

# Why we need more basic biology research, not less

David Botstein

Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544

**ABSTRACT** Much of the spectacular progress in biomedical science over the last half-century is the direct consequence of the work of thousands of basic scientists whose primary goal was understanding of the fundamental working of living things. Despite this, many politicians, funders, and even scientists have come to believe that the pace of successful applications to medical diagnosis and therapy is limited by our willingness to focus directly on human health, rather than a continuing deficit of understanding. By this theory, curiosity-driven research, aimed at understanding, is no longer important or even useful. What is advocated instead is “translational” research aimed directly at treating disease. I believe this idea to be deeply mistaken. Recent history suggests instead that what we have learned in the last 50 years is only the beginning. The way forward is to invest more in basic science, not less.

## Monitoring Editor

David G. Drubin  
University of California,  
Berkeley

Received: Aug 31, 2012

Accepted: Sep 5, 2012

## INTRODUCTION

The 50 years between the discovery of the DNA structure and the determination of the sequence of the human genome produced spectacular advances in knowledge and techniques that have thoroughly altered our understanding of basic biology and diseases of every kind. Biological science been transformed. DNA sequencing and DNA synthesis are routine techniques, often outsourced to commercial services. Virtually all laboratory studies today involve genetic manipulation, which has become standard for everything from protein purification (done from cloned and/or tagged genes) to determination of phenotype (done by construction mutants to specification). Naturally, this progress has raised expectations about applications to medical practice, not a few of which have been realized. One prominent example is the development of effective treatments for AIDS, a disease whose essential nature, apart from the viral etiology, could not have begun to be understood in the 1950s. Indeed, all of the products of the biotechnology industry could not have been imagined then. According to the Centers for Disease Control, the average life

expectancy in the United States has increased from 69.7 years in 1960 to 78.7 years in 2010, or 13%.



David Botstein

My thesis is that the progress outlined above is the direct consequence of the work of thousands of basic scientists whose primary goal was understanding the fundamental working of living things. Few, if any, of the scientists who discovered how to isolate genes, manipulate them in model organisms, and sequence DNA had any specific disease in mind, even though they knew that eventually knowledge and understanding would make applications to human health possible. I know this, because, until 1987, I was one of those scientists.

Regrettably, the recent, rapid increases in data volume (especially molecular sequences), coupled with the success of the biotechnology industry, has had a malign effect on science policy in the United States. Many politicians, policy makers, and even working scientists have come to believe that the pace of successful applications to diagnosis and therapy is limited by our willingness to focus on human health, rather than a continuing deficit of understanding. By this theory, curiosity-driven research aimed at understanding is no longer important or even useful; what is advocated instead is “translational” research aimed directly at treating disease. I believe this idea to be deeply mistaken. To me, recent history suggests instead that what we have learned in the past 50 years is only the beginning. The way forward is to invest more in basic science, not less.

DOI: 10.1091/mbc.E12-05-0406

Address correspondence to: David Botstein (botstein@princeton.edu).

Abbreviations used: NIH, National Institutes of Health.

© 2012 Botstein. This article is distributed by The American Society for Cell Biology under license from the author(s). Two months after publication it is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

“ASCB®,” “The American Society for Cell Biology®,” and “Molecular Biology of the Cell®” are registered trademarks of The American Society of Cell Biology.

## HOW DID WE GET HERE?

In the 1950s, most of the scientific community had only recently learned of the central role of DNA in inheritance. Nothing was

known of the role of ribosomes in protein synthesis, and the idea that there might be a genetic process that translates a DNA sequence into an amino acid sequence was still a hypothesis in search of evidence. It was not until the mid-1960s that a scientific consensus was reached on the genetic code and the concepts of "open reading frame," frameshift, and mRNA became generally accepted by the scientific community. The 1970s saw the development of DNA sequencing and "recombinant DNA" technology, which allowed the isolation and expression of individual genes from any organism in bacteria. All of this was worked out by basic scientists, who were motivated solely by the desire to understand the flow of information from genes to phenotype.

The success of these methods led to the emergence of the biotechnology industry. Within a decade, gene manipulation (i.e., the ability to make transgenic and mutant organisms at will) became routine in model organisms and soon became practical for mammals. This development, along with the introduction of gene mapping and gene isolation methods for human disease genes, had a transforming effect on thinking about drug discovery and development. Genetically engineered mouse models became indispensable tools not only in assessing gene function but also in drug development. Genomic sequences began to accumulate in the databases at astonishing rates, far outpacing the famously exponential rate of increase in electronic computing power known as Moore's law. Today, everyone recognizes that in a very few years treatment of cancer patients will routinely be guided by sequence analysis of their tumors. Not surprisingly, expectations for impact on medical practice rose dramatically during this period.

