcription of E2-responsive genes, increased R-loop formation that may result in genomic instability (Stork et al., 2016).

s-adRP mutations, R-loops and genomic instability

Knowing that transcriptional stalling can create RNA:DNA hybrids (Canugovi, Samaranayake, & Bhagwat, 2009), led us to think that an inefficient splicing that affects transcriptional elongation, would produce an accumulation of R-loop structures. This accumulation of R-loops may be higher in transcriptionally active tissues which would lead to a tissue-specific threat to genome integrity.

R-loops can act as genomic instability effectors in active transcriptional regions as they expose the single-strand-DNA (ssDNA) to DNA-damaging agents such as UV light (Aguilera & García-Muse, 2012). This would explain the sensitivity of *prp-8(RNAi)* animals to UV exposure (Fig.R.24). Moreover, R-loops are also a source of transcriptional associated mutagenesis (TAM), which is elevated under high transcriptional activity (N. Kim, Abdulovic, Gealy, Lippert, & Jinks-Robertson, 2007) and can also induce apoptosis (Hendriks, Jansen, Mullenders, & de Wind, 2010).

In addition to the formation of these DNA damage-prone sites, R-loops can contribute to genome instability because replication and transcriptional machineries could meet and collide at these R-loop sites leading to replicative stress (Brambati, Colosio, Zardoni, Galanti, & Liberi, 2015).

In *C. elegans* the collision with the replication machinery is feasible because hypodermal cells are actively dividing during larval stages and also suffer a round of endoreplication (Hedgecock & White, 1985; J. E. Sulston & Horvitz, 1977). This transcription-replication collision

would be consistent with the replicative stress sensitivity we observe in our CRISPR s-adRP mutants. It would also explain why *atl-1*/ATR is upregulated in s-adRP genes RNAi worms, as *atl-1*/ATR is known to respond to replicative stress induced DNA damage (Garcia-Muse & Boulton, 2005).

atl-1 and *egl-1*

Similarly to *s-adRP(RNAi)* animals, our CRISPR mutants also showed upregulation of *atl-1*/ATR and *egl-1* expression. *egl-1* transcription is dependent of *cep-1*/p53. It has been shown that upon DNA damage, the phosphorylation of Ser-15 may be a critical event in the activation and upregulation of *cep-1*/p53 and that this phosphorylation is carried out by *atl-1*/ATR (Tibbetts et al., 1999). Once *cep-1*/p53 is activated it can induce the transcription of *egl-1* to trigger apoptosis. We think that, at least in part, this could be the apoptotic pathway followed by some of the cells that die in s-adRP genes deficient background.

snrp-200(cer23) CRISPR mutant, which causes the amino acid change V683/676L located in the DExD/H Box domain, showed higher upregulation of *atl-1* and *egl-1*, as well as a higher sensitivity to replicative stress than *prp-8* CRISPR mutants. This would be consistent with the fact that the Jab1/MPN domain of PRP8 and its C-terminal tail act as regulators of SNRNP200/Brr2 activity. Thus, *prp-8* CRISPR mutants might have a milder effect because even when the regulation of *C. elegans* SNRP200 is compromised, its helicase activity still works being the effect more drastic when the actual active site is defective.

Genetic experiments should confirm this hypothesis since the roles of *atl-1* and *cep-1*/p53 in distinct types of somatic cells are still unclear.

D.5. Working model

So far our observations direct us towards a model in which mutations in s-adRP genes may produce an active but inefficient splicing that, in transcriptionally highly active tissues, unchains a number of unfortunate events that may include a decrease in the efficiency of transcription and RNA Pol II stalling, that would lead to R-Loops accumulation, replicative stress and genomic instability that would finally trigger the apoptotic response (Fig.D.2.).

In this model, replicative stress would play an important role. As mentioned in the introduction, seam cells undergo cell divisions in each larval stage transition (Fig.I.13), thus their DNA replicates. By contrast, photoreceptors do not divide in adult retina; actually, it is still not completely clear if any cell type within the retina has the capacity to do it.

However, the presence of ssDNA produced by R-loops during decades could be enough genomic instability to trigger apoptosis in photoreceptors. Thus, our model does not necessarily need a dividing tissue to explain the onset and development of the disease.

Then we should also discuss the fact that none other tissue seems to be affected in adRP patients. For example, epithelial cells from the skin are constantly dividing to renew layers (Le Bras & Le Borgne, 2014). But no effect of s-adRP mutations in this tissue has been reported or in other tissues with a high transcription activity like the testicles. In any case, the retina is 7-fold more transcriptionally active

than testicles (Tanackovic, Ransijn, Thibault, et al., 2011). Thus, a highly active gene expression machinery seems to be more relevant to the disease that the presence of dividing cells.

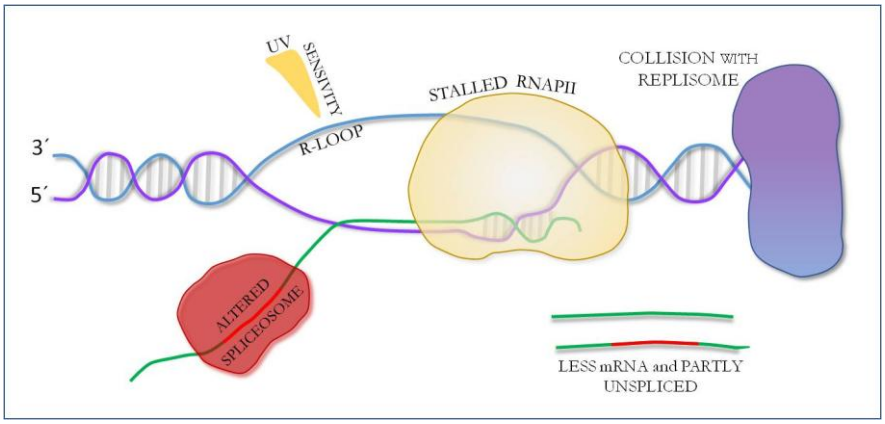

Fig. D. 2. Working model for tissue-specific apoptosis in s-adRP
When spliceosome is altered, the activity of the RNA Pol II is affected and R-loops (RNA:DNA hybrids) may accumulate, leaving single-strand DNA regions that are more sensitive to DNA insults. Moreover, R-loops may cause collisions between the transcriptional and replicative machineries.

R-loop detection

R-loop detection in our model system is of key importance to support our hypothesis. Basically, R-loops identification relies on the detection of these RNA:DNA hybrids by the specific antibody S9.6, which is commercially available. We made several attempts to perform immunofluorescence (IF) of embryos, gonads and larvae but we did not manage to get a reliable signal not even in a control strain carrying mutations for *met-2* and *set-25* that has been reported to present R-loops and showed IF signal in embryos (Zeller et al., 2016). We are also optimizing the dot-blot technique, in which samples are transferred into nylon membranes for posterior immunodetection of R-loops with the S9.6 antibody. Moreover, we plan to perform a

DRIP-seq (DNA-RNA hybrid immunoprecipitation) in the mutant samples. This technique will allow us to identify where the putative R-loops are accumulating, helping in the identification of sensitive spots in the genome. Finally, the presence of other proteins at R-loops (ex. Histone 3 phosphorylated at serine 10) may help as epitopes to study these DNA-RNA hybrids (Castellano-Pozo et al., 2013).

To date, we have at least an indication of the presence of R-loops in our experiment where we exposed *prp-8(RNAi)* worms to UV light. The sensitivity of this worms to UV, evidenced by a higher amount of *foci* formation that was not only present in germline but also in somatic cells, could be an indicative of ssDNA presence that would be a product of an accumulation of R-loops. In fact, indirect detection of R-loops by ssDNA detection has been reported before. Yu *et al*, used the mutation profile caused by sodium bisulfate on ssDNA where it acts specifically and converts C to U nucleotides. The mutation profile of the nontemplate ssDNA region allows the inference of the average length of R-loops as long as such a profile is abolished by RNase H1 treatment (Aguilera & García-Muse, 2012; Yu, Chedin, Hsieh, Wilson, & Lieber, 2003).

D.6. Present and future of Cellular and Gene therapies for Retinitis Pigmentosa

Stem Cell research is having a great development in the last few years. As signal of the potential of this research, there are many research institutes that were inaugurated just to work on this topic. Source of stem cells can be Embryonic Stem Cells (ESCs) but conveniently Induced Pluripotent stem cells (iPSCs) can be generated from a variety of somatic cells of patients, to later be differentiated to distinct cell types using a cocktail of transcription factors and even chemicals (Hou et al., 2013; M. Li & Belmonte, 2016). Cell types that can be obtained from iPSCs include RPE (Retinal Pigment Epithelium) and neuroretinal cells (Buchholz et al., 2009; Wiley, Burnight, Mullins, Stone, & Tucker, 2014). Such technology is a double-edge sword hitting both research and therapy.

Previously, to study retinitis pigmentosa in patient cells, lymphoblasts were used to obtain cells with an identical genetic background (Tanackovic, Ransijn, Thibault, et al., 2011). These days, by using iPSCs researchers have the opportunity of investigating cells with identical genetic background but also similar molecular pathways since iPSCs are differentiated *in vitro* to retinal cells. Moreover, iPSCs can be differentiated *in vitro* and stimulated to produce 3D organoids. These organoids will never resemble the defined retinal structure but it may be a useful near-physiological model to study RP disease mechanisms independent of retinal structures (Fatehullah, Tan, & Barker, 2016).

Besides its value as scientific tool, stem cells could be used in regenerative cell-based therapies once differentiated to the cell type that needs to be restored. There are ongoing clinical trials with ESCs and with iPSCs (Labrador-Velandia et al., 2016; Tang et al., 2017). If the source for differentiated retinal cells is ESCs, there are concerns about tumourigenicity and the continuous use of immunosupressors. Thus, novel sources of stem cells are being established as those derived from postnatal retina of animal models (RPECs) and developmentally mature organs (MSCs).

CRIPSR is being a revolution in all the scientific fields and this technology can be coupled to stem-cell based therapies since a mutation could be repaired leaving intact the remaining genetic background (Zheng, Li, & Tsang, 2015). Despite all these new technologies, correction of the vision in patients is still a challenge and different handicaps need to be solved such as immunomodulation, maintenance of the transcription factor activity to maintain the differentiated state, define management of vascular structures or fine tuning in the differentiation process to the distinct retinal cell types (Aharony, Michowiz, & Goldenberg-Cohen, 2017). However, in the near future this therapeutic pipeline may occur: adult cells from an RP patient are edited by CRISPR to restore missense mutations, and then derived to retinal cells that can maintain the differentiated state and can be implanted into the patient's eyes to recover, at least partially, his/her previous vision (Fig. D. 3).

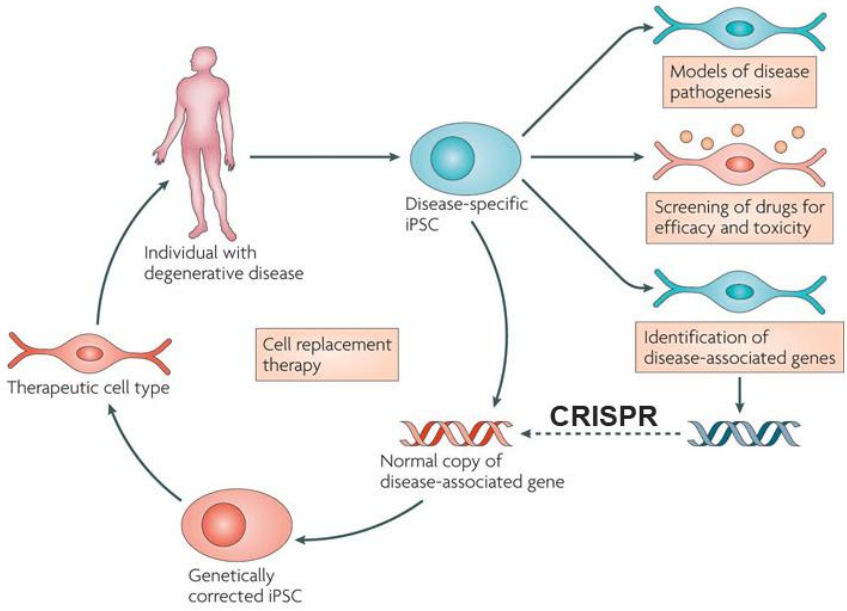

Fig. D. 3. iPSC use for research and therapy

Patient-derived iPSCs are a good system to investigate the mechanisms of the disease but also to screen for drugs and to develop gene therapies (*Adapted from* Fairchild, 2010).

