

PUBLISHED VERSION

Richards, Robert Ian; Samaraweera, Saumya Erandathi; van Eyk, Clare Louise; O'Keefe, Louise Veronica; Suter, Catherine

RNA pathogenesis via Toll-like receptor-activated inflammation in expanded repeat neurodegenerative diseases, *Frontiers in Molecular Neuroscience*, 2013; 6:25.

© 2013 Richards, Samaraweera, van Eyk, O'Keefe and Suter. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

Originally published by Frontiers Research Foundation –

http://www.frontiersin.org/Molecular_Neuroscience/10.3389/fnmol.2013.00025/abstract

PERMISSIONS

<http://www.frontiersin.org/Copyright.aspx>

Articles and other user-contributed materials may be downloaded and reproduced subject to any copyright or other notices.

As an author or contributor you grant permission to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the [Conditions for Authors](#) and [Conditions for Website Use](#) and subject to any copyright notices which you include in connection with such materials. The licence granted to third parties is a Creative Common Attribution ("CC BY") licence. Note that for articles published prior to July 2012, the licence may be different and you should check the pdf version of any article to establish what licence was granted. If an article carries only a non-commercial licence and you wish to obtain a commercial licence, please contact Frontiers.

4th December 2013

<http://hdl.handle.net/2440/81160>



RNA pathogenesis via Toll-like receptor-activated inflammation in expanded repeat neurodegenerative diseases

Robert I. Richards^{1*}, Saumya E. Samaraweera¹, Clare L. van Eyk¹, Louise V. O'Keefe¹ and Catherine M. Suter²

¹ Discipline of Genetics and Centre for Molecular Pathology, School of Molecular and Biomedical Science, The University of Adelaide, Adelaide, SA, Australia

² Victor Chang Cardiac Research Institute, Darlinghurst, NSW, Australia

Edited by:

Hermona Soreq, The Hebrew University of Jerusalem, Israel

Reviewed by:

Anthony J. Hannan, University of Melbourne, Australia

Alexander K. Murashov, East Carolina University, USA

*Correspondence:

Robert I. Richards, Discipline of Genetics and Centre for Molecular Pathology, School of Molecular and Biomedical Science, The University of Adelaide, North Terrace, Adelaide, SA 5000, Australia
e-mail: robert.richards@adelaide.edu.au

Previously, we hypothesized that an RNA-based pathogenic pathway has a causal role in the dominantly inherited unstable expanded repeat neurodegenerative diseases. In support of this hypothesis we, and others, have characterized *rCAG.rCUG*₁₀₀ repeat double-strand RNA (dsRNA) as a previously unidentified agent capable of causing pathogenesis in a *Drosophila* model of neurodegenerative disease. *Dicer*, *Toll*, and autophagy pathways have distinct roles in this *Drosophila* dsRNA pathology. *Dicer* dependence is accompanied by cleavage of *rCAG.rCUG*₁₀₀ repeat dsRNA down to *r(CAG)*₇ 21-mers. Among the “molecular hallmarks” of this pathway that have been identified in *Drosophila*, some [i.e., *r(CAG)*₇ and elevated tumor necrosis factor] correlate with observations in affected people (e.g., Huntington's disease and amyotrophic lateral sclerosis) or in related animal models (i.e., autophagy). The *Toll* pathway is activated in the presence of repeat-containing dsRNA and toxicity is also dependent on this pathway. How might the endogenously expressed dsRNA mediate *Toll*-dependent toxicity in neuronal cells? Endogenous RNAs are normally shielded from *Toll* pathway activation as part of the mechanism to distinguish “self” from “non-self” RNAs. This typically involves post-transcriptional modification of the RNA. Therefore, it is likely that *rCAG.rCUG*₁₀₀ repeat dsRNA has a characteristic property that interferes with or evades this normal mechanism of shielding. We predict that repeat expansion leads to an alteration in RNA structure and/or form that perturbs RNA modification, causing the unshielded repeat RNA (in the form of its *Dicer*-cleaved products) to be recognized by *Toll*-like receptors (TLRs), with consequent activation of the *Toll* pathway leading to loss of cell function and then ultimately cell death. We hypothesize that the proximal cause of expanded repeat neurodegenerative diseases is the TLR recognition (and resultant innate inflammatory response) of repeat RNA as “non-self” due to their paucity of “self” modification.

Keywords: RNA pathogenesis, Toll-like receptor, innate inflammation, expanded repeat diseases, neuro-degeneration

INTRODUCTION

Since the first discovery of trinucleotide repeat expansion as the basis for many important human genetic diseases (Kremer et al., 1991; La Spada et al., 1991; Yu et al., 1991; Richards and Sutherland, 1992), there has been a vast amount of research in this area (*PubMed* search “trinucleotide repeat disorders” gives >3,700 results). Much of this research is aimed at identifying the mechanism of pathogenesis underlying diseases caused by this form of mutation. Individual diseases can follow either dominant or recessive mode of inheritance indicating distinct pathogenic pathways. Repeat sequences that are expanded in copy number are the basis for ~20 dominantly inherited neurodegenerative diseases, including Huntington's disease (HD). Despite some of the responsible genes being identified as long as 20 years ago, the identity and nature of the disease-causing pathogenic pathway remains a gap in knowledge for these diseases, i.e., no definitive molecular pathway from the mutation to the clinical symptoms has yet

been identified. For at-risk individuals in families affected with dominantly inherited late-onset neurodegenerative diseases due to expanded repeats, the majority opt not to have the definitive pre-symptomatic diagnostic test. Their preference is to live with the uncertainty of not knowing, than the certainty of getting the disease, as no treatments are yet available. Therefore, determining the pathogenic pathway and identifying therapeutic targets for intervention is an urgent priority for reducing the impact of these devastating diseases. This understanding is essential for rational approaches to delay onset, slow progression, or ultimately effect cure.

MOLECULAR PATHWAY FROM REPEAT EXPANSION TO DISEASE

There are common properties exhibited by the various repeat expansions that give rise to human disease. The vast majority of these diseases originate from an existing repeat sequence that

exhibits copy number variation in the human population. In each case, the disease alleles arise when copy number increases beyond a critical threshold. The repeat composition varies, but most are trinucleotide repeats. In some diseases, the repeat expands to the point where gene expression at the expanded repeat locus is either substantially reduced or lost altogether, resulting in loss-of-function of the repeat-harboring gene. Typically such diseases are inherited in a recessive manner. Many repeat loci, however, give rise to dominantly inherited diseases in a manner that is not gene-dose dependent (i.e., two mutant alleles are no worse, and may be even better than one – see Carroll et al., 2013). This suggests that gain-of-function is the mechanism rather than haploinsufficiency. Repeat copy number in many cases is a major determinant of age at onset of clinical symptoms (referred to as “anticipation”) indicating that the repeat itself is a rate-limiting determinant of the pathogenic pathway. However, since pathology typically involves cell death and there are many ways in which cells die, identification of the disease-causing “toxic agent” has been problematic.

