

The 13th Annual International Mammalian Genome Society Conference: A Meeting Report

Leslie Sprunger,¹ Kent Hunter²

¹Department of Genetics, University of Michigan, 2806 Med Sci II, Ann Arbor, Michigan 48375, USA

²Laboratory of Population Genetics, National Cancer Institute, 41/D702, 41 Library Drive, Bethesda, Maryland 20892, USA

Received: 11 January 2000 / Accepted: 17 February 2000

The 13th International Mouse Genome Conference convened in the historic city of Philadelphia, Pennsylvania the first few days of November 1999. The meeting was honored by a number of distinguished speakers, including three NIH Institute Directors and the Directors of several Genome Sequencing Centers. With the rapid progress in the sequencing of the human genome and the beginnings of what we hope will be similarly rapid progress for the mouse, the single most often repeated phrase at this year's meeting was "the value of comparative sequence analysis." The tremendous potential of comparative sequence analysis for identification of novel genes, evolutionary analysis, facilitation of discovery of the molecular basis of mutations, and as a driving force for more sophisticated experimental biology was clearly and repeatedly demonstrated during the three full days of talks, poster sessions, and informal discussions of the over 400 attendees.

The meeting's first speaker was Francis Collins, the Director of the National Human Genome Research Institute. Dr. Collins spoke about the current status and future goals of the genome sequencing efforts. It is expected that approximately 90% of the draft sequence of human genome will be completed in the spring of 2000, with the complete finished sequence available sometime in 2002 or 2003. The sequence of human Chromosome (Chr) 22 was published during the writing of this review (Dunham et al. 1999)—the first mammalian chromosome to be completely sequenced. The status of the human sequencing project can be found at www.ncbi.nlm.nih.gov/genome/seq.

Dr. Collins and James Battey, the Director of the National Institute on Deafness and other Communication Disorders, speaking later in the conference, outlined the goals of the Trans-NIH Mouse Initiative (www.nih.gov/science/models/mouse). One of the goals of this initiative is to complete the sequence of the mouse genome, using the RPCI-23 C57BL/6J BAC library as a template. The sequencing centers will have the capacity to sequence approximately 500 BACs in the first year. Interested community members will have the opportunity to submit BAC clones of interest for the sequencing consortium (www.ncbi.nlm.nih.gov/genome/clone), and priority will be assigned by peer review. Information about the mouse sequencing project can be found at www.nhgri.nih.gov/NEWS/MouseRelease.htm and www.ncbi.nlm.nih.gov/genome/seq/MmHome.html.

The University of Pennsylvania Symposium on Genome Sequencing and Comparative Sequence Analysis included Directors of five Sequencing Centers discussing the status of their sequencing efforts and applications of comparative genome sequence analysis. Gerald Rubin of the University of California, Berkeley reported that over 26 megabase of the fruit fly have been completed, a total of 104 megabase sequenced, and 2.9 megabase

of the *Adh* region completely sequenced and annotated. Information about the fly sequencing project can be obtained from www.fruitfly.org. Elbert Branscomb of the Joint Genome Institute demonstrated the use of comparative sequence analysis to probe the evolution of gene clusters. He also raised the question of how to utilize the large-scale sequencing infrastructure after human and mouse, citing an estimated cost of \$70 million per vertebrate species for high-quality draft sequence and proposed the sequencing of one member of each branch of the eutherian animals. Eric Lander of the MIT/Whitehead Center for Genome Research presented the results of an integrated genetic, physical, and radiation hybrid simple sequence length polymorphism (SSLP) map of the mouse genome, as well as the development of a set of chromosome-substituted lines by using strains A/J and C57BL/6 for mapping polygenic traits. Dr. Lander noted the discovery, in the course of developing these consomic strains with Joe Nadeau, that A/J mice eat dixie cups, whereas B6 mice do not. (The genetic basis of this phenomenon has yet to be determined.) Steve Brown of the UK Mouse Genome Centre announced the goal of completing 50 Mb of mouse sequence by 2002, including the *Bpa-Str* region on the X Chr. Dr. Brown also demonstrated that comparative sequence analysis revealed that large conserved linkage groups can actually be patchworks of smaller regions and that comparative analysis can reveal important evolutionarily conserved regions, likely to be gene or chromosomal control regions. The final speaker of the symposium was Leroy Hood, from the University of Washington. Comparative sequence analysis of the mouse and human T cell receptor loci from Dr. Hood's group also demonstrated variation of the number of related genes between mouse and human T cell receptor gene clusters.

