Activation of the innate immune response and interferon signalling in myotonic dystrophy type 1 and type 2 cataracts

Jeremy D. Rhodes1,*, Martin C. Lott1,2, Sarah L. Russell1, Vincent Moulton2, Julie Sanderson3, I. Michael Wormstone1 and David C. Broadway4

1School of Biological Sciences, 2School of Computing Sciences and 3School of Pharmacy, University of East Anglia, Norwich, UK and 4Norfolk & Norwich University Hospital, Norwich, UK

Received September 20, 2011; Revised and Accepted November 1, 2011

Myotonic dystrophy (DM) is caused by a triplet repeat expansion in the non-coding region of either the DMPK (DM1) or CNBP (DM2) gene. Transcription of the expanded region causes accumulation of double-stranded RNA (dsRNA) in DM cells. We sought to determine how expression of triplet repeat RNA causes the varied phenotype typical of DM. Global transcription was measured in DM and non-DM cataract samples using Illumina Bead Arrays. DM samples were compared with non-DM samples and lists of differentially expressed genes (P ≤ 0.05) were prepared. Gene set enrichment analysis and the Interferome database were used to search for significant patterns of gene expression in DM cells. Expression of individual genes was measured using quantitative real-time polymerase chain reaction. DMPK and CNBP expression was confirmed in native lens cells showing that a toxic RNA gain of function mechanism could exist in lens. A high proportion, 83% in DM1 and 75% in DM2, of the significantly disregulated genes were shared by both forms of the disease, suggesting a common mechanism. The upregulated genes in DM1 and DM2 were highly enriched in both interferon-regulated genes (IRGs) and genes associated with the response to dsRNA and the innate immune response. The characteristic fingerprint of IRGs and the signalling pathways identified in lens cells support a role for dsRNA activation of the innate immune response in the pathology of DM. This new evidence forms the basis for a novel hypothesis to explain the complex mechanism of DM.

INTRODUCTION

Myotonic dystrophy type 1 (DM1) is the most common form of muscular dystrophy in adults and affects ~1 in 8000 people worldwide. The inheritance pattern is autosomal dominant, although this is complicated by the phenomenon of anticipation, where symptoms appear earlier and are more severe in successive generations. Phenotypes range from individuals who are only mildly affected in late adulthood to severely affected babies with the congenital form of the disease. DM1 is a highly variable, multisystemic disease and most organs and tissues of the body can be affected. Indeed, the clinical features of the disease are often extremely heterogeneous and include myotonia, progressive muscle weakness and atrophy, heart conduction defects, sleep apnea, mental retardation, testicular atrophy, insulin resistance, premature balding and cataracts (1). DM1 is in a group of diseases, including Huntington’s disease and Fragile-X syndrome, which are caused by unstable repeat expansions in the genome (2). The mutation responsible for DM1, discovered in 1992, is a CTG expansion in the 3′-untranslated region (UTR) of the dystrophia myotonica protein kinase (DMPK) gene on chromosome 19q13.3 (3). In the general population, repeats at the DM1 locus vary between 5 and 37 and are stable from generation to generation. Repeats of 38 or more, however, are unstable and in DM1 affected individuals CTG repeats between 50 and 4000 are recorded (4). More recently, the genetic basis of a second form of the disease (DM2) was found to be a CCTG repeat expansion in the first intron of the cellular nucleic acid binding protein (CNBP) gene, also known as
ZNF9, on chromosome 3q21 (5). The expansion size range in DM2, however, is much larger than for DM1 ranging between 75 and 11,000 repeats. Repeat instability in DM is complex and expansions between generations are frequent as is significant somatic instability during a patient’s lifetime. The clinical phenotypes of both types of DM are broadly similar and share many common features, although a congenital form of DM2 has not been described and DM2 has a generally milder phenotype than DM1 (4). The early appearance of cataract is a feature of both diseases.

The eye is severely affected in DM1 and symptoms can include ptosis, external ophthalmoplegia (paralysis of extraocular muscles), epiphora (excessive tear production), pupillary light-near dissociation, early onset cataracts, pigmentary retinopathy, bilateral optic nerve atrophy and low intraocular pressure (6,7). In its early stages, DM cataract manifests as fine iridescent dust-like opacities in the lens cortex. These develop into a stellate cataract in the posterior subcapsular region of the lens which impinges on the light path to the retina and usually requires cataract surgery. In their later stages, mature DM cataracts are often difficult to distinguish from other types of cortical cataract. In the majority of cataracts, the major damage to the lens occurs in the fibre cells. In DM1 cataracts, however, the density of cells in the lens epithelium is greatly reduced (8), indicating that the lens epithelial cells are also affected leading to a breakdown in lens homeostasis. DM1 and DM2 cataracts are generally regarded as indistinguishable from each other both in appearance and age of onset (1,9), although the authors are unaware of a specific study of cataract in DM1 or DM2.

