Cellular & Molecular Immunology (2013) 10, 284–285

www.nature.com/cmi

RESEARCH HIGHLIGHT

Of mice and men: what rodent models don't tell us

Adam C Drake

Cellular & Molecular Immunology (2013) **10,** 284–285; doi:10.1038/cmi.2013.21; published online 20 May 2013

Work recently published in *PNAS* by The Inflammation and Host Response to Injury, Large Scale Collaborative Research Program[1] compared the transcriptional responses in peripheral blood to inflammatory injuries (burns, blunt force trauma) and to endotoxin in human patients and in corresponding mouse models. The results were startling: while the human patients showed similar transcriptional responses to burns, trauma and endotoxemia, the mouse models showed little correlation either to each other or the human response. These differences extended beyond acute gene response and were also apparent in markedly slower recovery from injury, with gene expression not returning after weeks or months in patients as opposed to days in mouse models.

The results of this study shed light on an ongoing problem in therapeutic development. Approximately 90% of the drugs reported to have clinical efficacy in high profile journals based on animal models fail in clinical trials.[2] Drugs to treat sepsis and acute inflammation have a much worse track record in clinical trials, with all ~150 drugs trialed to date failing.[1] This report highlights a startling example of how the *in vivo* models used in preclinical drug development can bear little relationship to the human diseases at the transcriptional level. Given the similarity between the human responses to pro-inflammatory stimulation, and the divergent responses in mice it would be easy to conclude that in the ~75 million years[3] since humans and mice shared a common ancestor there has been tremendous divergence in the way in which inflammatory responses are generated and resolved. Indeed, there is strong evidence to support this view in the data mining reported by the authors, who show that systematic regulation of entire pathways associated with inflammatory responses occurs across human inflammation. This coordinated response does not occur across mouse models. The authors then examined other published datasets further strengthening the broad and general nature of their findings.

One of the most striking differences this group observed was in the regulation of the Toll-like receptors. These sensors of pathogen-associated molecular patterns exhibit broad transcriptional upregulation in all human inflammation examined, while the changes in expression levels in mice were variable, notably with very little response to endotoxin exposure. This difference could account for relatively greater sensitivity to endotoxic shock and severe pathology seen in human patients, which is absent from mouse models. The authors report upregulation of Toll-like receptor pathway genes following endotoxin exposure, which should result in increased sensitivity to subsequent pathogen-associated molecular patterns encounter, leading to further sensitization. This could result in both a cytokine storm and long-term alterations in the immune system—two of the major challenges of sepsis.[4]

As well as helping to explain differences in mouse and human responses, the data presented make a compelling case for caution when evaluating mouse models in the development of anti-inflammatory drugs. Medicines act at the molecular level and difference in the transcriptional response between human and mouse could, at the very least, reduce treatment effectiveness in humans. At worst, these differences could lead to unforeseen *in vivo* side effects.

However, there are significant caveats that need to be considered in the context of these results. Firstly, the patients studied were all receiving medical care. This variable combination of changes in activity, diet, medication and environment is not recapitulated in mouse studies and may account for some of the correlation between different human injuries. Secondly, only a single mouse strain, C57BL/6, was assessed and other strains may prove to be better models of human responses at the molecular level. Thirdly, total leukocytes were assessed. The composition of mouse and human blood leukocytes vary significantly—notably with dramatic differences in the circulating levels of neutrophils. These cells are also representative of only a single organ, and changes in gene expression in other tissues such as the liver, kidneys and vascular endothelium are also important for clinical outcome. Finally, this study assesses acute inflammatory conditions and does not speak to the predictive power of mouse models of chronic inflammation or models where inflammation is not a key feature of pathology.

Koch Institute for Integrative Cancer Research at MIT, Cambridge, MA, USA
Correspondence: Dr AC Drake, Koch Institute for Integrative Cancer Research at MIT, 500 Main st 76-253, Cambridge, MA, 02139, USA.
E-mail: adrake@mit.edu
Received: 15 April 2013; Accepted: 20 April 2013

Despite these limitations, the results of this study are extremely important, as commonly used models were assessed and found to be profoundly different to the corresponding human condition. This study is the first to systematically address at the molecular level why mouse models that faithfully anatomically mimic human trauma have lead to no clinically effective therapeutics. While mortality from acute sepsis in hospital has fallen from ~75% to ~25% this is largely due to improvements in supportive care rather than effective medicines.[5] The stark differences in gene expression between patients and mouse models helps to explain why there has been a persistent failure to address this major cause of morbidity.