Efforts by the mouse community itself will also provide a great deal of new resources and information over the next few years, as the presentations in the **Physical Mapping and Sequencing** session demonstrated. The European Consortium plans to map 8000–10,000 mouse ESTs on the radiation hybrid panels over the next two years and has already located 2000 (P. Avner; http://www.genoscope.cns.fr/externe/English/Projets/Projet_ZZZ/rhmap.html). William van Etten from the Whitehead/MIT Center for Genome Research (W/MCGR) reported on the status of their project to map 13,000 ESTs on the T31 Radiation Hybrid panel. At present 2486 loci have been mapped by W/MCGR, of which 2339 are SSLPs. Primers for more than 3400 Unigene clusters have been designed and pre-screened and are ready for analysis. Members of Riken in Japan have been isolating tens of thousands of full-length cDNAs with their cap-trapper technique and have performed more than 24,000 sequence reads to date (<http://genome.rtc.riken.go.jp>).

Large megabase contigs ready for sequencing or partially sequenced have been developed in a number of laboratories (R. Reeves, M. Platzer, M. Bucan), spanning more than 15 megabases

of mouse sequence. E. Green's laboratory is using the opposite approach. Rather than starting with a particular region in the mouse genome, James Thomas described the use of human sequence information from human Chr 7 to select specific orthologous mouse clones for comparative studies. To date more than 1750 clones have been identified from mouse libraries and assembled into more than 50 contigs for further analysis.

This session also hosted two talks about the next generation of genetic markers, single nucleotide polymorphisms (SNPs). Kirsten Lindblad presented the results of the WMCGR's efforts to develop a more efficient alternative to SSLP analysis. On the basis of fluorescent detection of single-base extension assays, Lindblad described the creation of 2848 multiplex loci for mouse genotyping. David Housman presented an alternative methodology based on hybridization. Using IRS-PCR to reduce genome complexity and allele-specific hybridization on microarrays of mouse DNAs, Housman announced the generation of more than 150 single nucleotide polymorphisms and their polymorphism rates in 96 strains of mice.

The **Chromosome Structure and Epigenetics** session was dominated by imprinting studies. Philip Avner led off with a presentation about the importance of the 3'UTR region of *Xist* for X Chr inactivation and demonstrated inactivation of a YAC transgene carrying *Xist* in male ES cells. Robert Nicholls spoke about the contiguous gene syndrome, Praeder-Willi, and the identification of one of the imprinted region boundaries as well as the location of an imprinting center in the 5' end of the first known mammalian bicistronic gene, *Snurf-SmN*. Tamara Davis presented evidence that methylation of the maternal and paternal *H19* alleles is acquired at different times during male germ cell development, indicating that the distinctions between maternal and paternal alleles are somehow maintained in spite of loss of methylation during gametogenesis. Shirley Tilghman continued the analysis of the *H19* locus with a series of elegant genetic studies. Using the Cre-lox system to induce specific chromosome translocations, Dr. Tilghman's laboratory demonstrated that a gene normally clustered in the *H19* imprinting region maintained its imprinted state despite being separated from the *H19* imprinting center. Unlike the Praeder-Willi region, the *H19* region genes therefore must have individual imprinting centers, raising the question of why the *H19* region imprinted genes are clustered.

Steve Hyman, Director of the National Institute of Mental Health, led off the **New Tools in Gene Mapping and Complex Genetics** session. Dr. Hyman noted that many behavioral disorders are probably due to interaction of environment and heredity, and that the genetic factors are likely to be epistatic in nature. The importance of mental health research was highlighted by the fact that of the top ten causes of disability in the United States, four are mental illnesses. The NIMH is therefore making a commitment in mouse genetics for better behavioral phenotyping to reduce variability between laboratories, supporting mutagenesis projects to develop new behavioral mutants for study, increasing the number and quality of transgenic and knock-out animal models, and promoting interdisciplinary training for scientists in behavior research. NIMH is also initiating the Brain Molecular Anatomy Project (BMAP) to increase the knowledge of the molecular architecture of the central nervous system. Dr. Hyman is optimistic that the NIMH commitment will help address the needs of the behavior research community by developing tighter control of gene expression in inducible systems, defining well-characterized reference strains for paradigm development and replication, as well as robust behavioral paradigms, and integrated analyses of genes, behavior, and the brain as mediator of genotype to phenotype.

