Thus, behavioral and metabolic responses of bacteria to the complex and heterogeneous structure of the organic matter field at the microscale influence ocean basin-scale carbon fluxes in all major pathways: microbial loop, sinking, grazing food chain, carbon storage, and carbon fixation itself. However, studying such varied influences of bacteria on organic matter, and their spatial-temporal variations, in piecemeal fashion will only result in a conceptual patchwork without a unifying framework or predictive power. A unifying theme should derive from applying robust principles of biochemical adaptation in a realistic microenvironmental context. Biogeochemical variability could then be considered as a consequence of adaptive responses to (micro)environmental variations. This approach should also serve as a framework to understand the maintenance of microbial diversity and to make predictions on the survival of specific bacterial species, including human pathogens such as *Vibrio cholerae*, in response to ecosystem perturbations (*21*). This framework, which includes bacteria-algae interactions, should also be relevant to the prediction of algal blooms, including toxigenic species. Powerful new approaches are enabling us to study microbial ecology, including consortial activities, in an ecosystem context. New techniques allow multiple interrogations—phylogeny, metabolism, growth—at the individual cell level. These ideas and approaches should lead to a synthesis of bacterial adaptation, evolution, ecology, and biogeochemistry, and should form a basis for integrating the roles of bacteria in predictive biogeochemical models.

References and Notes

1. W. C. Chin *et al.*, *Nature* **391**, 568 (1998).
2. F. Azam *et al.*, *Microbiol. Ecol.* **28**, 167 (1993).
3. F. Azam and D. C. Smith, in *Particle Analysis in Oceanography*, S. Demers, Ed. (Springer-Verlag, Berlin, 1991), pp. 213–236.
4. I. Koike *et al.*, *Nature* **345**, 242 (1990); M. L. Wells and E. D. Goldberg, *ibid.* **353**, 342 (1991); R. A. Long and F. Azam, *Aquat. Microb. Ecol.* **10**, 213 (1996).
5. J. H. Steele, in *The Structure of Marine Ecosystems* (Harvard Univ. Press, Cambridge, MA, 1974).
6. L. R. Pomeroy, *Bioscience* **24**, 499 (1974); J. E. Hobbie *et al.*, *Appl. Environ. Microbiol.* **33**, 1225 (1977).
7. F. Azam *et al.*, *Mar. Ecol. Prog. Ser.* **10**, 257 (1983).
8. J. J. Cole *et al.*, *ibid.* **43**, 1 (1988); J. A. Fuhrman and F. Azam, *Mar. Biol.* **66**, 109 (1982); A. Hagstrom *et al.*, *Appl. Environ. Microbiol.* **37**, 805 (1979).
9. H. W. Ducklow and C. A. Carlson, *Adv. Microb. Ecol.* **12**, 113 (1993).
10. H. W. Ducklow, *Deep Sea Res.* **40**, 753 (1993).
11. N. Williams, *Science* **279**, 483 (1998).
12. L. R. Pomeroy and D. Deibel, *ibid.* **233**, 359 (1986).
13. A. L. Alldredge *et al.*, *Deep Sea Res.* **40**, 1131 (1993).
14. J. G. Mitchell *et al.*, *Appl. Environ. Microbiol.* **61**, 4436 (1995); H.-P. Grossart and F. Azam, paper presented at the Ocean Sciences Meeting, San Diego, CA, 9 to 13 February 1998.
15. H. G. Hoppe *et al.*, *Mar. Ecol. Prog. Ser.* **93**, 277 (1993); J. T. Hollibaugh and F. Azam, *Limnol. Oceanogr.* **28**, 104 (1983).
16. E. F. DeLong *et al.*, *Limnol. Oceanogr.* **38**, 924 (1993); A.-S. Rehnstam *et al.*, *FEMS Microb. Ecol.* **102**, 161 (1993); J. Martinez *et al.*, *Aquat. Microb. Ecol.* **10**, 223 (1996).
17. H. Nissen *et al.*, *Mar. Ecol. Prog. Ser.* **16**, 155 (1984); G. G. Geesey and R. Y. Morita, *Appl. Environ. Microbiol.* **38**, 1092 (1979).
18. J. Stern, personal communication.
19. D. C. Smith *et al.*, *Nature* **359**, 139 (1992); A. L. Alldredge and C. C. Gotschalk, *Cont. Shelf. Res.* **10**, 41 (1990).
20. D. C. Smith *et al.*, *Deep Sea Res. II* **42**, 75 (1995); U. Passow and A. L. Alldredge, *ibid.*, p. 99.
21. P. R. Epstein, *Biosystems* **31**, 209 (1993); R. R. Colwell, *Science* **274**, 2025 (1996).

22. I thank H.-P. Grossart, R. A. Long, L. B. Fandino, K. D. Bidle, D. C. Smith, J. T. Hollibaugh, and F. Rassoulzadegan for discussions and suggestions. Research supported by NSF grants NSF OPP96-17045 and OPP95-40851.

