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 mouse[3]. 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 variants[1]. 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 strain[4]. 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 microsatellites[5] 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.
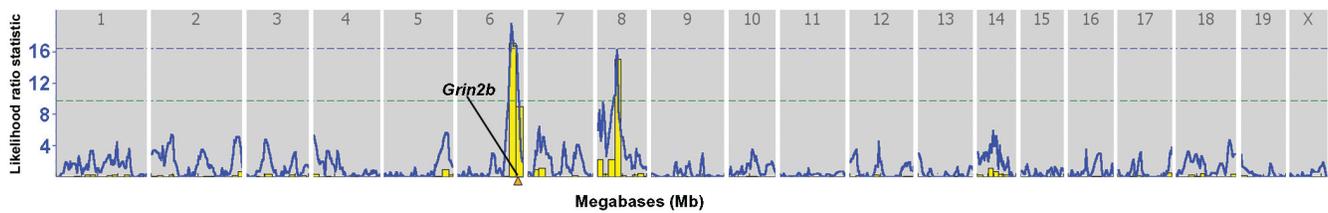e-mail: echesler@utmem.edu

Figure 1 A map of quantitative trait loci on chromosomes 6 and 8 regulating the expression of the ionotropic glutamate receptor gene *Grin2b*. All mouse chromosomes, with the exception of chromosome Y, are plotted along the *x*-axis. The orange triangle is the transcript location; the solid blue trace is the likelihood ratio statistic for association of the phenotype with genotype across the genome. The horizontal lines indicate permutation significance thresholds for significant ($P < 0.05$, blue) and suggestive ($P < 0.67$, green) loci. Yellow bars indicate the most likely location of QTL peaks by bootstrap analysis.

WebQTL. The abundance of *Grin2b* mRNA transcript in several forebrain data sets varies approximately two-fold across 35 strains. Half of the variation in expression is heritable; and this makes it practical to map the responsible QTLs. *Grin2b* has two major QTLs, one on chromosome 8 (**Fig. 1**), and another on chromosome 6 near the transcript itself—probably associated with one or more of the 556 SNPs in this gene. The QTL on chromosome 8 is particularly intriguing, but it would be a project in its own right to discover the single correct gene associated with this QTL among ~50 candidates.

WebQTL allows hundreds of covariates with high correlations to *Grin2b* to be rapidly extracted, analyzed and graphed. These include many ethanol-related phenotypes, as well as measures of locomotor activity, anxi-ety, maze learning, neuron cell numbers, hippocampal and cerebellar volumes and adult neurogenesis. Similarly, hundreds of transcripts with expression differences that covary with *Grin2b* expression can be extracted. These include *Mpdz*, which encodes a protein involved in the clustering and endocytosis of NMDA receptors; *Ag1g1*, which encodes an adaptor protein complex that is part of the clathrin coated pit; *Inpp4a*, which is involved in the cycling of clathrin to the Golgi; and *Nfm,* which encodes a neuro-filament protein. At least five members of the kinesin family of motor proteins, essential for transport of the endocytic vesicle through the axoplasm, also have expression levels that correlate positively with *Grin2b*. What is intriguing about this example is that genetic variation underlying receptor expression may simultaneously influence expression of several critical components of synaptic receptor cycling. Several members of this extended family of genes also have known relations to alcohol-related phenotypes[6]. This example is just one of many query paths that can be navigated rapidly in WebQTL to generate and test hypotheses using this reference set of RI strains.

1. Brem, R.B., Yvert, G., Clinton, R. & Kruglyak, L. *Science* **296**, 752–755 (2002).
2. Chesler, E.J. *et al. Neuroinformatics* **1**, 343–357 (2003).
3. Wang, J., Williams, R.W. & Manly, K.F. *Neuroinformatics* **1**, 299–308 (2003).
4. Taylor, B.A. *et al. Mamm. Genome* **10**, 335–348 (1999).
5. Williams, R.W., Gu, J., Qi, S. & Lu, L. *Genome Biol.* **2**, 0046.1–0046.18 (2001).
6. Fehr, C., Shirley, R.L., Belknap, J K., Crabbe, J.C. & Buck, K.J. *J. Neurosci.* **22**, 3730–3738 (2002).

Neurodatabase.org: networking the microelectrode

Daniel Gardner

Electrophysiological signals report activity of single neurons, neuronal arrays and networks. Our understanding of neural coding, information transmission and brain processes would be aided if such data could be made available for further analyses that could integrate and compare findings from individual laboratories, and test multiple hypotheses. Such analyses require access to actual sets of digitized data themselves, rather than to printed views of the data through static journal figures[1]. However, data sharing of this sort requires agreement on techniques, formats and ownership, as well as methods for classification and selection.

The Laboratory of Neuroinformatics, supported by the Human Brain Project of the US National Institutes of Health (NIH), has developed neurodatabase.org—a freely accessible database to aid sharing of neurophysiological data. This resource acquires, organizes, annotates, archives, delivers and displays single- and multi-unit neuronal data from mammalian cerebral cortex[1]. The database is presently populated with somatosensory recordings, with planned expansion to data from other regions. It is available to any user with a contemporary Java-enabled networked computer.

The underlying data model in the database provides standards for data archiving, description and exchange. Java-based tools dynamically reflect this data model and support multiplatform assembly, upload, annotation, search and acquisition. They also provide views of XML-wrapped data on user-controlled, expanded or contracted time scales, accompanied by descriptive metadata. Users can acquire actual datasets, supplied by multiple laboratories, recorded with different techniques from different preparations, all in a common format and annotated compatibly, for extended analyses. Additional user tools, such as Bruxton DataGlobe, under development as a joint academic–corporate effort, will ease reanalysis of shared data by permitting data upload from within data acquisition applications, and data download directly to standard acquisition routines.

Effective use and re-use of shared data requires methods for selecting specific datasets from among others in a repository. The neurodatabase.org QueryTool allows searching by data-descriptive metadata terms, specified by the submitters, from multiple hierarchies of controlled vocabularies. For example, data from particular cortical regions can be sought by either general terms, such as "somatosensory" or more spe-

Daniel Gardner is at the Laboratory of Neuroinformatics, Department of Physiology, Weill Medical College of Cornell University, New York, New York 10021, USA.
e-mail: dan@aplysia.med.cornell.edu