IS THERE A COMMON PATHOGENIC AGENT?

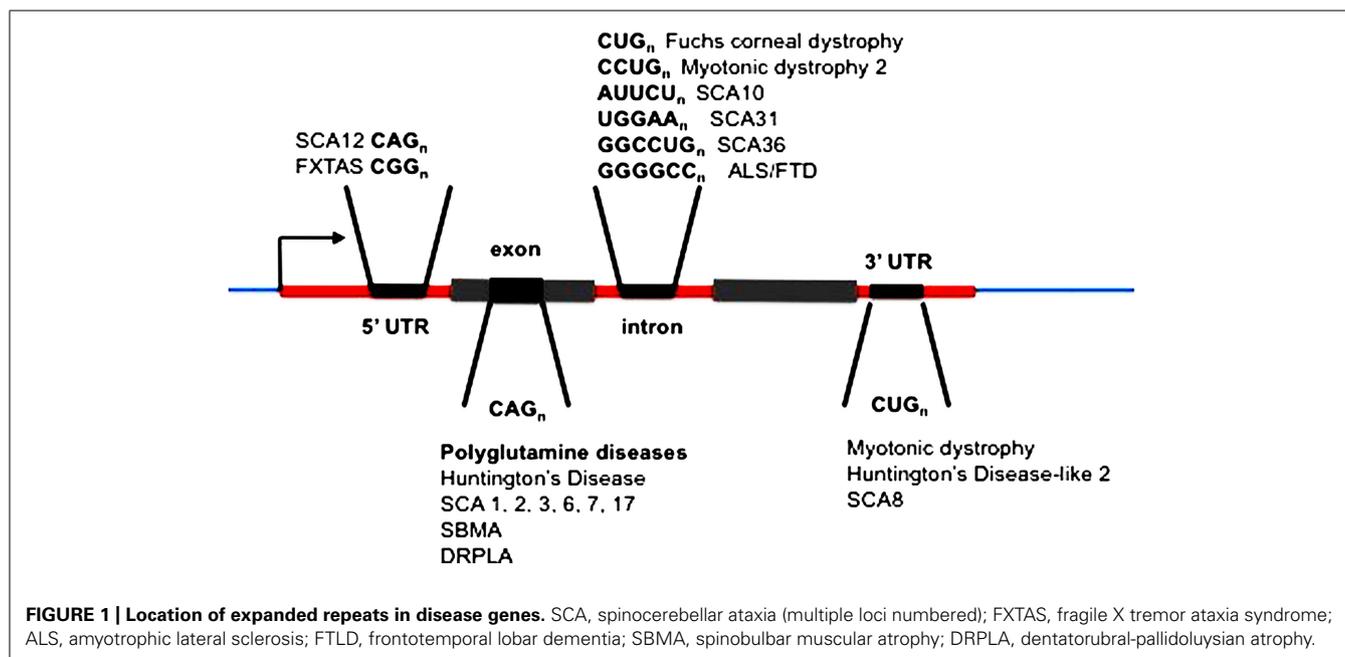
The unstable expanded repeat diseases (Figure 1) typically manifest as neurodegenerative and/or muscular diseases, some with a high degree of clinical overlap, despite affecting distinct proteins and unrelated loci. Where expanded repeats are translated, they generally code for polyglutamine; however, the proteins in which they are located are all unrelated in the remainder of their amino-acid sequence. Therefore, much attention has been focused on expanded polyglutamine as the common basis of pathology (McLeod et al., 2005; van Eyk et al., 2011, 2012). Some of these diseases, however, have repeat expansions located within untranslated RNAs and/or arise from repeat sequences that cannot encode polyglutamine (Figure 1; Richards, 2001; La

Spada and Taylor, 2010). Despite these significant differences in the location of the repeat in this family of diseases they exhibit overlapping symptoms resulting from neuronal loss of function and/or neurodegeneration. In addition, in most cases the polyglutamine and “untranslated” diseases have similar disease allele copy number repeat thresholds (*HD* CAG > 36, *SCA17* CAG > 47, *FXTAS* CGG > 55, *HDL2* CUG > 44, *DM1* CUG > 50, *SCA12* CAG > 66). This suggests that there may be a common pathogenic agent or agents in the translated and untranslated repeat diseases.

RNA MAY BE PATHOGENIC IN TRANSLATED REPEAT DISEASES

While there is growing consensus that RNA plays a causal role in “non-coding repeat expansion disorders,” its contribution when the repeat is located in coding regions (specifically polyglutamine disorders) is more controversial (Fischer and Krzyzosiak, 2013). Yet even here there is evidence that RNA is key. For example, intermediate copy number CAG alleles of *SCA2* that are below the threshold required to encode aggregate forming polyglutamine, increase the risk of amyotrophic lateral sclerosis (ALS; Elden et al., 2010). Furthermore, interruption of CAG repeat with CAA dramatically mitigates polyglutamine toxicity in a *Drosophila* model of *SCA3* (Li et al., 2008).

It is possible that multiple pathways (at least one of which is RNA mediated) contribute to progression of expanded repeat neurodegenerative diseases. In support of this possibility, ALS and *SCA7* both appear to involve two cell types (nerve cells and glial cells; Furrer et al., 2011; Polymenidou and Cleveland, 2011). Astrocytes and glial cells have both been shown to affect their neighboring neurons in individuals with repeat expansions, leading Ilieva et al. (2009) to hypothesize that the onset of the disease is determined in the nerve cell, and the progression of the disease determined in adjacent glial(-like) cells. But



importantly, there is consistent evidence implicating RNA as of principle importance as the originating causal event that initiates pathology.

MECHANISMS OF RNA-INITIATED PATHOLOGY

What precedents and potential mechanisms are there for RNA to initiate pathogenesis in human diseases? See **Figure 2**.

