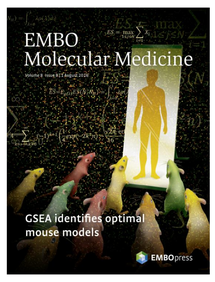
# Defining the optimal animal model for translational research using gene set enrichment analysis

[1] Department of Experimental Toxicology and ZEBET, German Federal Institute for Risk Assessment (BfR), Berlin, Germany.
[2] Department of Clinical Pharmacology and Toxicology, Charité-Universitätsmedizin Berlin, Berlin, Germany.

**Christopher Weidner[1], Matthias Steinfath[1], Elisa Opitz[1], Michael Oelgeschläger[1] and Gilbert Schönfelder[1,2]**




The mouse is the main model organism used to study the functions of human genes because most biological processes in the mouse are highly conserved in humans. Recent reports that compared identical transcriptomic datasets of human inflammatory diseases with datasets from mouse models using traditional gene-to-gene comparison techniques resulted in contradictory conclusions regarding the relevance of animal models for translational research (*Takao et. al* vs*. Seok et al*.). To reduce susceptibility to biased interpretation, all genes of interest for the biological question under investigation should be considered. Thus, standardized approaches for systematic data analysis are needed.

We analyzed the same datasets using gene set enrichment analysis (GSEA) focusing on pathways assigned to inflammatory processes in either humans or mice. The analyses revealed a moderate overlap between all human and mouse datasets, with average positive and negative predictive values of 48 and 57% significant correlations. Subgroups of the septic mouse models (i.e., *Staphylococcus aureus* injection) correlated very well with most human studies. These findings support the applicability of these targeted strategies to identify the optimal animal model and protocol to improve the success of translational research and to reduce the number of animal studies required.

## Results

### Correlation of inflammatory pathway regulation: human–human, mouse–mouse, and human–mouse comparisons

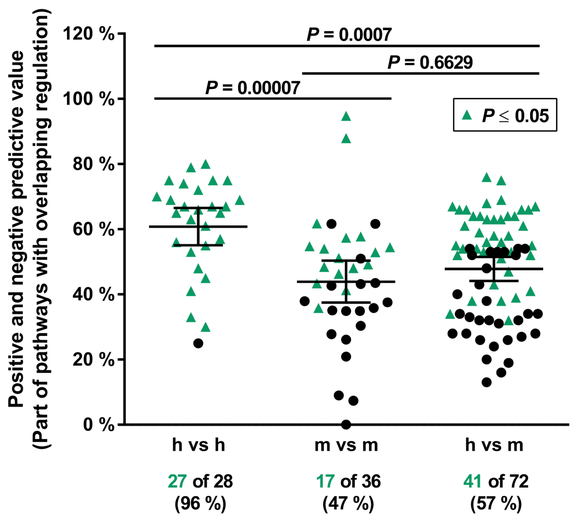



**Fig. 1.** The regulation of inflammatory pathways was assessed by gene set enrichment analysis (GSEA) using unfiltered gene expression data from 8 human and 9 mouse studies. Significantly regulated pathways were compared between two datasets from human (h) and/or mouse (m) studies. The degree of pathway overlap is depicted as the mean predictive values between these two datasets. Studies that revealed pathway overlap significantly greater than that expected by chance ($P \leq 0.05$) are labeled with green triangles. The numbers of significantly correlated studies are given below the datasets. Lines indicate the mean ± 95% confidence interval. The *P*-values for each pair of datasets were calculated using a chi-squared test, and the *P*-values for the comparison of species effects were calculated using the Kruskal–Wallis test followed by Dunn's multiple comparisons test and Bonferroni correction *(1)*.