The commitment for the development of better-characterized resources for mouse genetic research was re-iterated by Ken Paigen, the Director of The Jackson Laboratory. Dr. Paigen described a community effort that has been established to provide detailed phenotypic characterization of 48 inbred strains. Physiological

measurements will be performed in age-matched animals in at least two labs to establish reproducible baseline data. The 48 strains selected are in widespread use, are easy to handle and breed, and are genetically diverse to provide as much heterogeneity as possible for genotype-phenotype correlation. The data will incorporate as much old data as possible and will be accessible in tabular form through MGD. Raw data will be archived for future analysis if necessary. Dr. Paigen stressed that the Strain Characterization Project is a community effort and encouraged active participation by interested community members.

The six other speakers in the session presented new tools and strategies for dissecting complex genetic problems. The intricacy of complex traits was emphasized by the demonstration of inter-allelic epistatic interactions causing phenotypic variation or preventing generation of congenic strains, and multiple linked loci contributing to phenotypic variance. More optimistically, candidate gene analysis has identified a gene that may play a role in obesity, neuropeptide Y (*Npy*). NPY is thought to stimulate feeding and lower metabolism, and Ben Taylor and colleagues have identified a single base deletion in a highly conserved element in the 3'UTR. Although the analysis of this gene is not yet complete, the working hypothesis is that this deletion may disrupt normal mRNA regulation, resulting in overexpression of NPY. Dr. Joe Nadeau demonstrated an association of the hedgehog signaling pathways with the homocysteine biosynthetic pathway and showed how small, custom-made microarrays could detect subtle expression differences in specific metabolic pathways. Dr. David Beier described the application of interval mapping of complex traits to increase the efficiency of QTL identification.

The **Gene Discovery** session featured old friends as well as new mutations. Bernhard Herrmann presented a significant advance in our understanding of transmission ratio distortion in the identification of the *t*-complex responder locus as a member of a novel protein kinase gene family that affect sperm motility, dubbed *Smok*. Mary Lyon's earlier genetic analysis of the *t*-complex led her to formulate a model of differential effects of distorter loci on the wild-type vs. mutant responder, leading to distortion in favor of inheritance of the *t*-haplotype in heterozygotes, while homozygotes are sterile (Lyon 1986). Identification of the mode of inheritance of *Smok* mutants provides direct experimental support of this model. Additionally, Herrmann and colleagues showed that *Smok^{Tcr}* as a transgene can promote transmission ratio distortion for other chromosomes into which it is integrated. This constitutes a potential means of manipulating the inheritance of specific traits, which may be particularly useful in agricultural applications (Herrmann et al. 1999).

A number of other old friends were discussed in this session, including shaker with syndactylism (*sy*) (K. Steel), gunmetal (S. Kingsmore), pallid (L. Huang), and flailer (J. Jones). Shaker with syndactylism is a complex phenotype. The syndactylism is due to a contiguous gene syndrome. A mutation in an Na,K,2-Cl-cotransporter, *Slc12a2* (aka *Nkcc2* or *mBSC2*), results in deafness due to abnormal endolymph formation, and is deleted in the original radiation-induced *sy* mutant. Gunmetal is a splice site mutation in the Rab geranylgeranyl transferase gene (*Rabgga*), which disrupts normal platelet production and function. Pallid, a platelet storage pool deficiency mutation, is a nonsense mutation in a previously unknown gene. The novel protein, pallidin, was shown to interact with syntaxin-13, implicating it in vesicle docking and fusion. Flailer was found to be a fusion of *Myo5a* and an adjacent gene, *Gnb5*. The mutation, probably owing to unequal crossing over, fuses the 5' end of *Gnb5* to the 3' end of *Myo5a*, disrupting transport of organelles by a dose-dependent competition mechanism.

