Mice are not men

A vibrant discussion of the merits and limitations of animal models is long overdue. The limitation of space precludes addressing many of the questionable approaches and statements by Takao and Miyakawa (1).

Despite the different approaches used by Takao and Miyakawa (1), their results actually support the conclusion that "Genomic responses in mouse models poorly mimic human inflammatory diseases" (2). This can be best understood using the example provided by the authors: "13,586 and 3,116 genes are changed ($P < 0.05$ and fold-change > 1.2) in human burn conditions and mouse models of infection, respectively, and 1,992 among them are commonly changed in both humans and mice, among which 1,608 genes changed to the same direction (Fig. 2*E*)" (1). Therefore, among the 13,586 genes changed in the human disease, the model can reflect only 1,608 (or 12%) correctly to some extent. The appropriate conclusion should be that genomic responses in current models mimic roughly 12% of the genes in human inflammatory diseases. No matter the branch of science, this result indicates a poor model because it leaves close to 90% of the genes in the human disease not mimicked in the model.

We reported both Pearson's and Spearman's rank correlations of the human genes (false discovery rate < 0.001 and fold-change > 2), and the results show that the responses to human trauma and burns are highly correlated, whereas the murine models correlate poorly with the human conditions (figure 1 and supplemental figure 1 in ref. 2). Because Takao and Miyakawa restricted their analyses to the preselected 15% of genes that significantly changed in both mice and humans, it is not surprising that under such conditions, the overall correlations would be stronger, as shown in their figures 1 and 2 (1). In fact, the data in their figure 3 (1) corroborates our findings that the correlations between gene expression in humans (blue squares) with different injuries were dramatically better ($R = 0.8$ to <1.0) than between humans and mice genes (human: fold-change > 2.0; mouse model: fold change > 1.2) (blue circles), where the *R*s were from 0.15 to 0.30, which are comparable to the values we reported for R^2 (0.0–0.1). The authors used filters and approaches that selected for and maximized the similarities, as opposed to performing a genome-wide analysis for the human diseases.

Because drugs function at the molecular level, *P* values on the NextBio enrichments of the preselected small portion of genes over biogroups, such as "innate immune response" (figure 4*A*) (1), do not help the prioritization of therapeutic candidates.

The question that the science community should be asking is whether the appropriate measure of an animal model should be limited to highly selected genes that in this case retrospectively reflect the most similar common responses (in our view, a tautology) or whether the appropriate comparison is how all genes behave. If one limited the analysis to common entities and tried to understand a station wagon by studying a motorcycle, one would learn something about wheels and spark plugs but have no idea about steering wheels, airbags, and sunroofs, and the larger picture would be substantially missed. We agree much has been and can be learned from mouse models, in particular as to how specific gene manipulation may alter pathways and phenotype. However, the data we generated would suggest that the extent that such changes in specific genes or pathways reflect importance in human inflammatory disease should probably be verified in each case.

H. Shaw Warren[a], Ronald G. Tompkins[a], Lyle L. Moldawer[b], Junhee Seok[c], Weihong Xu[d], Michael N. Mindrinos[d], Ronald V. Maier[e], Wenzhong Xiao[a,1], and Ronald W. Davis[d,1]

[a]Departments of Surgery, Medicine, and Pediatrics, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114; [b]Department of Surgery, University of Florida, Gainesville, FL 32608; [c]School of Electrical Engineering, Korea University, Seoul 136-713, Korea; [d]Stanford Genome Technology Center, Stanford University, Palo Alto, CA 94305; and [e]Department of Surgery, Harborview Medical Center, Seattle, WA 98195

1 Takao K, Miyakawa T (2015) Genomic responses in mouse models greatly mimic human inflammatory diseases. *Proc Natl Acad Sci USA* 112:1167–1172.

2 Seok J, et al.; Inflammation and Host Response to Injury, Large Scale Collaborative Research Program (2013) Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proc Natl Acad Sci USA* 110(9):3507–3512.

Author contributions: H.S.W., R.G.T., L.L.M., J.S., W. Xu, M.N.M., R.V.M., W. Xiao, and R.W.D. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence may be addressed. Email: wenzhong.xiao@mgh.harvard.edu or jeannet8@stanford.edu.