



Drosophila melanogaster As a Model Organism to Study RNA Toxicity of Repeat Expansion-Associated Neurodegenerative and Neuromuscular Diseases

Alex C. Koon^{1,2} and Ho Yin Edwin Chan^{1,2,3,4,5*}

¹ Laboratory of *Drosophila* Research, Hong Kong, Hong Kong, ² Biochemistry Program, Hong Kong, Hong Kong, ³ Cell and Molecular Biology Program, Hong Kong, Hong Kong, ⁴ Molecular Biotechnology Program, Faculty of Science, School of Life Sciences, Hong Kong, Hong Kong, ⁵ School of Life Sciences, Gerald Choa Neuroscience Centre, The Chinese University of Hong Kong, Hong Kong, Hong Kong

OPEN ACCESS

Edited by:

Rosanna Parlato,
University of Ulm, Germany

Reviewed by:

Kah-Leong Lim,
National Neuroscience Institute,
Singapore
Kim A. Staats,
University of Southern California, USA

*Correspondence:

Ho Yin Edwin Chan
hyechan@cuhk.edu.hk

Received: 12 January 2017

Accepted: 27 February 2017

Published: 21 March 2017

Citation:

Koon AC and Chan HYE (2017)
Drosophila melanogaster As a Model
Organism to Study RNA Toxicity of
Repeat Expansion-Associated
Neurodegenerative and
Neuromuscular Diseases.
Front. Cell. Neurosci. 11:70.
doi: 10.3389/fncel.2017.00070

For nearly a century, the fruit fly, *Drosophila melanogaster*, has proven to be a valuable tool in our understanding of fundamental biological processes, and has empowered our discoveries, particularly in the field of neuroscience. In recent years, *Drosophila* has emerged as a model organism for human neurodegenerative and neuromuscular disorders. In this review, we highlight a number of recent studies that utilized the *Drosophila* model to study repeat-expansion associated diseases (READs), such as polyglutamine diseases, fragile X-associated tremor/ataxia syndrome (FXTAS), myotonic dystrophy type 1 (DM1) and type 2 (DM2), and C9ORF72-associated amyotrophic lateral sclerosis/frontotemporal dementia (C9-ALS/FTD). Discoveries regarding the possible mechanisms of RNA toxicity will be focused here. These studies demonstrate *Drosophila* as an excellent *in vivo* model system that can reveal novel mechanistic insights into human disorders, providing the foundation for translational research and therapeutic development.

Keywords: polyglutamine disease, Huntington's disease, C9orf72, amyotrophic lateral sclerosis (ALS), frontotemporal dementia (FTD), myotonic dystrophy, fragile X, spinocerebellar ataxia

INTRODUCTION

For nearly two decades, the fruit fly *Drosophila melanogaster* has been utilized as a model organism to study a number of human neurodegenerative and neuromuscular diseases (Chan and Bonini, 2000; Muqit and Feany, 2002; Zoghbi and Botas, 2002; Sang and Jackson, 2005; Marsh and Thompson, 2006; Yu and Bonini, 2011; Jaiswal et al., 2012; McGurk et al., 2015). The fruit fly offers multiple advantages for the investigation of the molecular mechanisms of diseases. Having short life cycle, high offspring numbers, low costs for maintenance, simple yet powerful genetic manipulations, the availability of mutants, and other genetic tools are only some of the many attractive features of *Drosophila* as a model system. *Drosophila* was introduced into

scientific research over a 100 years ago, and quickly became an invaluable tool that empowers our discoveries and understanding of a wide range of biological processes, such as embryogenesis, neural development, synaptic plasticity, and even complex behaviors such as decision-making and learning and memory (Bellen et al., 2010; Spindler and Hartenstein, 2010; Harris and Littleton, 2015). Sequencing of the *Drosophila* and the human genomes revealed remarkably high similarities between the fly and humans (Adams et al., 2000; Rubin, 2000), with ~75% of the genes implicated in human genetic disorders having at least one homolog in *Drosophila* (Rubin, 2000; Reiter et al., 2001). Most importantly, despite the obvious anatomic divergence between the fly and humans, the fundamental aspects of cell biology are highly conserved, including the regulation of gene expression, synaptogenesis, cell proliferation, cell differentiation, cell signaling, and cell death. Many genes and pathways that are being studied in mammals were originally identified in *Drosophila*. For instance, the mammalian *Wnt* gene was originally identified as *wingless* in *Drosophila* with its mutation resulting in flies having no wings (Sharma and Chopra, 1976). *Wnt/wingless* is now widely known for its involvement in a broad variety of conserved cellular processes and human diseases (Korkut and Budnik, 2009).

THE ADVANTAGES OF USING *DROSOPHILA* FOR STUDYING HUMAN DISEASE-ASSOCIATED GENES

Comparing with mammals, *Drosophila* has much simpler genetics. The fruit fly has only four pairs of homologous chromosomes, as compared to 20 in mice and 23 in humans. Flies also have a much simpler nervous system, with ~200,000 neurons in the fly brain comparing to ~100 billion neurons in humans. Despite its simplicity, the fly nervous system is comprehensive, allowing the fly to perform a wide variety of behaviors, such as feeding, walking, climbing, courtship, and communication. Fruit flies are even capable of being trained with fear/reward conditioning paradigms to examine complex biological processes like learning and memory. These features of *Drosophila* make it an ideal organism for modeling complex disorders because it can be used to model a specific subset of phenotypes associated with a particular disease, which will simplify the analysis.

A wide range of genetic manipulation techniques have been developed in *Drosophila* which are impractical to be implemented in mammals. For instance, the availability of *P*-elements mutants in libraries is certainly an advantage that is exclusive to the *Drosophila* community. Originally, *P*-elements are stretches of transposable DNA consisting of inverted repeats surrounding a transposase gene to allow hopping of the mobile element on chromosomes (Rubin and Spradling, 1982; O'Hare and Rubin, 1983). *Drosophila* researchers exploited its mobility, and designed *P*-elements without the transposase but with reporter genes. By pairing up this modified *P*-element with an inducible transposase, the *P*-element can then be randomly inserted into the genome to disrupt the transcription of

downstream genes. With this technique, libraries of “*P*-element mutants” for tenths of thousands of genes are constructed (Cooley et al., 1988; Bier et al., 1989; Spradling et al., 1999; Bellen et al., 2004). Researchers can now conveniently access the information of these mutants from FlyBase (<http://flybase.org>), and order them from resource centers such as the Bloomington *Drosophila* Stock Center. In addition, due to the nature that the mobilization of the *P*-elements are typically imprecise, these fly lines provide a means to generate excision mutants (Hummel and Klämbt, 2008). Lastly, the introduction of flipase recombination targets into *P*-elements have provided the option of generating precise chromosomal deletions between two *P*-elements (Hummel and Klämbt, 2008). The development of these genetic tools have greatly enhanced the process of genome editing in *Drosophila*.

Another popular technique used in *Drosophila* research is the GAL4/UAS binary transgene overexpression system, which is perhaps one of the most versatile expression system ever developed in *Drosophila* (Brand and Perrimon, 1993). GAL4 is a yeast transcription factor that drives the expression of a transgene downstream of an *Upstream Activation Sequence* (UAS). Usually, a GAL4 fly line expresses GAL4 under the control of a cell- or tissue-specific promoter. This can be achieved by enhancer-trap screens, or the fusion of identified promoters with the GAL4 gene and subsequently microinjecting the constructs into fly embryos to produce specific GAL4 driver lines (Duffy, 2002). For the operon, a human disease gene of interest is usually subcloned into an expression construct containing UAS, which is then also microinjected into fly embryos to produce transgenic fly lines. The UAS operon lines will be crossed to GAL4 fly lines, resulting in the overexpression of the human disease gene of interest in the desired cells/tissues (Duffy, 2002). If the expression of pathological human genes in the fly successfully yields an abnormal phenotype, such as degeneration in the photoreceptor neurons of the fly eye, genetic screens, and pharmacological screens can then be performed to search for genetic modifiers and small molecules that may impact the etiology of the corresponding human disorders.

One of the most prominent aspects of the fly model is its capacity to conduct genetic screens to identify novel molecular components of pathological pathways. Mice models of human diseases are absolutely vital to advancing our understanding of the diseases, but nevertheless with limitations. In *Drosophila*, the classical forward genetic screens can be performed using mutagens such as ethyl methane sulfonate to generate mutants (St Johnston, 2002; Yamamoto et al., 2014). Since this approach is unbiased, it allows the identification of novel genes from the entire genome, which would otherwise be extremely difficult to perform in mouse models. Due to having a short life cycle, the simplicity of having only four pairs of homologous chromosomes and the availability of balancer chromosomes for the suppression of recombination, the isolation of mutations from *Drosophila* is much easier than that from a mouse. Apart from forward genetic screens, tools for high throughput reverse genetic screens are also available. Genetic tools for RNAi-mediated knockdown are easily accessible from stock centers such as the Vienna *Drosophila* Resource Center and Harvard Transgenic RNAi Project, making

it possible to conveniently perform reverse genetic screens in the fly model (Dietzl et al., 2007).

External features of *Drosophila* makes it suitable for performing large scale screens of genes and pharmacologicals. Bristles, wing veins and the compound eye are all external structures of the organism that can be affected by genetic or cellular changes, and can easily be observed and scored under a stereomicroscope (Yamamoto et al., 2014). The compound eye is mostly composed of photoreceptor neurons, and is commonly used as a model to monitor neurodegeneration, allowing the identification of novel genetic modifiers and drugs in screens (St Johnston, 2002). Using the GAL4/UAS system, toxic RNAs or toxic proteins that commonly occurs in human neurodegenerative and neuromuscular diseases can be expressed in the fly eye, which often results in a rough-eye phenotype. Since the eye is not required for viability, this allows effective screening of genes or pharmacologicals that can alleviate or enhance the degeneration. Most compounds can be simply mixed into standard *Drosophila* food to be consumed by larvae and adult flies. Alternatively, compounds can be injected into adult flies as well if a high, acute dosage is needed in the adult stage. The identification of novel genetic modifiers or drugs may reveal the physiological role of disease-related genes and promote our understanding of the disease pathology (Zoghbi and Botas, 2002; Hirth, 2010; Pandey and Nichols, 2011; McGurk et al., 2015). Thus, *Drosophila* serves as an important platform for the discoveries of new components of pathological pathways and the development of pioneer therapeutic approaches.

THE LIMITATIONS OF MODELING HUMAN DISEASES IN THE FRUIT FLY

Although the use of classical genetic screens of *Drosophila* provides an unbiased approach to gain insights into human degenerative diseases, fly mutations may not be precise representations of human disease mutations. This is because the classical fly mutants from forward genetic screens are typically loss-of-function alleles. Yet in human diseases, mutations can have complex presentations, including both loss-of-function of the wild type allele and gain-of-function of the mutant allele. Most of the times, this gain-of-function effect can still be modeled in *Drosophila* by overexpressing the mutant protein using the GAL4/UAS system (St Johnston, 2002). But for some diseases which the loss-of-function of the wild type protein plays a major role in the disease etiology, if there are no *Drosophila* orthologs of the corresponding human disease-associated genes, then perhaps the fly is not a suitable model for the particular disease.

Furthermore, while the GAL4/UAS system is an extremely effective and versatile system, it is after all an overexpression system that may give rise to the typical caveats of overexpression studies. The magnitude of overexpression in a model can be tremendously different from the clinical situation (Floresco et al., 2005). Sometimes, even overexpressing a wild type version of a gene may cause a disease phenotype (Prelich, 2012). In fact, excess GAL4 protein in the *Drosophila* eye may lead to degeneration (Kramer and

Staveley, 2003). Hence, it is paramount for *Drosophila* researchers to set up appropriate controls to ensure that the phenotypes observed are not simply overexpression artifacts.

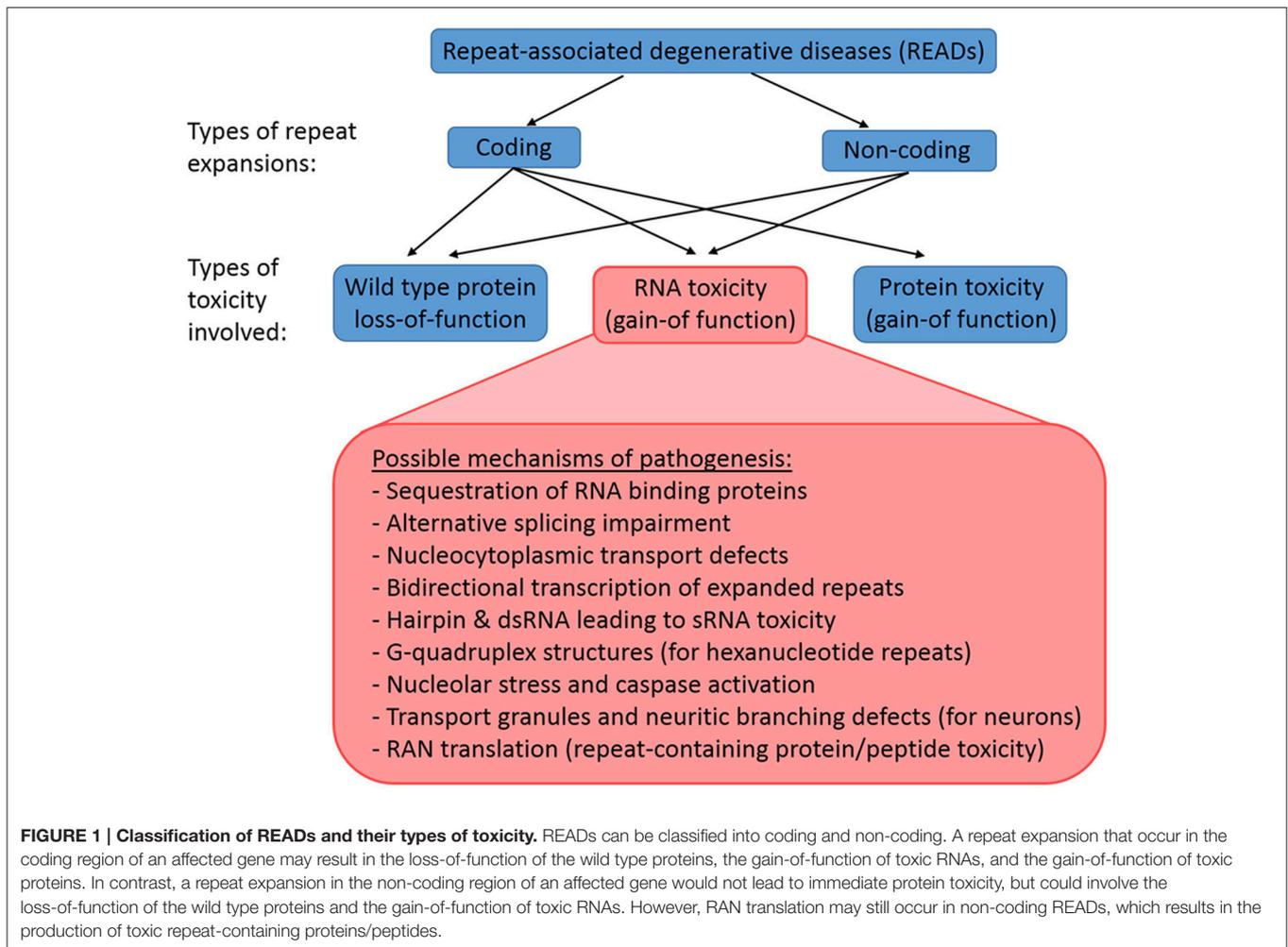
Lastly, it is a concern of whether the fly models can faithfully recapitulate the selective neuronal death and associated phenotypes of human neurological disorders. Although the *Drosophila* photoreceptor degeneration provides a convenient readout, it mostly reflects generic neurotoxicity instead of selective neurotoxicity that is disease-specific. Thus, it is important to verify relevant phenotypes using other systems that are more specific to the particular types of disease. For instance, for myotonic dystrophy, it is crucial to verify any degeneration phenotype in a neuromuscular system, like the *Drosophila* or mouse neuromuscular junction (de Haro et al., 2006; Panaite et al., 2013).

Despite all the limitations, *Drosophila* has so far proven to be a valid and useful tool in unraveling the etiology of human diseases and identifying compounds that can improve symptoms and modify the course of human diseases. Success examples include the rescuing of disease-like phenotypes in fly models of fragile X syndrome (Chang et al., 2008), prolonged survival of dopaminergic neurons in fly models of Parkinson's disease (Auluck et al., 2005; Faust et al., 2009) and lifespan extension in fly models of Alzheimer's disease (Rajendran et al., 2008).

REPEAT EXPANSION-ASSOCIATED NEURODEGENERATIVE AND NEUROMUSCULAR DISEASES

As early as 1918, unstable nucleotide (microsatellite) repeat expansion had already been identified as a mutational mechanism underlying several human disease (Fleischer, 1918). Originally, only trinucleotide repeats were associated with human diseases (Fu et al., 1991; La Spada et al., 1991). But over the years, other disease-causing expandable microsatellite repeats were gradually discovered including unstable tetra-, penta-, and hexa-nucleotide repeats (Mirkin, 2007). Remarkably, the majority of the repeat expansion-associated diseases (READs) manifest with neurological and neuromuscular symptoms. Until today, more than 20 neurodegenerative and neuromuscular disorders have been linked with unstable nucleotide repeat expansions in the human genome (Pearson et al., 2005; Brouwer et al., 2009; Todd and Paulson, 2010; Polak et al., 2013).

