Mouse Phenome Project: understanding human biology through mouse genetics and genomics

The Human Genome Project is generating vast amounts of new information at breakneck speed and causing a fundamental shift in disease research. Now with the availability of a nearly complete, high-accuracy sequence of the mouse genome (7), a new and powerful paradigm for biomedical research is established. The remarkable similarity of mouse and human genomes, in both synteny and sequence, unconditionally validates the mouse as an exceptional model organism for understanding human biology. The discovery among inbred mouse strains of defined regions of high and low genomic variation inherited primarily from two ancestral Mus subspecies (6) holds great promise to make mapping and positional cloning more rapid and feasible. Haplotype maps of inbred mouse strains combined with sophisticated delineation of their phenotypic variation and gene expression patterns will enable complex trait analysis on an unprecedented scale. This issue of Journal of Applied Physiology highlights inbred strain surveys exploring phenotypic variation in drug responses [see Crabbe et al. (1) and Watters et al. (8) in this issue]. These mouse initiatives demonstrate a viable, cost-effective alternative to human research requiring family studies, population linkage analysis, or genome-wide genotyping on a multitude of individuals for association mapping.

THE POWER OF INBRED MOUSE STRAINS

The laboratory mouse is an extraordinary genetic resource for complex trait analysis. Besides the immeasurable benefits of sequence data and comparative genomics, the mouse genome may be manipulated through selective breeding strategies or mutated with efficient techniques to create variants that otherwise are identical to the parental strain. The ever-expanding mutant resource joins the hundreds of distinct strains that have already been generated and described. For many years, inbred mouse strains have been key in the genetic analysis of Mendelian traits and have provided an excellent entry point in research for understanding complex metabolic and aberrant pathways (4), including those diseases with associated risk factors such as arthritis, diabetes, obesity, cancer, hypertension, and cardiovascular disease. Inbred strains make excellent animal models for a wide spectrum of diseases and conditions because, like humans, they differ at multiple genomic sites and are extremely diverse phenotypically. The wide array of phenotypes in inbred strains is due to the combinatorial potential of many variable sites derived from the admixture of ancestral Mus genomes as well as the naturally occurring mutations and compensatory mechanisms selected during inbreeding. Trait distributions can be quantified with robust phenotyping methods, and experimental complexity can be diminished by strictly controlling both the environment and heterozygosity, two unavoidable and significant variables in human studies that have great consequences on phenotype.

The phenotype-driven approach to complex trait analysis with inbred strains would be greatly facilitated in any laboratory that has access to universally available strain characteristic data generated under controlled conditions. This resource would provide great opportunity to identify new (or better) mouse models at the outset of a study, potentially accelerating the research pace and conserving resources. Comprehensive data are lacking for many existing and commonly used strains, however, leaving much of their potential for improving human health untapped. To fill this resource gap, the Mouse Phenome Project was launched in May 2000 at the request of the research community.

THE MOUSE PHENOME PROJECT

The Mouse Phenome Project is an ongoing international collaborative effort headquartered at The Jackson Laboratory that is designed to enhance the resources presently available for the laboratory mouse. With academic and industrial participation, the Mouse Phenome Project promotes the quantitative phenotypic characterization of a defined set of mouse strains under standardized conditions and makes the data publicly available through a web-accessible database. Data for a wide range of parameters are annotated and stored in the Mouse Phenome Database (MPD), along with submitter’s contact information, detailed protocols, and environmental parameters. Tools for data retrieval and analysis are available through a website interface (www.jax.org/phenome). Universal access to centralized strain data enables investigators to choose appropriate strains for many systems-based approaches, including modeling disease processes, physiological studies, toxicology, mutagenesis, and biodefense research.

With the guidance of an International Steering Committee, 40 inbred strains were carefully chosen and are recommended for testing; wild-derived inbred strains are included to bring in important new alleles not present in the more commonly used laboratory strains (3). The set of 40 priority strains and their reproducible derivatives (F1 hybrids, recombinant inbreds, consomics, and so forth) comprise a reference population to focus the Mouse Phenome Project on a fixed set of alleles that can be accessed in countless, but reproduc-
ible, combinations. As a community resource, data from other strains are also included in the MPD, provided that the strains are verifiable and readily available and that other recommendations are followed. Project guidelines are carefully formulated to standardize sampling and testing with the goal of generating high-quality data that can be reliably compared across laboratories and over time (feasible because proper inbreeding stabilizes the genome, providing the unique opportunity to repeatedly access the “fixed” genomes of the reference population). The framework for data consistency is a cost-effective strategy and key to the success of the Mouse Phenome Project because comprehensive strain profiles can be constructed from cumulative data contributed by the research community, effectively increasing the value of the strains characterized and the MPD. Genotypic data are also being deposited in the MPD. Predicting genotype-phenotype associations with quantitative data will accelerate the rate of identifying genes underlying normal metabolic and disease pathways and will facilitate efforts to determine gene function.