Several different hypotheses have been proposed to explain the complex aetiology of DM1. These include the potential for direct effects of the triplet repeat mutation on the expression of DMPK or neighbouring genes, such as SIX5 and DMWD. In DM1, CUG repeat-containing DMPK transcripts can be detected, by in situ hybridization, as accumulations in nuclear foci (10,11) and evidence that mutated DMPK mRNA was retained in the nucleus led to the DMPK haploinsufficiency hypothesis. However, since no other mutations in the DMPK gene have been linked with DM and hetero- and homozygous DMPK knockout mice do not reproduce the symptoms of DM1 (12,13), it now seems unlikely that nuclear retention of mutant DMPK transcripts alone plays a major part in the disease. Interestingly, heterozygous CNBP knockout mice display symptoms of DM, including myotonia, muscle wasting, heart abnormalities and cataracts of an apparently similar phenotype to that seen in patients (14). Studies in patient muscle samples and in cell lines have shown both normal and abnormal expression patterns (15–19), and whether or not haploinsufficiency of CNBP has a role in the mechanism of DM2 remains controversial. No other mutations, unrelated to the triplet repeat expansion in the CNBP gene, have yet been detected in DM2 patients. The presence of the triplet repeat in the promotor region of the downstream gene SIX5 may also influence its expression, and studies have indicated reduced levels of SIX5 expression in DM1 patients (20). The appearance of cataracts in SIX5 knockout mice led to the suggestion that haploinsufficiency of SIX5 expression was responsible for cataract formation in DM1 (21,22). It should be noted that the cataracts described in SIX5 knockout mice were nuclear and not cortical as seen in patients and are therefore likely to have had a different origin. All lens fibre cells, including those of the fetal lens, are retained throughout life, and as the lens grows, layers of new fibre cells, which differentiate from epithelial cells at the lens equator, are added to the outside of the existing fibre cell mass. The nuclear cataracts observed in SIX5 knockout mice could, therefore, indicate a potential role for SIX5 during early lens development. Also, SIX5 is not adjacent to the DM2 repeat expansion and the same cis-acting mechanism could not apply to DM2 cataracts. The similarity in the cataract phenotype between DM1 and DM2 suggests a common mechanism which is unrelated to SIX5 expression.

The discovery of DM2, which shares many of the symptoms of DM1 (12,23), has greatly reinforced the hypothesis that a common trans-acting, toxic gain of function by the mutant RNA is the basis for the mechanism in both diseases. Importantly, it has been shown in DM2 that the intronic repeat sequence can also be detected in nuclear foci and that CNBP protein levels are normal (19). The inappropriate expression in adults of a number of embryonic splice variants has been described in DM (23,24) and altered splicing of a muscle-specific chloride channel (CLCN1) and the insulin receptor have been linked directly to myotonia and diabetes in DM patients. The splicing factor, muscleblind like 1 (MBNL1) binds to expanded CUG repeats and co-localizes with nuclear foci in DM (25). One hypothesis proposes that MBNL1 is sequestered by triplet repeat-containing RNA in the nuclei of DM cells, thus reducing its activity as a splicing regulator. In support of this hypothesis, it has been shown that the muscle pathology and splicing pattern of DM is reproduced in Mbn1 knockout mice (26). Also over-expression of Mbn1 in the muscles of mice expressing triplet repeats, inserted into the 3′-UTR of a human skeletal α-actin (HSA) gene, reversed the mis-splicing effects (27). Two microarray studies carried out on skeletal muscle from HSA mice expressing CUG repeats have shown, however, that a significant proportion of deregulated genes could not be explained by the sequestration of MBNL1 (28,29). Levels of another splicing regulatory factor involved in RNA processing, CUGBP1, have been shown to increase in DM (30) and mice overexpressing CUGBP1 exhibit DM-like muscle abnormalities (30,31). Interestingly, the CUGBP1 protein is hyperphosphorylated and stabilized in DM cells by increased PKC activity (32), although the potential for triplet repeat RNA to activate cellular signalling pathways which include PKC activation has not yet been addressed. In addition to its effect on the activity of splicing regulation, other unrelated hypotheses have also been proposed. It has been suggested, for example, that the mutant CUG repeat-containing RNA could cause the sequestration of transcription factors, thus reducing the expression of regulated genes (33). Also double-stranded RNA (dsRNA) hairpin loops formed by CUG repeats have been shown to be substrates for Dicer, and downregulation of genes containing complementary repeats by an RNA interference mechanism has also been proposed as a mechanism of disease in DM (34).

