Human Molecular Genetics, 2005, Vol. 14, No. 6 **873–883**
doi:10.1093/hmg/ddi080
Advance Access published on February 9, 2005

Myotonic dystrophy associated expanded CUG repeat muscleblind positive ribonuclear foci are not toxic to *Drosophila*

Jonathan M. Houseley[1,†], Zongsheng Wang[1], Graham J. R. Brock[1,‡], Judith Soloway[1], Ruben Artero[2], Manuel Perez-Alonso[2], Kevin M. C. O'Dell[1] and Darren G. Monckton[1,*]

[1]Institute of Biomedical and Life Sciences, University of Glasgow, Anderson College Complex, 56 Dumbarton Road, Glasgow G11 6NU, UK and [2]Department of Genetics, University of Valencia, Dr Moliner 50, 46100 Burjasot, Spain

Received November 8, 2004; Revised January 21, 2005; Accepted February 1, 2005

Myotonic dystrophy type 1 is an autosomal dominant disorder associated with the expansion of a CTG repeat in the 3′ untranslated region (UTR) of the *DMPK* gene. Recent data suggest that pathogenesis is predominantly mediated by a gain of function of the mutant transcript. In patients, these expanded CUG repeat-containing transcripts are sequestered into ribonuclear foci that also contain the muscleblind-like proteins. To provide further insights into muscleblind function and the pathogenesis of myotonic dystrophy, we generated *Drosophila* incorporating CTG repeats in the 3′-UTR of a reporter gene. As in patients, expanded CUG repeats form discrete ribonuclear foci in *Drosophila* muscle cells that co-localize with muscleblind. Unexpectedly, however, foci are not observed in all cell types and muscleblind is neither necessary nor sufficient for their formation. The foci are dynamic transient structures with short half-lifes that do not co-localize with the proteasome, suggesting they are unlikely to contain mis-folded proteins. However, they do co-localize with non-A, the human orthologs of which are implicated in both RNA splicing and attachment of dsRNA to the nuclear matrix. Muscleblind is also revealed as having a previously unrecognized role in stabilizing CUG transcripts. Most interestingly, *Drosophila* expressing $(CUG)_{162}$ repeats has no detectable pathological phenotype suggesting that in contrast to expanded polyglutamine-containing proteins, neither the expanded CUG repeat RNA nor the ribonuclear foci are directly toxic.

INTRODUCTION

Myotonic dystrophy (DM) is the most common adult onset muscular dystrophy (1). It is an autosomal dominant disorder characterized by progressive myotonia and muscle weakness, although symptoms are not limited to muscle. Two transcribed, but not translated, mutations cause DM. Approximately 98% of DM patients carry the type 1 DM1 mutation, an expansion of CTG repeats in the 3′-UTR of the *DMPK* gene (2). Unaffected individuals in the general population usually have fewer than 30 CTG repeats. Late onset DM1 patients typically inherit from 60 to 90 repeats and often only present with cataracts. Adult onset cases usually have from 100 to 500 CTG repeats and develop myopathy, myotonia, cardiac conduction defects, insulin intolerance, infertility, behavioural abnormalities, including apathy and hypersomnia, and reduced life expectancy. Congenitally affected children typically inherit more than 700 CTG repeats and present with severe hypotonia at birth, are mentally retarded and go on to develop the symptoms typical of adult onset patients (1). The vast majority of DM patients who do not have the DM1 mutation have DM type 2 (DM2), caused by an expansion of CCTG repeats in intron 1 of the *ZNF9* gene (3).

Various hypotheses have been proposed to explain how untranslated mutations can lead to a dominant pathogenic phenotype (4). Although it is probable that defects in the levels of the protein products of flanking genes may contribute in part to the symptoms, two strong lines of evidence favour a gain of function of the expanded repeat RNA. First, DM1 and DM2 have similar phenotypes despite the two mutations

*To whom correspondence should be addressed. Tel: +44 1413306213; Fax: +44 1413306871; Email: d.monckton@bio.gla.ac.uk
†Present address: The Wellcome Trust Centre for Cell Biology, Michael Swann Building, Mayfield Road, Edinburgh EH9 3JR, UK.
‡Present address: Target and Drug Discovery, Ordway Research Institute, 150 New Scotland Avenue, Albany, NY 12208, USA.

© The Author 2005. Published by Oxford University Press. All rights reserved.
For Permissions, please email: journals.permissions@oupjournals.org

Downloaded from http://hmg.oxfordjournals.org/ at BBSRC on July 7, 2015

being related only by the nature of the transcribed, but untranslated, CTG/CCTG repeat. Secondly, mice expressing expanded CUG repeats in the 3′-UTR of an unrelated transgene develop myotonia and a DM-like myopathy (5).

Downstream pathology in DM is linked with splicing defects in a number of genes. Most convincingly, mis-splicing of the *chloride channel subunit 1 (CLC1)* and *insulin receptor* transcripts almost certainly underlie the observed myotonia (6) and insulin intolerance (7). The pathways that link these splicing defects to the primary CTG expansion are not yet completely understood but appear to involve two classes of proteins that can bind CUG repeats: the CUG-BP1 and ETR-3-like factors (CELF) and the muscleblind-like (MBNL) proteins (8). Both classes of proteins are regulators of alternative splicing with antagonistic effects on a subset of alternatively spliced genes. Consistent with a direct role for CUG-BP1 in DM pathogenesis, DM splicing defects are mirrored in normal cells over-expressing CUG-BP1 and nuclear levels of CUG-BP1 are increased in DM patient cells (9). Similarly, consistent with a direct role for the MBNLs in DM pathogenesis, mice lacking *Mbnl1* develop myopathy and myotonia and splicing defects in the *Clc1* transcript (10).

Precisely how the CELF and MBNL protein functions are perturbed in patient cells is not completely understood, but is assumed to be related to the observation that expanded CUG repeat RNA is trapped in discrete foci in the nuclei of patient cells (11). Although these ribonuclear foci do not contain CUG-BP1 (12), the MBNL proteins bind dsCUG RNA and co-localize with the ribonuclear foci, strongly supporting a role for MBNL titration in the pathogenic process (13,14). Nevertheless, mice homozygous null for *Mbnl1* are born healthy and do not present with the congenital form of DM. Thus, it remains unclear as to what extent MBNL titration contributes toward pathogenesis and, importantly, whether expanded CUG repeat RNA or ribonuclear foci are inherently noxious and have a direct toxic effect over and above dysregulation of alternative splicing. A more general toxic effect of expanded CUG repeat RNA might be mediated by the sequestration of transcription factors, as has been recently proposed (15), and is observed for highly toxic polyglutamine expansions (16).

The rate of progress in understanding fundamental mechanisms in DM is restricted by the complexity of analysing patient samples, the inherent limitations of cell culture models and the relative difficulty of generating additional mouse models. Significantly, a number of triplet repeat disorders have been successfully modelled in *Drosophila*, providing critical new insights into the molecular pathogenesis of the disease process (17–20). We have, therefore, created a *Drosophila* model of DM and provided insights into muscleblind function by expressing CTG repeats in the 3′-UTR of a marker gene.

RESULTS

Expanded CUG repeat tract expressing *Drosophila* are viable

To explore CUG repeat-mediated DM pathogenesis, *DM1* alleles of 11, 48, 56 and 162 CTG repeats flanked by ~100 bp of the human *DMPK* 3′-UTR were cloned into the 3′-UTR of a GFP marker gene under the control of a UAS promoter and used to generate transgenic *Drosophila* (Fig. 1 and Supplementary Material, Table S1). Preliminary observations revealed that multiple lines of *Drosophila* expressing CTG repeat transgenes either ubiquitously, under the control of the *da.G32*-GAL4 driver, or specifically in muscle, using the 24B-GAL4 driver, expressed high levels of GFP, but developed normally, could walk, jump and fly, could mate and were fertile.

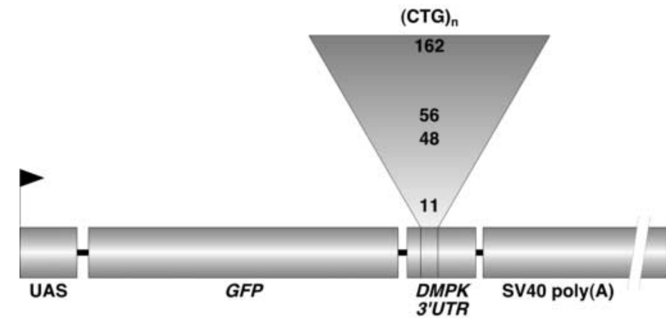

Figure 1. The myotonic dystrophy $(CTG)_n$ transgene. *DM1* alleles of 11, 48, 56 and 162 CTG repeats and 27 bp 5′ and 72 bp 3′ of the unique sequence DNA from the 3′-UTR of the human *DMPK* gene were cloned into the 3′-UTR of a *GFP* marker gene as part of a standard pP[UAST] *Drosophila* P-element transgene construct (42,43). RT–PCR and DNA sequencing of UAS driven transgenic transcripts confirmed the integrity of the transgenes and appropriate SV40 poly(**A**) mediated termination and polyadenylation (data not shown).

