

Natalia E Broude[†]
Charles R Cantor
[†]Author for correspondence
 Center for Advanced
 Biotechnology and
 Department of
 Biomedical Engineering,
 Boston University, Boston,
 MA, USA
 Tel.: +1 617 353 8490
 Fax: +1 617 353 8501
 nbroude@bu.edu

Neurological diseases and RNA-directed gene regulation: prospects for new diagnostics and therapy

'Despite ample work and enormous progress in the understanding of the etiology of neurological diseases, the precise mechanisms by which triplet repeat expansions lead to highly specific pathological development of neurons and muscle cells still remains largely unknown.'

Expert Rev. Mol. Diagn. 3(3), 269–274 (2003)

Since the discovery of the first expansions of the CGG trinucleotide repeat in the fragile X gene in 1991 [1,2], and of the CAG repeat in the androgen receptor gene linked to spinobulbar muscular atrophy [3], a new type of mutation, termed a dynamic mutation, has been recognized. Since then, the number of neurological diseases caused by dynamic mutations has grown, as has the list of short repeat sequences subject to expansion.

Despite ample work in this area and enormous progress in the understanding of the etiology of neurological diseases, the precise mechanisms by which triplet repeat expansions lead to highly specific pathological development of neurons and muscle cells still remains largely unknown. As a result, neither early diagnostics nor effective therapy are available for these diseases. However, extensive studies of the basic mechanisms leading to muscle and neurodegeneration support the hope that such diagnostics and cures will be available in the future.

How different & how common are neurological diseases at the molecular level?

There are several common features in the genetics and development of these neurological disorders. All of them are complex, progressive illnesses with multiple manifestations. Most of them are dominantly inherited and show genetic anticipation, which means that in the affected families the severity of the disease increases in subsequent generations or the age of onset decreases, or both. At the molecular level, a correlation between the

'From the diagnostic and therapeutic viewpoint, understanding of the mechanisms triggering muscle and neurodegeneration is absolutely essential for developing the most efficient ways of preventing and curing these debilitating diseases'.

severity of the illness and the length of the expanded repeat has been detected for most of the diseases [4,5]. It has also been demonstrated that the expanded repeat regions form stable DNA secondary structures that interfere with the replication machinery.

Dynamic mutations have been subdivided into two groups depending on the position of the mutation, and different mechanisms of pathogenesis have been proposed for these two

Table 1. Neurological diseases with coding location of mutations.

Disease	Repeat	Chromosome location	Inheritance type
Huntington's disease	CAG	4p16.3	AD
Spinocerebellar ataxia 1	CAG	6p23	AD
Spinocerebellar ataxia 2	CAG	12q24.1	AD
Spinocerebellar ataxia 3	CAG	14q32.1	AD
Spinocerebellar ataxia 6	CAG	19p13	AD
Spinocerebellar ataxia 7	CAG	3p12-13	AD
Dentatorubral-pallidoluysian atrophy	CAG	12p13.31	AD
Spinobulbar muscular atrophy	CAG	Xq13-21	X-linked
Oculopharyngeal muscular dystrophy	GCG	14q11	AD
Huntington's disease-like Type 2 [§]	CTG	16q24.3	AD

[§]In this disease, the CTG triplet can either be translated as polyileucine or polyalanine or untranslated depending on different splicing variants [23].
AD: Autosomal dominant.

groups. A group of diseases with exonic location of mutations, including Huntington's disease (HD) [5], is presented in TABLE 1. In all these cases, the expansions of the CAG repeat lead to proteins with expanded polyglutamine (polyQ) stretches. For the polyQ diseases, there is wide acceptance that the mechanisms leading to neurodegeneration involve abnormal polyQ protein conformations prone to form aggregates and precipitate [6-8]. Indeed, neuronal intranuclear inclusions and/or cytoplasmic aggregates containing mutant proteins have been demonstrated in transgenic mice bearing mutant HD genes, as well as in autopsied brains from patients with CAG-repeat diseases [9,10]. Recently, the repertoire of exonic mutations has been enlarged – a GCG-repeat expansion coding for the polyalanine string has been detected in the gene coding for a poly(A)-binding protein linked to the oculopharyngeal muscular atrophy (TABLE 1) [11].

