

endothelium and vascular smooth muscle, the energy metabolism of adipocytes and skeletal muscle, hematopoiesis, platelet activation and immunity^{3,4,9–11}. Indeed, this study joins many others in underscoring that last category, in which HDL is emerging as a potential regulator of not only innate inflammatory responses to acute stimuli (such as the TLR ligands used in the present study) but also adaptive immunity, including antigen presentation, lymphocyte proliferation and activation bias, and immunological signaling. Perhaps most importantly, HDL can also regulate the low-intensity and prolonged inflammatory states that pervade and drive chronic metabolic dysfunctions such as metabolic syndrome and type 2 diabetes, as well as carcinogenesis and other inflammation-associated disease processes^{3,4,12}. It is also important to note that even its canonical cholesterol-transport roles are being revised and rethought because ‘calculated HDL’ (the accepted standard for HDL measurement in the clinical literature) has proven to be an inadequate predictor of disease risk^{1,2}.

This relative abundance of newly reported functions and correlations stands in sharp contrast, however, to the relative dearth of thorough mechanistic studies defining how exactly HDL influences target cells to achieve those effects. Moreover, the independent and dependent influences of each of these activities on higher-order physiological processes such as cardiovascular and metabolic health, vascular maintenance and immunity remain almost entirely unknown. The work of De Nardo *et al.* raises each of those general questions in more specific terms⁵; for example, what is the mechanism by which the HDL complex activates *Atf3*, how does that contribute to higher-order physiological processes in health and disease and, most importantly, how can that be exploited for therapeutic effect? Moreover, might HDL use that pathway to regulate other target-cell functions or related pathways to similarly target inflammation? Whatever questions are raised, the implications of the accompanying paper are clear: HDL exerts a potent influence over target cells that indicates tantalizing therapeutic potential for the treatment

of inflammatory injury and endothelial damage, in addition to its role in cholesterol homeostasis. With many of the developed world’s major killers thus neatly tied to this particle, HDL seems poised for renewed and refocused therapeutic investigation, despite recent therapeutic disappointments.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Heinecke, J.W. *Nat. Med.* **18**, 1346–1347 (2012).
2. Rader, D.J. & Tall, A.R. *Nat. Med.* **18**, 1344–1346 (2012).
3. Murphy, A.J. *et al. Biochim. Biophys. Acta* **1821**, 513–521 (2012).
4. Zhu, X. & Parks, J.S. *Annu. Rev. Nutr.* **32**, 161–182 (2012).
5. De Nardo, D. *Nat. Immunol.* **15**, 152–160 (2014).
6. Gilchrist, M. *et al. Nature* **441**, 173–178 (2006).
7. Patodia, S. & Raivich, G. *Front. Mol. Neurosci.* **5**, 8 (2012).
8. Lerch, P.G., Spycher, M.O. & Doran, J.E. *Thromb. Haemost.* **80**, 316–320 (1998).
9. Drew, B.G. *Nat. Rev. Endocrinol.* **8**, 237–245 (2012).
10. Mineo, C. & Shaul, P.W. *Circ. Res.* **111**, 1079–1090 (2012).
11. Nofer, J.R., Brodde, M.F. & Kehrel, B.E. *Clin. Exp. Pharmacol. Physiol.* **37**, 726–735 (2010).
12. Norata, G.D. *et al. Atherosclerosis* **220**, 11–21 (2012).

Strength in numbers: comparing vaccine signatures the modular way

W Nicholas Haining

Gene-expression signatures of the human vaccine response can be complex and noisy. Li *et al.* develop a new collection of gene-expression modules and use it to compare the response to five different vaccines.