Repeat expansions can be coding or non-coding, and the underlying mechanism of a READ may involve (1) loss-of-function of wild-type protein, (2) gain-of-function of toxic RNAs, or (3) gain-of-function of toxic proteins (Figure 1). Coding expansion disorders usually result from smaller repeat copies (few to tens of copies) that manifest in exon coding regions of genes. This group of READs includes mostly polyglutamine (polyQ) diseases like Huntington's disease (HD) and several forms of spinocerebellar ataxia (SCA), which typically involves the translation of the gene products together with the repeat



expansions into toxic proteins with gain-of-function (La Spada et al., 1991; MacDonald et al., 1993; Orr et al., 1993; Kawaguchi et al., 1994; Imbert et al., 1996; Pulst et al., 1996; Sanpei et al., 1996; David et al., 1997; Zhuchenko et al., 1997; Nakamura et al., 2001). In contrast, the non-coding expansions disorders typically involve larger expansions (from ~100 to 1000 copies), which reside in the non-coding regions of genes, such as promoters, introns or untranslated regions (UTRs). This groups includes fragile X-associated tremor/ataxia syndrome (FXTAS), Friedrich's ataxia, C9-ALS/FTD, myotonic dystrophy types 1 and 2 (DM1 and DM2) and SCA8 (Brook et al., 1992; Montermini et al., 1997; Jin et al., 2003; Mutsuddi et al., 2004; DeJesus-Hernandez et al., 2011; Renton et al., 2011). Non-coding expansions usually results in repeat-containing transcripts with gain-of-function RNA toxicity. However, even non-coding expansions can still be translated into toxic peptides by repeat-associated non-ATG translation (Zu et al., 2011). Thus, the pathologies of READs are complex, as both coding and non-coding repeat expansions may involve a combination of mechanisms, including wild-type proteins loss-of-function, toxic RNA gain-of-function, and toxic protein gain-of-function.

RNA GAIN-OF-FUNCTION TOXICITY IN READS AND ITS PATHOGENIC MECHANISMS

The concept that RNA itself acts as a major toxic species in neurodegenerative/neuromuscular disorders was first described for DM1. DM1 is the most common form of muscular dystrophy in adults, caused by an expanded CTG repeat in the 3'UTR of the *dystrophia myotonica protein kinase (DMPK)* gene (Brook et al., 1992; Fu et al., 1992; Mahadevan et al., 1992). Ribonuclear foci were found to be forming in the nuclei of the affected cells, and these foci contain the RNA transcripts from the CTG repeat expansion (Taneja et al., 1995). In brief, the expanded repeats localized at the foci in turn sequester endogenous RNA binding proteins, impairing the alternatively splicing mechanisms and leading to cellular toxicity and degeneration (**Figure 2**; Renoux and Todd, 2012; Belzil et al., 2013; Fiszer and Krzyzosiak, 2013; Nalavade et al., 2013). This is the "sequestration hypothesis" or "alternative splicing impairment model," which is one of the most supported mechanisms of the emerging "RNA toxicity model." The ribonuclear foci are now a hallmark for most of the

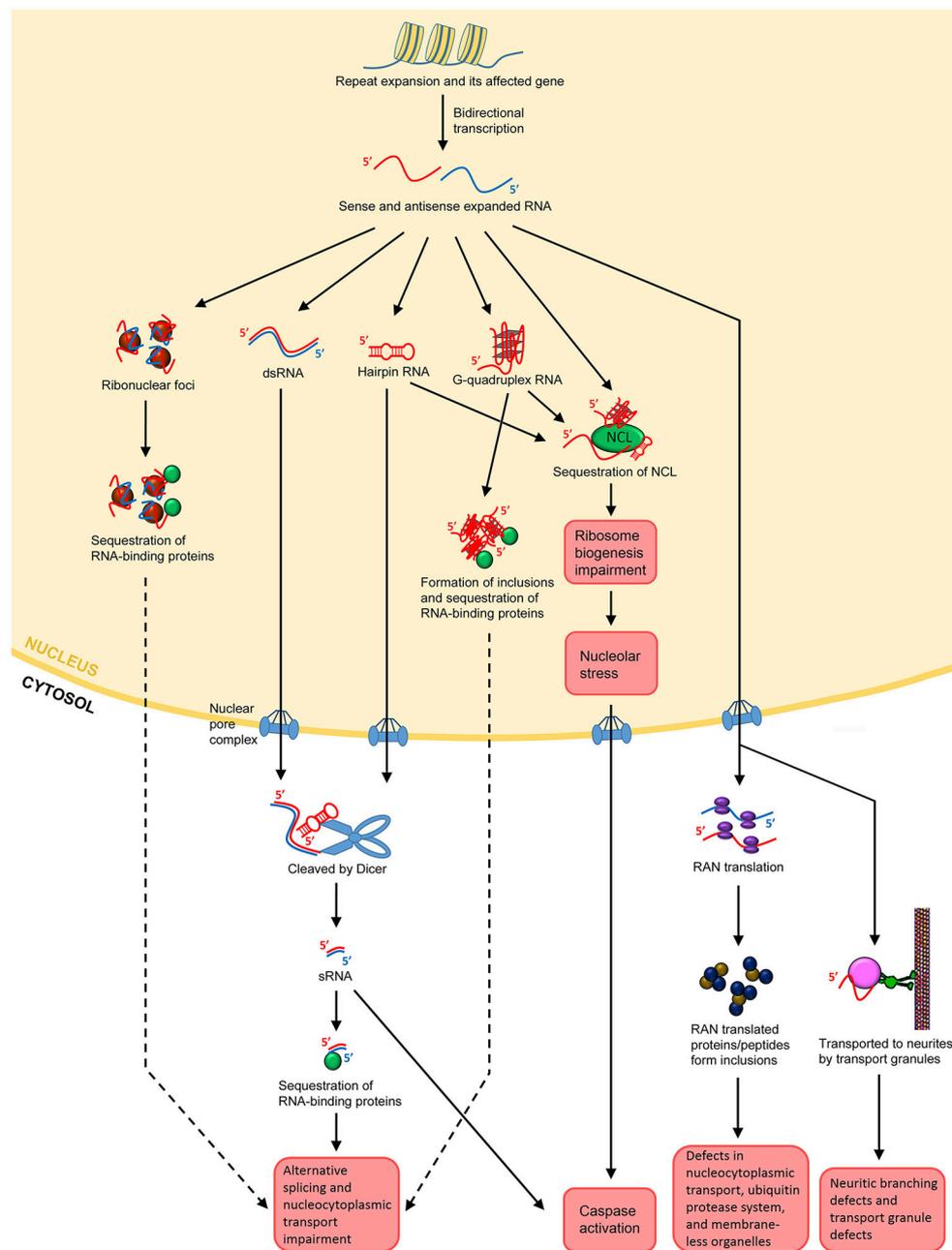


FIGURE 2 | Possible mechanisms of RNA toxicity in repeat expansion-associated diseases. Repeat expansion can occur in coding or non-coding regions of affected genes. Bidirectional transcription results in both sense and antisense transcripts containing the repeats. Both sense and antisense expanded RNAs can form ribonuclear foci. The foci may sequester RNA-binding proteins, such as MBNL1 in the case of DM1, leading to impairment of the alternative splicing and nucleocytoplasmic transport machineries. The sense and antisense ssRNAs can anneal to form dsRNAs. ssRNAs can also form hairpin structures by themselves, which are degradation resistant. Both dsRNAs and hairpin-forming RNAs may export the nucleus and be cleaved into sRNAs by Dicer. The sRNAs can then sequester RNA-binding proteins, leading to splicing and nucleocytoplasmic transport impairment. For hexanucleotide repeats such as GGGCC in C9-ALS/FTD, the ssRNA may form G-quadruplex structures. These structures facilitate DNA/RNA hybrid and impede transcription. They also form inclusions, leading to the sequestration of RNA-binding proteins as well. Repeat-containing transcripts, whether in hairpin or G-quadruplex structures, can bind to NCL, inducing nucleolar stress and caspase activation. Expanded transcripts that escape the nucleus associate themselves with ribosomes, leading to RAN translation. RAN translational products may form inclusions, which disrupts nucleocytoplasmic transport, impedes the ubiquitin protease system and impairs the assembly, dynamics, and functions of membrane-less organelles. In neurons, the expanded transcripts may also be actively transported to neurites, resulting in neuritic branching, and transport granule defects. It is postulated that the local translational machinery may also be disrupted.

READs associated with non-coding repeats (Wojciechowska and Krzyzosiak, 2011). Some even coined the term “RNAopathies” for these disorders due to the clear involvement of RNA toxicity. Noted that, despite the consensus that the formation of foci is a feature of pathogenesis, their role in pathogenesis is still unclear. The original belief was that they could be toxic since they harbor the toxic RNA transcripts. Yet, accumulating evidence from DM1, DM2 and C9-ALS/FTD studies now suggests that the foci *per se*, do not appear to be toxic, since cells with foci do not necessarily degenerate (Houseley et al., 2005; Yu et al., 2011, 2015; Mizielinska et al., 2014; Tran et al., 2015). The fact that cells with the same amount of foci could very well be alive or dying indicating that toxicity is not proportional to foci numbers (Mizielinska et al., 2014; Tran et al., 2015). Thus, ribonuclear foci could be neutral or even beneficial, at least in the beginning phase of the disease progression.

An alternative mechanism of RNA toxicity is bidirectional transcription. The original discovery for bidirectional transcription came from a study on SCA8. SCA8 is a type of spinocerebellar ataxia arises from a CTG expansion in the 3' end of a non-protein coding RNA, *Ataxin-8 Opposite Strand* (*ATXN8OS*; Koob et al., 1999). Its pathogenic mechanism was thought to be similar to DM1 since both disorders involves a CTG repeat. However, the antisense transcript, *Ataxin-8*, has a short cryptic open reading frame that results in CAG repeats in the glutamine reading frame (Moseley et al., 2006). Moreover, nuclear polyQ inclusions were found, indicating that both toxic RNA and toxic proteins gain-of-function were being produced and could be involved in the pathogenesis of SCA8 (Moseley et al., 2006). Bidirectional transcription was found in models of DM1, FXTAS and SCA7, suggesting that it is a common mechanism of RNA toxicity in READs (Cho et al., 2005; Ladd et al., 2007; Sopher et al., 2011).

A third mechanism of RNA toxicity is nucleolar stress and nucleocytoplasmic transport defects. Nucleolar stress is defined as the cellular pathways used by the nucleolus to communicate with cytosolic compartments in order to trigger apoptosis (Boulon et al., 2010; Lindenboim et al., 2011). Hence, nucleolar stress is a conserved and effective mechanism to eliminate ribosome biogenesis-defective cells. It has been reported that CAG repeat expansion in polyglutamine diseases interacts directly with nucleolin (NCL), a nucleolar protein that regulates rRNA transcription and ribosome biogenesis (Tsoi et al., 2012). The CAG repeat-induced impairment of ribosome biogenesis in turn increases the concentration of p53 in mitochondria, disrupts the interaction between anti- and pro-apoptotic factors such as Bcl-xl and Bak, causes cytochrome c release and caspase activation, and eventually leads to cell death (Tsoi et al., 2012). In polyQ diseases, the fact that CAG repeats-containing RNAs are retained in the nucleus indicates defective nucleocytoplasmic transport of the mutant transcripts (Tsoi et al., 2011). It was reported that CAG repeats can directly bind to the U2 small nuclear ribonucleoprotein auxiliary factor 65 (U2AF65), the fly homolog of human U2AF50 that is involved in the nuclear export of mRNA. U2AF65 is capable of forming a complex with the nuclear export receptor NXF1 to regulate nucleocytoplasmic transport of RNAs (Tsoi et al., 2011). Thus, the downregulation

of U2AF65 due to sequestration by CAG repeats results in the nuclear accumulation of RNAs (Tsoi et al., 2011). Furthermore, the researchers observed a neurodevelopmental downregulation of U2AF65 protein levels in a mouse model of HD, suggesting that the reduction in U2AF65 levels causes nuclear accumulation of expanded CAG-containing transcripts in HD. Apart from polyQ diseases, nucleocytoplasmic transport defects have also been demonstrated in models of C9-ALS/FTD (Freibaum et al., 2015; Zhang et al., 2015). These studies demonstrate the involvement of the nucleolar stress and nucleocytoplasmic transport defects in the RNA toxicity-mediated pathogenesis of READs.

A forth mechanism of RNA toxicity is mediated by hairpin-forming single stranded RNAs (ssRNAs) and double-stranded RNAs (dsRNAs) containing the expanded repeats. Both of them would eventually lead to the generating of toxic repeat-containing small RNAs (sRNAs), which is the true culprit of this mechanism of RNA toxicity in READs. It is known that hairpin structures can be formed by single stranded CAG repeats in polyQ diseases and CGG repeats in FXTAS (Sobczak et al., 2003, 2010). Complex secondary structures like hairpin loops can significantly alter the transcripts' processing, transport, translation, and interactions with RNA binding proteins. Hairpin structures also strongly protects the RNAs from attacks by nucleases, rendering them more resistant to degradation (Sobczak et al., 2010). Hence, the hairpin structure itself already conveys a type of toxicity. However, the more severe consequences come from the fact that it facilitates the production of sRNAs via the RNAi machinery. Dicer is an RNase III family ribonuclease that cleaves double-stranded precursor RNAs to generate sRNAs, and is a key enzyme involved in RNAi (Bernstein et al., 2001). Since expanded ssRNAs in READs form double-stranded hairpin structures, they are valid substrates of Dicer, and will be cleaved into sRNAs. Even though these repeat-containing sRNAs cannot form ribonuclear foci, they are still capable of binding to key proteins of the alternative splicing machinery, such as MBNL1, and cause defects in RNA processing (Mykowska et al., 2011). sRNAs with as few as seven CAG repeats have been shown to induce caspase activation and cellular toxicity in a disease model of HD (Bañez-Coronel et al., 2012). It has been reported that the knockdown of Dicer expression can reduce sRNA in HD and SCA1 models, supporting the involvement of RNAi in the sRNA-induced toxicity of READs (Krol et al., 2007; Banez-Coronel et al., 2012). Since bidirectional transcription is common with repeat expansions, the transcribed sense and antisense strands may form double-stranded RNAs (dsRNAs). Once again, dsRNA is a substrate of Dicer, thus it becomes a source of sRNAs. This mechanism was demonstrated in two *Drosophila* studies. One showed that that dsRNAs comprising CAG/CUG repeats can be a source of cellular toxicity via the Dicer pathway (Lawlor et al., 2011). The other study showed that the co-expression of CAG and CUG repeats results in enhanced RNA toxicity in a Dicer-dependent manner by generating more sRNAs, which impedes the expression of CAG-containing genes, such as *Ataxin-2* and *TATA-binding protein (TBP)* (Yu et al., 2011). Therefore, both ssRNA hairpins and dsRNAs can be a source of harmful sRNAs that results in cellular toxicity. For

hexanucleotide repeat expansions such as GGGGCC in C9-ALS/FTD, the single stranded RNA repeats can even form G-quadruplex structures (Reddy et al., 2013; Haeusler et al., 2014). G-quadruplex structures can induce a number of detrimental consequences in the cell including transcription impairment, formation of inclusions and sequestration of ribonucleoproteins to impeded alternative splicing and induce nucleolar stress (Haeusler et al., 2014; Conlon et al., 2016).

A fifth mechanism which non-coding expanded transcripts lead to cellular toxicity is by repeat-associated non-ATG (RAN) translation. Normally, protein synthesis is initiated by the start codon ATG, which leads to the translation of mRNA. Ranum and colleagues have discovered a novel mechanism which repeat expansions in READs can bypass the need of a start codon and mediate RAN translation of the repeat-containing transcripts (Zu et al., 2011). In RAN translation, the accumulated expanded repeat RNA may undergo protein translation in all three possible reading frames, resulting in the production of proteins consisting of different amino acid combinations (Zu et al., 2011). It has been shown in SCA8, FXTAS, and C9-ALS/FTD that the protein products from RAN translation are highly toxic and play an significant role in pathogenesis (Zu et al., 2011; Todd et al., 2013; Mizielinska et al., 2014; Freibaum et al., 2015; Oh et al., 2015; Tran et al., 2015; Lee et al., 2016).

A sixth mechanism of RNA toxicity involves the expanded RNAs being actively transported to neurites by transport granules, which in turn induces a neuritic branching defect in neurons (Burguete et al., 2015). This could be explained by the fact that the expanded RNA-bound transport granules cannot function normally, and the transport defects possibly interfere with the local translational machinery as well (Burguete et al., 2015).

MODELING THE RNA TOXICITY OF REPEAT EXPANSIONS IN *DROSOPHILA*

Recent work in *Drosophila* has begun to shed light on the emerging “RNA toxicity model.” One of the earliest *Drosophila* studies that demonstrated this model was pioneered by Rebay and colleagues in the pathogenesis of SCA8 (Mutsuddi et al., 2004). Unlike most other types of SCA, SCA8 is associated with a CTG repeat expansion within a non-coding region (Koob et al., 1999). A *Drosophila* model of SCA8 was developed by overexpressing the non-coding transcript of the human SCA8 locus in the fly eye, which results in retinal degeneration (Mutsuddi et al., 2004). The *Drosophila* RNA binding protein genes, *staufen*, *muscleblind (Mbl)*, *split ends*, and *CG3249*, were all subsequently identified as genetic modifiers of this degeneration, suggesting that RNA rather than expanded protein is the source of toxicity (Mutsuddi et al., 2004). DM1 is another non-coding READ demonstrated to be caused by RNA toxicity. DM1 is associated with a CUG repeat expansion in the 3' UTR of the *DMPK* gene. A *Drosophila* model of DM1 was established by expressing non-coding CUG repeats in either the *Drosophila* eyes or muscles. Degeneration was observed in both cases, indicating the presence of toxic RNA gain-of-function (de Haro et al., 2006).

In the case of FXTAS, a CGG repeat expansion occurs in the 5' UTR of the *FMR1* gene. In *Drosophila*, it was reported that the expression of non-coding CGG repeats was sufficient to induce retinal degeneration in the absence of expressed proteins (Jin et al., 2003). All these studies in *Drosophila* support the notion that RNA toxicity play a significant role in the pathologies of many READs. More importantly, these studies also demonstrate the fruit fly as an excellent model for studying human READs and RNA toxicity. Currently, fly models have been established for at least 13 different READs (Table 1). Below, we will cover a number of READs with RNA toxicity and highlight the discoveries made using their corresponding *Drosophila* models. These READs were selected base on the following two criteria: (i) It has been shown that gain-of-function RNA toxicity contributes to the pathology of these diseases, and (ii) there are established *Drosophila* models for studying the RNA toxicity of these diseases.

