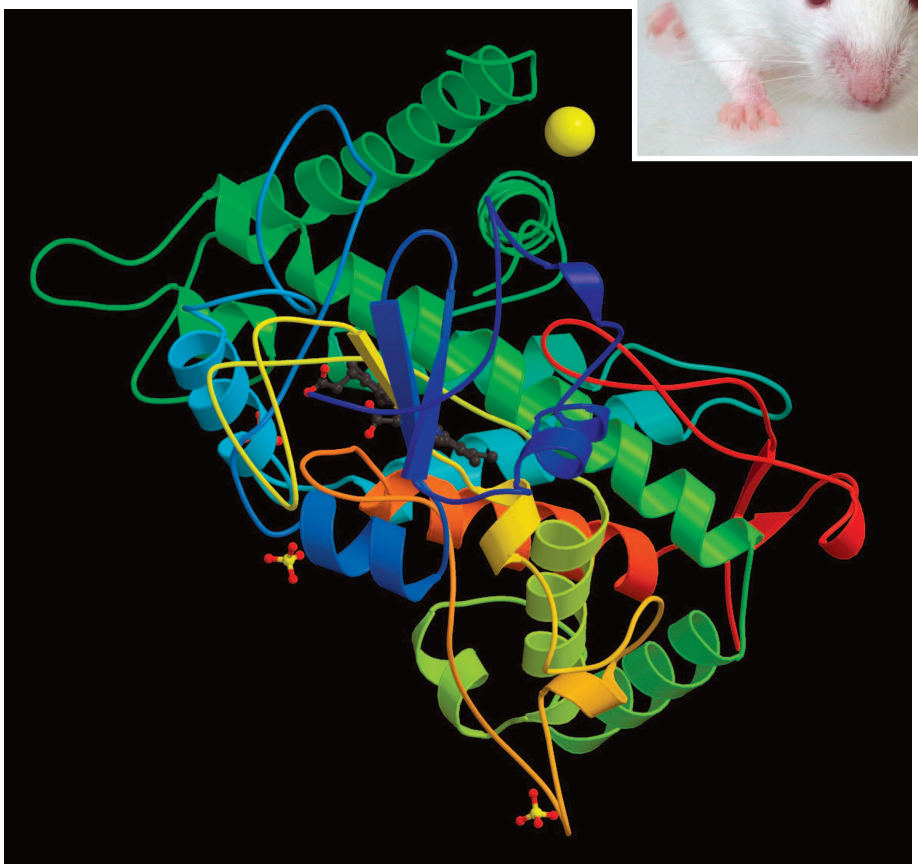
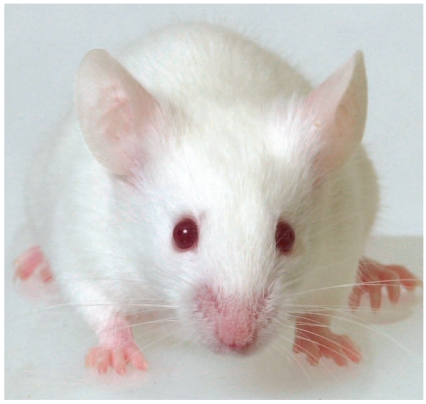
Sequencing *CYP*

More Data Allows Cross-Species Comparison

Although important elements of the human and mouse genome projects have been published and a sense of completion prevails, there remains much detail work to be done on both projects. This month, a team led by Haoyi Wang of Miami University in Oxford, Ohio, elegantly dissects the conservation of one cluster of highly duplicated genes across the two species. The team simultaneously resolves the sequence of one of the many remaining ambiguous regions of the mouse genome, finds new insights into evolution, and provides new tools for future work aimed at understanding how the cytochrome P450 (CYP) system works in the mouse model system, which is used heavily in research related to human health [***EHP* 111:1835–1842**].

Filling in the genomes' finest details, especially in regions with only slightly differentiated copies of the same information, will take far longer than the initial large-scale sequencing projects did. But it's a necessary task: in active clusters of duplicated genes, where repeated near-copies of the same sequence in the same region of DNA encode functional products, the genes are so similar that it is hard to be sure that the sequenced DNA has been assembled in the right order. Like putting together puzzle pieces of nearly the same shape and pattern, sequence assembly in repetitive regions can easily go astray. Furthermore, not knowing the exact sequence in such a region can confound efforts to reliably identify, knock out, or clone genes of interest for further study with biochemistry and genetics.





Of mice and men. New research maps the structure of the mouse *Cyp2* gene cluster that encodes the P450 enzymes responsible for metabolizing many xenobiotics. These new data will enable scientists to make cross-species comparisons to the previously sequenced human *CYP2* analog.

The *CYP* superfamily of genes is conserved across kingdoms, from the bacteria, plants, and fungi to the higher eukaryotes. It encodes a wide range of heme-thiolate monooxygenases, enzymes that carry a molecule of the pigment heme. Some of these enzymes catalyze specific oxidation reactions that detoxify a panoply of environmental compounds and drugs. Members of the *CYP* family are also involved in the metabolism of eicosanoids, steroids, and fatty acids.

Because the superfamily's conserved genes are also heavily duplicated within a given species, they provide an excellent system for examining how species, given the basic building blocks of the gene family, copy and rearrange the genes and thereby accumulate a set of enzymes appropriate for their own needs. The diversity is impressive. When sequencing is combined with other information, mice are shown to have 189 *CYP* genes and pseudogenes, and humans to have 115.

In 2001, principal investigator Susan M.G. Hoffman published the structure of a cluster of 13 genes from the *CYP2* family in humans. This month in *EHP*, she, Wang, and colleagues lay out the pattern of the corresponding cluster in mice, and they are working on similar analyses in other species. The *CYP2* genes are a substantial group, the largest P450 family in the mammals, with 15 subfamilies containing dozens of genes so far identified. Biochemical assays have shown that the CYP2 enzymes, together, are able to metabolize many chemicals, including more than half of all frequently prescribed drugs, as well as some steroids and arachidonic acid, the major precursor of several classes of signal molecules including the prostaglandins.

Wang and Hoffman's group combined work at the computer and at the bench to accurately assemble the mouse *Cyp2* cluster, located on chromosome 7, and to compare it with the *CYP2* cluster on human chromosome 19. Their new assembly solves the published sequence's gaps within the cluster, allowing them to identify new genes that had been fragmented or left out in earlier assemblies. At the same time, they have used cloned genomic DNA to painstakingly patch together an improved fine-structure physical map of the region.

Together, the reassembly and the new map have generated a more accurate picture of genes in the cluster. Coupled with the group's similar work with the related human *CYP2* gene cluster, the mouse mapping data show how—although not why—the two species have expanded their *CYP2* genes in the time since evolution separated them. The authors write, "This comparison should enable researchers to better utilize the mouse as a model system for the study of these *CYP* genes in humans and in other mammals." **–Victoria McGovern**

Top to bottom: Arnold Greenwell/*EHP*; Protein Data Bank