Vaccines have improved human health enormously. The trouble is, we are not quite sure how they work. While we know a great deal about some components of the vaccine response, the answers to many important questions remain murky: for example, which features of the vaccine response are required for immunological protection, or whether different vaccines induce similar patterns of immunity. Without answers to these questions, clinical vaccine discovery today looks more like the trial-and-error efforts of Pasteur’s era than the rational design approach of modern drug development. As a result, progress in coming up with effective

vaccines against diseases such as HIV and tuberculosis remains slow. To accelerate vaccine development, several groups have identified gene-expression signatures present in human peripheral blood mononuclear cells (PBMCs) that predict immune responses to yellow fever^{1,2} and influenza vaccines³. However, the extent of vaccine-induced change in PBMC profiles can be small, and the number of genes measured in a typical gene-expression profiling experiment is large⁴, making it difficult to distinguish the signal from the noise. In this issue, Li *et al.*⁵ provide a new computational resource to make identifying subtle signatures easier and use it to compare the signatures elicited by five different vaccines.

The simplest way to analyze gene-expression profiles is to identify individual genes that are differentially expressed between phenotypes or conditions of interest. This can be achieved by comparing profiles from, say, samples obtained before and after vaccination, or samples from

subjects with varying degrees of vaccine-induced antibody responses, and identifying genes whose difference in expression is greater than would be expected by chance. This gene-by-gene approach has been very useful in identifying specific genes that regulate the immune response⁶, but it tends to ‘ignore’ more complex and subtle patterns evident in genome-scale expression data⁷.

Li *et al.*⁵ have developed a new resource to help identify biologically meaningful patterns of gene expression in PBMC profiles from vaccinated subjects, and they use it to ask whether different vaccines elicit unique or shared patterns of immunity. Their study provides both a technical resource for transcriptional analysis and the first comparison of gene-expression signatures elicited by different vaccines.

The technical resource they have developed is a compendium of coordinately expressed modules of genes. They identified the modules by analyzing patterns of gene expression present

W. Nicholas Haining is at the Dana-Farber Cancer Institute and Harvard Medical School, Boston, Massachusetts, USA, and at the Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA.
e-mail: nicholas_haining@dfci.harvard.edu

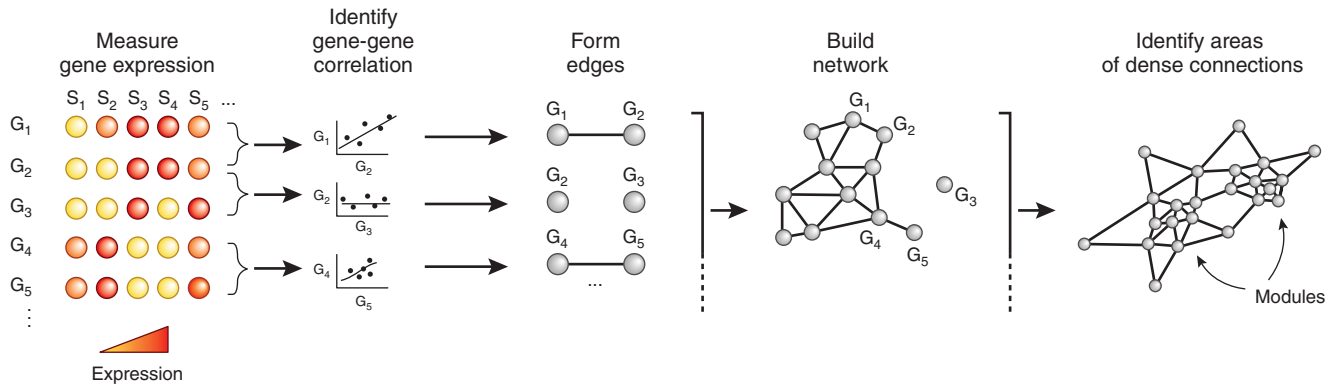


Figure 1 Identifying gene modules from gene-expression data. Biological functions of the cell are carried out by groups of genes—modules—that often share correlated levels of expression. In a typical experiment, expression levels of thousands of genes (G_1, G_2, \dots) are measured in many samples (S_1, S_2, \dots). Correlation in expression level of each gene and all others is tested with linear or nonlinear tests of association. Networks are modeled by connecting genes (represented as nodes) with lines (represented as edges) to correlated genes. Genes whose expression levels do not correlate are not connected. The interaction network is then compiled from all nodes and edges. Modules are identified as densely connected areas of the network.