A second group of dynamic mutations includes those located in the noncoding genomic regions (TABLE 2). Myotonic dystrophy (DM)1 is the best studied disease in this group [12,13]. It is a multisystemic neurological disease characterized by progressive muscle weakness, ocular cataracts and cardiac arrhythmias. In DM1, the CTG expansion is located in the 3'-untranslated region (UTR) of a protein kinase gene in the 19q13.3 region. The location of the mutation suggests that in DM1, the repeat expansion may be toxic not at the protein level but at the level of RNA synthesis, processing or transport.

A basic and very important practical question is whether a common ground exists for all known diseases caused by dynamic mutations independent of the location of the mutation. Two explanations appear *a priori* equally possible: short

Table 2. Neurological diseases with noncoding location of mutations.

Disease	Repeat	Chromosome location	Inheritance type
Fragile X syndrome	CGG	Xq27.3	X-linked
Friedreich's ataxia	GAA	9q13-21.1	AR
Myotonic dystrophy 1	CTG	19q13.3	AD
Myotonic dystrophy 2	CCTG	3q21	AD
Spinocerebellar ataxia 8	CTG	13q21	AD
Spinocerebellar ataxia Type 10	ATTCT	22q13	AD
Spinocerebellar ataxia Type 12	CTG	5q31-33	AD
Huntington's disease-like Type 2 [§]	CTG	16q24.3	AD

[§]In this disease, the CTG triplet can be either translated as polyileucine or polyalanine or untranslated depending on different splicing variants [23].
AD: Autosomal dominant; AR: Autosomal recessive.

repeat expansions in different genes rely on basically different mechanisms or there are common features in all diseases caused by triplet-repeat expansions but we do not fully understand them yet.

Who is in charge: RNA or protein, or both?

Several hypotheses have been suggested to explain dominant type inheritance in DM1 by gain-of-function at the RNA level. According to one model, triplet repeat expansions affect mRNA transport, which leads to accumulation of transcripts with CUG expansions as foci in the nucleus, while the abundance of poly(A)-mRNA in the cytoplasm decreases [13–15]. However, in DM1, the expression of several other genes upstream and downstream of the myotonic dystrophy protein kinase (*DMPK*) gene is also impaired [12,16,17]. The protein sequestration theory has been suggested to explain this phenomenon [18,19]. According to this theory, the expanded repeats within the *DMPK* transcripts specifically bind nuclear proteins and interfere with their normal function. Long-range *cis* autosomal inactivation

was suggested as another mechanism by which the expanded CTG repeats in the *DMPK* gene suppress distant genes [20].

Double-stranded RNA hairpin structures specific for the expanded CUG repeats have indeed been demonstrated *in vitro* [21] and proteins interacting with them in a length-dependent manner have been isolated and characterized [22]. Importantly, human analogs of *Drosophila* muscle-blind proteins have been identified among these proteins, which are necessary for terminal differentiation of muscle and photoreceptor cells [22]. One of the most convincing experiments pointing to a key role of the RNA repeat structures in DM1 was the introduction of an untranslated CUG repeat in an unrelated mRNA in transgenic mice, which resulted in myotonia and myopathy [23]. Another study used mouse C2C12 myoblasts transfected with reporter genes containing fragments of the 3'-UTR of *DMPK* gene with different numbers of CUG repeats [24]. Long CUG repeats resulted in decreased translation of the reporter genes and nuclear accumulation of the transcripts. These experiments clearly demonstrate a role of the expanded CUG structures at the RNA level and at least partly explain the striking similarity between the symptoms of DM1 and DM2, despite the fact that the expanded repeats are different, the affected genes are different, they are located on different chromosomes and are in different gene environments [25,26].

Recently, a more detailed mechanism involving mutant RNAs with expanded repeats have been revealed in DM1- and DM2-affected muscles. It has been demonstrated that splicing of several key proteins participating in muscle differentiation is impaired in myotonic muscles [27–29]. Thus, main chloride channel *ClC-1* is absent in DM1 and DM2 muscles [28,29], while myotubularin-related gene 1 is differentially spliced [30].

Thus, myotonic dystrophies can be considered as examples of neurological diseases where the RNA gain-of-function mechanisms caused by the triplet expansions might explain many of the available data [31,32].