### Certain mouse models correlated very well with most human studies

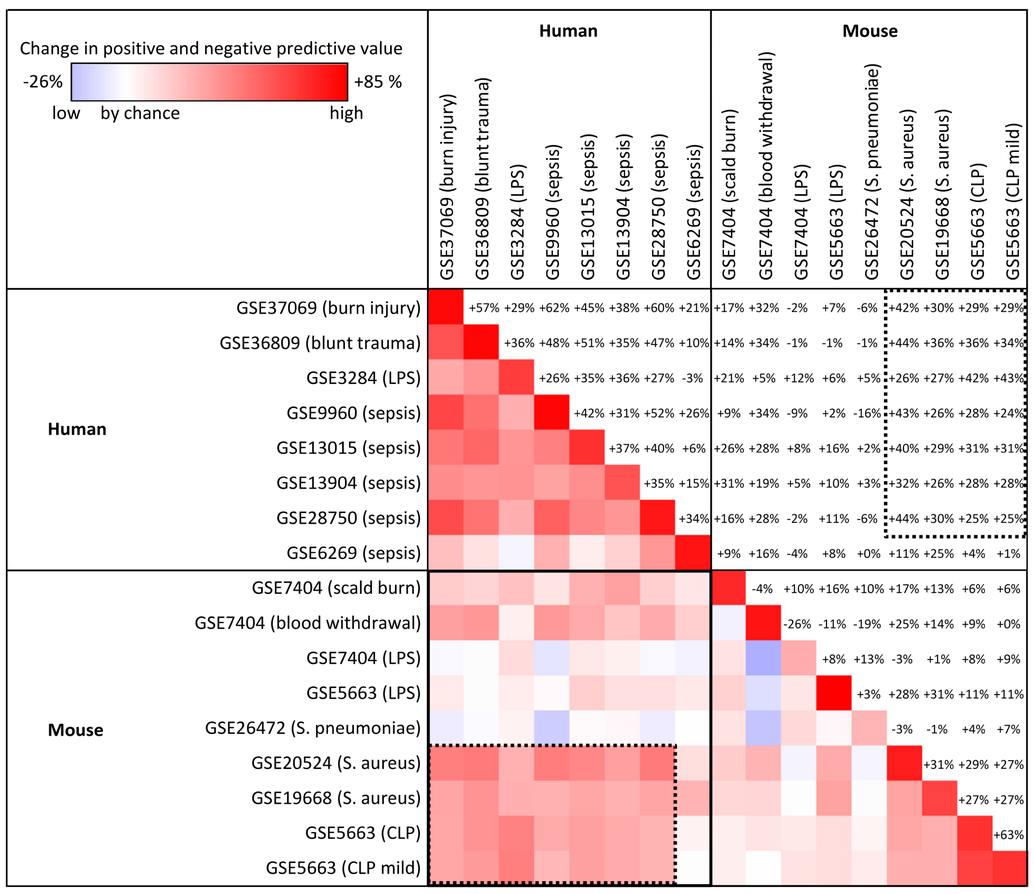



**Fig. 2. Correlation matrix of pathway comparisons among human and mouse inflammatory studies.** The overlap of pathway regulation is shown as the average change in the positive and negative predictive value over expectation by chance (blue, decrease, low correlation; red, increase, high correlation). The comparison of human with murine datasets revealed a subgroup of experimental murine models that were highly correlative to human clinical studies (dotted line), consisting of the *Staphylococcus aureus* injection and the cecal ligation and puncture (CLP) models. In contrast, lipopolysaccharide (LPS) gavage and intratracheal infection with *Streptococcus pneumoniae* showed no correlation to human inflammatory diseases *(1)*.

## Conclusion

A standardized, unbiased approach identifies the optimal animal model:

- Pathway-based gene set enrichment analysis (GSEA) overcome the limitations of biased single-gene analysis by avoiding the subjective setting of gene expression thresholds and gene filtering
- Parts of genomic responses in mouse models mimic human inflammatory diseases
- Here, infection models with *Staphylococcus aureus* injection or cecal ligation and puncture (CLP) correlated very well with most human studies (as opposed to LPS gavage or intratracheal *Streptococcus pneumoniae* administration)

➢ Applicability of targeted analysis strategies to define the optimal animal model and treatment protocol for a given human disorder

➢ Potential to improve the success of translational research and to reduce the number of animal studies required

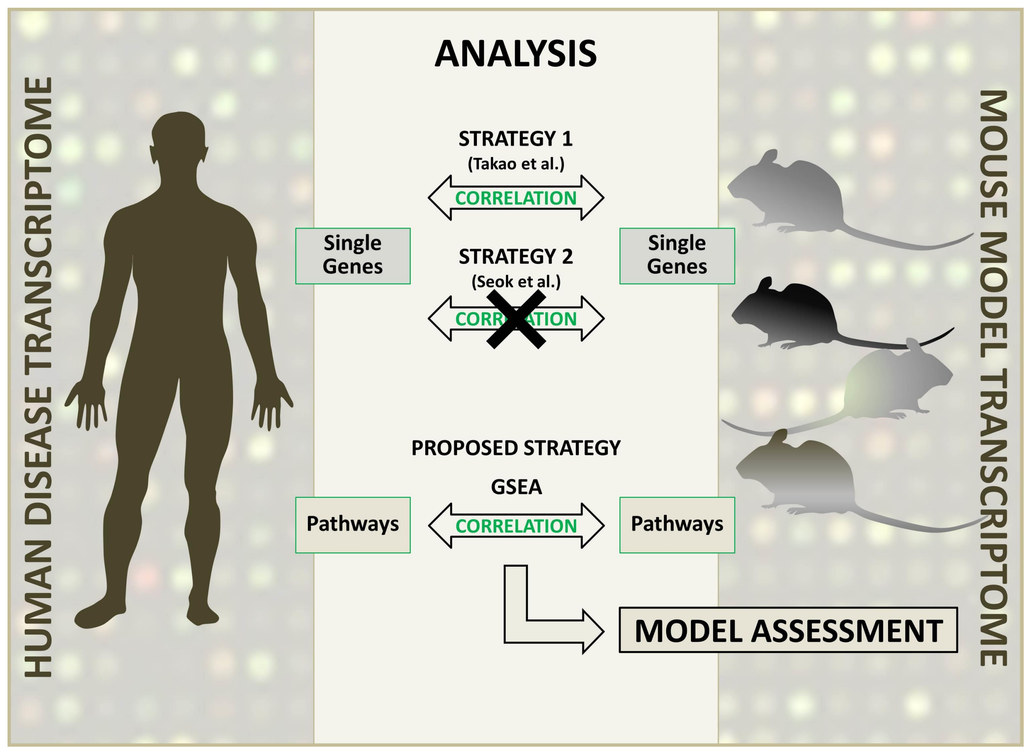



**Fig. 3. Standardized and unbiased approach to identify the optimal animal model** *(1)*. Video version

## References

*(1)* Weidner C, Steinfath M, Opitz E, Oelgeschläger M & Schönfelder G. Defining the optimal animal model for translational research using gene set enrichment analysis. *EMBO Mol Med* 8(8): 831-8 (2016).

*(2)* Takao K & Miyakawa T. Genomic responses in mouse models greatly mimic human inflammatory diseases. *Proc Natl Acad Sci USA* 112(4): 1167–1172 (2015).

*(3)* Seok J *et al*. Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proc Natl Acad Sci USA* 110(9): 3507–3512 (2013).

German Federal Institute for Risk Assessment • Max-Dohrn-Str. 8-10 • 10589 Berlin, GERMANY • Tel. + 49 30 - 184 12 - 0 • Fax + 49 30 - 184 12 - 47 41 • bfr@bfr.bund.de • www.bfr.bund.de