

A genome end-game: understanding gene function in the nervous system

Carol Bult, Warren A Kibbe, Jay Snoddy, Martha Vitaterna, Doug Swanson, Stephanie Pretel, Yanxia Li, Moses M Hohman, Eugene Rinchik, Joe S Takahashi, Wayne N Frankel & Dan Goldowitz

© 2004 Nature Publishing Group <http://www.nature.com/natureneuroscience>

The mouse has offered great promise as a model organism to study brain function and behavior; however, neurological phenotypes in mice are often detected by individual investigators in a low-throughput fashion, studying natural variants of mice or mice with spontaneous mutations or gene knock-downs. In 2000, because of the success of large-scale mutagenesis programs in other model organisms (the fruit fly *Drosophila melanogaster*, the nematode *Caenorhabditis elegans* and zebrafish) and the success of large-scale mouse mutagenesis efforts in Europe (www.mgu.har.mrc.ac.uk/mutbase/; www.gsf.de/ieg/groups/enomouse.html), the National Institutes of Health (NIH) began to support three mutagenesis centers in the US. Their collective purpose was to detect, characterize and distribute new mouse mutants with primarily neurological phenotypes. These centers are located at Northwestern University (<http://genome.northwestern.edu>), The Jackson Laboratory (JAX; <http://nmf.jax.org>) and the Tennessee Mouse Genome Consortium (TMGC; www.tnmouse.org).

These centers use *N*-ethyl-*N*-nitrosourea (ENU) to create mutations in the mouse genome and then screen the resulting mutants for neurological phenotypes. ENU is an ethylating agent that is both mutagenic and cytotoxic in mouse spermatogonial stem cells¹. ENU is particularly valuable for inducing allelic series of mutations because it causes primarily base-pair substitutions²,

resulting in a wide range of mutation outcomes. Importantly, of the 38,000 mutations in about 1,500 genes that cause aberrant phenotypes in humans, about 70% of these are of the single-base pair variety (<http://archive.uwcm.ac.uk/uwcm/mg/docs/hahaha.html>), and so it would appear that the ENU-based approach in mice mimics human genetic disease.

Producing mutant mice at the 'industrial' level brings considerable challenges in laboratory information management systems. It requires a database architecture that is flexible and scalable with reliable and secure data storage, and that allows acquisition of data from multiple locations. It should permit tracking of mice individually or collectively as families, and should allow the user to create automated reports and use statistical analysis tools. Effective sharing of data and information about mice and their phenotypes requires the use of standardized data exchange formats, the use of existing community databases and resources, and the development of common vocabularies for new types of data. To address these issues, the three NIH-funded neuroscience mutagenesis centers have formed a virtual distribution center (www.neuromice.org), which allows researchers to learn about and acquire mutant mice of interest. The website uses a

standardized XML data exchange format; each mouse file includes a description of the mutant mouse, heritability and mapping status information, distribution notes and links to the web page of the center from which the mutant was derived.

To help searches, a common vocabulary of phenotypic classification terms is used to group mice into meaningful bins. On the neuromice.org site, mouse models are classified by both phenotypic domain and by assay. The phenotypic domain terms include Aging, Epilepsy, Metabolism, Neuroendocrine, Social Behavior, Eye, Drug Abuse, Alcohol, Circadian Rhythm and Pain/Nociception. Within each of these domains, the mice are further subdivided according to the type of assay that was used to measure the phenotype. For example, in the phenotype domain of Alcohol there are mutants that have been identified using the following assays: activity after an injection of ethanol, and the two-bottle ethanol choice test. The use of controlled vocabularies to classify mice allows the individual centers to name mouse lines according to their own conventions while providing the research community with a consistent mechanism for searching for mice with specific characteristics.

Each group also regularly submits data and information on the mutant phenotypes they have discovered to well-established

Carol Bult, Stephanie Pretel and Wayne N. Frankel are at the Neuroscience Mutagenesis Facility, The Jackson Laboratory, Bar Harbor, Maine 04609, USA. Warren Kibbe, Martha Vitaterna, Yanxia Li, Moses Hohman and Joe S. Takahashi are at the Center for Functional Genomics, Howard Hughes Medical Institute, Northwestern University, Evanston, Illinois 60208, USA. Jay Snoddy and Eugene Rinchik are at the Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA. Doug Swanson and Dan Goldowitz are at the Center for Genomics and Bioinformatics, Department of Anatomy and Neurobiology, University of Tennessee Health Science Center, Memphis, Tennessee 38163, USA.
e-mail: dgold@nb.utmem.edu

community databases. For example, all of the mutants are being registered with the Mouse Genome Informatics (MGI) database at JAX (www.informatics.jax.org), which is the primary community database for the laboratory mouse³. Integrating the data allows researchers to get a comprehensive overview of the connections between genes, alleles and phenotypes in the mouse. In the longer term, the extensive strain-specific and mutant phenotypic data produced by each center will provide great synergy with other databases of mouse phenotypic data such as the Mouse Phenome Database at JAX (<http://aretha.jax.org/pub/cgi/phenome/mpd-cgi?rtn=docs/home>) and the gene expression profiling of BXD recombinant inbred strains at the University of Tennessee (<http://nervenet.org>).

