

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812

AUGUST 5, 2004

VOL. 351 NO. 6

Gene-Expression Patterns in Drug-Resistant Acute Lymphoblastic Leukemia Cells and Response to Treatment

Amy Holleman, B.Sc., Meyling H. Cheok, Ph.D., Monique L. den Boer, Ph.D., Wenjian Yang, Ph.D., Anjo J.P. Veerman, M.D., Ph.D., Karin M. Kazemier, Deqing Pei, M.Sc., Cheng Cheng, Ph.D., Ching-Hon Pui, M.D., Mary V. Relling, Pharm.D., Gritta E. Janka-Schaub, M.D., Ph.D., Rob Pieters, M.D., Ph.D., and William E. Evans, Pharm.D.*

ABSTRACT

BACKGROUND

Childhood acute lymphoblastic leukemia (ALL) is curable with chemotherapy in approximately 80 percent of patients. However, the cause of treatment failure in the remaining 20 percent of patients is largely unknown.

METHODS

We tested leukemia cells from 173 children for sensitivity *in vitro* to prednisolone, vincristine, asparaginase, and daunorubicin. The cells were then subjected to an assessment of gene expression with the use of 14,500 probe sets to identify differentially expressed genes in drug-sensitive and drug-resistant ALL. Gene-expression patterns that differed according to sensitivity or resistance to the four drugs were compared with treatment outcome in the original 173 patients and an independent cohort of 98 children treated with the same drugs at another institution.

RESULTS

We identified sets of differentially expressed genes in B-lineage ALL that were sensitive or resistant to prednisolone (33 genes), vincristine (40 genes), asparaginase (35 genes), or daunorubicin (20 genes). A combined gene-expression score of resistance to the four drugs, as compared with sensitivity to the four, was significantly and independently related to treatment outcome in a multivariate analysis (hazard ratio for relapse, 3.0; $P=0.027$). Results were confirmed in an independent population of patients treated with the same medications (hazard ratio for relapse, 11.85; $P=0.019$). Of the 124 genes identified, 121 have not previously been associated with resistance to the four drugs we tested.

CONCLUSIONS

Differential expression of a relatively small number of genes is associated with drug resistance and treatment outcome in childhood ALL.

From the Division of Pediatric Oncology–Hematology, Erasmus University Medical Center, Sophia Children’s Hospital, Rotterdam, the Netherlands (A.H., M.L.B., K.M.K., R.P.); the Departments of Pharmaceutical Sciences (M.H.C., W.Y., M.V.R., W.E.E.), Hematology–Oncology (C.-H.P.), and Biostatistics (D.P., C.C.), St. Jude Children’s Research Hospital, Memphis, Tenn.; the Pharmacogenetics of Anticancer Agents Research Group in the Pharmacogenetics Research Network, Memphis, Tenn. (W.Y., C.-H.P., M.V.R., W.E.E.); University of Tennessee Colleges of Pharmacy and Medicine, Memphis (C.-H.P., M.V.R., W.E.E.); Free University Medical Center, Department of Pediatric Hematology–Oncology, Amsterdam (A.J.P.V.); and the German Cooperative Study Group for Childhood Acute Lymphoblastic Leukemia (COALL), Department of Hematology–Oncology, Children’s University Hospital, Hamburg, Germany (G.E.J.-S.). Address reprint requests to Dr. Evans at St. Jude Children’s Research Hospital, Department of Pharmaceutical Sciences, 332 N. Lauderdale St., Memphis, TN 38105, or at william.evans@stjude.org.

N Engl J Med 2004;351:533-42.

Copyright © 2004 Massachusetts Medical Society.

*Ms. Holleman and Dr. Cheok contributed equally to the article, and Drs. Pieters and Evans contributed equally to the article.

IMPROVEMENTS IN THE TREATMENT OF childhood acute lymphoblastic leukemia (ALL) over the past four decades have resulted in rates of long-term disease-free survival of approximately 80 percent.^{1,2} We have shown that children whose ALL cells exhibit *in vitro* resistance to antileukemic agents have a substantially worse prognosis than children whose ALL cells are drug-sensitive.³⁻⁵ However, little is known about the genetic basis of resistance to chemotherapy. Multidrug-resistance genes⁶ and genes involved in cell-cycle progression,^{7,8} DNA repair,⁹ drug metabolism,⁹⁻¹¹ and apoptosis¹² have been associated with the prognosis of ALL, but their role in determining the sensitivity of ALL cells to individual antileukemic agents is not known. Gene products arising from rearrangements of the *TEL-AML1*,¹³ *BCR-ABL*,¹⁴ and *MLL*¹⁵ genes are also associated with prognosis and drug resistance, but for unknown reasons, many patients with a favorable genetic subtype (e.g., *TEL-AML1*) are not cured, whereas many with an unfavorable subtype (e.g., certain *MLL* rearrangements) are cured. Although it is likely that multiple pathways and genes contribute to the sensitivity of ALL cells to specific agents,¹⁶⁻¹⁸ all studies to date have focused on a small number of candidate genes instead of taking advantage of the genomic survey that is possible with the use of gene-expression profiling. Such profiles have been used successfully to investigate drug resistance in cancer cell lines^{19,20} and human tumor xenografts,²¹ but not in primary cancer cells.

Gene-expression profiles can differentiate lineage (T cell or B cell) and molecular subtypes of ALL²²⁻²⁵ and identify treatment-specific changes in gene expression in ALL cells.²³ However, it is not known whether gene-expression profiles of leukemia cells are associated with resistance to individual drugs. The present study was undertaken to identify genes that are differentially expressed in primary ALL cells exhibiting resistance or sensitivity to prednisolone, vincristine, asparaginase, or daunorubicin and to determine whether the expression of such genes influences the response to treatment.

METHODS

PATIENTS

The study population consisted of 271 children with newly diagnosed ALL: 173 were enrolled as part of the 9th ALL Dutch Childhood Oncology Group pro-

col at Erasmus Medical Center, Sophia Children's Hospital, in Rotterdam or treatment protocols 92 and 97 of the German Cooperative Study Group for Childhood Acute Lymphoblastic Leukemia in Hamburg, and 98 were enrolled as part of the Total Therapy protocols XIII A and XIII B of St. Jude Children's Research Hospital in Memphis, Tennessee.^{26,27} Patients were enrolled in the German protocol from 1992 to 2003, in the Dutch protocol from 1997 to 2004, and in the St. Jude protocols from 1991 to 1998. The original gene-profiling population consisted of the 173 children in the Dutch and German protocols, and the independent-validation population consisted of the 98 patients in the St. Jude protocols. The parents or guardians of the patients provided written informed consent, and the patients provided assent.