in a very large data set created by merging 30,000 PBMC gene-expression profiles from 500 published studies. They used this reference collection of PBMC expression profiles to identify groups of genes whose expression correlates with each other. The graphical representation of this long list of gene-gene correlations (termed interactions) can be visualized as the familiar ‘hairball’ beloved of systems biologists but few others. In interaction networks, each gene is represented by a node linked by connections (edges) to other genes with correlated expression levels (Fig. 1). Because the expression levels of many genes are often closely correlated, interaction networks form a dense thicket of connections. Modules can be identified within such a network as groups of genes that have more connections between them than to other genes. Using such an approach with the interaction network defined from 30,000 PBMC samples, Li *et al.*⁵ refined and extracted a set of 334 “blood transcriptional modules.”

Why did they go to the trouble of coming up with this set of modules? Because without them, changes elicited by some of the vaccines they were studying were too subtle to be easily detected. They analyzed PBMC expression profiles from 30 healthy donors vaccinated with one of two vaccines against the bacterial pathogen *Neisseria meningitidis*. The first vaccine (MPSV4) is a quadrivalent vaccine containing polysaccharides from four serogroups of the organism; the second (MCV4), a polysaccharide-protein conjugate vaccine, comprises conjugates of the same four polysaccharides together with diphtheria toxoid (DT) protein adjuvant. Initial analysis of the transcriptional response to each of these vaccines using a conventional gene-by-gene analysis yielded a disappointingly small number of differentially expressed genes. The authors

found many fewer changes, for instance, than are elicited by other vaccines such as YF-17D (an attenuated live viral vaccine against yellow fever virus)^{1,2} or the trivalent inactivated influenza vaccine (TIV)^{3,4}. Undaunted, Li and colleagues then applied their collection of blood transcriptional modules to identify subtle differences between vaccines based on changes in the aggregate expression of gene modules.

This time, they found much more striking differences. They saw three broad patterns of expression: a protein recall response that correlated with the antibody response to TIV and with the antibody response to the DT portion of the MCV4; a primary viral response elicited by YF-17D; and an anti-polysaccharide signature shared by the response to the polysaccharide portions of MCV4 and MPSV4. In addition, the nature of the modules shared by different vaccines also raised hypotheses about how different vaccines function. For instance, the modules that correlated with the anti-polysaccharide response suggested the involvement of myeloid dendritic cells (DCs). Subsequent experiments confirmed that human myeloid DCs and mouse CD11C⁺ DCs were efficiently activated by incubation with MPSV4. This study represents the first step toward identifying the molecular signatures that correlate with antibody responses induced by different classes of vaccines.

There are a number of reasons why analyzing complex data sets based on differences in modules of co-regulated genes—rather than individual genes—makes sense. First, a modular approach to studying gene expression is not just analytic grandstanding; rather, it reflects a general design principle found in nature⁸. As much as hairball figures might exasperate non-systems biologists, they represent a fair approximation of how the cell’s business is transacted.

Cellular functions do not result from the isolated activities of individual genes. Instead, they arise from the cooperative effects of groups of functionally and sometimes physically interacting gene products. Fortunately, the transcript abundance of genes that function in the same biological process is often co-regulated⁹. This means that functionally related groups of genes can be detected using statistical tests of association such as the Pearson correlation coefficient (for linear relationships) or by using mutual information (for nonlinear relationships, as used by Li *et al.*⁵). Analytic approaches that identify differences in the expression of groups of functionally related genes are therefore likely to stay closer to the underlying biology than those that evaluate genes one by one⁴.

A second advantage is that analysis using modules of genes can detect subtle changes in gene expression in many functionally related genes even when large shifts in the expression of smaller numbers of genes are absent. This is because differences between cells states are often manifest by small changes distributed across networks of genes¹⁰. This is particularly important in the analysis of data from genetically heterogeneous human subjects where modest biological signals generated by vaccination can be swamped by noise, as Li *et al.*⁵ demonstrate.