CONCLUSIONS

CONCLUSIONS

1. Partial inactivation by RNAi of U4 and U5 snRNP splicing factors related to autosomal dominant retinitis pigmentosa (s-adRP) permits the identification of two phenotypic clusters. This phenotypic distinction may be related with the function of each splicing factor within the splicing process.

2. RNAi of s-adRP genes produces (i) low intron retention events in *C. elegans*, which may be associated with transcriptional activity; and (ii) a tissue-specific apoptosis in *C. elegans´* hypodermal seam cells.

3. Apoptosis in *C. elegans´* seam cells upon RNAi of s-adRP genes is triggered by the upregulation of the pro-apoptotic gene *egl-1*, which encodes a conserved BH3-only domain protein. The pathway that activates *egl-1* expression remains unknown, however, the DNA damage sensor *atl-1*/ATR is also upregulated in these worms so there is the possibility that both genes work with the same pathway to trigger apoptosis.

4. C. elegans CRISPR mutants carrying specific s-adRP mutations can be maintain in homozygosis, thus, they are not lethal. s-adRP mutations produce phenotypes with different degrees of severity. This is consistent with the different onsets and time of progression that different s-adRP mutations produce in humans.

5. Similar to s-adRP genes RNAi treated worms, *C. elegans* carrying s-adRP mutations also present upregulation of *egl-1* and *atl-1*. Additionally, replicative stress sensitivity is also observed in these worms, which is an indicator of genomic instability.

6. Generation of CRISPR mutations exhibit the potential of a platform where to make high-throughput screenings looking for modifiers of the disease.

MATERIALS AND METHODS

MM.1. Strains and general methods

Worms were cultured and managed following standard methods (Brenner, 1974). Thus, worms were grown on NGM (Nematode Growth Media) agar plates previously seeded with an overgrown culture of the *Escherichia coli* strain OP50 at temperatures between 15 and 25°C depending on the experiment. Synchronization of worms was done following the sodium hypochlorite treatment (Porta-de-la-Riva et al., 2012).

Unless otherwise specified, Bristol N2 was used as the wild type strain. Mutant and transgenic strains can be found in table MM1.

Strain	Genotype	Source
TR1331	*smg-1 (r861)* I.	CGC
SS104	*glp-4(bn2)* I.	CGC
CER87	*unc-119(ed3)* III; *cerEx79* [P*prp-31::GFP·H2B::prp-31 3'UTR+pGH8+unc-119(+)*].	JC
CER148	*unc-119(ed3)* III; *cerIs05* [P*prp-31::GFP·H2B::prp-31 3'UTR +unc-119(+)*].	JC
CER149	*unc-119(ed3)* III; *cerIs06* [P*prp-31::GFP·H2B::prp-31 3'UTR +unc-119(+)*].	JC
VC2709	*prp-31(gk1094) I/hT2* [*bli-4(e937) let-?(q782) qIs48]* (I;III).	CGC
CER100	*prp-31(gk1094)/ unc-57(ad592) dpy-5(ed1)* I.	JC
CER101	*+/ unc-57(ad592) dpy-5(ed1)* I.	JC
VC3460	*prp-8(gk3511)* III/*hT2* [*bli-4(e937) let-?(q782) qIs48*] (I;III).	CGC
BC12486	*dpy-5(e907)* I; *sEx12486* [P*prp-8::GFP + pCeh361*]	CGC
WS1973	*unc-119(ed3)* III; *opIs56* [P*egl-1::2xNLS::GFP + unc-119(+)*]	CGC
MT4970	*ced-6(n2095)* III	CGC
NR222	*rde-1(ne219)* V; *kzIs9* [*pKK1260(lin-26p::nls::GFP) + pKK1253(lin-26p::rde-1) + pRF6(rol-6(su1006)*],	CGC
NR350	*rde-1(ne219)* V; *kzIs20* [*hlh-1p::rde-1 + sur-5p::NLS::GFP*]	CGC
WS4581	*unc-119(ed3)* III; *opIs263* [*rpa-1p::rpa-1::YFP+ unc-119(+)*].	CGC
ML1651	*mcIs46* [*dlg-1::RFP + unc-119(+)*]	CGC
JR667	*unc-119(e2498::Tc1)* III; *wIs51* [*SCMp::GFP + unc-119(+)*] V.	CGC
CER209	*opIs56* [P*egl-1::2xNLS::GFP + unc-119(+)*]; *mcIs46* [*dlg-1::RFP + unc-119(+)*]	JC
CER240	*prp-8(cer22)* III	JC
CER255	*prp-8(cer14)* III	JC
CER248	*snrp-200(cer24)* II	JC
CER256	*snrp-200(cer23)* II	JC

Table MM. 1. List of strains used in this study.
CGC: Caenorhabditis Genetic Center (Minnesota, USA), JC: Julián Cerón´s Lab

MM.1.2. Genotyping of mutant alleles

Preparation of worm lysates:

Single individuals or pools of worms were transferred into 10µl Lysis buffer with a worm picker. To crack the tissue, 3 freeze/thaw cycles were performed using dry ice combined with 96% EtOH, and a water bath at 37°C. The samples were subsequently incubated in a thermalcycler for 1h at 60°C and 15min. at 95°C to inactivate Proteinase K.

PCR for amplification of WT and mutant sequences:

For each sample a master mix was prepared. All required components came from the MyTaq™ Mix (Cat. No. BIO-25041), and were handled according to the manufacturer's instructions. PCR reactions were performed according to the annealing temperature of the respective primers and the length of the amplicon. Table MM.2 shows the list of primers used to genotype strains. Reference primers were used to genotype mutant alleles from CGC strains and new specific primers were designed to genotype strains generated in JC's lab.

Strain	Gene	Allele		5´-Sequence-3´
VC2709	*prp-31*	*gk1094*	Fw	TTTCATTCTCTTCGCGACCT
CER100			Rv	CACACTCCAGCACTGGAAAA
VC3460	*prp-8*	*gk3511*	Fw	ACAACAACTGGGAAACTCCG
			Rv	ACTGAACATGGGCATCAACA
WS1973	P*egl-1*		Fw	TCTGGATGTGTTCGTGGTGT
CER209	::GFP		Rv	GGTCTGCTAGTTGAACGC
CER240	*prp-8*	*cer22*	Fw	TATGGTGTATCGCCACCTGA
			Rv mut	GAAGTGAACCGGGCCGTG
			Rv wt	GGAAGTGAACCGGCCGATG
CR255	*prp-8*	*cer14*	Fw	TATGGTGTATCGCCACCTGA
			Rv mut	TTATGGAAGTGAACCGGCCC
			Rv wt	GGAAGTGAACCGGCCGATG
CER248	*snrp-200*	*cer24*	Fw	CTCATCACAAATCACTCGGAG
			Rv mut	TACTGCTGTTCCAGAGGGAG
			Rv wt	TATATTGTTGCTCGAGGGGAAC
CER256	*snrp-200*	*cer23*	Fw	GTGATCGTTTGTACGCCGGA
			Rv mut	GAAGAGTCGTCCGGCGAGT
			Rv wt	GAAGAGTCGTCCGGCACTC

Table MM. 2. Primers used in this study to genotype mutant and reporter strains.

Fw: forward; Rv: reverse; mut: mutant; wt: wild type

MM.2. RNA-mediated interference (RNAi)

By feeding

RNAi by feeding is a technique that permits the partial inactivation of a gene of interest by administration of the dsRNA through the food. When large number of animals needs to be treated at once, this delivery method is the one usually used.

To induce RNAi by feeding, nematode growth medium (NGM) plates were supplemented with 50 μg/mL ampicillin, 12.5 μg/mL tetracycline, and 1 mM IPTG. NGM-supplemented plates were seeded with the corresponding RNAi clone. All RNAi clones but the one corresponding to *snrp-200* were obtained either from the ORFeome library (Rual et al. 2004) or the Ahringer library (Kamath et al. 2003). RNAi cultures were prepared by growing a colony of each clone in 4ml of LB plus 50μg/ml of ampicillin and 12.5 μg/ml of tetracycline at 37°C O/N in agitation. 60mm plates were seeded with 400μl of the culture and induced O/N at room temperature. To study the RNAi phenotypes of s-adRP genes, a synchronized population of N2 worms at L1 stage was fed with each clone and grown at 20°C.

By microinjection

To induce RNAi by microinjection, dsRNA was synthesized from the RNAi feeding clone by using MEGA script T7 kit (Ambion Cat. No. AM1333). The vector L4440 used for both *C. elegans* RNAi libraries contains two RNA polymerase T7 promoters flanking the multicloning site that allow the synthesis of the transcribed *sense* and

antisense strands of the cloned fragment. Thus, RNA clones of the feeding libraries can be used as templates for the synthesis of dsRNA. Wild-type young adults were injected with 1 µg/µL of dsRNA.

MM.2.1. Generation of *snrp-200* clone by gateway

From cDNA of wild type worms we generated an RNAi clone (pCUC24) for *snrp-200* by Gateway cloning (amplicon size 1091 bp), which includes the spliced sequence from 106 to 1197 bp. In this case the destination vector was pGC49. The clone was validated by PCR and sequencing.

Primers to generate the cloning fragment for *snrp-200*		
Y46G5A.4	Fw attB1	ggggacaagtttgtacaaaaaagcaggctCCGACTGGCGAGGTGCTGC
	Rv attB2	ggggaccactttgtacaagaaagctgggtGTCGGCACGCGCGGAG

Table MM. 3. Primers used to obtain the cloning fragment for snrp-200.
PCR amplification fragments are indicated in capital letters, plasmid tails are in lowercase letters.

MM.3. Microinjection

The injection solution, that can be either, the dsRNA for RNAi, the transgenesis mix or the CRISPR/Cas9 mix, was loaded into the injection needle (Femtotips® microinjection capillaries, Eppendorf, Cat. No. 5242 952.008) using a special pipette tip (Microloader, Eppendorf, Cat. No. 5242 956.003). The needle was installed in the micromanipulator and the tip, if required, opened by pushing it slightly against a glass cover slight.

Animals in young adult stage were selected and prepared by washing them in M9 to get rid of bacteria and then seeded in NGM plates without food. Once they air dry, they are ready to microinject. Worms are transferred into a drop of Halocarbon oil 700 (Sigma; Cat. No. H8898) on an agarose pad. The worms were allowed to gravity settle in the drop and slightly pressed to the surface of the agarose pad with a picker to achieve their attachment to the agarose surface.
The injection solution was injected into the gonads of the immobilized worms.

After injecting, M9 was applied onto the Halocarbon oil drop, so that the worms could be released from the oil. The animals were carefully transferred into a drop M9 on a NGM plate and allowed to recover at 15 °C or 25°C o/n depending on the purpose of microinjection. The following day they were singled out and, in order to synchronize the progeny, transferred to new plates twice a day over two days. The phenotypes were studied in the F1 progeny.

MM.4. RNA sequencing

RNA-seq analyses were performed using RNA of the following population sets: synchronized N2 and *smg-1(r861)* L1 larvae fed for 24 h at 20°C with the RNAi clones of *prp-6*, *prp-8*, *prp-31*, and *gfp*; and 5-d adult *glp-4(bn2)* worms grown at 25°C and fed first, for 72 h with OP50 and next, for 48 more hours with the RNAi clones of *prp-8*, *prp-31*, and *gfp*.

RNA was isolated with TRI Reagent (MRC, Inc.) and purified by using the Purelink RNA Mini kit (Ambion) and Purelink DNAse (Ambion). RNA samples were subjected to quality and yield controls on the Agilent 2010 Bioanalyzer and Nanodrop Spectrophotometer, respectively. Poly(A)-enriched samples were multiplexed in libraries for paired-end RNA sequencing on Illumina Hiseq 2000 platform, through the CNAG (Centro Nacional de Análisis Genómico) sequencing facility.

