

*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.\*

## Supplementary Appendix 1

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Holleman A, Cheok MH, den Boer ML, et al. Gene-expression patterns in drug resistant acute lymphoblastic leukemia cells and response to treatment. N Engl J Med 2004;351:533-42.

## Table of Content

Supplemental Figure 1: Unsupervised hierarchical clustering discriminating drug resistant and drug sensitive ALL patients.....	5
Supplemental Figure 2: Data analysis flowchart. ....	6
Supplemental Figure 3: Supervised hierarchical clustering discriminating drug resistant and drug sensitive ALL patients (B- and T-lineage ALL). ....	11
Supplemental Figure 4: Principal component analysis of drug resistant and sensitive ALL samples for the four antileukemic agents (B- and T-lineage ALL). ....	12
Supplemental Figure 5: Supervised hierarchical clustering discriminating prednisolone resistant and sensitive B-lineage ALL.....	13
Supplemental Figure 6: Supervised hierarchical clustering discriminating vincristine resistant and sensitive B-lineage ALL. ....	13
Supplemental Figure 7: Supervised hierarchical clustering discriminating L-asparaginase resistant and sensitive B-lineage ALL.....	14
Supplemental Figure 8: Supervised hierarchical clustering discriminating daunorubicin resistant and sensitive B-lineage ALL.....	14
Supplemental Figure 9: Supervised hierarchical clustering discriminating prednisolone resistant and sensitive ALL (B- and T-lineage ALL). ....	15
Supplemental Figure 10: Supervised hierarchical clustering discriminating vincristine resistant and sensitive ALL (B- and T-lineage ALL). ....	15
Supplemental Figure 11: Supervised hierarchical clustering discriminating L-asparaginase resistant and sensitive ALL (B- and T-lineage ALL). ....	16
Supplemental Figure 12: Supervised hierarchical clustering discriminating daunorubicin resistant and sensitive ALL (B- and T-lineage ALL). ....	16
Supplemental Figure 13: Gene Ontology classification of genes discriminating drug resistance in patients with acute lymphoblastic leukemia (B- and T-lineage ALL). ....	17

Supplemental Figure 14: Pathway analysis of genes associated with L-asparaginase resistance in B-lineage ALL. .... 19

Supplemental Table 1: LC<sub>50</sub> values for classification of resistant and sensitive ALL for each chemotherapeutic agent. .... 4

Supplemental Table 2: Patient characteristics. .... 4

Supplemental Table 3: Permutation analysis and false discovery rate of probe sets selected by Wilcoxon rank sum test and t-test. .... 6

Supplemental Table 4: Number of significant probe sets discriminating drug resistance for each individual antileukemic agent. .... 8

Supplemental Table 5: Prediction accuracy using gene expression profiles for classification of drug resistant and sensitive acute lymphoblastic leukemia. .... 9

Supplemental Table 6: Gene expression scores for the intermediate sensitivity group, using genes selected to discriminate resistant and sensitive B-lineage ALL. .... 10

Supplemental Table 7: Genes previously linked to drug resistance or prognosis in ALL. .... 18

Supplemental Table 8: Multivariate analysis of gene expression and known prognostic factors (age, WBC count) to discriminate in vitro drug resistance. .... 20

Supplemental Table 9: Association of the LC50 scores with gene expression scores and other known prognostic variables. .... 20

Supplemental Table 10: Association of the drug resistance gene expression scores and minimal residual disease (MRD). .... 21

**Supplemental Table 1: LC<sub>50</sub> values for classification of resistant and sensitive ALL for each chemotherapeutic agent.**

LC<sub>50</sub> by MTT as described in Pieters and Kaspers et al<sup>1,2</sup>. Classification based on LC<sub>50</sub>-values previously associated with treatment outcome, as described in den Boer et al<sup>3</sup>.

<u>Drug</u>	<u>Sensitive</u>	<u>Resistant</u>
<b>Prednisolone</b>	≤ 0.100µg/ml	≥ 150µg/ml
<b>Vincristine</b>	≤ 0.391µg/ml	≥ 1.758µg/ml
<b>Asparaginase</b>	≤ 0.033IU/ml	≥ 0.912IU/ml
<b>Daunorubicin</b>	≤ 0.075µg/ml	≥ 0.114µg/ml

**Supplemental Table 2: Patient characteristics.**

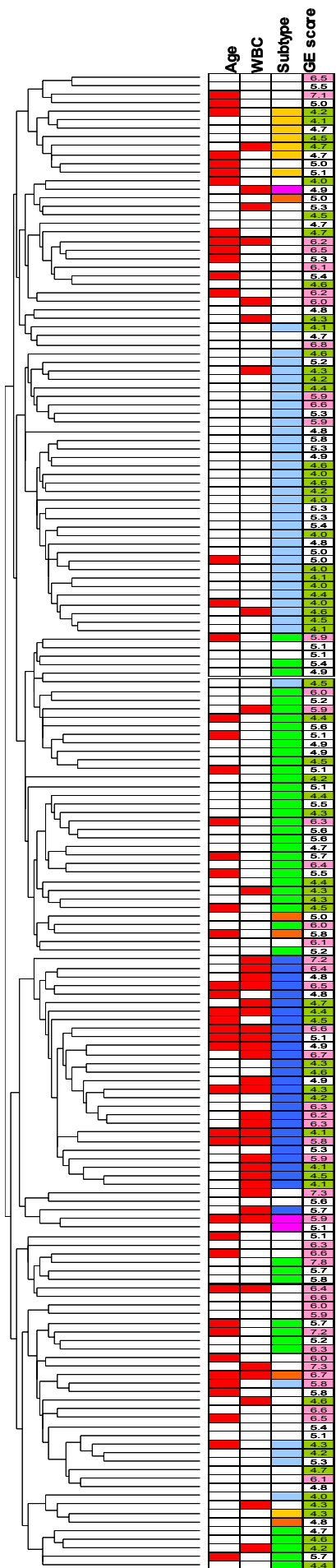
Included were 173 patients categorized as resistant or sensitive to one or more of the four drugs. The sample size and the distribution of known prognostic factors (i.e., white blood cell count (WBC) and age at diagnosis; gender, and lineage) for each antileukemic agent are shown for (a) all patients and (b) the B-lineage ALL patients.

