Dynamic mutations on the move in Banff

Albert R La Spada, Robert I Richards & Bé Wieringa

Earlier this year, the 4th International Meeting on Unstable Microsatellites and Human Disease highlighted how far and fast the research of diseases associated with expanded repeats has advanced, and spotlighted the remaining recalcitrant problems.

The mechanisms involved in microsatellite repeat instability and the consequent pathways that lead to disease have fascinated investigators from a wide range of disciplines for over a decade. This fascination was sparked by the discovery of dynamic mutations as the cause of a number of important and enigmatic human genetic diseases, including fragile X syndrome (FRAXA), myotonic dystrophy and Huntington disease. The fourth in a series of biannual meetings on Unstable Microsatellites and Human Disease* was recently held in Banff, Canada, and approximately 120 scientists gathered to discuss progress in understanding the mechanisms of repeat expansion, the molecular basis of disease pathology and possible ways to treat these disorders.

Mechanisms of repeat expansion

A primary focus of current research that was discussed at the conference is the molecular basis of microsatellite repeat instability. Both the composition of the repeat sequence and the length of the repeat are primary determinants of the likelihood that a repeat will expand. Recent work has focused on genomic context, and additional evidence was presented to reinforce the emerging view that sequences outside the repeat regulate instability. Sarah Ennis (of Patricia Jacobs' laboratory) presented the results of an ambitious linkage disequilibrium analysis at the *FRAXA* locus that pinpointed a region ~50 kb upstream of the gene as a key determinant for repeat instability of *FRAXA* alleles. Al La Spada presented evidence of a particular sequence motif (a CTCF binding site) in the vicinity of the spinocerebellar type 7 (SCA7) CAG tract that may control repeat instability at this locus, emphasizing the probable importance of chromatin-packaging changes in regulating repeat instability. Evidence for a reciprocal mechanism (*i.e.*, influence of repetitive DNA sequences on the potential formation of heterochromatin, as in position effect variegation) was also presented. Richard Festenstein showed that relatively short GAA (Friedreich ataxia) or CTG (myotonic dystrophy) repeats can recruit heterochromatin-like silencing, which is sensitive to heterochromatin protein 1 dosage, regardless of the site of integration of the transgene containing the triplet repeat in the mouse genome. The relationship between microsatellite sequences and fragile sites in cancer further illustrates the complex role of repeat elements in epigenetic regulation. Rob Richards reported that DNA instability at common fragile site loci in cancer cells is associated with *de novo* generation of repeat sequences. Richards proposed that DNA instability associated with fragile sites could be an adaptive response to genotoxic stress involving both chromosomal fragile sites and the genes that span them.

In addition to homing in on potential *cis*-acting 'instability elements', numerous presentations at the Banff Conference dealt with the identification and characterization of *trans*-acting factors involved in repeat instability. In proliferating cells, initiation of DNA replication and regulation of the progression of the bidirectional replication fork require the action of distinct protein complexes composed of many different proteins, of which more than 30 are now known. The probability of stalling of the replication machinery on anomalous repeat structures is determined by the nature of the DNA helix, the chromatin packaging in the region and the number of proteins available to the locally active replication machinery. A number of presenters provided some clues as to which proteins might be involved in the resolution and processing of DNA stretches composed of tandem trinucleotide repeats by modeling instability in the yeast *Saccharomyces cerevisiae*. Yeon-Soo Seo presented work implicating dysfunction of the Dna2 protein in triplet-repeat instability during processing of Okazaki fragments consisting of trinucleotide repeats. Such processing also requires flap endonuclease 1 and the single-stranded DNA binding protein RPA, a guidance and assembly protein. Data from Robert Lahue and Guy Richard Franck argued for involvement of Srs2 helicase as a safeguard against expansion. Although these replication dependent processes were studied in cycling cells (yeast), the prospect of their involvement in postmitotic or growth-arrested noncycling cells is conceivable, as the action of mismatch repair enzymes or endonucleases on aberrant microsatellite-expanded DNA structures may provide a point of entry (**Fig. 1**). Some meetings ago, Anne Messer's group pointed out that tissue repeat instability (somatic mosaicism) in a mouse model of Huntington disease requires the presence of the mismatch repair protein Msh2. The evidence for participation of the mismatch repair proteins Msh2 and Msh3 (and perhaps others) in the recognition and