SINGLE-STRANDED RNA TOXICITY

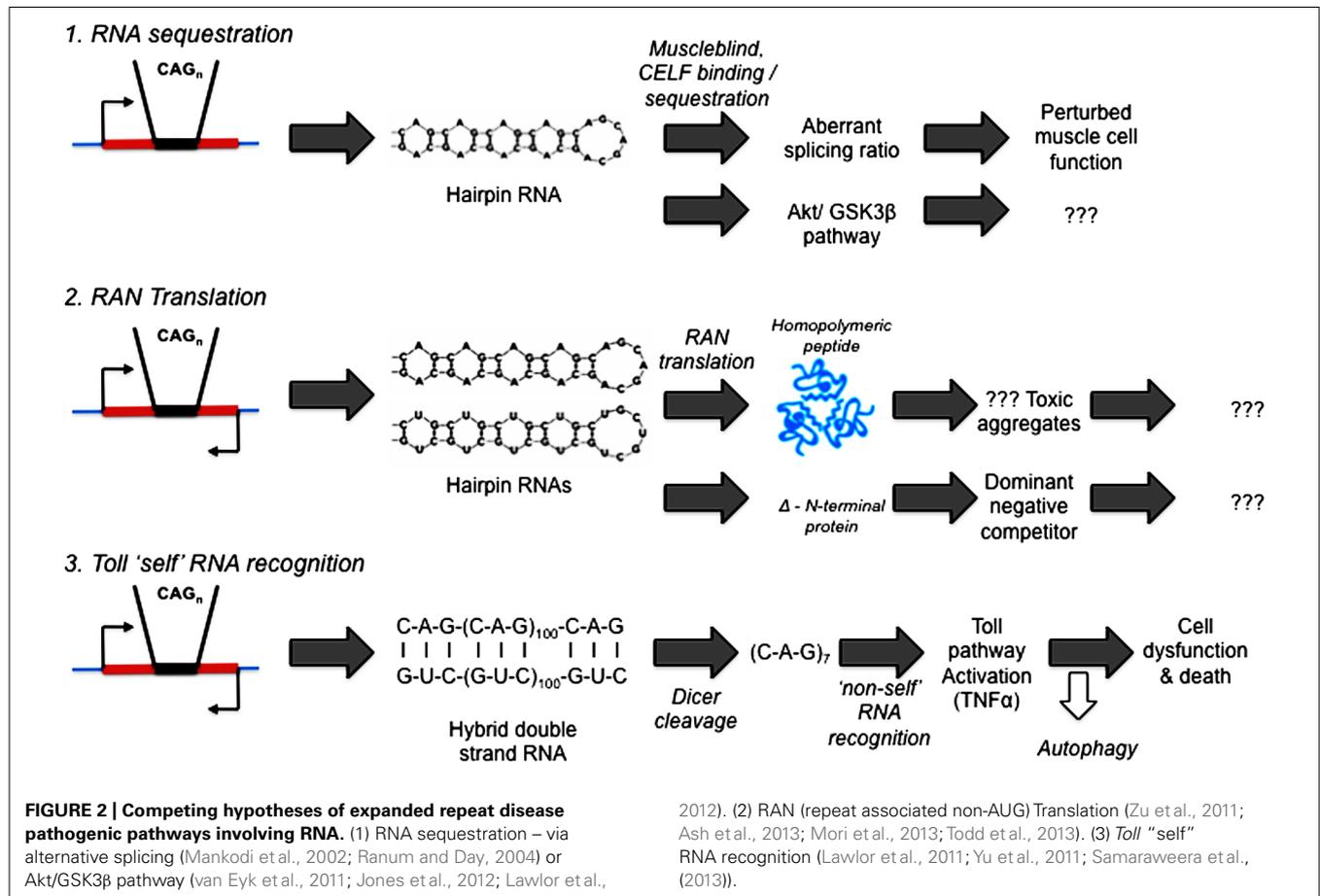
Precedence for expanded repeat RNA being a disease-causing entity in its own right first came from the *DM1* and *DM2* repeat expansions that both give rise to myotonic dystrophy (Ranum and Day, 2004). The repeat expansions in these diseases are similar, but importantly, not identical (CUG vs. CCUG) and are located in untranslated regions (3'UTR or intron) of two otherwise unrelated genes (*DMPK* and *ZNF9*). In muscle cells, RNAs from expanded alleles of either repeat are able to bind and sequester alternative splicing factors (muscleblind and CUG-BP) and in so doing, perturb the splicing pathways of proteins for which alternative splicing is a necessary step for their complete range of functions (Mankodi et al., 2002; Ranum and Day, 2004). It is now generally accepted that RNA is the common pathogenic agent in these diseases most likely through its impact on alternative splicing, although this has recently been challenged with evidence that GSK3 β mediates at least some aspects of the RNA-based pathology in myotonic

dystrophy (Jones et al., 2012) and in a *Drosophila* model (van Eyk et al., 2011).

Evidence for a more widespread role for RNA in neurodegenerative diseases has been steadily accumulating. *SCA31* and *SCA36* are due to large expansions of *de novo* 5 bp TGGAA repeat and an existing 6 bp GGCCTG repeat, respectively – both located within introns of different genes (Sato et al., 2009; Kobayashi et al., 2011). An expanded GGGGCC repeat has recently been found to cause a substantial proportion of cases of ALS and frontotemporal lobar dementia (FTLD; DeJesus-Hernandez et al., 2011; Renton et al., 2011). As indicated by others (Orr, 2011) “The location of this repeat within an intron of the *C9ORF72* gene along with some evidence for alternative splicing of *C9ORF72* transcripts brings in to play a prominent aspect of non-coding repeat expansion disorders – the role of RNA metabolism in pathogenesis.”

REPEAT ASSOCIATED NON-AUG TRANSLATION

The hairpin structure of expanded repeat RNA is such that it can enable the initiation of translation in the absence of the normal requirement of an AUG start codon (Zu et al., 2011). Although this mechanism involves conversion of the RNA into peptides, thereby rendering the RNA no longer “untranslated,” the phenomenon can occur to RNA sequences that do not normally appear in protein-coding sequences, i.e., RNA from introns or 5' or 3' untranslated regions of mRNAs. The resultant translated



polypeptides can initiate from within the repeat sequence and in any reading frame, therefore, a single strand containing repeat RNA sequence can encode three different polypeptide sequences. Since expanded repeat sequences are typically located in regions of bi-directional transcription (Batra et al., 2010), the resultant transcripts from both strands potentially enable the production of six different peptide sequences, any of which may be toxic to the cell. Such polypeptides have now been detected in pathology specimens from individuals affected with a number of different expanded repeat diseases including DM1, fragile X syndrome (FRAXA) and ALS/FTLD (Zu et al., 2011; Ash et al., 2013; Mori et al., 2013; Todd et al., 2013). Of particular note, two recent publications (Ash et al., 2013; Mori et al., 2013) have identified repeat associated non-AUG (RAN)-translation of the GGGGCC expanded repeats that cause ALS/FTLD into polypeptides that also form aggregates in affected tissues. However, these aggregates are confined to nerve cells and are absent from adjacent glial cells that are also involved in the pathology. On the other hand, the absence of visible aggregates does not prove the absence of toxic peptides.

These RAN translation results have suggested that an aggregate polypeptide analogous to polyglutamine could be neurotoxic in diseases where the causative repeat expansion cannot encode polyglutamine. However, this is doubtful in the cases of ALS and FTLD because of the observations that mutations in either of two RNA-binding proteins, FUS and TDP-43, can also cause disease (Rutherford et al., 2008; Van Langenhove et al., 2010). In individuals affected due to these mutations, no such expanded polyGly-Pro polypeptide is evident, therefore while polyGly-Pro may lead to subtle differences in pathology (Ash et al., 2013; Mori et al., 2013), it appears to play a modifying role at most.