Drosophila Models of Polyglutamine Diseases

Polyglutamine diseases are caused by the expansion of unstable CAG repeats in the protein-coding region of the respective disease gene. The CAG repeats are translated into a string of glutamine in normal functioning proteins. Patients of polyQ diseases often develop progressive neurodegeneration and motor impairment (Rub et al., 2013). PolyQ diseases include HD, spinobulbar muscular atrophy (SBMA; a.k.a. Kennedy disease), dentatorubral-pallidoluyian atrophy (DRPLA), spinocerebellar ataxia (SCA) types 1, 2, 3 (a.k.a. Machado-Joseph disease or MJD), 6, 7, and 17 (La Spada et al., 1991; MacDonald et al., 1993; Orr et al., 1993; Kawaguchi et al., 1994; Imbert et al., 1996; Pulst et al., 1996; Sanpei et al., 1996; David et al., 1997; Zhuchenko et al., 1997; Nakamura et al., 2001; Warrick et al., 2005). HD is caused by an abnormal expansion of CAG repeats in exon 1 of the *Huntingtin (HTT)* gene (MacDonald et al., 1993). For SBMA, the CAG repeat expansion lies in the N-terminus of the coding region of the androgen receptor gene (La Spada et al., 1991; Zhou et al., 1994). For DRPLA, the unstable CAG expansion occurs in the middle of the Atrophin-1 gene (Koide et al., 1994). As for SCA1, 2, 3, 6, 7, and 17, the CAG expansion are located in the coding regions of *Ataxin-1*, *Ataxin-2*, *Ataxin-3*, *CACNA1A*, *Ataxin-7*, and *TATA-binding protein (TBP)*, respectively (Nakamura et al., 2001; Rub et al., 2013). Despite the fact that the CAG repeat expansions in these polyQ diseases affect different genes, there are several shared features among the majority of polyQ diseases. First, a common hallmark is the abnormal aggregation of nuclear inclusions containing the mutant proteins. These inclusions that contain abnormal aggregated proteins were once regarded as a potential cause of cellular toxicity. However, a number of studies have demonstrated that the formation of inclusions could actually be a cellular protective response to mitigate the toxicity (Stefanis et al., 2001; Petrucelli et al., 2002; Arrasate et al., 2004). Until today, the general belief is that the formation of aggregates is a dynamic process. Smaller oligomeric inclusions are likely to toxic, whereas the large aggregates might actually be protective (Ross and Poirier, 2005). Further investigations will be required to elucidate their true nature. Second, the CAG repeat length is inversely

TABLE 1 | Published transgenic *Drosophila* lines for modeling repeat expansion-associated diseases.

READ	Affected human gene	Expansion region of the human gene	Nucleotide repeat	Established <i>Drosophila</i> lines for modeling READs (under UAS control)
HD	<i>Huntingtin (HTT)</i>	Coding region (Exon 1)	CAG	<p>With an exon1-derived HTT fragment (Jackson et al., 1998): HTT^{1-142-Q2} HTT^{1-142-Q75} HTT^{1-142-Q120}</p> <p>(Steffan et al., 2001): Htt-ex1p-Q20 Htt-ex1p-Q93 UAS-Q48 + myc/flag</p> <p>(Doumanis et al., 2009): Nhtt-18Q-EGFP Nhtt-48Q-EGFP Nhtt-98Q-EGFP Nhtt-152Q-EGFP Nhtt-48Q-EGFP^{NLS}</p> <p>With a HTT fragment of amino acids 1-336 (Kaltenbach et al., 2007): 128Q-htt^{M64}</p> <p>With a HTT fragment of amino acids 1-548 (Lee et al., 2004): Htt-Q0 Htt-Q128</p> <p>With a HTT fragment of amino acids 1-588 (Weiss et al., 2012): mRFP-Htt-Q15 mRFP-Htt-Q138 EGFP-Htt-Q15 EGFP-Htt-Q138 EGFP-Htt-Q15-mRFP EGFP-Htt-Q138-mRFP Htt^{1-81-Q96-GFP} (with HTT exon 1 amino acids 1-81)</p> <p>With full length HTT (3144 amino acids) (Romero et al., 2008): 16Q-htt^{FL} 128Q-htt^{FL}</p>
SBMA	<i>Androgen receptor (AR)</i>	Coding region (N-terminus)	CAG	(Takeyama et al., 2002): AR-Q0 AR-Q52
DRPLA	<i>Atrophin-1 (AT1)</i>	Coding region	CAG	(Nisoli et al., 2010): At-1-wt At-1-65Q At-1-65ΔC Atro-wt Atro-75QN Atro-66QC
SCA1	<i>Ataxin-1 (ATXN1)</i>	Coding region	CAG	(Fernandez-Funez et al., 2000): Ataxin-1 Q30 Ataxin-1 Q82
SCA3 (MJD)	<i>Ataxin-3 (ATXN3)</i>	Coding region	CAG	(Warrick et al., 1998): SCA3tr-Q27 SCA3tr-Q78 <p>(Warrick et al., 2005): SCA3-Q27 SCA3-Q78 SCA3-Q84 SCA3-Q27-UIM* SCA3-Q80-UIM*</p>

(Continued)

TABLE 1 | Continued

READ	Affected human gene	Expansion region of the human gene	Nucleotide Repeat	Established <i>Drosophila</i> lines for modeling READs (under UAS control)
				SCA3-Q27-C14A SCA3-Q88-C14A SCA3-delta (Li et al., 2008): SCA3tr-Q61 SCA3tr-Q78CAA/G DsRed-CAG0 DsRed-CAG100 DsRed-CAG250
SCA6	<i>CACNA1A</i>	Coding region (exon 47)	CAG	(Tsou et al., 2015): α 1ACT-Q11 α 1ACT-Q70 (Tsou et al., 2016): α 1ACT-Q33
SCA7	<i>Ataxin-7 (ATXN7)</i>	Coding region	CAG	(Jackson et al., 2005): SCA7-Q90 (Latouche et al., 2007): ATXN7T-Q10 ATXN7T-Q102
SCA8	<i>ATXN8OS (non-protein coding Ataxin-8 opposite strand)</i>	3'UTR	CTG	(Mutsuddi et al., 2004): SCA8(CTG) ₉ SCA8(CTG) ₁₁₂
SCA17	<i>TATA binding protein (TBP)</i>	Coding region	CAG	(Ren et al., 2011): TBP-Q34 TBP-Q54 TBP-Q80
FXTAS	<i>Fragile X mental retardation 1 (FMR1)</i>	5'UTR	CGG	(Jin et al., 2003): (CGG) ₆₀ -GFP (CGG) ₉₀ -GFP (Todd et al., 2013): 5'UTR: (CGG) ₁₀₀ -(+1)-ATG-GFP 5'ATG: ATG-(CGG) ₁₀₀ -(+1)-ATG-GFP 3'UTR: ATG-GFP-STOP-(CGG) ₁₀₀ -(+1) 5' Stop: STOP-(CGG) ₁₀₀ -(+1)-ATG-GFP
DM1	<i>Dystrophia myotonica protein kinase (DMPK)</i>	3'UTR	CTG	(Houseley et al., 2005): GFP-(CTG) ₁₁ GFP-(CTG) ₄₈ GFP-(CTG) ₅₆ GFP-(CTG) ₁₆₂ (de Haro et al., 2006): i(CTG) ₄₈₀ (Garcia-Lopez et al., 2008): (CTG) ₆₀ i(CTG) ₄₈₀ (Yu et al., 2011): (CTG) ₁₉ (CTG) ₁₃₀ (CTG) ₂₀₀ (CTG) ₂₃₀ (CTG) ₂₅₀ (CTG) ₂₇₀

(Continued)

TABLE 1 | Continued

READ	Affected human gene	Expansion region of the human gene	Nucleotide Repeat	Established <i>Drosophila</i> lines for modeling READs (under UAS control)
DM2	<i>Zinc finger protein 9 (ZNF9)</i>	Intron 1	CCTG	(Picchio et al., 2013): i(CTG) ₂₄₀ i(CTG) ₄₈₀ i(CTG) ₆₀₀ i(CTG) ₉₆₀ ("i" indicates "interrupted repeats"). (Yu et al., 2015): (CCTG) ₁₆ (CCTG) ₂₀₀ (CCTG) ₄₇₅ (CCTG) ₅₂₅ (CCTG) ₇₀₀ (CCTG) ₇₂₀
C9ORF72-associated ALS-FTD	<i>C9ORF72</i>	Promoter and intron 1	GGGGCC	(Xu et al., 2013): (GGGGCC) ₃ (GGGGCC) ₃₀ (Mizielinska et al., 2014): (GGGGCC) ₃ Pure (GGGGCC) ₃₆ Pure (GGGGCC) ₃₆ RO (GGGGCC) ₁₀₈ RO (GGGGCC) ₂₈₈ RO (PA) ₃₆ (GA) ₃₆ (PR) ₃₆ (GR) ₃₆ (Freibaum et al., 2015): (GGGGCC) ₈ (GGGGCC) ₂₈ (GGGGCC) ₅₈ GFP-(GP) ₄₇ GFP-(GA) ₅₀ GFP-(GR) ₅₀ (Iran et al., 2015): (GGGGCC) _{5R} (flanked by human intronic and exonic sequences) (GGGGCC) _{80R} (flanked by human intronic and exonic sequences) (GGGGCC) _{160R} (flanked by human intronic and exonic sequences)

UIM* indicates that the highly conserved serine residues in the ubiquitin interaction motifs are mutated to alanine to disrupt ubiquitin binding.

proportional to the age of disease onset. Longer CAG lengths cause earlier onset of polyQ diseases (Zoghbi and Orr, 2000; Ross and Poirier, 2005). This is because once the polyQ repeat length exceeds a pathogenic threshold (typically ~35), the mutant proteins are likely to mis-fold into abnormal conformations that can resist normal cellular degradation machineries such as the ubiquitin-proteasome system (Sisodia, 1998; Chan et al., 2000; Ross and Poirier, 2005).

PolyQ diseases were among the first neurodegenerative diseases modeled in *Drosophila*. The earliest *Drosophila* models include HD and SCA3. Zipursky and colleagues fused a truncated form of wild-type human HTT with a polyQ tract of 2, 75, or 120 repeats (HTT¹⁻¹⁴²-Q2, HTT¹⁻¹⁴²-Q75 and HTT¹⁻¹⁴²-Q120; Jackson et al., 1998), whereas Bonini and colleagues generated truncated forms of human Ataxin-3 with

a polyQ tract of 27 or 78 repeats (SCA3tr-Q27 and SCA3tr-Q78) (Warrick et al., 1998). When expressed in the fly eye, both studies found that only the constructs with pathogenic lengths (HTT¹⁻¹⁴²-Q75, HTT¹⁻¹⁴²-Q120, and SCA3tr-Q78) induced nuclear inclusions and degeneration, recapitulating the pathogenic features of polyQ diseases. These studies, for the first time, demonstrated that *D. melanogaster* can be a powerful model to study human neurodegenerative diseases. Subsequently, more *Drosophila* models of HD and SCA3 (Steffan et al., 2001; Lee et al., 2004; Warrick et al., 2005; Kaltenbach et al., 2007; Li et al., 2008; Romero et al., 2008; Doumanis et al., 2009; Weiss et al., 2012), as well as other polyQ diseases, including SCA1 (Fernandez-Funez et al., 2000), SCA6 (Tsou et al., 2015, 2016), SCA7 (Jackson et al., 2005; Latouche et al., 2007), SCA17 (Ren et al., 2011), SBMA (Takeyama et al., 2002; Pandey et al., 2007; Nedelsky et al., 2010),

and DRPLA (Nisoli et al., 2010; Napoletano et al., 2011) were developed. All these studies demonstrate that polyQ diseases, in most cases, can yield parallel neural degenerative effects in the fruit fly.

Besides protein toxicity, there is now substantial evidence for RNA toxicity in polyQ diseases as well. One of the most compelling studies was performed by Bonini and colleagues in the studying of SCA3 in *Drosophila* (Li et al., 2008), which they showed that the expression of both translated and untranslated CAG repeats caused neural degeneration, whereas the expression of CAA-interrupted CAG repeats resulted in significantly less pronounced neurodegenerative effects (Li et al., 2008). Similar to CAG repeats, CAA repeats also encode for glutamine. Yet CAA repeats are unable to form a hairpin structure (Sobczak et al., 2010). As the hairpin structure contributes to CAG repeat-containing sRNA generation, it is one of the major mechanisms of RNA toxicity (Krol et al., 2007; Bañez-Coronel et al., 2012). Therefore, the above findings suggest that RNA toxicity contributes significantly to neurodegeneration in SCA3.

Apart from hairpin structures, it is also known that CAG repeats in polyQ diseases, similar to CUG repeats in DM1, can form ribonuclear foci that colocalize with RNA-binding proteins such as muscleblind-like 1 (MBNL1), and cause misregulation in alternative splicing (Ho et al., 2004, 2005; Yuan et al., 2007; Hsu et al., 2011; Mykowska et al., 2011; Wang et al., 2011). In fact, Krzyzosiak and colleagues demonstrated that expanded CAG repeats present in HTT and ATXN3 transcripts colocalize with MBNL1 in ribonuclear foci, further confirming the accumulation of mutant RNA in the nucleus (Mykowska et al., 2011). Interestingly, although both CAG and CUG repeats form foci, sequester MBNL1 and affect alternative splicing, the splicing of some pre-mRNAs are only affected by CUG repeats but not CAG repeats (Ho et al., 2005). Hence, despite the overlapping mechanisms of involvement of MBNL1, there are still differences in how splicing is disrupted by CAG repeats in polyQ diseases and CUG repeats in DM1.

In addition to the sequestering of RNA-binding proteins, RNA toxicity in polyQ diseases may also arise from the perturbation of RNA export mechanisms and the induction of nucleolar stress (Tsoi et al., 2011, 2012). In the *Drosophila* model of SCA3, it has been demonstrated that U2AF65, which is involved in the nuclear export of mRNA, can bind directly to CAG repeats and form a complex with the nuclear export receptor NXF1 (Tsoi et al., 2011), and the downregulation of U2AF65 correlates with the nuclear accumulation of expanded CAG RNA (Tsoi et al., 2011). Apart from U2AF65, CAG repeat expansion can also interact directly with NCL, and induce apoptosis by the activation of the nucleolar stress pathway (Tsoi et al., 2012). NCL is an essential protein involved in RNA polymerase I-mediated rRNA transcription in the nucleolus. The binding of expanded CAG repeats perturbs ribosome biogenesis, which in turn increases the levels of mitochondrial p53, and causes cytochrome c release and caspase activation (Tsoi et al., 2012).

Utilizing the established fly models of SCA3, much advances have been made in the search of genetic modifiers of polyQ toxicity and CAG-repeat RNA toxicity. Using the fly model of SCA3, HSP70 was found to colocalize with polyQ inclusion

bodies, and was identified as a strong suppressor of polyQ toxicity (Warrick et al., 1999; Chan et al., 2000). A genome-wide genetic screen was also performed, and identified a number of other chaperones and ubiquitin pathway-related molecules as genetic modifiers of polyQ toxicity (Bilen and Bonini, 2007). As for RNA toxicity in SCA3, a microarray analysis was performed to search for genes which expression levels are affected by CAG repeats (Shieh and Bonini, 2011). Several iron ion binding and nucleotide binding molecules were found to be affected. Interestingly, HSP70, co-chaperone Trp2, the transcriptional regulator Dpld, and the RNA-binding protein Orb2 were able to modify both polyQ and RNA toxicity, suggesting that there are overlapping mechanisms of RNA and protein based toxicity (Shieh and Bonini, 2011).

All the above studies demonstrated that both protein and RNA toxicity contribute to neurodegeneration in polyQ diseases, and both types of toxicity can involve a number of different underlying molecular mechanisms which can be studied using *Drosophila* models.

***Drosophila* Models of Fragile X Syndrome and Fragile X-Associated Tremor/Ataxia Syndrome**

Fragile X syndrome (FXS) is the most common cause of inherited intellectual disability, often associated with autism spectrum disorders (Bhakar et al., 2012; Santoro et al., 2012). The disease was first described as mental retardation segregated in X-linked fashion (Martin and Bell, 1943). The chromosomal variation was later mapped to Xq27.3 and was dubbed “the fragile X chromosome” (Harrison et al., 1983). In FXS, a trinucleotide CGG expansion in the 5' UTR of the *FMR1* gene results in transcriptional silencing and loss of the corresponding protein, Fragile X Mental Retardation Protein (FMRP; Fu et al., 1991; Kremer et al., 1991). Normal alleles contain 6–55 repeats, premutation alleles contain 55–200, and the disease alleles contain over 200 repeats. FMRP is an RNA-binding protein which negatively regulates the translation of mRNAs, especially at synapses in neurons. Loss of FMRP impairs synaptic plasticity, which is believed to be the molecular basis for the intellectual disability in FXS patients (Bassell and Warren, 2008).

The *Drosophila* genome contains a highly-conserved *FMR1* gene, *dfmr1* (Wan et al., 2000). The first *Drosophila* FXS model was generated by imprecise excision of a *P*-element located upstream of the *dfmr1* locus to produce a null mutant (Zhang et al., 2001). Subsequently, a wide array of *dfmr1* alleles and transgenes have been created to facilitate FXS research in this classic genetic model (Dockendorff et al., 2002; Morales et al., 2002; Lee et al., 2003). Overexpression of either the human FMRP (*FMR1* gene) or the native *Drosophila* dFMRP (*dfmr1* gene) fully rescues all detectable disease-like phenotypes in the *Drosophila* FXS model (Coffee et al., 2010), confirming the evolutionary conserved functions of *FMR1* in the nervous system.