Going forward this study will provide a general template for future evaluation of preclinical data and a starting point for experimental strategies to further understand inflammation. These points should lead to the development of better models, yielding more relevant preclinical data, leading to new therapeutics. Each of these points is worthy of consideration in greater detail. By highlighting the differences in gene expression with respect to induction/repression, magnitude and duration of response between patients and animal models, the authors have provided a new standard for testing models and interpreting drug efficacy results. While future studies will build on this initial work by refining the fractionation of cells and broadening the tissues examined the basic template has now been defined: regulatory authorities should begin asking for gene expression data validating the relevance of *in vivo* preclinical studies and research funding bodies should consider devoting resources to developing on the datasets presented here to further clarify what particular model systems offer.

The combination of the caveats regarding interpretation and the clear differences in gene expression point to several important directions where new model systems are needed or where clarification is vital to understanding the relevance of preclinical data. Firstly, experiments are needed to assess the effects of supportive care and identify which pathways need to be targeted to compliment existing hospital treatment. Secondly, all rodent strains and other animal models need to be assessed in the same way as the common C57BL/6 mouse model was in this report. Directing drug development to use the best existing models or utilizing less widely used strains could yield substantially better outcomes without extensive development and testing. Thirdly, a more nuanced analysis is needed where fractionated cell subtypes are assessed and the differences in expression phenotype are attributed to specific populations rather than the stoichiometry of the mixture of leukocytes isolated. Finally, the methods laid out in this study need to be replicated for models of other conditions, especially where failure rates in clinical trials are high. For diseases such as cancer, these sorts of efforts are already underway.

Based on the results of Seok *et al.*, it is also clear that there is an opportunity to develop new models for inflammation. While work continues to develop and refine *in vitro* models none envisaged to date can recapture the complexity of a living organism. Since human and mouse leukocytes respond so differently, *in vivo* models where the immune cells more closely resemble those in patients are particularly attractive. For example, humanized mice [6,7] offer many of the advantages of small animal models with the majority of leukocytes being of human origin. However, these cells having developed in a murine environment have their own set of limitations imposed by incompatibilities between graft and host. Primate models offer excellent similarity due to their close evolutionary relationship with humans; this, however, comes at the expense of using outbred animals with far slower generation times and greater housing and care requirements. Whatever models are proposed will have to exceed the existing rodent systems in a rigorous cost/benefit analysis before being adopted. The comparative study of Seok *et al.* provides a framework for molecular assessment to complement existing pathological and anatomical criteria.

In conclusion, this paper highlights the fact that there are major differences in gene expression between patients and existing murine inflammation models. This raises a number of important questions about why the models were so poorly predictive and provides a baseline to assess whether new models come closer to clinical reality. Perhaps the strongest lesson is an old one that is all to often forgotten in the rush of progress: all models are flawed, imperfect representations of reality. They are still useful if their limitations are understood and respected. Seok *et al.* have highlighted unappreciated weaknesses of a broad class of models and begun a process of adjustment that will hopefully lead to new treatments.

1 Seok J, Warren HS, Cuenca AG, Mindrinos MN, Baker HV, Xu W *et al*. Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proc Natl Acad Sci USA* 2013; **110**: 3507–3512.
2 van der Worp HB, Howells DW, Sena ES, Porritt MJ, Rewell S, O'Collins V *et al*. Can animal models of disease reliably inform human studies? *PLoS Med* 2010; **7**: e1000245.
3 Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P *et al*. Initial sequencing and comparative analysis of the mouse genome. *Nature* 2002; **420**: 520–562.
4 Wiersinga WJ. Current insights in sepsis: from pathogenesis to new treatment targets. *Curr Opin Crit Care* 2011; **17**: 480–486.
5 Mitka M. Drug for severe sepsis is withdrawn from market, fails to reduce mortality. *JAMA* 2011; **306**: 2439–2440.
6 Shultz LD, Brehm MA, Garcia-Martinez JV, Greiner DL. Humanized mice for immune system investigation: progress, promise and challenges. *Nat Rev Immunol* 2012; **12**: 786–798.
7 Drake AC, Chen Q, Chen J. Engineering humanized mice for improved hematopoietic reconstitution. *Cell Mol Immunol* 2012; **9**: 215–224.