Returning to the polyQ diseases, increasing data are accumulating suggesting that the separation of neurological diseases into two groups with different mechanisms of pathogenesis may not be very productive [33]. The toxic effect of the polyQ proteins may not be the only mechanism of triplet-expansion toxicity and may be not a mechanism that triggers neurodegeneration. Several lines of evidence argue against considering the polyQ neurological diseases as a distinct group and suggest a role for RNA in these diseases. The data suggesting common etiology of all triplet repeat expansion-based disorders are summarized in BOX 1.

Concluding remarks

All these data suggest the possibility that different neurological diseases share a common ground not only by having expanded repeats at the DNA level – as the underlying molecular cause – and dominant inheritance and anticipation, but also by mechanisms involving RNAs (FIGURE 1). The common molecular feature of all neurological diseases independent of the position of the expansion could be the presence in the nucleus of RNA

Box 1. Arguments supporting common etiology of dynamic mutation-based diseases and suggesting a role for RNA in polyQ neurological diseases.

- The dominant type of inheritance in most neurological diseases independent of the location of the mutation suggests common mechanisms of pathogenesis in both groups (TABLES 1 & 2).
- Protein aggregation caused by polyglutamine (polyQ) is neither necessary nor sufficient for neuronal dysfunction [7,34]. It was demonstrated, for example, that the formation of intranuclear inclusions might not be a critical step in Huntington-induced neurodegeneration [34].
- The clinical features of spinocerebellar ataxia (SCA)8 are similar to those of other SCAs. However, the SCA8 transcript with an expanded CTG repeat is an antisense RNA [35], while in all other spinocerebellar ataxias the expanded CAG repeat is translated and codes for polyQ. Thus, it apparently does not matter which type of repeat is expanded, CAG or CTG. In both cases there is significant damage of the cerebellar system. The discovery of new nontranslated mutations in SCA10 and SCA12 [36,37] and of Huntington's disease Type 2 [38] offers more support to this notion.
- Nuclear inclusions in polyQ diseases have been shown to contain transcriptional components [39]. This may be a result of interaction of transcription factors with the polyQ proteins but it may also indicate the presence of mRNA in nuclear inclusions. The presence of mutant RNAs with expanded repeats in the nucleus is much easier to explain than the presence of the polyQ proteins, which are synthesized in the cytoplasm and then transported back into the nucleus.
- It has been demonstrated that genes encoding proteins with RNA-binding domains modify SCA1 degeneration [40] and polyQ cytotoxicity can be rescued by dsRNA-mediated RNA interference [41].

transcripts with the expanded repeat regions. These regions are most likely to form hairpin structures and are capable of strong protein binding. Interference at the RNA level may be responsible for highly cell-specific gene expression regulation leading to neuronal and muscle degeneration.

Intensive studies are required to answer several basic questions. First, neuronal intranuclear inclusions in HD- or SCA1-affected brains have to be tested for the presence of RNA transcripts with expanded repeats. Second, gel retardation tests for long CUG (or CAG) tracts should be performed with cell-specific lysates in order to find proteins interacting with repeats in a length-dependent manner. Additionally, analyses of large genomic regions with expanded repeats may reveal genes for the noncoding RNAs participating in gene regulation. The first example in this direction is the finding that the SCA8 transcript is a *cis*-antisense RNA to a gene encoding a protein

likely belonging to a family of actin-organizing proteins [35]. A recently found link between fragile X protein and components of RNA interference machinery may open a new chapter in mechanisms leading to neurodegeneration, where RNA regulation plays an important role [42,43].

From the diagnostic and therapeutic viewpoint, understanding of the mechanisms triggering muscle and neurodegeneration is absolutely essential for developing the most efficient ways of preventing and curing these debilitating diseases.

Acknowledgments

The authors are thankful to Vadim Demidov, Sergei Mirkin, Arturas Petronis and Maria Carmo-Fonseca for encouragement and critical reading of the manuscript. This work was supported by a sponsored research agreement from Hamilton-Thorne Biosciences, Inc. (Beverly, MA, USA).