Albert R. La Spada is in the Department of Laboratory Medicine at the University of Washington, Box 357110, Room NW120, Seattle, Washington 98195-7110, USA. Robert I. Richards is at the Centre for Molecular Genetics of Development, The University of Adelaide, Victoria Drive, Gate 8, Adelaide, SA, 5005 Australia. Bé Wieringa is in the Department of Cell Biology, NCMLS, UMC University of Nijmegen, PO Box 9101 6500 HB, Nijmegen, The Netherlands.
e-mail: laspada@u.washington.edu

processing of putative abnormal helical structures of repeat-expanded DNA remains strong, as was highlighted in presentations from Genevieve Gourdon, Cynthia McMurray and Be Wieringa. It is now known that the molecular mechanisms involved in somatic repeat instability are developmentally and cell-cycle regulated and have species-specific characteristics. Future progress in this area will depend on our ability to deconstruct the various mechanistic steps and to compare them between different models (yeast, flies, mice, etc.) of human repeat instability. As a first step towards this goal, Christopher Pearson, working with a human cell-free system to process artificially generated DNA slipped-strand structures, showed that the processing outcome varied according to the DNA composition of the slipped strand. Systematic testing of individual protein components in such cell-free systems seems to be the logical next step.

A common molecular pathway

In the area of disease pathogenesis, the repeat expansion field has witnessed a coalescence of various diseases into certain classes according to their molecular basis. For example, for the nine diseases associated with CAG polyglutamine expansion, pathogenesis stems from the production of proteins containing abnormally long glutamine tracts whose misfolding and accumulation seem to be the crux of the initiating pathology. At least 11 other diseases are caused by the expansion of repeats that are incapable of encoding polyglutamine, a few of which are actually clinically indistinguishable from certain diseases associated with polyglutamine expansion.

In the case of myotonic dystrophy type 1 (DM1), the most common form of muscular dystrophy in humans, the reason that CTG repeat expansions in the 3′ untranslated region of a protein kinase gene are harmful was not known for years after their initial discovery. The existence of a nearly indistinguishable second genetic form of myotonic dystrophy (DM2) offered an opportunity to gain insight into the molecular pathogenesis of both DM1 and DM2. When DM2 was found to be caused by a CCTG expansion in the first intron of the gene *ZNF9*, investigators focused on mutant *DM1* and *ZNF9* (also called *DM2*) RNAs containing abnormally large CUG and CCUG repeats, respectively. The RNA gain-of-function hypothesis thus took center stage. In two studies done in mice, first overexpression of a CUG repeat expansion tract and then elimination of synthesis of the CUG repeat binding protein muscleblind-like 1 (MBNL1) were shown to be sufficient to recapitulate several characteristic aspects of the myotonic dystrophy phenotype. The latest progress in the molecular dissection of myotonic dystrophy was provided by Maurice Swanson, who explained that MBNL1 may act as a splicing factor and compete with the CUG-binding protein CUGBP1, a member of the CELF family of proteins, to regulate alternative splicing of specific exons. This competition could be a key point of regulation for the control of alternative splicing during development. Thus, in DM1 and DM2, the balance of alternative splicing may be aberrantly tipped, owing to the accumulation and sequestration of MBNL1 in nuclear ribonucleoprotein foci containing CUG expansions, resulting in a widespread splicing derangement.

One potential problem with the currently existing mouse models of myotonic dystrophy is that the tissue distribution patterns of repeat overexpression and factor depletion do not match the distribution profiles of expanded *DMPK* (DM1) or *ZNF9* (DM2) alleles in humans. How do we study certain classic features of myotonic dystrophy, such as cognitive impairment, if the mouse models do not accurately recapitulate where pathologically expanded repeats accumulate in humans? Charles Thornton showed that one can overcome this problem by directly studying material from affected individuals. He presented data showing the existence of prominent nuclear ribonucleoprotein foci in cerebral structures (but not in the cerebellum) of individuals with congenital DM1,