In this study, we sought to investigate, using native lens cells from DM cataract patients, how accumulation of triplet repeat-containing RNA causes the varied symptoms that are
typical of DM. We have used oligonucleotide gene arrays to record global changes in the expression of genes in lens samples from both DM1 and DM2 cataract patients. We have discovered that a very high proportion of disregulated genes were common to both DM1 and DM2, providing evidence in favour of there being a common mechanism for both types of the disease. Using gene set enrichment analysis (GSEA), we searched for common patterns of gene expression which could point to the underlying mechanism of DM. Interferon-regulated genes (IRGs) were upregulated and evidence for type 1 interferon (IFN) signalling pathway activity was found in both DM1 and DM2. Based on these new data, we propose that the innate immune response plays an important and potentially fundamental role in the mechanism of both types of DM.

RESULTS

Expression of DMPK, SIX5 and CNBP in native DM lens epithelial cells

To investigate the possibility that a toxic gain of function by triplet repeat-containing RNA could also occur in the lens, it was important to show that both DMPK and CNBP (ZNF9) were expressed. The expression of DMPK and SIX5 genes was detected in lens epithelial samples from DM1 and control cataract patients (Fig. 1A and B). The mean DMPK mRNA level was greater in the DM1 samples compared with the control cataracts, although the difference was not significant. The SIX5 mRNA level was similar in both the DM1 and control samples. CNBP gene expression was measured in lens epithelial samples from DM2 and control cataract patients (Fig. 1C), and the CNBP mRNA level in the DM2 samples was found to be similar to that of the control cataract samples. It should be noted that DM2 is less common than DM1 and only two DM2 patient samples were available for this part of the study. We also checked the CNBP mRNA expression data obtained using gene arrays and this showed a slight but not significantly increased fold change (FC) in the DM2 samples compared with the controls, confirming the PCR data (Fig. 1D).

Genes disregulated in lens epithelial samples from DM cataract patients

To investigate global gene expression in DM cataract, samples of lens epithelium were obtained from three DM1 and three DM2 cataract patients. Using Illumina Human WG-6 bead arrays, global mRNA levels were analysed by comparing the DM samples with samples obtained from control non-DM cataract patients (n = 4). Hierarchical cluster analysis of all samples based on all transcripts on the array (Fig. 2) revealed that the DM and control samples clustered separately. There was, however, no distinct segregation between the DM1 and DM2 samples. Of the 382 genes which showed a significant (P ≤ 0.05) change in the mRNA steady-state level in the DM1 cataract patient samples compared with the control non-DM cataract group, 317 increased in expression (FC > 1.7) and 65 decreased in expression (FC < −1.7) (Supplementary data). Of the 419 genes which showed a significant...
(P ≤ 0.05) change in the mRNA steady-state level in the DM2 cataract patient samples compared with the control group, 351 increased in expression and 68 decreased in expression (Supplementary data). Table 1 shows the 10 most up- and downregulated genes in DM1 and DM2, respectively. It can be seen that there is a remarkable similarity in the deregulated genes between both types of DM with ~80% of the genes the same. Interestingly, 317 genes (265 upregulated and 52 downregulated) which showed a significant change in expression were common to both DM1 and DM2 (Fig. 3A and B) representing 83 and 75% of the significant genes, respectively. This strongly suggests a common disease mechanism in both types of DM. We used quantitative real-time polymerase chain reaction (QRT-PCR) to confirm the expression pattern of selected genes (Fig. 4) and in each case the genes were up- or downregulated in the same direction as that reported in the gene arrays.