ISOLATION OF LEUKEMIA CELLS

Bone marrow and peripheral blood were obtained before treatment, and mononuclear cells were isolated by means of sucrose density-gradient centrifugation (density, 1.077 g per milliliter; Lymphoprep, Nycomed Pharma) within 24 hours. Cells were resuspended in RPMI 1640 medium (GIBCO BRL) supplemented with 20 percent fetal-calf serum (Integro), 2 mM L-glutamine, 200 µg of gentamicin per milliliter (GIBCO BRL), 100 IU of penicillin per milliliter, 100 µg of streptomycin per milliliter, 0.125 µg of fungizone per milliliter (GIBCO BRL), 5 µg of insulin per milliliter, 5 µg of transferrin per milliliter, and 5 ng of sodium selenite per milliliter (ITS media supplement, Sigma-Aldrich Chemie). If necessary, ALL samples were further enriched to achieve more than 90 percent blasts by removing nonmalignant cells with the use of immunomagnetic beads (DynaBeads).

DRUG-RESISTANCE ASSAY

The sensitivity of leukemia cells to prednisolone (Bufa Pharmaceutical Products), vincristine (TEVA Pharma), asparaginase (Paronal, Christiaens), and daunorubicin (Cerubidine, Rhône-Poulenc Rorer) was determined with the use of the four-day *in vitro* 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-tetrazoliumbromide (MTT) drug-resistance assay.³ The following concentrations of each drug were tested: 0.008 to 250 µg of prednisolone per milliliter, 0.05 to 50 µg of vincristine per milliliter, 0.003 to 10 IU of asparaginase per milliliter, and 0.002 to 2.0 µg of daunorubicin per milliliter. The drug con-

centration lethal to 50 percent of the leukemia cells (the LC_{50} value) was used as the measure of drug resistance. The LC_{50} values used to define cells as sensitive or resistant to each agent were those previously associated with a good or bad treatment outcome in patients with ALL (see Table 1 in the Supplementary Appendix, available with the full text of this article at www.nejm.org).³⁻⁵

PURIFICATION, LABELING, AND HYBRIDIZATION OF RNA

Total cellular RNA was extracted from a minimum of 5×10^6 leukemia cells with the use of Trizol reagent (GIBCO BRL), RNA was additionally purified with phenol-chloroform-isoamylalcohol (25:24:1), and RNA integrity was assessed as previously described.^{23,24} RNA processing and hybridization to the U133A GeneChip oligonucleotide microarray (Affymetrix) were performed according to the manufacturer's protocol.

STATISTICAL ANALYSIS

Gene-expression values were calculated with the use of Affymetrix Microarray Suite version 5.0.^{23,24} Expression signals were scaled to the target intensity of 2500 and log-transformed. Arrays were omitted if the scaling factor exceeded 3 SD of the mean or if the ratio of 3' to 5' messenger RNA for β -actin or glyceraldehyde-3-phosphate dehydrogenase was greater than 3. From the total of 22,283 probe sets, those expressed in fewer than five patients were omitted, leaving 14,550 probe sets for subsequent analyses.

For each antileukemic agent, we identified genes that were most discriminative for resistance and sensitivity using the Wilcoxon rank-sum test and t-test for each probe set and estimated the false discovery rate using the q value according to Storey and Tibshirani.²⁸ At the selected P value (alpha) for ranked discriminating genes (e.g., $P < 0.001$), the overall significance of the estimated false discovery rate was computed as the probability of observing equal or lower false discovery rates on the basis of 1000 random permutations.

To assess the predictive accuracy using the top 30, 50, and 100 discriminating genes for drug sensitivity as compared with drug resistance, for each drug, we randomly divided the patients with drug-sensitive leukemic cells and the patients with drug-resistant leukemic cells into two groups, using two thirds to build the model and one third to assess the

Glossary.

Bagging algorithm: An iterative procedure (also called "bootstrap averaging") that reduces the variability of estimating predictive accuracy. Models based on the bootstrap sample set are used to predict the value of interest for samples not included in the bootstrap sample set.

False discovery rate: The proportion of false positive results incurred in multiple testing among selected variables (e.g., gene-probe sets), expressed as the q value.

Gene Ontology (GO) classification: A set of controlled categories linking gene products with biologic processes, molecular functions, and cellular components, allowing exploration of common functional categories in a selected set of genes.

Genetic subtypes of ALL: The t(9;22)(q34;q11) translocation is a translocation of parts of chromosomes 9 and 22 to create the *BCR-ABL* fusion gene, which is associated with an unfavorable prognosis in ALL. The t(1;19)(q23;p13.3) translocation creates the *E2A-PBX1* fusion gene and occurs in approximately 5 percent of cases of pediatric ALL. The t(4;11)(q21;q23) translocation creates the *MLL-AF4* fusion gene; this is the most common genetic abnormality in infant ALL but occurs in only 5 percent of cases of childhood ALL. Approximately 40 different partners are known to be fused to the *MLL* gene; any *MLL* rearrangement carries an unfavorable prognosis. The t(12;21)(p13;q22) translocation creates the *TEL-AML1* fusion gene, the most common genetic abnormality in childhood ALL, which is associated with a relatively favorable prognosis.

Hierarchical clustering: A computational method that uses similarity among data to group genes or samples into clusters. The result is typically represented by a cluster tree or dendrogram.

Hyperdiploidy: ALL cells with more than 50 chromosomes; a relatively favorable prognostic feature.

Permutation analysis: An iterative procedure that measures the probability of observing a finding by chance alone. In each iteration, a permuted data set is generated by randomly assigning class labels (e.g., drug resistance) to each sample. The statistic is computed on the basis of the permuted sample set. Given a large number of iterations (e.g., 1000), the probability of observing a given finding by chance is computed (e.g., 7 of 1000 = 0.007) and used to determine the level of significance.

Principal-component analysis: A technique for visualizing high-dimensional data sets. The top components usually explain most of the variation. Data points are more similar to each other if they are shown closer together. Each component is a weighted linear combination of the original selected variables (e.g., expression of selected genes).