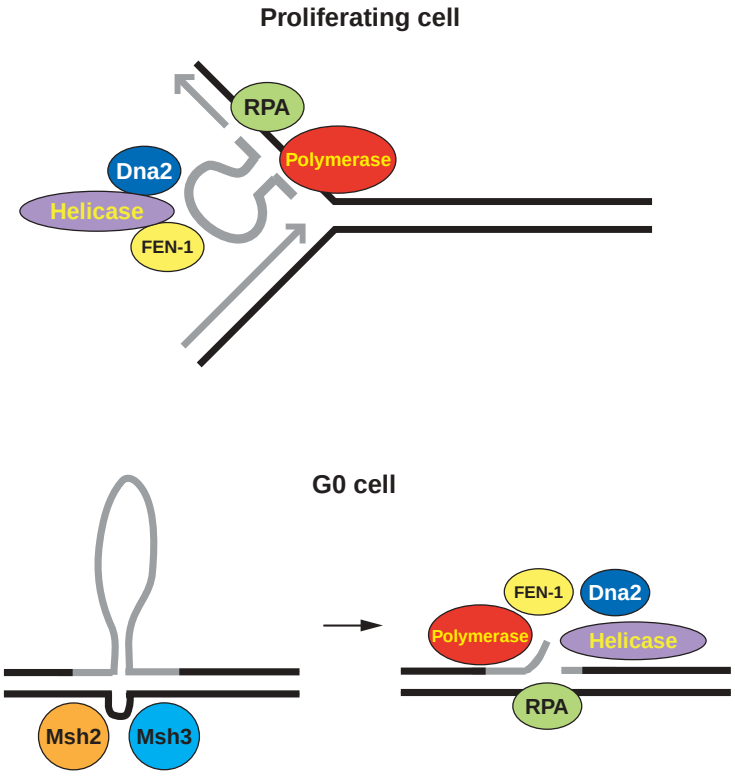

Figure 1 Models of microsatellite instability. Do repair and replication enzymes have a fundamental role? The tendency of expanded microsatellite DNA tracts to form anomalous DNA structures (slipped strand structures, loop-outs, cruciforms, etc.) is a key prerequisite for dynamic repeat instability. In cycling cells, repeat tracts may form single-stranded folded flap structures on Okazaki fragments during discontinuous lagging-strand replication or formation of 'chicken foot' structures in stalled replication forks. For resolution or repair of these structures, recognition by RPA, unwinding by helicases (such as Srs2 / Sgs1) and processing by other proteins with nuclease activity (*e.g.*, FEN-1, Dna2 or a DNA polymerase such as δ or ε) are needed. In noncycling cells (G0 cells), these structures are recognized by Msh2 and Msh3 and perhaps by other components of the mismatch repair machinery. Many investigators in the instability field posit aberrant replication-resolution or repair activity as a key event in the repeat expansion process. Cell-cycle or differentiation phase, cell type, age and the nature of the repeat structure may account for differences in the composition of the recruited DNA replication and repair machinery and thereby determine whether repeat tracts change in length and, if so, in what direction and to what extent.

noting that aggregate formation leads to depletion of MBNL from the neuronal nucleoplasm. Aggregates formed in both oligodendrocytes and in neurons and seem to also contain other members of the MBNL protein family, such as muscleblind-like 2. His analysis of the nuclear RNA aggregates detected proteasome subunits, raising the possibility that disorders characterized by nuclear protein aggregates and mutant RNA foci may have more in common than initially surmised (**Fig. 2**). Thornton closed his presentation by suggesting that preliminary evidence for abnormal splicing of the tau protein and NMDA receptor in the brains of individuals with congenital DM1 could be a clue to the basis of central nervous system involvement. Another interesting demonstration that abnormal RNAs are the culprit in a repeat expansion disease was provided by Laura Ranum, whose laboratory has generated a number of *SCA8* transgenic mouse models by introducing BACs containing *Klhl1as* (also called *Sca8*) with a disease-associated 107-CTG repeat. Although certain aspects of SCA8 disease causality are controversial, the expanded *Klhl1as* gene (which is transcribed into RNA but encodes no protein!) was sufficient to produce dosage-dependent neurological disease. In high-copy number lines, a progressive and fatal phenotype results, whereas in low-copy number lines, motor dysfunction characterized by an unsteady gait is observed. Ranum thus provided the final vertex in what she called the 'Bermuda triangle' group of diseases associated with RNA CTG (CUG)-repeat gain of function.

Innocuous premutations no longer

That the RNA gain-of-function toxicity model applies to more than just CTG expansions has also recently come to the forefront in the repeat disease field. Paul Hagerman and colleagues have been instrumental in describing a neurological disorder of intention tremor, gait ataxia and parkinsonism that occurs in male carriers of a *FRAXA* premutation. The Hagerman laboratory proposed that this fragile X-associated tremor and ataxia syndrome (FXTAS) is due to a toxic gain-of-function of expanded *FRAXA* RNAs with premutation-size CGG repeats (55-200 CGG repeats) that accumulate in intranuclear inclusions in the brains of affected individuals. Hagerman reported on biophysical approaches to purify and analyze inclusion material from individuals with FXTAS and from the FXTAS mouse model with Fourier-transform mass spectroscopy. His preliminary data were suggestive of differences in protein composition, as well as neural cell-type distribution and modification, between mouse and human. Peng Jin (of Steve Warren's laboratory) provided an update on their model of FXTAS and on the function of FRAXA as a translational repressor, two lines of investigation that used *Drosophila*. They reported that FRAXA binding to mRNA creates a 'core' that recruits additional factors such as Dicer, EIF2C2, RISC and components of the microRNA pathway. Assembly of this ribonucleoprotein complex in the cytosol leads to translational suppression due to ribosome stalling and is an aspect of 'normal' FRAXA function. Loss of this function results in classical FRAXA mental retardation in individuals with full mutations; an altogether different event, production of premutation-expanded FRAXA mRNAs, yields the FXTAS phenotype.

