Physiological Genomics: Who we are and where we’re going

THREE YEARS AGO, the American Physiological Society created a journal devoted to the advancement of the new and exciting field of physiological genomics. The journal has grown at a steady pace with a strong trajectory especially in the last 6 months. Since Physiological Genomics (“PG”) went online, 132 high-quality papers have been published that cover a broad array of subjects ranging from bioinformatics to proteomics. The journal is now cited by Index Medicus as well as by Medline. Thanks to an entirely online manuscript submission, review, and publishing process the journal can offer speedy review, as well as the latest innovation of immediate online publication upon acceptance.

This journal has a bright future. In this editorial, we reaffirm our goals and identify specific research areas that the journal wishes to emphasize.

Recent developments in genomics have spawned novel technologies in identification, quantification, and comparison of multiple genes. With this profusion of new technology comes a torrent of original research papers in functional genomics. While a number of journals cover the field, Physiological Genomics is distinguished by its strong link to physiology. PG papers will help to define the genetic pathways and protein interactions that mediate physiological responses.

Physiological Genomics will continue to focus on four major fields: research on gene expression and profiling utilizing microarrays; research on model systems such as yeast, Escherichia coli, rats, mice, Drosophila, and zebrafish; mapping of complex traits; and proteomics. Each of these areas brings us closer to understanding the intricacies of physiology in a postgenomics era.

Microarrays

The current workhorse of the genomics revolution is the DNA microarray. Rapid advances in high-throughput technology have brought down the cost of producing and analyzing microarray data. This makes it possible for even small laboratories to compete with larger research groups in analyzing the changes in numerous transcripts simultaneously. Just as cDNA libraries have long been commercially available for a variety of tissues, disease states, and organisms, whole genome arrays are becoming standard and affordable. Microarrays can identify changes in gene regulation over developmental time courses or over the progression of a disease. Microarrays can also display subtle differences between research cohorts under particular treatment regimens. Scientists will be able to correlate changes in specific regulatory pathways that are important in physiological control. Thus the complex interactions of homeostatic mechanisms responsible for physiology may become more transparent as their interconnected nature is revealed.

Microarray analysis is a frequently employed research method in the Journal (accounting for ~25% of original research articles), as befits studies of functional genomics. PG has published a number of papers using microarray data in “expression profiling,” in which transcripts from a particular tissue in a certain disease or treatment state are compared with a normal or wild-type state. Such comparisons may reveal important genetic factors associated with disease onset or progression. Some specific examples of uses of microarray technology include a paper describing the use of support vector machines to discriminate among different cancer tissue samples (7); and a paper describing a DNA array capable of presenting verifiable differences in regulation of members of the cytochrome P-450 family in rat liver upon treatment with a battery of pharmaceuticals (2).

This technology will be adopted by a wide range of biomedical researchers. To facilitate the understanding and use of this technology, PG will publish “toolbox” articles, which will present research methodologies, and meeting reports, such as one describing a Nature Genetics microarray meeting in 2000 (1). The Journal encourages the submission of papers in this field and provides online storage space for the megabytes of supplementary data that such research invariably produces.

Making supplementary data available to readers provides an opportunity for independent confirmation of published results and encourages further scientific exploration.

Model Systems

In bridging the gap between genotype and phenotype, scientists must understand transcriptional and translational regulation, posttranslational modifications, and protein-protein interactions. Many of these basic aspects are best studied in organisms with well-
defined genetics, a host of research tools, and a short generation time. *Physiological Genomics* is committed to publishing research on the genetics of such “model organisms,” including *E. coli*, yeast, *Caenorhabditis elegans*, *Drosophila*, zebrafish, mice, rats, and other vertebrates including larger animals such as dogs and nonhuman primates. Rats and mice are the standard research organisms for studies of physiology. The *C. elegans* and *Drosophila* genomes are complete, and the mouse genome is reportedly near completion, as are the genome sequences of several other organisms. Combining model organism genetics, genomics, and molecular biology will provide a powerful armamentarium to dissect the underlying mechanisms of disease states and of treatment regimens that may be applicable to humans. Indeed, given the high degree of evolutionary conservation of gene families and protein pathways throughout the eukaryotes, a comparative genomics approach is a necessity.

Key technologies developed for particular organisms are often transferable to other species, allowing researchers to utilize a number of tools at once. Certain organisms, such as *E. coli* and yeast, for example, are well suited for studies of signal transduction, DNA replication, and cell cycle regulation, all of which are key to understanding normal development and cancer. The developmental genetics of mice and *Drosophila* has revealed underlying, common pathways shared by all higher eukaryotes; the main challenge is determining how particular organisms such as humans have modified such pathways, for example, by adding more proteins. A variety of higher vertebrates share genetic regulation and physiology with humans and are the starting point for medical research. The similarities of genome and, presumably, proteome in these animals promise to validate cross-species comparisons of physiology.

Many newly developed techniques of transcript profiling and proteomics were initially developed in yeast. In a *PG* paper published in 2000, Jia and coworkers (6) used DNA microarrays to examine the expression profiles of *Saccharomyces cerevisiae* genes in strains treated with an herbicide, sulfonylurea methyl, which interferes with certain amino acid biosynthetic pathways. As the authors note, an understanding of the response of these simple fungi to chemicals may provide insight into metabolic processes of higher organisms. Indeed, characterizing such responses may become a powerful new strategy for identifying novel pharmaceutical agents.