Furthermore, inhibition of an RNA lariat debranching enzyme has recently been shown to suppress TDP-43 toxicity in ALS disease models (Armakola et al., 2012). These observations reinforce the view that RNA has a central role to play in this disease. While the role of such polypeptides in disease pathogenesis is unclear, for example, whether their aggregation may actually be protective rather than pernicious, they are a curious set of products driven by the unusual structure of expanded repeat RNAs. One possibility is that rather than the homopolymeric polypeptides themselves being toxic, the initiation of translation within the repeats could give rise to *N*-terminal truncated proteins devoid of upstream functional domains that could then act as dominant negative competitors for the full-length functionally intact proteins.

DOUBLE-STRANDED EXPANDED REPEAT RNA IS PATHOGENIC

Drosophila models of expanded repeat diseases have been described that specifically investigate the intrinsic toxicity of both translated and untranslated expanded repeat sequences (Lawlor et al., 2011; van Eyk et al., 2011, 2012; Samaraweera et al., 2013). In one study (Lawlor et al., 2011), a single line of *Drosophila* expressing untranslated CAG was identified with a marked degenerative phenotype (whereas multiple other random insertion lines of the same transgene had no such phenotype). Upon detailed characterization, this degenerative phenotype line was found to have the repeat transgene inserted into an endogenous gene (*cheerio*) in the opposite orientation to normal transcription.

Transcripts containing expanded repeats would, therefore, originate from both strands via bi-directional transcription. This finding coincided with numerous reports in the literature that expanded repeat disease loci are typically transcribed from both DNA strands (see Batra et al., 2010). Therefore, this *Drosophila* line mimicked a previously uncharacterized property of these disease genes. Bi-directional transcription was subsequently modeled in a controlled manner by co-expression from two different transgenes of expanded $rCAG_{\sim 100}$ together with $rCUG_{\sim 100}$ [giving rise to $rCAG.rCUG_{\sim 100}$ or double-strand RNA (dsRNA)] to produce repeat-containing dsRNA (Lawlor et al., 2011). Flies expressing dsRNA showed *Dicer*-dependent toxicity. Additionally dsRNA expression throughout the nervous system caused an age-dependent neurodegenerative phenotype. An abundance of $r(CAG)_7$ also implicated specific *Dicer* processing of the $rCAG.rCUG_{\sim 100}$ dsRNA as a pathogenic pathway in this model (Lawlor et al., 2011). Similar findings have also been reported in an independent *Drosophila* model (Yu et al., 2011). There are, with all animal models, caveats. In order to manifest a phenotype in the time frame of laboratory experiments, these *Drosophila* (and other animal) models employ copy numbers well in excess of those that cause pathology (after several decades) in some of these diseases. This is thought to be due to an inverse relationship between repeat copy number and age-at-onset, the basis of which could be somatic repeat instability over time (see Figure 2 in Richards, 2001 and Swami et al., 2009). Furthermore, the level of expression of the repeat RNAs required to give an early phenotype in animal models (Lawlor et al., 2011; Yu et al., 2011) may be well in excess of that of the endogenous human disease gene. Importantly, however, examination of HD patient samples (Bañez-Coronel et al., 2012) revealed the presence of the same $r(CAG)_7$ cleavage product seen in the *Drosophila* models, providing evidence in support of the activity of this pathway in HD pathogenesis.

In an effort to identify further components of expanded repeat RNA pathogenesis in *Drosophila*, microarray analyses of *Drosophila* expressing $rCAG.rCUG_{\sim 100}$ dsRNA have been undertaken (Samaraweera et al., 2013). Changes in transcription profiles revealed candidate pathways for mediating the resultant pathogenesis. Alterations in transcripts common to several pathways were detected, including components of inflammation and innate immunity. Hallmarks of immune activation, including elevated plasma tumor necrosis factor (TNF), appear prior to clinical symptoms of dominantly inherited expanded repeat human diseases (Moreau et al., 2005; Björkqvist et al., 2008). Therefore, the *Drosophila* model expressing $rCAG.rCUG_{\sim 100}$ dsRNA was utilized to test two key elements of immune activation – the *Toll* and autophagy pathways for their contribution to expanded repeat RNA pathogenesis. *Toll* signaling pathway was identified as essential for dsRNA pathogenesis and autophagy was found to reduce toxicity in this model (Samaraweera et al., 2013). Furthermore, multiple reports implicate glial cells in the pathology of expanded repeat diseases. Neurons are dependent upon glial cell function that includes the destruction and removal of the carcasses of dead neurons. The $rCAG.rCUG_{\sim 100}$ dsRNA was found to impact nerve cell function even when exclusively expressed in glial cells (Samaraweera et al., 2013), providing evidence that dsRNA pathology in

Drosophila is, like the human expanded repeat diseases, non-cell autonomous (Ilieva et al., 2009; Furrer et al., 2011).

The requirement for *Toll* signaling pathway in this *Drosophila* model is intriguing. *Toll*-like receptors (TLRs) function in normal biology to protect an organism from infection by viruses and bacteria. They recognize foreign pathogen molecules including DNA and RNA through specific receptors (such as endosomal *TLR3*) and can distinguish these nucleic acids (as “non-self”) from the endogenous nucleic acids (“self”). Therefore, while the *rCAG.rCUG~100* dsRNA is being expressed endogenously in this *Drosophila* model, it is being recognized by the *Toll* signaling pathway as foreign or “non-self” – a recognition that then activates innate inflammatory regulatory pathways, ultimately leading to cell death.

PATHOGENIC MUTATIONS IN PROTEINS THAT FUNCTIONALLY INTERACT WITH RNA

While it can be difficult to ascribe specific functions to RNA in pathogenic pathways, there are some noteworthy instances of disease-causing mutations in proteins that functionally interact with RNAs. By implication, the RNAs that these proteins normally act upon are, therefore, likely contributors to and/or mediators of the relevant pathogenic process.

RNA-BINDING MOTIFS – THE RNAs THAT HAVE THEM AND THE PROTEINS THAT RECOGNIZE THEM

Recent discoveries regarding the importance of RNA–protein recognition in disease pathogenesis have led to a renewed interest in the role that these interactions play in biological processes. While they have long been recognized as key regulators of gene expression, only a small fraction have been functionally characterized. A recent compendium of RNA-binding motifs (Ray et al., 2013) highlighted both the significance and scope of these interactions. The human genome encodes at least 400 known or predicted RNA-binding proteins with a diverse array of RNA sequence-binding motifs. Indeed the number of such human RNA-binding proteins appears to be much higher than this, with 860 identified in HeLa cells alone (Castello et al., 2012). The scope and specificity of RNA recognition is determined both by the number and variety of RNA-binding proteins and by the number and variety of RNA-sequence motifs that they bind.