The prospects for therapy

Given the advances in understanding the mechanisms of disease pathogenesis in recent

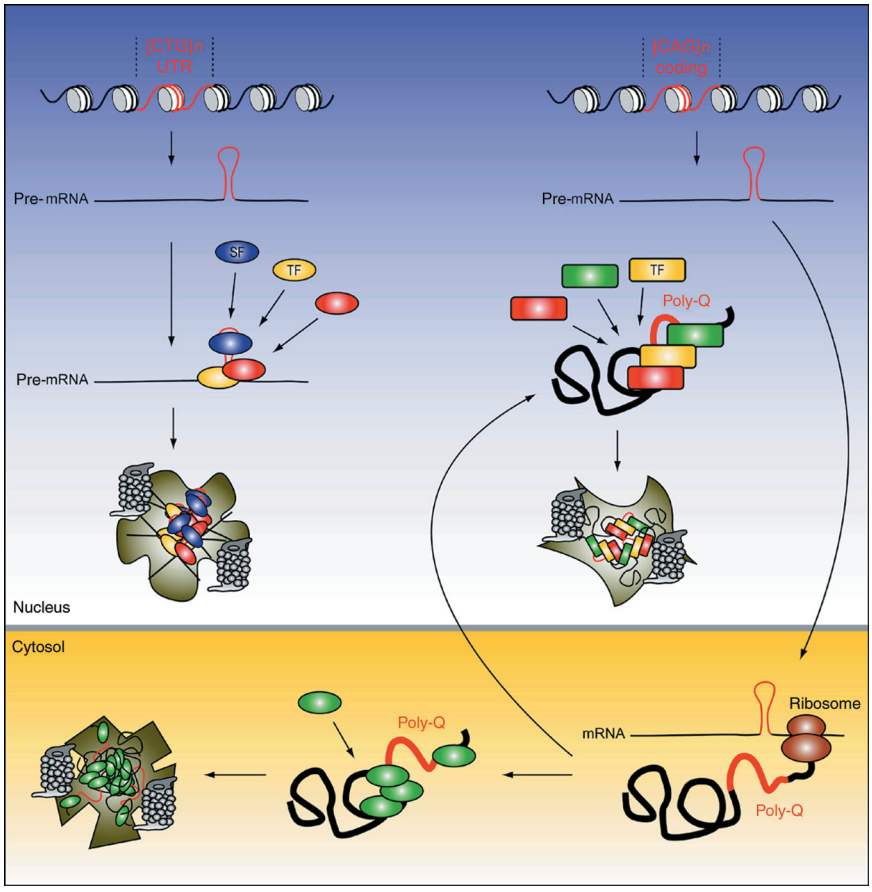

Figure 2 Aggregate formation in diseases associated with RNA expansion and with CAG or polyglutamine (poly-Q) expansion. Work presented at the Banff Conference suggests that different diseases (DM1, DM2, FXTAS, SCA8 and perhaps even SCA10) are characterized by RNAs with repeat expansion segments in their untranslated regions (UTRs). These RNAs may recruit nuclear factors (such as splicing factors (SF) and transcription factors (TF)), ultimately resulting in the formation of large ribonucleoprotein aggregates (RNA foci). Attempted degradation of RNA foci by the cell is indicated by the presence of ubiquitin and components of the proteasome. In disorders associated with CAG or polyglutamine repeats (*e.g.*, Huntington disease and certain SCAs), the production of a protein with an abnormal conformation is thought to precipitate aggregate formation. Depending on the subcellular routing of the polyglutamine-expanded proteins, different aggregation pathways, resulting in complexes of different make-up, yield inclusions in the cytosol and nucleus. The composition of the aggregates may also vary according to cell type; ubiquitin and proteasomal subunits are typical components. A parallel process of aggregation thus seems to characterize diseases associated with RNA expansion and with polyglutamine expansion, suggesting that (ribonucleo)protein inclusions may act as 'sinks' where essential factors are sequestered and ultimately depleted from the cell, resulting in toxicity. The susceptibility of neurons to this process is a common theme, not just for these diseases, but also for a wide range of neurological diseases, and is probably a feature of the aging process of this cell type.