**MOUSE PHENOME DATABASE AS A RESEARCH TOOL**

Beyond data storage and retrieval, the MPD is an integral tool for the analysis of complex biological systems. For example, after an appropriate model is identified via up-to-date strain profile data, MPD analysis tools can access key characteristics of the model strain and identify all measurements in the MPD with compelling correlations to help reveal overlapping phenotypic networks and elucidate genetic components with pleiotropic effects. Gleaning this information may instantly provide clues for developing strategies that could alleviate a genetically determined condition or that apply to risk assessment, diagnosis, prevention, or other treatment options for the trait of interest. On the other hand, if an appropriate mouse model cannot be found, it would be prudent to choose strains for selective breeding strategies to “build” a mouse with traits of interest. This approach is feasible when using MPD strain profile data and tapping into the rich and ever-expanding mouse mutant resources.

In addition to revealing pleiotropic effects, the MPD can assist with other aspects of complex trait analysis, including detecting modifier genes, predicting haplotype associations, and determining the influence of environment on phenotype. The MPD is only one of many tools for dissecting complex traits, however. Multiple, community-wide research approaches and resources will be required to reveal the full spectrum of genomic components and environmental factors that comprise a complex trait. Genotype- and phenotype-driven approaches, sequence and gene expression data, microarray technology, comparative genomics, and cleverly designed second-generation strain resources are all essential. Robust computational methods and informatics are necessary to accommodate, mine, and connect the vast quantities of data from diverse sources. To avoid duplication of effort and to maximize resources, communication and standardization are essential throughout the research community. The Mouse Phenome Project is coordinating with the Complex Trait Consortium (www.complextrait.org, Ref. 5), Mouse Genome Informatics (www.informatics.jax.org), and large-scale genotyping consortia. The Mouse Phenome Project is eager to establish additional partnerships with individuals, groups, and organizations willing to participate in building and improving community resources for the mouse. It is imperative to also coordinate and integrate with relevant databases for other model organisms and of course with human research resources, such as the recently proposed Human Phenome Project (2).

**MOUSE PHENOME PROJECT STATUS**

The Mouse Phenome Project has moved beyond the pilot phase with widespread support from the research community. A number of participating investigators are affiliated with the Mouse Phenome Project Collaborations Program, which fosters three-way partnerships among the research community, funding partners, and the Mouse Phenome Project to support typing efforts in highly qualified laboratories through a peer-review process. Investigators from large-scale phenotyping centers and independent laboratories are voluntarily contributing datasets for inclusion in the MPD. Presently, there are about 400 measurements for phenotypes relevant to atherosclerosis, blood disorders, cancer susceptibility, neurological and behavioral disorders, sensory function defects, gallstones, infectious disease susceptibility, pulmonary responsiveness, hypertension, osteoporosis, and obesity. Over 150,000 single-nucleotide polymorphisms have been identified in a subset of priority strains and contributed to the MPD for worldwide access (6, 9). The number of measurements is expected to more than double in 2003, along with the acquisition of genome-wide single nucleotide polymorphisms for more than 45 strains. Because the power of the MPD depends on data quality and density, the database will be regularly updated and expanded as more phenotypic domains are incorporated and higher-resolution phenotypic data are generated from new, more sophisticated technologies.

This is the first issue of the *Journal of Applied Physiology* to highlight research papers from Mouse Phenome Project participants. As an ongoing series, these and future articles will be marked with the Mouse Phenome Project banner. In this issue, John Crabbe, Doug Wahlsten, and colleagues (1) report coordinated studies from two laboratories (Oregon Health and Science University and University of Alberta) on the differential effects of ethanol on inbred strains, providing important conclusions about the validity of particular task parameters and critical issues that affect data interpretation. In a separate paper, collaborators from Washington University, Howard Mcleod, Timothy Graubert, James Watters, and colleagues (8), explore strain and sex differences in re-
sponse to cyclophosphamide, a widely used but unpredictable chemotherapy agent, to identify genetic factors that may help determine whether this drug is likely to be effective or toxic in human patients. These studies underscore the utility and potential of the MPD, which provides the framework for understanding complex traits, like drug responses and toxicity, with applications for pharmacogenomics and gene discovery. As with all Mouse Phenome Project initiatives, these comprehensive phenotyping studies are relevant to human health and well being, and significant findings should translate well to human research.

Molly Bogue
The Jackson Laboratory, Bar Harbor, Maine 04609

REFERENCES