Probe set: A collection of oligonucleotide-probe pairs designed to detect the level of expression of a specific gene. Two oligonucleotides (25 bases) designed to be complementary to the target sequence, of which one has a mismatch in the central (13th) position, define a probe pair. The mismatch probe serves as an internal control. A probe set includes 11 probe pairs on the Affymetrix U133A GeneChip.

Supervised methods: The use of gene expression to distinguish among known phenotypes (also called class prediction). The predefined phenotype (e.g., drug resistance) is used to guide the selection of gene-expression patterns.

Support vector machine: A classification tool similar to logistic regression but designed to exploit the nonlinear relationship among data points.

Unsupervised analysis: A method used to identify relationships in a microarray data set (also called class discovery) without reference to the phenotype.

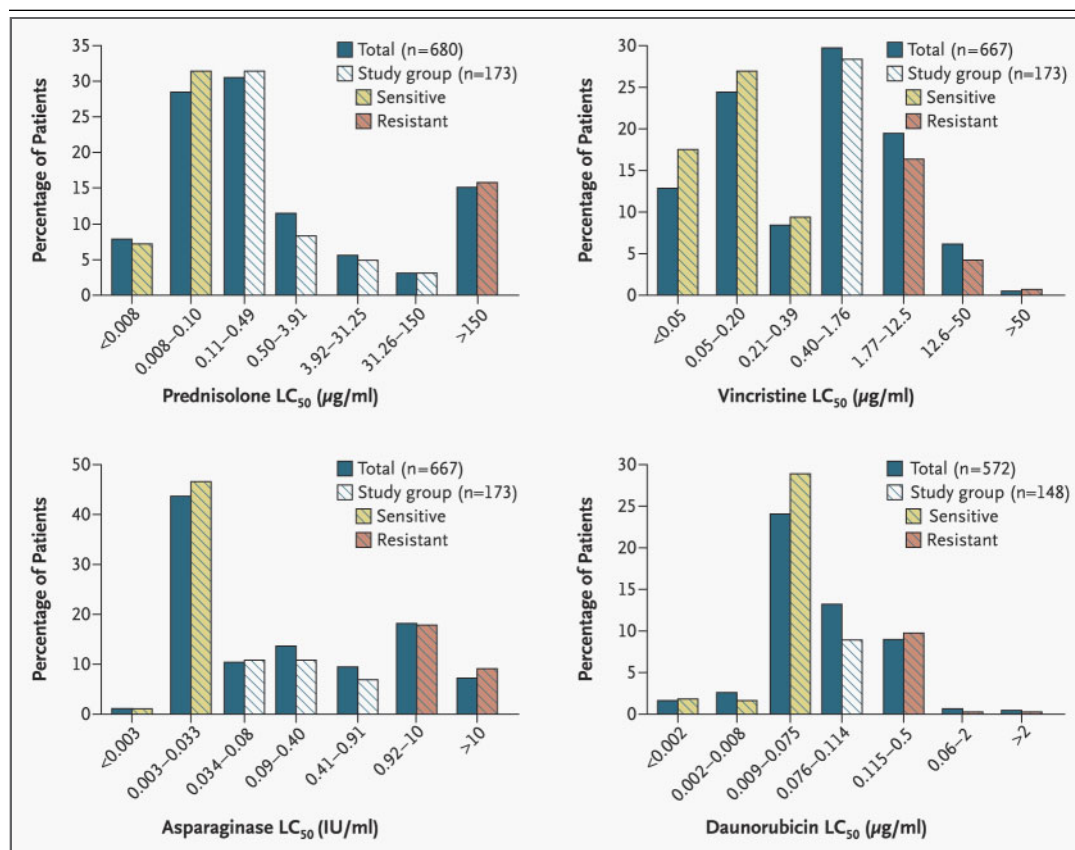


Figure 1. Distribution of the Drug Concentrations Lethal to 50 Percent of Primary Leukemia Cells (LC₅₀) in the Study Group and in the Larger Population of Children with ALL.

The study group comprised 173 patients whose leukemia-cell samples were selected for gene-expression analysis from the total group of approximately 700 patients whose ALL blasts had been assessed at diagnosis for sensitivity to a panel of four antileukemic agents. The distribution of LC₅₀ values between the study group and the corresponding total group did not differ significantly for any of the drugs: P=0.89 for prednisolone, P=0.63 for vincristine, P=0.89 for asparaginase, and P=0.22 for daunorubicin (by the chi-square test).

accuracy of the model. This process was repeated 1000 times; in each case we reselected a fixed number of probe sets to build a prediction model using support vector machines. Predictive accuracies of the various gene-expression profiles with respect to the sensitivity of each antileukemic agent and their confidence intervals were computed with the use of data from the 173 Dutch and German patients.

In the outcome analysis, we computed drug-resistance gene-expression scores for the 173 Dutch and German patients in the original population and the 98 St. Jude patients²⁵ in the validation population on the basis of the 172 gene-probe sets that discriminated between leukemic cells that were sensitive and those that were resistant to each of the four drugs. The scores were computed with the use

of bagging algorithms.²⁹ For each of the four drugs, we assigned each patient a score of 1 if the cells were predicted to be sensitive and 2 if the cells were predicted to be resistant. After 1000 iterations, the average scores for each of the four drugs for each patient were combined as the final drug-resistance gene-expression score and used in the outcome analysis. For the analysis of disease-free survival, any type of leukemia relapse was considered. The duration of disease-free survival was defined as the time from diagnosis until the date of treatment failure. Data were censored at the time of the last follow-up visit in the absence of treatment failure. Cox proportional-hazards regression analysis was used to assess the association between the combined gene-expression score and treatment