To generate BAM files, more than 50 million reads (length of 73 bp) for sample were mapped against the *C. elegans* worm version WS236 following the GEMTools pipeline (http://gemtools.github.io/). These BAM files were analyzed using the SeqSolve software, which use Cufflinks/Cuffdiff for differential gene/transcript expression analyses (Trapnell et al. 2012). A full 98.7% of the reads from the control *gfp(RNAi)* sample mapped in exons.
The sequence data for the 11 transcriptomes analyzed in this study are available at the NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) un- der accession number GSE72952.

MM.5. Quantification of gene expression by real time PCR

RNA isolation:

Worms were harvested with M9 and washed several times (1.500rpm; 1min.). To remove bacteria from the gut, the animals were incubated for 30min at RT with agitation and washed again. The worm pellet was resuspended in 7x volume of TRI Reagent® (Cat. No. TR-118) and vigorously vortexed. Afterwards 5 freeze/thaw cycles were performed using dry ice combined with 96% EtOH and a water bath at 37°C. The suspension was vigorously vortexed for 30sec., and let stand at RT for 30sec. This step was repeated 5 times.

After 5min. incubation at RT, the sample was mixed with chloroform and vortexed. The quantity of chloroform depended on the amount of TRI Reagent® initially used (0.2ml chloroform/1ml TRI Reagent®).

Phase separation was achieved by incubation of 15min. at RT, followed by centrifugation (13.200rpm; 15min; 4°C). The aqueous phase was transferred to a new Eppendorf, united with isopropanol (0.5ml/ml TRI Reagent®) and incubated at RT for 10min. Afterwards the RNA was pelleted (13.200rpm; 30min.; 4°C) and washed in 75% ethanol (diluted with DEPC-water) (9.000rpm; 5min.; 4°C). The pellet was air-dried and subsequently resuspended in 30µl DEPC-water. The concentration and purity of the RNA was determined using a Nanodrop ND-1000 spectrophotometer.

DNase treatment:

To eliminate DNA contaminations, the DNAse I Amplification Grade system (Invitrogen™, Cat. No. 18068-015) was used. 1µg RNA was united with 1µl DNAse I in 1xDNAse I Reaction buffer (total reaction volume 10µl) and incubated for 30min. at 37°C. Afterwards 1µl 25mM EDTA was added and the sample incubated for 10min. at 65°C to stop the reaction.

cDNA synthesis:

For each sample a master mix was prepared using the RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Scientific, # K1612) and incubated for 1h at 42°C and 5min. at 25°C.

Amplification of cDNA:

To avoid the amplification of contaminant genomic DNA at least one primer was designed to span an exon/exon junction. cDNA of heterozygote or wild type animals was used as control.

Preparation of samples:

cDNA samples are diluted 1:4 for genes of interest and 1:50 for reactions targeting housekeeping genes.

Primer design:

Primers were designed that span exon/exon junctions to avoid the amplification of genomic DNA. All primers had approximately the same annealing temperatures (ideally 60°C), a GC-content of 40-60% and produced amplicons of roughly the same length (100-120bp).

RNAi efficiency quantification		
prp-8 qRT FW1	GTGGAATCACACCACTGCTC	Genomic DNA library
prp-8 qRT RV1 (Exon junction)	TGACAGCAGCACGTAATTCA	
prp-6 qRT FW1	ATTGGAGACGTACAACTGAAGAG	cDNA library
prp-6 qRT RV1 (3'UTR)	AAAGTGAGAAAAAAGGAAAAATCGATAG	
prp-31 qRT FW1	GGTCTAGAGATTATCAATCCAG	
prp-31 qRT RV1 (3'UTR)	GAATTCAGGAAGAGGAAAGTAAC	
snrp-200 qPCR Fw1 (Exon junction)	CCGATTCAGACACAGGTATTC	
snrp-200 qPCR Rv1	TAGACTGCCTTGGCTTCTGG	
Quantification of gene expression		
tbb-2 RTFW	TATGTGCCACGCGCCGTGTT	
tbb-2 RTREV	TTTCCGGCTCCGCTTTGTCCG	
egl-1 qPCRFW	ATGCTGATGCTCACCTTTGCC	
egl-1 qPCRREV	TGAGACGAGGAGTAGAACATG	
atl-1 qPCRFW	CTGTGATTGTAAGATTCTCAGTGG	
atl-1 qPCRREV	AATGAAGAACCTTGACTTTGTCGC	

Table MM. 4. List of primers used in quantitative PCR assays.

qPCR:

For the quantification of gene expression levels, Roche LightCycler 480 Instrument I and the LightCycler 480 SYBR green I Master Kit (Cat. No. 04 707 516 001) were used according to the manufacturer's instructions. Gene expression was normalized to transcript levels of the housekeeping gene *tbb-2*.

To exclude contaminations, for each primer pair a water sample was run in parallel. All experiments were performed in triplicates.

MM.6. Generation of transgenic reporter strains

The molecular construct to study the expression and localization of *prp-31* (pCUC32) was made using MultiSite Gateway Three-Fragment Vector Construction Kit (Invitrogen Cat. No. 12537-023).

This technology allows obtaining an expression vector generating and combining three Donor vectors (5', middle and 3') by a sequential series of recombination reactions. We combined three vectors: pDONRp4-P1R, that contains 664 bp of *prp-31* promoter region, pDONR201, containing the gfp·H2B (pCM1.35 plasmid), and finally pDONRP2R-P3, which contains 3´entry clone with 77 bp of the 3´UTR sequence of *prp-31*.

All reactions performed to generate and combine these vectors to eventually obtain the P*prp-31:gfp·H2B:prp-31 3'UTR* final construct (pCUC32 in the destination vector pCFJ150) were performed according to manufacturer's instructions (Invitrogen).

This final vector was transformed in Library Efficiency DH5α competent bacteria (Invitrogen), purified and verified by sequencing using M13 primers.

PCR amplification primers for *prp-31* fragements		
prp-31 PROMOTER	attB4 Fw	ggggacaactttgtatagaaaagttgattctcttcgcgacctgaag
	attB1R Rv	ggggactgctttttttgtacaaacttgtatttttaggtagtttaaagattaactg
prp-31 3´UTR	attB2R Fw	GGGGACAGCTTTCTTGTACAAAGTGGTATATTTATTTGTTACTTTCCTCTTCCTG
	attB3 Rv	GGGGACAACTTTGTATAATAAAGTTGtaaatgaaaaagtcgattcatttgg

Table MM. 5. Primers used to generate *prp-31* trascriptional reporter.
Blue: fragments corresponding to *prp-31* sequence

Plasmid	Expressing	Vector
pCUC30	*prp-31* promoter	pDONRp4-P1R
pCUC31	*prp-31* 3'UTR	pDONRP2R-P3
pCM1.35	GFP:H2B from pKS111-His	pDONR201
pCUC32	P*prp-31:gfp·H2B:prp-31 3'UTR*	pCFJ150

Table MM. 6. List of plasmids generated to study *prp-31* expression *in vivo* in *C. elegans*

Microinjection

Transgenic animals were generated by microinjecting 2ng/µL of pCUC32 [P*prp-31::GFP::H2B::prp-31 3´UTR+unc-119(+)*] together with 20ng/µL of the red neuronal marker pGH8 [P*rab-3::mCherry::unc-54utr*] into DP38 [*unc-119(ed3)*] young adult worms. 78ng/µL of 1 Kb Plus DNA ladder (Invitrogen) were used as carrier. Wild-type animals were selected over uncoordinated (Unc) worms, and mCherry expression was checked under the dissecting microscope. GFP expression was observed at high magnification using an inverted fluorescence Nikon ECLIPSE Ti-S microscope.

Bombardment

One and a half milligrams of gold powder (Chempur, gold powder 0.3–3.0 micron 99.9%) was transferred to a low protein binding 1.5-mL tube, resuspended in 150 µL of 50 mM spermidine, and sonicated for 10 sec. Twenty-four micrograms of pCUC32 DNA was added to the mixture, deionized water was added up to 540 µL, and the mixture was incubated for 10 min. Next, we added 150 µL of 1 M $CaCl_2$. The mixture was incubated for another 10 min and centrifuged at maximum speed for 30 sec. After three washes with 96% ethanol, the pellet was resuspended in 300 µL PVP 0.1mg/mL in EtOH. Shots with 20 µL of the suspension were performed using a Caenotec gene

gun on pre-chilled 35-mm plates with 20–30μL of DP38 [*unc-119(ed3)*] young adults. Then the NGM agar of each plate was sliced into six pieces, and each slice was placed on a 90mm plate.

These plates were incubated at 20°C for 2 wk and then screened for transgenic worms that rescued the *unc-119(ed3)* phenotype.

MM.7. Induction of replicative stress.

To induce replicative stress we used hydroxyurea (HU) (Sigma-Aldrich, Cat# H8627).

Replicative stress induction in *egl-1*::GFP Reporter

In all cases 25 mM HU plates were used and prepared adding 500 µL of 0.5 M HU to plates containing 10 mL of NGM medium and seeded with E. coli OP50. To observe *egl-1* expression, synchronized P*egl-1*::GFP (WS1973) L1 worms were grown in either HU or standard plates for 24 and 48 h at 20°C. After each time point, worms were recovered with M9 buffer, washed, and mounted using 0.3mMof levamisole for in vivo observation through an inverted fluorescence microscope.

Quantification of gene expression

To quantify the expression levels of atl-1 and egl-1 in a replicative stress context, synchronized N2 L1 worms were grown at 20°C for 24 h in HU plates and then recovered with M9 buffer for total RNA isolation.

Replicative stress sensitivity assay

Synchronized populations of wild type N2, *cer22*, *cer23*, *cer24*, and the control strain *met-2,set-25* were exposed to HU pulse 0, 2.5, 5, 10 and 20mM during 16h. Dilutions of HU were prepared in M9 with 400µl of concentrated OP50 for a 5ml final volume. Once dilutions were prepared, around 400 worms were seeded in each tube. They were left in agitation at 20°C for 16 hours. Afterwards, worms were washed 3

times with M9 and 120-150 worms were seeded in NGM plates with OP50 and let recover for 3 days at 20°C before scoring.

At the third day, plates were scored and the total number of recovered adults (gravid healthy adults), sick and dead worms was scored for every strain and every concentration. 20mM concentration of HU showed the strongest effect, for simplicity, only those results are being shown. Experiment was repeated 3 times each one with a different biological replicate.

MM.8 DNA damage induction through UV-C exposure

Synchronized worms were exposed to UV-C light (254 nm) at different time points depending on the experiment. A UV crosslinker (model 2400, Stratagene) was used to apply 100 J/m2 in all the cases. Before the exposure to UV-C, worms were washed several times to get rid of bacteria and placed in a bacteria-free NGM plate.

Quantification of gene expression

To quantify the expression levels of *atl-1* and *egl-1* upon DNA damage induced by UV-C exposure, synchronized L1 N2 worms were grown in standard conditions at 20°C. UV-C exposure was performed 18 h post L1 and total RNA isolation 6 h later.

Foci quantification

For RPA-1 foci observation synchronized P*rpa-1*::YFP L1 worms were grown at 20°C for 24 h. Then, they were exposed to UV-C and later placed again in control or *prp-8(RNAi)* plates for an additional 24 h. Worms were then mounted using 0.3 mM of levamisole and observed using a confocal fluorescence microscope. RPA-1::YFP foci formation was quantified by counting the number of bright foci present in germline and somatic cells.

MM.9. Developmental assay

To asses if our CRISPR mutations had any effect in the worms development, we study the developmental timing of each mutant strain and compared it with wild types N2 worms.

Worms were synchronized using the sodium hypochlorite treatment (Porta-de-la-Riva et al., 2012). Each strain was seeded in an OP50 plate and let recover for 1 h at 20°C and 25°C. After that, they were singled onto small plates containing OP50 bacteria and grown at the corresponding temperature.

The stage of the worms was determined every 24 h during a 3 d period. Stage determination was based on size and development of body structures like vulva and the size of the gonad.

M.10. Genome editing by CRISPR/Cas9 system

The CRISPR/Cas9 system is based in a commercial kit from Integrated DNA Technologies. The method detailed bellow follows recommendations from publications of CRISPR/Cas9 optimization techniques in *C. elegans* (Dickinson & Goldstein, 2016; Paix et al., 2015).

