

Finding common ground in cancer research

Reproducibility in science is a prominent topic in both lay and scientific press. But a new facet of this discussion has arisen in a recent comparison of two pharmacogenomic studies, and it calls for an evaluation of how we interpret science in the face of discrepant results.

In 2012, two independent reports in *Nature* presented comprehensive *in vitro* analyses of the genomic features of hundreds of cancer cell lines, coupled with their sensitivities to numerous anticancer drugs. By analyzing DNA sequence, copy number variation and gene expression, they identified predictors of known and previously unknown sensitivities to certain anticancer drugs, which the authors proposed might be translated to new cancer treatment options in the clinic (*Nature* 483, 570–575, 2012; *Nature* 483, 603–607, 2012).

Their simultaneous publication and their overlapping analysis of 15 drugs and 471 cell lines also presented the opportunity for their cross-validation. John Quackenbush and his colleagues recently published just such an analysis and found that the genomic data—mutations and gene expression profiles of the 471 cell lines—were highly consistent between the two studies, but, surprisingly, the drug sensitivities of the cell lines were not (*Nature* 504, 389–393, 2013). Moreover, when they assessed the association of genomic data (such as gene expression) with drug sensitivity, they again found poor correlations between the two reports—a finding that undermines the rationale of the studies, which aimed to identify genomic profiles predictive of drug response.

Notably, the two pharmacogenomic papers used different assays for determining cell viability after drug treatment (one measured intracellular ATP concentration, whereas the other used a dye that binds nucleic acid in living cells), and Quackenbush and his co-workers concluded that differences in these methods probably account for the discrepancies between the studies. However, even in the one instance when both groups tested one drug (camptothecin) using the same assay and cell lines, Quackenbush and colleagues found that the IC_{50} values measured by each group showed only fair correlation, indicating that additional methodological variables may contribute to the discordant results.

The need for improved reproducibility of findings in cancer research has been underscored by two commentaries detailing the inability of scientists at pharmaceutical companies to reproduce results from the published literature (*Nat. Rev. Drug Discov.* 10, 712, 2011; *Nature* 483, 531–533, 2012). However, those commentaries did not identify the targets pursued, the nature of the experiments performed, the specific discrepancies in the data obtained and their potential basis in the variability of reagents or assays, and the criteria for a determination of nonreproducibility. As the data from the validation attempts are not available for peer review, the conclusions from the commentaries—although worrisome—are based on anecdotal evidence. Thus, the availability of the extensive data in the 2012

Nature papers and their overlapping analysis of a subset of the drugs and cell lines permit an objective and quantitative assessment of the concordance between related studies and, by extension, of the robustness of their conclusions.

To be fair, the authors of the cancer pharmacogenomic studies acknowledged that although their findings suggest potential new therapeutic leads for certain cancer types, their results are limited to an *in vitro* assessment of drug sensitivity and do not provide definitive insight into the likely efficacy of drugs *in vivo*. However, the analysis by Quackenbush and his colleagues raises a more basic concern regarding the ability of different labs to generate comparable results when using the same reagents in arguably straightforward assays and the extent to which interlab assay variability affects research results in general.

In contrast, standardization of microarray analyses has, over time, resulted in increasingly reproducible results, as is clear from the concordance of the genomic data between the two *Nature* pharmacogenomics studies. But the need for standardized tools was apparent soon after the widespread adoption of microarray technology because of the high variability obtained with different methods of reagent preparation, microarray platforms and software, and even different hands. In contrast, assays of cell viability are not new, and the incentive to troubleshoot an older technique is substantially lower. These findings should alert academia and the companies marketing phenotypic assays that the variability of results using these tests may mitigate their value.

The mechanisms underlying the drug sensitivities of cancer remain poorly understood, making it difficult to predict the response of individual patients to a given treatment. Although some cancers—such as BCR-ABL⁺ chronic myelogenous leukemia—are associated with defining mutations that generate a gene product that is highly susceptible to a specific inhibitor, most are not. Thus, identifying gene expression responses or genomic alterations associated with *in vitro* drug response that can be translated to more informed and ultimately more personalized treatment decisions is an attractive solution for cancer therapy. But the technological advances achieved in genomics may be overshadowed by a lack of comparable development and standardization of cell-based assays that are robustly reproducible and of assays that provide some measure of *in vivo* insight. The wealth of genomic data generated from cell lines and primary tissues has exponentially increased the numbers of testable hypotheses in cancer and other fields. Now we need tools to make common sense of the data.