In 2003, Warren and colleagues expressed 90 CGG repeats with a reporter in the *Drosophila* compound eye and demonstrated retinal degeneration in the absence of expressed proteins (Jin et al., 2003). These findings suggest

that, apart from loss-of-function of wild-type FMRP, there exists a toxic RNA gain-of-function component associated with the repeat expansion that occurs at the 5' UTR of the *FMR1* gene. This RNA toxicity causes a late-onset neural degeneration phenotype, which is now identified as Fragile X-associated tremor/ataxia syndrome (FXTAS). While FXS is mostly caused by the loss-of-function of wild-type FMRP, which results in neurodevelopmental and behavioral disorders characterized by mental retardation, autism, anxiety and mood instability, FXTAS is mostly caused by the CGG repeat expansion that produces gain-of-function RNA toxicity in the CNS, which results in a late-onset neurodegenerative disorder, characterized by progressive intention tremor, gait ataxia, parkinsonism, and cognitive decline (Jacquemont et al., 2007; Tassone and Hagerman, 2012). FXTAS is found mostly in male premutation carriers after the age of 50 (Tassone and Hagerman, 2012). Although at a much lower rate, female carriers may also develop FXTAS. The neuropathological hallmark of FXTAS is the ubiquitin-positive intranuclear inclusion, present in both neurons and astrocytes throughout the brain (Greco et al., 2002). Warren and colleagues found that in their (CGG)₉₀ fly model of FXTAS, the inclusion bodies in the degenerating neurons contain HSP70 (Jin et al., 2003). Furthermore, when they overexpress HSP70, the CGG repeat-induced degeneration was suppressed (Jin et al., 2003), which is similar to the case of SCA3 where HSP70 also presents suppressor activity of CAG repeat-induced toxicity (Warrick et al., 1999; Chan et al., 2000). Two subsequent studies have utilized this *Drosophila* model of FXTAS and identified a number of tropomyosin and RNA-binding proteins as genetic modifiers of the degeneration phenotype (Jin et al., 2007; Sofola et al., 2007). These findings strongly suggest that the RNA processing machinery play an important role in the pathogenesis of FXTAS. Interestingly, both studies have found that one of these RNA-binding proteins that interacts with the CGG repeats to be CUG-binding protein 1 (CUGBP1), which is a protein also involved in the pathogenesis of another READ, DM1. Indeed, CGG repeats have been shown to form RNA hairpin structures *in vitro*, similar to CAG repeats in SCA3 and CUG repeats in DM1 (Sobczak et al., 2003, 2010). Thus, it is possible that these RNA-binding proteins bind to the CGG repeats in an effort to disassemble these potentially toxic hairpins, but in the process neglected their other cellular functions.

Recent studies in neurodegenerative and neuromuscular diseases have advanced toward the investigation of cellular transcriptional and translational profiles. Nelson and colleagues discovered that mice Purkinje neurons ectopically expressing 90 CGG repeats exhibit a dramatic change in their translational profile even prior to the onset of CGG repeat-induced phenotypes (Galloway et al., 2014). Interestingly, the *Tardbp* gene, which encodes the TDP-43 protein was found to be reduced in Purkinje neurons (Galloway et al., 2014). TDP-43 is a RNA-binding protein commonly found in the inclusion bodies in ALS (Xu and Yang, 2014). In the fly model of FXTAS, overexpression of TDP-43 suppresses neurodegeneration in the fly eye, whereas knockdown of *Drosophila* TDP-43 results in enhancement of the degeneration phenotype (Galloway et al., 2014). Todd and colleagues also discovered that overexpression

of TDP-43 was capable of suppressing CGG repeat toxicity (He et al., 2014). Interestingly, they found that the CGG repeat RNA levels and RAN translation levels are both not affected by TDP-43. Instead, they found TDP-43 alters the composition, behavior and distribution of multimeric RNA-binding protein complexes, preventing the sequestration of hnRNP A2/B1, and rescuing the alternative splicing machinery (He et al., 2014). They also tested an ALS mutant form of TDP-43, which can still bind to hnRNP A2/B1, and yet it failed to rescue the CGG repeat toxicity, suggesting that wild type TDP-43 functions are necessary for this suppression of CGG repeat toxicity (He et al., 2014).

In the study above by Nelson and colleagues, the researchers discovered ~500 transcripts that are differentially associated with ribosomes in the CGG repeats-expressing Purkinje neurons (Galloway et al., 2014). Theoretically, if ribosomes are recruited to the repeat expansions, translation of the repeats would be facilitated. Indeed, RAN translation has also been demonstrated in FXTAS models. Paulson and colleagues have demonstrated that CGG repeats lead to RAN translation of a cryptic polyglycine-containing protein, FMRpolyG, which is present in inclusion bodies in FXTAS fly models, mouse models and FXTAS patient brains (Todd et al., 2013). The researchers constructed a *Drosophila* line expressing CGG repeats preceded by a stop codon just 5' to the repeats. Production of FMRpolyG is inhibited by this stop codon, but the CGG repeat-containing RNA remained. This line gave rise to a much milder phenotype comparing with CGG repeats without the stop codon, indicating RAN translated FMRpolyG contribute to toxicity (Todd et al., 2013). A follow-up *Drosophila* study revealed that FMRpolyG impairs the ubiquitin protease system (Oh et al., 2015). Inhibition of RAN translation can reduce inclusion formation and attenuate the impairment of the ubiquitin protease system (Oh et al., 2015).

Taken together, these studies suggest that FXTAS pathogenesis may be mechanistically related to the polyQ diseases, myotonic dystrophy, and ALS. Thus, different READs that arise from different causes may converge and share some common molecular mechanisms in their pathogenesis.

***Drosophila* Models of Myotonic Dystrophy Type 1 and Type 2**

Myotonic dystrophy is the most common form of muscular dystrophy affecting about 1 in 8,000 adults (Harper, 2001). It is a multisystemic, autosomal dominant neuromuscular disorder, characterized by progressive myotonia (inability of muscles to relax after contraction), muscle degeneration, iridescent cataracts, cardiac arrhythmias, testicular atrophy, and symptoms of neuropathology (Harper, 2001). DM1 is caused by a CTG repeat expansion in the 3' UTR of the *DMPK* gene, which encodes a serine-threonine protein kinase (Brook et al., 1992). Normal individuals typically have fewer than 37 CTG repeats, whereas DM1 patients can range from 50 to 4,000 repeats (Thornton et al., 1994). The mutant transcripts containing CUG repeats fold into RNA hairpins that are retained in the nucleus (Davis et al., 1997; Koch and Leffert, 1998), whereas the levels of DMPK protein are correspondingly reduced (Fu et al., 1993). *Dmpk* knockout mice have reduced force generation in skeletal muscles (Reddy et al., 1996) and abnormal cardiac conduction

(Berul et al., 1999), which suggests that the loss of DMPK function may contribute to at least the muscle weakness and cardiac defects in DM1. However, *Dmpk* knockout mice do not display myotonia and muscle degeneration (Reddy et al., 1996), which are some of the hallmark features of DM1. In contrast, transgenic mice expressing expanded CTG repeats display myotonia, degeneration, and several other muscular pathological abnormalities (Mankodi et al., 2000). These results demonstrated that the expanded CTG repeats by themselves could account for at least some of the major pathological features of DM1 independent of the *DMPK* locus. It was eventually demonstrated that the CUG-repeat-containing transcripts retained in the nucleus aggregate into ribonuclear foci (Taneja et al., 1995; Mankodi et al., 2000). These foci binds to RNA-binding proteins like MBNL1 and CUGBP1 (Miller et al., 2000; Kanadia et al., 2003), which in turn compromises the RNA splicing machinery (Philips et al., 1998; Ho et al., 2004) and even disrupts the regulation of transcription (Ebrilidze et al., 2004).

Monckton and colleagues developed a *Drosophila* model of DM1 by fusing 162 non-coding CTG repeats to a reporter gene and overexpressing the repeats in fly muscles (Houseley et al., 2005). Although 162 CUG repeats results in the formation of ribonuclear foci, it did not result in detectable cellular toxicity. Botas and colleagues further examined the effect of non-coding CUG repeats by increasing the repeat number to 480 (interrupted repeats). With this amount of repeats, vacuolization and degeneration were then detected in adult fly muscles (de Haro et al., 2006). Expression of 480 CUG repeats in the adult eye also results in disorganization of ommatidia and reduction of the eye size, suggesting gain-of-function RNA toxicity (de Haro et al., 2006; Garcia-Lopez et al., 2008). Similar to mammals, the molecular mechanisms of this CUG-induced toxicity involve the sequestration of *Drosophila* Mbl and CUGBP1 by the CUG repeat-containing ribonuclear foci, resulting in defective alternative splicing (de Haro et al., 2006; Garcia-Lopez et al., 2008). However, in contrast to CUGBP1, which has been shown to interact with CGG repeats and is involved in FXTAS, no evidence for the genetic interaction of muscleblind family proteins and CGG repeats had been found in *Drosophila* (Sofola et al., 2007). Thus, although the pathogenesis of FXTAS and myotonic dystrophy may share some commonalities, there are certainly distinct mechanisms for the two READs as well. Utilizing this fly model of DM1, Botas and colleagues identified SMAUG, a myosin-binding protein, as suppressor of CUG repeat-induced toxicity (de Haro et al., 2013). SMAUG interacts genetically and physically with CUGBP1, and its overexpression restores the translational activity of CUGBP1 in this DM1 model (de Haro et al., 2013).

Refinements of the DM1 fly models were made in some recent studies, with new discoveries being yield. Jagla and colleagues constructed fly lines with 240, 480, 600, and 960 interrupted CTG repeats, and established a novel *Drosophila* larval model of DM1 (Picchio et al., 2013). Using this model, they constructed transcription profiles of DM1, and uncovers a number of splice-independent deregulated genes (Picchio et al., 2013). Bonini and colleagues cloned 19, 130, 200, 230, 250, and 270 uninterrupted, pure CTG repeats into the 3' UTR of the *DsRed* gene to

better recapitulate the pathological condition of the disease (Yu et al., 2011). Repeat length-dependent toxicity was detected when the repeat length reaches 200 or more. Interestingly, the researchers discovered that there was a synergistic effect on the enhancement of toxicity when CAG repeats are co-expressed together with CUG repeats due to the biogenesis of sRNAs (Yu et al., 2011). This enhancement was Dicer-dependent, indicating the involvement of the RNAi machinery (Yu et al., 2011).

DM2 is caused by a tetranucleotide CCTG repeat expansion in the first intron of the *zinc finger protein 9 (ZNF9)* gene, which encodes a RNA-binding protein (Liquori et al., 2001). Affected individuals can have 75 to 11,000 CCTG repeats, with a mean of ~5,000 (Liquori et al., 2001). DM2, much like DM1, involves the aggregation of expanded mutant transcripts into ribonuclear foci, which interfere with key molecules of the splicing machinery, such as MBNL1 and CUGBP1 (Liquori et al., 2001). Although DM1 and DM2 share much similarities in their pathologies, DM2 has less-pronounced neurological involvement and cannot be congenital (Minnerop et al., 2011). To develop a fly model of DM2, Bonini and colleagues generated constructs with 16, 200, 475, 525, 700, and 720 non-coding CCTG repeats to investigate length-dependent RNA toxicity (Yu et al., 2015). Expression of 200 repeats was sufficient to induce the formation of ribonuclear foci, but impairment of alternative splicing and retinal degeneration were only observed when the repeat length was 475 or above. In addition, the researchers demonstrate that the expression of two different isoforms of MBNL1 (with or without the linker region between zinc finger domain 2 and 3) can lead to cleavage of non-coding CUG or CCUG RNA repeats, and mildly rescue CCUG-induced degeneration in the fly eye (Yu et al., 2015).

***Drosophila* Models of C9ORF72-Associated Amyotrophic Lateral Sclerosis/Frontotemporal Dementia**

Amyotrophic lateral sclerosis (ALS), also known as Lou Gehrig's disease, is an adult-onset neurodegenerative disease characterized by the progressive loss of motor neurons in the cortex, brainstem, and spinal cord. Symptoms include the loss of movement control, difficulty in speaking, swallowing, and eventually breathing, which leads to fatality (Kurtzke, 1982). ALS affects about 2 per 100,000 individuals in Europe and the United States (Kiernan et al., 2011). Most cases of ALS are deemed sporadic, but 5–10% cases are inherited (Hand and Rouleau, 2002). Mutations from more than 20 different genes have been linked to familial ALS (Chen et al., 2013; Ajroud-Driss and Siddique, 2015). Despite some of these identified genes play a role in RNA metabolism and axonal trafficking (Strong, 2010; Ikenaka et al., 2012; Millecamps and Julien, 2013; Droppelmann et al., 2014), most of them apparently do not function in the same molecular pathway, complicating the underlying mechanisms of the disease (Chen et al., 2013; Ajroud-Driss and Siddique, 2015).

In 2011, the expansion of a GGGGCC hexanucleotide repeat in the non-coding region of *C9ORF72* was found to be associated

with ALS combined with frontotemporal dementia (DeJesus-Hernandez et al., 2011; Renton et al., 2011). Frontotemporal dementia is a group of disorders caused by the progressive loss of neurons predominantly at the frontal and/or temporal lobes of the brain resulting in cognitive and behavioral impairment, and is one of the most common causes of dementia second only to Alzheimer's disease (Mercy et al., 2008; Perry and Miller, 2013; Baizabal-Carvallo and Jankovic, 2016). Remarkably, over 6% of sporadic ALS/FTD cases and over 37% familial ALS/FTD cases are caused by the expansion of the GGGGCC hexanucleotide repeat in the first intron of the *C9ORF72* gene (Majounie et al., 2012).

Jin and colleagues pioneered the first *Drosophila* model of C9-ALS/FTD by expressing 30 GGGGCC repeats [(GGGGCC)₃₀] with a reporter in the compound eye, which resulted in severe retinal degeneration (Xu et al., 2013). When (GGGGCC)₃₀ was expressed in motor neurons, it resulted in defective locomotor activity; whereas when it was expressed in the central nervous system using a pan-neuronal driver, it resulted in lethality (Xu et al., 2013). One important question to be answered was how GGGGCC repeats actually cause the disease. The pathogenesis could be due to: (1) loss-of-function of wild-type C9ORF72 protein, (2) gain-of-function of toxic RNA consisting of the GGGGCC repeats, and/or (3) gain-of-function of toxic dipeptide repeat (DPR) proteins generated by RAN translation. Although the mild phenotypes in a homozygous *C9ORF72* mutation case and the lack of *C9ORF72* coding mutations suggest that wild-type protein loss-of-function is unlikely the primary cause of C9-ALS/FTD (Fratta et al., 2013; Harms et al., 2013), loss-of-function of the wild type protein is still a possible mechanism that contribute to the disease pathogenesis. Since *Drosophila* does not have an ortholog of C9ORF72, the fly may not be a suitable model organism for investigating the haploinsufficiency aspect of C9ORF72. The fly, however, proved a valuable tool in determining the gain-of-function toxicity of GGGGCC repeats in C9-ALS/FTD. In the attempt to differentiate between repeat RNA and DPR protein toxicity, Isaacs and colleagues generated stop codon-interrupted RNA-only (RO) repeats of GGGGCC and compared it with pure GGGGCC repeats which can be RAN-translated into DPR proteins in *Drosophila* (Mizielinska et al., 2014). Isaacs and colleagues found that, the expression of RO (GGGGCC)₃₆ or pure (GGGGCC)₃₆ results in the same amount of ribonuclear foci formation. However, the expression of all RO constructs, including RO (GGGGCC)₃₆, RO (GGGGCC)₁₀₈, and even RO (GGGGCC)₂₈₈, had no detectable effect on the fly eye. In contrast, both pure (GGGGCC)₃₆ and pure (GGGGCC)₁₀₃ caused degeneration. These findings suggest that, although both RNA and DPR proteins may contribute to toxicity, degeneration was primarily attributable to DPR proteins (Mizielinska et al., 2014). To further investigate whether DPR protein expression alone was sufficient to cause degeneration, they generated four "protein-only" constructs by using alternative codons. When expressed in the fly eye, only glycine-arginine (GR)₃₆ and proline-arginine (PR)₃₆ constructs resulted in degeneration, whereas glycine-alanine (GA)₃₆ and proline-alanine (PA)₃₆ had no detectable effect. Even at the repeat length of 100, same results were obtained for the above DPR proteins, indicating

that high arginine-containing DPR protein is the major cause of neurodegeneration in C9-ALS/FTD (Mizielinska et al., 2014).

To determine the molecular mechanism that underlies DPR proteins-induced degeneration, Taylor, Gao and colleagues generated reporter-tagged (GGGGCC)₅₈ flies, and performed a screen to look for genetic modifiers of GGGGCC-induced eye degeneration (Freibaum et al., 2015). Eighteen modifier genes were identified with the majority of them being transcription/export (TREX) complex proteins and nuclear pore complex proteins, involving particularly nuclear export of RNA. Lastly, Taylor, Gao and colleagues predicted that (GGGGCC)₅₈ can produce poly(GA), poly(GR), or poly(GP) DPR proteins by RAN translation. Consistent with an aforementioned study, they found that the arginine-containing poly(GR) to be highly toxic (Freibaum et al., 2015). In summary, the researchers demonstrated that GGGGCC repeat expansion compromises nucleocytoplasmic transport, facilitating nuclear retention of toxic RNAs. Whereas the build-up of toxic poly(GR) DPR proteins by RAN translation is a major cause of GGGGCC-induced neurodegeneration. Rothstein and colleagues also conducted a screen to search for interactors for GGGGCC RNA, and uncovered RanGAP as a potent suppressor of neurodegeneration (Zhang et al., 2015). RanGAP is a key regulator of nucleocytoplasmic transport, supporting that the nucleocytoplasmic transport mechanisms are involved in the neurodegeneration in C9-ALS/FTD (Zhang et al., 2015). In the same year, another screen performed in yeast by Gitler and colleagues identified effectors of Ran-mediated nucleocytoplasmic transport and karyopherins as modifiers of GGGGCC-induced DPR protein toxicity (Jovicic et al., 2015). A recent study by Van Den Bosch and colleagues also reported the results of a *Drosophila* genetic screen, which identified importins, exportins, Ran-GTP cycle regulators, nuclear pore components and arginine methylases as potent modifiers of DPR protein toxicity (Boeynaems et al., 2016). Taking a different approach, Taylor and colleagues analyzed the interactome of all DPR proteins in human cells, and found that arginine-containing DPR proteins (GR and PR) interact with RNA-binding proteins with low complexity sequence domains (LCDs), such as TDP-43, FUS, hnRNPA1, hnRNPA2B1, Matrin-3, and Ataxin-2 (Lee et al., 2016). Using the *Drosophila* eye, they confirmed that these proteins are genetic modifiers of DPR protein-induced toxicity. Remarkably, most of these genetic modifiers encode components of membrane-less organelles, including the nucleoli, the nuclear pore complexes and stress granules. The researchers showed that GR and PR bind to the LCD-containing RNA-binding proteins, alter their biophysical properties, and lead to impairment of their assembly, dynamics, and functions (Lee et al., 2016).