MM.10.1.CRISPR/Cas9 – crRNA - tracrRNA

Cas9 protein: From IDTDNA: Alt-R™ S.p. Cas9 Nuclease 3NLS (61 μM)

IDT Cas9 is a recombinant *S. pyogenes* Cas9 nuclease, purified from *E. coli* strain expressing codon optimized Cas9. To improve efficiency this Cas9 contains 1 N-terminal nuclear localization sequence (NLS), 2 C-terminal NLSs, and a C-terminal 6-His tag. 100 μg of Cas9 nuclease = 610 pmol.

tracrRNA: From IDTDNA: Alt-R™ CRISPR-Cas9 tracrRNA, 20 nmol

The universal 67mer tracrRNA that contains proprietary chemical modifications conferring increased nuclease resistance. Hybridizes to crRNA to activate the Cas9 enzyme.

The same tracrRNA is used for all the injection mixes. Upon receipt, briefly spin the tube and resuspend in 62,5μl IDTE nuclease-free buffer (320 μM stock solution). Store at -20°C.

crRNA

- Pick about 150-200 nucleotide of genomic sequence around the site that you want to cut, delete or edit.
- Paste the sequence in any of these online tools that search crRNAs and identified putative off-target. Our favorites are:

http://crispr.cos.uni-heidelberg.de

http://www.benchling.com/

http://crispr.mit.edu/

MM.10.2. Considerations for crRNA designing

- Sequence composition:
  - high GC content (50-75%) and high Tm: it should help crRNA to anneal faster (online tools should filter this)
  - If possible, chose (N18)**GG**NGG crRNA – higher efficiency according to Farboud and Meyer 2015.
  - Avoid C at position 20. (N19)**C**NGG crRNAs have been predicted to be less efficient.
- Efficiency of the guide RNA (on-target score): use predicted values provided by online tools, although they have limited value.
- Specificity of the guide RNA (off-target score): it is not a big problem in *C. elegans* if off-targets are in different chromosomes. To make sure that we do not have off-target mutations, we usually keep at least 2 independent lines and outcross them at least 2 times. You can always PCR and sequence a potential off-target that can influence in your phenotype.
- To introduce point mutations, the proximity of the cut (*cleavage site is 3 bp upstream of the PAM sequence*) to the desired modification site is very important. For insertions up to 1kb (short-range homology-directed repair) using an ssDNA or PCR product repair template, the distance between the cleavage site and the edit should be **<10 bases**. Insertions at higher distances can occur but often lead to partial recombination events.

MM.10.3. Repair template design

General designing rules

Repair templates should contain around 35bp homology arms (Paix et al. 2014), starting from the first and the last modification towards each site. Homology arms are unmodified sequences that flank the 5' and 3' ends of the edit, and are homologous to the genomic DNA.

To prevent Cas9 recutting of the edit after integration in the genome, introduce silent mutations in the PAM and/or the protospacer sequence. Mutation of the PAM is sufficient to completely block Cas9 cleavage; however, this is not always possible depending on the reading frame.

When silent mutations can't be introduced in the PAM, and also to facilitate PCR screening of the edits, add silent mutations in the protospacer sequence. 4 silent mutations should be enough, but we usually add as many mutations as possible to increase primer specificity. Consult a codon usage table to introduce synonym codons that are used at a frequency similar to the original codon.

If the cut site is >10 bp far from the codon you want to mutate, introduce silent mutations between the cut site and the edit to prevent partial recombination events.

	Initial concentration		Final concentration		
	µM	ng/µL	µM	ng/µL	Volume (µl)
Cas9 labomics	30	5000	4,5	750	1,5
ALT-R tracrRNA	320	7098,34	32	709,83	1
ALT-R *dpy-10* crRNA	100	1150,99	10	115,09	1
ALT-R *target gene* crRNA	100	x	35	x	3,5
ssODN *dpy-10(cn64)* repair template	16,3	500	1,22	37,5	0,75
ssODN *target gene* repair template	x	1000	x	175	1,75
Nuclease free water					0,5
Total volume					**10**

Table MM. 7. CRISPR/Cas9 mix preparation

Before injecting, spin for 10 min at 13000 rpm and incubate at 37°C for 15 minutes.

Keep the injection mix on ice. Store at -20°C. Injection mixes can be reused.

MM. Recipes for general methods

MM.1. Recipes

Nematode Growth Media agar (NGM)

For 1 liter of plates:

NaCl	3 g
Peptone	2,5 g
Agar	17 g
H_2O	975 ml

Autoclave, cool to 55°C and then add the following reagents mixing after every addition:

Cholesterol (5mg/ml in EtOH)	1 ml
1M $CaCl_2$	1 ml
1M $MgSO_4$	1 ml
1M Kalium phosphate buffer	25 ml

1M Kalium phosphate buffer

For 1 liter:

KH_2PO_4	108,3 g
K_2HPO_4	35,6 g
H_2O	975 ml

Luria Bertani (LB)

For 1 liter:

Tryptone	10 g
Yeast extract	5 g
NaCl	10 g
H_2O	950 ml

MM.2. Recipes

1x Lysis buffer

H_2O	135µl
10x NH4 Reaction Buffer*	15µl
Proteinase K (10mg/ml)	1µl

* BIOTAQ™ PCR Kit

REFERENCES

REFERENCES

Aballay, A., & Ausubel, F. M. (2001). Programmed cell death mediated by ced-3 and ced-4 protects Caenorhabditis elegans from Salmonella typhimurium-mediated killing. *Proceedings of the National Academy of Sciences*, *98*(5), 2735–2739. https://doi.org/10.1073/pnas.041613098

Aguilera, A., & García-Muse, T. (2012). R Loops: From Transcription Byproducts to Threats to Genome Stability. *Molecular Cell*, *46*(2), 115–124. https://doi.org/10.1016/j.molcel.2012.04.009

Aharony, I., Michowiz, S., & Goldenberg-Cohen, N. (2017). The promise of stem cell-based therapeutics in ophthalmology. *Neural Regeneration Research*, *12*(2), 173–180. https://doi.org/10.4103/1673-5374.200793

Ameur, A., Zaghlool, A., Halvardson, J., Wetterbom, A., Gyllensten, U., Cavelier, L., & Feuk, L. (2011). Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. *Nature Structural & Molecular Biology*, *18*(12), 1435–1440. https://doi.org/10.1038/nsmb.2143

AP. Page, I. J. (2007). The cuticle. *Wormbook*.

Barberan-Soler, S., Lambert, N. J., & Zahler, A. M. (2009). Global analysis of alternative splicing uncovers developmental regulation of nonsense-mediated decay in C. elegans. *RNA*, *15*(9), 1652–1660. https://doi.org/10.1261/rna.1711109

Basu, U., Meng, F.-L., Keim, C., Grinstein, V., Pefanis, E., Eccleston, J., … Alt, F. W. (2011). The RNA Exosome Targets the AID Cytidine Deaminase to Both Strands of Transcribed Duplex DNA Substrates. *Cell*, *144*(3), 353–363. https://doi.org/10.1016/j.cell.2011.01.001

Benaglio, P., Mcgee, T. L., Capelli, L. P., Harper, S., Berson, E. L., & Rivolta, C. (2011). Next generation sequencing of pooled samples

reveals new SNRNP200 mutations associated with retinitis pigmentosa. *Human Mutation*, *32*(6). https://doi.org/10.1002/humu.21485

Berget, S. M., Moore, C., & Sharp, P. A. (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proceedings of the National Academy of Sciences of the United States of America*, *74*(8), 3171–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/269380

Binder, G., Brown, M., & Parks, J. S. (1996). Mechanisms responsible for dominant expression of human growth hormone gene mutations. *The Journal of Clinical Endocrinology & Metabolism*, *81*(11), 4047–4050. https://doi.org/10.1210/jcem.81.11.8923859

Black, D. L. (2003). Mechanisms of Alternative Pre-Messenger RNA Splicing. *Annual Review of Biochemistry*, *72*(1), 291–336. https://doi.org/10.1146/annurev.biochem.72.121801.161720

Blaxter, M. (2011). Nematodes: the worm and its relatives. *PLoS Biology*, *9*(4), e1001050. https://doi.org/10.1371/journal.pbio.1001050

Boon, K.-L., Grainger, R. J., Ehsani, P., Barrass, J. D., Auchynnikava, T., Inglehearn, C. F., & Beggs, J. D. (2007). prp8 mutations that cause human retinitis pigmentosa lead to a U5 snRNP maturation defect in yeast. *Nature Structural & Molecular Biology*, *14*(11), 1077–1083. https://doi.org/10.1038/nsmb1303

Brambati, A., Colosio, A., Zardoni, L., Galanti, L., & Liberi, G. (2015). Replication and transcription on a collision course: eukaryotic regulation mechanisms and implications for DNA stability. *Frontiers in Genetics*, *6*, 166. https://doi.org/10.3389/fgene.2015.00166

Brenner, S. (1974). The genetics of Caenorhabditis elegans. *Genetics*, *77*(1), 71–94. https://doi.org/10.1002/cbic.200300625

Briese, M., Esmaeili, B., & Sattelle, D. B. (2005). Is spinal muscular atrophy the result of defects in motor neuron processes? *BioEssays*, *27*(9), 946–957. https://doi.org/10.1002/bies.20283

Buchholz, D. E., Hikita, S. T., Rowland, T. J., Friedrich, A. M., Hinman, C. R., Johnson, L. V., & Clegg, D. O. (2009). Derivation of Functional Retinal Pigmented Epithelium from Induced Pluripotent Stem Cells. *Stem Cells*, *27*(10), 2427–2434. https://doi.org/10.1002/stem.189

C. elegans Sequencing Consortium. (1998). Genome sequence of the nematode C. elegans: a platform for investigating biology. *Science (New York, N.Y.)*, *282*(5396), 2012–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9851916

Canugovi, C., Samaranayake, M., & Bhagwat, A. S. (2009). Transcriptional pausing and stalling causes multiple clustered mutations by human activation-induced deaminase. *FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology*, *23*(1), 34–44. https://doi.org/10.1096/fj.08-115352

Castellano-Pozo, M., García-Muse, T., & Aguilera, A. (2012). R-loops cause replication impairment and genome instability during meiosis. *EMBO Reports*, *13*(10), 923–929. https://doi.org/10.1038/embor.2012.119

Castellano-Pozo, M., Santos-Pereira, J., Rond??n, A. G., Barroso, S., And??jar, E., P??rez-Alegre, M., … Aguilera, A. (2013). R loops are linked to histone H3 S10 phosphorylation and chromatin condensation. *Molecular Cell*, *52*(4), 583–590. https://doi.org/10.1016/j.molcel.2013.10.006

Chakarova, C. F., Hims, M. M., Bolz, H., Abu-Safieh, L., Patel, R. J., Papaioannou, M. G., … Bhattacharya, S. S. (2002). Mutations in HPRP3, a third member of pre-mRNA splicing factor genes, implicated in autosomal dominant retinitis pigmentosa. *Human Molecular Genetics*, *11*(1), 87–92. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11773002

Chalfie, M. (1994). Green fluorescent protein as a marker for gene expression. *Trends in Genetics*, *10*(5), 151. https://doi.org/10.1016/0168-9525(94)90088-4

Chan, Y. A., Aristizabal, M. J., Lu, P. Y. T., Luo, Z., Hamza, A., Kobor, M. S., … Hieter, P. (2014). Genome-Wide Profiling of Yeast DNA:RNA Hybrid Prone Sites with DRIP-Chip. *PLoS Genetics*, *10*(4), e1004288. https://doi.org/10.1371/journal.pgen.1004288

Chang, Y.-F., Imam, J. S., & Wilkinson, M. F. (2007). The Nonsense-Mediated Decay RNA Surveillance Pathway. *Annual Review of Biochemistry*, *76*(1), 51–74. https://doi.org/10.1146/annurev.biochem.76.050106.093909

Cogan, J. D., Prince, M. A., Lekhakula, S., Bundey, S., Futrakul, A., McCarthy, E. M., & Phillips, J. A. (1997). A novel mechanism of aberrant pre-mRNA splicing in humans. *Human Molecular Genetics*, *6*(6), 909–12. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9175738

Conradt, B., & Horvitz, H. R. (1998). The C. elegans Protein EGL-1 is required for programmed cell death and interacts with the Bcl-2-like protein CED-9. *Cell*, *93*(4), 519–529. https://doi.org/10.1016/S0092-8674(00)81182-4