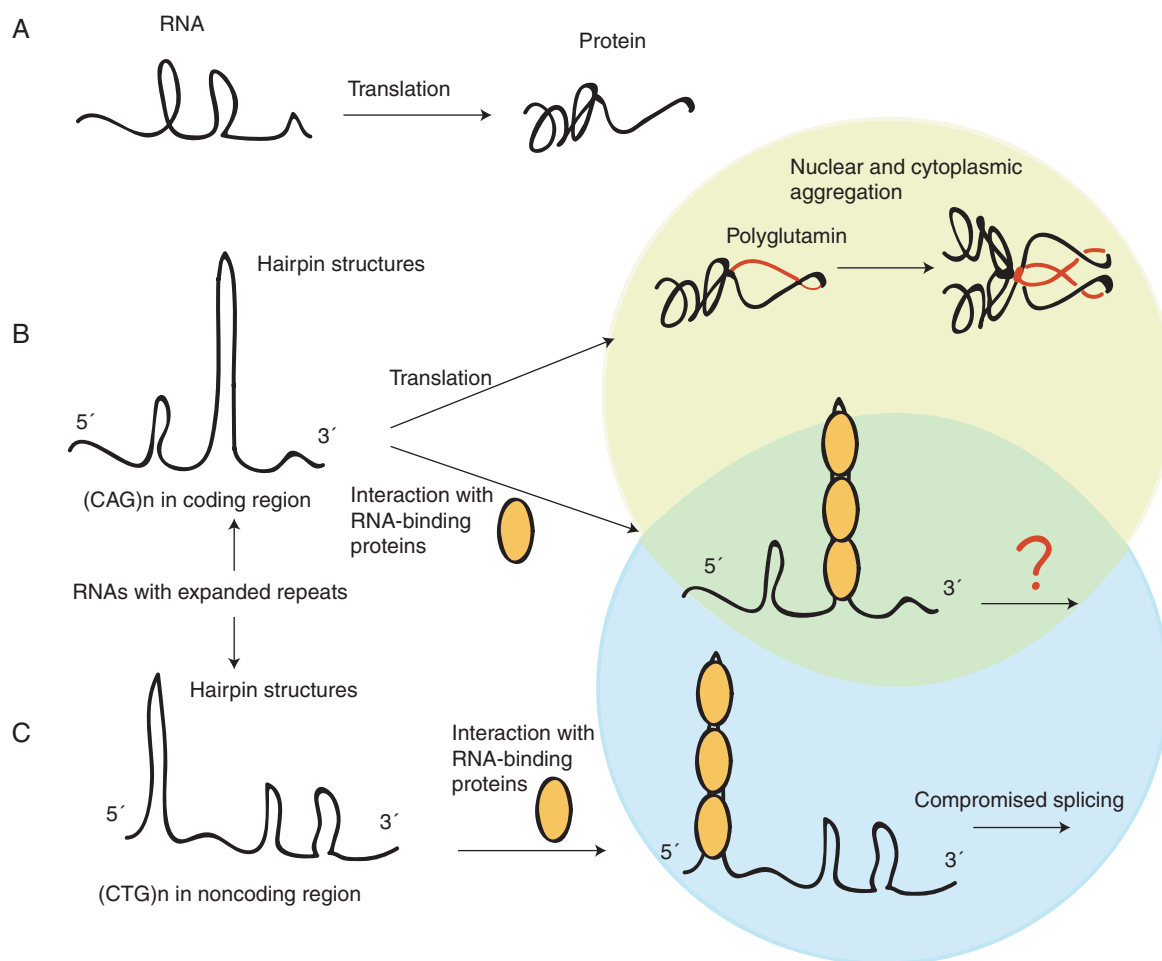


Figure 1. The expanded repeats in RNA transcripts independent of their location within the RNA molecule recruit RNA-binding proteins and participate in processes not characteristic of RNAs with short repeats [33]. (A) Normal RNA transcript is translated into normal protein. **(B)** RNA transcript with an expanded repeat within the coding region is translated into protein with an enlarged polyglutamine tract and is prone to nuclear and cytoplasmic precipitation. In parallel, RNA transcripts with the expanded repeats form RNA-protein complexes, which then participate in new RNA-protein interactions. This results in pathological development. **(C)** RNA transcripts with the repeats in noncoding regions interact with RNA-binding proteins in the nucleus and trigger aberrant splicing, which results in muscle degeneration.