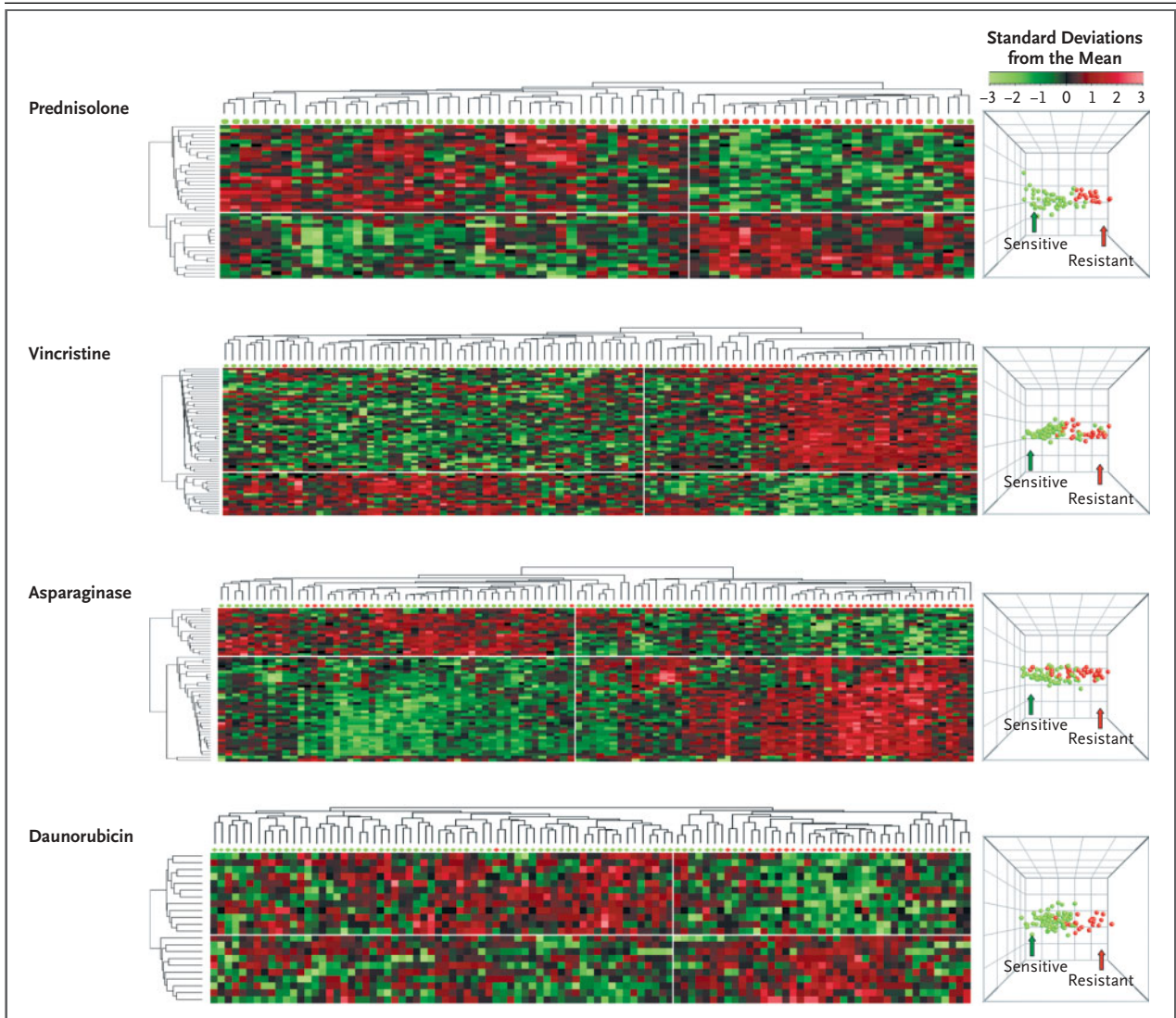


Figure 2. Results of Supervised Hierarchical-Clustering and Principal-Component Analyses with the Use of Genes That Discriminate between Drug-Resistant and Drug-Sensitive B-Lineage ALL with Respect to Prednisolone, Vincristine, Asparaginase, and Daunorubicin.

The Wilcoxon rank-sum test and t-test were used to identify genes that were differentially expressed in sensitive and resistant ALL ($P < 0.001$). Each column represents an ALL sample, labeled according to whether it was sensitive (green) or resistant (red) to a given drug, and each row represents a probe set. The “heat” maps on the left side of the figure indicate a high (red) or low (green) level of expression relative to the number of standard deviations from the mean. For prednisolone, 42 probe sets were found to discriminate resistant leukemia cells from sensitive leukemia cells (33 genes and 3 complementary DNA [cDNA] clones); for vincristine, 59 such probe sets were identified (40 genes and 14 cDNA clones); for asparaginase, 54 such probe sets were identified (35 genes and 10 cDNA clones); and for daunorubicin, 22 such probe sets were identified (20 genes and 2 cDNA clones). The three-dimensional plots on the right show three principal components based on the significant discriminating genes for each drug. Each circle represents a patient with leukemia; red circles indicate those with drug-resistant ALL, and green circles those with drug-sensitive ALL.

outcome. Leukemia-free survival was analyzed with the use of Fine and Gray’s estimator accounting for competing events.³⁰

We used Fisher’s exact test to determine the degree of overrepresentation or underrepresenta-

tion of discriminating genes in specific functional groups as compared with the genes on the U133A GeneChip, using the Gene Ontology database (<http://www.geneontology.org/>). Probe sets with the same gene symbol were counted as one. Primary

data are available through the GeneExpression Omnibus of the National Center for Biotechnology Information at <http://www.ncbi.nlm.nih.gov/geo/> (Platform, GPL91; Sample, GSM9653 to 9934; Series, GSE635 to 660). Additional information concerning the methods used is available at www.stjuderresearch.org/data/ALL4/, at www.eur.nl/fgg/kgk/, and in the Supplementary Appendix.