New mutations include Long Bones (*LOBO*) and *Lvis*'s. *LOBO* was induced by a transgene insertion, resulting in dominant abnormal bone growth (A. Rump). The transgene insertion has disrupted the transcription of a 3.8-kb gene with high homology to the

yeast heterotrimeric G-protein *Dis3*. *Dis3* maps to 2q37 in humans, and makes *LOBO* a strong candidate for Albright hereditary osteodystrophy. *Lvis*'s are genes associated with retroviral insertions in B-cell leukemias (M. Justice). Using a PCR-based technique called VISA, the Justice laboratory has identified a number of common insertions in spontaneous B-cell tumors in AKXD recombinant inbred mice and developed a viable tumor bank that, in combination with the VISA gene discovery technique, can be used to further probe the role of novel genes in oncogenesis.

ENU mutagenesis screens, chips, and knockouts were the order of the day in the **Functional Genomics** session. Mary Bedell described efforts to characterize ENU and spontaneous mutations of the Steel locus, and how the analysis of these mutants would lead to a better understanding of the in vivo structure-function relationships of mast cell growth factor. Yuan Zhuang used ENU mutagenesis to investigate the role of the E2A pathway in lymphocyte development. A mutation in the gene *Moel* was found to abrogate development of mature lymphocytes and will lead to greater understanding of the role of the bHLH transcription factors in the immune system. Larger-scale mutagenesis projects included searches for mutations affecting eye or ear development. Jack Favor described the results of screening over 500,000 mutagenized mice for cataract development; 84 mutants have been identified to date, mapping to 22 loci, including 5 independent mutations mapping to the connexin 50 locus. Rachel Hardisty reported the results of a 10,000 mouse screen for deafness mutants. To date, nine mutants have been identified with vestibular defects, two with hearing defects, and one with vestibular and hearing defects. Five of the nine vestibular mutants map near the Wheels locus, a semi-dominant mutation with complex behavior abnormalities, and may represent new alleles. In addition, several potential novel *Shl* mutants are currently being tested.

Arthur Sands of Lexicon Genetics announced the current status of the gene-trapped ES cell bank, Omnibank. Using a retroviral vector to tag genes, Lexicon Genetics has developed a bank of over 17,000 trapped genes. Sequence analysis of the regions flanking the retroviral insert indicated that 50% of the trapped sequence is not currently available in the public databases. Interested researchers can search through Omnibank for their genes of interest at www.lexgen.com.

Three presentations dealt with gene chip technology. Tetsuya Tanaka described the efforts of the NIA to develop a gene chip for 15,275 unique genes expressed in embryonic tissues. cDNAs were isolated from pre-implantation, E7.5 embryonic and extra-embryonic, E12.5 mesonephros tissues, as well as newborn ovary. 3331 (22%) corresponded to known genes, 3526 (23%) were related to known genes, and 8418 (55%) were novel genes. Information about the NIA chip can be obtained at lgsun.grc.nia.nih.gov. The Riken group (genome.rtc.riken.go.jp) has also developed a high-complexity chip (Y. Okazaki). 20,000 full-length cDNAs selected from both adult and embryonic tissues have been used to develop a chip and produce a database of the spatial and temporal expression of these genes in mice. On the other end of the scale, the NHGRI has been working with smaller scale chips to interrogate specific pathways. Using tissue-specific ESTs on their custom arrays, Stacie Loftus and co-workers have demonstrated specific differential expression of neural crest genes, which could not be detected on arrays without tissue-specific bias.

This year's **Verne Chapman Memorial Lecture** was delivered by Janet Rossant, at the start of the session on **Developmental Genetics**. Dr. Rossant discussed her work on fibroblast growth factor receptor signaling pathways in embryogenesis. The function of *FGFR1* as a posteriorizing signal in the antero-posterior development pathway was established by using mutagenesis, reporter constructs, expression arrays, novel imaging techniques, and careful morphologic analysis. Additionally, *FGFR2* and *FGFR4* have been shown to be involved in the required signal from the embryo to continue trophoblast proliferation. Dr. Rossant's talk also served

to re-emphasize the value of chimeras for study of cell autonomy and interaction, and the historical complementarity of genetics and developmental biology.