FRAGILE X SYNDROME IS DUE TO LOSS OF RNA-BINDING PROTEIN FUNCTION

Fragile X syndrome is a striking example of the role of an RNA-binding protein in human disease. FRAXA is due to the expanded CGG repeat that is responsible for the *FRAXA* rare, folate-sensitive chromosomal fragile site (Kremer et al., 1991), located in the 5'UTR of the *FMR1* gene (Verkerk et al., 1991). Expansion of the repeat beyond ~230 copies results in inactivation of the gene and consequent loss of encoded FMRP (fragile X mental retardation protein) function (Pieretti et al., 1991). The FMRP is an RNA-binding protein with KH- and RGG-binding motifs (Ashley et al., 1993). The loss of function of this protein is responsible for the clinical symptoms as rare cases of point mutation or deletion of the *FMR1* gene have similar clinical symptoms. Indeed one of these pathogenic point mutations is at a highly conserved amino

acid in a KH domain of FMRP highlighting the significance of the role of RNA interaction in FMRP function (De Boulle et al., 1993). The FMRP has an impact on the translation of the mRNAs with which it interacts (Darnell et al., 2001) and, therefore, its absence leads to the dysregulation of the translation of these specific mRNAs. This is thought to be the proximal cause of the symptoms of FRAXA.

THE INTRIGUING PATHOGENESIS OF AICARDI–GOUTIÈRES SYNDROME

Aicardi–Goutières syndrome (AGS) is a genetically heterogeneous disorder that is due (at least in a substantial proportion of cases) to the mutation of various nucleic acid-metabolizing enzymes, including various subunits of ribonuclease H2 or the RNA-editing enzyme ADAR1 (see Crow and Rehwinkel, 2009 and OMIM #225750). AGS is characterized, in its more severe forms, by severe neurological dysfunction in infancy that includes progressive microcephaly, spasticity, dystonic posturing, profound psychomotor retardation, and often death in early childhood (OMIM #225750) (**Figure 3**). In its milder forms, these neurological symptoms are diminished or even absent, but peripheral symptoms outside the nervous system are common to the phenotypic spectrum and include thrombocytopenia, hepatosplenomegaly, and elevated hepatic transaminases along with intermittent fever. Chilblains are also a typical feature. Together these symptoms demonstrate phenotypic overlap both with systemic lupus erythematosus and with the sequelae of congenital infection (Crow and Rehwinkel, 2009). The disease, therefore, appears to be due to defects in the processes that remove and/or modify endogenous nucleic acids. These endogenous unmodified nucleic acids then accumulate and are sensed as “non-self” by TLRs, that, in turn, activate innate inflammatory regulatory pathways. This bears a striking resemblance to mechanisms we have identified as responsible for dsRNA pathogenesis in the *Drosophila* model of expanded repeat neurodegenerative diseases.

HYPOTHESIS

EXPANDED REPEAT RNAs AS PATHOGENIC AGENTS BY *TOLL* “SELF” RNA RECOGNITION

Repeat RNA sequences represent a pivotal point of potential weakness in processes that utilize RNA–protein recognition, as the repeat RNA sequence will harbor either a paucity or excess of sequence-binding motifs. Expansion of repeat RNA sequences, therefore, clearly has the potential to give rise to too much or too little of an interaction that is a rate-limiting factor in a crucial biological process. RNA modification is one process that is sequence motif-dependent and known to be key to the distinction between “self” and “non-self” by components of the innate immune pathways. Indeed, it has been shown that exogenous “non-self” RNAs require *in vitro* modification in order to escape innate immune recognition and activation when transferred *in vivo* (Warren et al., 2007; Pan, 2013). The exposure of the innate immune activators to unmodified nucleic acids, including RNA, appears to be the proximal cause of AGS. We, therefore, hypothesize that this provides a clear molecular mechanism for the ability of expanded repeat RNA sequences, through their paucity of RNA modification, to initiate pathogenesis in the dominantly inherited, expanded repeat neurodegenerative diseases (**Figure 4**).

Aicardi-Goutières Syndrome (AGS) – Mutations in RNA Metabolizing Genes in Bold

Locus – mutated gene

AGS1 – TREX1	– DNA 3' to 5' exonuclease, prevents autoimmunity caused by endogenous retroelements.
AGS2 – RNASEH2B	– subunit B of the human ribonuclease H2 enzyme complex – ribonuclease H2 cleaves ribonucleotides from RNA:DNA duplexes.
AGS3 – RNASEH2C	– subunit C of the human ribonuclease H2 enzyme complex.
AGS4 – RNASEH2A	– subunit A of the human ribonuclease H2 enzyme complex.
AGS5 – SAMHD1	– converts deoxynucleoside triphosphates to constituent deoxynucleoside and inorganic triphosphate. Restriction factor that renders human dendritic and myeloid cells largely refractory to HIV-1 infection.
AGS6 – ADAR1	– converts adenosine to inosine in double strand RNA

FIGURE 3 | Pathogenic mutations in Aicardi-Goutières syndrome. Mutations in genes in at least six distinct loci are able give rise to the constellation of symptoms that defines Aicardi-Goutières syndrome. Four of these (AGS2, AGS3, AGS4, and AGS6) are in genes that encode RNA-metabolizing proteins. The remaining two that have been identified

(AGS1 and AGS5) are also in enzymes that have roles in nucleic acid metabolism. Deficiencies in any one of these enzymes are thought to result in the accumulation of endogenous nucleic acids that are sensed as “non-self” by Toll-like receptors, that in turn activate innate inflammatory pathways (Crow and Rehwinkel, 2009).

