EDITORIAL

Omics and Its 15 Minutes

Mark W. Duncan
Associate Editor

It has been over a decade since the introduction of the word *proteomics* (1), and some suggest that the last 10 years has seen us fall way short of delivering on its potential. Others claim that too much was promised and expected in the first place. It is therefore timely to examine what has been delivered and perhaps to redefine the promise of proteomics and related omic strategies. It is also a good opportunity to assess the important elements of omics submissions that might be published in journals such as *Experimental Biology and Medicine*.

Most life scientists operate in an environment where data are hard to come by. They have been trained to construct focused, falsifiable hypotheses, to test one or a couple of variables at a time, and in this manner to incrementally advance our knowledge of a specific issue. By contrast, omic techniques offer the potential to interrogate thousands of independent variables in a single study, and thereby promise expedited approaches to advancing biomarker discovery, identifying novel drug targets, and understanding the underlying mechanisms of health and disease. But are we applying these methods appropriately, or are we simply enamored with these tools to the point where we have lost sight of the biologic imperative underlying the work? Have we equated the acquisition of massive data sets with the practice of science believing that biomarkers, therapeutic targets, and mechanistic insights would flow from these studies like water from a fountain?

Our scientific colleagues in other disciplines would attest to the fact that observation and measurement are major components of the scientific process but that data acquisition does not equate to knowledge acquisition. After all, people have been "looking" at thousands of variables for centuries (including biologists, astronomers, and geologists), but precious few have been blessed with the insight to advance novel, falsifiable hypotheses, test them, and then repeat and refine the process.

To my mind, the most critical issues as they relate to the application of the omics methods in experimental biology and medicine can be summarized in a few key points.

1. The Biology not the Methods: The omics technologies are simply problem-solving tools. Most are blunt, but they can be powerful in a specific and well-defined context: They are not all things in all situations. While potentially powerful, the tools themselves have gained disproportionate attention and are frequently the focus of method-oriented rather than problem-oriented publications. While the review of these methods can represent a valuable contribution to the literature, there has been more discussion of potential than objective reports of their application to experimental biology and clinical research. For the biologist, and most of the readership of *EBM*, these tools are only as useful as the problems they can solve. Less infatuation with the technology and a greater focus on practical problem solving is therefore appropriate, especially in submissions to *EBM*.
2. Data Quality, not Quantity: Generating enormous data sets is impressive, but quantity is no substitute for quality. If results are of dubious precision and accuracy, if we fail to replicate our studies, or if we do not frame our findings in a meaningful biologic context, the exercise may be worthless.
3. Hypotheses and Alternative Interpretations: Science can begin with vaguely stated, multicomponent hypotheses. Typical omic studies do just this: They address thousands of separate, parallel hypotheses in an untargeted manner. The experimental design aims to "see" whether differences in metabolite, message, or expression levels distinguish between control and test samples (or populations). Studies of this nature are often described as discovery, but in the context of science they are better defined as observational investigations that can aid in the generation of novel hypotheses. Unfortunately, we have been too keen to promote these applications of our blunt tools as a single-step strategy to science: i.e., discovery, testing, and proof. While the data may flow quickly

1535-3702/07/2324-0471$15.00

> in these instances, the short history of omics demonstrates that we are slow to develop falsifiable hypotheses and thoroughly interrogate them. Consequently, the potential of these studies is rarely realized, and follow-up is uncommon. Equally, we often fail to acknowledge the limitations in the data themselves (e.g., the uncertainties in the identifications made and the imprecision in measures of quantification) and to offer alternative interpretations of our data. By failing to take these critical steps we inadvertently misdirect future scientific endeavors and help construct a house of cards.

Perhaps at different points in time and in response to different circumstances, scientists should revisit some central issues. In 1964 John Platt wrote of "strong inference" (2): i.e., (i) devising alternative hypotheses; (ii) devising a crucial experiment (or several of them), with alternative possible outcomes, each of which will, as nearly as possible, exclude one or more of the hypotheses; (iii) carrying out the experiment so as to get a clean result; (iv) recycling the procedure, making subhypotheses or sequential hypotheses to refine the possibilities that remain; and so on. Building on the works of Bacon, Popper, and others Platt defines these steps as critical components in the scientific endeavor.

We need to remember that the omics methods do not profoundly change the scientific process. The great value of these data, just like the images from an interplanetary probe such as Hubble, is that they offer a powerful alternative first step to hypothesis generation because they are independent of existing knowledge and less dependent on insight, instinct, and experience. A single omics study may present us with data from which we can formulate dozens of testable hypotheses and, as T. C. Chamberlin suggested (3), when we have multiple lines of independent investigation, it limits our potential to embrace a single hypothesis with too much affection, to press our theory to fit the facts and to press the facts to make them fit our theory. Multiple independent hypotheses therefore help us to be more objective, to distribute our effort and divide our affections. Each hypothesis then suggests " ... its own criteria, its own means of proof, its own method of developing the truth, and if a group of hypotheses encompass the subject on all sides, the total outcome of means and of methods is full and rich" (3). But, as Platt suggests, each hypothesis needs to be investigated: no corners can be cut; no alternative explanations of the findings overlooked.

The omic methods will be powerful aids to understanding biology and advancing clinical practice, but it was never going to be as easy as some led us to believe. The early promise we offered our colleagues, collaborators, and granting agencies was false—a few months' work and a couple dozen retrospective collected samples were never going to deliver a handful of markers ready for routine clinical application. Highly specific and sensitive markers do not fall in your lap; rigorous testing is required. Scientific investigations over the centuries prove that: (i) there is no substitute for formulating, testing and retesting falsifiable hypotheses; (ii) we must eliminate the potential for bias and chance to confound our studies; and (iii) we must have an intimate understanding of the quality of our data and the sources of errors associated with it.

In summary, as we generate our data and prepare it for publication, we should not forget the principles that underpin science and the tried and tested approaches to advancing knowledge. We need to ask good questions; design and conduct good studies; define the limits of our methods; interpret our data in an honest and objective manner; and, wherever possible, we should offer alternative explanations for our observations. No study will be perfect. We will rarely deliver the bottom line (e.g., validated biomarkers), but elegant and well-conceived studies should be our goal; massive, overly ambitious studies are more likely to confound rather than advance our knowledge.

1. Wasinger VC, Cordwell SJ, Cerpa-Poljak A, Yan JX, Gooley AA, Wilkins MR, Duncan MW, Harris R, Williams KL, Humphery-Smith I. Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. Electrophoresis 16:1090–1094, 1995.
2. Platt JR. Strong inference. Science 146:347–353, 1964.
3. Chamberlin TC. The method of multiple working hypotheses. J Geol 5: 837–848, 1897.