

Oral (talk requested)

Cold Spring Harbor Meeting, May 2001

**LARGE-SCALE MAMMALIAN PHENOTYPE AND GENOTYPE ANALYSIS:
TOWARDS INTEGRATED COMPUTATIONAL AND EXPERIMENTAL RESEARCH****Erich Baker^{1,3}, Miriam Land¹, Doug Hyatt^{1,3}, Stefan Kirov^{1,3}, Philip LoCascio¹,
Barbara Jackson¹, Denise Schmoyer¹, Manesh Shah^{1,3}, Jay Snoddy^{1,2,3}, the ORNL
Computational Biology Section, and the Tennessee Mouse Genome Consortium.**¹Oak Ridge National Laboratory ²Univ. of Tennessee Center of Excellence in Genomics
and Bioinformatics ³UT-ORNL Graduate School of Genome Science and Technology
<http://www.tnmouse.org> <http://compbio.ornl.gov>

The Tennessee Mouse Genome Consortium (TMGC) is establishing collaborative research efforts in the mouse. A recently established TMGC project is screening for neurological phenotypes in mice after ethylnitrosourea (ENU)-mutagenesis. This project is currently screening approximately 10% of the mouse genome from portions of chromosomes 7, 15, 10 and X. The project's approach has several advantages. One advantage is that the screen finds recessive mutants. Another general TMGC advantage is that it brings together researchers and resources from several institutions. (See www.tnmouse.org for researchers and institutions.) The integration of phenotyping effort from geographically-distributed domain experts in different biological systems requires integrative bioinformatics support. The overall approach and the specific information systems under development may be reusable for other geographically-distributed research projects. The data the TMGC wishes to collect include molecular phenotypes (e.g. data from RNA expression and proteomics). We will present the current and planned phenotype analysis, available data, and several information systems that are being built to support TMGC research.

We are coupling the approaches of the TMGC with an evolving computational biology program for mammalian Comparative Phenotype and Complex Biosystem Analysis. This adds a number of computational scientists from several institutions to the TMGC experimental research. Computational testbeds for this group can include the genes, phenotypes, and biosystems encoded by the genome regions that are being experimentally analyzed by the TMGC in the mouse and their counterparts in other species, including the human. We are exploring a number of approaches, some of which are only feasible with complete or near-complete genomes and the application of high-performance computing. We will present progress toward this analysis, including cross-species comparison of conserved non-coding sequences (CNSs) between mouse and man and comparative alternative splicing. We believe that this new effort and other integrative efforts that bring together experimental and computational research are a core foundation to build from if we are to begin to discover, model, and understand the molecular networks and complex biosystems that create phenotype from genotype.

The TMGC includes Vanderbilt Medical Center, Meharry Medical Center, Univ. of Tennessee, East Tennessee State Univ., Univ. of Memphis, ORNL, and St. Jude's Children's Research Hospital.

Research sponsored by the National Institutes of Health and, in part, by the US Department of Energy under contract DE-AC05-00OR22725 with UT-Battelle, LLC