A consensus from all these studies was that the DPR proteins that cause toxicity in C9-ALS/FTD contain arginine. Arginine is a unique amino acid that has as many as five potential hydrogen-bond donors that are readily available to react with biological hydrogen-bond acceptors. In DNA-protein complexes, arginine residues are the most frequent hydrogen bond donors to backbone phosphate groups (Luscombe et al., 2001). Networks of hydrogen bonds can be formed by arginine residues with phosphate groups of RNA loops as well (Calnan

et al., 1991). Whereas among the interactions between amino acids, the arginine-aspartate double hydrogen-bond interaction is also an extremely stable structure (Mitchell et al., 1992). Each methylation of an arginine residue removes a potential hydrogen-bond donor, and changes the shape of the amino acid. Thus, methylation can reduce the chance of arginine bonding to its interactors, which in turn increases the chance of an arginine-containing DPR protein to be degraded. This is supported by the finding that arginine methylases are potent modifiers of DPR protein toxicity (Boeynaems et al., 2016).

Arginine-containing DPR proteins have been proven toxic in a number of studies by various groups. But the question remained whether the nuclear GGGGCC repeat-containing RNA induces toxicity. The study above by Isaacs and colleagues did not detect any significant degeneration when as many as 288 repeats (RO) were expressed in the fly eye (Mizielinska et al., 2014). Gao and colleagues further explored this important question, and tested the effects of a *Drosophila* model expressing a *C9ORF72* minigene with 160 GGGGCC repeats (160R; Tran et al., 2015). The 160R in this model is flanked by human intronic and exonic sequences of *C9ORF72*. Intriguingly, even though the spliced intronic 160R formed abundant nuclear GGGGCC sense RNA foci, they gave rise to no detectable neurodevelopmental defects. Cells were also normal in their rRNA biogenesis and mRNA processing (Tran et al., 2015). When 160R flies were reared at 29°C instead of the normal 25°C, a modest level of toxicity was detected. But since the higher temperature induced 4-folds more DPR proteins, while the number of RNA foci did not change, it is likely that this toxicity was due to the increased DPR proteins instead of the RNA foci (Tran et al., 2015).

Summarizing the above studies, the GGGGCC repeat expansion in C9-ALS/FTD share much commonalities with other READs such as the formation of foci, sequestration of RNA-binding proteins, and induction of nucleolar stress. Yet, unlike most other READs described here, the RNA transcripts containing the GGGGCC repeat expansions and the foci in C9-ALS/FTD do not appear to be the main source of toxicity. Rather, it is the DPR proteins produced by RAN translation (which can still arguably be called a type of RNA toxicity) that are the major determinants of toxicity and degeneration in this disease.

Nevertheless, just when everyone thought GGGGCC repeat-containing RNA does not give rise to any kind of toxicity, Bonini and colleagues discovered a new form of GGGGCC RNA toxicity that may be independent of DPR proteins. In cultured mice spinal cord neurons, GGGGCC and CAG repeat RNAs were found to localize to transport granules that travel to distal neuritic segments (Burguete et al., 2015). Expression of the GGGGCC repeat-containing RNA gave rise to neuritic branching defects. Using a *Drosophila* model, the researchers found that this branching defects can be enhanced by the upregulation or suppressed by the downregulation of *dfmr1* or *orb2* (Burguete et al., 2015). The researchers speculated that the presence of GGGGCC repeat RNA in neurites might result in local RAN translation of toxic DPR proteins. However, they were unable to detect RAN peptides in their study (Burguete et al., 2015). These findings demonstrate that GGGGCC expanded repeat RNAs can cause transport granule dysfunction, which could be a novel RNA

toxicity mechanism independent of DPR proteins, contributing to the neuronal defects in C9-ALS/FTD.

It is important to note that, hexanucleotide repeats such as GGGGCC, are capable of forming distinct structures called DNA and RNA G-quadruplexes (Reddy et al., 2013; Haeusler et al., 2014). These structures facilitate DNA/RNA hybrids, impedes transcription and binds to ribonucleoproteins such as NCL, inducing nucleolar stress (Haeusler et al., 2014). They also form inclusions and sequester hnRNP H, which impedes alternative splicing (Conlon et al., 2016). The real impacts of these G-quadruplex structures on the pathogenesis of C9-ALS/FTD are only just beginning to be elucidated.

THERAPEUTIC DEVELOPMENT TO COMBAT RNA TOXICITY IN READS

Much progress has been made in the past decade regarding potential therapeutic strategies against the RNA toxicity in READs. The current three main streams of therapeutics are oligonucleotide-based therapeutics, peptoid-based therapeutics, and small molecule-based therapeutics. In oligonucleotide-based therapeutics, antisense-oligos (ASOs) have been demonstrated to disrupt foci formation and reduce RNA toxicity in both DM1 and C9-ALS/FTD (Mulders et al., 2009; Wheeler et al., 2012; Donnelly et al., 2013; Lagier-Tourenne et al., 2013; Sareen et al., 2013; Riboldi et al., 2014; Wojtkowiak-Szlachcic et al., 2015). Peptoid-based therapeutics have shown some promising results in mitigating RNA toxicity in DM1 models by reducing the interaction between CUG repeats and MBNL1 and decreasing foci formation (Pushechnikov et al., 2009; García-López et al., 2011; Rzuczek et al., 2013). It has also been reported that a 13-amino-acid peptide can bind to CAG repeats and suppress nucleolar stress and neurotoxicity in a *Drosophila* model of SCA3 (Zhang et al., 2016). Last but not least, designs in small molecules designs have achieved much successes in combating DM1 and C9-ALS/FTD as well. Non-steroidal anti-inflammatory drugs like ketoprofen, histamine receptor inhibitors like orphenadrine and sodium/calcium metabolism modifiers like clenbuterol and spironolactone have been shown in *Drosophila* models of DM1 to mitigate the effects of CUG toxicity (Garcia-Lopez et al., 2008). Pentamidine and neomycin B have been shown to disrupt MBNL1 binding to CUG repeats (Warf et al., 2009). A number of small molecules, including bisamidinium inhibitors, can target toxic RNA-protein interactions and alleviate the sequestration of MBNL1 by repeat expansions (Gareiss et al., 2008; Kumar et al., 2012; Parkesh et al., 2012; Wong et al., 2014; Luu et al., 2016). Porphyrins that distorts RNA G-quadruplex structures can ablate the GGGGCC repeat expansion's interaction with hnRNPA1 or ASF/SF2 (Zamiri et al., 2014). There are small molecules that can even reduce DPR protein production in a model of C9-ALS/FTD (Su et al., 2014). Histone deacetylase (HDAC) inhibitors like sodium butyrate and suberoylanilide hydroxamic acid had been shown to suppress degeneration in a *Drosophila* model of HD (Steffan et al., 2001). Although, this suppression may involve protein toxicity instead of RNA toxicity. Furthermore, in a *Drosophila* model of SBMA and a *Drosophila* model of FXTAS,

overexpression of the HDACs themselves actually suppressed the degeneration phenotypes (Pandey et al., 2007; Todd et al., 2010). Thus, the effects of HDACs inhibitors could be different for different types of READs.

The newest trend in the designing of small molecules is driving toward compounds with multiple modes of actions in combating the pathogenesis of the disease. Zimmerman and colleagues have reported a type of small molecule can target DMI1 pathology by acting as transcription inhibitors, by inhibiting aberrant protein binding to toxic RNA, and by acting as RNase mimics to degrade the toxic RNA (Nguyen et al., 2015). Gitler and colleagues discovered that the transcription elongation factor Spt4 selectively decreased both sense and antisense GGGGCC expanded transcripts, and partially rescued retinal degeneration in the fly eye model of C9-ALS/FTD (Kramer et al., 2016). Knockdown of *SUPT4H1*, the human ortholog of *Spt4*, decreased sense and antisense RNA foci and DPR protein production in C9-ALS/FTD patient-derived cells (Kramer et al., 2016). The study demonstrated Spt4 as a new potential therapeutic target since the disruption of Spt4 can ameliorate degeneration induced by different modes of toxicity sources. Animal models such as the fly models described in this review are valuable tools in the testing of novel compounds and therapeutic targets before clinical trials. They will also continue to serve as important platforms for the discoveries of unknown mechanisms of diseases that may facilitate the development of new therapeutic approaches.

CONCLUDING REMARKS

In this review, we highlighted *Drosophila* as an excellent model organism for the studying of READs. The studies mentioned above have revealed comprehensive information regarding the mechanisms of neurodegeneration in READs, demonstrating the elegant use of the fly models in pioneering discoveries of RNA toxicity. Despite the fact that protein toxicity is a major factor that contributes to the pathogenesis of many READs, we are now certain that RNA toxicity also plays a major role in inducing neural and muscle degeneration via a number of possible ways, including but not limited to the impairment of alternative splicing, the promotion of bidirectional transcription, the induction of nucleolar stress and disruption of nucleocytoplasmic transport, the generation of sRNA from hairpins and dsRNA, and RAN translation of repeat-containing peptides. Although the formation of visible ribonuclear foci is still a hallmark of pathogenesis, more, and more evidence is indicating that the initial emergence of foci does not equal to immediate detectable toxicity. The foci do not initiate the disease process *per se*. In fact, the formation of foci could possibly be a healthy response to temporarily sequester the toxic RNA. Although in the long run, foci may still contribute to the progression of the disease by sequestering RNA-binding proteins, and gradually lead to degeneration.

Apart from the traditional advantages in genetics, the recent advances in CRISPR/Cas9 system and bioinformatics have become a powerful driving force that will surely aid *Drosophila*

research to ascend to the next level. For instance, Liu and colleagues have developed an improved RNA injection-based CRISPR/Cas9 system that is highly efficient for creating desired mutagenesis in the *Drosophila* genome with a mutagenesis induction rate up to 88% (Bassett et al., 2013). O'Connor-Giles and colleagues have expanded the system through homology-directed repair to facilitate complex genome engineering, enabling the insertion of sizable DNA sequences to create targeted knock-ins and conditional knockouts (Gratz et al., 2014). The CRISPR/Cas9 system can also be used for tissue-specific tagging of endogenous proteins (T-STEP; Koles et al., 2015). In contrast to the GAL4/UAS overexpression system, which sometimes may change the localization patterns of proteins, T-STEP retains the endogenous protein localization patterns. Furthermore, it achieves tissue-specific labeling, allowing researcher to dissect tissue-specific functions of the protein of interest. Next generation sequencing is now widely available, and the access of disease information is also becoming more convenient for *Drosophila* geneticists as well. Online databases such as FlyBase has even introduced "Human Disease Model Reports" recently, further facilitating the studies of human diseases using the fruit fly (Millburn et al., 2016). With the development of these techniques and databases, we expect many new and improved fly models of human diseases to emerge soon.

In conclusion, *Drosophila* is a useful *in vivo* model system that can reveal novel mechanistic insights into human disorders, providing the foundation for translational research and therapeutic development. Looking into the future, we anticipate the tiny fruit fly to continue playing an important role in gene and drug discoveries in the ongoing battle between humans and neurodegenerative and neuromuscular diseases.

AUTHOR CONTRIBUTIONS

ACK wrote the manuscript. HVEC contributed to the writing of the manuscript.

ACKNOWLEDGMENTS

This work was supported by the General Research Fund (460712, 461013, 14100714, and 14122815), ANR/RGC Joint Research Scheme (A-CUHK401/14) and Collaborative Research Fund (CUHK1/CRF/13G) of the Hong Kong Research Grants Council; CUHK Lui Che Woo Institute of Innovative Medicine BRAIN Initiative (8303404); Direct Grant for Research from the Research Committee of CUHK (4053148); CUHK Group Research Scheme (3110102); Vice-Chancellor's One-Off Discretionary Fund (4930713); One-off Funding for Joint Lab/Research Collaboration (3132980); and Gerald Choa Neuroscience Centre (7105306); and donations from Chow Tai Fook Charity Foundation (6903898); and Hong Kong Spinocerebellar Ataxia Association (6903291).

REFERENCES

- Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., et al. (2000). The genome sequence of *Drosophila melanogaster*. *Science* 287, 2185–2195. doi: 10.1126/science.287.5461.2185
- Ajrroud-Driss, S., and Siddique, T. (2015). Sporadic and hereditary amyotrophic lateral sclerosis (ALS). *Biochim. Biophys. Acta* 1852, 679–684. doi: 10.1016/j.bbadis.2014.08.010
- Arrasate, M., Mitra, S., Schweitzer, E. S., Segal, M. R., and Finkbeiner, S. (2004). Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature* 431, 805–810. doi: 10.1038/nature02998
- Auluck, P. K., Meulener, M. C., and Bonini, N. M. (2005). Mechanisms of suppression of [alpha]-synuclein neurotoxicity by geldanamycin in *Drosophila*. *J. Biol. Chem.* 280, 2873–2878. doi: 10.1074/jbc.M412106200
- Baizabal-Carvalho, J. F., and Jankovic, J. (2016). Parkinsonism, movement disorders and genetics in frontotemporal dementia. *Nat. Rev. Neurol.* 12, 175–185. doi: 10.1038/nrnneurol.2016.14
- Bañez-Coronel, M., Porta, S., Kagerbauer, B., Mateu-Huertas, E., Pantano, L., Ferrer, I., et al. (2012). A pathogenic mechanism in Huntington's disease involves small CAG-repeated RNAs with neurotoxic activity. *PLoS Genet.* 8:e1002481. doi: 10.1371/journal.pgen.1002481
- Bassell, G. J., and Warren, S. T. (2008). Fragile X syndrome: loss of local mRNA regulation alters synaptic development and function. *Neuron* 60, 201–214. doi: 10.1016/j.neuron.2008.10.004
- Bassett, A. R., Tibbit, C., Ponting, C. P., and Liu, J. L. (2013). Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. *Cell Rep.* 4, 220–228. doi: 10.1016/j.celrep.2013.06.020
- Bellen, H. J., Levis, R. W., Liao, G., He, Y., Carlson, J. W., Tsang, G., et al. (2004). The BDGP gene disruption project: single transposon insertions associated with 40% of *Drosophila* genes. *Genetics* 167, 761–781. doi: 10.1534/genetics.104.026427
- Bellen, H. J., Tong, C., and Tsuda, H. (2010). 100 years of *Drosophila* research and its impact on vertebrate neuroscience: a history lesson for the future. *Nat. Rev. Neurosci.* 11, 514–522. doi: 10.1038/nrn2839
- Belzil, V. V., Gendron, T. F., and Petrucelli, L. (2013). RNA-mediated toxicity in neurodegenerative disease. *Mol. Cell. Neurosci.* 56, 406–419. doi: 10.1016/j.mcn.2012.12.006
- Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409, 363–366. doi: 10.1038/35053110
- Berul, C. I., Maguire, C. T., Aronovitz, M. J., Greenwood, J., Miller, C., Gehrman, J., et al. (1999). DMPK dosage alterations result in atrioventricular conduction abnormalities in a mouse myotonic dystrophy model. *J. Clin. Invest.* 103, R1–R7. doi: 10.1172/JCI5346
- Bhakar, A. L., Dölen, G., and Bear, M. F. (2012). The pathophysiology of fragile X (and what it teaches us about synapses). *Annu. Rev. Neurosci.* 35, 417–443. doi: 10.1146/annurev-neuro-060909-153138
- Bier, E., Vaessin, H., Shepherd, S., Lee, K., McCall, K., Barbel, S., et al. (1989). Searching for pattern and mutation in the *Drosophila* genome with a P-lacZ vector. *Genes Dev.* 3, 1273–1287. doi: 10.1101/gad.3.9.1273
- Bilen, J., and Bonini, N. M. (2007). Genome-wide screen for modifiers of ataxin-3 neurodegeneration in *Drosophila*. *PLoS Genet.* 3:e177. doi: 10.1371/journal.pgen.0030177
- Boeynaems, S., Bogaert, E., Michiels, E., Gijssels, I., Sieben, A., Jovicic, A., et al. (2016). *Drosophila* screen connects nuclear transport genes to DPR pathology in c9ALS/FTD. *Sci. Rep.* 6:20877. doi: 10.1038/srep20877
- Boulton, S., Westman, B. J., Hutten, S., Boisvert, F. M., and Lamond, A. I. (2010). The nucleolus under stress. *Mol. Cell* 40, 216–227. doi: 10.1016/j.molcel.2010.09.024
- Brand, A. H., and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 118, 401–415.
- Brook, J. D., McCurrach, M. E., Harley, H. G., Buckler, A. J., Church, D., Aburatani, H., et al. (1992). Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* 69, 385. doi: 10.1016/0092-8674(92)90154-5
- Brouwer, J. R., Willemsen, R., and Oostra, B. A. (2009). Microsatellite repeat instability and neurological disease. *Bioessays* 31, 71–83. doi: 10.1002/bies.080122
- Burguete, A. S., Almeida, S., Gao, F. B., Kalb, R., Akins, M. R., and Bonini, N. M. (2015). GGGGCC microsatellite RNA is neuriticly localized, induces branching defects, and perturbs transport granule function. *Elife* 4:e08881. doi: 10.7554/eLife.08881
- Calnan, B. J., Tidor, B., Biancalana, S., Hudson, D., and Frankel, A. D. (1991). Arginine-mediated RNA recognition: the arginine fork. *Science* 252, 1167–1171. doi: 10.1126/science.252.5009.1167
- Chan, H. Y., and Bonini, N. M. (2000). *Drosophila* models of human neurodegenerative disease. *Cell Death Differ.* 7, 1075–1080. doi: 10.1038/sj.cdd.4400757
- Chan, H. Y., Warrick, J. M., Gray-Board, G. L., Paulson, H. L., and Bonini, N. M. (2000). Mechanisms of chaperone suppression of polyglutamine disease: selectivity, synergy and modulation of protein solubility in *Drosophila*. *Hum. Mol. Genet.* 9, 2811–2820. doi: 10.1093/hmg/9.19.2811
- Chang, S., Bray, S. M., Li, Z., Zarnescu, D. C., He, C., Jin, P., et al. (2008). Identification of small molecules rescuing fragile X syndrome phenotypes in *Drosophila*. *Nat. Chem. Biol.* 4, 256–263. doi: 10.1038/nchembio.78
- Chen, S., Sayana, P., Zhang, X., and Le, W. (2013). Genetics of amyotrophic lateral sclerosis: an update. *Mol. Neurodegener.* 8:28. doi: 10.1186/1750-1326-8-28
- Cho, D. H., Thienes, C. P., Mahoney, S. E., Analau, E., Filippova, G. N., and Tapscott, S. J. (2005). Antisense transcription and heterochromatin at the DMI CTG repeats are constrained by CTCF. *Mol. Cell* 20, 483–489. doi: 10.1016/j.molcel.2005.09.002
- Coffee, R. L. Jr., Tessier, C. R., Woodruff, E. A. III., and Broadie, K. (2010). Fragile X mental retardation protein has a unique, evolutionarily conserved neuronal function not shared with FXR1P or FXR2P. *Dis. Model. Mech.* 3, 471–485. doi: 10.1242/dmm.004598
- Conlon, E. G., Lu, L., Sharma, A., Yamazaki, T., Tang, T., Shneider, N. A., et al. (2016). The C9ORF72 GGGGCC expansion forms RNA G-quadruplex inclusions and sequesters hnRNP H to disrupt splicing in ALS brains. *Elife* 5:e17820. doi: 10.7554/eLife.17820
- Cooley, L., Kelley, R., and Spradling, A. (1988). Insertional mutagenesis of the *Drosophila* genome with single P elements. *Science* 239, 1121–1128. doi: 10.1126/science.2830671
- David, G., Abbas, N., Stevanin, G., Dürr, A., Yvert, G., Cancel, G., et al. (1997). Cloning of the SCA7 gene reveals a highly unstable CAG repeat expansion. *Nat. Genet.* 17, 65–70. doi: 10.1038/ng0997-65
- Davis, B. M., McCurrach, M. E., Taneja, K. L., Singer, R. H., and Housman, D. E. (1997). Expansion of a CUG trinucleotide repeat in the 3' untranslated region of myotonic dystrophy protein kinase transcripts results in nuclear retention of transcripts. *Proc. Natl. Acad. Sci. U.S.A.* 94, 7388–7393. doi: 10.1073/pnas.94.14.7388
- de Haro, M., Al-Ramahi, I., De Gouyon, B., Ukani, L., Rosa, A., Faustino, N. A., et al. (2006). MBNL1 and CUGBP1 modify expanded CUG-induced toxicity in a *Drosophila* model of myotonic dystrophy type 1. *Hum. Mol. Genet.* 15, 2138–2145. doi: 10.1093/hmg/ddl137
- de Haro, M., Al-Ramahi, I., Jones, K. R., Holth, J. K., Timchenko, L. T., and Botas, J. (2013). Smaug/SAMD4A restores translational activity of CUGBP1 and suppresses CUG-induced myopathy. *PLoS Genet.* 9:e1003445. doi: 10.1371/journal.pgen.1003445
- DeJesus-Hernandez, M., Mackenzie, I. R., Boeve, B. F., Boxer, A. L., Baker, M., Rutherford, N. J., et al. (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* 72, 245–256. doi: 10.1016/j.neuron.2011.09.011
- Dietzl, G., Chen, D., Schnorrer, F., Su, K. C., Barinova, Y., Fellner, M., et al. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* 448, 151–156. doi: 10.1038/nature05954
- Dockendorff, T. C., Su, H. S., McBride, S. M., Yang, Z., Choi, C. H., Siwicki, K. K., et al. (2002). *Drosophila* lacking dfmr1 activity show defects in circadian output and fail to maintain courtship interest. *Neuron* 34, 973–984. doi: 10.1016/S0896-6273(02)00724-9
- Donnelly, C. J., Zhang, P. W., Pham, J. T., Haeusler, A. R., Mistry, N. A., Vidensky, S., et al. (2013). RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention. *Neuron* 80, 415–428. doi: 10.1016/j.neuron.2013.10.015