BIOMEDICINE

Triplet-Repeat Transcripts: A Role for RNA in Disease

Robert H. Singer

A set of baffling human diseases—including myotonic dystrophy and Huntington's disease—are caused by expansion of a repeated sequence of three nucleotides within almost a dozen genes identified to date (*1*). With each generation, these repeats are replicated and the sequence gets longer, eventually compromising the function of the gene. The effect is dominant—only one of the two alleles of the gene need be expanded to result in the full pathology. Furthermore, the severity of the disease can be proportional to the length of the repeat expansion.

Unlike most genetic diseases, which are a result of a mutation that impairs or eliminates a gene product, the triplet-repeat diseases are likely caused by a "gain of function," in which a new function arises from the genetic defect. The new function is most easily understood when the expansion of the triplet CAG, which encodes the amino acid glutamine, occurs within the coding frame of a gene and creates a new protein with a polyglutamine tract. Huntington's disease is one example of such an expansion and is typical in that it exhibits a central nervous system pathology, as do all the polyglutamine diseases. Normal individuals can withstand a few repeats of glutamine at this position in their genes. As the polyglutamine expands, however, it disrupts the protein and affects cellular functions, possibly due to the high charge density of the expanding repeat. The new, gain-of-function protein can wreak havoc on cellular processes such as nuclear export, RNA and DNA binding, or membrane transport.

Expansions of these triplet repeats can also occur outside the coding region, but in these instances new proteins are not produced. Within this group of disorders, myotonic dystrophy is a particular curiosity. In this disease, the expansion occurs in the untranslated region of the transcript, after protein coding has occurred, and can increase the mRNA by 6 kilobases or more. As this expansion gets progressively larger in one of the alleles, the resulting pathology of the disease becomes proportionately more severe. This behavior begs for a new model of molecular cytopathology. Such a model is provided by Philips, Timchenko, and Cooper on page 737 of this issue. They propose that the gain of function in myotonic dystrophy is at least in part a result of disrupted activity of a CUG-binding protein induced by the repeats in the RNA, which prevents it from doing its normal job of splicing a certain family of genes.

The new proposal is not the first; various models have been readily forthcoming since the first description of these diseases. None has been sufficient to explain the molecular etiology. The expansion most likely occurs initially in the germ cells or early embryos, where the DNA polymerase may "stutter" on the repeats and, in doing so, expand them. Some models rely on DNA-based mechanisms to explain the pathology of these diseases—changes in chromatin organization because of nucleosome positioning, stalling of the RNA polymerase, or suppression of other genes nearby. But these models cannot explain the trans-dominant effect of the allele, the effect of one abnormal gene on the functioning of the whole cell.

Another DNA mechanism is possible.

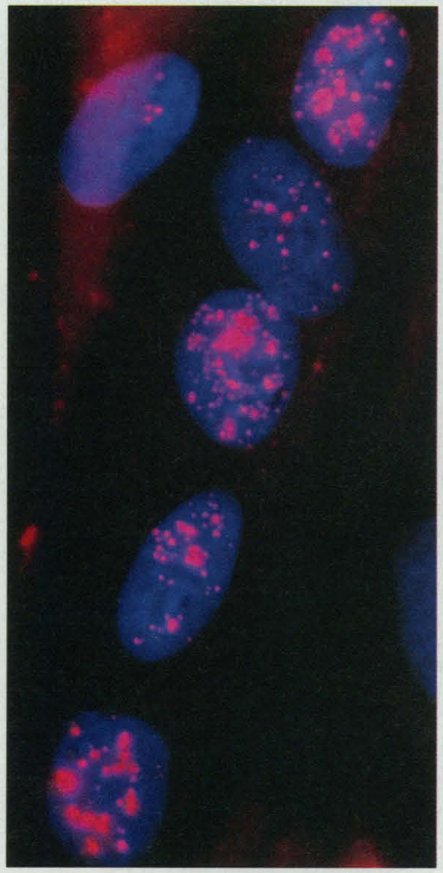
The author is in the Department of Anatomy and Structural Biology and Cell Biology, Institute for Molecular Medicine, Albert Einstein College of Medicine, Bronx, NY 10461, USA. E-mail: rhsinger@aecom.yu.edu

The disrupted expression of one allele could result in haplo-insufficiency of the protein (not enough protein is produced because in effect there is only a haploid dose of the gene); in the case of myotonic dystrophy the protein kinase (DMPK) synthesized from the remaining allele would be present in the cell in half the normal amount. It is difficult, however, to imagine how such extreme variability in the disease—from mild throughout adulthood to fatal early in life—could result from a protein that varies from a half to full dose. In addition, no point mutations in this gene reproduce the disease, indicating that the pathology is not primarily due to inadequacies of the protein product. [Data from animal models that resemble myotonic dystrophy suggest that a disruption of one or both *DM* genes can cause some muscle pathology (*2*), but gene dosage alone cannot explain the entire effect.]