Expanded repeat RNA forms ribonuclear foci in *Drosophila*

A hallmark of DM pathogenesis is the presence of CUG/CCUG repeat ribonuclear foci in the nuclei of cells expressing expanded repeat RNA. Whole mount *in situ* hybridization using a Cy3 labelled $(CAG)_{10}$ probe in *Drosophila* ubiquitously expressing the $(CTG)_{162}$ transgene revealed ribonuclear foci in muscle cell nuclei at various developmental time points (Fig. 2A). Ribonuclear foci were not present in *Drosophila* expressing 11, 48 or 56 CUG repeat RNA. Identical results were obtained using a unique sequence riboprobe to *GFP* (Fig. 2B). As this probe can only bind to the transgene RNA in single copy, the inability to detect ribonuclear foci in lines with smaller repeat tracts is not an artefact of the number of copies of the repeat in the target sequence. Ribonuclear foci were not detectable in embryos but were observed in all muscle cells of first instar larvae. The number and intensity of foci increased throughout development with late Third instar larval muscle cells containing numerous intense foci. Ribonuclear foci were observed in all cells of larval muscles. The number of foci-positive cells decreased during pupation, concomitant with muscle remodelling. Nonetheless, ribonuclear foci were present at eclosion and persisted for the life of the adult fly. In adults, ribonuclear foci were restricted to abdominal, cranial and pleurosternal muscles (where they were present in all cells) and were absent in all leg and indirect

Downloaded from http://hmg.oxfordjournals.org/ at BBSRC on July 7, 2015

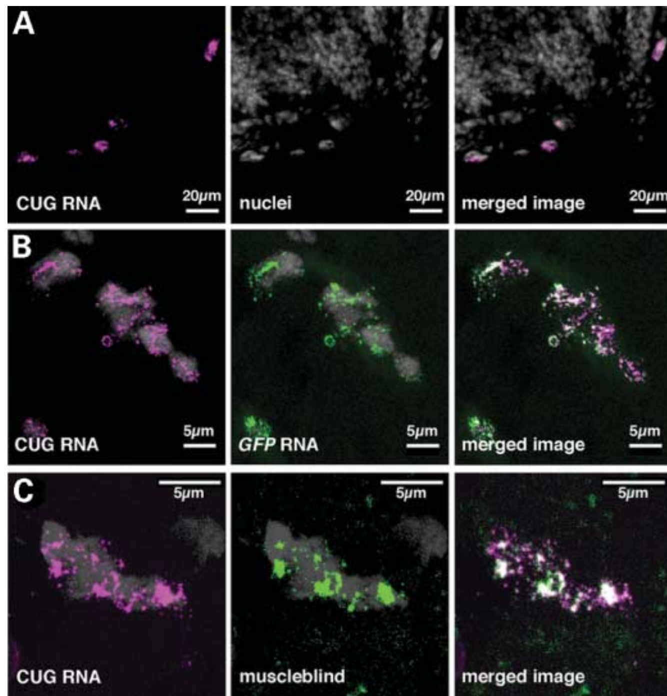

Figure 2. $(CUG)_{162}$ RNA forms ribonuclear foci that co-localize with muscleblind. (**A**) Muscle cell-specific ribonuclear foci. Ribonuclear foci (purple) are present in the nuclei (grey) of body wall muscle cells but are not present in the adjacent nuclei of neurons from the ventral nerve cord. (**B**) Ribonuclear foci contain *GFP* RNA as well as CUG repeat RNA. CUG repeat RNA (purple) and *GFP* RNA (green) co-localize (white) in muscle nuclei. (**C**) Muscleblind co-localizes with ribonuclear foci. Ribonuclear foci (purple) co-localize (white) with muscleblind (green) in muscle nuclei. All images are from third instar larvae ubiquitously expressing *GFP*$(CTG)_{162}$. Ribonuclear foci were detected using a Cy3-$(CAG)_{10}$ probe, *GFP* RNA with a DIG-labelled riboprobe, muscleblind with polyclonal sheep anti-muscleblind and nuclei with DAPI, which for clarity has been omitted from some of the merged images.

flight muscle cells. Interestingly, $(CUG)_{162}$-specific transgene RNA was also detected in the nuclei of all larval and adult salivary gland cells. However, rather than the discrete foci observed in muscle cells, the $(CUG)_{162}$-RNA accumulated in the inter-chromatin space (data not shown). Despite ubiquitous expression of the transgene, ribonuclear foci were not detectable in the brain, peripheral nervous system or any other internal organs (Fig. 2A). Strong transgenic GFP fluorescence was observed in tissue with (e.g. larval muscle) and without ribonuclear foci (e.g. larval neurons and adult indirect flight muscle) in both $(CTG)_{162}$ and $(CTG)_{11}$ lines. The expression level and pattern of all GAL4 driver lines were also tested with β-galactosidase reporter genes and levels of expression were shown to be indistinguishable in many ribonuclear-positive and ribonuclear-negative cell types. Thus, expanded CUG repeat RNA can exit the nucleus and be translated, even though some transcripts in some cell types are sequestered into ribonuclear foci. These data also reveal that the expression of expanded CUG repeat RNA alone is not sufficient to drive ribonuclear foci formation, indicating a requirement for other cell type-specific factors.

Expression of expanded CUG repeat RNA does not compromize locomotor activity, muscle histology, life span, eye development or female fertility

To determine whether flies expressing expanded repeat tracts that result in ribonuclear foci develop muscle defects, locomotor reactivity and spontaneous activity were assessed. No statistically significant expanded repeat length-specific differences in activity were observed and $(CUG)_n$ expressing flies maintained a normal circadian rhythm (Supplementary Material, Fig. S1 and Tables S2 and S3). Similarly, no perturbation of normal muscle structure was observed in haematoxylin and eosin stained sections or by polarized light microscopy. To determine if expression of expanded CUG arrays results in shortened life expectancy, *Drosophila* life span was assessed. Contrary to the expectation of a shortened life span (assuming a toxic effect of the transgene), the mean life span of $(CTG)_{162}$ repeat-expressing *Drosophila* was extended by $\sim$15% ($P < 0.001$, Supplementary Material, Fig. S2 and Table S4).

Although the presence of muscle specific ribonuclear foci appears not to be associated with a pathogenic effect in $(CUG)_{162}$ repeat-expressing flies, we sought to determine whether other tissues might be more vulnerable to $(CUG)_{162}$ repeat-induced toxicity. The *Drosophila* eye is a complex, developmentally sensitive organ and all reported studies of triplet repeat-mediated pathogenesis in *Drosophila* have demonstrated neurodegeneration in the eye and disruption and loss of ommatidia (17–20). However, expression of $(CTG)_{162}$ expanded repeat transgenes from two independent lines using two separate eye-specific GAL4 drivers (*ninaE*-GAL4 and *ey*-GAL4) failed to produce any disruption of ommatidial patterning even in 2-month-old *Drosophila* (data not shown).

In DM patients, CUG-BP1 activity is increased and associated with DM-specific RNA splicing errors (9). The two closest orthologues of *CUG-BP1* in the *Drosophila* genome are *aret* and *bruno2*. The role of *bruno2* is unknown, but the *aret* gene product, Bruno, is a well characterized gonad-specific protein for which both loss and over expression are detrimental to *Drosophila* oogenesis (21). Therefore, to determine if expression of expanded CUG tracts might interfere with Bruno function, fertility of female *Drosophila* expressing CUG repeat RNA ubiquitously was quantitatively assessed. These analyses revealed no defects in female fertility in expanded CUG repeat-expressing lines (Supplementary Material, Table S5).

Ribonuclear foci co-localize with muscleblind protein

In mammalian cells expressing large expanded CUG repeat arrays, the MBNL proteins co-localize with ribonuclear foci (14). Likewise, *Drosophila* muscleblind protein demonstrated excellent co-localization with $(CTG)_{162}$ mediated ribonuclear foci within cells (Fig. 2C). However, unlike the MBNLs, which are almost ubiquitously distributed in mammalian tissues (14), we only observed muscleblind in muscle, salivary gland and imaginal discs in wild-type larvae. Thus, there is a good correlation between ribonuclear foci formation and muscleblind expression. Nonetheless, ribonuclear foci were not observed in larval imaginal discs or adult indirect flight

Downloaded from http://hmg.oxfordjournals.org/ at BBSRC on July 7, 2015

muscle, despite the presence of muscleblind protein and GFP in these tissues.