The mission of the three centers is to provide the scientific community with new mouse models for understanding gene function in the nervous system. To date, over 100 new mouse mutants relevant to neurological disorders in humans have been generated by these centers. The mutants include mice with defects in balance, blindness, susceptibility to seizures and abnormalities in circadian rhythm, open field behavior, pain responses and hearing. These mutant lines can be used as models to study disorders of neural function. For example, the Center for Functional Genomics at Northwestern University discovered a new mutant named 'overtime' that defines a clock locus that maps to a region of mouse chromosome 14 where there are no known circadian genes. The Neuroscience Mutagenesis Facility at JAX, using an electro-

convulsive threshold screen, has identified two new mutant alleles in the *Kcnq2* gene, whose human homolog is mutated in a form of human epilepsy. Finally, the Neuromutagenesis Program of the TMGC, using tail suspension and open-field behavioral screens, has identified several distinct anxiety/depression or emotional behavior mutants, four of which are localized to mouse chromosome 7 and one to mouse chromosome 15. The discovery of the mutant genes that give rise to these and other mutant phenotypes is another powerful strategy for the functional annotation of the genome.

1. Russell, W.L. *et al. Proc. Natl. Acad. Sci. USA* **76**, 5818–5819 (1979).
2. Noveroske, J.K., Weber, J.S. & Justice, M.J. *Mamm. Genome* **11**, 478–483 (2000).
3. Bult, C.J. *et al. Nucleic Acids Res.* **32**, D476–D481 (2004).

WebQTL: rapid exploratory analysis of gene expression and genetic networks for brain and behavior

Elissa J Chesler, Lu Lu, Jintao Wang, Robert W Williams & Kenneth F Manly

Brain mRNA expression is modulated by numerous genetic factors and often varies substantially between strains of mice that have been reared in a standard laboratory environment. Examples include members of the NMDA receptor family that are critical in learning and memory, and genes involved in synaptic vesicle trafficking. Molecular variation of this type is often heritable and is produced by genetic polymorphisms at many locations across the genome. Differences in both alleles and mRNA levels will often produce significant behavioral, pharmacological and neuroanatomical variants¹. Over the past several years, with support from the NIH Human Brain Project, we have assembled a suite of databases and web-based analysis software called WebQTL (www.webqtl.org). WebQTL is a freely accessible system that

exploits sophisticated gene mapping methods^{2,3} to rapidly perform whole-genome analysis at many levels—from differences in NR2B mRNA levels to differences in open-field activity levels.

WebQTL has three major applications: exploring variation in gene expression using a panel of more than 30 recombinant inbred strains and several different tissues (for example, forebrain, cerebellum, hematopoietic stem cells); mapping upstream gene loci that modulate transcript levels; and studying networks of genetic correlations among ~100,000 transcript assays and 650 published phenotypes. Additional features include tools for the simultaneous analysis of groups of traits, custom annotation of Affymetrix probes and probe sets, and external links to the Gene Ontology Machine (<http://genereg.ornl.gov/gotm>), the Gene Expression Atlas (<http://expression.gnf.org>), NCBI (www.ncbi.nlm.nih.gov) and the Genome Browser (<http://genome.ucsc.edu>). The integration of diverse data types provides a powerful resource for exploratory systems biology.

Data in WebQTL have been acquired from two common progenitor strains, C57BL/6J (B) and DBA/2J (D), their F1 hybrid, and a set of different BXD recombinant inbred

(RI) strains. The two progenitor strains, B and D, have both been sequenced and are known to differ at roughly 1.8 million single-nucleotide polymorphisms (SNPs) across the mouse genome. This amounts to an average of one SNP every 1,500 base pairs. Each of the BXD strains is a unique 'mosaic' of chromosomal segments inherited from either the B or D progenitor strain⁴. About 34 BXD strains are available from The Jackson Laboratory, and an additional 45 strains will soon be available from The University of Tennessee. A wide range of phenotypes seen in the BXD reference population are also incorporated in WebQTL (see WebQTL's *Published Phenotypes* database). WebQTL also includes high-density marker maps based on 779 microsatellites⁵ and SNPs. By testing the association of genetic markers with variation in transcript levels and other traits, WebQTL maps the quantitative trait loci (QTLs) that are likely to contain modulators of these complex phenotypes. The value of this BXD reference population to the research community grows multiplicatively as additional phenotypes are collected and integrated into WebQTL.

The gene encoding the NMDA NR2B receptor subunit *Grin2b* provides an example of the type of analysis possible using

All authors are at the Center for Genomics and Bioinformatics, University of Tennessee Health Science Center, 855 Monroe Avenue, Memphis, Tennessee 38163, USA. Elissa J. Chesler, Lu Lu and Robert W. Williams are in the Department of Anatomy and Neurobiology, and Kenneth F. Manly and Jintao Wang are in the Department of Pathology and Laboratory Medicine.
e-mail: echesler@utm.edu