<b>a.</b>		<b>N</b>	<b>Age at diagnosis</b> [median in years]	<b>Gender</b>	<b>WBC at diagnosis</b> [10 <sup>9</sup> /L]	<b>Lineage</b>
<b>Prednisolone</b>	Sensitive	66	5.9	36M 30F	43.0	54B 12T
	Resistant	27	8.0	17M 10F	19.9	20B 7T
			P=0.007*	P=0.497 <sup>‡</sup>	P=0.314*	P=0.408 <sup>‡</sup>
<b>Vincristine</b>	Sensitive	89	6.0	54M 35F	34.7	75B 14T
	Resistant	36	6.7	20M 16F	15.4	29B 7T
			P=0.651*	P=0.689 <sup>‡</sup>	P=0.172*	P=0.606 <sup>‡</sup>
<b>asparaginase</b>	Sensitive	81	4.4	47M 34F	30.0	66B 15T
	Resistant	45	7.3	24M 21F	28.3	40B 5T
			P=0.004*	P=0.708 <sup>‡</sup>	P=0.729*	P=0.320 <sup>‡</sup>
<b>Daunorubicin</b>	Sensitive	94	4.6	57M 37F	40.0	82B 12T
	Resistant	28	7.7	17M 11F	17.6	23B 5T
			P=0.031*	P=1.0 <sup>‡</sup>	P=0.378*	P=0.537 <sup>‡</sup>

<b>b.</b>		<b>N</b>	<b>Age at diagnosis</b> [median in years]	<b>Gender</b>	<b>WBC at diagnosis</b> [10 <sup>9</sup> /L]
<b>Prednisolone</b>	Sensitive	54	5.8	28M 26F	33.1
	Resistant	20	8.9	12M 8F	8.3
			P=0.016*	P=0.605 <sup>‡</sup>	P=0.012*
<b>Vincristine</b>	Sensitive	75	6.0	46M 29F	25.6
	Resistant	29	6.3	15M 14F	9.9
			P=0.888*	P=0.358 <sup>‡</sup>	P=0.016*
<b>asparaginase</b>	Sensitive	66	4.2	38M 28F	23.2
	Resistant	40	7.3	21M 19F	18.5
			P=0.003*	P=0.688 <sup>‡</sup>	P=0.732*
<b>Daunorubicin</b>	Sensitive	82	4.4	49M 33F	26.8
	Resistant	23	7.2	13M 10F	16.3
			P=0.117*	P=0.814 <sup>‡</sup>	P=0.275*

M=male, F=female; P-values determined by \*Wilcoxon rank sum test <sup>‡</sup>Fisher's exact test