In humans, MBNL1 is distributed throughout the cytoplasm and nucleus within wild-type muscle cells but is recruited to the ribonuclear foci in DM patient cells (13). This situation was replicated in *Drosophila* third instar larvae (Fig. 3A), but in wild-type and $(CTG)_{11}$ adults, muscleblind was clearly located in nuclear foci in the absence of expanded repeat RNA (Fig. 3B). This suggests that either muscleblind intracellular localization is developmentally controlled by other proteins or that developmentally regulated muscleblind isoforms differ in their localization. Four different *muscleblind* mRNA isoforms have been identified in *Drosophila*: *mbl-A*, *mbl-B*, *mbl-C* and *mbl-D* (22). Only *mbl-C* is expressed in adult *Drosophila*, whereas isoforms *mbl-A*, *-B*, *-C* and *-D* are all present in third instar larvae (Houseley *et al.*, unpublished data). These data suggest that isoform mbl-C may normally exist in foci, whereas other isoforms are normally distributed throughout the cell but can be titrated into ribonuclear foci. Northern blot and RT–PCR analyses revealed no detectable alteration in *muscleblind* splicing patterns or levels in larvae or adults expressing expanded CUG repeats (data not shown).

Ribonuclear foci are transient and can be muscleblind independent

The early expression of *muscleblind* is critical for terminal differentiation of muscles and *muscleblind* mutants are embryonic lethal (23). However, despite 24B-GAL4 driven embryonic expression of $(CTG)_{162}$ transgene transcripts, ribonuclear foci were not observed until the first larval instar, several hours after the point of co-expression with muscleblind. This implies that foci formation either requires additional cell type-specific factors or is a relatively slow process compared with the very rapid rate of *Drosophila* muscle differentiation in the embryo. To gain a better understanding of ribonuclear foci dynamics, $(CTG)_n$ transgenes were expressed in adult flies using a heat shock-activated GAL4 driver (Fig. 4). Transgene expression peaked within 8–24 h and was almost undetectable in whole flies by 4 days after heat shock. Ribonuclear foci were only observed in $(CTG)_{162}$ lines and were present in most cell types, except neurons. These experiments revealed ribonuclear foci as dynamic transient structures. Their number and intensity peaked at ~8 h and foci were mostly lost within 2–4 days, although their precise dynamics were highly cell-type dependent (Supplementary Material, Fig. S3). For example, ribonuclear foci were observed in follicle cells within 1 h, but remained detectable there for <48 h (Supplementary Material, Fig. S3A). In contrast, ribonuclear foci were not detected in indirect flight muscles until 24 h after heat-shock induction, where they could last for at least 10 days (Supplementary Material, Fig. S3B). By 19 days, post-heat-shock ribonuclear foci had been lost from all tissues in which they were absent pre-heat-shock. Western blot analysis of transgenic samples for GFP after heat shock demonstrated that ribonuclear foci formation did not appreciably deplete the quantity of mRNA available for translation (Fig. 4A). Very interestingly, it was also observed that ribonuclear foci were formed in most cell types in the absence of detectable

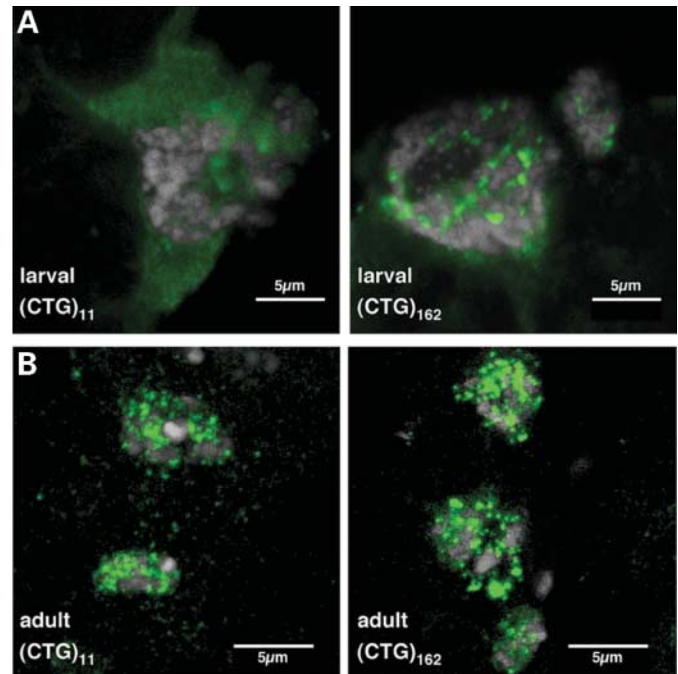

Figure 3. Intracellular muscleblind distribution varies during the *Drosophila* lifecycle. (**A**) Intracellular muscleblind distribution in larvae. In body wall muscle cells from third instar larvae, muscleblind (green) is widely dispersed in the cytoplasm and nucleus (grey) in *GFP*$(CTG)_{11}$, but relocalizes to discrete foci in *GFP*$(CTG)_{162}$ expressing animals. (**B**) Intracellular muscleblind distribution in adults. In abdominal muscle nuclei of adult wild-type *Drosophila* (data not shown), or flies expressing either *GFP*$(CTG)_{11}$ or *GFP*$(CTG)_{162}$, muscleblind (green) is present in nuclear foci. The nuclei from *GFP*$(CTG)_{162}$ expressing *Drosophila* also contain ribonuclear foci that co-localize with the muscleblind foci (data not shown).

muscleblind, as determined by standard immunohistochemistry (Fig. 4B). When detectable in other tissues, however, muscleblind did co-localize with the ribonuclear foci (Fig. 4C).

MBNL1 can promote ribonuclear foci formation in some cells

We did not observe ribonuclear formation in neuronal cells when driving ubiquitous expression of the $(CTG)_{162}$ transgene either continuously or transiently with heat shock. Likewise, muscleblind expression was not detected in neurons. To ascertain whether muscleblind is sufficient to promote ribonuclear foci formation, UAS transgenics were used to ectopically express *mbl-A*, *mbl-C* and *MBNL1* (24) in larval motorneurons using the D42-GAL4 driver, in combination with *GFP*$(CTG)_{48/162}$ transgenes. Neuronal ribonuclear foci were only observed in larvae expressing *MBNL1* and *GFP*$(CTG)_{162}$ (Fig. 5). Thus, human MBNL1 can drive ribonuclear foci formation in some cells, although the number of foci-positive neurons (~5–10%) was far fewer than the number of cells in which the GAL4 driver was active. In flies expressing *GFP*$(CTG)_{162}$ transgenes in muscle, ribonuclear foci are only observed in a subset of adult muscles and are absent in leg and indirect flight muscles. To further test whether mbl-A or mbl-C could stimulate foci formation, their over-expression was driven in adult muscle using

Downloaded from http://hmg.oxfordjournals.org/ at BBSRC on July 7, 2015

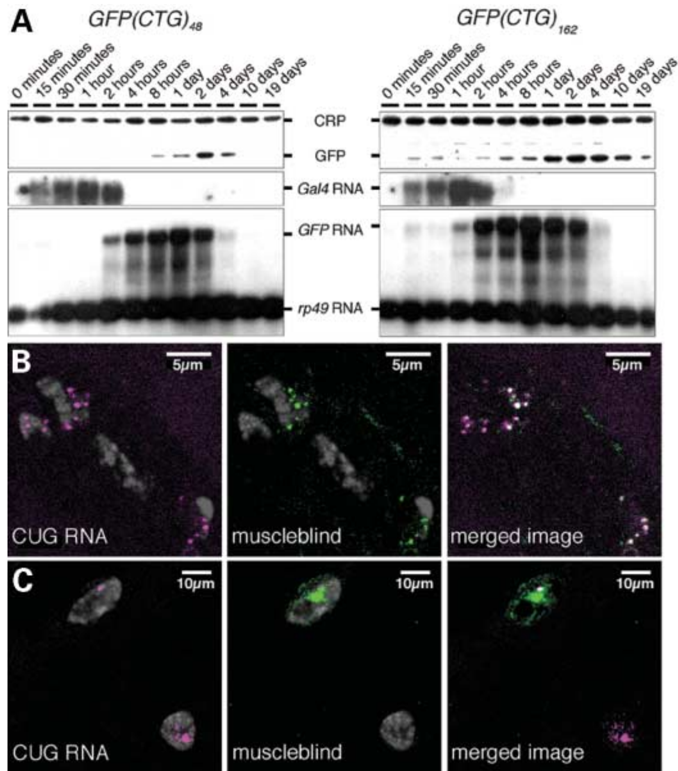

Figure 4. Ribonuclear foci are transiently induced after heat-shock activation. (**A**) The dynamics of transgene mRNA and protein levels after heat-shock induction. Adult *Drosophila* expressing *GFP*(CTG)$_{48}$ or *GFP*(CTG)$_{162}$ driven by *hs*-GAL4 were raised at 18°C, heat shocked for 30 min at 37°C and returned to 18°C. Protein and DNA samples were collected at various time points after heat-shock induction. Levels of transgene RNA and GFP were analyzed by northern (middle and bottom panels) and western blotting (top panel). Loading controls were *rp49* and CRP (a non-heat shock responsive cross-reacting protein) for northern and western blots, respectively. (**B**) Ribonuclear foci induced in *GFP*(CTG)$_{162}$ muscle nuclei by heat shock co-localize with muscleblind. Repeat RNA (purple) colocalises (white) with muscleblind (green) in muscle cell nuclei (grey). These images were taken 4 h after a 30 min heat shock. (**C**) Ribonuclear foci induced in *GFP*(CTG)$_{162}$ non-muscle nuclei by heat shock do not co-localize with muscleblind. The nucleus in the upper left of the panel is from a muscle cell and shows co-localization of ribonuclear foci and muscleblind. The nucleus in the lower right of the panel is from a non-muscle cell, which despite the presence of several intense ribonuclear foci, contains no muscleblind. The two nuclei shown are from the same image, but have been moved closer together for the sake of clarity. No differential image processing has been applied.

