Either way, we should be careful neither to conflate real clinical science with mere testing of technologies (such as drug trials) nor with outcomes research (*7*, *8*). We need all these activities, and there is room for real intellectual innovation in the way we carry them out. But clinical science remains distinct—namely, solving disease based on the experience of seeing, thinking about, and treating individual patients. How did we forget?

References and Notes

1. J. L. Goldstein, M. S. Brown, *J. Clin. Invest.* **99**, 2803 (1997).
2. D. F. Horrobin, *J. R. Soc. Med.* **93**, 341 (2000).
3. _____, *Lancet*, in press.
4. J. Drews, S. Ryser, *Nature Biotechnol.* **15**, 1318 (1997).
5. L. E. Rosenberg, *Science* **283**, 331 (1999).
6. D. J. Weatherall, *Br. Med. J.* **302**, 1002 (1991).
7. M. Angell, *N. Engl. J. Med.* **342**, 1516 (2000).
8. D. A. Grimes, K. F. Schulz, *Lancet* **359**, 57 (2002).
9. H. A. Simon, *Models of My Life* (Basic Books, New York, 1991).
10. A. N. Schechter, *Nature* **401**, 424 (1999).
11. P. Janssen, in *The Psychopharmacologists II*, *Interviews by David Healy*, D. Healy, Ed. (Oxford Univ. Press, New York, 1999), chap. 3.
12. N. A. Holtzman, T. M. Marteau, *N. Engl. J. Med.* **343**, 141 (2000).
13. K. M. Weiss, J. D. Terwilliger, *Nature Genet.* **26**, 151 (2000).
14. R. Bellamy, *Thorax* **53**, 588 (1998).
15. K. J. Rothman, S. Greenland, *Modern Epidemiology* (Williams & Wilkins, Philadelphia, ed. 2, 1998).
16. J. L. Hopper, in *Encyclopedia of Biostatistics*, P. Armitage and T. Colton, Eds. (Wiley, Chichester, 1998).
17. S. Shuster, *Triangle* **26**, 125 (1987).
18. K. A. Dill, *Nature* **400**, 309 (1999).
19. G. L. Peck *et al.*, *N. Engl. J. Med.* **300**, 329 (1979).
20. B. J. Marshall, J. R. Warren, *Lancet* **1**, 1311 (1984).
21. S. Shuster, *Br. J. Dermatol.* **111**, 235 (1984).
22. H. van Weelden, H. B. De la Faille, E. Young, J. C. van der Leun, *Br. J. Dermatol.* **119**, 11 (1988).
23. J. A. Parrish, T. B. Fitzpatrick, L. Tanenbaum, M. A. Pathak, *N. Engl. J. Med.* **291**, 1207 (1974).
24. J. L. Rees, *Clin. Med.* **1**(5), 393, 2001.
25. P. W. Anderson, *Science* **177**, 393 (1972).
26. E. L. Altschuler, R. Charlton, *Lancet* **356**, 1360 (2000).
27. F. J. Dyson, *Science* **280**, 1014 (1998).

28. I thank S. Shuster, B. Charlton, D. Horrobin, and N. Hastie for many discussions on this topic; and the Wellcome Trust for support.

VIEWPOINT

Maneuvering in the Complex Path from Genotype to Phenotype

Richard Strohman

Human disease phenotypes are controlled not only by genes but by lawful self-organizing networks that display system-wide dynamics. These networks range from metabolic pathways to signaling pathways that regulate hormone action. When perturbed, networks alter their output of matter and energy which, depending on the environmental context, can produce either a pathological or a normal phenotype. Study of the dynamics of these networks by approaches such as metabolic control analysis may provide new insights into the pathogenesis and treatment of complex diseases.

Cell and molecular biology, in conjunction with new theoretical developments, have, in the past decade, taken us from a grossly naïve view of genetic determinism (that complex traits are caused by a single gene) to the stark reality that almost all human diseases are complex context-dependent entities to which our genes make a necessary, but only partial, contribution (*1*). Molecular biologists have rediscovered the profound complexity of the genotype-phenotype relationship, but are unable to explain it: Something is missing. The missing element was described 35 years ago by Michael Polanyi, who characterized live mechanisms and information in DNA as "boundary conditions with a sequence of boundaries above them" (*2*).