References

- 1 Fu YH, Kuhl DP, Pizzuti A *et al.* Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* 67, 1047–1058 (1991).
- 2 Kremer EJ, Pritchard M, Lynch M *et al.* Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)_n. *Science* 252, 1711–1714 (1991).
- 3 La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischbeck. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* 352, 77–79 (1991).
- 4 Cummings CJ, Zoghbi HY. Trinucleotide repeats: mechanisms and pathophysiology. *Ann. Rev. Genomics Hum. Genet.* 1, 281–328 (2000).
- 5 Fischbeck KH. Polyglutamine expansion neurodegenerative disease. *Brain Res. Bull.* 56, 161–163 (2001).
- 6 Paulson HL, Bonini NM, Roth KA. Polyglutamine disease and neuronal cell death. *Proc. Natl Acad. Sci. USA* 97, 12957–12958 (2000).
- 7 Zoghbi HY, Orr HT. Glutamine repeats and neurodegeneration. *Ann. Rev. Neurosci.* 23, 217–47 (2000).
- 8 Zoghbi HY, Botas J. Mouse and fly models of neurodegeneration. *Trends Genet.* 18, 463–471 (2002).
- 9 Davies SW, Turmaine M, Cozens BA *et al.* Formation of neuronal intranuclear inclusions underlies the neurological dysfunction in mice transgenic for the HD mutation. *Cell* 90, 537–548 (1997).
- 10 DiFiglia M, Sapp E, Chase KO *et al.* Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science* 277, 1990–1993 (1997).
- 11 Nakamoto M, Nakano S, Kawashima S *et al.* Unequal crossing-over in unique PABP2 mutations in Japanese patients: a possible cause of oculopharyngeal muscular dystrophy. *Arch. Neurol.* 59, 474–477 (2002).
- 12 Larkin K, Fardaei M. Myotonic dystrophy – a multigene disorder. *Brain Res. Bull.* 56, 389–395 (2001).
- 13 Wang J, Pegoraro E, Menegazzo E *et al.* Myotonic dystrophy: evidence for a possible dominant-negative RNA mutation. *Hum. Mol. Genet.* 4, 599–606 (1995).
- 14 Taneja KL, McCurrach M, Schalling M, Housman D, Singer RH. Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J. Cell Biol.* 128, 995–1002 (1995).
- 15 Davis BM, McCurrach ME, Taneja KL, Singer RH, Housman DE. 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 (1997).
- 16 Khajavi M, Tari AM, Patel NB *et al.* 'Mitotic drive' of expanded CTG repeats in myotonic dystrophy Type 1 (DM1). *Hum. Mol. Genet.* 10, 855–863 (2001).
- 17 Timchenko NA, Iakova P, Cai ZJ, Smith JR, Timchenko LT. Molecular basis for impaired muscle differentiation in myotonic dystrophy. *Mol. Cell Biol.* 21, 6927–6938 (2001).
- 18 Caskey CT, Swanson MS, Timchenko LT. Myotonic dystrophy: discussion of molecular mechanism. *Cold Spring Harb. Symp. Quant. Biol.* 61, 607–614 (1996).
- 19 Timchenko LT, Miller JW, Timchenko NA *et al.* Identification of a (CUG)_n triplet repeat RNA-binding protein and its expression in myotonic dystrophy. *Nucleic Acids Res.* 24, 4407–4414 (1996).
- 20 Junghans RP, Ebralidze A, Tiwari B. Does (CUG)_n repeat in DMPK mRNA 'paint' chromosome 19 to suppress distant genes to create the diverse phenotype of myotonic dystrophy? A new hypothesis of long-range *cis* autosomal inactivation. *Neurogenetics* 3, 59–67 (2001).
- 21 Napierala M, Krzyosiak WJ. CUG repeats present in myotonin kinase RNA form metastable 'slippery' hairpins. *J. Biol. Chem.* 272, 31079–31085 (1997).
- 22 Miller JW, Urbinati CR, Teng-Umuay P *et al.* Recruitment of human muscleblind proteins to (CUG)_n expansions associated with myotonic dystrophy. *EMBO J.* 19, 4439–4448 (2000).
- 23 Mankodi A, Logigian E, Callahan L *et al.* Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science* 289, 1769–1773 (2000).
- 24 Amack JD, Reagan SR, Mahadevan MS. Mutant DMPK 3'-UTR transcripts disrupt C2C12 myogenic differentiation by compromising MyoD. *J. Cell Biol.* 159, 419–429 (2002).
- 25 Ranum LP, Day JW. Dominantly inherited, non-coding microsatellite expansion disorders. *Curr. Opin. Genet. Dev.* 12, 266–271 (2002).
- 26 Ranum LP, Day JW. Myotonic dystrophy: clinical and molecular parallels between myotonic dystrophy Type 1 and Type 2. *Curr. Neurol. Neurosci. Rep.* 2, 465–470 (2002).
- 27 Tiscornia G, Mahadevan MS. Myotonic dystrophy: the role of the CUG triplet repeats in splicing of a novel DMPK exon and altered cytoplasmic DMPK mRNA isoform ratios. *Mol. Cells* 5, 959–967 (2000).
- 28 Mankodi A, Takahashi MP, Jiang H *et al.* Expanded CUG repeats trigger aberrant splicing of ClC-1 chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy. *Mol. Cells* 10, 35–44 (2002).
- 29 Charlet BN, Savkur RS, Singh G, Philips AV, Grice EA, Cooper TA. Loss of the muscle-specific chloride channel in Type 1 myotonic dystrophy due to misregulated alternative splicing. *Mol. Cells* 10, 45–53 (2002).
- 30 Fardaei M, Rogers MT, Thorpe HM *et al.* 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 (2002).
- 31 Tapscott SJ, Thornton CA. Reconstructing myotonic dystrophy. *Science* 293, 816–817 (2001).
- 32 Tapscott SJ. Deconstructing myotonic dystrophy. *Science* 289, 1701–1702 (2000).
- 33 Galvao R, Mendes-Soares L, Camara J, Jaco I, Carmo-Fonseca M. Triplet repeats, RNA secondary structure and toxic gain-of-function models for pathogenesis. *Brain Res. Bull.* 56, 191–201 (2001).
- 34 Saudou F, Finkbeiner S, Devys D, Greenberg ME. Huntingtin acts in the nucleus to induce apoptosis but death does not correlate with the formation of intranuclear inclusions. *Cell* 95, 55–66 (1998).
- 35 Nemes JP, Benzow KA, Moseley ML, Ranum LP, Koob MD. The SCA8 transcript is an antisense RNA to a brain-specific transcript encoding a novel actin-binding protein (KLHL1). *Hum. Mol. Genet.* 9, 1543–1551 (2000).
- 36 Holmes SE, O'Hearn EE, McInnis MG *et al.* Expansion of a novel CAG trinucleotide repeat in the 5' region of PPP2R2B is associated with SCA12. *Nature Genet.* 23, 391–392 (1999).
- 37 Matsuura T, Yamagata T, Burgess DL *et al.* Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia Type 10. *Nature Genet.* 26, 191–194 (2000).

