

What Does it All Mean to You?

Floyd E. Bloom

Department of Neuropharmacology, The Scripps Research Institute, La Jolla, California 92037

Nearly 20 years ago, I introduced a special series of articles for *Trends in Neurosciences* (5:295, 1982) that sought to illustrate some of the then-latest developments in molecular neuroscience. That was the time when neuroscientists were beginning to borrow heavily from the nascent field of molecular genetics, just as we had previously borrowed flagrantly from the previously powerful fields of biophysics and cell biology. The tools of that day seemed quite novel and powerful: the ability to make cDNA libraries carrying modest-sized inserts of mammalian mRNAs, the ability to isolate and amplify these cloned segments of active genes, and the ability to make what seemed to be rapid determinations of the sequences of those mRNAs and deduce the likely amino acid sequences of their gene products. Even more exciting then was the prospect, realized in less than 2 years, of detecting gene mutations that cause inheritable neurological diseases such as Huntington's.

We fast-forward our retrospectroscope to the present and recognize what primitive times we were then entering, to which the intervening progress has added PCR to amplify genetic information and to do so quantitatively, the ability to identify disease-associated or disease-causing genetic mutations and to develop transgenic animals that express those genes, the ability to knock out the expression of their natural counterparts, the ability to apply a variety of strategies to turn up or down the expression of specific genes in specific neurons at specific times after birth, and the ability to determine what unknown proteins interact with proteins just discovered to mediate, for example the process by which synaptic vesicles store their transmitter, discharge it under Ca-dependent, voltage-dependent conditions, and then reconstitute themselves for reloading with fresh transmitter molecules. Implicit in all these advances is the automation of data collection and analysis and the ability to enlist the computer as a user-friendly wizard to search the data and deduce useful patterns within them. And as a vaudevillian performer might have been tempted to announce, "We ain't seen nothing yet!"

The single largest bolus of molecular information ever infused into the research community was ushered in with a publicly proclaimed political event in June of 2000 and finally presented in published form this past February: namely, the initial "draft" inventory of the human genome. What this phrase means in simple terms is that by using advanced versions of the powerful methods of molecular biology, several large scientific teams have been able to take apart human DNA in very refined ways, amplify the amounts of the pieces, determine the order of the nucleic acid bases in each fragment, and then put all the fragments back

together again in proper order across the 23 pairs of human chromosomes.

However, having determined the sequences of the nucleic acids, it was possible to train computers to read the sequence information and spot the specific signals that identify the beginnings and endings of sequences likely to encode proteins and hence the genes for those proteins. In addition, the computer systems could then sort among those proteins based on similar sequences (motifs) of their amino acid building blocks and assign them to families of similar proteins whose functions had already been established. In this way, scientists were able to determine approximately how many proteins the human genome (all of the genes a human has) could encode. Similar routines allowed the determination of how many of those newly recognized genes were similar to genes we have already recognized in the smaller genomes of other organisms previously mapped out [for example, the yeast, the worm (*Caenorhabditis elegans*), and the fruitfly (*Drosophila melanogaster*)] or of organisms about to be mapped out (*Mus musculus*) and how many other genes had never been encountered previously.

Scientifically, this state of information has been termed a draft because it is based on a very dense but not yet complete sample of the whole genome. What has been determined still contains a very large number of interruptions and gaps. Nevertheless, compared with completed genomes of other organisms, the human has greatly increased its representation of genes related to nervous system function, which had already been thought to contain at least half of the genome's treasure. Importantly for diseases of the nervous system characterized by premature death of neurons, such as Alzheimer's and Parkinson's diseases, there is a very much larger genetic representation than had been expected of programmed cell death-related genes. Comparing the consensus sequence of human DNA with other samples entered previously in publicly accessible databases, it has been estimated that there may be >2 million sites at which significant variations between individuals (so-called "polymorphisms") exist that may (or may not) alter the functions of the encoded proteins and render an individual more or less vulnerable to specific disorders.

Two major future vistas can then be inferred. To create organisms as complex as humans with a relatively small number of genes (at least compared with worms and fruit flies) probably means that the rich diversity of the proteins necessary for human structure and function is based on their modifications either during transcription of the gene or after translation of the intermediate mRNA into the protein. Second, as far as this observer can determine, it confirms that although compiling this draft inventory represents a stunning technical achievement, for many neuroscientists the next steps in the harvesting of this information for influencing their own specific experimental objectives may be quite unclear. From my perspective, determining where in the

Correspondence should be addressed to Floyd E. Bloom, Department of Neuropharmacology, The Scripps Research Institute, La Jolla, CA 92037. E-mail: fbloom@scripps.edu.

brain's circuits specific genes are normally expressed, and how that expression pattern may be altered by the demands of illness or an unfriendly environment, represents an enormously daunting task. That task, at present, is one for which there are as yet no tools equivalently as powerful as those used to acquire the enormous amount of nucleotide sequence data that we have now amassed.

The stage of understanding at which we are now has been referred to as the end of "naive reductionism." To fulfill the promise of the enormously rich mother lode of genetic information already in hand, we must next determine where these genes are expressed, what functions they control, and what sorts of controls other gene products can exert over them. In the nervous system, where cell-to-cell interaction is the main operating system in relating molecular events to functional behavioral events, the still murky properties of activity-dependent gene expression will certainly require enormous investment by the neuroscience community. This special section of the *Journal of Neuroscience* was commissioned to help make more apparent the opportunities that lie in the new technology, which continues to evolve its powers almost as fast as applications of the technology to specific experimental problems have begun to render early solutions. The authors of these brief reviews were asked not only to communicate the excitement of the strategies they have themselves helped evolve but also to extol their shortcomings as well as their special virtues and to illustrate their descriptions with problems being solved by them.

Although each brief review is concise enough and accessible to every reader of this journal, it may be useful to indicate how they relate to each other to depict some selected views of the molecular neurosciences in late-2001.

Sutcliffe reviews the several current strategies of gene detection and presumptive functional characterization, from differential display technologies to DNA chip array technologies to highly

automated rapid open-ended detection of genes whose mRNAs differ in selected experimental perturbations.

Eberwine et al. review just how much can be done in characterizing the differences in gene expression within single cells *in vitro*, what might be expected from *in vivo* extensions, and how these living molecular experiments can provide quantitative insights.

Grant and Blackstock describe how similar high-throughput, highly specific mass spectrometry combined with computer-intensified molecular database mining can be used to define the complexes of proteins (proteomics) that work as molecular ensembles in, for example, the signal transduction pathways of neurotransmitter receptors. The final two articles demonstrate how these approaches have begun to be applied to human and animal models of human diseases.

Blakely reviews the molecular foundations of two neurotransmitter transporter systems, that for norepinephrine and that for serotonin, long associated with depression and the actions of anti-depressant drugs, and suggests how the field may in fact advance more rapidly by analysis of rare functional polymorphisms first recognized from dysfunctions of these transporters in the peripheral nervous system.

Last, Nestler demonstrates how the genomic and proteomic approaches can be applied to the science and the societal consequences of substance abuse to define vulnerabilities, resiliences, and the basis by which the brain adapts to prolonged drug exposures.

As the editorial organizer for these brief reviews, it is my hope that many more neuroscientists will find these methods and, more importantly, the quality of the floods of data they produce, conducive to new and more powerful collaborations. However, if the past is any indication, today's tools and methods of data analysis will likely be surpassed by new surprises soon.