One might think the natural interpretation of this history is that investment in basic science aimed simply at understanding is a successful strategy for finding applications to disease. Important applications can be made relatively easily once the basic science is clear. For example, once sequencing became practical and affordable, the path to application of sequence analysis to guiding cancer therapy was obvious. One might also think the natural policy implication of this history is to support basic research strongly going forward in the expectation that more good stuff is on the way, even if we do not quite know what applications will be found. Indeed, it is abundantly clear that the basic understanding of biology offered by the genomic sequences is very far from complete; we have only scratched the surface. Consider the pace of discovery of new kinds of functional RNAs or the number and diversity of epigenetic mechanisms, to cite just two examples of discoveries being made weekly.

Unfortunately, this seemingly natural interpretation and logical policy implication of recent history have not won favor with some politicians, scientific funders, and policy makers. Instead, the new concept of "translational research" gained popularity, especially in the medical establishment. This powerful but deeply mistaken idea confuses knowledge of facts with understanding of principle, process, and mechanism. It is used to justify policies that restrict curiosity-driven basic research aimed at understanding in favor of research aimed at applying the data and concepts already in hand. This idea has now infected the peer-review systems. I fear that the eloquent defense of basic science by Francis Collins, the director of the National Institutes of Health (NIH), in a recent issue of *Science* (Collins, 2012) is unlikely to stem the growing reluctance of study sections and councils to consider work that cannot be directly tied to patient care in the short (one grant cycle) or medium (two grant cycles) term. Many of my colleagues have noted that the scores for "Impact" in the new NIH scoring system for grants are being seen as a proxy for assessments of translational relevance, rather than being used for their original intention, which

explicitly included, and indeed emphasized, assessment of impact on basic understanding.

Worse, the current enthusiasm for translation over discovery and understanding disproportionately affects our younger scientists, most of whom understandably perceive this idea as writing on the wall aimed at them. They sense that it might be prudent to abandon basic science for translational research in order to safeguard their careers. Many of the most basic areas of research, the very ones that provided us with progress, have come under threat.

## HOW MUCH OF WHAT IS "KNOWN" DO WE ACTUALLY UNDERSTAND?

These developments are particularly unfortunate, because, in reality, only a very modest fraction of human genes is securely annotated with any functional information, and most of these annotations are simply transferred from model organisms (yeast, flies, worms, and, in the best case, mice) by sequence-similarity algorithms. Even in the model organisms, the functions of most genes, proteins, and RNAs are understood only sketchily, and less is known about their interactions. The study of the latter is clearly still in its infancy. Is this really a sufficient basis for finding cures for disease? How much do we really understand? If we do not understand the basic text, how will we translate this into cures for disease?

My reaction to the sequencing revolution was one of amazement, not at how much we have learned, but rather at how much we have yet to understand. Every poorly annotated gene is a challenge to our science. Knowing that there are hundreds of transcription factors, protein kinases, microRNAs, etc., of which we have a realistic understanding of only a tiny fraction, means that there is much more left to do than has already been done in the way of discovery of function and mechanism. Do we know how each of these factors, kinases, and RNAs acts in development, let alone in disease? Do we know how they might interact? Can we predict very much about their behavior? Can we understand their diversity? Do we really believe that the way to face this mountain of ignorance is to deemphasize basic research, which is the only proven path to achieving understanding?

I believe that the way forward is to continue to maintain a balance between research aimed at basic understanding and research aimed at more direct application to disease. Both types of research have their place.

Of course, the purpose of the NIH is to address disease. The official NIH website ([www.nih.gov/about/mission.htm](http://www.nih.gov/about/mission.htm)) makes the order of priorities quite clear:

NIH's mission is to seek fundamental knowledge about the nature and behavior of living systems and the application of that knowledge to enhance health, lengthen life, and reduce the burdens of illness and disability.

The first task inevitably will be to understand the roles and activities of as many as possible of the genes and gene products we so recently have discovered and so few of which we really understand. This might well require more basic science effort going forward, not less.

In a way, the "translation" metaphor is helpful. If there is no basic text to translate, how can it be translated to another language? If there is inadequate basic understanding, redoubling efforts to apply what we know to disease will have no good result. We need basic research to help us understand biology and disease. Only when we do that, will we truly succeed. To use another apt metaphor, let us not eat our seed corn—if we hope to have a bigger harvest, we will need more seeds, not fewer.

## REFERENCE

Collins FS (2012). NIH basics. *Science* 337, 503.