**GSEA and Interferome analysis**

Having established a common gene expression profile in DM1 and DM2, a third gene list was produced by combining all the DM samples (n = 6) and comparing them to the control cataract group (n = 4). GSEA was used to identify potential disease mechanisms in DM. Analysis of the gene sets associated with transcription factor targets (Molecular Signatures Database—MSigDB, C3: motif gene sets—Transcription factor targets) identified 44 highly significant (familywise error rate—FWER ≤ 0.001) gene sets of known transcription factors that were enriched in the DM samples (Table 2). A relatively large number of transcription factors were associated with IFN/cytokine signalling pathways (e.g. STAT1), immune responses (e.g. PU.1), transcriptional repression (e.g. YY1) and cell cycle arrest (e.g. C/EBP alpha). A search of the MsigDB (C3: motif gene sets—transcription factor targets) found 10 gene sets which were directly associated with IFN transcription factor targets and all of these were significantly (false discovery rate—FDR-q < 0.02) enriched in the DM gene list (Supplementary data). A search of all gene sets in the MsigDB (C2 curated gene sets) identified 65 associated with IFN and 58 of these were significantly (FDR-q < 0.25) enriched in the DM gene set (Supplementary data).

To further investigate the potential role of IFN-activated pathways in DM, we searched the significantly deregulated genes for type 1 IRGs using the Interferome database (35). Of the most upregulated genes in the combined DM list, 15.1% were IRGs compared with <6% of the downregulated genes. In the top 100 upregulated genes, 27 were IRGs. In order to confirm that IRGs formed a similar proportion of the deregulated genes in both types of DM, we did the same analysis on the separate DM1 and DM2 gene lists. We found that 14.5 and 15.3% of the upregulated genes were IRGs, respectively, compared with 4.6 (DM1) and 5.8% (DM2) of the downregulated genes. Of the top 100 upregulated genes, 28 and 25% (Table 3) were IRGs in the DM1 and DM2 gene lists, respectively.

### Table 1. The 10 most up- and downregulated genes in DM1 and DM2

<table>
<thead>
<tr>
<th>Gene</th>
<th>Fold change</th>
<th>P-value</th>
<th>q-value</th>
<th>Entrez gene id</th>
</tr>
</thead>
<tbody>
<tr>
<td>DM1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HNRNPA1L2</td>
<td>13.108</td>
<td>3.05e</td>
<td>9.90e</td>
<td>6</td>
</tr>
<tr>
<td>PGAM1</td>
<td>8.817</td>
<td>1.40e</td>
<td>5.31e</td>
<td>6</td>
</tr>
<tr>
<td>SNURF</td>
<td>7.247</td>
<td>5.19e</td>
<td>2.37e</td>
<td>8</td>
</tr>
<tr>
<td>CRIM1</td>
<td>7.092</td>
<td>2.38e</td>
<td>5.02e</td>
<td>5</td>
</tr>
<tr>
<td>AQP1</td>
<td>6.677</td>
<td>2.64e</td>
<td>2.53e</td>
<td>3</td>
</tr>
<tr>
<td>TIMP3</td>
<td>6.492</td>
<td>3.65e</td>
<td>1.34e</td>
<td>7</td>
</tr>
<tr>
<td>SEPP1</td>
<td>6.209</td>
<td>4.11e</td>
<td>2.32e</td>
<td>2</td>
</tr>
<tr>
<td>TMEM123</td>
<td>5.594</td>
<td>6.17e</td>
<td>1.76e</td>
<td>5</td>
</tr>
<tr>
<td>SNURF</td>
<td>5.514</td>
<td>6.17e</td>
<td>1.76e</td>
<td>5</td>
</tr>
<tr>
<td>DM2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HNRNPA1L2</td>
<td>9.811</td>
<td>5.35e</td>
<td>1.18e</td>
<td>4</td>
</tr>
<tr>
<td>PGAM1</td>
<td>7.115</td>
<td>2.05e</td>
<td>5.24e</td>
<td>5</td>
</tr>
<tr>
<td>SNURF</td>
<td>5.614</td>
<td>2.20e</td>
<td>5.59e</td>
<td>5</td>
</tr>
<tr>
<td>TME1M123</td>
<td>5.604</td>
<td>6.00e</td>
<td>2.00e</td>
<td>5</td>
</tr>
<tr>
<td>AQP1</td>
<td>5.33</td>
<td>2.12e</td>
<td>1.47e</td>
<td>2</td>
</tr>
<tr>
<td>CRIM1</td>
<td>5.275</td>
<td>9.33e</td>
<td>1.14e</td>
<td>3</td>
</tr>
<tr>
<td>SEPP1</td>
<td>5.137</td>
<td>1.60e</td>
<td>7.50e</td>
<td>2</td>
</tr>
<tr>
<td>CDC14B</td>
<td>4.884</td>
<td>8.34e</td>
<td>6.97e</td>
<td>2</td>
</tr>
<tr>
<td>EEF1A1</td>
<td>4.467</td>
<td>5.61e</td>
<td>4.94e</td>
<td>2</td>
</tr>
<tr>
<td>DAAM2</td>
<td>4.413</td>
<td>1.56e</td>
<td>1.72e</td>
<td>3</td>
</tr>
<tr>
<td>TRIM16L</td>
<td>-5.672</td>
<td>1.18e</td>
<td>2.25e</td>
<td>10</td>
</tr>
<tr>
<td>PCP4</td>
<td>-4.711</td>
<td>3.19e</td>
<td>8.07e</td>
<td>1</td>
</tr>
<tr>
<td>SALL3</td>
<td>-3.053</td>
<td>2.06e</td>
<td>3.55e</td>
<td>4</td>
</tr>
<tr>
<td>ZNF660</td>
<td>-2.989</td>
<td>1.60e</td>
<td>6.97e</td>
<td>6</td>
</tr>
<tr>
<td>FCGRT</td>
<td>-2.873</td>
<td>7.48e</td>
<td>9.57e</td>
<td>2</td>
</tr>
<tr>
<td>PPID</td>
<td>-2.736</td>
<td>5.15e</td>
<td>15.97e</td>
<td>10</td>
</tr>
<tr>
<td>SIX3</td>
<td>-2.653</td>
<td>9.04e</td>
<td>2.96e</td>
<td>1</td>
</tr>
<tr>
<td>GNPTG</td>
<td>-2.613</td>
<td>2.35e</td>
<td>1.58e</td>
<td>2</td>
</tr>
<tr>
<td>HCRT</td>
<td>-2.587</td>
<td>1.43e</td>
<td>6.51e</td>
<td>8</td>
</tr>
<tr>
<td>LRBP3</td>
<td>-2.572</td>
<td>6.90e</td>
<td>3.77e</td>
<td>2</td>
</tr>
</tbody>
</table>