24B-GAL4 with *GFP*(CTG)$_{48/162}$ transgenes (over-expression of MBNL1 in muscle was associated with embryonic/early larval lethality). No ribonuclear foci were detected in any adult skeletal muscle in which they were not observed in the absence of mbl-A/C over-expression. These data illustrate that neither mbl-A nor mbl-C over-expression can mediate foci formation in larval motorneurons or adult leg and indirect flight muscles, indicating that some other factor is rate limiting in these tissues.

Muscleblind and MBNL1 increase steady state levels of CUG repeat RNA

Larvae over-expressing *mbl-A/C* and CUG$_{11/48/162}$ repeat RNA driven by 24B-GAL4 showed greater GFP fluorescence in the salivary gland than those expressing CUG repeat RNA alone (Fig. 6A). Higher levels of transgenic RNA were detected by *in situ* hybridization in salivary glands (Fig. 6A) and in whole larvae by northern blotting (Fig. 6B). Over-expressed mbl-A and -C accumulated in the cytoplasm of salivary gland cells and resulted in the formation of ribocytoplasmic foci in larvae expressing *GFP*(CUG)$_{162}$ (Fig. 6A). Despite the fact that all CUG repeat sizes were stabilized by over-expression of mbl-A and mbl-C, these ribocytoplasmic foci were not observed in lines expressing *GFP*(CUG)$_{11/48}$. The effects of MBNL1 were not analysed in this experiment, because its over-expression from the 24B-GAL4 driver was lethal. However, the D42-GAL4 driver also mediates expression in the salivary glands. Co-expression of *mbl-A/C* and expanded repeat RNA from this GAL4 driver gave similar results to expression driven from 24B-GAL4. However, co-expression of MBNL1 with *GFP*(CUG)$_{162}$ in salivary gland by D42-GAL4 led primarily to ribonuclear foci formation, in addition to the formation of some ribocytoplasmic foci (Fig. 6C). Co-expression of MBNL1 with *GFP*(CUG)$_{48}$ in salivary gland driven by D42-GAL did not produce ribonuclear or ribocytoplasmic foci, but resulted in higher levels of RNA and GFP fluorescence than that obtained with *GFP*(CUG)$_{48}$ alone (data not shown). Thus, MBNL1 is also capable of stabilizing the transgenic RNA. The stabilization of expanded repeat RNA, as detected by both enhanced GFP fluorescence and northern blotting, was not restricted to salivary glands as it was also observed when mbl-A/mbl-C/MBNL1 transgenes were co-expressed with (CUG)$_{48}$ from the *elav*-GAL4 driver (Fig. 6B). This GAL4 driver expresses in larval neurons, but not salivary glands, and does not cause ribonuclear foci formation with any combination of (CUG)$_{48/162}$ and mbl-A/-C/MBNL1 over-expression.

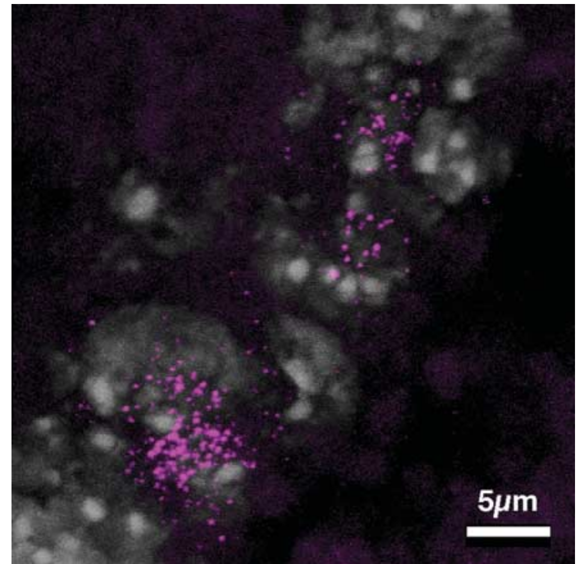

Figure 5. MBNL1 induced ribonuclear foci in motorneurons. Ribonuclear foci are present in the nucleus of neurons when *GFP*(CUG)$_{162}$ RNA is co-expressed with MBNL1. Image shows neurons of the ventral nerve cord (grey) and ribonuclear foci in purple. The final image was generated by merging a small number of images from a confocal stack by the maximum projection method.

Downloaded from http://hmg.oxfordjournals.org/ at BBSRC on July 7, 2015

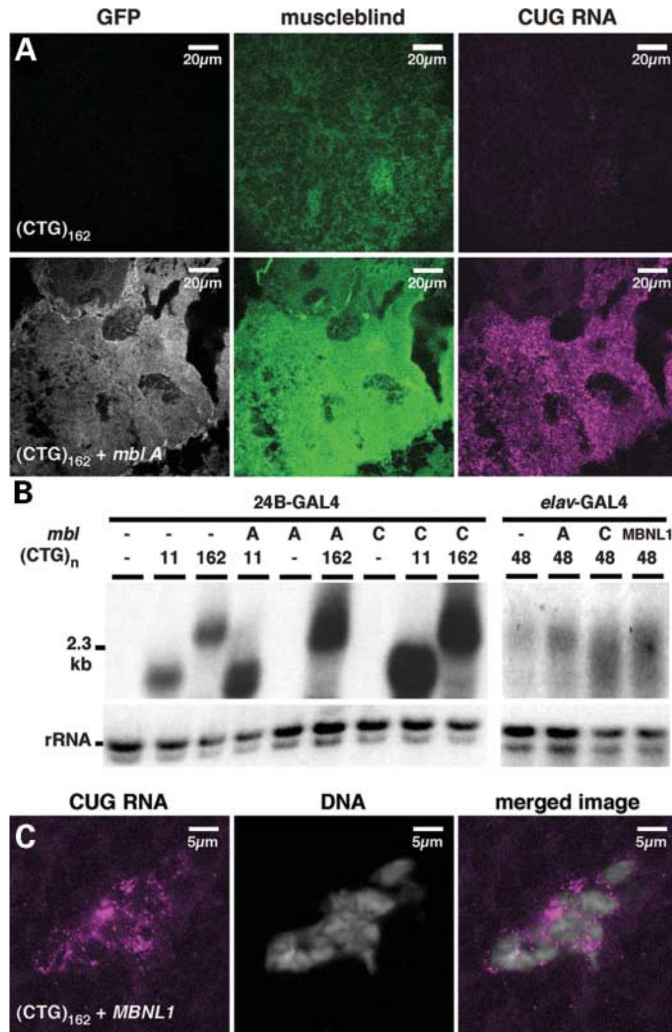

Figure 6. Muscleblind/MBNL1 and CUG repeat RNA interactions in salivary glands. (**A**) muscleblind A stabilizes cytoplasmic CUG repeat RNA. Salivary glands of third instar larvae expressing *GFP*$(CUG)_{162}$ RNA alone or in conjunction with a UAS-*mbl-A* transgene driven by 24B-GAL4. Images were captured at identical laser settings and camera exposure times to allow direct comparison of signal intensities. Note that although cytoplasmic levels of CUG repeat (purple) are increased in the presence of mbl-A, no nuclear ribonuclear foci were detected. Muscleblind was detected with an AMCA conjugated secondary antibody (green) because the exceptional levels of GFP (grey) produced by the stabilized transgenic RNA were not quenched during the *in situ* hybridization procedure. (**B**) mbl-A, mbl-C and MBNL1 all stabilize CUG repeat RNA. Northern blot analysis of transgene RNA levels in whole larvae expressing *GFP*$(CTG)_{11/48/162}$ RNA alone or in conjunction with a UAS-*mbl-A* or *-C* or UAS-*MBNL1* transgene. Blots were hybridized with a *GFP* specific probe. The ribosomal RNA loading control levels were revealed by ethidium bromide staining (image inverted). Transgene expression was driven in the salivary gland and muscle by 24B-GAL4 and in neurons by *elav*-GAL4. (**C**) Co-expression of *GFP*$(CTG)_{162}$ and UAS-*MBNL1* leads to the formation of ribonuclear foci in salivary gland nuclei. Transgene expression was driven by D42-GAL4. Repeat RNA forms ribonuclear and ribocytoplasmic foci (purple) in the nuclei (grey).