Biologists today who work on systems biology refer to these boundary conditions as levels of constraints, or control constraints, outlined in Table 1. Molecular biology has shown that in the progression from genotype to phenotype, many levels of control are introduced. Each control level is defined by a dynamic system of self-organizing proteins, the output of which is governed by laws that are still poorly understood. Polanyi illustrated his concept of levels of control with a metaphor from the game of chess: "The strategy of the player imposes boundaries on the several moves which follow the laws of chess, but our interest [in experimental biology] lies in the boundaries, that is, in the strategy, not in the several moves as exemplifications of the laws." Molecular biology, in identifying control levels, has focused on the "moves" of genes and proteins but has largely ignored the strategy used by dynamic protein networks that generate phenotype from genotype. Systems biology is all about finding the strategy used by cells and at higher levels of organization (tissue, organ, and whole organism) to produce orderly adaptive behavior in the face of widely varying genetic and environmental conditions (*3*). At the center of this effort is a need to understand the formal relationship between genes and proteins as agents, and the dynamics of the complex systems of which they are composed. Much effort has been spent in attempts to predict phenotype, first from genomic, and then from proteomic, databases. But these databases do not contain sufficient information to specify the behavior of a complex system. The "systems" relationship between genotype and phenotype is perhaps best represented in the formulation by Howard Pattee (*4*).

Dynamics describes laws (operating rules) controlling the behavior (the phenotype) of any self-organizing system of gene-encoded proteins. Therefore, we expect that the various transitions shown in Table 1 will involve laws governing an orderly interaction between proteins; between proteins and environmental signals; and, in the case of DNA binding proteins (level 1, Table 1), between proteins and critical small molecules such as nicotinamide adenine dinucleotide (NAD)/NADH, where molecular concentrations are symbols of the entire bioenergetic state of the cell. DNA binding proteins that also sense levels of NAD/NADH are able to transmit that "sense" of energy readiness of a cell to information that changes the pattern of gene expression and, therefore, changes the energy-dependent cellular phenotype (*5*–*7*). The systems controlling transitions from transcriptome to proteome (level 2) and from proteome to complex systems (level 3) are presently foci of intense research activity, but we are still mostly ignorant of the laws governing the context dependency and integration of environmental signals into the output patterns of those systems.

In contrast, at the level of metabolic networks, it is clear that the phenotype (the output of energy and matter) is predictable from known laws of chemistry: laws of kinetics and thermodynamics (*8*). Metabolic

Department of Molecular and Cell Biology, 229 Stanley Hall, No. 3206, University of California at Berkeley, Berkeley, CA 94720–3206, USA. E-mail: strohman@uclink4.berkeley.edu

networks such as glycolysis and the mitochondrial tricarboxylic acid (TCA) cycle and electron transport system are common to all cells. Metabolic control analysis (MCA), discussed below, may reveal, on a cell- and tissue-specific level, changes in redox potential and in key redox-sensing proteins that in turn are related to changes in gene expression and therefore to disease phenotype (Table 1).

Metabolic Control Analysis and Complex Disease

Metabolic systems are identical in all human cells. They are responsible for the conversion of matter into the energy of adenosine triphosphate (ATP) and therefore into the many work functions necessary for health and for life. When these systems are impaired by defective genes or proteins or by environmental conditions, they fail to supply energy at the level demanded by workloads, and the organism as a whole fails as a robust healthy entity.

These systems, as one might expect, are also equipped with a variety of redundant mechanisms: alternative pathways through which normal levels of "bioenergetic potential" may be restored and maintained even in the presence of genetic or environmental insult. Metabolic control theory has established two fundamental concepts of metabolic systems regulation: distributed control and supramolecular organization of the many enzymes that constitute a given metabolic process. Together they state that the control of the overall output of a metabolic pathway is distributed among all the enzymes in that pathway, that any one enzyme or several enzymes may become rate limiting depending on local conditions, and that the interconnectedness between the enzymes is such that it is not possible to change the activity of one without affecting the entire system. These two concepts have been confirmed for the pathways of glycolysis and the TCA cycle of mitochondria (*9*, *10*). Context-dependent regulation of multienzyme systems creates a formidable problem: How does one go about analyzing such a complex system with so many interactions and variables? MCA provides a theoretical basis for measuring the sensitivity of enzymes to many variables (such as substrate concentration, ions, etc.). Using these quantitative measurements from human tissue and standard equations of kinetics and thermodynamics, it is possible to identify not only the key control points but also the effects of altered genes and environments on the overall bioenergetic processes of cells in normal or diseased tissues. In short, the phenotype—the flux of matter and energy through the system—is predictable from quantitative measurements fitted to kinetic and thermodynamic equations; no additional modeling or computational strategies are required.