Figure 2. A hierarchical cluster plot for all genes, where the standard deviation divided by the mean was >0.1, in the global gene expression profiles of DM and control cataract samples. The dendrogram shows the results of unsupervised clustering analysis of three DM1, three DM2 and four control (CON) non-DM cataract samples. Global gene expression was measured using Illumina gene arrays (see Materials and Methods for details).
DISCUSSION

In this study, we present evidence in support of a common mechanism of cataract formation in type 1 and type 2 DM. It has previously been proposed that haploinsufficiency of SIX5 expression could explain cataract formation in DM1 patients (21, 22) and reports that SIX5 expression could explain cataract formation in DM1.

It has previously been proposed that haploinsufficiency of SIX5 expression could explain cataract formation in type 1 and type 2 DM. In this study, we present evidence in support of a common mechanism of cataract formation in type 1 and type 2 DM.

Analysis of the global mRNA profiles of DM-affected lens cells has revealed significant indicators for the involvement of the innate immune response in the mechanism of the disease. First, Interferome (35) analysis showed that there was a high proportion of disregulated genes (83% of DM1 and 75% of DM2), identified using gene array analysis, which were common to both diseases, suggests that there were common pathways regulating their expression that are independent of SIX5 expression.

Analysis of the global mRNA profiles of DM-affected lens cells has revealed significant indicators for the involvement of the innate immune response in the mechanism of the disease. First, Interferome (35) analysis showed that there was a high proportion of disregulated genes (83% of DM1 and 75% of DM2), identified using gene array analysis, which were common to both diseases, suggests that there were common pathways regulating their expression that are independent of SIX5 expression.