Ribonuclear foci co-localize with non-A

To gain further insights into the organization of ribonuclear foci, their spatial position in muscle nuclei was analysed. Ribonuclear foci occupied nuclear regions lightly stained with DAPI, but they were not observed in the nucleolus. They therefore shared the inter-chromatin space with the splicing and mRNA export machinery. The relative location of ribonuclear foci and molecular markers of spliceosomes and exosomes were analysed. No co-localization was observed (Fig. 7A). The relative location of the proteasome was also investigated to determine whether the proteins present in the ribonuclear foci contain mis-folded proteins and are targeted for degradation, and again no co-localization was observed (Fig. 7A).

One evolutionarily conserved cellular strategy for dealing with nuclear dsRNA is adenosine to inosine RNA editing (25). In vertebrate cells, it is known that editing is followed by attachment to the nuclear matrix by PSF and $p54^{nrb}$ (26). Although dsCUG repeat RNA should be immune to such editing, the hairpin structure may nonetheless interact with proteins in this pathway, so the relative locations of ribonuclear foci and non-A, the *Drosophila* PSF/$p54^{nrb}$ orthologue, were analysed (Fig. 7B). Co-localization of these elements was very good, indicating that non-A and expanded repeat RNA occupy the same nuclear regions and probably interact, directly or indirectly, *in vivo*. To determine whether expanded CUG containing transcripts were edited by the non-A pathway, RT–PCR amplified transcripts were cloned and sequenced from $(CTG)_{162}$ expressing flies. RNA editing and reverse transcription converts edited adenosines to guanosines in the cDNA (27). Sequencing of the CUG repeat tract and 120 bp 5′ and 280 bp 3′ of the flanking sequence failed to identify any adenosine to guanine mutations in six independent clones.

DISCUSSION

One of the most distinctive and consistent features of affected muscle in DM1 and DM2 patients is the formation of discrete ribonuclear foci of the expanded repeat RNA. These ribonuclear foci have become a hallmark of DM pathogenesis, yet their precise role in the disease pathway remains unclear. We have shown that the processes underlying ribonuclear foci formation are conserved in *Drosophila*, which develop reproducible patterns of ribonuclear foci when expressing a transgene containing 162 CUG repeats in the 3′-UTR of a *GFP* reporter gene. Moreover, these ribonuclear foci co-localize with muscleblind, the *Drosophila* orthologue of the human MBNLs. However, despite expression of the transgene ubiquitously and in several defined tissues, and the careful examination of a variety of phenotypes, we were unable to detect a deleterious affect of expanded CUG repeat expression on *Drosophila*. This included morphological analysis of the developmentally highly sensitive compound eye and a detailed examination of muscle using both functional and histological analyses. Muscle is the primary affected tissue in DM and the main tissue in which we observed ribonuclear foci in our *Drosophila* model. These data therefore demonstrate that neither the expanded $(CUG)_{162}$ repeat RNA nor the ribonuclear foci that are characteristic of DM are directly toxic to *Drosophila*. These data are in stark contrast to expanded polyglutamine-containing proteins that appear to be inherently toxic in most, if not all, cell types, including *Drosophila* neurons (17,18). Polyglutamine toxicity is mediated by

Downloaded from http://hmg.oxfordjournals.org/ at BBSRC on July 7, 2015

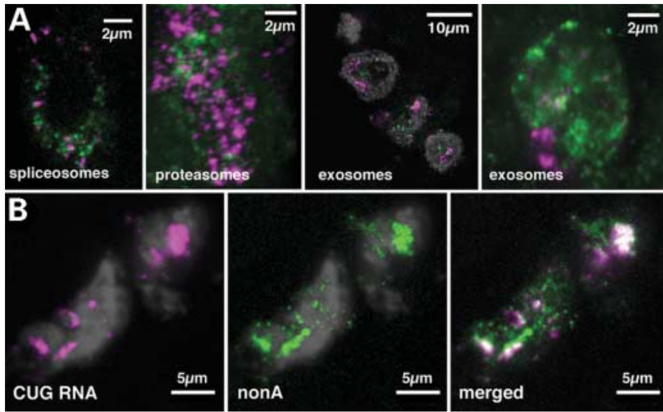

Figure 7. Ribonuclear foci and other intranuclear structures. (**A**) Ribonuclear foci do not co-localize with spliceosomal, proteasomal or exosomal markers. Muscle nuclei of a third instar larvae expressing *GFP*$(CTG)_{162}$ in muscle, showing ribonuclear foci (purple) and nuclear DNA (grey). The spliceosome was detected with antibodies for SC-35 (green), the proteasome with antibodies for dRegγ (green) and the exosome with antibodies for dSpt6 (green, left) and dRrp6 (green, right). (**B**) Ribonuclear foci co-localize with non-A in *Drosophila.* Muscle nuclei of a third instar larvae expressing *GFP*$(CTG)_{162}$ showing ribonuclear foci (purple) and non-A detected with anti-non-A (green).

disruption of a number of important cellular pathways via a mechanism that is assumed to involve aberrant interactions between the mis-folded polyglutamine tract and a host of normal cellular proteins (16). We have shown that ribonuclear foci are dynamic structures that do not co-localize with the proteasome and are unlikely to contain mis-folded proteins. These data therefore suggest that the expanded CUG repeat toxicity observed in humans and mice is mediated by a limited set of very precise RNA–protein interactions and that expanded CUG repeat RNA is not inherently toxic to cells *in vivo.* This observation indicates that pathogenicity is unlikely to extend beyond MBNL sequestration, which in turn suggests that therapeutic strategies that restored MBNL function to normal are likely to be highly beneficial.

Of course, it could always be hypothesized that an even larger $(CUG)_n$ repeat tract might be toxic to *Drosophila*, and this can only be definitively answered by direct experimentation. Nonetheless, several considerations lead us to conclude that the absence of a detectable pathological effect elicited by $(CUG)_{162}$ still represents a significant insight. First, an allele of $(CTG)_{162}$ repeats is at least three times the size of the smallest disease causing allele in humans (1), the same size as an allele shown to be toxic to spermatogenesis in mice (28) and not that much smaller than those of the $(CTG)_{250}$ repeats shown to produce myotonia and myopathy in mice (5). However, it should be noted that preliminary data suggest that the expanded array in the $(CTG)_{162}$ flies is somatically stable (Houseley *et al.*, unpublished data) in contrast to the high levels of expansion-biased somatic mosaicism observed in DM1 patients (29). Secondly, the GFP levels observed in muscle cells of $(CUG)_{162}$ and $(CUG)_{11}$ lines were indistinguishable, suggesting that the capacity to trap expanded CUG repeat RNA within the *Drosophila* cell nucleus is already saturated and arguing against any additional effect of a further expanded array. Although ribonuclear foci are observed with as few as 57 repeats in a mammalian cells (30), it is possible that $(CUG)_{162}$ is simply too short to become efficiently trapped within the *Drosophila* cell nucleus. These data, therefore, do not preclude the possibility that longer alleles may produce a pathogenic phenotype in *Drosophila.* Regardless, however, $(CUG)_{162}$ transcripts did generate abundant ribonuclear foci, thus demonstrating directly that ribonuclear foci *per se* are not inherently toxic to cells. In these experiments, we have already described the largest simple sequence repeat array yet incorporated into the *Drosophila* genome. Attempts to generate animal models with longer alleles are likely to be limited by the problems associated with cloning large $(CTG)_n$ tracts. Although cell culture models have been reported in which such an obstacle appears to have been overcome, the expanded arrays used are actually composed of synthetic concatamers of $[(CTG)_{20}$-$CTCGA]_n$ (9). Transcripts derived from such sequences would be predicted to have different structural and protein binding properties than those of $(CUG)_n$ arrays and have not been demonstrated to be functionally equivalent.

