DOI 10.1002/art.10827

The promise and limitations of DNA microarray analysis: comment on the editorial by Firestein and Pisetsky

To the Editor:

A pinpoint-accurate editorial on quality requirements for complementary DNA (cDNA) array technology in science was presented by Firestein and Pisetsky (1). As they note, as "with all 'hot' new methods," many years will pass before the strengths and weaknesses are recognized, microarray technology "represents another cutting-edge technology that promises extraordinary advances in the study of disease," the "siren song of microarrays is indeed strong and tantalizing," and in the "face of such power and precision, it may appear as if hypotheses are superfluous to the pursuit of research and that specific questions need no longer be asked." Criteria such as reproducibility, detailed biostatistics, reasonable origin of the material, and external confirmation represent central issues for assuring scientific quality rather than merely amassing data. There is nothing to be added or withdrawn from the attempt to build state-of-the-art standards as proposed in the editorial, although the technical difficulties currently associated with this technology have to be acknowledged. We would, however, like to comment on a few points made by Firestein and Pisetsky.

It is questionable that Northern blot and quantitative polymerase chain reaction (PCR) approaches per se are more reliable than cDNA array experiments for determination of messenger RNA (mRNA) expression levels. Clearly these techniques, if performed carefully, are less likely to be affected by nonspecific signal reactions. However, the values obtained with these assays also do not reflect absolute values of gene expression levels, and thus, quantification in these systems relies very much on the availability of so-called "housekeeping genes" for "normalization." To date, though, no housekeeping gene that is not up- or down-regulated in some conditions has been found (1,2), as has now also become particularly apparent with gene array technology. In this respect, cDNA array experiments potentially represent a preferable option, however, since they allow a normalization approach based on a much larger panel of available gene expression data (3). Consequently, GAPDH ratios found by, for example, Northern blot and cDNA array do not need to be similar. Clearly, none of the techniques currently available for quantification of mRNA expression is optimal, and certainly a method in which a large amount of data are produced in parallel will not easily match the technical accuracy of time-consuming technologies such as online quantitative PCR, but expression analysis by cDNA arrays offers its own strengths. It should be viewed with the same caution and enthusiasm as all other techniques.

A potential shortcoming of the editorial is that it is too focused on the conventional, straightforward use of microarrays, since it only considers the "purpose of microarrays as a screening test to identify the best targets for hypothesis-driven studies" instead of exploiting more intertwined combinations of hypothesis-driven research and large-scale screening. At present, for various reasons (mainly related to the shortage of suitable computational tools [4]), cDNA arrays are primarily used for screening of genes at a high throughput level. However, much more strength will come from the promises of this technology to address biologic complexity at a new level, it opens up the possibility of intricate computational analysis in terms of depicting molecular networks as well as classifying diseases on a molecular basis, an issue of high interest for diseases showing such a wide and heterogeneous spectrum as rheumatoid arthritis and osteoarthritis. Thus, as one approach, one could attempt to map the identified genes onto precompiled biochemical networks (5), thereby displaying the putative target genes in their known or postulated metabolic and/or signaling context. This allows determination of metabolic pathways and signaling cascades, which provide a biologic background and context, on the basis of which parts of the experimental data can be explained.

As a second, converse approach, one could use the available network knowledge for investigation of possible pathways as biologic entities which are subject to a coordinated up- or down-regulation in a certain biologic context. This would allow evaluation of a large number of potentially biologically meaningful pathways with respect to their significance given the available (expression) data. Also, for osteoarthritis research, this would support the interpretation of gene regulation within the context of the biology of the disease (for review, see ref. 4), rather than simply providing large amounts of (expression) "naked" data.

Overall, high-throughput gene expression analysis has great potential as a technological tool, but even more, it opens up a great future for molecular research. At present, we are still struggling mainly with the technologic and computational limitations, and it must be emphasized that most results have to be considered to be somewhat preliminary since technical tools and analysis methods are only now emerging and require more validation. Tomorrow, however, this methodology will be one pillar in our attempts to understand the biology and pathobiology of molecules, cells, tissues, and diseases. The editorial by Firestein and Pisetsky is an important and groundbreaking step.

Thomas Aigner, MD
University of Erlangen-Nürnberg
Nurenberg, Germany
Ralf Zimmer, PhD
University of Munich
Munich, Germany

1. Firestein GS, Pisetsky DS. DNA microarrays: boundless technology or bound by technology? Guidelines for studies using microarray technology [editorial]. Arthritis Rheum 2002;46:859–61.
2. Sive JI, Baird P, Jeziorsk M, Watkins A, Hoyland JA, Freemont AJ. Expression of chondrocyte markers by cells of normal and degenerate intervertebral discs. Mol Pathol 2002;55:91–7.
3. Zien A, Aigner T, Zimmer R, Lengauer T. Centralization: a new paradigm for the normalization of gene expression data. Bioinformatics 2001;17 Suppl:S323–S31.
4. Aigner T, Zien A, Bartnik E, Zimmer R. Functional genomics of osteoarthritis. Pharmacogenomics. 2002;3:635–50.
5. Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, et al. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. Science 2001;292:929–34.