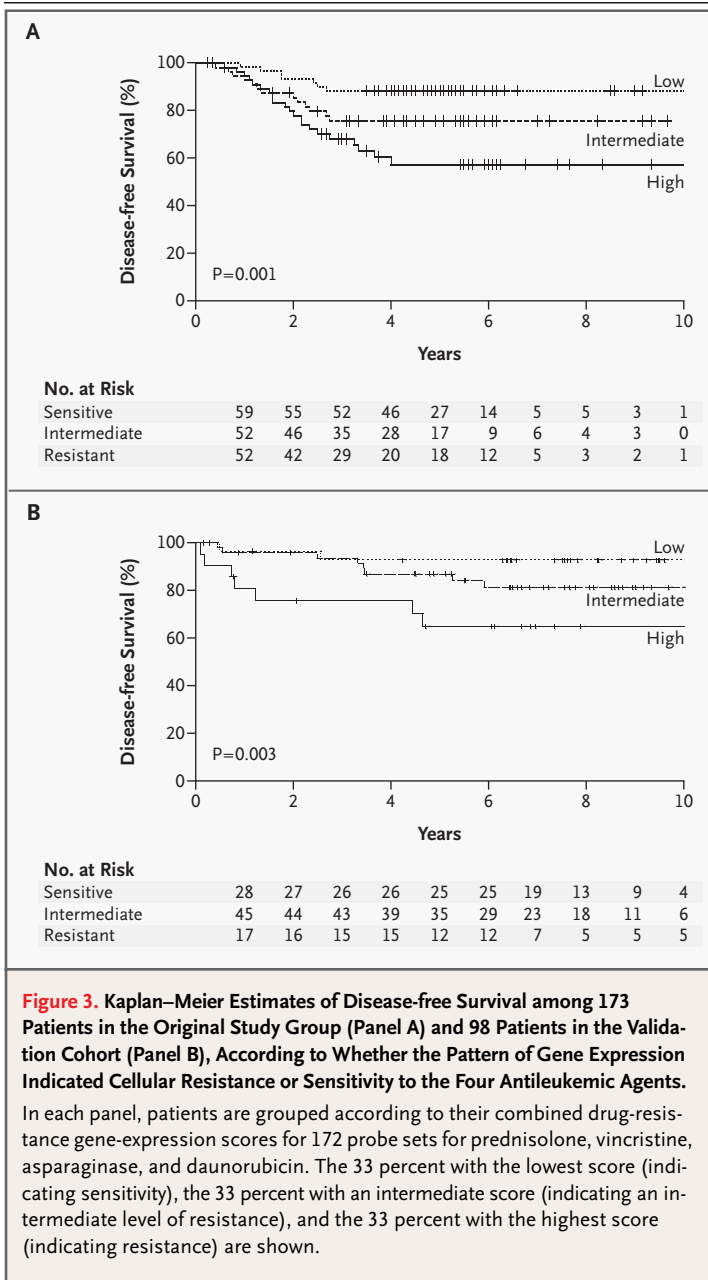
RESULTS

Gene expression was determined in ALL cells from 173 patients with newly diagnosed disease whose leukemia cells were either sensitive or resistant to prednisolone, vincristine, asparaginase, or daunorubicin, as assessed by the in vitro MTT assay. The distribution of LC₅₀ values (the drug concentration lethal to half the cultured lymphoblasts) in our study population did not differ significantly from that of the entire population of approximately 700 patients for whom we had previously determined the sensitivity status to each of these antileukemic agents (Fig. 1). Likewise, the proportion of patients classified as having sensitive or resistant leukemia cells, according to previously defined LC₅₀ values (Table 1 in the Supplementary Appendix)³⁻⁵ did not differ significantly between the study group and the entire population (Fig. 1). The leukocyte counts, age at diagnosis, proportions of girls and boys, and immunophenotypes in the drug-sensitive and drug-resistant groups for each antileukemic agent are summarized in Table 2 in the Supplementary Appendix.

PREDICTION OF SENSITIVITY AND RESISTANCE WITH THE USE OF DIFFERENTIALLY EXPRESSED GENES

Unsupervised hierarchical clustering, which groups patients according to the predominant similarities in gene expression, did not cluster patients according to their resistance to any of the four antileukemic agents. Rather, patients were clustered predominantly according to immunophenotype or ALL genetic subtype (Fig. 1 in the Supplementary Appendix).²⁴ Because cases of T-cell ALL have a strong gene-expression signature, subsequent analyses were performed with the use of all samples or only the samples of B-lineage ALL (Table 2 in the Supplementary Appendix). At 28, the number of cases of T-cell ALL was too small for a separate analysis.

Supervised methods (i.e., the Wilcoxon rank-sum test or t-test) were used to identify genes associated with resistance or sensitivity to each drug (Fig. 2). The Wilcoxon rank-sum test and t-test yielded similar results. The results of permutation analyses of gene-probe sets associated with resistance to prednisolone, vincristine, and asparaginase were significant overall (all P<0.05) (Table 3 in the Supplementary Appendix) in the total popula-



tion and within the B-lineage ALL group, whereas those of analyses of gene-probe sets associated with resistance to daunorubicin were significant in the B-lineage ALL group, but not at the level of $P < 0.05$ in the group as a whole. The false discovery rate was higher for daunorubicin than for the other three drugs. For all drugs, the false discovery rates were lower in the B-lineage ALL group than in the total group and highest for daunorubicin (Table 3 in the Supplementary Appendix). Using the top 30, 50, and 100 discriminating genes for each drug yielded predictive accuracies of 67 to 73 percent. For B-lineage ALL, the estimated predictive accuracies were higher, ranging from 71 to 76 percent (Table 5 in the Supplementary Appendix).

SUPERVISED CLUSTERING AND PRINCIPAL-COMPONENT ANALYSIS

The number of genes used to build drug-resistance models for each antileukemic agent was based on the false discovery rate and predictive accuracy (Tables 3, 4, and 5 in the Supplementary Appendix). This determination resulted in 172 probe sets corresponding to 124 unique genes and 28 complementary DNA clones (some genes are represented on the array by multiple probe sets) that were differentially expressed in sensitive and resistant B-lineage ALL. This included 42 gene-probe sets for prednisolone, 59 for vincristine, 54 for asparaginase, and 22 for daunorubicin. Hierarchical clustering with the use of these probe sets correctly as-