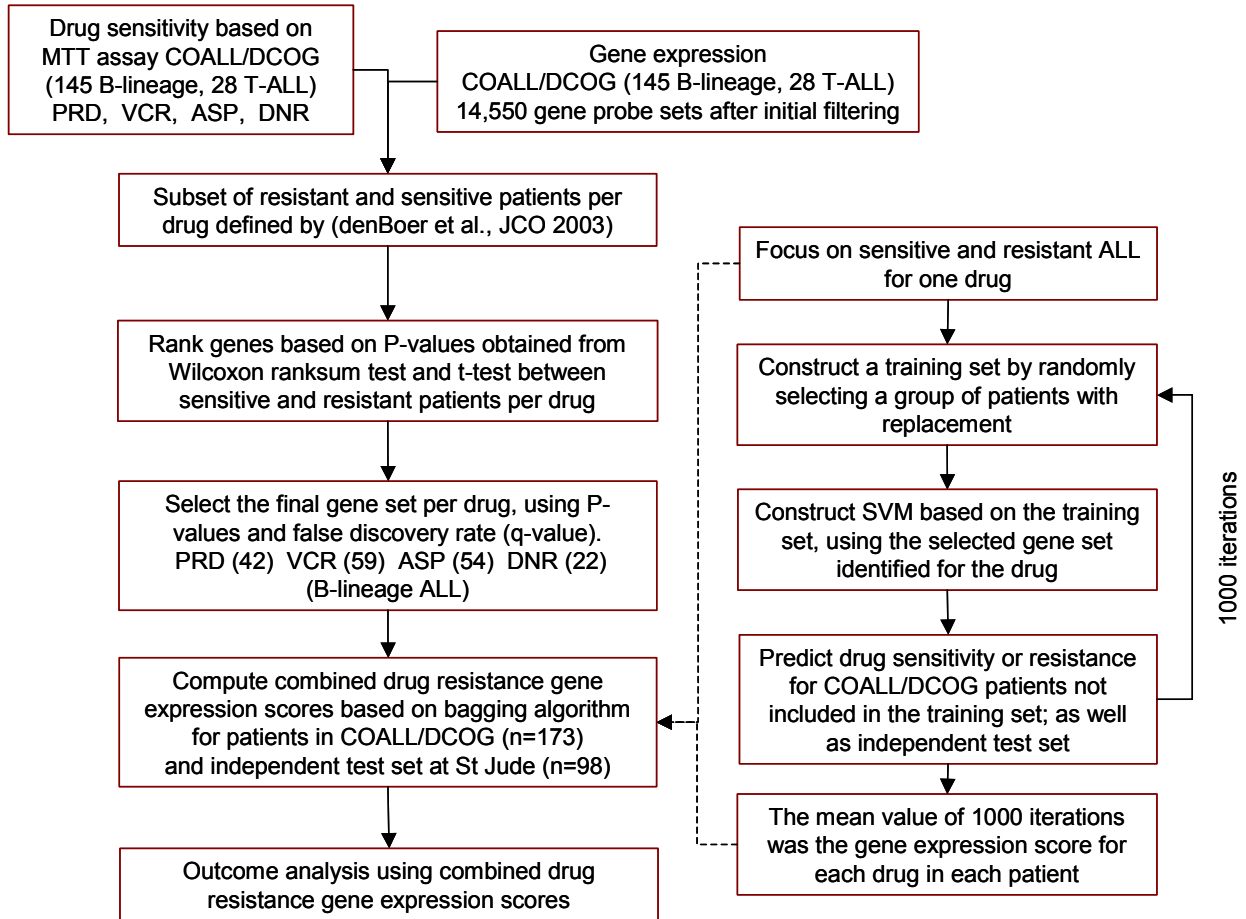


**Supplemental Figure 1: Unsupervised hierarchical clustering discriminating drug resistant and drug sensitive ALL patients.**

Unsupervised hierarchical clustering of 173 patients was performed using 14,550 probe sets. Age (red= >10 years), white blood cell count (WBC, red= >100 x10<sup>9</sup>/L), genetic or lineage subtype (see legend) and the combined drug resistance gene expression (GE) score for prednisolone, vincristine, asparaginase and daunorubicin are shown for each patient (high/resistant in red= GE score>5.58 and low/sensitive green= GE score<4.70). Unsupervised clustering is largely driven by ALL subtype, e.g., T- and B-lineage ALL. None of the other known prognostic features (patient age, WBC) clustered in an unsupervised analysis, as was also the case for the drug resistance gene expression score.

**Legend**

- Age
  - >10 years
  - <10 years
  
- WBC
  - >100 x10<sup>9</sup>/L
  - <100 x10<sup>9</sup>/L
  
- Subtype
  - BCRABL
  - E2APBX1
  - MLL rearranged
  - T-ALL
  - TELAML1
  - hyperdiploid (HD)
  - B-ALL other
  
- GE score
  - <4.70
  - >5.58



**Supplemental Figure 2: Data analysis flowchart.**

**Supplemental Table 3: Permutation analysis and false discovery rate of probe sets selected by Wilcoxon rank sum test and t-test.**

Permutation analysis (N=1000) was computed for each dataset (prednisolone (PRED), vincristine (VCR), asparaginase (ASP), daunorubicin (DNR) using (a) all patients and (b) only patients with B-lineage acute lymphoblastic leukemia. For each P-value ( $\alpha$ ) using Wilcoxon rank sum test, the number of probe sets (N), the false discovery rate (FDR) and the P-value are listed. Numbers in red indicate the subset of genes chosen for further analyses (Supplemental Table 4) based on a low P-value, an FDR less than 10 percent when possible,

and between 25 and 100 genes for each drug. Similar results were obtained with probe sets selected by t-test using (c) all patients and (d) only patients with B-lineage.

**a. All patients**

<i>drug</i>	$\alpha = 0.0001$			$\alpha = 0.0005$			$\alpha = 0.001$		
	N	FDR [%]	P-value	N	FDR [%]	P-value	N	FDR [%]	P-value
<b>PRED</b>	11	10	0.021	32	18	0.017	<b>53</b>	<b>22</b>	<b>0.019</b>
<b>VCR</b>	7	13	0.028	76	15	<0.001	<b>76</b>	<b>16</b>	<b>&lt;0.001</b>
<b>ASP</b>	24	5	0.007	<b>91</b>	<b>6</b>	<b>&lt;0.001</b>	135	9	<0.001
<b>DNR</b>	2	22	0.060	11	45	0.080	<b>27</b>	<b>45</b>	<b>0.065</b>

**b. B-lineage**

<i>drug</i>	$\alpha = 0.0001$			$\alpha = 0.0005$			$\alpha = 0.001$		
	N	FDR [%]	P-value	N	FDR [%]	P-value	N	FDR [%]	P-value
<b>PRED</b>	19	6	0.007	<b>57</b>	<b>10</b>	<b>0.002</b>	92	13	0.003
<b>VCR</b>	22	6	0.004	<b>74</b>	<b>8</b>	<b>&lt;0.001</b>	138	10	<0.001
<b>ASP</b>	<b>67</b>	<b>2</b>	<b>0.002</b>	202	3	<0.001	279	4	<0.001
<b>DNR</b>	5	22	0.029	25	26	0.011	<b>38</b>	<b>32</b>	<b>0.014</b>

**c. All patients**

<i>drug</i>	$\alpha = 0.0001$			$\alpha = 0.0005$			$\alpha = 0.001$		
	N	FDR [%]	P-value	N	FDR [%]	P-value	N	FDR [%]	P-value
<b>PRED</b>	10	12	0.058	33	17	0.021	<b>49</b>	<b>24</b>	<b>0.034</b>
<b>VCR</b>	7	8	0.032	55	11	0.012	<b>91</b>	<b>13</b>	<b>0.012</b>
<b>ASP</b>	21	5	0.028	<b>84</b>	<b>7</b>	<b>&lt;0.001</b>	120	9	<0.001
<b>DNR</b>	3	29	0.556	19	34	0.188	<b>34</b>	<b>37</b>	<b>0.164</b>

