nature medicine
VOLUME 7 • NUMBER 7 • JULY 2001

Much ado about data

We are on the verge of a new period of biomedical research. No longer is the bewildering complexity of biology impenetrable. A few biomedical researchers are beginning to reassess health and disease in the context of truly complex systems. The rest of the biomedical research community must catch up.

Although Hippocrates is credited with freeing medicine from the direct clutches of religion and imbuing an ethic that has largely stood the test of time, he, his fellow ancient Greeks and later the Romans based their biomedical thinking on the theory of 'humors', knew precious little about anatomy and perhaps did more harm than good for individual patients. With the Renaissance came progress in anatomy, epitomized by the breathtaking drawings of first DaVinci and later Vesalius. Anatomy begat physiology and medicine could for the first time claim an objective and scientific basis for subsequent theories. Nonetheless, biomedical investigation remained largely observational and it was not until the second half of the 20th century that tools to tackle mechanism were developed.

Once nucleic acids were established as the hereditary material, and the wherewithal to isolate, chop and recombine them was established, biologists started describing and experimenting with the mechanisms that control health and disease. But for 50 years this experimentation has been characterized by deconstruction and reductionism. In genetics, for example, the central dogma (genes are the heredity unit; a single gene codes for a single protein) became the pervasive paradigm of the field, and individual genes were located, cloned, mutated, knocked in and knocked out, one at a time. Endless components in artificially isolated cellular pathways were described and poked and prodded to unearth their influence on direct neighbors in the pathway. Whether we were studying cancer or neurodegeneration, our approach to understanding disease (and therefore to designing therapeutic interventions) was always to reduce the process to a series of manageable components and to investigate the influence and parameters of each component.

Our half-century fascination with reductionism has served us well. But quite suddenly, biomedical research has become a science of complex systems. This movement has not sprung from a new appreciation of the complexity of biology—that of course has long been understood. What is new is a technical adroitness that enables us to generate and manage huge data sets. This high-throughput science brings two obvious opportunities. First, we can continue to practice reductionist research as we have done in the past, but at a much greater rate. Far more exciting, however, is the possibility that by generating large and complex data sets we will gradually develop the tools and eventually the understanding that will allow us to exploit these data to the fullest.

For many in the biomedical research community, and particularly those involved in the more medically applied sciences, this exploration of unknown consequences calls for a leap of faith. Contemporary biomedical research has been under pressure to deliver on specific goals and often has sought only those data pertinent to those goals. For many researchers it is still counter-intuitive to expect that the tools and understanding will develop only after the data are assembled. (The lesson of genomics should persuade them otherwise—ten years ago, who would have guessed that tumor profiling using thousands of genes would be commonplace today?)

Exciting as this shift to complex biology is, to take full advantage of these remarkable advances, the biomedical research community must go one step further and fully embrace interdisciplinary research. And this collaboration must go beyond the well-rehearsed pattern of, for example, statisticians helping numerically challenged biologists. New research groups must be assembled and given radically new laboratory environments in which to work.

Few have been as quick to recognize the potential of this approach as Andrew Murray, the newly appointed Director of the Harvard Center for Genomics Research. Recognizing that the real world of biology is a complex environment of buffers, amplifiers, redundancy, flexibility and variation, and that we need to develop a new set of tools to manipulate and explore biological systems, Murray and a few like-minded visionaries are creating a new research paradigm, building an institute that will physically and intellectually place biologists, physicists, chemists, mathematicians, computer scientists, and engineers together. It is Murray's hope that collectively they will "develop experimental and theoretical approaches to understanding biological questions" and help identify the "general principles that will give insight into the structure, behavior, and evolution of cells and organisms."

To complete the puzzle we now need adventurous funding agencies to recognize the potential of this brave new world of complexity. Although narrowly defined goals of this new genre of research may still be elusive compared with the hypothesis-driven culture of today's biomedical research, funding agencies must take their own leap of faith by supporting and encouraging the pioneers that will create the infrastructure and physical environments that will allow a new generation of biomedical researchers to explore complexity.