An important point worth emphasizing is that metabolism presents itself as a universal and predictable process underlying all phenotypes. MCA would appear, therefore, to be an essential aspect of clinical analysis of human diseases. This predictability of a complex biological system stands in contrast to other systems and levels of regulation described recently in this journal (*3*), in which fundamental laws and their equations are not known and extensive computational and modeling strategies must be attempted.

MCA principles have been confirmed and applied, directly or indirectly, to living systems, including cellular models of two neurodegenerative diseases: Alzheimer's disease (AD) and Parkinson's disease. The sought-after research end point is an understanding of these two diseases in terms of their metabolic control and possible metabolic therapy. It is surprising, perhaps, that this research path into neurodegenerative diseases begins with observations on such disparate topics as starving humans, motile sperm, and failing hearts with and without insulin (*11*). Common to all these is the fact that normal levels of ketone bodies, beta hydroxy butyrate (β-OHB), and acetoacetate (both normally occurring metabolites) restore or enhance bioenergetic function in (i) starving humans experiencing abnormally low levels of glucose and insulin (*12*); (ii) bull sperm experiments in which, when β-OHB is the added metabolite, sperm motility increases at the same time that overall metabolic efficiency is increased (*13*); and (iii) congestive heart failure, in which perfused rat hearts respond to ketone bodies added to the perfusion chamber with increased cardiac efficiency (30%) made possible by ketone-induced lowered oxygen consumption (*9*, *10*). These effects were all attributed to a common mechanism as follows: Ketone bodies enter the energy-generating pathways of the TCA cycle and the mitochondria through a pathway that bypasses the glucose entry through the major pyruvate dehydrogenase (PDH) multienzyme complex. In addition, it was shown that the ketone body pathway is thermodynamically about 25% more efficient than glucose alone (*9*–*11*). Clearly, ketone bodies, when applied at physiological concentrations known to be well tolerated by humans (no acid ketosis) (*12*), are an alternative, bioenergetically more efficient fuel (as compared with glucose) for a variety of cells and tissues. In fact, they are being studied for use in the treatment of wounds and other physical trauma (*14*). It is surprising to learn, therefore, that this entire line of research is much neglected by mainstream biotechnology, which emphasizes gene- and protein-based research in the identification of likely targets for which drugs may be developed; an emphasis now being questioned by many scientists and executives in a troubled pharmaceutical industry (*15*, *16*). A summary of all the MCA work described above, carried out by Richard Veech's laboratory at the National Institutes of Health, was published in 2001 (*11*).

The possibility of therapeutic applications of ketone bodies has now been extended to treatment of neurodegenerative diseases. Both AD and Parkinson's disease are characterized by neuronal death in brain areas where there are deposits of amyloid or alpha synuclein peptides. Amyloid peptides are thought to be neurotoxic, although the mechanism underlying their toxicity is not well understood. Given these facts, it is interesting to consider one proposed mechanism for amyloid toxicity that is fundamentally metabolic and involves amyloid peptides as inhibitors of the same PDH multienzyme complex discussed above.

Evidence supporting a metabolic deficiency basis for amyloid toxicity in AD was

Table 1. Levels of regulation—loci of control constraints—above the genome.