**d. B-lineage**

<i>drug</i>	$\alpha = 0.0001$			$\alpha = 0.0005$			$\alpha = 0.001$		
	N	FDR [%]	P-value	N	FDR [%]	P-value	N	FDR [%]	P-value
<b>PRED</b>	26	4	0.014	<b>69</b>	<b>8</b>	<b>0.003</b>	98	12	0.003
<b>VCR</b>	33	3	0.016	<b>101</b>	<b>5</b>	<b>&lt;0.001</b>	166	7	<0.001
<b>ASP</b>	<b>71</b>	<b>2</b>	<b>&lt;0.001</b>	184	3	<0.001	269	4	<0.001
<b>DNR</b>	8	12	0.140	34	18	0.043	<b>50</b>	<b>26</b>	<b>0.068</b>

In each random permutation, the class label (resistant or sensitive) was randomly assigned to each patient and genes were reselected using Wilcoxon rank sum test and t-test based on the random labels. The overall significance ( $P_\alpha$ ) of the model was estimated using the following formula:

$$P_\alpha = \frac{\text{number of permutations with } q\text{-value}_\alpha^{\text{random}} \leq q\text{-value}_\alpha^{\text{obs}}}{\text{Total number of permutations}}$$

Where  $\alpha$  is the P-value using Wilcoxon rank sum test or t-test;  $q\text{-value}_\alpha^{\text{random}}$  is estimated q-value given  $\alpha$  using random class label;  $q\text{-value}_\alpha^{\text{obs}}$  is the observed q-value given  $\alpha$  using the observed class label.

**Supplemental Table 4: Number of significant probe sets discriminating drug resistance for each individual antileukemic agent.**

Probe sets were rank-ordered according to their P-values, with the smallest P-values indicating the strongest statistical difference between resistant and sensitive patients. Among predefined significance levels ( $\alpha$ ), we chose the cut-offs that provided a low FDR rate (less than 10 percent when possible), high prediction accuracy and we limited the number of genes so that they did not exceed the number of samples, to facilitate future statistical analyses. The one exception was DNR ( $\alpha = 0.001$ , FDR=32%), because the FDR was never less than 10 percent. The final genes selected were required to be significant by both Wilcoxon rank sum test and t-test, using the selected P-value cut-offs (indicated in red in Supplemental Table 3). Listed below are the number of probe sets with a P-value lower than or equal to the value indicated ( $\alpha$ ), determined by Wilcoxon rank sum test and t-test, which yielded above for (a) all samples, including 153 probe sets (146 unique) and 107 known unique genes and 29 cDNA clones were selected by both methods for all four drugs; and (b) only the precursor-B lineage ALL, including 177 probe sets (172 unique) and 124 known unique genes and 28 cDNA clones were selected by both methods. Indicated in red are the numbers of top gene probes that were used to build the final drug resistance models.

**a. All patients**

	Wicoxon rank sum test	t-test	both tests
<b>PRED</b>	53*	49*	32*
<b>VCR</b>	76*	91*	54*
<b>ASP</b>	91 <sup>†</sup>	81 <sup>†</sup>	52 <sup>†</sup>
<b>DNR</b>	27*	34*	15*
<b>sum</b>			153

**b. B-lineage**

	Wicoxon rank sum test	t-test	both tests
	57 <sup>†</sup>	69 <sup>†</sup>	42 <sup>†</sup>
	74 <sup>†</sup>	101 <sup>†</sup>	59 <sup>†</sup>
	67 <sup>‡</sup>	71 <sup>‡</sup>	54 <sup>‡</sup>
	38*	50*	22*
			177

<sup>‡</sup> $\alpha < 0.0001$ , <sup>†</sup> $\alpha < 0.0005$ , \* $\alpha < 0.001$

**Supplemental Table 5: Prediction accuracy using gene expression profiles for classification of drug resistant and sensitive acute lymphoblastic leukemia.**

Prediction accuracy in randomly selected cohorts comprising 1/3 of the training set for each antileukemic agent using 30, 50 or 100 reselected top probe sets. The median prediction accuracy is shown with corresponding P-values and the 95 percent confidence interval (C.I.) for prednisolone (PRED), vincristine (VCR), asparaginase (ASP) and daunorubicin (DNR) (a) based on all patients and (b) for only patients with B-lineage acute lymphoblastic leukemia. The prediction accuracies were even higher (c, d), when genes were selected only once from the entire training set, and then used to predict sensitive or resistant phenotypes in randomly selected cohorts comprising 1/3 of the population.