Third, analysis at the module level is likely to be a more tolerant of measurement inaccuracies. For instance, measurement of differences in the expression of a few genes of interest in vaccinated subjects may be affected by technical variability from one study to the next. However, detection of the coordinated up- or downregulation of hundreds of genes is much less likely to be derailed by inaccurate measurement of a small number of them. Indeed

the modular arrangement of genes involved in important cellular functions may accomplish a similar error-reduction role in nature, insulating the cell from the consequences of aberrant expression of a rogue gene or two¹¹.

Much as Li *et al.*⁵ have done, another group has also developed a complementary resource of modules and used it to analyze vaccine responses¹². Together, these companion studies provide immunologists with a powerful set of resources for analyzing human PBMC gene-expression profiles. Helpfully, Li *et al.* have developed an intuitive web interface (<http://www.immuneprofiling.org/papers/meni/>) and application programming interface (API) to enable their modular resource to be easily used and integrated into other analysis software.

Still, a number of caveats remain to this modular approach to analysis. While the existence of modules of genes may be statistically evident, their function is not always so clear. Assigning a 'name' to a gene module is therefore dependent on expert knowledge that tends to introduce subjectivity into the process. A useful alternative is to identify sets of genes from those differentially expressed in specific, well-annotated experiments—an approach

that does not require manual annotation¹⁰. Other strategies to identify sets of genes from well-annotated experiments in a way that does not require manual annotation have proven a useful alternative¹⁰. In addition, a general concern when studying expression profiles from complex mixtures of cells such as those present in PBMC is that it is difficult to distinguish between changes in the biology of one cell state and changes in the relative frequencies of cells in a mixed population. However, approaches to deconvolve mixed cell populations analytically¹³ or to measure expression profiles in individual cells will help address this concern.

Finally, if we are to truly understand the mechanistic basis of the vaccine response, simply cataloging differences in transcriptional modules elicited by vaccination is unlikely to be sufficient by itself. Rather, we will need to combine observation of changes in modular gene expression in humans with directed experiments *in vitro* or in animal models to understand the biology of the underlying transcriptional circuits. For instance, perturbing putative regulators of expression modules can identify genes that control the transcriptional response to TLR

signaling¹⁴. The study by Li *et al.*⁵ is therefore important not only because it is the first comparative molecular analysis of five different vaccines but also because it starts to identify the transcriptional modules whose biology we must define in order to understand how vaccines work.

COMPETING FINANCIAL INTERESTS

The author declares no competing financial interests.

1. Querec, T.D. *et al.* *Nat. Immunol.* **10**, 116–125 (2009).
2. Gaucher, D. *et al.* *J. Exp. Med.* **205**, 3119–3131 (2008).
3. Nakaya, H.I. *et al.* *Nat. Immunol.* **12**, 786–795 (2011).
4. Tan, Y. *et al.* *Eur. J. Immunol.* doi:10.1002/eji.201343657 (29 November 2013).
5. Li, S. *et al.* *Nat. Immunol.* **15**, 195–204 (2014).
6. Ravindran, R. *et al.* *Science* doi:10.1126/science.1246829 (5 December 2013).
7. Haining, W.N. & Wherry, E.J. *Immunity* **32**, 152–161 (2010).
8. Hartwell, L.H., Hopfield, J.J., Leibler, S. & Murray, A.W. *Nature* **402**, C47–C52 (1999).
9. Segal, E. *et al.* *Nat. Genet.* **34**, 166–176 (2003).
10. Mootha, V.K. *et al.* *Nat. Genet.* **34**, 267–273 (2003).
11. Wagner, G.P., Pavlicev, M. & Cheverud, J.M. *Nat. Rev. Genet.* **8**, 921–931 (2007).
12. Obermoser, G. *et al.* *Immunity* **38**, 831–844 (2013).
13. Shen-Orr, S.S. *et al.* *Nat. Methods* **7**, 287–289 (2010).
14. Amit, I. *et al.* *Science* **326**, 257–263 (2009).