Table 1. Multivariate Proportional-Hazards Analysis of the Risk of Relapse.*

Variable	Study Group (N=173)			Validation Cohort (N=98)		
	No. of Patients	Hazard Ratio (95% CI)	P Value	No. of Patients	Hazard Ratio (95% CI)	P Value
Age						
1–10 yr	126	1.0†		62	1.0†	
<1 yr	0			7	6.48 (0.60–69.79)	0.12
>10 yr	47	1.37 (0.68–2.75)	0.37	29	7.61 (1.23–46.95)	0.029
Subtype						
B-lineage	46	1.0†		24	1.0†	
Other						
BCR-ABL	5	2.08 (0.56–7.63)	0.27	8	14.17 (2.59–77.55)	0.002
E2A-PBX1	8	1.75 (0.47–6.52)	0.40	12	0.63 (0.07–6.29)	0.69
MLL-AF4	3	13.63 (2.98–62.27)	0.001	14	1.23 (0.22–6.76)	0.81
TEL-AML1	43	0.13 (0.03–0.58)	0.007	16	1.40 (0.09–21.36)	0.81
Hyperdiploid‡	40	0.34 (0.09–1.24)	0.10	15	0.92 (0.10–8.02)	0.94
T-cell	28	0.63 (0.24–1.66)	0.35	9	4.03 (0.60–26.89)	0.15
White-cell count						
<10,000/mm ³	38	1.0†		25	1.0†	
10,000–49,999/mm ³	68	0.84 (0.30–2.34)	0.74	27	1.93 (0.24–15.15)	0.53
50,000–100,000/mm ³	28	0.97 (0.28–3.35)	0.96	20	12.02 (1.30–110.79)	0.028
>100,000/mm ³	39	2.74 (0.92–8.12)	0.069	26	8.35 (0.84–83.11)	0.070
Combined drug-resistance gene-expression score§						
Sensitive (<4.7)	60	1.0†		29	1.0†	
Intermediate resistance (4.7–5.58)	56	2.58 (0.97–6.87)	0.07	48	4.00 (0.58–27.73)	0.16
Resistant (>5.58)	57	3.0 (1.13–7.96)	0.027	21	11.85 (1.51–93.12)	0.019

* CI denotes confidence interval.

† This group served as the reference group.

‡ Cytogenetic analysis revealed more than 50 chromosomes.

§ Scores reflect the values for prednisolone, vincristine, asparaginase, and daunorubicin.

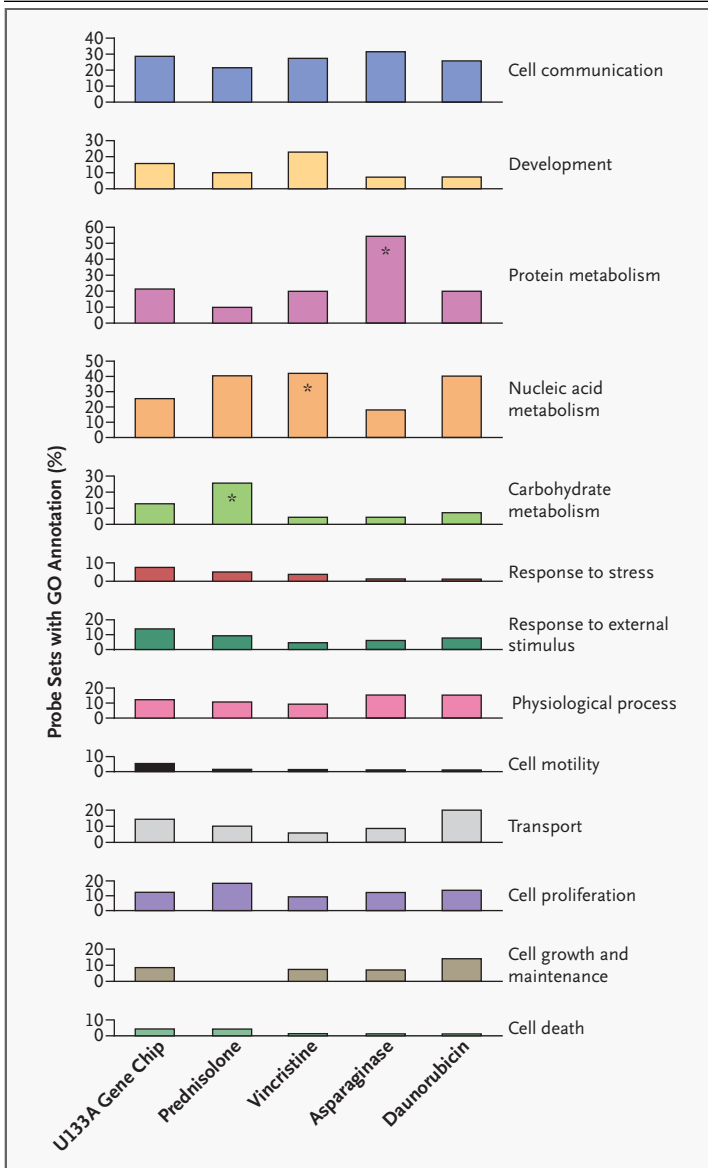


Figure 4. Gene Ontology (GO) Functional Classification of Genes That Discriminated between Drug-Sensitive and Drug-Resistant B-Lineage ALL.

The functional GO classification of genes identified by the probe sets as discriminating B-lineage ALL cells that are resistant to each of the antileukemic agents, as compared with the entire genome as represented by all probe sets on the U133A GeneChip (22,283 probe sets, 12,983 with GO annotation), is shown. For prednisolone, 42 probe sets were found to discriminate between resistant and sensitive ALL cells; for vincristine, 59 such probe sets were identified; for asparaginase, 54 such probe sets were identified; and for daunorubicin, 22 such probe sets were identified. There were 25, 35, 39, and 16 probe sets annotated in the GO database for prednisolone, vincristine, asparaginase, and daunorubicin, respectively. Functional categories that are proportionally overrepresented in the probe sets, as compared with the entire genome, are indicated by an asterisk ($P < 0.05$ by Fisher's exact test).

signed the drug-sensitivity status (as sensitive or resistant) of 66 of 74 cases with respect to prednisolone, 84 of 104 with respect to vincristine, 83 of 106 with respect to asparaginase, and 86 of 105 with respect to daunorubicin (Fig. 2) (Table 4 in the Supplementary Appendix). Similarly, principal-component analyses correctly grouped samples from most patients into the resistant or sensitive cluster for each of the four antileukemic agents (Fig. 2). Hierarchical clustering and principal-component analyses involving all 173 patients gave similar results (Fig. 3 and 4 in the Supplementary Appendix). The probe-set identification, gene names, annotations, and the gene-expression ratio in resistant as compared with sensitive leukemia cells for discriminating genes are shown for each drug in Figures 5, 6, 7, and 8 (B-lineage ALL) and 9, 10, 11, and 12 (B-lineage and T-cell ALL combined) in the Supplementary Appendix.