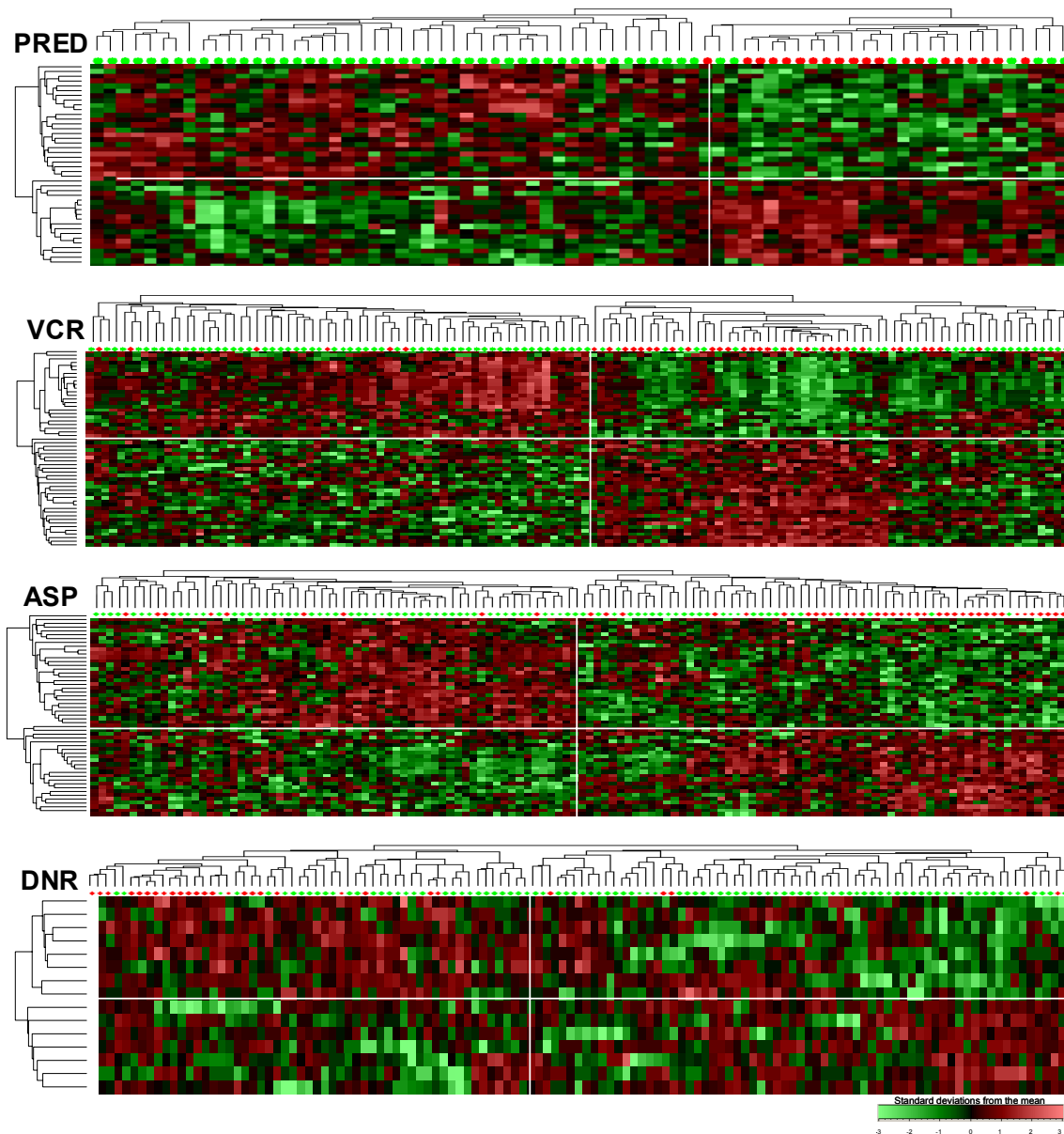
<b>a. All patients</b>					<b>b. B-lineage</b>		
	<b>N</b>	<b>accuracy [%]</b>	<b>95% C.I.</b>	<b>P-value</b>	<b>accuracy [%]</b>	<b>95% C.I.</b>	<b>P-value</b>
<b>PRED</b>	30	71	58-81	0.031	75	63-88	0.025
	50	71	58-81	0.031	75	63-88	0.025
	100	71	58-81	0.031	75	67-87	0.025
<b>VCR</b>	30	68	59-78	0.029	74	62-85	0.010
	50	71	59-79	0.012	76	65-85	0.004
	100	71	59-79	0.012	76	65-85	0.004
<b>ASP</b>	30	67	52-76	0.045	71	60-83	0.018
	50	67	52-79	0.045	71	57-83	0.018
	100	67	55-78	0.045	71	60-80	0.018
<b>DNR</b>	30	73	60-80	0.007	76	65-85	0.004
	50	73	60-80	0.007	76	65-85	0.004
	100	73	58-76	0.007	76	67-84	0.004

<b>c. All patients</b>				<b>d. B-lineage</b>	
	<b>N</b>	<b>accuracy [%]</b>	<b>95% C.I.</b>	<b>accuracy [%]</b>	<b>95% C.I.</b>
<b>PRED</b>	30	95	88-100	98	90-100
	50	93	86-100	98	90-100
<b>VCR</b>	30	81	71-88	81	71-88
	50	81	74-90	81	71-91
<b>ASP</b>	30	81	71-90	83	71-90
	50	83	71-91	86	74-93
<b>DNR</b>	30	90	83-95	95	90-100
	50	90	86-98	95	90-100

**Supplemental Table 6: Gene expression scores for the intermediate sensitivity group, using genes selected to discriminate resistant and sensitive B-lineage ALL.**

We randomly split the resistant and sensitive patients into 2/3 training set and 1/3 test set. Additionally, all patients with an intermediate drug sensitivity phenotype were included in the test set. In each random split, probe sets were reselected based on the training set. A patient in the test set was assigned a score of 1 if classified by the model as sensitive and 2 if resistant; no score was assigned to patients in the training set. The above was repeated 1000 times to compute a gene expression score for each patient as the average of scores assigned to the patient. Median gene expression scores of resistant (R), sensitive (S) and intermediate (I) samples as well as the range in parentheses are presented based on 50 and 100 probe sets. The P-values are given for each pair-wise comparison, using Wilcoxon rank sum test. Patients with intermediate drug sensitivity had median scores that were between the median gene expression score of the drug sensitive and drug resistant groups for all four antileukemic agents. For asparaginase and prednisolone, the gene expression scores of the intermediate group were significantly different from both the sensitive group and the resistant group ( $P < 0.05$ , Wilcoxon rank sum test). For daunorubicin and vincristine, the intermediate group was significantly different from the sensitive group, but not from the resistant group.

