Genetica **122:** 71–74, 2004.

The Mouse Phenome Project

Molly A. Bogue & Stephen C. Grubb
The Jackson Laboratory, 600 Main Street, Bar Harbor, Maine 04609 USA (Phone: 207 288 6016; Fax: 207 288 6079; E-mail: mollyb@jax.org)

Key words: bioinformatics, database, inbred mouse strains, phenotyping

Abstract

The laboratory mouse is the organism of choice for many studies in biology and medicine. Reliable phenotypic data are essential for the full utility of genotypic information emerging from efforts to sequence human and mouse genomes. The Mouse Phenome Project has been organized to help accomplish this task by establishing a collection of baseline phenotypic data on commonly used and genetically diverse inbred mouse strains and making this information publicly available through a web-accessible database. The Mouse Phenome Database (MPD) is being developed to manage these data and to provide researchers with tools for exploring both raw phenotypic data and comparative summary analyses. The MPD serves as a repository for detailed protocols and raw data. This resource enables investigators to identify appropriate strains for (1) physiological testing, (2) drug discovery, (3) toxicology studies, (4) mutagenesis, (5) modeling human diseases, (6) QTL analyses and identification of new genes and (7) unraveling the influence of environment on genotype.

Abbreviations: MPD – Mouse Phenome Database (www.jax.org/phenome), QTL – quantitative trait loci.

Introduction

Inbred strains are obligate homogygotes and collectively exhibit enormous phenotypic diversity, indicating the presence of many allelic variants. Their fixed genomes can be accessed repeatedly over time and in multiple laboratories, making them a powerful research tool. Numerous inbred strains are currently available and are characterized to some degree – some phenotypic information is available, and in some cases the genes responsible for variation have been identified. Unfortunately, systematic and comprehensive phenotypic characterization of these inbred strains is lacking. To fill this information gap, the concept of 'a mouse phenome project' was presented in a review article (Paigen and Eppig, 2000). Together with data generated from mouse mutagenesis studies, strain characteristics data provides a much needed and an extremely rich source of phenotypic data on a large and growing set of mouse strains.

The notion of the Mouse Phenome Project was the topic of the Strain Characterization Workshop held at The Jackson Laboratory in May, 1999 where 37 scientists, representing 17 research institutions and corporations assembled. Workshop participants concluded that:

- Comprehensive phenotypic information on inbred mouse strains is urgently needed because the laboratory mouse, with its hundreds of inbred, specialized, and mutant strains, serves as the primary animal model for exploring genetic variation and human biology.
- Reliable phenotypic data are essential for realizing the full utility of genomic information that will emerge from sequencing the mouse genome.

- The scope of this large-scale collaborative project requires international cooperation and both academic and industrial participation.
- Experts in diverse fields of biomedical science should generate this phenotypic data.
- A central, web-accessible database should be developed and housed at The Jackson Laboratory so that it can be integrated with the Mouse Genome Database (MGD).

Because of its broad scope and nature, workshop participants recognized that the success of this project is dependent on an international collaborative effort. Appropriately, an international steering committee was recommended to spearhead the project. An eleven-member Steering Committee (Table 1), representing five countries in both the academic and corporate sectors is overseeing this challenging endeavor. The Jackson Laboratory has taken a lead role in facilitating this international effort by recruiting a project director and database manager/software engineer.

Table 1. Mouse Phenome Project: members of the steering committee

Rudi Balling, PhD	Institute of Mammalian Genetics GBF, Germany
Steve D.M. Brown, PhD	Medical Research Council Harwell, United Kingdom
Allen Cowley, Jr., PhD	Medical College of Wisconsin USA
Kenneth Fasman, PhD	AstraZeneca USA
Jean-Louis Guenet, PhD	Institut Pasteur France
Steve Kaminsky, PhD	Uniformed Services University USA
Miriam Meisler, PhD	Univeristy of Michigan Medical School USA
Kazuo Moriwaki, PhD	The Graduate University for Advanced Studies Japan
Kenneth Paigen, PhD	The Jackson Laboratory USA
Joseph Takahashi, PhD	Howard Hughes Medical Investigator Northwestern University, USA
Richard Woychik, PhD	The Jackson Laboratory USA

The Mouse Phenome Database (MPD) serves as a repository for protocols and raw data, which are available for on-screen viewing and downloading. The MPD website provides current project information, focal points for community participation, and navigation routes to phenotypic data. Query tools are available to compare data from diverse sources. The MPD is populated with data generated through the efforts of the Mouse Phenome Project, data contributed from the scientific community, and data published in the scientific literature. The MPD is updated regularly, making this a comprehensive source of phenotypic information for inbred strains of mice.

Project recommendations

General recommendations

Recommendations made by workshop participants and approved by Steering Committee members were carefully devised to provide consistency among investigators to produce reliable data from multiple laboratories. The rationale for drawing up a set of recommendations ensures that all studies are conducted under standardized conditions, effectively reducing the number of variables and making the data as useful as possible.

Genetic purity of inbred strains is critical for this project; therefore, it is highly recommended that mice are received from a reputable breeding source. An acclimation period of at least 2 weeks is recommended before testing. Phenotyping assays should be performed on groups of inbred strains (see priority groups below), at an age of 10–14 weeks on 10 males and 10 females of each strain. Raw data should be submitted along with detailed protocols and animal documentation (information about source, diet, environment, and health status of the mice). Undoubtedly there are circumstances where it is impractical to follow these recommendations; in these instances, modifications are reported and posted with the data.

Assays should be carried out by at least two independent laboratories for data validation. Concurring data collected from independent sources will be annotated accordingly. This class of data is expected to be the most useful and reliable. Likewise, appropriate annotations will be assigned for data generated from projects following the

project recommendations that are in agreement with data from existing literature or other phenotyping projects. Validation annotations will serve as indicators for data quality and reliability.