The most likely explanation for the pathology of myotonic dystrophy is that the expansion in the untranslated region of the DMPK mRNA is like the expansion in the coding region: It leads mainly to a gain of function. Transcripts from the expanded gene accumulate in the nucleus of both cultured cells and biopsied tissues from patients with myotonic dystrophy (see the figure) (*3*). In the nucleus, these transcripts cannot make the appropriate cytoplasmic protein, and therefore their sequestration results in a haplo-insufficiency. But in addition, these repeat expansions could also cause a new problem: They can build up within the nucleus, possibly to toxic levels, exerting a trans-dominant effect on cellular processes. Their influence would be particularly acute in differentiated cells, such as muscle and nerve, which not only express high levels of this transcript, but which no longer divide and cannot therefore dispose of this accumulated RNA during the breakdown of the nucleus during mitosis. If the toxicity results from the burden of these excess CUG repeats, the correlation between disease severity and repeat length is easy to explain—more CUG repeats cause a more severe gain-of-function phenotype.

But how might the repeats result in a gain of function? One hypothesis proposes that the repeats bind to a protein required for normal cellular function (*4*), and in fact they are in the nucleus in an insoluble form, not easily extracted by reagents that dissolve nuclear components such as DNA or DNA-bound proteins (*5*). And the new results by Philips *et al.* relate protein binding to the repeat to a specific gain of function. The authors implicate a CUG-binding protein (CUG-BP) in the specific splicing of a family of genes containing a few CUG repeats near an exon; one member of this category is the cardiac form of troponin T (*6*).

Two patients with myotonic dystrophy were analyzed, one with congenital disease and the other homozygous for the expansion. Cells from both patients contained increased amounts of an aberrant transcript of cardiac troponin T (cTNT) that inappropriately included exon 5, normally spliced out during embryogenesis. A constitutively spliced minigene containing a genomic fragment containing exon 5 of the human cardiac TNT was transfected into normal or DM primary muscle cells. Transcripts containing exon 5 were increased in DM cells relative to normal cells, and this increase was abolished when the CUG was mutated to CAG. The inclusion could be mimicked in normal cells by expressing the CUG-BP, or by transfecting increasing amounts of repeats, along with the minigene. Further evidence that the CUG-BP is the trans-acting factor comes from its in vitro binding to CUG motifs located within a 34-nucleotide region downstream of the alternative exon. Interestingly, an expansion in an intron may also disrupt splicing in the newest addition to the repeat diseases, Friedrich's ataxia (*7*). It is not known whether RNA-binding proteins may be implicated in this or other triplet-repeat diseases.



Stuck in the nucleus. In situ hybridization in muscle cells derived from a patient with myotonic dystrophy reveals aggregates of a trinucleotide repeat in the nucleus. See (*3*, *5*) for details. [Photo: K. L. Taneja]

All of this suggests that something is going on in the nucleus of cells from these patients with myotonic dystrophy. Possibly, this nuclear CUG-BP may be titrated out by the excess nuclear CUGs of the expansion, resulting in defective regulation of those genes that depend on the CUG-BP for their proper processing. Yet, because increased CUG-BP is associated with the disease, it is difficult to reconcile this mechanism with a trans-dominant model in which repeat length is correlated with disease severity. The CUG-binding protein accumulates in the nucleus of affected cells from patients (*8*), and this may be correlated in some as yet unknown way with the repeat expansion. Additional modifications to the CUG-BP, or its isoforms, such as phosphorylation may interfere with its intrinsic splicing activity. The CUG-BP is present in most cells of the organism, consistent with the multisystemic nature of the disease—both neuromuscular pathology and myocardiopathy are well described in this disease. The results predict that exons that contain this common CUG-repeat motif required for splicing would be present in genes expressed in nervous tissue as well.

Still unexplained is how the mechanism may relate to disease pathology and whether it is a cause or an effect. Although there is nuclear build-up of CUG repeat transcripts, the CUG-BP should colocalize with these foci of transcript aggregation (*4*)—which might be expected of a CUG-binding protein. Intranuclear inclusions of protein, however, are found in brain neurons of patients with Huntington's disease and in a transgenic animal model of the same disease, also caused by a repeated polyglutamate tract (*9*). A common feature of triplet-repeat diseases may be protein or protein-RNA complexes in the nucleus. All of these considerations solidify the gain-of-function model, but still leave room for many more questions to be answered.

References

1. P. S. Reddy and D. E. Housman, *Curr. Opin. Cell Biol.* **9**, 364 (1997).
2. S. Reddy *et al.*, *Nature Genet.* **13**, 325 (1996).
3. K. Taneja, M. McCurrach, M. Schalling, D. E. Housman, R. H. Singer, *J. Cell Biol.* **128**, 995 (1995).
4. L. T. Timchenko *et al.*, *Nucleic Acids Res.* **24**, 447 (1996).
5. B. M. Davis. M. E. McCurrach, K. L. Taneja, R. H. Singer, D. E. Housman, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7388 (1997).
6. A. V. Philips, L. T. Timchenko, T. A. Cooper, *Science* **280**, 737 (1998).
7. V. Campuzano *et al.*, *ibid.* **271**, 1423 (1996).
8. R. Roberts *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 13221 (1997).
9. B. W. Davies *et al.*, *Lancet* **351**, 131 (1998).