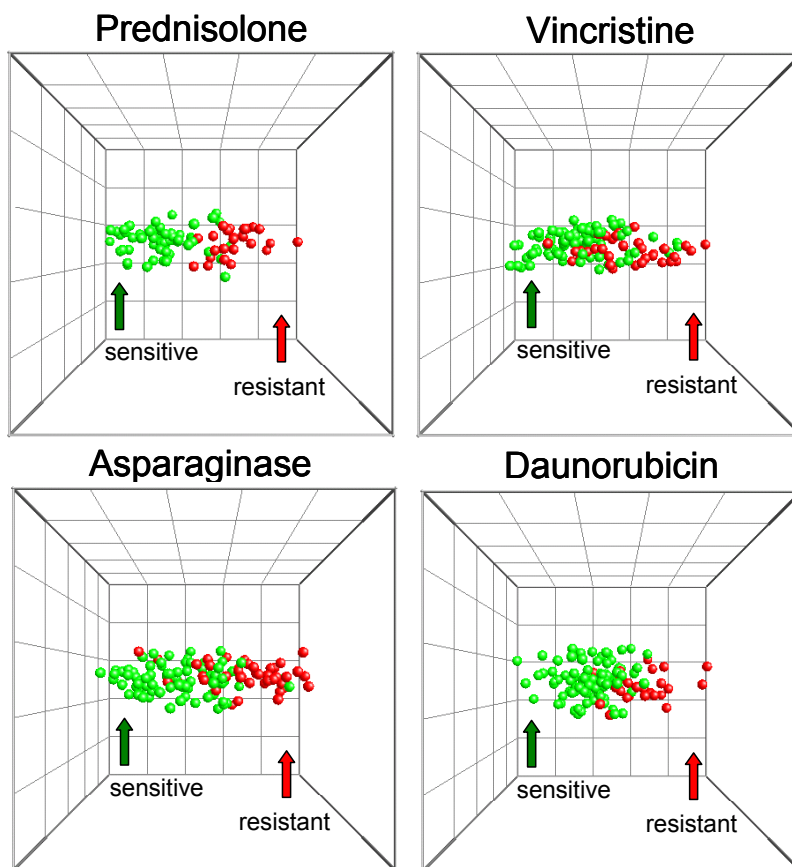
	N	Score Resistant (R)	Score Intermediate (I)	Score Sensitive (S)	P-value R vs. S	P-value R vs. I	P-value I vs. S
<b>PRED</b>	50	1.28 (1.01-1.96)	1.04 (1-1.88)	1.02 (1-1.65)	<0.001	0.002	0.044
	100	1.29 (1-1.99)	1.01 (1-1.92)	1.01 (1-1.72)	<0.001	0.001	0.035
<b>VCR</b>	50	1.28 (1.01-1.88)	1.29 (1.01-1.61)	1.11 (1.01-1.9)	0.006	0.659	0.002
	100	1.26 (1-1.93)	1.24 (1-1.68)	1.08 (1-1.94)	0.002	0.518	0.002
<b>ASP</b>	50	1.53 (1.03-1.91)	1.31 (1.02-1.88)	1.14 (1.01-1.97)	<0.001	0.013	0.004
	100	1.51 (1.02-1.91)	1.31 (1-1.92)	1.1 (1-1.98)	<0.001	0.024	0.008
<b>DNR</b>	50	1.16 (1-1.5)	1.11 (1-1.71)	1.04 (1-1.6)	0.021	0.404	0.09
	100	1.15 (1-1.47)	1.07 (1-1.81)	1.01 (1-1.56)	0.019	0.583	0.045



**Supplemental Figure 3: Supervised hierarchical clustering discriminating drug resistant and drug sensitive ALL patients (B- and T-lineage ALL).**

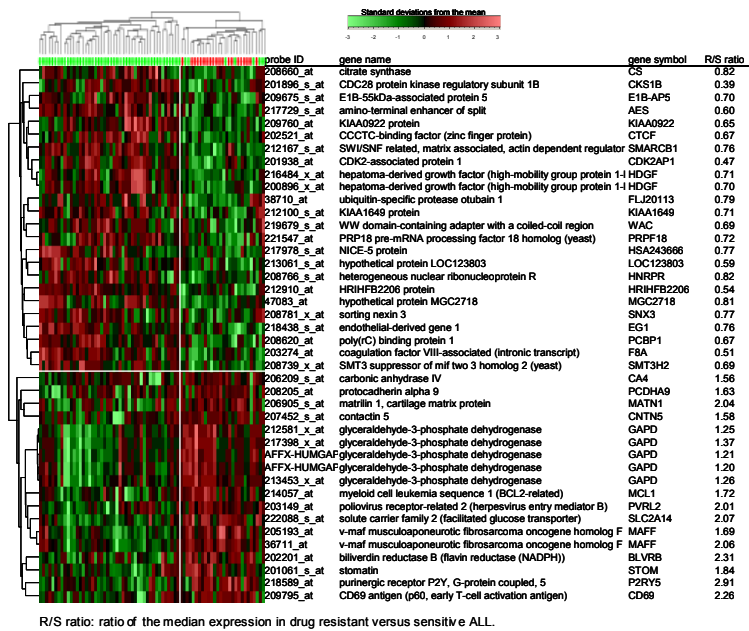
Genes that were significant ( $P \leq 0.001$ ) by both Wilcoxon rank sum test and t-test, were used for hierarchical clustering of all 173 patients, according to sensitivity to prednisolone (PRED), vincristine (VCR), asparaginase (ASP) and daunorubicin (DNR). Each column represents an ALL sample, labeled according to drug sensitivity (red= resistant, green= sensitive) and each row depicts a gene probe set. The heat maps indicate high (red) or low (green) levels of expression relative to the mean, according to the scale shown Fig 2. For PRED, 32