RESISTANCE GENES, COMBINED GENE-EXPRESSION SCORES, AND TREATMENT OUTCOME

For the 173 patients treated according to the Dutch and German protocols, the median follow-up was 4.2 years; 132 patients remained in continuous complete remission, 40 patients relapsed, and 1 patient had a second cancer, at which time data on this patient were censored. A high combined gene-expression score indicative of resistance to the four drugs was associated with a significantly increased risk of relapse ($P=0.001$) (Fig. 3A). The combined drug-resistance gene-expression score also predicted the outcome of treatment in a multivariate analysis that included the patient's age, ALL genetic subtype, ALL lineage, and leukocyte count at diagnosis (hazard ratio for relapse with a high score as compared with a low score, 3.0; $P=0.027$) (Table 1).

The combined gene-expression score was tested in an independent cohort of 98 U.S. patients who had been treated with these four drugs, but according to a different protocol. The median follow-up of these patients was 7.0 years; 17 patients relapsed, 9 had competing events (7 had second cancers, and remission failed in 2), and 72 remained in continuous complete remission. As in the training set, a high combined drug-resistance gene-expression score was associated with a significantly increased risk of relapse ($P=0.003$) (Fig. 3B). When the patient's age, genetic subtype of ALL, ALL lineage, and leukocyte count at diagnosis were included in a multivariate analysis, a high combined drug-

resistance gene-expression score was independently associated with a higher probability of relapse than was a low score (hazard ratio, 11.85; $P=0.019$) (Table 1).

ONTOLOGY CLASSIFICATION OF DISCRIMINATING GENES

Genes that could be used to identify B-lineage ALL that was resistant to each antileukemic agent were grouped into functional categories according to the Gene Ontology database (Fig. 4). As compared with the entire array, the 42 gene-probe sets related to prednisolone sensitivity had a higher percentage of genes involved in carbohydrate metabolism (25 percent vs. 11 percent, $P=0.039$). As compared with the entire array, the gene-probe sets related to vincristine sensitivity had a higher percentage of genes involved in nucleic acid metabolism (39 percent vs. 23 percent, $P=0.021$), and the gene-probe sets related to asparaginase sensitivity had a higher percentage of protein metabolism genes (53 percent vs. 20 percent, $P<0.001$).

GENES PREVIOUSLY LINKED WITH DRUG RESISTANCE OR PROGNOSIS IN ALL

Of the 124 differentially expressed genes, to our knowledge 121 have not previously been linked to resistance to the four agents investigated. Only three genes for which results were significant in our analyses (*RPL6*, *ARHA*, and *SLC2A14*) have previously been associated with resistance to doxorubicin (*RPL6*³¹ and *ARHA*³²) or vincristine (*SLC2A14*³³). Other genes previously associated with drug resistance or prognosis were not associated with sufficient statistical significance (i.e., $P<0.001$) for inclusion in our models (Tables 4 and 7 in the Supplementary Appendix).

DISCUSSION

We have identified genes that are differentially expressed in ALL cells with resistance to four antileukemic drugs and have shown that the pattern of expression of these genes is related to the outcome of treatment. The expression of 172 gene-probe sets (representing 124 unique known genes and 28 complementary DNA clones) in primary B-lineage leukemia cells was associated with resistance to prednisolone (42 probe sets), vincristine (59 probe sets), asparaginase (54 probe sets), and daunorubicin (22 probe sets). Of these 124 genes, to

our knowledge 121 have not previously been associated with resistance to the four agents. Twelve other genes that have previously been associated with drug resistance or prognosis in ALL were differentially expressed in sensitive and resistant ALL but not at the level required for inclusion in our models ($P<0.001$). No universal cross-resistance gene was identified, since no single gene was associated with resistance to all four drugs. Discriminating genes belong to numerous functional groups, and specific functional categories were significantly overrepresented for some antileukemic agents (Fig. 4). These findings document that resistance to mechanistically distinct antileukemic agents is associated with the expression of different functional groups of genes and support the use of combination chemotherapy for cancer treatment.

Our findings point to previously unrecognized potential targets for new agents to augment the efficacy of current chemotherapy for ALL. For example, in prednisolone-resistant ALL there was overexpression of the antiapoptosis gene *MCL1* and underexpression of several transcription-associated genes (e.g., *SMARCB1*, *PRPF18*, and *CTCF*), in asparaginase-resistant ALL there was overexpression of several ribosomal protein genes (e.g., *RPL3*, *RPL4*, *RPL5*, *RPL6*, and *RPL11*), and in vincristine-resistant ALL there was altered expression of cytoskeleton and extracellular-matrix genes (e.g., *TMSB10*, *PDLIM1*, and *DSC3*). It will be important to determine whether modulation of the proteins encoded by these genes will enhance treatment efficacy in patients with drug-resistant ALL.

It is noteworthy that the gene-expression signatures associated with resistance to individual antileukemic agents were also related to the response to treatment. The robustness of these signatures was validated in an independent population of patients who were treated with these same drugs, but in a different country and according to a different protocol. In a multivariate analysis that included the patient's age, ALL genetic subtype, ALL lineage, and leukocyte count, the combined gene-expression score remained significantly related to the risk of relapse in both the training and validation populations (Table 1). This indicates that the expression of genes associated with drug resistance has an independent influence on the outcome of treatment in ALL. Because genes associated with sensitivity or resistance differ for each antileukemic agent,

our findings point to strategies whereby one could modulate specific components of therapy to which an individual patient is resistant.

Supported in part by grants from the National Institutes of Health (R37 CA36401, R01 CA78224, R01 CA51001, U01 GM61393, and U01 GM61394), a support grant (P30 CA21765) from the National Cancer Institute, the American Cancer Society F.M. Kirby Clinical Research Professorship (to Dr. Pui), the American Lebanese Syrian Associated Charities, the Pediatric Oncology Foundation Rotter-

dam, the Nijbakker-Morra Foundation, and the René Vogels Stipendium 2002 (to Ms. Holleman).

We are indebted to Jessica Gladdines, Sanne Lugthart, John Morris, Michael Shipman, and Mark Wilkinson, as well as to Dr. Clayton Naeve and his staff in the Hartwell Center for Bioinformatics and Biotechnology at St. Jude Children's Research Hospital, for outstanding technical support; to the clinical staff who cared for the patients; to the patients and parents for their participation in these studies; and to Drs. Charles Sherr, John Cleveland, and James Downing for providing critical feedback that helped shape the manuscript.