To encourage publication in this area of model organism research, recently Monica Riley, a senior scientist focusing on *E. coli* genomics at Woods Hole, and Marco Marra, of the British Columbia Cancer Research Centre, who works on apoptosis in *Drosophila*, have been added as Associate Editors of *Physiological Genomics*. There are mini-reviews in preparation on *Drosophila* genetics in apoptosis and cancer and on the neurophysiology of learning and memory in *Drosophila*. In addition, our Journal’s call for more proteomics research should result in an increase in submission and publication of original research studies of yeast and bacteria, whose genomes are particularly well suited to broad surveys.

**Mapping Complex Traits**

In classic, “forward genetics,” selective breeding and targeted mutagenesis is combined with chromosomal mapping and molecular biology to determine the location and sequence of candidate genes. The genetic investigation of complex traits exploits these techniques along with quantitative trait locus (QTL) mapping. With a complete roadmap of the genome, the path from the identification of segregating QTLS to testable candidate genes is becoming shorter. Since there is widespread chromosomal synteny homology between rats, mice, and humans, information on candidate loci from one organism may be readily applied to another.

Our current understanding of the genome is far from complete, with each untested, predicted gene posing its own catalog of questions calling for new experiments. However, in the so-called “postgenomics” era, it will become easier to tackle the genetic basis of multigenic traits, since the transcription and expression levels of numerous genes can be measured simultaneously. Microarray technology allows researchers to identify clusters of genes whose expression levels show parallel changes in genetically distinct strains. This, then, forms the first step in the characterization of genes that are modulated in association with complex traits. Indeed, since many such transcripts will stem from uncharacterized open-reading frames, labs will utilize both forward and reverse genetics routinely in the same study. With libraries of single nucleotide polymorphisms (SNPs) available, scientists will be able to investigate haplotype-phenotype associations more comprehensively. This paves the way for the development of more targeted treatment of various disorders.

Even without the benefit of a completely sequenced or mapped genome, it is possible to take advantage of statistical tools developed for analyzing complex traits. One such example is a *PG* paper published in 1999 discussing QTLS associated with milk production in Holstein cattle (5). Using microsatellite markers and a large genetic population, the authors describe a number of loci that will have potential for candidate gene identification. These sites have statistically significant associations with complex physiological traits such as percentage of body fat and milk yield.

*PG* also publishes research on complex human genetic traits. In one such report, Hegele et al. (4) describe the association between an SNP in exon 10 of the lamin A gene and measurements of plasma leptin and other indices of obesity in a population of Canadian Oji-Cree. As would be expected, based on our conceptual understanding of polygenic syndromes such as obesity, the polymorphism accounts for only 25% of the overall variance of the log of the body mass index.
Proteomics

All biological changes, whether accompanying normal development or disease processes, involve highly complex interactions between proteins and other proteins or peptides, nucleic acids, or small molecules. Adding to this complexity, we have come to recognize the old dictum “one gene, one protein” as simplistic and highly reductionist. It is time to begin a systematic effort aimed at cataloging and understanding the entirety of proteins in a living organism, along with their respective interactions. Proteomics, the term coined to encompass these endeavors, thus refers to the study of all the proteins found within a cell at a particular functional state, including splice variants and post-translational modifications.

Using a variety of evolving methodological approaches to fractionate the biological material to be studied and the proteins contained therein, scientists are aiming for a comprehensive map of a cell’s or an organism’s protein expression. These measurements are made both qualitatively, with regard to the modifications of any one protein present, and quantitatively, with regard to the relative abundance of any given subspecies. Some of this information may be of direct functional relevance; thus assessment of the phosphorylation status of numerous proteins simultaneously may provide a “snapshot” of protein activation. Accurate quantification will be important, as seemingly minor differences in protein levels may be associated with major differences in the functional state of a cell or an organ. Here, again, early inroads have been made using lower organisms, such as yeast; and experience gained from such studies will be instrumental in up-scaling to the vastly more complex task of studying more physiologically relevant models.

PG recently published a paper describing differences in the plasma membrane proteome between a metastatic breast cancer cell line and fibroblasts (3). Using mass spectrometry to analyze proteins separated by two-dimensional gel electrophoresis, the authors identified a number of plasma membrane-associated proteins specifically expressed in the cancer cell lines. The use of proteomic information as a starting point to identify statistically significant correlations with altered cellular function may contribute importantly to a better understanding of human disease. The journal has recruited several Editors who are pioneers in proteomics, including Ruedi Aebersold and Andrew Link. There will be mini-reviews on new proteomics database initiatives as well as on the latest advances in technology.

Conclusions

Progress in our understanding of the complex interaction of genes, RNAs, proteins, and organ systems that comprise an organism’s physiology will rely on the creative use of the approaches outlined above. The journal focuses on innovations in genomics and proteomics with real promise for medical research. Physiological Genomics will serve as a resource for scientists interested in using an integrative and comparative approach in their work, an approach that holds real promise for future breakthroughs. Furthermore, researchers who submit to the journal are assured a rapid, thorough, and fair peer review. The opportunities that lie ahead for this journal, and for biology, are tremendous indeed.

For the Senior and Associate Editors

Victor J. Dzau
Editor-in-Chief
Susan Glueck
Deputy Editor
Physiological Genomics
December 2001, Volume 7

REFERENCES