significant probe sets (26 genes, 6 ESTs), for VCR 54 probe sets (37 genes, 12 ESTs), for ASP 52 probe sets (35 genes, 14 ESTs) and for DNR 15 probe sets (13 genes, 2 ESTs) were found to significantly discriminate resistant from sensitive ALL samples.



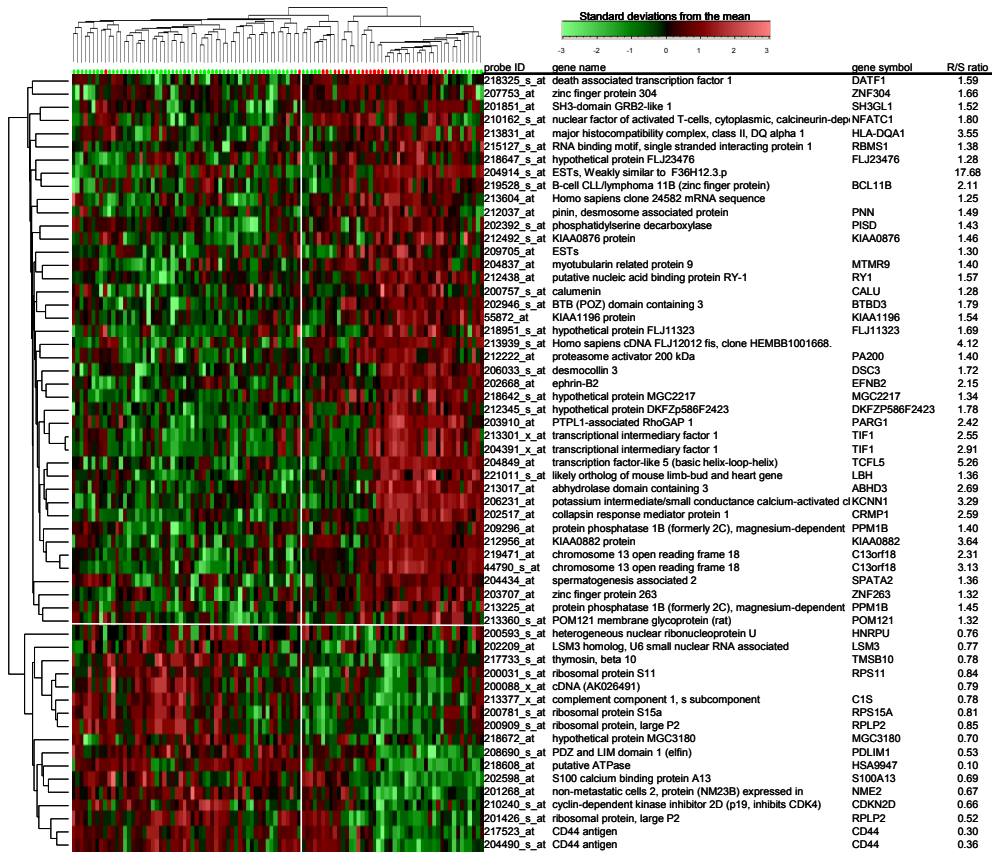
**Supplemental Figure 4: Principal component analysis of drug resistant and sensitive ALL samples for the four antileukemic agents (B- and T-lineage ALL).**

Illustrated are the top three principal components based on the significant discriminating genes for prednisolone (PRED), vincristine (VCR), asparaginase (ASP) and daunorubicin (DNR) in all 173 patients. Each sphere represents an ALL patient, labeled according to drug sensitivity (red= resistant, green= sensitive). Axes are represented in the same units as indicated by the gridlines.