Another highlight of this session was Colin Bishop's presentation of the role of *Sox9* in sex determination. XX mice with a mutation upstream of the *Sox9* gene developed as phenotypic males, despite the absence of *Sry*, pointing to *Sox9* as the controller of Sertoli cell development and consistent with the model that depression of *Sox9* results in phenotypic maleness. The value of an allelic series was demonstrated in the presentation of a collection of engineered *Pitx2* mutants with varying levels of retained function of the gene product (D. Martin). Phenotypic variation in these mice mimics some of the heterogeneity of human patients with Reiger syndrome owing to mutations in *PITX2*, indicating that the mice should be useful for identification of modifier genes. Cerebellar development was the topic of two talks, with Kathy Hentges reporting on an ENU-induced mutant called Flat-top, and Thomas Vogt presenting the effects of knockouts of *Hbnf* and *Mdk*. Loss of function of either of these genes singly results in mild effects, but the double knockout showed profound effects on cerebellar foliation and greatly reduced proliferation of granule cells. Elizabeth Lacy reported on progress toward characterizing a likely candidate gene for amnionless. The candidacy of a gene disrupted by the transgene insertion was bolstered by a successful BAC rescue experiment. The function of the gene is unknown, but it shows appropriate temporal and spatial expression for an involvement in gastrulation. The use of a conditional deletion mutant to overcome the early embryonic lethality of *Bmpr* knockouts was described (E.B. Crenshaw). Restricting the deletion to brain and spinal cord allows further testing of hypotheses concerning *Bmpr* signaling in CNS development.

For the second year, in addition to platform presentations and traditional poster sessions, the IMGS Conference also featured interactive poster sessions. Select posters from related topics were presented, followed by a discussion of topics of interest related to the focus of the interactive poster session. The three topics of the interactive poster sessions this year were Bioinformatics, QTL mapping, and Mutagenesis.

The **Bioinformatics Interactive Poster Session** provided information about useful tools available for the mouse community and future plans for improving electronic resources. An overview of the Mouse Genome Informatics website (www.informatics.jax.org) was presented (J. Blake) to help users navigate the increasing volume of data and information. A number of new links (the Mouse Tumor Database; Rat Genome Database; Flybase; Mouse Facts Page) and other enhancements of MGI were described (J. Richardson). The Mouse Tumor Database (C. Lutz) provides on-line resources for using the mouse as a model for oncogenesis and includes information about incidence, latency, and pathology of tumors in various inbred strains (tumor.informatics.jax.org). The Rat Genome Database (www.informatics.jax.org/rat/) contains information about polymorphisms and positions of 5256 loci in the rat genome, as well as links to the rat Radiation Hybrid mapping panel (R. Palazola). The Jackson Laboratory mouse interspecific backcrosses and compiled radiation hybrid maps (L. Rowe) currently have over 5000 and 3100 loci mapped, respectively, and an integrated framework map is under construction. The backcross and RH data can be accessed through www.jax.org/resources/documents/cmdata. The new organization of the Mouse Locus Catalog was described (A. Davis), dividing the phenotype and allele database into the five categories of gene family/protein domains; mouse model/disease; gene/gene product information; expression; and phenotype/alleles and variants. As part of the ongoing attempts to improve accessibility and enhance searches for interrelationships of the burgeoning amounts of data, Janan Eppig described efforts to standardize terms used for phenotypic descriptions. Standardization of gene symbols and assignment of genes into gene families is the current focus of the Nomenclature Com-

mittees (L. Maltais). Information about assigning new gene symbols and links to gene family webpages can be accessed at www.gene.ucl.ac.uk/nomenclature/ or www.informatics.jax.org/support/nomen/.