Accumulating data strongly implicate the MBNL proteins as critical targets of expanded CUG repeat RNA in man and mice (8,10). Given that muscleblind function is absolutely required for *Drosophila* development (23) and muscleblind is recruited to larval ribonuclear foci, why are there no pathogenic effects in our model? One possibility is the absence of ribonuclear foci during the critical period in embryonic development when muscleblind function is essential (23). The lack of ribonuclear foci during this short developmental window might be mediated by the lag between expression of muscleblind/$(CUG)_{162}$ transcripts and foci formation, or the absence of some other critical factor needed to form foci. In humans and mice, the period of muscle development is more prolonged and may be more susceptible to ribonuclear foci-mediated MBNL sequestration. Alternatively, despite co-localization with the ribonuclear foci, *Drosophila* muscleblind may not have as high an affinity for CUG RNA as the human MBNLs (discussed later), and its function may not be significantly affected by $(CUG)_{162}$ RNA.

The lack of a detectable toxic effect for the $(CUG)_{162}$ transgenes described here is also in contrast to two recent reports of *Drosophila* expressing untranslated triplet repeats. In the first, fragile X related $(CGG)_{90}$ repeats were expressed in the 5′-UTR of an *EGFP* transgene resulting in ubiquitinated intranuclear HSP70-positive protein aggregates resembling those adopted by expanded polyglutamine-containing proteins (20). It seems likely that the CGG repeats interact with a yet to be identified protein, leading to aggregation of mis-folded proteins and downstream dysfunction similar to that observed in the polyglutamine disorders. In the second model, the full length human *SCA8* transcript, which also contains an untranslated CUG repeat tract, was shown to cause *Drosophila* neurodegeneration (19). However, the neurodegeneration observed was as potent in *Drosophila* expressing $(CUG)_9$ as those expressing the expanded $(CUG)_{112}$ transcript, demonstrating that the toxicity of the RNA was not mediated by the expanded repeat tract but by some other element within the transcript. It is of course possible that additional sequence elements within the *DMPK* transcript absent in our system might contribute toward the pathology of the DM1 expanded CUG repeat

Downloaded from http://hmg.oxfordjournals.org/ at BBSRC on July 7, 2015

array in DM1 patients. Indeed, it has been shown that additional sequence elements within the *DMPK* 3′-UTR modify the effects of the expanded CUG repeat on myoblast differentiation (31). Nonetheless, the DM2 mutation (3) and the CUG repeat-expressing myotonic mice (5) demonstrate that an expanded repeat tract in the absence of additional *DMPK* sequences is sufficient to mediate DM pathology in mammalian cells.

In $(CTG)_{162}$ *Drosophila* expressing the transgene ubiquitously, ribonuclear foci were observed only in salivary glands, larval muscle cells and a subset of adult muscles. Thus, ribonuclear foci formation is not an obligate manifestation of expressing large expanded CUG repeat arrays suggesting that the hairpin structures such RNAs adopt (12) are not sterically blocked from exiting the nucleus, as has been previously proposed (32). These data therefore indicate that other cell type-specific factors must be important in mediating ribonuclear foci formation. In the salivary gland and muscle cells in which ribonuclear foci were observed, they co-localized with muscleblind. However, the relationship between muscleblind and ribonuclear foci proved to be more complex than this. For instance, endogenous muscleblind was detected in larval imaginal discs and adult indirect flight muscle, but no ribonuclear foci were detected in these tissues. Similarly, ectopic over-expression of muscleblind isoforms A and C in a tissue where muscleblind is normally absent (larval motorneurons) was unable to mediate ribonuclear foci formation. Conversely, heat-shock induction of $(CTG)_{162}$ transgene expression resulted in apparently muscleblind-negative ribonuclear foci formation in many tissues. Although it remains theoretically possible that these foci contained very low levels of muscleblind undetectable using fluorescent microscopy, this seems unlikely given that using the same methodology, muscleblind foci were readily observed in muscle cells with a similar signal intensity to the RNA component of the foci. Because ribonuclear foci were usually absent in the majority of these tissues, their formation must therefore be mediated by either the very high expression levels produced from the *hs*-GAL4 element or the induction of another protein able to aide foci formation. Thus, muscleblind is neither necessary nor sufficient to mediate ribonuclear foci of expanded CUG repeat RNAs in *Drosophila*.

Very interestingly though, ectopic expression of human MNBL1 in *Drosophila* was able to mediate ribonuclear foci formation in a subset of motorneurons and in salivary gland cells. Moreover, the ribonuclear foci initiated by MBNL1 were slightly smaller and more compact than those observed in the absence of MBNL1. These data suggest MBNL1 has a greater affinity for expanded CUG repeat RNA than muscleblind and/or a lesser dependence on additional factors to mediate foci formation. Another difference between the behaviour of MBNL1 and muscleblind was its normal intracellular location. In mammalian cells and *Drosophila* larvae MBNL/muscleblind is distributed throughout the cell. In adult *Drosophila* cells, muscleblind was clustered in intranuclear foci even in the absence of expanded CUG repeat RNA. These muscleblind foci were capable of recruiting expanded CUG repeat RNA, although this process did not alter their localization. Identification of the factors that mediate intranuclear muscleblind clustering in wild-type adult *Drosophila* cells may shed light on the formation of expanded CUG repeat RNA-mediated ribonuclear foci.

The MBNLs have recently been established as regulators of alternative splicing (8). The ability of muscleblind and MBNL1 to increase steady state levels of CUG repeat-containing transcripts in the cytoplasm was an unexpected finding that suggests a second role for muscleblind/MBNLs in translational control through the modulation of RNA stability. Although an effect of muscleblind/MBNL1 over-expression on transcription rates cannot be formally excluded, several lines of evidence argue against this possibility. Most significantly, the subcellular co-localization of muscleblind/CUG RNA in the muscle cell nuclei in wild-type flies (including sequestration of muscleblind into the nucleus in larval muscle cells), the subcellular co-localization of muscleblind/CUG RNA in the cytoplasm of salivary gland cells over-expressing muscleblind isoforms mbl-A and mbl-C and the subcellular co-localization of MBNL1/CUG RNA in the nucleus of salivary gland cells over-expressing MBNL1 all support a direct molecular interaction between muscleblind/MBNL1 and the CUG RNA. An additional cytoplasmic role for the muscleblind/MBNLs is also supported by its cytoplasmic location in mammalian cells (13) and *Drosophila* larvae and its preferential binding affinity for RNA as opposed to DNA (33). It is well known that AU-rich elements in the 3′-UTR of transcripts can modulate their stability (34). CUG repeats in the 3′-UTR presumably form a similar modulatory element. Because this effect was also observed for transcripts containing $(CUG)_{11}$, which are not uncommon in higher eukaryotic genomes, such a regulatory process could be physiologically relevant in the normal function of the MBNLs. Moreover, it suggests that MBNL sequestration mediated pathology in DM will not be limited to aberrant alternative splicing but will also include defects associated with aberrant transcript stability/translational control.

In our *Drosophila* model, the ribonuclear foci did not co-localize with the spliceosomal marker SC-35. Although it is assumed that RNA splicing can occur within or near such structures (35), it is probable that some aspects of RNA splicing also occur at additional sites within the nucleus. Interestingly, in our *Drosophila* model, the ribonuclear foci co-localized with non-A, a protein with mammalian orthologues (PSP1, PSF and $p54^{nrb}$) that are implicated in a number of processes, including RNA splicing. $p54^{nrb}$ and PSF share a consensus binding sequence for the U5 snRNA (36), and $p54^{nrb}$ has been shown to bind polypyrimidine tracts (37). PSP1 and $p54^{nrb}$ are present in paraspeckles, which are closely juxtaposed with spliceosomes (38). An interesting PSF-associated factor is PTB, which acts as an antagonist to CELF proteins in splice site choice (39). If ribonuclear foci interfere with PTB function, this could contribute to the splicing defects observed in DM. It is possible that the MBNLs are involved in the binding of PTB to splicing enhancers and, therefore, may function as part of a PSF/PTB complex. Consistent with this, model both PTB and PSF have been previously observed to associate with the *DMPK* 3′-UTR (40). Also of note, PSF and $p54^{nrb}$ are associated with matrin 3 and mediate attachment of dsRNAs to the nuclear matrix (26). These data therefore provide a potential physical link

Downloaded from http://hmg.oxfordjournals.org/ at BBSRC on July 7, 2015

between expanded CUG repeat RNA and nuclear matrix. We propose that in normal cells, some MBNL is associated with the nuclear matrix via PSF/p54^{nrb} and that transcripts containing short $(CUG)_n$ tracts bind to a single muscleblind/MNBL protein which is then released from the matrix and may be involved in mRNA export, stabilization and/or translational control. However, large expanded $(CUG)_n$ repeat tracts would be expected to bind more than one molecule of muscleblind/MBNL. This would delay dissociation from the matrix, because multiple muscleblind/MBNL molecules would need to dissociate simultaneously. Vacant muscleblind binding sites in the expanded $(CUG)_n$ tract would recruit newly synthesized or re-imported shuttling muscleblind, eventually culminating in the observed ribonuclear foci.