- Doumanis, J., Wada, K., Kino, Y., Moore, A. W., and Nukina, N. (2009). RNAi screening in *Drosophila* cells identifies new modifiers of mutant huntingtin aggregation. *PLoS ONE* 4:e2725. doi: 10.1371/journal.pone.0007275
- Droppelmann, C. A., Campos-Melo, D., Ishtiaq, M., Volkening, K., and Strong, M. J. (2014). RNA metabolism in ALS: when normal processes become pathological. *Amyotroph. Lateral Scler. Frontotemporal Degener.* 15, 321–336. doi: 10.3109/21678421.2014.881377
- Duffy, J. B. (2002). GAL4 system in *Drosophila*: a fly geneticist's Swiss army knife. *Genesis* 34, 1–15. doi: 10.1002/gene.10150
- Ebralidze, A., Wang, Y., Petkova, V., Ebralidze, K., and Junghans, R. P. (2004). RNA leaching of transcription factors disrupts transcription in myotonic dystrophy. *Science* 303, 383–387. doi: 10.1126/science.1088679
- Faust, K., Gehrke, S., Yang, Y., Yang, L., Beal, M. F., and Lu, B. (2009). Neuroprotective effects of compounds with antioxidant and anti-inflammatory properties in a *Drosophila* model of Parkinson's disease. *BMC Neurosci.* 10:109. doi: 10.1186/1471-2202-10-109
- Fernandez-Funez, P., Nino-Rosales, M. L., de Gouyon, B., She, W. C., Luchak, J. M., Martinez, P., et al. (2000). Identification of genes that modify ataxin-1-induced neurodegeneration. *Nature* 408, 101–106. doi: 10.1038/35040584
- Fiszler, A., and Krzyzosiak, W. J. (2013). RNA toxicity in polyglutamine disorders: concepts, models, and progress of research. *J. Mol. Med.* 91, 683–691. doi: 10.1007/s00109-013-1016-2
- Fleischer, B. (1918). Über myotonische Dystrophie mit Katarakt. *Albrecht. Von Graefes. Arch. Klin. Exp. Ophthalmol.* 96, 91–133. doi: 10.1007/BF02018704
- Floresco, S. B., Geyer, M. A., Gold, L. H., and Grace, A. A. (2005). Developing predictive animal models and establishing a preclinical trials network for assessing treatment effects on cognition in schizophrenia. *Schizophr. Bull.* 31, 888–894. doi: 10.1093/schbul/sbi041
- Fratta, P., Poulter, M., Lashley, T., Rohrer, J. D., Polke, J. M., Beck, J., et al. (2013). Homozygosity for the C9orf72 GGGGCC repeat expansion in frontotemporal dementia. *Acta Neuropathol.* 126, 401–409. doi: 10.1007/s00401-013-1147-0
- Freibaum, B. D., Lu, Y., Lopez-Gonzalez, R., Kim, N. C., Almeida, S., Lee, K. H., et al. (2015). GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature* 525, 129–133. doi: 10.1038/nature14974
- Fu, Y. H., Friedman, D. L., Richards, S., Pearlman, J. A., Gibbs, R. A., Pizzuti, A., et al. (1993). Decreased expression of myotonin-protein kinase messenger RNA and protein in adult form of myotonic dystrophy. *Science* 260, 235–238. doi: 10.1126/science.8469976
- Fu, Y. H., Kuhl, D. P., Pizzuti, A., Pieretti, M., Sutcliffe, J. S., Richards, S., et al. (1991). Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* 67, 1047–1058. doi: 10.1016/0092-8674(91)90283-5
- Fu, Y. H., Pizzuti, A., Fenwick, R. G. Jr., King, J., Rajnarayan, S., Dunne, P. W., et al. (1992). An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* 255, 1256–1258. doi: 10.1126/science.1546326
- Galloway, J. N., Shaw, C., Yu, P., Parghi, D., Poidevin, M., Jin, P., et al. (2014). CGG repeats in RNA modulate expression of TDP-43 in mouse and fly models of fragile X tremor ataxia syndrome. *Hum. Mol. Genet.* 23, 5906–5915. doi: 10.1093/hmg/ddu314
- García-López, A., Llamusi, B., Orzáez, M., Pérez-Payá, E., and Artero, R. D. (2011). *In vivo* discovery of a peptide that prevents CUG-RNA hairpin formation and reverses RNA toxicity in myotonic dystrophy models. *Proc. Natl. Acad. Sci. U.S.A.* 108, 11866–11871. doi: 10.1073/pnas.1018213108
- García-Lopez, A., Monferrer, L., Garcia-Alcover, I., Vicente-Crespo, M., Alvarez-Abril, M. C., and Artero, R. D. (2008). Genetic and chemical modifiers of a CUG toxicity model in *Drosophila*. *PLoS ONE* 3:e1595. doi: 10.1371/journal.pone.0001595
- Gareiss, P. C., Sobczak, K., McNaughton, B. R., Palde, P. B., Thornton, C. A., and Miller, B. L. (2008). Dynamic combinatorial selection of molecules capable of inhibiting the (CUG) repeat RNA-MBNL1 interaction *in vitro*: discovery of lead compounds targeting myotonic dystrophy (DM1). *J. Am. Chem. Soc.* 130, 16254–16261. doi: 10.1021/ja804398y
- Gratz, S. J., Ukken, F. P., Rubinstein, C. D., Thiede, G., Donohue, L. K., Cummings, A. M., et al. (2014). Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in *Drosophila*. *Genetics* 196, 961–971. doi: 10.1534/genetics.113.160713
- Greco, C. M., Hagerman, R. J., Tassone, F., Chudley, A. E., Del Bigio, M. R., Jacquemont, S., et al. (2002). Neuronal intranuclear inclusions in a new cerebellar tremor/ataxia syndrome among fragile X carriers. *Brain* 125(Pt 8), 1760–1771. doi: 10.1093/brain/awf184
- Haeusler, A. R., Donnelly, C. J., Periz, G., Simko, E. A., Shaw, P. G., Kim, M. S., et al. (2014). C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* 507, 195–200. doi: 10.1038/nature13124
- Hand, C. K., and Rouleau, G. A. (2002). Familial amyotrophic lateral sclerosis. *Muscle Nerve* 25, 135–159. doi: 10.1002/mus.10001
- Harms, M. B., Cady, J., Zaidman, C., Cooper, P., Bali, T., Allred, P., et al. (2013). Lack of C9ORF72 coding mutations supports a gain of function for repeat expansions in amyotrophic lateral sclerosis. *Neurobiol. Aging* 34, 2234.e13–2234.e19. doi: 10.1016/j.neurobiolaging.2013.03.006
- Harper, P. S. (2001). *Myotonic Dystrophy, 3rd Eds.* London: W.B. Saunders Company.
- Harris, K. P., and Littleton, J. T. (2015). Transmission, development, and plasticity of synapses. *Genetics* 201, 345–375. doi: 10.1534/genetics.115.176529
- Harrison, C. J., Jack, E. M., Allen, T. D., and Harris, R. (1983). The fragile X: a scanning electron microscope study. *J. Med. Genet.* 20, 280–285. doi: 10.1136/jmg.20.4.280
- He, F., Krans, A., Freibaum, B. D., Taylor, J. P., and Todd, P. K. (2014). TDP-43 suppresses CGG repeat-induced neurotoxicity through interactions with HnRNP A2/B1. *Hum. Mol. Genet.* 23, 5036–5051. doi: 10.1093/hmg/ddu216
- Hirth, F. (2010). *Drosophila melanogaster* in the study of human neurodegeneration. *CNS Neurol. Disord. Drug Targets* 9, 504–523. doi: 10.2174/187152710791556104
- Ho, T. H., Charlet, B. N., Poulos, M. G., Singh, G., Swanson, M. S., and Cooper, T. A. (2004). Muscleblind proteins regulate alternative splicing. *EMBO J.* 23, 3103–3112. doi: 10.1038/sj.emboj.7600300
- Ho, T. H., Savkur, R. S., Poulos, M. G., Mancini, M. A., Swanson, M. S., and Cooper, T. A. (2005). Colocalization of muscleblind with RNA foci is separable from mis-regulation of alternative splicing in myotonic dystrophy. *J. Cell Sci.* 118(Pt 13), 2923–2933. doi: 10.1242/jcs.02404
- Houseley, J. M., Wang, Z., Brock, G. J., Soloway, J., Artero, R., Perez-Alonso, M., et al. (2005). Myotonic dystrophy associated expanded CUG repeat muscleblind positive ribonuclear foci are not toxic to *Drosophila*. *Hum. Mol. Genet.* 14, 873–883. doi: 10.1093/hmg/ddi080
- Hsu, R. J., Hsiao, K. M., Lin, M. J., Li, C. Y., Wang, L. C., Chen, L. K., et al. (2011). Long tract of untranslated CAG repeats is deleterious in transgenic mice. *PLoS ONE* 6:e16417. doi: 10.1371/journal.pone.0016417
- Hummel, T., and Klämbt, C. (2008). P-element mutagenesis. *Methods Mol. Biol.* 420, 97–117. doi: 10.1007/978-1-59745-583-1_6
- Ikenaka, K., Katsuno, M., Kawai, K., Ishigaki, S., Tanaka, F., and Sobue, G. (2012). Disruption of axonal transport in motor neuron diseases. *Int. J. Mol. Sci.* 13, 1225–1238. doi: 10.3390/ijms13011225
- Imbert, G., Saudou, F., Yvert, G., Devys, D., Trottier, Y., Garnier, J. M., et al. (1996). Cloning of the gene for spinocerebellar ataxia 2 reveals a locus with high sensitivity to expanded CAG/glutamine repeats. *Nat. Genet.* 14, 285–291. doi: 10.1038/ng1196-285
- Jackson, G. R., Salecker, I., Dong, X., Yao, X., Arnheim, N., Faber, P. W., et al. (1998). Polyglutamine-expanded human huntingtin transgenes induce degeneration of *Drosophila* photoreceptor neurons. *Neuron* 21, 633–642. doi: 10.1016/S0896-6273(00)80573-5
- Jackson, S. M., Whitworth, A. J., Greene, J. C., Libby, R. T., Baccam, S. L., Pallanck, L. J., et al. (2005). A SCA7 CAG/CTG repeat expansion is stable in *Drosophila melanogaster* despite modulation of genomic context and gene dosage. *Gene* 347, 35–41. doi: 10.1016/j.gene.2004.12.008
- Jacquemont, S., Hagerman, R. J., Hagerman, P. J., and Leehey, M. A. (2007). Fragile-X syndrome and fragile X-associated tremor/ataxia syndrome: two faces of FMR1. *Lancet Neurol.* 6, 45–55. doi: 10.1016/S1474-4422(06)70676-7
- Jaiswal, M., Sandoval, H., Zhang, K., Bayat, V., and Bellen, H. J. (2012). Probing mechanisms that underlie human neurodegenerative diseases in *Drosophila*. *Annu. Rev. Genet.* 46, 371–396. doi: 10.1146/annurev-genet-110711-155456
- Jin, P., Duan, R., Qurashi, A., Qin, Y., Tian, D., Rosser, T. C., et al. (2007). Pur alpha binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a *Drosophila* model of fragile X tremor/ataxia syndrome. *Neuron* 55, 556–564. doi: 10.1016/j.neuron.2007.07.020
- Jin, P., Zarnescu, D. C., Zhang, F., Pearson, C. E., Lucchesi, J. C., Moses, K., et al. (2003). RNA-mediated neurodegeneration caused by the