REFERENCES

- Pui CH, Evans WE. Acute lymphoblastic leukemia. *N Engl J Med* 1998;339:605-15.
- Pui CH, Campana D, Evans WE. Childhood acute lymphoblastic leukaemia — current status and future perspectives. *Lancet Oncol* 2001;2:597-607.
- Den Boer ML, Harms DO, Pieters R, et al. Patient stratification based on prednisolone-vincristine-asparaginase resistance profiles in children with acute lymphoblastic leukemia. *J Clin Oncol* 2003;21:3262-8.
- Kaspers GJ, Veerman AJ, Pieters R, et al. In vitro cellular drug resistance and prognosis in newly diagnosed childhood acute lymphoblastic leukemia. *Blood* 1997;90:2723-9.
- Pieters R, Huismans DR, Loonen AH, et al. Relation of cellular drug resistance to long-term clinical outcome in childhood acute lymphoblastic leukaemia. *Lancet* 1991;338:399-403.
- van den Heuvel-Eibrink MM, Sonneveld P, Pieters R. The prognostic significance of membrane transport-associated multidrug resistance (MDR) proteins in leukemia. *Int J Clin Pharmacol Ther* 2000;38:94-110.
- Gu L, Findley HW, Zhou M. MDM2 induces NF-kappaB/p65 expression transcriptionally through Sp1-binding sites: a novel, p53-independent role of MDM2 in doxorubicin resistance in acute lymphoblastic leukemia. *Blood* 2002;99:3367-75.
- Ramakers-van Woerden NL, Pieters R, Slater RM, et al. In vitro drug resistance and prognostic impact of p16INK4A/P15INK4B deletions in childhood T-cell acute lymphoblastic leukaemia. *Br J Haematol* 2001;112:680-90.
- Krajinovic M, Labuda D, Mathonnet G, et al. Polymorphisms in genes encoding drugs and xenobiotic metabolizing enzymes, DNA repair enzymes, and response to treatment of childhood acute lymphoblastic leukemia. *Clin Cancer Res* 2002;8:802-10.
- Kearns PR, Pieters R, Rottier MM, Pearson AD, Hall AG. Raised blast glutathione levels are associated with an increased risk of relapse in childhood acute lymphocytic leukemia. *Blood* 2001;97:393-8.
- McLeod HL, Krynetski EY, Relling MV, Evans WE. Genetic polymorphism of thiopurine methyltransferase and its clinical relevance for childhood acute lymphoblastic leukemia. *Leukemia* 2000;14:567-72.
- Prokop A, Wieder T, Sturm I, et al. Relapse in childhood acute lymphoblastic leukemia is associated with a decrease of the Bax/Bcl-2 ratio and loss of spontaneous caspase-3 processing in vivo. *Leukemia* 2000;14:1606-13.
- Ramakers-van Woerden NL, Pieters R, Loonen AH, et al. TEL/AML1 gene fusion is related to in vitro drug sensitivity for L-asparaginase in childhood acute lymphoblastic leukemia. *Blood* 2000;96:1094-9.
- Hongo T, Okada S, Inoue N, et al. Two groups of Philadelphia chromosome-positive childhood acute lymphoblastic leukemia classified by pretreatment multidrug sensitivity or resistance in in vitro testing. *Int J Hematol* 2002;76:251-9.
- Stam RW, den Boer ML, Meijerink JP, et al. Differential mRNA expression of Ara-C-metabolizing enzymes explains Ara-C sensitivity in MLL gene-rearranged infant acute lymphoblastic leukemia. *Blood* 2003;101:1270-6.
- Evans WE, Relling MV. Pharmacogenomics: translating functional genomics into rational therapeutics. *Science* 1999;286:487-91.
- Idem*. Moving toward individualized medicine with pharmacogenomics. *Nature* 2004;429:464-8.
- Pieters R, den Boer ML. Molecular pharmacodynamics in childhood leukemia. *Int J Hematol* 2003;78:402-13.
- Staunton JE, Slonim DK, Collier HA, et al. Chemosensitivity prediction by transcriptional profiling. *Proc Natl Acad Sci U S A* 2001;98:10787-92.
- Weinstein JN, Myers TG, O'Connor PM, et al. An information-intensive approach to the molecular pharmacology of cancer. *Science* 1997;275:343-9.
- Zembutsu H, Ohnishi Y, Tsunoda T, et al. Genome-wide cDNA microarray screening to correlate gene expression profiles with sensitivity of 85 human cancer xenografts to anticancer drugs. *Cancer Res* 2002;62:518-27.
- Armstrong SA, Staunton JE, Silverman LB, et al. MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat Genet* 2002;30:41-7.
- Cheok MH, Yang W, Pui CH, et al. Treatment-specific changes in gene expression discriminate in vivo drug response in human leukemia cells. *Nat Genet* 2003;34:85-90. [Erratum, *Nat Genet* 2003;34:231.]
- Yeoh EJ, Ross ME, Shurtleff SA, et al. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell* 2002;1:133-43.
- Ross ME, Zhou X, Song G, et al. Classification of pediatric acute lymphoblastic leukemia by gene expression profiling. *Blood* 2003;102:2951-9.
- Harms DO, Gobel U, Spaar HJ, et al. Thioguanine offers no advantage over mercaptopurine in maintenance treatment of childhood ALL: results of the randomized trial COALL-92. *Blood* 2003;102:2736-40.
- Pui CH, Sandlund JT, Pei D, et al. Results of therapy for acute lymphoblastic leukemia in black and white children. *JAMA* 2003;290:2001-7.
- Storey JD, Tibshirani R. Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A* 2003;100:9440-5.
- Dudoit S, Fridlyand J. Bagging to improve the accuracy of a clustering procedure. *Bioinformatics* 2003;19:1090-9.
- Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc* 1999;94:496-509.
- Du JP, Jin XH, Shi YQ, et al. Differential expression RPL6/Taxreb107 in drug resistant gastric cancer cell line SGC7901/ADR and its correlation with multiple-drug resistance. *Zhonghua Zhong Liu Za Zhi* 2003;25:21-5. (In Chinese.)
- Van Aelst L, D'Souza-Schorey C. Rho GTPases and signaling networks. *Genes Dev* 1997;11:2295-322.
- Martell RL, Slapak CA, Levy SB. Effect of glucose transport inhibitors on vincristine efflux in multidrug-resistant murine erythroleukaemia cells overexpressing the multidrug resistance-associated protein (MRP) and two glucose transport proteins, GLUT1 and GLUT3. *Br J Cancer* 1997;75:161-8.

Copyright © 2004 Massachusetts Medical Society.