A number of new or improved informatics tools were also described. A prototype program, Genome Spatial Information System or GenoSIS, has been developed for the graphical representation of genomic sequence features, including genes, promoters, repeat sequences, etc. (C Bult). A new version of the mapping program Map Manager QTX (K. Manley) has been under development to facilitate genetic analysis of classical and quantitative traits on PC platforms as well as Macintosh. The MRC UK Mouse Genome Centre at Harwell (R. Selley) has developed a database for tracking animal husbandry, mutagenesis, and behavioral analysis of the UK mouse mutagenesis project. The latest information about this project is available at www.mgu.har.mrc.ac.uk/mutabase/. Two different EST mapping tools were also described. The European Bioinformatics Institute developed a set of tools for the alignment and clustering of ESTs, selection of ESTs for mapping, management of the RH database, and a public repository of the data (www.ebi.ac.uk). The MRC UK Mouse Genome Centre has developed EMPATH, a laboratory management system developed from the earlier HOSEPIPE system, and based on web-accessible relational databases. A central repository of annotated and curated genome sequence is being maintained at the European Bioinformatics Institute in collaboration with the DNA Data Bank of Japan and Genbank. Sequence data are deposited in the public database as soon as they are available from the sequencing centers. Individuals can submit data by contacting datasub@ebi.ac.uk. The EBI is also maintaining a genome monitoring table (Genome MOT) to enable community members to track the amount of sequence deposited from a number of organisms (www.ebi.ac.uk/~sterk/genome-MOT/). Finally, to integrate the enormous amount of sequence data that will be generated, The Jackson Laboratory MGI Group is developing a data integration system to connect genome sequence features to function, expression, homology, mapping, and phenotypic data present in MGD and GXD.

In the **QTL Interactive Poster Session**, the value of congenics, consomics, and progeny testing for localization of QTLs was stressed, as well as the value of multiple crosses to evaluate a single trait. Extensive phenotyping was also suggested as a means to gain additional mileage from a single cross by mapping modifiers of multiple phenotypes. The utility of very large crosses was debated: although they permit detection of multiple epistatic interactions, the general consensus was that once a QTL has been localized, other techniques such as recombinant congenic strains are more efficient.

Taken together, the posters presented in the **Mutagenesis Interactive Poster Session** amount to the beginnings of a remarkable resource for virtually every subdiscipline of biology. A number of laboratories reported on screens for neurologic and/or behavioral mutants. Brian Libby presented results from a screen for gametogenesis mutants, an example of an important biological issue that requires more specific screening methods. Yijing Chen and John Schimenti reported on two projects to conduct ENU and EMS mutagenesis in embryonic stem cells rather than in male mice as a means of increasing efficiency and reducing animal costs. Updates were provided on the large scale mutagenesis projects at MRC Harwell and GSF Munich, and the Frozen Embryo and Sperm Archive (FESA) at MRC Harwell, an important adjunct to the large-scale production of new mutants. In an effort to increase the number of mouse models of human deafness, a deafness screen has been added to the Harwell and GSF projects, and a dozen new mutants have been isolated. This approach of incorporating phenotype-specific screens into existing programs may prove to be a way to increase the feasibility of mutant screens for smaller laboratories.

Involvement of smaller laboratories in mutagenesis projects was encouraged, so that phenotype-specific screens are conducted by groups with a high level of expertise and interest in a particular physiologic system. The consensus was that such smaller scale screens would be greatly facilitated by the availability of ENU-treated mice from larger centers. In addition, there is a perceived need to persuade funding agencies that small screens are feasible and necessary in addition to the larger screening projects. Finally, the importance of efficient information management in both small and large-scale screens was emphasized along with the potential value of software distribution by the large centers.

Overall, this conference reflected momentum and excitement in the prospects for progress in mouse genetics and genomics. The 14th International Mouse Genome Conference will be held Nov. 6–10, 2000, in Narita, Japan (genome.rtc.riken.go.jp/IMG2000.html). Individuals interested in membership in the International Mammalian Genome Society can visit the IMGS website at <http://img.org>. Abstracts for the meeting are also available.

References

- Dunham I, Shimizu N, Roe BA, Chissole S, Hunt AR et al. (1999) The DNA sequence of human chromosome 22. *Nature* 402, 489–495
- Herrmann BG, Koschorz B, Wertz K, McLaughlin KJ, Kispert A (1999) A protein kinase encoded by the *t complex responder* gene causes non-mendelian inheritance. *Nature* 402, 141–146
- Lyon MF (1986) Male sterility of the mouse *t*-complex is due to homozygosity of the *Distorter* genes. *Cell* 44, 357–363