Conradt, B., Wu, Y.-C., & Xue, D. (2016). Programmed Cell Death During Caenorhabditis elegans Development. *Genetics*, *203*(4), 1533–1562. https://doi.org/10.1534/genetics.115.186247

Corrionero, A., Minana, B., & Valcarcel, J. (2011). Reduced fidelity of branch point recognition and alternative splicing induced by the anti-tumor drug spliceostatin A. *Genes & Development*, *25*(5), 445–459. https://doi.org/10.1101/gad.2014311

Corsi, A. K. (2006). A biochemist's guide to Caenorhabditis elegans. *Analytical Biochemistry*, *359*(1), 1–17. https://doi.org/10.1016/j.ab.2006.07.033

Corvelo, A., Hallegger, M., Smith, C. W. J., & Eyras, E. (2010). Genome-wide association between branch point properties and alternative splicing. *PLoS Computational Biology*, *6*(11), 12–15. https://doi.org/10.1371/journal.pcbi.1001016

Cox, G. N., & Hirsh, D. (1985). Stage-specific patterns of collagen gene expression during development of Caenorhabditis elegans. *Molecular and Cellular Biology*, *5*(2), 363–72. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2983191

Cox, G. N., Kusch, M., DeNevi, K., & Edgar, R. S. (1981). Temporal regulation of cuticle synthesis during development of Caenorhabditis elegans. *Developmental Biology*, *84*(2), 277–85. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20737865

Culetto, E., & Sattelle, D. B. (2000). A role for Caenorhabditis elegans in understanding the function and interactions of human disease genes. *Human Molecular Genetics*, *9*(6), 869–77. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10767309

Daiger, S. P., Sullivan, L. S., & Bowne, S. J. (2013). Genes and mutations causing retinitis pigmentosa. *Clinical Genetics*, *84*(2), 132–41. https://doi.org/10.1111/cge.12203

Dang, V. T., Kassahn, K. S., Marcos, A. E., & Ragan, M. A. (2008). Identification of human haploinsufficient genes and their genomic proximity to segmental duplications. *European Journal of Human Genetics*, *16111*(10), 1350–1357. https://doi.org/10.1038/ejhg.2008.111

Das, R., Dufu, K., Romney, B., Feldt, M., Elenko, M., & Reed, R. (2006). Functional coupling of RNAP II transcription to spliceosome assembly. *Genes & Development*, *20*(9), 1100–1109. https://doi.org/10.1101/gad.1397406

David, C. J., Boyne, A. R., Millhouse, S. R., & Manley, J. L. (2011). The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65-Prp19 complex. *Genes & Development*, *25*(9), 972–983. https://doi.org/10.1101/gad.2038011

De Sousa Dias, M., Hernan, I., Pascual, B., Borràs, E., Mañé, B., Gamundi, M. J., & Carballo, M. (2013). Detection of novel mutations that cause autosomal dominant retinitis pigmentosa in candidate genes by long-range PCR amplification and next-generation sequencing. *Molecular Vision*, *19*, 654–64. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/23559859

Deaton, A. M., & Bird, A. (2011). CpG islands and the regulation of transcription. *Genes & Development*, *25*(10), 1010–22. https://doi.org/10.1101/gad.2037511

Dickinson, D. J., & Goldstein, B. (2016). CRISPR-Based Methods for Caenorhabditis elegans Genome Engineering. *Genetics*, *202*(3), 885–901. https://doi.org/10.1534/genetics.115.182162

Drolet, M., Bi, X., & Liu, L. F. (1994). Hypernegative supercoiling of the DNA template during transcription elongation in vitro. *The Journal of Biological Chemistry*, *269*(3), 2068–74. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8294458

Dvinge, H., Kim, E., Abdel-Wahab, O., & Bradley, R. K. (2016). RNA splicing factors as oncoproteins and tumour suppressors. *Nature Reviews Cancer*, *16*(7), 413–430. https://doi.org/10.1038/nrc.2016.51

El Hage, A., Webb, S., Kerr, A., Tollervey, D., & Andujar, E. (2014). Genome-Wide Distribution of RNA-DNA Hybrids Identifies RNase H Targets in tRNA Genes, Retrotransposons and Mitochondria. *PLoS Genetics*, *10*(10), e1004716. https://doi.org/10.1371/journal.pgen.1004716

Ellis, H. M., & Horvitz, H. R. (1986). Genetic control of programmed cell death in the nematode C. elegans. *Cell*, *44*(6), 817–29. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3955651

Fabrizio, P., Dannenberg, J., Dube, P., Kastner, B., Stark, H., Urlaub, H., & Lührmann, R. (2009). The Evolutionarily Conserved Core Design of the Catalytic Activation Step of the Yeast Spliceosome. *Molecular Cell*, *36*(4), 593–608. https://doi.org/10.1016/j.molcel.2009.09.040

Fackenthal, J. D., Cartegni, L., Krainer, A. R., & Olopade, O. I. (2002). BRCA2 T2722R Is a Deleterious Allele That Causes Exon Skipping. *The American Journal of Human Genetics*, *71*(3), 625–631. https://doi.org/10.1086/342192

Fairchild, P. J. (2010). The challenge of immunogenicity in the quest for induced pluripotency. *Nature Reviews Immunology*, *10*(12), 868–875. https://doi.org/10.1038/nri2878

Fatehullah, A., Tan, S. H., & Barker, N. (2016). Organoids as an in vitro model of human development and disease. *Nature Cell Biology*, *18*(3), 246–254. https://doi.org/10.1038/ncb3312

Faustino, N. A., Cooper, T. a, & Andre, N. (2003). Pre-mRNA splicing and human disease. *Genes & Development*, *17*(4), 419–437. https://doi.org/10.1101/gad.1048803.snRNP

Ferrari, S., Di Iorio, E., Barbaro, V., Ponzin, D., Sorrentino, F. S., & Parmeggiani, F. (2011). Retinitis pigmentosa: genes and disease mechanisms. *Current Genomics*, *12*(4), 238–49. https://doi.org/10.2174/138920211795860107

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., & Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, *391*(6669), 806–811. https://doi.org/10.1038/35888

Firestein, B. L., & Rongo, C. (2001). DLG-1 is a MAGUK similar to SAP97 and is required for adherens junction formation. *Molecular Biology of the Cell*, *12*(November), 3465–3475. Retrieved from file:///Unknown/DLG-1 Is a MAGUK Similar to SAP97 and Is Required for Adherens Junction Formation - 0.pdf

Fong, Y. W., & Zhou, Q. (2001). Stimulatory effect of splicing factors on transcriptional elongation. *Nature*, *414*(6866), 929–933. https://doi.org/10.1038/414929a

Fontrodona, L., Porta-De-La-Riva, M., Morá N, T. S., Niu, W., Nica Díaz, M., Aristizá Bal-Corrales, D., … Ceró N, J. N. (2013). RSR-2, the Caenorhabditis elegans Ortholog of Human Spliceosomal Component SRm300/SRRM2, Regulates Development by

Influencing the Transcriptional Machinery. https://doi.org/10.1371/journal.pgen.1003543

Frand, A. R., Russel, S., & Ruvkun, G. (2005). Functional genomic analysis of C. elegans molting. *PLoS Biology*, *3*(10), e312. https://doi.org/10.1371/journal.pbio.0030312

Frézal, L., Félix, M.-A., Baird, S., Félix, M.-A., Haag, E., Adair, R., … Waterston, R. (2015). *C. elegans* outside the Petri dish. *eLife*, *4*, 997–1012. https://doi.org/10.7554/eLife.05849

Friedland, A. E., Tzur, Y. B., Esvelt, K. M., Colaiácovo, M. P., Church, G. M., & Calarco, J. A. (2013). Heritable genome editing in C. elegans via a CRISPR-Cas9 system. *Nature Methods*, *10*(8), 741–743. https://doi.org/10.1038/nmeth.2532

Frøkjær-Jensen, C. (2013). Exciting prospects for precise engineering of Caenorhabditis elegans genomes with CRISPR/Cas9. *Genetics*, *195*(3), 635–642. https://doi.org/10.1534/genetics.113.156521

Frøkjær-Jensen, C., Wayne Davis, M., Hopkins, C. E., Newman, B. J., Thummel, J. M., Olesen, S.-P., … Jorgensen, E. M. (2008). Single-copy insertion of transgenes in Caenorhabditis elegans. *Nature Genetics*, *40*(11), 1375–1383. https://doi.org/10.1038/ng.248

Galej, W. P., Hoang, T., Nguyen, D., Newman, A. J., & Nagai, K. (2014). Structural studies of the spliceosome: zooming into the heart of the machine. *Current Opinion in Structural Biology*, *25*, 57–66. https://doi.org/10.1016/j.sbi.2013.12.002

Gamundi, M. J., Hernan, I., Muntanyola, M., Maseras, M., López-Romero, P., Álvarez, R., … Carballo, M. (2008). Transcriptional expression of cis-acting and trans-acting splicing mutations cause autosomal dominant retinitis pigmentosa. *Human Mutation*, *29*(6), 869–878. https://doi.org/10.1002/humu.20747

Gan, W., Guan, Z., Liu, J., Gui, T., Shen, K., Manley, J. L., & Li, X. (2011). R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes & Development*, *25*(19), 2041–56. https://doi.org/10.1101/gad.17010011

Garcia-Muse, T., & Boulton, S. J. (2005). Distinct modes of ATR activation after replication stress and DNA double-strand breaks in Caenorhabditis elegans. *The EMBO Journal*, *24*(24), 4345–55. https://doi.org/10.1038/sj.emboj.7600896

Gartner, A., Boag, P. R., & Blackwell, T. K. (2008). Germline survival and apoptosis. *WormBook : The Online Review of C. Elegans Biology*, 1–20. https://doi.org/10.1895/wormbook.1.145.1

Gartner, A., Milstein, S., Ahmed, S., Hodgkin, J., & Hengartner, M. O. (2000). A Conserved Checkpoint Pathway Mediates DNA Damage–Induced Apoptosis and Cell Cycle Arrest in C. elegans. *Molecular Cell*, *5*, 435–443.

Gasiunas, G., Barrangou, R., Horvath, P., & Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(39), E2579–86. https://doi.org/10.1073/pnas.1208507109

Ghosh, S., & Garcia-Blanco, M. A. (2000). Coupled in vitro synthesis and splicing of RNA polymerase II transcripts. *RNA (New York, N.Y.)*, *6*(9), 1325–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10999609

Ginno, P. A., Lott, P. L., Christensen, H. C., Korf, I., & Chédin, F. (2012). R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters. *Molecular Cell*, *45*(6), 814–825. https://doi.org/10.1016/j.molcel.2012.01.017

Grainger, R. J. (2005). Prp8 protein: At the heart of the spliceosome. *RNA*, *11*(5), 533–557. https://doi.org/10.1261/rna.2220705

Grishok, A. (2005). RNAi mechanisms in Caenorhabditis elegans. *FEBS Letters*, *579*(26), 5932–5939. https://doi.org/10.1016/j.febslet.2005.08.001

Gumienny, T. L., Lambie, E., Hartwieg, E., Horvitz, H. R., & Hengartner, M. O. (1999). Genetic control of programmed cell death in the Caenorhabditis elegans hermaphrodite germline.