Inbred mouse strains

Forty inbred strains were chosen and prioritized by the research community and approved by the Steering Committee according to their usage and genetic diversity (Table 2). The priority status of some strains is expected to change over time, depending on research trends and community input. *Priority Group A* strains are highest priority and are recommended for all projects; these strains are generally easy to maintain with good reproductive performance. Baseline phenotypic information from this group is critical because these strains are widely used with available genetic and phenotypic information, providing useful data for comparison and validation. Present in this group are strains that will be sequenced by the Mouse Sequencing Consortium (C57BL/6J) or Celera (A/J, DBA/2J, 129), strains that are progenitors in transgenesis or mutagenesis studies, and strains that are progenitors of recombinant inbred, consomic, or congenic strains. The strains in this group are genetically diverse, particularly with the inclusion of the wild-derived strains, CAST/Ei and SPRET/Ei. *Priority Group B* strains have intermediate use in the community and have good reproductive performance. The 10 strains in this list are encouraged for all projects. *Priority Groups C and D* strains represent more depth in genetic diversity and include additional wild-derived inbred strains. Derivatives of the priority strains (F1 hybrids, recombinant inbred, mutant mice, etc.) can be phenotyped along with inbred strains as well. These valuable data will be included in the MPD.

Table 2. Mouse phenome project priority strains

A	B	C	D
129S1/SvImJ	AKR/J	BUB/BnJ	BTBR T+ tf/tf
A/J	C57L/J	C57BL/10J	C57BR/cdJ
BALB/cByJ	C58/J	C57BLKS/J	CE/J
C3H/HeJ	MOLF/Ei	CBA/J	I/LnJ
C57BL/6J	NOD/LtJ	CZECHII/Ei	JF1/Ms
CAST/Ei	NZB/BlNJ	KK/HlJ	MA/MyJ
DBA/2J	PERA/Ei	LP/J	NON/LtJ
FVB/NJ	PL/J	MSM/Ms	NZW/LacJ
SJL/J	SM/J	RIIIS/J	PWK/Ph
SPRET/Ei	SWR/J	WSB/Ei	SEA/GnJ

Listed by priority groups A > B > C > D.
Strain nomenclature has been updated since submission of this manuscript.

Phenotyping assays

Phenotyping assays have been prioritized by the Steering Committee according to merit with considerations on time, cost, and effort required for completion. Phase 1 tests include basic clinical measurements, histology, and some neurosensory/behavioral assays while Phase 2 tests are more specific to certain disciplines and are generally more time consuming and costly. More information may be access through the MPD website. Multiple sampling from the same animal is encouraged if measurement validity is not affected; and measurements in duplicate or triplicate should be obtained if the assay permits. It should be stressed that prioritizing these assays into Phase 1 and 2 tests is only a recommendation – the rationale is to obtain useful data of broad interest as quickly as possible. Data from other phenotyping assays are valuable for the research community, including data generated from treatment studies (aging, drug studies, diet effects, etc.) or from new phenotyping technologies.

Animal documentation

Detailed information about the mice should be submitted for each project. Source of mice, diet, environment, and health status are critical when studying gene-environment interactions (or when attempting to sort out data discrepancies between independent laboratories).

Data submission and project webpages

One of the goals of the Mouse Phenome Project is to make raw phenotypic data from many sources available for interactive viewing and downloading in a standard way. Flexible submission guidelines have been devised in order to produce clean, uniform data sets that can be understood and used by others. Detailed guidelines are posted on the website.

For each individual project, a webpage is created where animal documentation, protocol information, and raw data are accessible. Raw data are

summarized in standard tabular and graphical formats. It is possible to generate customized analyses and reports based on user specifications. For any given project, it is also possible to post investigator's comments (such as any discussion, conclusions, or recommendations to MPD users), funding sources or other acknowledgments, citation information for publications, and URLs for relevant websites.

Each project is assigned a unique identifier (MPD accession number). This accession number can then be referenced in journal publications, utilizing the MPD as a repository for raw data and detailed protocols.

Mouse Phenome Database capabilities

The following list of capabilities have been implemented or will be possible in the near future:

- On-screen viewing and/or downloading of raw data.
- Access to animal documentation and protocol information.
- Viewing summary graphs of measurements (selected strain or all strains).
- Viewing summary table of all strain outliers for a particular measurement.
- Ability to define high- and low-end outlier values.
- Searching by key word, investigator, measurement, phenotyping assay, or strain name.
- Browsing indices for individual projects, investigators, subjects (key words), measurements, strains, and status/current recommendations.
- Ability to perform log transformation on measurement data and graphical representations.
- Correlation determination (scatterplot with correlation coefficient) between measurements with ability to select strains and sex.

Progress report

The MPD contains several phenotyping projects from a variety of disciplines. Contents are maintained on the website for overall status of the Mouse Phenome Project, status of individual projects, current assay recommendations, availability of projects, and links to relevant information.

The Mouse Phenome Project is driven by the needs of the research community. Requests are welcome; please forward to phenome@jax.org.

Acknowledgments

The initial stages of this project have been funded by generous contributions from: AstraZeneca, BD Biosciences, Burroughs Wellcome Fund, Aventis, Hoffman-La Roche, Merck Genome Research Institute, and The Jackson Laboratory.

PHENOME-LIST has been initiated as a forum for discussions related to the Mouse Phenome Project: http://aretha.jax.org/pub-cgi/phenome/mpdcgi?rtn=docs/phenomelist) The MPD homepage is http://www.jax.org/phenome

References

Paigen, K. & J. Eppig, 2000. A Mouse Phenome Project. Mamm. Genome. 11: 715–717.