years, much attention has turned to how this knowledge can be used to develop useful therapies. One important question is how resilient injured neurons are in neurodegenerative diseases. Using the powerful approach of tetracycline-regulated transgene expression, Harry Orr found that severely dysfunctional Purkinje cells with considerable amounts of accumulated mutant ataxin-1 protein can nearly fully recover after shutting off expression of the transgene encoding mutant ataxin-1. Functional repair of Purkinje cells in the spinocerebellar ataxia type 1 mouse model, as presented by Orr, together with previously published work on conditional reversal in a mouse model of Huntington disease, bode well for effective therapeutic intervention in even advanced-stage disease. Optimism about developing the possible cures for repeat disorders also prevailed at the meeting. Anne Messer described a strategy involving the use of intrabodies (single-chain Fv antibodies) directed to amino acid epitopes adjacent to the glutamine expansion tract in the disease-causing mutant huntingtin protein in Huntington disease. Potential approaches using ribozyme- or siRNA-targeted therapies in myotonic dystrophy were explained by Jack Puyrimat, and Bronwen Connor presented some data (assembled with colleagues Richard Faull and Maurice Curtis) suggesting that progenitor cell proliferation and increased neurogenesis do occur in diseased adult human brain. Consequently, one can envision activating 'self-repair' pathways once we can block production of toxic oligomers. Following a completely different line of thinking, Bob Korneluk proposed another approach that may hold therapeutic promise. He presented data on the use of XIAP, an antiapoptotic protein known to be capable of inhibiting both extrinsic and intrinsic apoptotic pathways. Korneluk argued that targeted increases in XIAP could protect against degeneration in neuromuscular tissues, as he showed for the *mdx* mouse model (of Duchenne muscular dystrophy). Whether XIAP will be a panacea for diseases associated with ribonucleoprotein inclusions has yet to be determined; his group intends to test XIAP delivery in existing fly and mouse models for myotonic dystropyhy, FRAXA and Huntington disease to find out.

Unanswered questions

One question that remained largely unanswered during the conference is how incremental increases in repeat length are coupled to graded effects in disease severity and progression in microsatellite-associated disorders. Here we need a convergence of molecular and cellular studies, because persistently altered cellular dysfunction and an increasing propensity for apoptotic activation probably combine to ultimately lead to irreversible neurodegeneration. Understanding the myriad pathological cascades may require that we inventory the biological functions of the components involved in the formation of aggregates to identify the key players. In light of powerful emerging genomic and proteomic technologies, this may be feasible. Approaches in model organisms provide alternate genetic strategies that, when combined with cataloguing of involved factors, may permit more rapid progress on this front. For future design of therapeutic strategies, thorny questions still remain. Why are there threshold effects in repeat instability and in ribonucleoprotein inclusion processes? How are the events involved in the formation of the 'crystallization cores' and expanding foci ordered, and why is there cell-type specificity (with neurons seeming to be particularly susceptible)? Progress is also needed in understanding the ultimate fate of the RNA and protein constituents, as well as the different cellular stress responses. Finally, it would be of interest to learn whether components are continuously accumulating or whether there is an 'in-transit' flux of material through nuclear and cytoplasmic inclusions.

Summary

Because a universal pathogenic pathway for diseases that have a common mutation type and, often, a common pathology has thus far been elusive, most investigators classify the diseases associated with unstable repeats according to the nature of the mutation mechanism (loss of function, protein gain of function, RNA gain of function). The emerging (and somewhat surprising) role of RNA as the pathogenic agent in DM1, DM2, SCA8 and FXTAS allows the inevitable speculation that RNA accumulation and coalescence of abnormal RNA-protein complexes could be involved in other dominant dynamic mutation disorders not associated with polyglutamine expansion. Mechanistically similar, but uniquely protein-based, mechanisms may have a key role in diseases associated with polyglutamines, although involvement of RNA has not been completely excluded for all disorders in this category. Certainly, a role for RNA toxicity in the polyglutamine-associated disease SCA6 and perhaps even in the Huntington disease-like-2 repeat disorder seems plausible. As for the basis of repeat instability, the interplay between *cis* elements and *trans*-acting factors is still shrouded in mystery, and so there is much work to be done. One thing can be concluded with certainty, however. In a field that has been yielding surprises for more than a decade, there are many more in store for the participants of the next International Conference on Unstable Microsatellites and Human Disease, to be held in 2006. The place at which the investigators will then aggregate is yet to be localized, so stay tuned! ■

**The 4th International Conference on Unstable Microsatellites and Human Disease was held in Banff, Canada, from 28 February to 4 March 2004. Meeting details are available at http://www.microsatellites.ca/.*