In DM patients, the stability and longevity of ribonuclear foci is not known. In our *Drosophila* model, foci were revealed as dynamic structures that formed rapidly during development or after heat-shock induction, but subsequently disappeared over the period of a few days after transient heat-shock-mediated expression. *In vitro* experiments have previously been used to demonstrate that the ribonuclear foci are lost during cell division in fibroblasts (41). Although it is possible that heat-shock-induced ribonuclear foci were lost from some cells because of cell-division, foci were also lost over the period of a few days in non-dividing cells such as muscle. These data are encouraging for DM therapeutic strategies aimed at down-regulating *DMPK/ZNF9* expression, because they suggest that pre-existing ribonuclear foci would be rapidly lost upon *DMPK/ZNF9* down regulation and MBNL sequestration reversed.

To summarize, we have created transgenic *Drosophila* incorporating CTG tracts of up to 162 repeats, the largest expanded simple sequence repeat yet incorporated into the fly genome. *Drosophila* expressing expanded CUG repeat RNA have abundant ribonuclear foci but no pathophysiological phenotype. This demonstrates that expanded CUG repeat RNA and ribonuclear foci are not directly toxic, at least to *Drosophila.* Ribonuclear foci are transient entities that normally contain muscleblind, which can aid their formation, although, surprisingly, it is not absolutely required. Moreover, muscleblind appears to have an additional role in the stabilization of $(CUG)_n$ transcripts in the cytoplasm. Ribonuclear foci exist in the inter-chromatin space and overlap to a significant extent with non-A in *Drosophila*, suggesting a direct link with the nuclear matrix and the ability to interact with additional factors involved in splice site selection. The *Drosophila* model presented in this study has already provided new insights into the basic biology underlying DM and muscleblind function and should prove an invaluable resource in addressing many of the remaining uncertainties.

MATERIALS AND METHODS

Drosophila transgenics

A human DM1 allele of 11 CTG repeats was PCR amplified using primers DM-CXH and DM-DRXB (Supplementary Material, Table S6 for oligonucleotide sequences). DM-CRX contains a single G/A mismatch with human genomic DNA that generates an *Xho* I site in the resultant amplification product. DM-DRXB contains a 12 nucleotide 5′ extension that generates an *Xba* I site in the resultant amplification product. The 11 repeat product was digested and cloned into the *Xho* I/*Xba* I sites immediately downstream of *GFP* in UAS-*GFP*, in the multiple cloning site of pP[UAST] (42,43). Human DM1 alleles of 56 and 162 CTG repeats were PCR amplified using primers DM-CXH and DM-DR and cloned into the pGEM T-easy (Promega) multiple cloning site. The 56 and 162 CTG repeat alleles were excised from pGEM T-easy using *Xho* I and *Spe* I and similarly cloned into the *Xho* I/*Xba* I sites in the 3′-UTR of *GFP* in pP[UAST]. During subcloning of the 162 repeat allele into pP[UAST] in *Escherichia coli*, a *de novo* deletion derivative allele of 48 CTG repeats was also generated. Transgenic *Drosophila* containing CTG repeat tracts of 11, 48, 56 and 162 were created from these pP[UAST] constructs by standard P-element mediated transformation. Details of the GAL4 driver lines used to mediate transgene expression and other UAS lines are provided in Supplementary Material, Tables S7 and S8.

Behavioural and histological analysis

See Supplementary Material for behavioural and histological analysis.

In situ hybridization

Oligonucleotide probes. Ten micrometre *Drosophila* sections were fixed in 4% paraformaldehyde for 15 min, washed three times with PBS, pre-hybridized with 2× SSC/40% formamide and hybridized at 37°C with 50 fM μl^{-1} Cy3-$(CAG)_{10}$ probe in 40% formamide, 1 mg ml^{-1} *E. coli* tRNA, 1 mg ml^{-1} sonicated salmon sperm DNA, 10% dextran sulphate and 0.2% BSA in 2× SSC. Slides were washed twice in 2× SSC and twice in 0.5× SSC for 15 min at 37°C, then mounted in Vectashield (Vector) with 2 $\mu g\ ml^{-1}$ DAPI.

RNA probes. Sections were fixed and washed as described above, permeabilized with 0.1 $\mu g\ ml^{-1}$ proteinase K in TE for 30 min, fixed for 5 min with 4% paraformaldehyde and washed three times with PBS. After dehydration through an ethanol gradient, slides were hybridized at 42°C with 300 pg μl^{-1} DIG-labelled antisense GFP RNA probe in 50% formamide, 1× Denhardt, 1 mM DTT, 1 mg ml^{-1} *E. coli* tRNA, 1 mg ml^{-1} sonicated salmon sperm DNA, 10% dextran sulphate in 4× SSC, then washed as for oligonucleotide probes. Slides were equilibrated in buffer 1 (0.1 M Tris–HCl pH 7.5/0.15 M NaCl), blocked with 0.5% blocking reagent (Roche) in buffer 1, stained with 1:200 FITC anti-DIG (Roche) and 0.5% blocking reagent in buffer 1 and mounted in Vectashield (Vector) with 2 $\mu g\ ml^{-1}$ DAPI after three washes with buffer 1. All *in situ* hybridizations were replicated in multiple sections from multiple animals and reproduced in their entirety at least twice.

Immunohistochemistry

Ten micrometre sections were fixed for 15 min with 4% paraformaldehyde, followed by three washes with PBS.

Downloaded from http://hmg.oxfordjournals.org/ at BBSRC on July 7, 2015

After blocking for 30 min with 5% normal serum (from the species in which the secondary antibody was generated) in PBS, samples were incubated with the primary antibody for 1 h in the same solution, followed by three washes with PBS. Samples were incubated for 1 h with the secondary antibody and 0.5% blocking reagent (Roche) in PBS, followed by three washes with PBS and mounting in Vectashield (Vector) with 2 $\mu g\ ml^{-1}$ DAPI. For combined *in situ* hybridization and immunohistochemistry, the *in situ* protocol was stopped after the last wash, samples were equilibrated in PBS and then the immunohistochemistry protocol was followed from the blocking step. Primary antibody details are supplied in Supplementary Material, Table S9. Fluorescent secondary antibodies were purchased from Jackson Immunochemicals and Molecular Probes. All immunohistochemical analyses were replicated in multiple sections from multiple animals and reproduced in their entirety at least twice.

Creation of antisera

ORFs of *muscleblind A*, *B* and *C* were cloned into pGEX-6P 1 (Amersham) to create N-terminal GST fusion constructs. Protein expression was induced for 4 h in Top10F′ *E. coli* (Invitrogen) at 37°C using 1 mM IPTG, and protein aggregates isolated using a protein refolding kit (Novagen). Fusion proteins were then purified by SDS–PAGE electrophoresis and band extraction, and an equal mixture of all three isoforms injected into a sheep by Diagnostics Scotland. The second bleed serum was used for all stainings.

RNA extraction and analysis

For northern blots, 1–10 μg total RNA, extracted using Tri-Reagent (Sigma), was separated on 1.5% agarose gels with guanidine thiocyanate, capillary blotted and probed with DIG labelled anti-sense riboprobes in EasyHyb (Roche) or with ^{32}P-labelled DNA probes in ExpressHyb (Clontech).

Protein extraction and analysis

Proteins were extracted with Tri-Reagent (Sigma), separated on 4–12% Bis-Tris Nu-PAGE gels, and blotted as per the manufacturer's instructions (Invitrogen). Antibody detection was performed by standard methods.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG Online.

ACKNOWLEDGEMENTS

We would like to thank Patrick Young, Fred Winston, Erik Andrulis and Harold Saumweber for provision of antisera, Andrea Brand for provision of the pP[UAST] GFP plasmid and Stephen Goodwin for provision of developmentally-staged *Drosophila* cDNA and fly stocks. Useful advice with *Drosophila* muscle histology and some behavioural experiments was provided by John Sparrow and Upendra Nongthomba. We would also like to thank the Dynamic Mutation Group and the *Drosophila* Fly Neurobiology Groups of the University of Glasgow for helpful discussion throughout the course of this work, Richard H. Wilson for statistical advice and Sarah L. Mole and Jane Evans for preliminary analyses of *Drosophila* courtship behaviour. This research was supported by awards from the Lister Institute and the Wellcome Trust. R.D.A. was supported by a contract from the 'Ramón y Cajal' program of the Ministerio de Educación y Ciencia.