Levels and transitions	Dynamic regulatory system
1. Genome to transcriptome	Epigenetic regulation of gene expression (*5*). Includes pathways that detect energy levels (redox levels) and repress DNA transcription when cellular NADH levels are increased (*6*, *7*).
2. Transcriptome to proteome	Regulatory constraints include posttranslational modification of proteins.
3. Proteome to dynamic system	Metabolic networks of glycolysis and mitochondrial oxidation-reduction are the dynamic systems presently the best understood in terms of both mechanism of formation and operating principles (*8*). They display control distributed over all enzymes of a network, and their phenotype includes cellular redox potential.
4. Dynamic systems to phenotype	Control of global phenotype such as disease may be localized to a single regulatory system (such as metabolic, hormone signaling, etc.) or be distributed over many systems and levels.

first presented by Hochi *et al.* in 1996 (*17*, *18*), who showed that a fragment of the beta chain of amyloid, Aβ 1-42, stimulates glycogen kinase activity, leading to phosphorylation and thus to the inhibition of the pyruvate dehydrogenase multienzyme complex in cultures of hippocampal neurons. This blockade results in lowered ATP production and in downstream effects, including the inhibition of acetylcholine synthesis, as shown in studies of cultured septal neurons (*18*). In turn, this inhibition proceeds from a decreased intracellular citrate concentration, a source of acetylcholine, caused by the inhibition of PDH by the Aβ 1-42 fragment.

Following up on this earlier work, Veech's group later showed that cultured hippocampal cells die when exposed to the Aβ 1-42 fragment and that addition of the ketone body β-OHB protects these neurons from Aβ 1-42–induced cell death (*19*). The likely explanation for the ketone rescue of amyloid-poisoned cells is that ketone metabolism bypasses the Aβ 1-42 blockade of PDH, restores the normal metabolic supply of energy, and also restores citrate concentrations necessary for acetylcholine synthesis essential to brain cell activity. The authors also reported that primary cultures of mesencephalic dopaminergic neurons die when exposed to the drug MMP+, which causes symptoms similar to Parkinson's disease, and that neuronal death can be prevented by addition of β-OHB.

Based on these results, one is led to consider a simple metabolic deficiency model for AD: amyloid 1-42 blockade of the PDH multienzyme complex is the source of amyloid toxicity and leads to loss of bioenergetic potential and neurotransmitter production, which ultimately leads to neuronal death. Ketone bodies, because they enter the TCA cycle at a point below pyruvate, effectively bypass the amyloid blockade of PDH and so restore normal levels of acetyl CoA, thereby restoring normal redox potential and acetylcholine levels. In contrast to more conventional hypotheses, which derive from the standard practice of identifying a variety of cascading molecular events as causal mechanisms and as targets for drug design, the metabolic hypothesis treats these events as epiphenomenona of a primary toxicity of metabolic origin. The changes in redox couples (NAD/NADH) induced by ketone bodies also provide a thermodynamic explanation for the ability of ketones to ameliorate the free radical damage thought to play a role in the pathogenesis of Parkinson's disease. Could it be that simple?

It is hard to be sure, because the MCA work summarized here has never been duplicated at the level of detail carried out by the biochemists involved, in part because it is simply too labor intensive. Why? It took the Veech group 4 years to complete their MCA on rat heart preparations; they did all the measurements "by hand," because no technology exists to do for MCA what DNA amplification (polymerase chain reaction) and sequencing have done for molecular genetics. Without a technology able to automate the gathering of massive amounts of detailed information, this powerful approach to understanding complex diseases must languish. For example, technology is needed that would automate data collection from small samples (human biopsies and/or cell cultures) taken from normal and diseased individuals, which could be manipulated while quantitative measurements of all enzymes and their substrates were performed and as genetic and environmental conditions were changed.

Lack of technology is not the only force impeding a transition to a biology that includes dynamic systems. Other factors include departmentalized university systems unfriendly to interdisciplinary studies, and economic pressures on universities to establish relations with corporate biotechnology groups already committed to large investments in agent-based diagnostics and therapy. Also counterproductive are outmoded research funding patterns that continue to see complex phenotypes as primarily derivable from genomic and proteomic databases. In the end, the most serious impediments may be insufficient training of biomedical scientists and clinical investigators in mathematics, statistics, kinetics, and thermodynamics. Without these quantitative backgrounds, scientists reviewing grant applications based on systems theory and quantitative biochemistry will, predictably, fail to appreciate their value.