Expanded Repeat RNAs as Pathogenic Agents by Toll ‘self’ RNA recognition

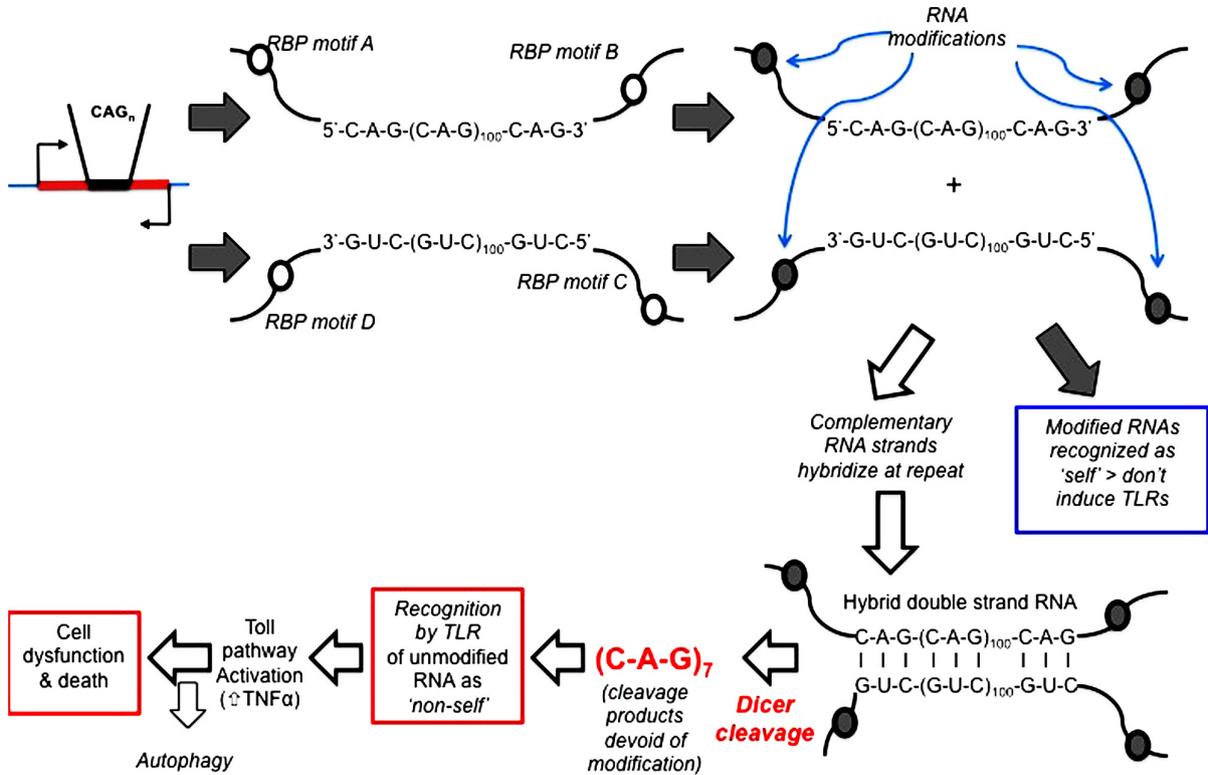


FIGURE 4 | Hypothesis: expanded repeat neurodegenerative diseases are caused by the TLR recognition (and resultant innate inflammatory response) of repeat RNA as “non-self” due to their paucity of “self” modification that is exposed upon Dicer processing of double-strand RNA. Open circles represent sequence motifs for RNA modifying proteins; filled circles represent the modification of RNA at these specific sequence motifs (e.g., by methylation or A > I editing). Dicer is required for pathology in

the *Drosophila* model and cleaves long high copy number repeat RNA down to 21mers [mainly *r(CAG)7* mers; Lawlor et al., 2011]. These *r(CAG)7* mers are, therefore, unmodified and recognized by TLRs as “non-self.” Toll-like receptor pathways (most probably the endosomal TLR3 receptor) are required for pathology (Samaraweera et al., 2013), through activation of the innate inflammatory pathway. Autophagy reduces pathology, possibly by metabolizing *r(CAG)7* mers.

ACTIVITY OF *TOLL* “SELF” RNA RECOGNITION IN NEURODEGENERATIVE DISEASES

Double strand expanded repeat RNA pathology has been modeled in *Drosophila*. What evidence is there that this pathway of TLR recognition of expanded repeat RNA and subsequent activation of the innate inflammatory cascade is active in the human dominantly inherited neurodegenerative diseases due to expansion of repeat sequences?

One of the key steps in dsRNA pathology is the generation of *r(CAG)₇* 21mers from the much greater copy number double strand repeat RNA by *Dicer*. This *r(CAG)₇* 21mer has been identified in the brain RNA of individuals affected with HD (Bañez-Coronel et al., 2012). The activity of *Dicer* is crucial to the observed pathology in the *Drosophila* model (Lawlor et al., 2011) and, therefore, it would appear that this step is a likely proximal event in the observed phenotype. The appearance of *r(CAG)₇* 21mers in HD brain is therefore an important “molecular hallmark” of this pathway and a key indicator of its activity in the human disease. Another, albeit less direct, indicator of this pathway is seen in the increased activity of components of the innate inflammatory response mechanism in human diseases associated with expanded repeats. Elevated TNF is seen in the *Drosophila* model as one read-out of innate immune activation (Samaraweera et al., 2013) and both TNF and various interleukins (i.e., IL-4, IL-5, IL-6, IL-8, and IL-10) have been found to be elevated in people affected with the repeat expansion responsible for HD even before clinical manifestation of the disease (Björkqvist et al., 2008). Another indicator of innate immune activation in HD is the abnormal peripheral chemokine profile that has been observed in HD (Wild et al., 2011). Various reports indicate activation of innate adaptive immunity via TLR signaling in ALS (Casula et al., 2011; Sta et al., 2011) – a disease that has recently been found, at least in a proportion of instances, to be also due to an expanded repeat (DeJesus-Hernandez et al., 2011; Renton et al., 2011).

CONCLUSION

A growing body of literature indicates a consistent association between innate immunity, neuroinflammation and

neurodegeneration (Shastri et al., 2013). Where there are exogenous causes (e.g., trauma or infection), activation of the TLR pathway can be attributed to an external agent (e.g., bacterial lipopolysaccharide or viral RNA); however, a causal basis for this relationship has not been clear when there is an endogenous basis to the disease, e.g., expansion of a repeat sequence beyond a pathogenic threshold. Recognition by the *Toll* receptor pathway of expanded repeat RNA as “non-self” and consequent activation of the innate immune inflammatory cascade provides a mechanism and a common pathogenic pathway for the neurodegenerative diseases due to expanded repeats. This new understanding, once proven in the relevant human diseases, will provide new targets for intervention and ultimately, we hope, therapeutic targets for drugs to delay onset and/or alleviate disease progression.

AUTHOR CONTRIBUTIONS

Robert I. Richards drafted the initial version of the manuscript, including the hypothesis, then edited in the additions and changes made by the other co-authors. Saumya E. Samaraweera and Clare L. van Eyk provided original unpublished data on which the manuscript and its hypothesis is based, as well as contributions to the development of the hypothesis, the text and figures. Louise V. O’Keefe contributed to the development of the hypothesis and additional text and revision of the manuscript. Catherine M. Suter contributed to information on RNA-binding proteins, the development of the hypothesis and content of the text and figures.

ACKNOWLEDGMENTS

This work is funded in part by a Project Grant (627183) from the National Health and Medical Research Council of Australia and a post-doctoral fellowship to Clare L. van Eyk from the National Ataxia Foundation (USA). Robert I. Richards wishes to thank Dan Kastner, Ivona Aksentijevich, and Massimo Gadina (NIH) for valuable discussions and Sarah Robertson, Amanda Choo, Danielle Fornarino, and Cheng Shou Lee for helpful and constructive criticism of drafts of this manuscript.