- 38 Margolis RL, O'Hearn E, Rosenblatt A *et al.* A disorder similar to Huntington's disease is associated with a novel CAG repeat expansion. *Ann. Neurol.* 50, 373–380 (2001).
- 39 Chai Y, Shao J, Miller VM, Williams A, Paulson HL. Live-cell imaging reveals divergent intracellular dynamics of polyglutamine disease proteins and supports a sequestration model of pathogenesis. *Proc. Natl Acad. Sci. USA* 99, 9310–9315 (2002).
- 40 Fernandez-Funez P, Nino-Rosales ML, de Gouyon B *et al.* Identification of genes that modify ataxin-1-induced neurodegeneration. *Nature* 408, 101–106 (2000).
- 41 Caplen NJ, Taylor JP, Statham VS, Tanaka F, Fire A, Morgan RA. Rescue of polyglutamine-mediated cytotoxicity by double-stranded RNA-mediated RNA interference. *Hum. Mol. Genet.* 11, 175–184 (2002).
- 42 Cady AA, Myers M, Hannon GJ, Hammond SM. Fragile X-related protein and VIG associate with the RNA interference machinery. *Genes Dev.* 16, 2491–2496 (2002).
- 43 Ishizuka A, Siomi MC, Siomi HA. Drosophila fragile X protein interacts with components of RNAi and ribosomal proteins. *Genes Dev.* 16, 2497–2508 (2002).

Affiliations

- *Natalia E Broude, PhD*
Senior Scientist
Center for Advanced Biotechnology,
Boston University,
36 Cummington St,
Boston, MA 02215, USA
Tel.: +1 617 353 8490
Fax: +1 617 353 8501
nebroude@bu.edu
- *Charles R Cantor, PhD*
Chief Scientific Officer
Sequenom Inc.,
3595 John Hopkins Court
San Diego, CA 92121, USA
Tel.: +1 858 202 9000
Fax: +1 858 202 9001
ccantor@sequenom.com