Concluding Thoughts

Molecular biology still commands the lion's share of attention and funding, which will not, and should not, be easily surrendered by the scientists involved. After all, whereas systems biology, including MCA, makes a strong case for itself and for a shift in scientific emphasis, it too may ultimately fail to close the gap between genotype and phenotype. Nevertheless, it does seem to this observer that a crisis is developing in medical biotechnology as it moves forward with continued emphasis on a scientific paradigm that largely omits a dynamic systems component (*20*).

This brings us to Thomas Kuhn, the late philosopher and historian of science, and the lessons to be learned from the history of science when conflicts arise concerning new directions and policy decisions crucial to making a choice of research direction (whether to fund particular research). At the present juncture, requests for new funding patterns will be seen by the dominant group as a deflection from their hard-won struggle to bring "the century of the gene" (*21*, *22*) to its completion in the form of computational biology. To those who see genomics and proteomics as necessary but insufficient, their attempts to bring systems biology into prominence might well be seen in terms of Kuhn's observation concerning similar junctures of crises and revolutions in the history of the physical sciences: "Confronted with anomaly or with crisis, scientists take a different attitude toward existing paradigms, and the nature of their research changes accordingly. The proliferation of competing articulations, the willingness to try anything, the expression of explicit discontent, the awareness of philosophy and to debate over fundamentals, all these are symptoms of a transition from normal to extraordinary research. It is upon their existence more than upon that of revolutions that the notion of normal science depends" (*23*).

To the extent that the perceived crisis in modern biology signals a transition from normal to extraordinary science, perhaps that perception will propel a revolutionary change in our approach to understanding the nature of the genotype-phenotype relationship; one in which genes and dynamic systems would have equal standing. Will it happen soon? One thing is certain: The development of technology that is able to provide quantitative answers to dynamic systems questions will make all the difference. As Ludwig Wittgenstein put it: "Where there are no answers there are no questions."

References and Notes

1. D. J. Weatherall, *Nature Rev, Genet.* **2**, 245 (2001).
2. M. Polanyi, *Science* **160**, 1308 (1968).
3. Special section on Systems Biology, *Science* **295**, 1661 (2002).
4. H. Pattee, in *Causation, Control and the Evolution of Complexity*, in *Downward Causation*, P. B. Anderson. C. Emmeche, N. O. Finnemann, P. V. Christiansen, Eds. (Aahrus Univ. Press, Langelandsgade, Denmark, 2000), pp. 63–78.
5. Special section on Epigenetics, *Science* **293**, 1063 (2001).
6. J. Rutter, M. Reick, L. C. Wu, S. L. McKnight, *Science* **293**, 510 (2001).
7. Q. Zhang, D. W. Piston, R. H. Goodman, *Science* **295**, 1895 (2002).
8. D. Fell, *Understanding the Control of Metabolism* (Portland Press, London, 1997).
9. Y. Kashiwaya *et al.*, *J. Biol. Chem.* **269**, 25502 (1994).
10. K. Sato *et al.*, *FASEB J.* **9**, 651 (1995).
11. R. L. Veech, B. Chance, Y. Kashiwaya, H. A. Lardy, G. Cahill Jr., *Int. Union Biochem. Mol. Biol. Life* **51**, 241 (2001).
12. G. F. Cahill, *N. Engl. J. Med.* **282**, 668 (1970).
13. H. A. Lardy, R. G. Hanson, P. H. Phillips, *Arch. Biochem.* **6**, 41 (1945).
14. R. Veech, personal communication.
15. K. Garber, *Nature Biotechnol.* **20**, 207 (2002).
16. D. Horrobin, *Nature Biotechnol.* **19**, 1099 (2002).
17. M. Hoshi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2719 (1996).
18. M. Hoshi *et al.*, *J. Biol. Chem.* **272**, 2038 (1997).
19. Y. Kashiwaya, T. Takeshima, N. Mori, K. Clarke, R. Veech, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5440 (2000).
20. R. Strohman, *Nature Biotechnol.* **15**, 194 (1997).
21. E. Fox-Keller, *The Century of the Gene* (Harvard Univ. Press, Cambridge, MA, 2001).
22. R. Lewontin, *The Triple Helix* (Harvard Univ. Press, Cambridge, MA, 2000).
23. T. Kuhn, *The Structure of Scientific Revolutions* (Univ. of Chicago Press, Chicago, IL, ed. 3, 1996), chap. 8.