REFERENCES

- Armakola, M., Higgins, M. J., Figley, M. D., Barmada, S. J., Scarborough, E. A., Diaz, Z., et al. (2012). Inhibition of RNA lariat debranching enzyme suppresses TDP-43 toxicity in ALS disease models. *Nat. Genet.* 44, 1302–1309. doi: 10.1038/ng.2434
- Ash, P. E., Bieniek, K. F., Gendron, T. F., Caulfield, T., Lin, W. L., DeJesus-Hernandez, M., et al. (2013). Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* 77, 1–8. doi: 10.1016/j.neuron.2013.02.004
- Ashley, C. T., Jr., Wilkinson, K. D., Reines, D., and Warren, S. T. (1993). FMR1 protein: conserved RNP family domains and selective RNA-binding. *Science* 262, 563–566. doi: 10.1126/science.7692601
- 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
- Batra, R., Charizanis, K., and Swanson, M. S. (2010). Partners in crime: bidirectional transcription in unstable microsatellite disease. *Hum. Mol. Genet.* 19, R77–R82. doi: 10.1093/hmg/ddq132
- Björkqvist, M., Wild, E. J., Thiele, J., Silvestroni, A., Andre, R., Lahiri, N., et al. (2008). A novel pathogenic pathway of immune activation detectable before clinical onset in Huntington’s disease. *J. Exp. Med.*, 205, 1869–1877. doi: 10.1084/jem.20080178
- Carroll, J. M., Quaid, K. A., Stone, K., Jones, R., Schubert, F., and Griffith, C. B. (2013). Two is better than one: a case of homozygous myotonic dystrophy type 1. *Am. J. Med. Genet.* 161, 1763–1767. doi: 10.1002/ajmg.a.35967
- Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B. M., Strein, C., et al. (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* 149, 1393–1406. doi: 10.1016/j.cell.2012.04.031
- Casula, M., Iyer, A. M., Spliet, W. G., Anink, J. J., Steentjes, K., Sta, M., et al. (2011). Toll-like receptor signaling in amyotrophic lateral sclerosis spinal cord tissue. *Neuroscience* 179, 233–243. doi: 10.1016/j.neuroscience.2011.02.001
- Crow, Y. J., and Rehwinkel, J. (2009). Aicardi–Goutières syndrome and related phenotypes: linking nucleic acid metabolism with autoimmunity. *Hum. Mol. Genet.* 18, R130–R136. doi: 10.1093/hmg/ddp293
- Darnell, J. C., Jensen, K. B., Jin, P., Brown, V., Warren, S. T., and Darnell, R. B. (2001). Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function. *Cell* 107, 489–499. doi: 10.1016/S0092-8674(01)00566-9
- De Boule, K., Verkerk, A. J. M. H., Reyniers, E., Vits, L., Hendrickx, J., Van Roy, B., et al. (1993). A point mutation in the FMR-1 gene associated with fragile X mental retardation. *Nat. Genet.* 3, 31–35. doi: 10.1038/ng0193-31