dsRNA, as a replication intermediate, is a common feature of many viruses and is exploited by mammalian cells to detect infection and activate antiviral countermeasures through type 1 IFN production and the innate immune response (41). It has been shown that CUG and CCUG repeat RNA can form dsRNA (42) and, therefore, a possible consequence of transcription of either the DMPK gene in DM1 or the CNBP gene in DM2 could be the aberrant activation of the innate immune response. dsRNA is sensed by several intracellular receptors, including the dsRNA dependent protein kinase (PKR), 2′-5′-oligoadenylate synthetase (OAS) proteins, toll-like receptor 3 (TLR3) and the Rig-I-like helicases (RLHs), RIG-1 and MDA-5 (Fig. 5). PKR can be activated by relatively short (30 bp) fragments of dsRNA which on binding cause homodimerization followed by autophosphorylation. Although its primary effect is to inhibit protein synthesis via the phosphorylation of the translation initiation factor eIF-2α, PKR also has roles in signal transduction pathways, including JNK and NF-κB, regulation of cell growth and induction of apoptosis (43). The four members of the 2′-5′OAS group of proteins bind dsRNAs longer than 15 bp and when activated convert ATP to 2′–5′-linked oligo adenylates (2′–5′A) up to 30-mers in size. 2′–5′A binds to and activates RNase L which catalyses the degradation of both viral and cellular RNAs, also causing a reduction in general protein synthesis (44). The more recent discovery of TLR3 indicated the existence of further dsRNA sensors in addition to PKR and the 2′–5′A/RNase L system. TLR3 is confined to intracellular...
vesicles, such as endosomes and endoplasmic reticulum, and recognizes dsRNA. TLR3 initiates the TLR adaptor molecule 1-dependent pathway to induce inflammatory cytokines and type 1 IFN via the transcription factors NF-κB and IRF3 (45). The RLH proteins include RIG-1 and MDA-5 and they comprise a RNA helicase domain and a caspase recruitment domain (CARD). A CARD adaptor protein called mitochondrial anti-viral sensor transmits the signal from RLHs to IRF3 and NF-κB, which in turn activate the type 1 IFN promoter (46). Although multiple cytokines and chemokines can be produced by host cells in response to infections, type 1 IFNs (IFNα/β) are the principle cytokines induced by dsRNA. Once synthesized, IFN α or β is released by the cell where they act in an autocrine or paracrine way to stimulate a common heterodimeric receptor (IFNAR) which after dimerization activates the downstream JAK/STAT pathway. A heterotrimeric complex between STAT1 and 2 and IRF9 binds to the IFN-stimulated response element of IRGs stimulating their transcription (38). Several hundred IRGs have been identified and their expression in response to IFN is responsible for inducing an antiviral state in infected cells. Cell cycle arrest, transcriptional repression, inhibition of protein synthesis, release of inflammatory cytokines and apoptosis are all characteristics of the antiviral response (41).

Many of the transcription factors identified in Table 3 have direct or indirect links to the immune response, to viral infection and/or to viral gene expression. Many, for example, are directly targeted by viral oncoproteins (47,48) and are implicated in cellular antiviral responses. In addition to transcription factor targets directly associated with the innate immune response and cytokine signalling, a high proportion were associated with transcriptional repression and cell cycle control (YY1, E4BP4, E4F1, E2F1 and TFDP1). Type 1 IFN has been shown to be highly antiproliferative and both
transcriptional repression and cell cycle control are characteristic of the antiviral state (41). YY1, for example, has a critical role in regulating the expression of a very large number of genes (49) and significantly it also binds to the IFNβ promoter to repress or activate its expression (50) and thus may have a role in modulating the cellular response to dsRNA.

It has previously been shown in vitro that CUG repeat RNA can activate PKR (51) and is a substrate for Dicer (34), and studies of the structural diversity of triplet repeat RNA confirm that CUG and CCUG triplet repeats form stable, double-stranded hairpin loops (42,52). It is therefore conceivable that the presence of triplet repeat-containing dsRNA could activate the innate immune response in DM-affected cells. Interestingly, it has recently been reported that transcriptional repression of 76% of transcripts occurred in DM1 myoblast cell lines which were not the result of altered levels of MBNL1 or CUGBP1 (53). A microarray study carried out on skeletal muscle from HSA mice expressing MBNL1 or CUGBP1 (53). A microarray study carried out on skeletal muscle from HSA mice expressing CCUG repeats found that a large set of genes associated with the extracellular matrix (ECM) function was disregulated which could not be explained by the sequestration of MBNL1 (28). Dysregulation of cell adhesion molecules is a feature of the antiviral response (41) and on closer inspection it can be seen that ~43% of the ECM genes disregulated in the study of Du et al. (28) were IRGs supporting a role for IFN signalling in DM. Cell cycle arrest is also a prominent feature of the IFN-induced antiviral state and reduced cell density in the epithelia of lenses from DM1 patients (8), reduced proliferation rates in human lens cell lines derived from DM cataract patients (37) and human DM myoblasts (54) have also been reported. Furthermore, cell cycle arrest inducing premature senescence through increased levels of p16 has been reported in human DM.
DM also produces dsRNA and could activate all or some of these antiviral muscle precursor cells (55). PKR expression is induced by members of the 2-phorylation of the initiation factor eIF2. When activated by dsRNA, PKR has been detected in the nucleus (57), many of the dsRNA-sensing proteins described above are cytoplasmic in their distributions (46). Whether or not triplet repeat RNA is exclusively retained in the nucleus is controversial, and several studies indicate that it is also present in the cytoplasm (10,58,59). Importantly, during cell division, the nuclear membrane disappears and in consequence the nuclear contents, including foci of dsRNA, are dispersed in the cytoplasm. It can, therefore, be argued that it is likely that at some stage during the cell cycle, dsRNA, produced as a consequence of either DMPK or CNBP transcription, will be detected by one or all of the many dsRNA receptors present in cells.