REFERENCES

1. Harper, P.S. (2004) *Myotonic Dystrophy*, 3rd edn. WB Saunders Co., London.
2. Brook, J.D., McCurrach, M.E., Harley, H.G., Buckler, A.J., Church, D., Aburatani, H., Hunter, K., Stanton, V.P., Thirion, J.-P., Hudson, T. *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*, **68**, 799–808.
3. Liquori, C.L., Ricker, K., Moseley, M.L., Jacobsen, J.F., Kress, W., Naylor, S.L., Day, J.W. and Ranum, L.P. (2001) Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science*, **293**, 864–867.
4. Ranum, L.P. and Day, J.W. (2004) Myotonic dystrophy: RNA pathogenesis comes into focus. *Am. J. Hum. Genet.*, **74**, 793–804.
5. Mankodi, A., Logigian, E., Callahan, L., McClain, C., White, R., Henderson, D., Krym, M. and Thornton, C.A. (2000) Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science*, **289**, 1769–1773.
6. Charlet, B.N., Savkur, R.S., Singh, G., Philips, A.V., Grice, E.A. and Cooper, T.A. (2002) Loss of the muscle-specific chloride channel in type 1 myotonic dystrophy due to misregulated alternative splicing. *Mol. Cell*, **10**, 45–53.
7. Savkur, R.S., Philips, A.V. and Cooper, T.A. (2001) Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. *Nat. Genet.*, **29**, 40–47.
8. 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.
9. 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.
10. Kanadia, R.N., Johnstone, K.A., Mankodi, A., Lungu, C., Thornton, C.A., Esson, D., Timmers, A.M., Hauswirth, W.W. and Swanson, M.S. (2003) A muscleblind knockout model for myotonic dystrophy. *Science*, **302**, 1978–1980.
11. 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.
12. Michalowski, S., Miller, J.W., Urbinati, C.R., Paliouras, M., Swanson, M.S. and Griffith, J. (1999) Visualization of double-stranded RNAs from the myotonic dystrophy protein kinase gene and interactions with CUG-binding protein. *Nucleic Acids Res.*, **27**, 3534–3542.
13. Miller, J.W., Urbinati, C.R., Teng-Umnuay, P., Stenberg, M.G., Byrne, B.J., Thornton, C.A. and Swanson, M.S. (2000) Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *EMBO J.*, **19**, 4439–4448.
14. Fardaei, M., Rogers, M.T., Thorpe, H.M., Larkin, K., Hamshere, M.G., Harper, P.S. and Brook, J.D. (2002) Three proteins, MBNL, MBLL and MBXL, co-localize *in vivo* with nuclear foci of expanded-repeat transcripts in DM1 and DM2 cells. *Hum. Mol. Genet.*, **11**, 805–814.
15. Ebralidze, A., Wang, Y., Petkova, V., Ebralidse, K. and Junghans, R.P. (2004) RNA leaching of transcription factors disrupts transcription in myotonic dystrophy. *Science*, **303**, 383–387.
16. Lipinski, M.M. and Yuan, J. (2004) Mechanisms of cell death in polyglutamine expansion diseases. *Curr. Opin. Pharmacol.*, **4**, 85–90.
17. Warrick, J.M., Paulson, H.L., Gray-Board, G.L., Bui, Q.T., Fischbeck, K.H., Pittman, R.N. and Bonini, N.M. (1998) Expanded

Downloaded from http://hmg.oxfordjournals.org/ at BBSRC on July 7, 2015

polyglutamine protein forms nuclear inclusions and causes neural degeneration in *Drosophila*. *Cell*, **93**, 939–949.
18. Jackson, G.R., Salecker, I., Dong, X., Yao, X., Arnheim, N., Faber, P.W., MacDonald, M.E. and Zipursky, S.L. (1998) Polyglutamine-expanded human huntingtin transgenes induce degeneration of *Drosophila* photoreceptor neurons. *Neuron*, **21**, 633–642.
19. 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.
20. Jin, P., Zarnescu, D.C., Zhang, F., Pearson, C.E., Lucchesi, J.C., Moses, K. and Warren, S.T. (2003) RNA-mediated neurodegeneration caused by the fragile X premutation rCGG repeats in *Drosophila*. *Neuron*, **39**, 739–747.
21. Filardo, P. and Ephrussi, A. (2003) Bruno regulates gurken during *Drosophila* oogenesis. *Mech. Dev.*, **120**, 289–297.
22. Begemann, G., Paricio, N., Artero, R., Kiss, I., Perez-Alonso, M. and Mlodzik, M. (1997) *muscleblind*, a gene required for photoreceptor differentiation in *Drosophila*, encodes novel nuclear Cys3His-type zinc-finger-containing proteins. *Development*, **124**, 4321–4331.
23. Artero, R., Prokop, A., Paricio, N., Begemann, G., Pueyo, I., Mlodzik, M., Perez-Alonso, M. and Baylies, M.K. (1998) The *muscleblind* gene participates in the organization of Z-bands and epidermal attachments of *Drosophila* muscles and is regulated by *Dmef2*. *Dev. Biol.*, **195**, 131–143.
24. Garcia-Casado, M.Z., Artero, R.D., Paricio, N., Terol, J. and Perez-Alonso, M. (2002) Generation of GAL4-responsive muscleblind constructs. *Genesis*, **34**, 111–114.
25. Reenan, R.A. (2001) The RNA world meets behavior: A–>I pre-mRNA editing in animals. *Trends Genet.*, **17**, 53–56.
26. Zhang, Z. and Carmichael, G.G. (2001) The fate of dsRNA in the nucleus: a p54(nrb)-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. *Cell*, **106**, 465–475.
27. Emeson, R.B. and Singh, M. (2001) Adenosine-to-inosine RNA editing: substrates and consequences. In Bass, B.L. (ed.), *RNA Editing*. Oxford University Press, Oxford, pp. 109–138.
28. Monckton, D.G., Ashizawa, T. and Siciliano, M.J. (1998) Murine models for myotonic dystrophy. In Wells, R.D. and Warren, S.T. (eds), *Genetic instabilities and hereditary neurological diseases*. Academic Press, San Diego, pp. 181–193.
29. Monckton, D.G., Wong, L.-J.C., Ashizawa, T. and Caskey, C.T. (1995) Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males: small pool PCR analyses. *Hum. Mol. Genet.*, **4**, 1–8.
30. Amack, J.D., Paguio, A.P. and Mahadevan, M.S. (1999) Cis and trans effects of the myotonic dystrophy (DM) mutation in a cell culture model [*Erratum (1999) Hum. Mol. Genet*. **8**(13):2573]. *Hum. Mol. Genet.*, **8**, 1975–1984.
31. Amack, J.D. and Mahadevan, M.S. (2001) The myotonic dystrophy expanded CUG repeat tract is necessary but not sufficient to disrupt C2C12 myoblast differentiation. *Hum. Mol. Genet.*, **10**, 1879–1887.
32. 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.
33. Kino, Y., Mori, D., Oma, Y., Takeshita, Y., Sasagawa, N. and Ishiura, S. (2004) Muscleblind protein, MBNL1/EXP, binds specifically to CHHG repeats. *Hum. Mol. Genet.*, **13**, 495–507.
34. Tebo, J., Der, S., Frevel, M., Khabar, K.S., Williams, B.R. and Hamilton, T.A. (2003) Heterogeneity in control of mRNA stability by AU-rich elements. *J. Biol. Chem.*, **278**, 12085–12093.
35. Wang, J., Cao, L.G., Wang, Y.L. and Pederson, T. (1991) Localization of pre-messenger RNA at discrete nuclear sites. *Proc. Natl Acad. Sci. USA*, **88**, 7391–7395.
36. Peng, R., Dye, B.T., Perez, I., Barnard, D.C., Thompson, A.B. and Patton, J.G. (2002) PSF and p54nrb bind a conserved stem in U5 snRNA. *RNA*, **8**, 1334–1347.
37. Hallier, M., Tavitian, A. and Moreau-Gachelin, F. (1996) The transcription factor Spi-1/PU.1 binds RNA and interferes with the RNA-binding protein p54nrb. *J. Biol. Chem.*, **271**, 11177–11181.
38. Fox, A.H., Lam, Y.W., Leung, A.K., Lyon, C.E., Andersen, J., Mann, M. and Lamond, A.I. (2002) Paraspeckles: a novel nuclear domain. *Curr. Biol.*, **12**, 13–25.
39. Charlet, B.N., Logan, P., Singh, G. and Cooper, T.A. (2002) Dynamic antagonism between ETR-3 and PTB regulates cell type-specific alternative splicing. *Mol. Cell*, **9**, 649–658.
40. Tiscornia, G. and Mahadevan, M.S. (2000) Myotonic dystrophy: the role of the CUG triplet repeats in splicing of a novel DMPK exon and altered cytoplasmic DMPK mRNA isoform ratios. *Mol. Cell*, **5**, 959–967.
41. 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. USA*, **94**, 7388–7393.
42. Brand, A. (1995) GFP in *Drosophila*. *Trends Genet.*, **11**, 324–325.
43. 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.

Downloaded from http://hmg.oxfordjournals.org/ at BBSRC on July 7, 2015