Development (Cambridge, England), *126*(5), 1011–22. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9927601

Hamel, C. (2006). Retinitis pigmentosa. *Orphanet Journal of Rare Diseases*, *1*, 40. https://doi.org/10.1186/1750-1172-1-40

Hamperl, S., & Cimprich, K. A. (2014). The contribution of co-transcriptional RNA:DNA hybrid structures to DNA damage and genome instability. *DNA Repair*, *19*, 84–94. https://doi.org/10.1016/j.dnarep.2014.03.023

Hamperl, S., Cimprich, K. A., Taschner, M., Reid, J., Walker, J., Erdjument-Bromage, H., … al., et. (2016). Conflict Resolution in the Genome: How Transcription and Replication Make It Work. *Cell*, *167*(6), 1455–1467. https://doi.org/10.1016/j.cell.2016.09.053

Hartong, D. T., Berson, E. L., & Dryja, T. P. (2006). Retinitis pigmentosa. *Lancet*, *368*(9549), 1795–1809. https://doi.org/10.1016/S0140-6736(06)69740-7

Hedgecock, E. M., & White, J. G. (1985). Polyploid tissues in the nematode Caenorhabditis elegans. *Developmental Biology*, *107*(1), 128–33. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2578115

Hendriks, G., Jansen, J. G., Mullenders, L. H. F., & de Wind, N. (2010). Transcription-coupled repair and apoptosis provide specific protection against transcription-associated mutagenesis by ultraviolet light. *Transcription*, *1*(2), 95–8. https://doi.org/10.4161/trns.1.2.12788

Hengartner, M. O., Ellis, R., & Horvitz, R. (1992). Caenorhabditis elegans gene ced-9 protects cells from programmed cell death. *Nature*, *356*(6369), 494–499. https://doi.org/10.1038/356494a0

Hengartner, M. O., & Horvitz, H. R. (1994). C. elegans cell survival gene ced-9 encodes a functional homolog of the mammalian proto-oncogene bcl-2. *Cell*, *76*(4), 665–676. https://doi.org/10.1016/0092-8674(94)90506-1

Hirose, Y., Tacke, R., & Manley, J. L. (1999). Phosphorylated RNA polymerase II stimulates pre-mRNA splicing. *Genes & Development*, *13*(10), 1234–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10346811

Hodgkin, J., Papp, A., Pulak, R., Ambros, V., & Anderson, P. (1989). A new kind of informational suppression in the nematode Caenorhabditis elegans. *Genetics*, *123*(2), 301–13. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2583479

Hofmann, E. R., Milstein, S., Boulton, S. J., Ye, J., Hofmann, J. J., Stergiou, L., … Hengartner, M. O. (2002). Caenorhabditis elegans HUS-1 is a DNA damage checkpoint protein required for genome stability and induction of EGL-1. *European Worm Meeting*, *12*(02), 1908–1918. https://doi.org/10.1016/S0960-9822(02)01262-9

Horvitz, H. R. (2003). Worms, Life, and Death (Nobel Lecture). *ChemBioChem*, *4*(8), 697–711. https://doi.org/10.1002/cbic.200300614

Hou, P., Li, Y., Zhang, X., Liu, C., Guan, J., Li, H., … Deng, H. (2013). Pluripotent Stem Cells Induced from Mouse Somatic Cells by Small-Molecule Compounds. *Science*, *341*(6146), 651–654. https://doi.org/10.1126/science.1239278

Huertas, P., & Aguilera, A. (2003). Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination. *Molecular Cell*, *12*(3), 711–21. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14527416

Hunt-Newbury, R., Viveiros, R., Johnsen, R., Mah, A., Anastas, D., Fang, L., … Moerman, D. G. (2007). High-Throughput In Vivo Analysis of Gene Expression in Caenorhabditis elegans. *PLoS Biology*, *5*(9), e237. https://doi.org/10.1371/journal.pbio.0050237

Ivings, L., Towns, K. V, Matin, M. A., Taylor, C., Ponchel, F., Grainger, R. J., … Inglehearn, C. F. (n.d.). Evaluation of splicing efficiency in lymphoblastoid cell lines from patients with splicing-factor retinitis pigmentosa. Retrieved from

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2603472/pdf/mv-v14-2357.pdf

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, *337*(6096). Retrieved from http://science.sciencemag.org/content/337/6096/816.long

Johnson, S. J., & Jackson, R. N. (2013). Ski2-like RNA helicase structures: common themes and complex assemblies. *RNA Biology*, *10*(1), 33–43. https://doi.org/10.4161/rna.22101

Johnstone, I. L. (1994). The cuticle of the nematode Caenorhabditis elegans: A complex collagen structure. *BioEssays*, *16*(3), 171–178. https://doi.org/10.1002/bies.950160307

Kadandale, P., Chatterjee, I., & Singson, A. (2009). Germline transformation of Caenorhabditis elegans by injection. *Methods in Molecular Biology (Clifton, N.J.)*, *518*, 123–33. https://doi.org/10.1007/978-1-59745-202-1_10

Kaletta, T., & Hengartner, M. O. (2006). Finding function in novel targets: C. elegans as a model organism. *Nature Reviews. Drug Discovery*, *5*(5), 387–98. https://doi.org/10.1038/nrd2031

Kamath, R. S., & Ahringer, J. (2003). Genome-wide RNAi screening in Caenorhabditis elegans. *Methods (San Diego, Calif.)*, *30*(4), 313–21. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12828945

Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., … Ahringer, J. (2003). Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. *Nature*, *421*(6920), 231–237. https://doi.org/10.1038/nature01278

Kastan, M. B., & Bartek, J. (2004). Cell-cycle checkpoints and cancer. *Nature*, *432*(7015), 316–23. https://doi.org/10.1038/nature03097

Kim, N., Abdulovic, A. L., Gealy, R., Lippert, M. J., & Jinks-Robertson, S. (2007). Transcription-associated mutagenesis in

yeast is directly proportional to the level of gene expression and influenced by the direction of DNA replication. *DNA Repair*, *6*(9), 1285–96. https://doi.org/10.1016/j.dnarep.2007.02.023

Kim, S. K. (2001). Http://C. elegans: mining the functional genomic landscape. *Nature Reviews. Genetics*, *2*(9), 681–9. https://doi.org/10.1038/35088523

Korir, P. K., Roberts, L., Ramesar, R., & Seoighe, C. (2014). A mutation in a splicing factor that causes retinitis pigmentosa has a transcriptome-wide effect on mRNA splicing. *BMC Research Notes*, *7*(1), 401. https://doi.org/10.1186/1756-0500-7-401

Krawczak, M., Reiss, J., & Cooper, D. N. (1992). The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: Causes and consequences. *Human Genetics*, *90*(1-2), 41–54. https://doi.org/10.1007/BF00210743

Kurtovic-Kozaric, a, Przychodzen, B., Singh, J., Konarska, M. M., Clemente, M. J., Otrock, Z. K., … Padgett, R. a. (2014). PRPF8 defects cause missplicing in myeloid malignancies. *Leukemia*, *29*(April), 1–11. https://doi.org/10.1038/leu.2014.144

Labrador-Velandia, S., Alonso-Alonso, M. L., Alvarez-Sanchez, S., González-Zamora, J., Carretero-Barrio, I., Pastor, J. C., … Srivastava, G. K. (2016). Mesenchymal stem cell therapy in retinal and optic nerve diseases: An update of clinical trials. *World Journal of Stem Cells*, *8*(11), 376–383. https://doi.org/10.4252/wjsc.v8.i11.376

Le Bras, S., & Le Borgne, R. (2014). Epithelial cell division – multiplying without losing touch. *Journal of Cell Science*, *127*(24). Retrieved from http://jcs.biologists.org/content/127/24/5127

Lettre, G., & Hengartner, M. O. (2006). Developmental apoptosis in C. elegans: a complex CEDnario. *Nature Reviews. Molecular Cell Biology*, *7*(2), 97–108. https://doi.org/10.1038/nrm1836

Levitan, D., & Greenwald, I. (1995). Facilitation of lin-12-mediated signalling by sel-12, a Caenorhabditis elegans S182 Alzheimer's

disease gene. *Nature*, *377*(6547), 351–354. https://doi.org/10.1038/377351a0

Li, M., & Belmonte, J. C. I. (2016). Looking to the future following 10 years of induced pluripotent stem cell technologies. *Nature Protocols*, *11*(9), 1579–1585. https://doi.org/10.1038/nprot.2016.108

Li, X., Niu, T., & Manley, J. L. (2007). The RNA binding protein RNPS1 alleviates ASF/SF2 depletion-induced genomic instability. *RNA (New York, N.Y.)*, *13*(12), 2108–15. https://doi.org/10.1261/rna.734407

Lim, L. P., & Burge, C. B. (2001). A computational analysis of sequence features involved in recognition of short introns. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(20), 11193–8. https://doi.org/10.1073/pnas.201407298

Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S., & Fu, X.-D. (2008). The splicing factor SC35 has an active role in transcriptional elongation. *Nature Structural & Molecular Biology*, *15*(8), 819–826. https://doi.org/10.1038/nsmb.1461

Liu, H. X., Cartegni, L., Zhang, M. Q., & Krainer, A. R. (2001). A mechanism for exon skipping caused by nonsense or missense mutations in BRCA1 and other genes. *Nature Genetics*, *27*(1), 55–58. https://doi.org/10.1038/83762

Liu, H. X., Zhang, M., & Krainer, A. R. (1998). Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. *Genes & Development*, *12*(13), 1998–2012. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9649504

Liu, S., Mozaffari-Jovin, S., Wollenhaupt, J., Santos, K. F., Theuser, M., Dunin-Horkawicz, S., … Wahl, M. C. (2015). A composite double-/single-stranded RNA-binding region in protein Prp3 supports tri-snRNP stability and splicing. *eLife*, *4*, e07320. https://doi.org/10.7554/eLife.07320

Liu, S., Rauhut, R., Vornlocher, H.-P., & Lührmann, R. (2006). The network of protein-protein interactions within the human U4/U6.U5 tri-snRNP. *RNA (New York, N.Y.)*, *12*(7), 1418–30. https://doi.org/10.1261/rna.55406

Liu, T., Jin, X., Zhang, X., Yuan, H., Cheng, J., Lee, J., … Wang, W. (2012). A Novel Missense SNRNP200 Mutation Associated with Autosomal Dominant Retinitis Pigmentosa in a Chinese Family. *PLoS ONE*, *7*(9), e45464. https://doi.org/10.1371/journal.pone.0045464

Longman, D., Johnstone, I. L., & Cáceres, J. F. (2000). Functional characterization of SR and SR-related genes in Caenorhabditis elegans. *The EMBO Journal*, *19*(7), 1625–37. https://doi.org/10.1093/emboj/19.7.1625

López-Bigas, N., Audit, B., Ouzounis, C., Parra, G., & Guigó, R. (2005). Are splicing mutations the most frequent cause of hereditary disease? *FEBS Letters*, *579*(9), 1900–1903. https://doi.org/10.1016/j.febslet.2005.02.047

Losson, R., & Lacroute, F. (1979). Interference of nonsense mutations with eukaryotic messenger RNA stability. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(10), 5134–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/388431

Maeder, C., Kutach, A. K., & Guthrie, C. (2009). ATP-dependent unwinding of U4/U6 snRNAs by the Brr2 helicase requires the C terminus of Prp8. *Nature Structural & Molecular Biology*, *16*(1), 42–48. https://doi.org/10.1038/nsmb.1535

Makarova, O. V., Makarov, E. M., Liu, S., Vornlocher, H., & Lührmann, R. (2002). Protein 61K, encoded by a gene (PRPF31) linked to autosomal dominant retinitis pigmentosa, is required for U4/U6middle dotU5 tri-snRNP formation and pre-mRNA splicing. *The EMBO Journal*, *21*(5), 1148–1157. https://doi.org/10.1093/emboj/21.5.1148

Martínez-Gimeno, M., Gamundi, M. J., Hernan, I., Maseras, M., Millá, E., Ayuso, C., … Carballo, M. (2003). Mutations in the Pre-mRNA Splicing-Factor Genes PRPF3 , PRPF8 , and PRPF31 in

Spanish Families with Autosomal Dominant Retinitis Pigmentosa. *Investigative Opthalmology & Visual Science*, *44*(5), 2171. https://doi.org/10.1167/iovs.02-0871

McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., … Bentley, D. L. (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature*, *385*(6614), 357–361. https://doi.org/10.1038/385357a0

McKie, A. B., McHale, J. C., Keen, T. J., Tarttelin, E. E., Goliath, R., van Lith-Verhoeven, J. J., … Inglehearn, C. F. (2001). Mutations in the pre-mRNA splicing factor gene PRPC8 in autosomal dominant retinitis pigmentosa (RP13). *Human Molecular Genetics*, *10*(15), 1555–1562. https://doi.org/10.1093/hmg/10.15.1555

Merkhofer, E. C., Hu, P., & Johnson, T. L. (2014). Introduction to cotranscriptional RNA splicing. *Methods in Molecular Biology (Clifton, N.J.)*, *1126*, 83–96. https://doi.org/10.1007/978-1-62703-980-2_6