We have identified an expression profile in the lens epithelial cells of DM1 and DM2 cataract patients, which suggests that dsRNA-induced IFN signalling pathways are involved in a common mechanism of cataract formation. Although cataracts are frequently associated with infections inside the eye causing uveitis, little is known about the mechanism and a link to IFN in the eye has not yet been established. The characteristic fingerprint of IRGs among the deregulated genes and the pathways identified by GSEA provide evidence that dsRNA activation of the innate immune response may play an important role in the pathology of DM not only in the eye but in the many other affected parts of the body. This new evidence forms the basis for a novel and coherent hypothesis to explain the complexities of this disease. Nearly 30 triplet repeat diseases have been identified to date and of these 8, in addition to DM1 and DM2, also have repeat sequences in the non-coding regions of the affected genes (60). Our hypothesis could, therefore, be applied more widely to include other triplet repeat diseases where transcription of the mutated genes could also produce dsRNA with the potential to activate the innate immune response.

**MATERIALS AND METHODS**

**Patients and samples**

Lens epithelial samples were obtained from DM patients who were undergoing cataract surgery. For the microarrays, three DM1 samples (two male and one female) were obtained from patients between the ages of 54 and 67, three DM2 samples (two female and one male) were from patients between 40 and 58 years old and four control samples were obtained from female patients between the ages of 61 and 82 years. An additional four DM1 (age 32–64), two DM2 (age 40–42) and four control (age 69–80) samples were obtained for validation using QRT–PCR. All the DM patients had the adult onset form of the disease and cortical cataracts. The control samples obtained from the Norfolk and Norwich University Hospital NHS Trust, Norwich, UK were from lenses with non-specific, age-related cortical cataracts. Samples were placed immediately after removal from the patient’s eye into RNA later (Qiagen, UK) and sent directly to the laboratory. Ethical approval (reference: 2010/11-032) for the study was obtained from the Faculty of Medicine and Health Sciences Research Ethics Committee, University of East Anglia, Norwich, UK.

**RNA extraction and microarray processing**

Total RNA was extracted using Qiagen MicroRNA extraction kits following the manufacturer’s protocols. On column DNA digestion was performed using a RNase-free DNase I kit (Qiagen). RNA samples were frozen at −80°C until required. Gene expression analysis was carried out using Illumina Gene Expression Arrays HumanWG-6 v3 BeadChips. Array processing was performed by Source BioScience Plc. (Nottingham, UK).