- fragile X premutation rCGG repeats in *Drosophila*. *Neuron* 39, 739–747. doi: 10.1016/S0896-6273(03)00533-6
- Jovicic, A., Mertens, J., Boeynaems, S., Bogaert, E., Chai, N., Yamada, S. B., et al. (2015). Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nat. Neurosci.* 18, 1226–1229. doi: 10.1038/nn.4085
- Kaltenbach, L. S., Romero, E., Becklin, R. R., Chettier, R., Bell, R., Phansalkar, A., et al. (2007). Huntingtin interacting proteins are genetic modifiers of neurodegeneration. *PLoS Genet.* 3:e82. doi: 10.1371/journal.pgen.0030082
- Kanadia, R. N., Johnstone, K. A., Mankodi, A., Lungu, C., Thornton, C. A., Esson, D., et al. (2003). A muscleblind knockout model for myotonic dystrophy. *Science* 302, 1978–1980. doi: 10.1126/science.1088583
- Kawaguchi, Y., Okamoto, T., Taniwaki, M., Aizawa, M., Inoue, M., Katayama, S., et al. (1994). CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. *Nat. Genet.* 8, 221–228. doi: 10.1038/ng1194-221
- Kiernan, M. C., Vucic, S., Cheah, B. C., Turner, M. R., Eisen, A., Hardiman, O., et al. (2011). Amyotrophic lateral sclerosis. *Lancet* 377, 942–955. doi: 10.1016/S0140-6736(10)61156-7
- Koch, K. S., and Leffert, H. L. (1998). Giant hairpins formed by CUG repeats in myotonic dystrophy messenger RNAs might sterically block RNA export through nuclear pores. *J. Theor. Biol.* 192, 505–514. doi: 10.1006/jtbi.1998.0679
- Koide, R., Ikeuchi, T., Onodera, O., Tanaka, H., Igarashi, S., Endo, K., et al. (1994). Unstable expansion of CAG repeat in hereditary dentatorubral-pallidolusian atrophy (DRPLA). *Nat. Genet.* 6, 9–13. doi: 10.1038/ng0194-9
- Koles, K., Yeh, A. R., and Rodal, A. A. (2015). Tissue-specific tagging of endogenous loci in *Drosophila melanogaster*. *Biol. Open* 5, 83–89. doi: 10.1242/bio.016089
- Koob, M. D., Moseley, M. L., Schut, L. J., Benzow, K. A., Bird, T. D., Day, J. W., et al. (1999). An untranslated CTG expansion causes a novel form of spinocerebellar ataxia (SCA8). *Nat. Genet.* 21, 379–384. doi: 10.1038/7710
- Korkut, C., and Budnik, V. (2009). WNTs tune up the neuromuscular junction. *Nat. Rev. Neurosci.* 10, 627–634. doi: 10.1038/nrn2681
- Kramer, J. M., and Staveley, B. E. (2003). GAL4 causes developmental defects and apoptosis when expressed in the developing eye of *Drosophila melanogaster*. *Genet. Mol. Res.* 2, 43–47.
- Kramer, N. J., Carlomagno, Y., Zhang, Y. J., Almeida, S., Cook, C. N., Gendron, T. F., et al. (2016). Spt4 selectively regulates the expression of C9orf72 sense and antisense mutant transcripts. *Science* 353, 708–712. doi: 10.1126/science.aaf7791
- Kremer, E. J., Pritchard, M., Lynch, M., Yu, S., Holman, K., Baker, E., et al. (1991). Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)_n. *Science* 252, 1711–1714. doi: 10.1126/science.1675488
- Krol, J., Fiszler, A., Mykowska, A., Sobczak, K., de Mezer, M., and Krzyzosiak, W. J. (2007). Ribonuclease dicer cleaves triplet repeat hairpins into shorter repeats that silence specific targets. *Mol. Cell* 25, 575–586. doi: 10.1016/j.molcel.2007.01.031
- Kumar, A., Parkesh, R., Sznajder, L. J., Childs-Disney, J. L., Sobczak, K., and Disney, M. D. (2012). Chemical correction of pre-mRNA splicing defects associated with sequestration of muscleblind-like 1 protein by expanded r(CAG)-containing transcripts. *ACS Chem. Biol.* 7, 496–505. doi: 10.1021/cb200413a
- Kurtzke, J. F. (1982). Epidemiology of amyotrophic lateral sclerosis. *Adv. Neurol.* 36, 281–302.
- Ladd, P. D., Smith, L. E., Rabaia, N. A., Moore, J. M., Georges, S. A., Hansen, R. S., et al. (2007). An antisense transcript spanning the CGG repeat region of FMR1 is upregulated in premutation carriers but silenced in full mutation individuals. *Hum. Mol. Genet.* 16, 3174–3187. doi: 10.1093/hmg/ddm293
- Lagier-Tourenne, C., Baughn, M., Rigo, F., Sun, S., Liu, P., Li, H. R., et al. (2013). Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proc. Natl. Acad. Sci. U.S.A.* 110, E4530–E4539. doi: 10.1073/pnas.1318835110
- La Spada, A. R., Wilson, E. M., Lubahn, D. B., Harding, A. E., and Fischbeck, K. H. (1991). Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* 352, 77–79. doi: 10.1038/352077a0
- Latouche, M., Lasbleiz, C., Martin, E., Monnier, V., Debeir, T., Mouatt-Prigent, A., et al. (2007). A conditional pan-neuronal *Drosophila* model of spinocerebellar ataxia 7 with a reversible adult phenotype suitable for identifying modifier genes. *J. Neurosci.* 27, 2483–2492. doi: 10.1523/JNEUROSCI.5453-06.2007
- Lawlor, K. T., O'Keefe, L. V., Samaraweera, S. E., van Eyk, C. L., McLeod, C. J., Maloney, C. A., et al. (2011). Double-stranded RNA is pathogenic in *Drosophila* models of expanded repeat neurodegenerative diseases. *Hum. Mol. Genet.* 20, 3757–3768. doi: 10.1093/hmg/ddr292
- Lee, A., Li, W., Xu, K., Bogert, B. A., Su, K., and Gao, F. B. (2003). Control of dendritic development by the *Drosophila* fragile X-related gene involves the small GTPase Rac1. *Development* 130, 5543–5552. doi: 10.1242/dev.00792
- Lee, K. H., Zhang, P., Kim, H. J., Mitrea, D. M., Sarkar, M., Freibaum, B. D., et al. (2016). C9orf72 dipeptide repeats impair the assembly, dynamics, and function of membrane-less organelles. *Cell* 167, 774–788.e17. doi: 10.1016/j.cell.2016.10.002
- Lee, W. C., Yoshihara, M., and Littleton, J. T. (2004). Cytoplasmic aggregates trap polyglutamine-containing proteins and block axonal transport in a *Drosophila* model of Huntington's disease. *Proc. Natl. Acad. Sci. U.S.A.* 101, 3224–3229. doi: 10.1073/pnas.0400243101
- Li, L. B., Yu, Z., Teng, X., and Bonini, N. M. (2008). RNA toxicity is a component of ataxin-3 degeneration in *Drosophila*. *Nature* 453, 1107–1111. doi: 10.1038/nature06909
- Lindenboim, L., Borner, C., and Stein, R. (2011). Nuclear proteins acting on mitochondria. *Biochim. Biophys. Acta* 1813, 584–596. doi: 10.1016/j.bbamcr.2010.11.016
- Liquori, C. L., Ricker, K., Moseley, M. L., Jacobsen, J. F., Kress, W., Naylor, S. L., et al. (2001). Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science* 293, 864–867. doi: 10.1126/science.1062125
- Luscombe, N. M., Laskowski, R. A., and Thornton, J. M. (2001). Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res.* 29, 2860–2874. doi: 10.1093/nar/29.13.2860
- Luu, L. M., Nguyen, L., Peng, S., Lee, J., Lee, H. Y., Wong, C. H., et al. (2016). A potent inhibitor of protein sequestration by expanded triplet (CUG) repeats that shows phenotypic improvements in a *Drosophila* model of myotonic dystrophy. *ChemMedChem* 11, 1428–1435. doi: 10.1002/cmdc.201600081
- MacDonald, M. E., Ambrose, C. M., Duyao, M. P., Myers, R. H., Lin, C., and Srinidhi, L., et al. (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. *Cell* 72, 971–983. doi: 10.1016/0092-8674(93)90585-e
- Mahadevan, M., Tsilfidis, C., Sabourin, L., Shutler, G., Amemiya, C., Jansen, G., et al. (1992). Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. *Science* 255, 1253–1255. doi: 10.1126/science.1546325
- Majounie, E., Renton, A. E., Mok, K., Dopper, E. G., Waite, A., Rollinson, S., et al. (2012). Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol.* 11, 323–330. doi: 10.1016/S1474-4422(12)70043-1
- Mankodi, A., Logigian, E., Callahan, L., McClain, C., White, R., Henderson, D., et al. (2000). Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science* 289, 1769–1773. doi: 10.1126/science.289.5485.1769
- Marsh, J. L., and Thompson, L. M. (2006). *Drosophila* in the study of neurodegenerative disease. *Neuron* 52, 169–178. doi: 10.1016/j.neuron.2006.09.025
- Martin, J. P., and Bell, J. (1943). A pedigree of mental defect showing sex-linkage. *J. Neurol. Psychiatry* 6, 154–157. doi: 10.1136/jnnp.6.3-4.154
- McGurk, L., Berson, A., and Bonini, N. M. (2015). *Drosophila* as an *in vivo* model for human neurodegenerative disease. *Genetics* 201, 377–402. doi: 10.1534/genetics.115.179457
- Mercy, L., Hodges, J. R., Dawson, K., Barker, R. A., and Brayne, C. (2008). Incidence of early-onset dementias in Cambridgeshire, United Kingdom. *Neurology* 71, 1496–1499. doi: 10.1212/01.wnl.0000334277.16896.f
- Millburn, G. H., Crosby, M. A., Gramates, L. S., Tweedie, S., and FlyBase, C. (2016). FlyBase portals to human disease research using *Drosophila* models. *Dis. Model. Mech.* 9, 245–252. doi: 10.1242/dmm.023317
- Millecamps, S., and Julien, J. P. (2013). Axonal transport deficits and neurodegenerative diseases. *Nat. Rev. Neurosci.* 14, 161–176. doi: 10.1038/nrn3380
- Miller, J. W., Urbinat, C. R., Teng-Ummuay, P., Stenberg, M. G., Byrne, B. J., Thornton, C. A., et al. (2000). Recruitment of human muscleblind proteins

- to (CUG)_n expansions associated with myotonic dystrophy. *EMBO J.* 19, 4439–4448. doi: 10.1093/emboj/19.17.4439
- Minnerop, M., Weber, B., Schoene-Bake, J. C., Roeske, S., Mirbach, S., Anspach, C., et al. (2011). The brain in myotonic dystrophy 1 and 2: evidence for a predominant white matter disease. *Brain* 134(Pt 12), 3530–3546. doi: 10.1093/brain/awr299
- Mirkin, S. M. (2007). Expandable DNA repeats and human disease. *Nature* 447, 932–940. doi: 10.1038/nature05977
- Mitchell, J. B., Thornton, J. M., Singh, J., and Price, S. L. (1992). Towards an understanding of the arginine-aspartate interaction. *J. Mol. Biol.* 226, 251–262. doi: 10.1016/0022-2836(92)90137-9
- Mizielinska, S., Gronke, S., Niccoli, T., Ridler, C. E., Clayton, E. L., Devoy, A., et al. (2014). C9orf72 repeat expansions cause neurodegeneration in *Drosophila* through arginine-rich proteins. *Science* 345, 1192–1194. doi: 10.1126/science.1256800
- Montermini, L., Andermann, E., Labuda, M., Richter, A., Pandolfo, M., Cavalcanti, F., et al. (1997). The Friedreich ataxia GAA triplet repeat: premutation and normal alleles. *Hum. Mol. Genet.* 6, 1261–1266. doi: 10.1093/hmg/6.8.1261
- Morales, J., Hiesinger, P. R., Schroeder, A. J., Kume, K., Verstreken, P., Jackson, F. R., et al. (2002). *Drosophila* fragile X protein, DFXR, regulates neuronal morphology and function in the brain. *Neuron* 34, 961–972. doi: 10.1016/S0896-6273(02)00731-6
- Moseley, M. L., Zu, T., Ikeda, Y., Gao, W., Mosemiller, A. K., Daughters, R. S., et al. (2006). Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. *Nat. Genet.* 38, 758–769. doi: 10.1038/ng1827
- Mulders, S. A., van den Broek, W. J., Wheeler, T. M., Croes, H. J., van Kuik-Romeijn, P., de Kimpe, S. J., et al. (2009). Triplet-repeat oligonucleotide-mediated reversal of RNA toxicity in myotonic dystrophy. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13915–13920. doi: 10.1073/pnas.0905780106
- Muqit, M. M., and Feany, M. B. (2002). Modelling neurodegenerative diseases in *Drosophila*: a fruitful approach? *Nat. Rev. Neurosci.* 3, 237–243. doi: 10.1038/nrn751
- Mutsuddi, M., Marshall, C. M., Benzow, K. A., Koob, M. D., and Rebay, I. (2004). The spinocerebellar ataxia 8 noncoding RNA causes neurodegeneration and associates with staufen in *Drosophila*. *Curr. Biol.* 14, 302–308. doi: 10.1016/j.cub.2004.01.034
- Mykowska, A., Sobczak, K., Wojciechowska, M., Kozlowski, P., and Krzyzosiak, W. J. (2011). CAG repeats mimic CUG repeats in the misregulation of alternative splicing. *Nucleic Acids Res.* 39, 8938–8951. doi: 10.1093/nar/gkr608
- Nakamura, K., Jeong, S. Y., Uchihara, T., Anno, M., Nagashima, K., Nagashima, T., et al. (2001). SCA17, a novel autosomal dominant cerebellar ataxia caused by an expanded polyglutamine in TATA-binding protein. *Hum. Mol. Genet.* 10, 1441–1448. doi: 10.1093/hmg/10.14.1441
- Nalavade, R., Griesche, N., Ryan, D. P., Hildebrand, S., and Krauss, S. (2013). Mechanisms of RNA-induced toxicity in CAG repeat disorders. *Cell Death Dis.* 4, e752. doi: 10.1038/cddis.2013.276
- Napoletano, F., Occhi, S., Calamita, P., Volpi, V., Blanc, E., Charroux, B., et al. (2011). Polyglutamine Atrophin provokes neurodegeneration in *Drosophila* by repressing fat. *EMBO J.* 30, 945–958. doi: 10.1038/emboj.2011.1
- Nedelsky, N. B., Pennuto, M., Smith, R. B., Palazzolo, I., Moore, J., Nie, Z., et al. (2010). Native functions of the androgen receptor are essential to pathogenesis in a *Drosophila* model of spinobulbar muscular atrophy. *Neuron* 67, 936–952. doi: 10.1016/j.neuron.2010.08.034
- Nguyen, L., Luu, L. M., Peng, S., Serrano, J. F., Chan, H. Y., and Zimmerman, S. C. (2015). Rationally designed small molecules that target both the DNA and RNA causing myotonic dystrophy type 1. *J. Am. Chem. Soc.* 137, 14180–14189. doi: 10.1021/jacs.5b09266
- Nisoli, I., Chauvin, J. P., Napoletano, F., Calamita, P., Zanin, V., Fanto, M., et al. (2010). Neurodegeneration by polyglutamine Atrophin is not rescued by induction of autophagy. *Cell Death Differ.* 17, 1577–1587. doi: 10.1038/cdd.2010.31
- O'Hare, K., and Rubin, G. M. (1983). Structures of P transposable elements and their sites of insertion and excision in the *Drosophila melanogaster* genome. *Cell* 34, 25–35. doi: 10.1016/0092-8674(83)90133-2
- Oh, S. Y., He, F., Krans, A., Frazier, M., Taylor, J. P., Paulson, H. L., et al. (2015). RAN translation at CGG repeats induces ubiquitin proteasome system impairment in models of fragile X-associated tremor ataxia syndrome. *Hum. Mol. Genet.* 24, 4317–4326. doi: 10.1093/hmg/ddv165
- Orr, H. T., Chung, M. Y., Banfi, S., Kwiatkowski, T. J. Jr., Servadio, A., Beaudet, A. L., et al. (1993). Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nat. Genet.* 4, 221–226. doi: 10.1038/ng0793-221
- Panaite, P. A., Kuntzer, T., Gourdon, G., Lohrinus, J. A., and Barakat-Walter, I. (2013). Functional and histopathological identification of the respiratory failure in a DMSXL transgenic mouse model of myotonic dystrophy. *Dis. Model. Mech.* 6, 622–631. doi: 10.1242/dmm.010512
- Pandey, U. B., and Nichols, C. D. (2011). Human disease models in *Drosophila melanogaster* and the role of the fly in therapeutic drug discovery. *Pharmacol. Rev.* 63, 411–436. doi: 10.1124/pr.110.003293
- Pandey, U. B., Nie, Z., Batlevi, Y., McCray, B. A., Ritson, G. P., Nedelsky, N. B., et al. (2007). HDAC6 rescues neurodegeneration and provides an essential link between autophagy and the UPS. *Nature* 447, 859–863. doi: 10.1038/nature05853
- Parkesh, R., Childs-Disney, J. L., Nakamori, M., Kumar, A., Wang, E., Wang, T., et al. (2012). Design of a bioactive small molecule that targets the myotonic dystrophy type 1 RNA via an RNA motif-ligand database and chemical similarity searching. *J. Am. Chem. Soc.* 134, 4731–4742. doi: 10.1021/ja210088v
- Pearson, C. E., Nichol Edamura, K., and Cleary, J. D. (2005). Repeat instability: mechanisms of dynamic mutations. *Nat. Rev. Genet.* 6, 729–742. doi: 10.1038/nrg1689
- Perry, D. C., and Miller, B. L. (2013). Frontotemporal dementia. *Semin. Neurol.* 33, 336–341. doi: 10.1055/s-0033-1359316
- Petrucci, L., O'Farrell, C., Lockhart, P. J., Baptista, M., Kehoe, K., Vink, L., et al. (2002). Parkin protects against the toxicity associated with mutant alpha-synuclein: proteasome dysfunction selectively affects catecholaminergic neurons. *Neuron* 36, 1007–1019. doi: 10.1016/S0896-6273(02)01125-X
- Philips, A. V., Timchenko, L. T., and Cooper, T. A. (1998). Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science* 280, 737–741. doi: 10.1126/science.280.5364.737
- Picchio, L., Plantie, E., Renaud, Y., Poovthumkadavil, P., and Jagla, K. (2013). Novel *Drosophila* model of myotonic dystrophy type 1: phenotypic characterization and genome-wide view of altered gene expression. *Hum. Mol. Genet.* 22, 2795–2810. doi: 10.1093/hmg/ddt127
- Polak, U., McIvor, E., Dent, S. Y., Wells, R. D., and Napierala, M. (2013). Expanded complexity of unstable repeat diseases. *Biofactors* 39, 164–175. doi: 10.1002/biof.1060
- Prelich, G. (2012). Gene overexpression: uses, mechanisms, and interpretation. *Genetics* 190, 841–854. doi: 10.1534/genetics.111.136911
- Pulst, S. M., Nechiporuk, A., Nechiporuk, T., Gispert, S., Chen, X. N., Lopes-Cendes, I., et al. (1996). Moderate expansion of a normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2. *Nat. Genet.* 14, 269–276. doi: 10.1038/ng1196-269
- Pushchynikov, A., Lee, M. M., Childs-Disney, J. L., Sobczak, K., French, J. M., Thornton, C. A., et al. (2009). Rational design of ligands targeting triplet repeating transcripts that cause RNA dominant disease: application to myotonic muscular dystrophy type 1 and spinocerebellar ataxia type 3. *J. Am. Chem. Soc.* 131, 9767–9779. doi: 10.1021/ja9020149
- Rajendran, L., Schneider, A., Schlechtingen, G., Weidlich, S., Ries, J., Braxmeier, T., et al. (2008). Efficient inhibition of the Alzheimer's disease beta-secretase by membrane targeting. *Science* 320, 520–523. doi: 10.1126/science.1156609
- Reddy, K., Zamiri, B., Stanley, S. Y., Macgregor, R. B. Jr., and Pearson, C. E. (2013). The disease-associated r(GGGGCC)_n repeat from the C9orf72 gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures. *J. Biol. Chem.* 288, 9860–9866. doi: 10.1074/jbc.C113.452532
- Reddy, S., Smith, D. B., Rich, M. M., Leferovich, J. M., Reilly, P., Davis, B. M., et al. (1996). Mice lacking the myotonic dystrophy protein kinase develop a late onset progressive myopathy. *Nat. Genet.* 13, 325–335. doi: 10.1038/ng0796-325
- Reiter, L. T., Potocki, L., Chien, S., Gribskov, M., and Bier, E. (2001). A systematic analysis of human disease-associated gene sequences in *Drosophila melanogaster*. *Genome Res.* 11, 1114–1125. doi: 10.1101/gr.169101
- Ren, J., Jegga, A. G., Zhang, M., Deng, J., Liu, J., Gordon, C. B., et al. (2011). A *Drosophila* model of the neurodegenerative disease SCA17 reveals a role of RBP-J/Su(H) in modulating the pathological outcome. *Hum. Mol. Genet.* 20, 3424–3436. doi: 10.1093/hmg/ddr251

