Predicting drug sensitivity

| ArticleContext | : | 13059221212 |

Agnieszka M Lichanska

Abstract

Transcriptional profiling of 60 human cancer lines has been used to assess chemosensitivity to 232 chemical compounds.

Significance and context

Pharmacogenomics aims to predict an individual patient's response to particular drugs on the basis of their genotype, in order to design tailor-made treatments. These are particularly needed in cancer, where the ease with which drug resistance develops is most variable and unpredictable. A known effective drug is expected to kill the tumor, but most of these drugs can also have serious side effects on the patient's metabolism. The development of oligonucleotide microarrays that enable analysis of the expression of many genes simultaneously will make it easier to predict which cellular functions are likely to be affected by a drug.

Key results

Staunton *et al*. have developed a new algorithm for analyzing gene expression with respect to the drug sensitivity of the cells. The authors combined information from the analysis of gene expression in untreated cells and from drug-sensitivity testing (232 drugs out of 5,084 initially assayed) in 60 tumor cell lines to find sets of genes (classifiers) that describe a cell's sensitivity. Generation of the classifiers was based on training sets, and experimental results from the remaining cell lines were used to test the accuracy of the predictions.

Statistical analysis showed that the classification based on gene expression is significantly non-random. The validity of the classifiers was evaluated on a test set with 38% of the expression-based classifiers performing with an accuracy of 64-92%. This indicates that there is a significant group of drugs for which expression data have adequate predictive power.

The authors show example results for cytochalasin D, for which a 120-gene classifier of 80% accuracy predicted sensitivity in 20 of the cell lines. Of the classifier genes, 24% were cytoskeleton and extracellular matrix (ECM) genes, in keeping with the known mode of drug function. In contrast, the highly accurate (87%) gene classifier for the antifolate drug NSC633713 also includes more than 20%

cytoskeletal/ECM genes. These results were unexpected, and suggest that cytoskeleton components might influence cell sensitivity to drugs independent of their cellular target.

Methodological innovations

A new algorithm was developed to classify the chemosensitivity of the cell lines. Training sets were selected by picking the most sensitive and most resistant cell lines within each tissue. The remaining cell lines were used as test samples for evaluation of the classifiers. A set of marker genes or a classifier was identified by using weighted voting classification. The only gene expression differences considered significant and used for voting were those greater than fivefold and 500 average difference units (calculated according to the chip manufacturer's instructions) across all training cell lines and greater than twofold in each pair of training cell lines. The class of the cell line was described by summing the votes for each marker gene. To ensure that the small sample size does not interfere with the accuracy, classifiers were optimized by cross-validation. This was done by taking one of the cell lines out of the training set and training the classifier on the other cell lines; it was then used to predict the sensitivity of the withdrawn cell line, and the process repeated with other cell lines.

Links

Expression data, cross-validation accuracy rates and classifier genes and weights are available at the authors' webpage Supplemental data for Staunton *et al*.

Conclusions

Staunton *et al*. conclude that their new algorithm for predicting cells' reactions to chemical compounds has considerable accuracy: 88 out of 232 gene-based classifiers predicted accurately, with only 12 expected to do so by chance. Their results suggest that gene-expression-based predictions of sensitivity are possible for at least some drugs.

Reporter's comments

This work is the first step towards designing ways of predicting an individual's response to chemotherapies. It is clear that although the degree of sensitivity to all drugs is not going to be predictable, a significant proportion of them can be analyzed in such a manner. There is still a lot to be

done in regard to validating the prediction algorithm on real tissue samples; an increase in its accuracy will be important. Possible problems with the method, such as small datasets (only two to nine cell lines), are highlighted by the authors and can lead to overestimation of accuracy; this problem was partially addressed by cross-validation. Use of larger numbers of cell lines and larger arrays might also help improve the performance of the method. The problem of the tissue specificity of drug action was touched on, but not expanded on in the analysis and discussion, and this is an important issue requiring further study.

Table of links

Proceedings of the National Academy of Sciences of the United States of America

Supplemental data for Staunton *et al*

References

1. Staunton JE, Slonim DK, Coller HA, Tamayo P, Angelo MJ, Park J, Scherf U, Lee JK, Reinhold WO, Weinstein JN, et al: Chemosensitivity prediction by transcriptional profiling. Proc Natl Acad Sci U S A. 2001, 98: 10787-10792. 0027-8424

This PDF file was created after publication.