Metzstein, M. M., Stanfield, G. M., & Horvitz, H. R. (1998). Genetics of programmed cell death in C. elegans: past, present and future. *Trends in Genetics : TIG*, *14*(10), 410–416. https://doi.org/10.1016/S0168-9525(98)01573-X

Mischo, H. E., Gómez-González, B., Grzechnik, P., Rondón, A. G., Wei, W., Steinmetz, L., … Proudfoot, N. J. (2011). Yeast Sen1 Helicase Protects the Genome from Transcription-Associated Instability. *Molecular Cell*, *41*(1), 21–32. https://doi.org/10.1016/j.molcel.2010.12.007

Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., & Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, *155*(3), 733–740. https://doi.org/10.1099/mic.0.023960-0

Montecucco, A., & Biamonti, G. (2013). Pre-mRNA processing factors meet the DNA damage response. *Frontiers in Genetics*, *4*. https://doi.org/10.3389/fgene.2013.00102

Mordes, D., Luo, X., Kar, A., Kuo, D., Xu, L., Fushimi, K., … Wu, J. Y. (2006). Pre-mRNA splicing and retinitis pigmentosa. *Molecular Vision*, *12*(312), 1259–71. https://doi.org/10.1055/s-0029-1237430.Imprinting

Morthorst, T. H., & Olsen, A. (2013). Cell-nonautonomous inhibition of radiation-induced apoptosis by dynein light chain 1 in Caenorhabditis elegans. *Cell Death & Disease*, *4*(9), e799. https://doi.org/10.1038/cddis.2013.319

Mortillaro, M. J., Blencowe, B. J., Wei, X., Nakayasu, H., Du, L., Warren, S. L., … Berezney, R. (1996). A hyperphosphorylated form of the large subunit of RNA polymerase II is associated with splicing complexes and the nuclear matrix. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(16), 8253–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8710856

Moseley, C. T., Mullis, P. E., Prince, M. A., & Phillips, J. A. (2002). An Exon Splice Enhancer Mutation Causes Autosomal Dominant GH Deficiency. *The Journal of Clinical Endocrinology & Metabolism*, *87*(2), 847–852. https://doi.org/10.1210/jcem.87.2.8236

Mozaffari-Jovin, S., Wandersleben, T., Santos, K. F., Will, C. L., Lührmann, R., & Wahl, M. C. (2013). Inhibition of RNA helicase Brr2 by the C-terminal tail of the spliceosomal protein Prp8. *Science (New York, N.Y.)*, *341*(6141), 80–4. https://doi.org/10.1126/science.1237515

Mozaffari-Jovin, S., Wandersleben, T., Santos, K. F., Will, C. L., Lührmann, R., & Wahl, M. C. (2014). Novel regulatory principles of the spliceosomal Brr2 RNA helicase and links to retinal disease in humans. *RNA Biology*, *11*(4), 298–312. https://doi.org/10.4161/rna.28353

Nehme, R., & Conradt, B. (2009). egl-1: a key activator of apoptotic cell death in C. elegans. *Oncogene*, *27*, 30–40. https://doi.org/10.1038/onc.2009.41

Nilsen, T. W., & Graveley, B. R. (2010). Expansion of the eukaryotic proteome by alternative splicing. *Nature*, *463*(7280), 457–63. https://doi.org/10.1038/nature08909

Nissim-Rafinia, M., & Kerem, B. (2002). Splicing regulation as a potential genetic modifier. *Trends in Genetics*, *18*(3), 123–127. https://doi.org/10.1016/S0168-9525(01)02619-1

Nottrott, S., Urlaub, H., & Lührmann, R. (2002). Hierarchical, clustered protein interactions with U4/U6 snRNA: a biochemical role for U4/U6 proteins. *The EMBO Journal*, *21*(20), 5527–38. https://doi.org/10.1093/emboj/cdf544

Ogg, S., Paradis, S., Gottlieb, S., Patterson, G. I., Lee, L., Tissenbaum, H. A., & Ruvkun, G. (1997). The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in C. elegans. *Nature*, *389*(6654), 994–999. https://doi.org/10.1038/40194

Paix, A., Folkmann, A., Rasoloson, D., & Seydoux, G. (2015). High Efficiency, Homology-Directed Genome Editing in Caenorhabditis elegans Using CRISPR-Cas9 Ribonucleoprotein Complexes. *Genetics*, *201*(1), 47–54. https://doi.org/10.1534/genetics.115.179382

Papasaikas, P., Tejedor, J. R., Vigevani, L., & Valcárcel, J. (2015). Functional splicing network reveals extensive regulatory potential of the core spliceosomal machinery. *Molecular Cell*, *57*(1). https://doi.org/10.1016/j.molcel.2014.10.030

Pena, V., Liu, S., Bujnicki, J. M., Lührmann, R., & Wahl, M. C. (2007). Structure of a Multipartite Protein-Protein Interaction Domain in Splicing Factor Prp8 and Its Link to Retinitis Pigmentosa. *Molecular Cell*, *25*(4), 615–624. https://doi.org/10.1016/j.molcel.2007.01.023

Philips, A. V, & Cooper, T. A. (2000). Review RNA processing and human disease. *CMLS, Cell. Mol. Life Sci*, *57*, 235–249.

Podbilewicz, B., & White, J. G. (1994). Cell fusions in the developing epithelial of C. elegans. *Developmental Biology*, *161*(2), 408–424. https://doi.org/10.1006/dbio.1994.1041

Porta-de-la-Riva, M., Fontrodona, L., Villanueva, A., & Cerón, J. (2012). Basic Caenorhabditis elegans Methods: Synchronization and Observation. *Journal of Visualized Experiments*, (64), e4019–e4019. https://doi.org/10.3791/4019

Poulos, M. G., Batra, R., Charizanis, K., & Swanson, M. S. (2011). Developments in RNA splicing and disease. *Cold Spring Harbor Perspectives in Biology*, *3*(1), 1–14. https://doi.org/10.1101/cshperspect.a000778

Praitis, V., & Maduro, M. F. (2011). Transgenesis in C. elegans. *Methods in Cell Biology*, *106*, 159–185. https://doi.org/10.1016/B978-0-12-544172-8.00006-2

Purves, D., Augustine, G. J., Fitzpatrick, D., Katz, L. C., LaMantia, A.-S., McNamara, J. O., & Williams, S. M. (2001). The Retina. In *Neuroscience* (2nd editio). Sunderland (MA): Sinauer Associates. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK10885/

Qadota, H., Inoue, M., Hikita, T., Köppen, M., Hardin, J. D., Amano, M., … Kaibuchi, K. (2007). Establishment of a tissue-specific RNAi system in C. elegans. *Gene*, *400*(1-2), 166–73. https://doi.org/10.1016/j.gene.2007.06.020

Ramani, A. K., Calarco, J. A., Pan, Q., Mavandadi, S., Wang, Y., Nelson, A. C., … Fraser, A. G. (2011). Genome-wide analysis of alternative splicing in Caenorhabditis elegans. *Genome Research*, *21*(2), 342–348. https://doi.org/10.1101/gr.114645.110

Ramani, A. K., Nelson, A. C., Kapranov, P., Bell, I., Gingeras, T. R., & Fraser, A. G. (2009). High resolution transcriptome maps for wild-type and nonsense-mediated decay-defective Caenorhabditis elegans. *Genome Biology*, *10*(9), R101. https://doi.org/10.1186/gb-2009-10-9-r101

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat. Protocols*, *8*(11), 2281–2308. https://doi.org/10.1038/nprot.2013.143\rhttp://www.nature.com/nprot/journal/v8/n11/abs/nprot.2013.143.html#supplementary-information

Ranganathan, R., Sawin, E. R., Trent, C., & Horvitz, H. R. (2001). Mutations in the Caenorhabditis elegans serotonin reuptake transporter MOD-5 reveal serotonin-dependent and -independent activities of fluoxetine. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *21*(16), 5871–84. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11487610

Riddle, D. L. (1997). *C. elegans II*. Cold Spring Harbor Laboratory Press. Retrieved from http://www.cshlpress.org/default.tpl?cart=149489634726846464&fromlink=T&linkaction=full&linksortby=oop_title&--eqSKUdatarq=38

Rio Frio, T., Wade, N. M., Ransijn, A., Berson, E. L., Beckmann, J. S., & Rivolta, C. (2008). Premature termination codons in PRPF31 cause retinitis pigmentosa via haploinsufficiency due to nonsense-mediated mRNA decay. *Journal of Clinical Investigation*, *118*(4), 1519–1531. https://doi.org/10.1172/JCI34211

Robertson, A. M. G., & Thomson, J. N. (1982). Morphology of programmed cell death in the ventral nerve cord of Caenorhabditis elegans larvae. *Embryol Exp. Morph*, *67*, 89–100.

Rose, A. M., & Bhattacharya, S. S. (2016). Variant haploinsufficiency and phenotypic non-penetrance in PRPF31-associated retinitis pigmentosa. *Clinical Genetics*, *90*(2), 118–126. https://doi.org/10.1111/cge.12758

Rose, A. M., Shah, A. Z., Venturini, G., Krishna, A., Chakravarti, A., Rivolta, C., & Bhattacharya, S. S. (2016). Transcriptional regulation of PRPF31 gene expression by MSR1 repeat elements causes incomplete penetrance in retinitis pigmentosa. *Scientific Reports*, *6*(1), 19450. https://doi.org/10.1038/srep19450

Rual, J. F., Ceron, J., Koreth, J., Hao, T., Nicot, A. S., Hirozane-Kishikawa, T., … Vidal, M. (2004). Toward improving Caenorhabditis elegans phenome mapping with an ORFeome-based RNAi library. *Genome Research*, *14*(10 B), 2162–2168. https://doi.org/10.1101/gr.2505604.7

Rubio-Peña, K., Fontrodona, L., Aristizábal-Corrales, D., Torres, S., Cornes, E., García-Rodríguez, F. J., … Cerón, J. (2015). Modeling of autosomal-dominant retinitis pigmentosa in *Caenorhabditis elegans* uncovers a nexus between global impaired functioning of certain splicing factors and cell type-specific apoptosis. *Rna*, *21*(12), 2119–2131. https://doi.org/10.1261/rna.053397.115

Ruzickova, S., & Stanek, D. (2016). Mutations in spliceosomal proteins and retina degeneration. *RNA Biology*, *6286*(November), 1–9. https://doi.org/10.1080/15476286.2016.1191735

Sanford, J. R., & Caceres, J. F. (2004). Pre-mRNA splicing: life at the centre of the central dogma. *Journal of Cell Science*, *117*(Pt 26), 6261–6263. https://doi.org/10.1242/jcs.01513

Santos-Pereira, J. M., Herrero, A. B., García-Rubio, M. L., Marín, A., Moreno, S., & Aguilera, A. (2013). The Npl3 hnRNP prevents R-loop-mediated transcription-replication conflicts and genome instability. *Genes & Development*, *27*(22), 2445–58. https://doi.org/10.1101/gad.229880.113

Schreiber, F., Patricio, M., Muffato, M., Pignatelli, M., & Bateman, A. (2014). TreeFam v9: a new website, more species and orthology-on-the-fly. *Nucleic Acids Research*, *42*(D1), D922–D925. https://doi.org/10.1093/nar/gkt1055

Schumacher, B., Hofmann, K., Boulton, S., & Gartner, A. (2001). The C. elegans homolog of the p53 tumor suppressor is required for DNA damage-induced apoptosis. *Current Biology : CB*, *11*(21), 1722–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11696333

Scotti, M. M., & Swanson, M. S. (2016). RNA mis-splicing in disease. *Nature Reviews. Genetics*, *17*(1), 19–32. https://doi.org/10.1038/nrg.2015.3

Shaham, S., & Horvitz, H. R. (1996). Developing Caenorhabditis elegans neurons may contain both cell-death protective and killer activities. *Genes & Development*, *10*(5), 578–91. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8598288

Sharp, P. A. (2005). The discovery of split genes and RNA splicing. *Trends in Biochemical Sciences*, *30*(6), 279–281. https://doi.org/10.1016/j.tibs.2005.04.002

Singh, R., & Sulston, J. (1978). Some Observations On Moulting in Caenorhabditis Elegans. *Nematologica*, *24*(1), 63–71. https://doi.org/10.1163/187529278X00074