- Renoux, A. J., and Todd, P. K. (2012). Neurodegeneration the RNA way. *Prog. Neurobiol.* 97, 173–189. doi: 10.1016/j.pneurobio.2011.10.006
- Renton, A. E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J. R., et al. (2011). A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* 72, 257–268. doi: 10.1016/j.neuron.2011.09.010
- Riboldi, G., Zanetta, C., Ranieri, M., Nizzardo, M., Simone, C., Magri, F., et al. (2014). Antisense oligonucleotide therapy for the treatment of C9ORF72 ALS/FTD diseases. *Mol. Neurobiol.* 50, 721–732. doi: 10.1007/s12035-014-8724-7
- Romero, E., Cha, G. H., Verstreken, P., Ly, C. V., Hughes, R. E., Bellen, H. J., et al. (2008). Suppression of neurodegeneration and increased neurotransmission caused by expanded full-length huntingtin accumulating in the cytoplasm. *Neuron* 57, 27–40. doi: 10.1016/j.neuron.2007.11.025
- Ross, C. A., and Poirier, M. A. (2005). Opinion: what is the role of protein aggregation in neurodegeneration? *Nat. Rev. Mol. Cell Biol.* 6, 891–898. doi: 10.1038/nrm1742
- Rüb, U., Schöls, L., Paulson, H., Auburger, G., Kermer, P., Jen, J. C., et al. (2013). Clinical features, neurogenetics and neuropathology of the polyglutamine spinocerebellar ataxias type 1, 2, 3, 6 and 7. *Prog. Neurobiol.* 104, 38–66. doi: 10.1016/j.pneurobio.2013.01.001
- Rubin, G. M. (2000). Biological annotation of the *Drosophila* genome sequence. *Novartis Found Symp.* 229, 79–82. discussion: 82–73.
- Rubin, G. M., and Spradling, A. C. (1982). Genetic transformation of *Drosophila* with transposable element vectors. *Science* 218, 348–353. doi: 10.1126/science.6289436
- Rzuczek, S. G., Gao, Y., Tang, Z. Z., Thornton, C. A., Kodadek, T., and Disney, M. D. (2013). Features of modularly assembled compounds that impart bioactivity against an RNA target. *ACS Chem. Biol.* 8, 2312–2321. doi: 10.1021/cb400265y
- Sang, T. K., and Jackson, G. R. (2005). *Drosophila* models of neurodegenerative disease. *NeuroRx* 2, 438–446. doi: 10.1602/neurorx.2.3.438
- Sanpei, K., Takano, H., Igarashi, S., Sato, T., Oyake, M., Sasaki, H., et al. (1996). Identification of the spinocerebellar ataxia type 2 gene using a direct identification of repeat expansion and cloning technique, DIRECT. *Nat. Genet.* 14, 277–284. doi: 10.1038/ng1196-277
- Santoro, M. R., Bray, S. M., and Warren, S. T. (2012). Molecular mechanisms of fragile X syndrome: a twenty-year perspective. *Annu. Rev. Pathol.* 7, 219–245. doi: 10.1146/annurev-pathol-011811-132457
- Sareen, D., O'Rourke, J. G., Meera, P., Muhammad, A. K., Grant, S., Simpkinson, M., et al. (2013). Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. *Sci. Transl. Med.* 5, 208ra149. doi: 10.1126/scitranslmed.3007529
- Sharma, R. P., and Chopra, V. L. (1976). Effect of the Wingless (*wg1*) mutation on wing and haltere development in *Drosophila melanogaster*. *Dev. Biol.* 48, 461–465. doi: 10.1016/0012-1606(76)90108-1
- Shieh, S. Y., and Bonini, N. M. (2011). Genes and pathways affected by CAG-repeat RNA-based toxicity in *Drosophila*. *Hum. Mol. Genet.* 20, 4810–4821. doi: 10.1093/hmg/ddr420
- Sisodia, S. S. (1998). Nuclear inclusions in glutamine repeat disorders: are they pernicious, coincidental, or beneficial? *Cell* 95, 1–4. doi: 10.1016/S0092-8674(00)81743-2
- Sobczak, K., de Mezer, M., Michlewski, G., Krol, J., and Krzyzosiak, W. J. (2003). RNA structure of trinucleotide repeats associated with human neurological diseases. *Nucleic Acids Res.* 31, 5469–5482. doi: 10.1093/nar/gkg766
- Sobczak, K., Michlewski, G., de Mezer, M., Kierzek, E., Krol, J., Olejniczak, M., et al. (2010). Structural diversity of triplet repeat RNAs. *J. Biol. Chem.* 285, 12755–12764. doi: 10.1074/jbc.M109.078790
- Sofola, O. A., Jin, P., Qin, Y., Duan, R., Liu, H., de Haro, M., et al. (2007). RNA-binding proteins hnRNP A2/B1 and CUGBP1 suppress fragile X CAG premutation repeat-induced neurodegeneration in a *Drosophila* model of FXTAS. *Neuron* 55, 565–571. doi: 10.1016/j.neuron.2007.07.021
- Sopher, B. L., Ladd, P. D., Pineda, V. V., Libby, R. T., Sunkin, S. M., Hurley, J. B., et al. (2011). CTCF regulates ataxin-7 expression through promotion of a convergently transcribed, antisense noncoding RNA. *Neuron* 70, 1071–1084. doi: 10.1016/j.neuron.2011.05.027
- Spindler, S. R., and Hartenstein, V. (2010). The *Drosophila* neural lineages: a model system to study brain development and circuitry. *Dev. Genes Evol.* 220, 1–10. doi: 10.1007/s00427-010-0323-7
- Spradling, A. C., Stern, D., Beaton, A., Rhem, E. J., Laverty, T., Mozden, N., et al. (1999). The Berkeley *Drosophila* Genome Project gene disruption project: single P-element insertions mutating 25% of vital *Drosophila* genes. *Genetics* 153, 135–177.
- Stefanis, L., Larsen, K. E., Rideout, H. J., Sulzer, D., and Greene, L. A. (2001). Expression of A53T mutant but not wild-type alpha-synuclein in PC12 cells induces alterations of the ubiquitin-dependent degradation system, loss of dopamine release, and autophagic cell death. *J. Neurosci.* 21, 9549–9560.
- Steffan, J. S., Bodai, L., Pallos, J., Poelman, M., McCampbell, A., Apostol, B. L., et al. (2001). Histone deacetylase inhibitors arrest polyglutamine-dependent neurodegeneration in *Drosophila*. *Nature* 413, 739–743. doi: 10.1038/35099568
- St Johnston, D. (2002). The art and design of genetic screens: *Drosophila melanogaster*. *Nat. Rev. Genet.* 3, 176–188. doi: 10.1038/nrg751
- Strong, M. J. (2010). The evidence for altered RNA metabolism in amyotrophic lateral sclerosis (ALS). *J. Neurol. Sci.* 288, 1–12. doi: 10.1016/j.jns.2009.09.029
- Su, Z., Zhang, Y., Gendron, T. F., Bauer, P. O., Chew, J., Yang, W. Y., et al. (2014). Discovery of a biomarker and lead small molecules to target r(GGGGCC)-associated defects in c9FTD/ALS. *Neuron* 83, 1043–1050. doi: 10.1016/j.neuron.2014.07.041
- Takeyama, K., Ito, S., Yamamoto, A., Tanimoto, H., Furutani, T., Kanuka, H., et al. (2002). Androgen-dependent neurodegeneration by polyglutamine-expanded human androgen receptor in *Drosophila*. *Neuron* 35, 855–864. doi: 10.1016/S0896-6273(02)00875-9
- Taneja, K. L., McCurrach, M., Schalling, M., Housman, D., and Singer, R. H. (1995). Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J. Cell Biol.* 128, 995–1002. doi: 10.1083/jcb.128.6.995
- Tassone, F., and Hagerman, R. (2012). The fragile X-associated tremor ataxia syndrome. *Results Probl. Cell Differ.* 54, 337–357. doi: 10.1007/978-3-642-21649-7_18
- Thornton, C. A., Johnson, K., and Moxley, R. T. III. (1994). Myotonic dystrophy patients have larger CTG expansions in skeletal muscle than in leukocytes. *Ann. Neurol.* 35, 104–107. doi: 10.1002/ana.410350116
- Todd, P. K., Oh, S. Y., Krans, A., He, F., Sellier, C., Frazer, M., et al. (2013). CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. *Neuron* 78, 440–455. doi: 10.1016/j.neuron.2013.03.026
- Todd, P. K., Oh, S. Y., Krans, A., Pandey, U. B., Di Prospero, N. A., Min, K. T., et al. (2010). Histone deacetylases suppress CGG repeat-induced neurodegeneration via transcriptional silencing in models of fragile X tremor ataxia syndrome. *PLoS Genet.* 6:e1001240. doi: 10.1371/journal.pgen.1001240
- Todd, P. K., and Paulson, H. L. (2010). RNA-mediated neurodegeneration in repeat expansion disorders. *Ann. Neurol.* 67, 291–300. doi: 10.1002/ana.21948
- Tran, H., Almeida, S., Moore, J., Gendron, T. F., Chalasani, U., Lu, Y., et al. (2015). Differential toxicity of nuclear RNA foci versus dipeptide repeat proteins in a *Drosophila* model of C9ORF72 FTD/ALS. *Neuron* 87, 1207–1214. doi: 10.1016/j.neuron.2015.09.015
- Tsoi, H., Lau, C. K., Lau, K. F., and Chan, H. Y. (2011). Perturbation of U2AF65/NXF1-mediated RNA nuclear export enhances RNA toxicity in polyQ diseases. *Hum. Mol. Genet.* 20, 3787–3797. doi: 10.1093/hmg/ddr297
- Tsoi, H., Lau, T. C., Tsang, S. Y., Lau, K. F., and Chan, H. Y. (2012). CAG expansion induces nucleolar stress in polyglutamine diseases. *Proc. Natl. Acad. Sci. U.S.A.* 109, 13428–13433. doi: 10.1073/pnas.1204089109
- Tsou, W. L., Hosking, R. R., Burr, A. A., Sutton, J. R., Ouyang, M., Du, X., et al. (2015). DnaJ-1 and karyopherin alpha3 suppress degeneration in a new *Drosophila* model of Spinocerebellar Ataxia Type 6. *Hum. Mol. Genet.* 24, 4385–4396. doi: 10.1093/hmg/ddv174
- Tsou, W. L., Qiblawi, S. H., Hosking, R. R., Gomez, C. M., and Todi, S. V. (2016). Polyglutamine length-dependent toxicity from alpha1ACT in *Drosophila* models of spinocerebellar ataxia type 6. *Biol. Open* 5, 1770–1775. doi: 10.1242/bio.021667
- Wan, L., Dockendorff, T. C., Jongens, T. A., and Dreyfuss, G. (2000). Characterization of dFMR1, a *Drosophila melanogaster* homolog of the fragile X mental retardation protein. *Mol. Cell Biol.* 20, 8536–8547. doi: 10.1128/MCB.20.22.8536-8547.2000
- Wang, L. C., Chen, K. Y., Pan, H., Wu, C. C., Chen, P. H., Liao, Y. T., et al. (2011). Muscleblind participates in RNA toxicity of expanded CAG and

- CUG repeats in *Caenorhabditis elegans*. *Cell. Mol. Life Sci.* 68, 1255–1267. doi: 10.1007/s00018-010-0522-4
- Warf, M. B., Nakamori, M., Matthys, C. M., Thornton, C. A., and Berglund, J. A. (2009). Pentamidine reverses the splicing defects associated with myotonic dystrophy. *Proc. Natl. Acad. Sci. U.S.A.* 106, 18551–18556. doi: 10.1073/pnas.0903234106
- Warrick, J. M., Chan, H. Y., Gray-Board, G. L., Chai, Y., Paulson, H. L., and Bonini, N. M. (1999). Suppression of polyglutamine-mediated neurodegeneration in *Drosophila* by the molecular chaperone HSP70. *Nat. Genet.* 23, 425–428. doi: 10.1038/70532
- Warrick, J. M., Morabito, L. M., Bilen, J., Gordesky-Gold, B., Faust, L. Z., Paulson, H. L., et al. (2005). Ataxin-3 suppresses polyglutamine neurodegeneration in *Drosophila* by a ubiquitin-associated mechanism. *Mol. Cell* 18, 37–48. doi: 10.1016/j.molcel.2005.02.030
- Warrick, J. M., Paulson, H. L., Gray-Board, G. L., Bui, Q. T., Fischbeck, K. H., Pittman, R. N., et al. (1998). Expanded polyglutamine protein forms nuclear inclusions and causes neural degeneration in *Drosophila*. *Cell* 93, 939–949. doi: 10.1016/S0092-8674(00)81200-3
- Weiss, K. R., Kimura, Y., Lee, W. C., and Littleton, J. T. (2012). Huntingtin aggregation kinetics and their pathological role in a *Drosophila* Huntington's disease model. *Genetics* 190, 581–600. doi: 10.1534/genetics.111.133710
- Wheeler, T. M., Leger, A. J., Pandey, S. K., MacLeod, A. R., Nakamori, M., Cheng, S. H., et al. (2012). Targeting nuclear RNA for *in vivo* correction of myotonic dystrophy. *Nature* 488, 111–115. doi: 10.1038/nature11362
- Wojciechowska, M., and Krzyzosiak, W. J. (2011). Cellular toxicity of expanded RNA repeats: focus on RNA foci. *Hum. Mol. Genet.* 20, 3811–3821. doi: 10.1093/hmg/ddr299
- Wojtkowiak-Szlachcic, A., Taylor, K., Stepniak-Konieczna, E., Sznajder, L. J., Mykowska, A., Sroka, J., et al. (2015). Short antisense-locked nucleic acids (all-LNAs) correct alternative splicing abnormalities in myotonic dystrophy. *Nucleic Acids Res.* 43, 3318–3331. doi: 10.1093/nar/gkv163
- Wong, C. H., Nguyen, L., Peh, J., Luu, L. M., Sanchez, J. S., Richardson, S. L., et al. (2014). Targeting toxic RNAs that cause myotonic dystrophy type 1 (DM1) with a bisamidinium inhibitor. *J. Am. Chem. Soc.* 136, 6355–6361. doi: 10.1021/ja5012146
- Xu, Z., Poidevin, M., Li, X., Li, Y., Shu, L., Nelson, D. L., et al. (2013). Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. *Proc. Natl. Acad. Sci. U.S.A.* 110, 7778–7783. doi: 10.1073/pnas.1219643110
- Xu, Z., and Yang, C. (2014). TDP-43-The key to understanding amyotrophic lateral sclerosis. *Rare Dis.* 2:e944443. doi: 10.4161/21675511.2014.944443
- Yamamoto, S., Jaiswal, M., Charng, W. L., Gambin, T., Karaca, E., Mirzaa, G., et al. (2014). A *drosophila* genetic resource of mutants to study mechanisms underlying human genetic diseases. *Cell* 159, 200–214. doi: 10.1016/j.cell.2014.09.002
- Yu, Z., and Bonini, N. M. (2011). Modeling human trinucleotide repeat diseases in *Drosophila*. *Int. Rev. Neurobiol.* 99, 191–212. doi: 10.1016/B978-0-12-387003-2.00008-2
- Yu, Z., Goodman, L. D., Shieh, S. Y., Min, M., Teng, X., Zhu, Y., et al. (2015). A fly model for the CCUG-repeat expansion of myotonic dystrophy type 2 reveals a novel interaction with MBNL1. *Hum. Mol. Genet.* 24, 954–962. doi: 10.1093/hmg/ddu507
- Yu, Z., Teng, X., and Bonini, N. M. (2011). Triplet repeat-derived siRNAs enhance RNA-mediated toxicity in a *Drosophila* model for myotonic dystrophy. *PLoS Genet.* 7:e1001340. doi: 10.1371/journal.pgen.1001340
- Yuan, Y., Compton, S. A., Sobczak, K., Stenberg, M. G., Thornton, C. A., Griffith, J. D., et al. (2007). Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs. *Nucleic Acids Res.* 35, 5474–5486. doi: 10.1093/nar/gkm601
- Zamiri, B., Reddy, K., Macgregor, R. B. Jr., and Pearson, C. E. (2014). TMPyP4 porphyrin distorts RNA G-quadruplex structures of the disease-associated r(GGGGCC)_n repeat of the C9orf72 gene and blocks interaction of RNA-binding proteins. *J. Biol. Chem.* 289, 4653–4659. doi: 10.1074/jbc.C113.502336
- Zhang, Y. Q., Bailey, A. M., Matthies, H. J. G., Renden, R. B., Smith, M. A., Speese, S. D., et al. (2001). *Drosophila* fragile X-related gene regulates the MAP1B homolog Futsch to control synaptic structure and function. *Cell* 107, 591–603. doi: 10.1016/S0092-8674(01)00589-X
- Zhang, K., Donnelly, C. J., Haeusler, A. R., Grima, J. C., Machamer, J. B., Steinwald, P., et al. (2015). The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature* 525, 56–61. doi: 10.1038/nature14973
- Zhang, Q., Tsoi, H., Peng, S., Li, P. P., Lau, K. F., Rudnicki, D. D., et al. (2016). Assessing a peptidic inhibitor-based therapeutic approach that simultaneously suppresses polyglutamine RNA- and protein-mediated toxicities in patient cells and *Drosophila*. *Dis. Model. Mech.* 9, 321–334. doi: 10.1242/dmm.022350
- Zhou, Z. X., Wong, C. I., Sar, M., and Wilson, E. M. (1994). The androgen receptor: an overview. *Recent Prog. Horm. Res.* 49, 249–274. doi: 10.1016/b978-0-12-571149-4.50017-9
- Zhuchenko, O., Bailey, J., Bonnen, P., Ashizawa, T., Stockton, D. W., Amos, C., et al. (1997). Autosomal dominant cerebellar ataxia (SCA6) associated with small polyglutamine expansions in the alpha 1A-voltage-dependent calcium channel. *Nat. Genet.* 15, 62–69. doi: 10.1038/ng0197-62
- Zoghbi, H. Y., and Botas, J. (2002). Mouse and fly models of neurodegeneration. *Trends Genet.* 18, 463–471. doi: 10.1016/S0168-9525(02)02729-4
- Zoghbi, H. Y., and Orr, H. T. (2000). Glutamine repeats and neurodegeneration. *Annu. Rev. Neurosci.* 23, 217–247. doi: 10.1146/annurev.neuro.23.1.217
- Zu, T., Gibbens, B., Doty, N. S., Gomes-Pereira, M., Huguet, A., Stone, M. D., et al. (2011). Non-ATG-initiated translation directed by microsatellite expansions. *Proc. Natl. Acad. Sci. U.S.A.* 108, 260–265. doi: 10.1073/pnas.1013343108

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Koon and Chan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.