**Microarray data analysis**

Illumina microarray (BeadArray) unnormalized probe profile data were analysed using the Bioconductor package (http://www.bioconductor.org) in R (http://www.r-project.org). First, the data from different chips were loaded into R to be background corrected, quartile normalized and variance stabilized (61). Quartile normalization involves sorting the expression levels of each sample and then setting the value at each position i in the lists to be the average of the values at position i across all samples. Finally, since the variance of each list may be affected by the mean, variance stabilization was performed to ensure that the variance of the expression levels is independent of the mean. The normalized and raw data from all the arrays are available in the Array Express database (www.ebi.ac.uk/arrayexpress) (62) under the accession number: E-MEXP-3365. As a quality-control check, a hierarchical clustering plot for all genes, where the standard
deviation divided by the mean was >0.1, was used to visualize whether patient samples grouped by phenotype. To determine which genes were differentially expressed between two given phenotypes, we first computed the expression level of each gene according to the probe(s) to which that gene corresponded. More specifically, each probe was first converted to its nucleotide identifier (nuID) (55), as defined by the Human WG6 v3 genome chip, and the nuID was used to look up the gene name in the lumiHumanIDMapping database. Lists of differentially expressed genes were computed by using the TREAT statistic (63) to compute a P-value. However, in contrast to a basic t-test, the t-test we perform is relative to a FC threshold and this allows us to formally test the hypothesis that a gene is more differentially expressed than a given FC, that is the ratio of average expression level between two phenotypes. The FC threshold used for all P-values which result from the TREAT statistic throughout this paper is 1.7 which corresponds to a 70% increase in average expression. Since there are multiple comparisons, the P-value adjusted for multiple testing is also computed to control the expected likelihood that a gene which appears up/downregulated in our lists has been falsely identified as being differentially expressed; this is known as the FDR. The basic t-test is known to give a high FDR and is only weakly related to FC. In contrast, the TREAT statistic has been shown to improve upon the FDR of existing methods (63). We use the Benjamini and Hochberg method (64) to compute adjusted P-values (q values) and control the FDR rate. Given a cut-off of 0.05 on the q-values of each gene, the expected proportion of genes falsely identified to be differentially expressed is controlled to be less than the threshold value, in this case 5% (65).

The array data were further analysed using GSEA (66) and the Molecular Signatures Database (MSigDB; http://www.broadinstitute.org/gsea). GSEA is a computational method that enabled us to compare the lists of differentially expressed genes with previously defined sets of genes in the MSigDB and identify gene sets which were significantly enriched in our experimental data. The MSigDB gene sets are divided into five major collections and in this study we probed our gene lists using the C2: curated gene sets (3272 gene sets) and the C3: motif gene sets (836 gene sets). GSEA takes as its input the entire set of probes together with their normalized, logged expression levels in each data set. We considered any gene sets with a FDR q-value of <25% to be significant, although in most cases a higher level of significance was obtained and the more stringent FWER was also used. The FWER is the probability of stating a set of genes that are differentially expressed when they are in fact not, that is, a false-positive result. In addition to the GSEA analysis, the differentially expressed genes were further analysed using the InterFerome database (35) (http://www.interfomer.org) to search for type I IRGs.

Quantitative real-time PCR

Array expression data were validated in similar cataract samples to those used to perform the gene array analysis by quantitative RT–PCR (QRT–PCR). Total RNA was extracted using an RNase-free DNase I kit (Qiagen) following the manufacturer’s protocol along with on-column DNA digestion performed using an RNase-free DNase I kit (Qiagen). RNA was quantified using a Nanodrop ND-1000 spectrophotometer (Nanodrop, Wilmington, DE, USA) and frozen at −80°C until required. One hundred to 300 ng of RNA were reverse transcribed on a PTC-200 DNA engine (MJ Research, Reno, NV, USA) using SuperScript II reverse transcriptase (Invitrogen, Paisley, Renfrewshire, UK) following the manufacturer’s instructions. Expression of selected genes was measured using 5 ng cDNA per reaction with the following predesigned primers and probes (Assay-on-Demand, Applied Biosystems, Warrington, Cheshire, UK): DMPK—Hs01094334_g1, HNRNPA1L2—Hs00402055_m1, PGAM1—Hs01652468_g1, TRIM16L—Hs02598492_mH, PCP4—Hs01113638_m1, CNBP (ZNF9)—Hs00231535_m1, AQP1—Hs01028916_m1, 18S—Part Number 4308329. SIX5 expression was measured using an Assay-by-Design (Applied Biosystems) as previously reported (37). QRT–PCR was performed using a standard run on a 7500 Fast Real-Time PCR system. Threshold cycle values were converted to nominal input values using previously constructed standard curves and normalized to the 18S expression value of each sample (37).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS

The authors would like to thank Dr Penny Powell (University of East Anglia, UK) for many useful discussions during the progress of this study. We would also like to thank the staff of the Eye Department, Norfolk and Norwich University Hospital for their assistance with obtaining samples.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by The Humane Research Trust (UK), the Myotonic Dystrophy Support Group (UK) and the John and Pamela Salter Charitable Trust (UK).

REFERENCES

myotonia after muscleblind overexpression in a mouse poly(CUG) model for myotonic dystrophy. PNAS, 103, 11748–11753.