Skourti-Stathaki, K., & Proudfoot, N. J. (2014). A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. *Genes & Development*, *28*(13), 1384–96. https://doi.org/10.1101/gad.242990.114

Skourti-Stathaki, K., Proudfoot, N. J., & Gromak, N. (2011). Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Molecular Cell*, *42*(6), 794–805. https://doi.org/10.1016/j.molcel.2011.04.026

Sollier, J., Stork, C. T., García-Rubio, M. L., Paulsen, R. D., Aguilera, A., & Cimprich, K. A. (2014). Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Molecular Cell*, *56*(6), 777–85. https://doi.org/10.1016/j.molcel.2014.10.020

Sönnichsen, B., Koski, L. B., Walsh, A., Marschall, P., Neumann, B., Brehm, M., … Echeverri, C. J. (2005). Full-genome RNAi profiling of early embryogenesis in Caenorhabditis elegans. *Nature*, *434*(7032), 462–9. https://doi.org/10.1038/nature03353

Stergiou, L., & Hengartner, M. O. (2004). Death and more: DNA damage response pathways in the nematode C. elegans. *Cell Death and Differentiation*, *11*(1), 21–28. https://doi.org/10.1038/sj.cdd.4401340

Sternberg, P. W., & Horvitz, H. R. (1984). The Genetic Control of Cell Lineage During Nematode Development. *Annual Review of Genetics*, *18*(1), 489–524. https://doi.org/10.1146/annurev.ge.18.120184.002421

Stork, C. T., Bocek, M., Crossley, M. P., Sollier, J., Sanz, L. A., Chédin, F., … Cimprich, K. A. (2016). Co-transcriptional R-loops are the main cause of estrogen-induced DNA damage. *eLife*, *5*. https://doi.org/10.7554/eLife.17548

Suetomi, K., Mereiter, S., Mori, C., Takanami, T., & Higashitani, A. (2013). Caenorhabditis elegans ATR checkpoint kinase ATL-1 influences life span through mitochondrial maintenance. https://doi.org/10.1016/j.mito.2013.02.004

Sulston, J. (1988). 05 Cell Lineage. In W. B. Wood. (Ed.), *The Nematode Caenorhabditis elegans* (pp. 123–156). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. https://doi.org/10.1101/087969307.17.123

Sulston, J. E., & Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. *Developmental Biology*, *56*(1), 110–56. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/838129

Sulston, J. E., Schierenberg, E., White, J. G., & Thomson, J. N. (1983). The embryonic cell lineage of the nematode Caenorhabditis elegans. *Developmental Biology*, *100*(1), 64–119. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6684600

Sun, H., & Chasin, L. A. (2000). Multiple splicing defects in an intronic false exon. *Molecular and Cellular Biology*, *20*(17), 6414–25. https://doi.org/10.1128/MCB.20.17.6414-6425.2000

Sundaram, M., & Greenwald, I. (1993). Suppressors of a lin-12 hypomorph define genes that interact with both lin-12 and glp-1

in Caenorhabditis elegans. *Genetics*, *135*(3), 765–83. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8293978

Tanackovic, G., Ransijn, A., Ayuso, C., Harper, S., Berson, E. L., & Rivolta, C. (2011). A missense mutation in PRPF6 causes impairment of pre-mRNA splicing and autosomal-dominant retinitis pigmentosa. *American Journal of Human Genetics*, *88*(5), 643–649. https://doi.org/10.1016/j.ajhg.2011.04.008

Tanackovic, G., Ransijn, A., Thibault, P., Abou Elela, S., Klinck, R., Berson, E. L., … Rivolta, C. (2011). PRPF mutations are associated with generalized defects in spliceosome formation and pre-mRNA splicing in patients with retinitis pigmentosa. *Human Molecular Genetics*, *20*(11), 2116–30. https://doi.org/10.1093/hmg/ddr094

Tang, Z., Zhang, Y., Wang, Y., Zhang, D., Shen, B., Luo, M., & Gu, P. (2017). Progress of stem/progenitor cell-based therapy for retinal degeneration. *Journal of Translational Medicine*, *15*(1), 99. https://doi.org/10.1186/s12967-017-1183-y

Tibbetts, R. S., Brumbaugh, K. M., Williams, J. M., Sarkaria, J. N., Cliby, W. A., Shieh, S. Y., … Abraham, R. T. (1999). A role for ATR in the DNA damage-induced phosphorylation of p53. *Genes & Development*, *13*(2), 152–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9925639

Tina L. Gumienny, Eric Lambie, Erika Hartwieg, H. R. H. and M. O. H. (1999). Genetic control of programmed cell death in the Caenorhabditis elegansINTRODUCTION. *Development*, *1022*(126), 1011–1022.

Towns, K. V., Kipioti, A., Long, V., McKibbin, M., Maubaret, C., Vaclavik, V., … Inglehearn, C. F. (2010). Prognosis for splicing factor PRPF8 retinitis pigmentosa, novel mutations and correlation between human and yeast phenotypes. *Human Mutation*, *31*(5), 1361–1376. https://doi.org/10.1002/humu.21236

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., … Pachter, L. (2012). Differential gene and transcript expression

analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*, *7*(3), 562–78. https://doi.org/10.1038/nprot.2012.016

Tuduri, S., Crabbé, L., Conti, C., Tourrière, H., Holtgreve-Grez, H., Jauch, A., … Pasero, P. (2009). Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription. *Nature Cell Biology*, *11*(11), 1315–24. https://doi.org/10.1038/ncb1984

Vermezovic, J., Stergiou, L., Hengartner, M. O., & d'Adda di Fagagna, F. (2012). Differential regulation of DNA damage response activation between somatic and germline cells in Caenorhabditis elegans. *Cell Death and Differentiation*, *19*(11), 1847–1855. https://doi.org/10.1038/cdd.2012.69

Vithana, E. N., Abu-Safieh, L., Allen, M. J., Carey, A., Papaioannou, M., Chakarova, C., … Bhattacharya, S. S. (2001). A human homolog of yeast pre-mRNA splicing gene, PRP31, underlies autosomal dominant retinitis pigmentosa on chromosome 19q13.4 (RP11). *Molecular Cell*, *8*(2), 375–81. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11545739

Vithana, E. N., Abu-Safieh, L., Pelosini, L., Winchester, E., Hornan, D., Bird, A. C., … Bhattacharya, S. S. (2003). Expression of PRPF31 mRNA in patients with autosomal dominant retinitis pigmentosa: A molecular clue for incomplete penetrance? *Investigative Ophthalmology and Visual Science*, *44*(10), 4204–4209. https://doi.org/10.1167/iovs.03-0253

Waaijers, S., & Boxem, M. (2014). Engineering the Caenorhabditis elegans genome with CRISPR/Cas9. *Methods*, *68*(3), 381–388. https://doi.org/10.1016/j.ymeth.2014.03.024

Wahba, L., Amon, J. D., Koshland, D., & Vuica-Ross, M. (2011). RNase H and Multiple RNA Biogenesis Factors Cooperate to Prevent RNA:DNA Hybrids from Generating Genome Instability. *Molecular Cell*, *44*(6), 978–988. https://doi.org/10.1016/j.molcel.2011.10.017

Wahl, M. C., Will, C. L., & Lührmann, R. (2009). The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell*, *136*(4), 701–718. https://doi.org/10.1016/j.cell.2009.02.009

Wang, G.-S., & Cooper, T. a. (2007). Splicing in disease: disruption of the splicing code and the decoding machinery. *Nature Reviews. Genetics*, *8*(october), 749–761. https://doi.org/10.1038/nrg2164

White, J. (1988). 04 The Anatomy. In W. B. Wood. (Ed.), *The Nematode Caenorhabditis elegans* (pp. 81–122). New York: ing Harbor Laboratory Press, Cold Spring Harbor, NY.

Wildwater, M., Sander, N., de Vreede, G., & van den Heuvel, S. (2011). Cell shape and Wnt signaling redundantly control the division axis of C. elegans epithelial stem cells. *Development*, *138*(20), 4375–4385. https://doi.org/10.1242/dev.066431

Wiley, L. A., Burnight, E. R., Mullins, R. F., Stone, E. M., & Tucker, B. A. (2014). Stem cells as tools for studying the genetics of inherited retinal degenerations. *Cold Spring Harbor Perspectives in Medicine*, *5*(5), a017160. https://doi.org/10.1101/cshperspect.a017160

Wilkie, S. E., Vaclavik, V., Wu, H., Bujakowska, K., Chakarova, C. F., Bhattacharya, S. S., … Hunt, D. M. (2008). Disease mechanism for retinitis pigmentosa (RP11) caused by missense mutations in the splicing factor gene PRPF31. *Molecular Vision*, *14*, 683–90. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18431455

Will, C. L., & Lührmann, R. (2006). Spliceosome Structure and Function. In R. F. et al. Gesteland (Ed.), *The RNA World* (3rd ed., pp. 369–400). New York: Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. Retrieved from http://cshmonographs.org/index.php/monographs/article/viewFile/3739/2959

Will, C. L., & Lührmann, R. (2011). Spliceosome Structure and Function. *Cold Spring Harbor Perspectives in Biology*, *3*(7), a003707–a003707. https://doi.org/10.1101/cshperspect.a003707

Wirth, B. (2000). An update of the mutation spectrum of the survival motor neuron gene (SMN1) in autosomal recessive spinal muscular atrophy (SMA). *Human Mutation*, *15*(3), 228–237. https://doi.org/10.1002/(SICI)1098-1004(200003)15:3<228::AID-HUMU3>3.0.CO;2-9

Yoshida, K., & Ogawa, S. (2014). Splicing factor mutations and cancer. *Wiley Interdisciplinary Reviews: RNA*, *5*(4), 445–459. https://doi.org/10.1002/wrna.1222

Yu, K., Chedin, F., Hsieh, C.-L., Wilson, T. E., & Lieber, M. R. (2003). R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nature Immunology*, *4*(5), 442–451. https://doi.org/10.1038/ni919

Yuan, J., & Horvitz, H. R. (1992). The Caenorhabditis elegans cell death gene ced-4 encodes a novel protein and is expressed during the period of extensive programmed cell death. *Development (Cambridge, England)*, *116*(2), 309–20. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1286611

Yuan, J., Shaham, S., Ledoux, S., Ellis, H. M., & Horvitz, H. R. (1993). The C. elegans cell death gene ced-3 encodes a protein similar to mammalian interleukin-1 beta-converting enzyme. *Cell*, *75*(4), 641–52. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8242740

Zeller, P., Padeken, J., van Schendel, R., Kalck, V., Tijsterman, M., & Gasser, S. M. (2016). Histone H3K9 methylation is dispensable for Caenorhabditis elegans development but suppresses RNA:DNA hybrid-associated repeat instability. *Nature Genetics*, *48*(11), 1385–1395. https://doi.org/10.1038/ng.3672

Zhao, C., Bellur, D. L., Lu, S., Zhao, F., Grassi, M. A., Bowne, S. J., … Larsson, C. (2009). Autosomal-Dominant Retinitis Pigmentosa Caused by a Mutation in SNRNP200, a Gene Required for Unwinding of U4/U6 snRNAs. *American Journal of Human Genetics*, *85*(5), 617–627. https://doi.org/10.1016/j.ajhg.2009.09.020

Zheng, A., Li, Y., & Tsang, S. H. (2015). Personalized therapeutic strategies for patients with retinitis pigmentosa. *Expert Opinion on Biological Therapy*, *15*(3), 391–402. https://doi.org/10.1517/14712598.2015.1006192

Zhou, T., Hasty, P., Walter, C. A., Bishop, A. J. R., Scott, L. M., & Rebel, V. I. (2013). Myelodysplastic syndrome: An inability to appropriately respond to damaged DNA? *Experimental Hematology*, *41*(8), 665–674. https://doi.org/10.1016/j.exphem.2013.04.008

Rubio-Peña K, Fontrodona L, Aristizábal-Corrales D, Torres S, Cornes E, García-Rodríguez FJ, et al. Modeling of autosomal-dominant retinitis pigmentosa in Caenorhabditis elegans uncovers a nexus between global impaired functioning of certain splicing factors and cell type-specific apoptosis. RNA. 2015 Dec;21(12):2119–31. DOI: 10.1261/rna.053397.115