- 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 non-coding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* 72, 245–256. doi: 10.1016/j.neuron.2011.09.011
- Elden, A. C., Kim, H. J., Hart, M. P., Chen-Plotkin, A. S., Johnson, B. S., Fang, X., et al. (2010). Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature* 466, 1069–1075. doi: 10.1038/nature09320
- Fiszer, 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
- Furrer, S. A., Mohanachandran, M. S., Waldherr, S. M., Chang, C., Damian, V. A., Sopher, B. L., et al. (2011). Spinocerebellar ataxia type 7 cerebellar disease requires the coordinated action of mutant ataxin-7 in neurons and glia, and displays non-cell-autonomous Bergmann glia degeneration. *J. Neurosci.* 31, 16269–16278. doi: 10.1523/JNEUROSCI.4000-11.2011
- Ilieva, H., Polymenidou, M., and Cleveland, D. W. (2009). Non-cell autonomous toxicity in neurodegenerative disorders: ALS and beyond. *J. Cell Biol.* 187, 761–772. doi: 10.1083/jcb.200908164
- Jones, K., Wei, C., Iakova, P., Bugiardi, E., Schneider-Gold, C., Meola, G., et al. (2012). GSK3 β mediates muscle pathology in myotonic dystrophy. *J. Clin. Invest.* 122, 4461–4472. doi: 10.1172/JCI64081
- Kobayashi, H., Abe, K., Matsuura, T., Ikeda, Y., Hitomi, T., Akechi, Y., et al. (2011). Expansion of intronic GGCCTG hexanucleotide repeat in NOP56 causes SCA36, a type of spinocerebellar ataxia accompanied by motor neuron involvement. *Am. J. Hum. Genet.* 89, 121–130. doi: 10.1016/j.ajhg.2011.05.015
- Kremer, E., Pritchard, M., Lynch, M., Yu, S., Holman, K., Warren, S., et al. (1991). DNA instability at the fragile X maps to a trinucleotide repeat sequence p(CCG) $_n$. *Science* 252, 1711–1714. doi: 10.1126/science.1675488
- La Spada, A. R., and Taylor, J. P. (2010). Repeat expansion disease: progress and puzzles in disease pathogenesis. *Nat. Rev. Genet.* 11, 247–258. doi: 10.1038/nrg2748
- 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
- Lawlor, K. T., O'Keefe, L. V., Samaraweera, S., van Eyk, C., McLeod, C. J., Maloney, C., 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
- Lawlor, K. T., O'Keefe, L. V., Samaraweera, S. E., van Eyk, C. L., and Richards, R. I. (2012). Ubiquitous expression of CUG or CAG trinucleotide repeat RNA causes common morphological defects in a *Drosophila* model of RNA-mediated pathology. *PLoS ONE* 7:e38516. doi: 10.1371/journal.pone.0038516
- 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
- Mankodi, A., Takahashi, M. P., Jiang, H., Beck, C. L., Bowers, W. J., Moxley, R. T., et al. (2002). Expanded CUG repeats trigger aberrant splicing of CIC-1 chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy. *Mol. Cell* 10, 35–44. doi: 10.1016/S1097-2765(02)00563-4
- McLeod, C., O'Keefe, L., and Richards, R. I. (2005). The pathogenic agent in *Drosophila* models of 'polyglutamine' diseases. *Hum. Mol. Genet.* 14, 1041–1048. doi: 10.1093/hmg/ddi096
- Moreau, C., Devos, D., Brunaud-Danel, V., Defebvre, L., Perez, T., Destée, A., et al. (2005). Elevated IL-6 and TNF-alpha levels in patients with ALS: inflammation or hypoxia? *Neurology* 65, 1958–1960. doi: 10.1212/01.wnl.0000188907.97339.76
- Mori, K., Weng, S. M., Arzberger, T., May, S., Rentzsch, K., Kremmer, E., et al. (2013). The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTL/ALS. *Science* 339, 1335–1338. doi: 10.1126/science.1232927
- Orr, H. T. (2011). FTD and ALS: genetic ties that bind. *Neuron* 72, 189–190. doi: 10.1016/j.neuron.2011.10.001
- Pan, T. (2013). N6-methyl-adenosine modification in messenger and long non-coding RNA. *Trends Biochem. Sci.* 38, 204–209. doi: 10.1016/j.tibs.2012.12.006
- Pieretti, M., Zhang, F., Fu, Y.-H., Warren, S. T., Oostra, B. A., Caskey, C. T., et al. (1991). Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell* 66, 817–822. doi: 10.1016/0092-8674(91)90125-I
- Polymenidou, M., and Cleveland, D. W. (2011). The seeds of neurodegeneration: prion-like spreading in ALS. *Cell* 147, 498–508. doi: 10.1016/j.cell.2011.10.011
- Ranum, L. P., and Day, J. W. (2004). Myotonic dystrophy: RNA pathogenesis comes into focus. *Am. J. Hum. Genet.* 74, 793–804. doi: 10.1086/383590
- Ray, D., Kazan, H., Cook, K. B., Weirauch, M. T., Najafabadi, H. S., Li, X., et al. (2013). A compendium of RNA-binding motifs for decoding gene regulation. *Nature* 499, 172–177. doi: 10.1038/nature12311
- 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
- Richards, R. I. (2001). Dynamic mutations: a decade of unstable expanded repeats in human disease. *Hum. Mol. Genet.* 10, 2187–2194. doi: 10.1093/hmg/10.20.2187
- Richards, R. I., and Sutherland, G. R. (1992). Dynamic mutations: a new class of mutations causing human disease. *Cell* 70, 709–712. doi: 10.1016/0092-8674(92)90302-S
- Rutherford, N. J., Zhang, Y. J., Baker, M., Gass, J. M., Finch, N. A., Xu, Y. F., et al. (2008). Novel mutations in TARDBP (TDP-43) in patients with familial amyotrophic lateral sclerosis. *PLoS Genet.* 4:e1000193. doi: 10.1371/journal.pgen.1000193
- Samaraweera, S. E., O'Keefe, L. V., Price, G. R., Venter, D. J., and Richards, R. I. (2013). Distinct roles for Toll and autophagy pathways in double-stranded RNA toxicity in a *Drosophila* model of expanded repeat neurodegenerative diseases. *Hum. Mol. Genet.* 22, 2811–2819. doi: 10.1093/hmg/ddt130
- Sato, N., Amino, T., Kobayashi, K., Asakawa, S., Ishiguro, T., Tsunemi, T., et al. (2009). Spinocerebellar ataxia type 31 is associated with "inserted" penta-nucleotide repeats containing (TGGAA) $_n$. *Am. J. Hum. Genet.* 85, 544–557. doi: 10.1016/j.ajhg.2009.09.019
- Shastri, A., Bonfati, D. M., and Kishore, U. (2013). Innate immunity and neuroinflammation. *Mediators Inflamm.* 2013, 342931. doi: 10.1155/2013/342931
- Sta, M., Sylva-Steenland, R. M., Casula, M., de Jong, J. M., Troost, D., Aronica, E., et al. (2011). Innate and adaptive immunity in amyotrophic lateral sclerosis: evidence of complement activation. *Neurobiol. Dis.* 42, 211–220. doi: 10.1016/j.nbd.2011.01.002
- Swami, M., Hendricks, A. E., Gillis, T., Massood, T., Mysore, J., Myers, R. H., et al. (2009). Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset. *Hum. Mol. Genet.* 18, 3039–3047. doi: 10.1093/hmg/ddp242
- 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
- van Eyk, C., McLeod, C. J., O'Keefe, L. V., and Richards, R. I. (2012). Comparative Toxicity of polyglutamine, polyalanine and polyisoleucine tracts in *Drosophila* models of expanded repeat disease. *Hum. Mol. Genet.* 21, 536–547. doi: 10.1093/hmg/ddr487
- van Eyk, C. L., O'Keefe, L. V., Lawlor, K. T., Samaraweera, S. E., McLeod, C. J., Price, G. R., et al. (2011). Perturbation of the Akt/Gsk3- β signalling pathway is common to *Drosophila* expressing expanded untranslated CAG, CUG and AUUCU repeat RNAs. *Hum. Mol. Genet.* 20, 2783–2794. doi: 10.1093/hmg/ddr177
- Van Langenhove, T., van der Zee, J., Sleegers, K., Engelborghs, S., Vandenberghe, R., Gijssels, I., et al. (2010). Genetic contribution of FUS to frontotemporal lobar degeneration. *Neurology* 74, 366–371. doi: 10.1212/WNL.0b013e3181ccc732
- Verkerk, A. J., Pieretti, M., Sutcliffe, J. S., Fu, Y. H., Kuhl, D. P., Pizzuti, A., et al. (1991). Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* 65, 905–914. doi: 10.1016/0092-8674(91)90397-H
- Warren, L., Manos, P. D., Ahfeldt, T., Loh, Y. H., Li, H., Lau, F., et al. (2007). Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell* 7, 618–630. doi: 10.1016/j.stem.2010.08.012
- Wild, E., Magnussen, A., Lahiri, N., Krus, U., Orth, M., Tabrizi, S. J., et al. (2011). Abnormal peripheral chemokine profile in Huntington's disease. *PLoS Curr.* 3:RRN1231. doi: 10.1371/currents.RRN1231

- Yu, S., Kremer, E., Pritchard, M., Lynch, M., Nancarrow, J., Baker, E., et al. (1991). The fragile X genotype is characterized by an unstable region of DNA. *Science* 252, 1179–1182. doi: 10.1126/science.252.5009.1179
- 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
- Zu, T., Gibbens, B., Doty, N. S., Gomes-Pereira, M., Huguette, 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.
- Received: 22 July 2013; paper pending published: 04 August 2013; accepted: 14 August 2013; published online: 05 September 2013.
- Citation: Richards RI, Samaraweera SE, van Eyk CL, O'Keefe LV and Suter CM (2013) RNA pathogenesis via Toll-like receptor-activated inflammation in expanded repeat neurodegenerative diseases. *Front. Mol. Neurosci.* 6:25. doi: 10.3389/fnmol.2013.00025
- This article was submitted to the journal *Frontiers in Molecular Neuroscience*.
- Copyright © 2013 Richards, Samaraweera, van Eyk, O'Keefe and Suter. 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.