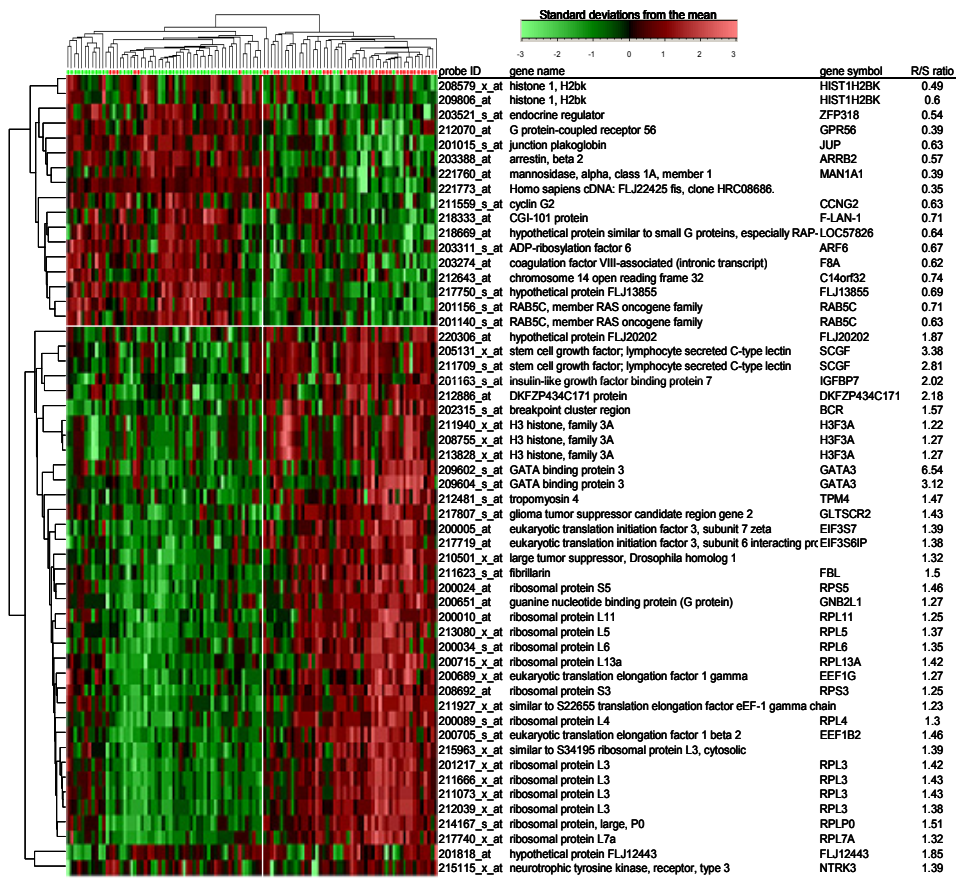
R/S ratio: ratio of the median expression in drug resistant versus sensitive ALL.

**Supplemental Figure 5: Supervised hierarchical clustering discriminating prednisolone resistant and sensitive B-lineage ALL.**



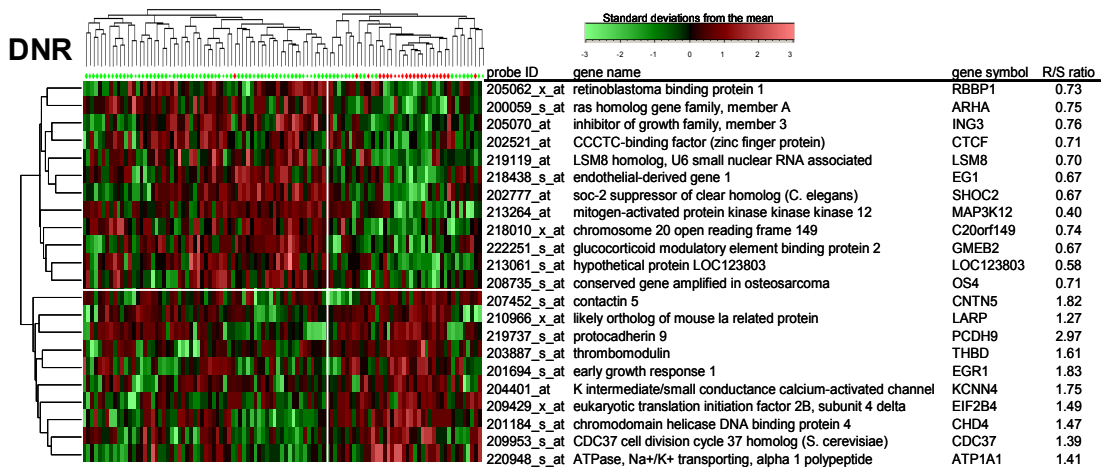
R/S ratio: ratio of the median expression in drug resistant versus sensitive ALL.

**Supplemental Figure 6: Supervised hierarchical clustering discriminating vincristine resistant and sensitive B-lineage ALL.**

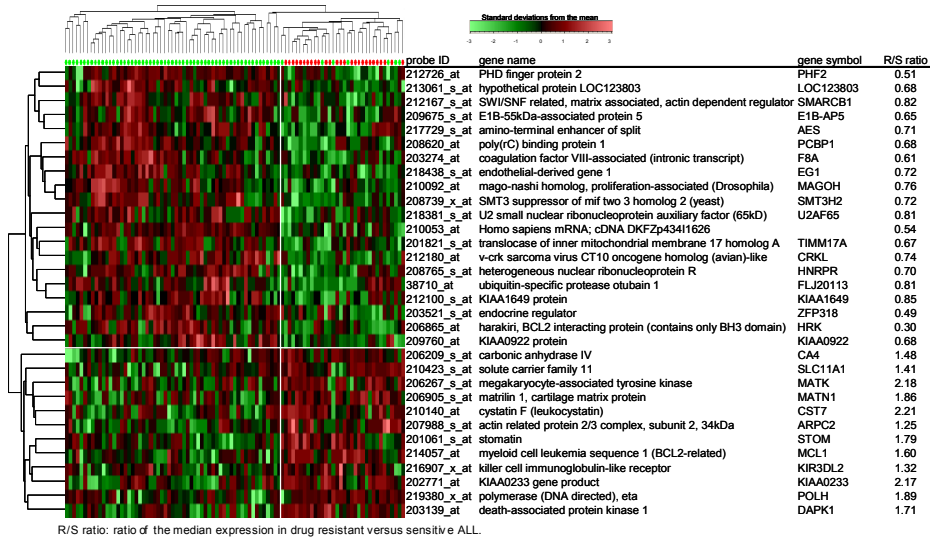


R/S ratio: ratio of the median expression in drug resistant versus sensitive ALL.

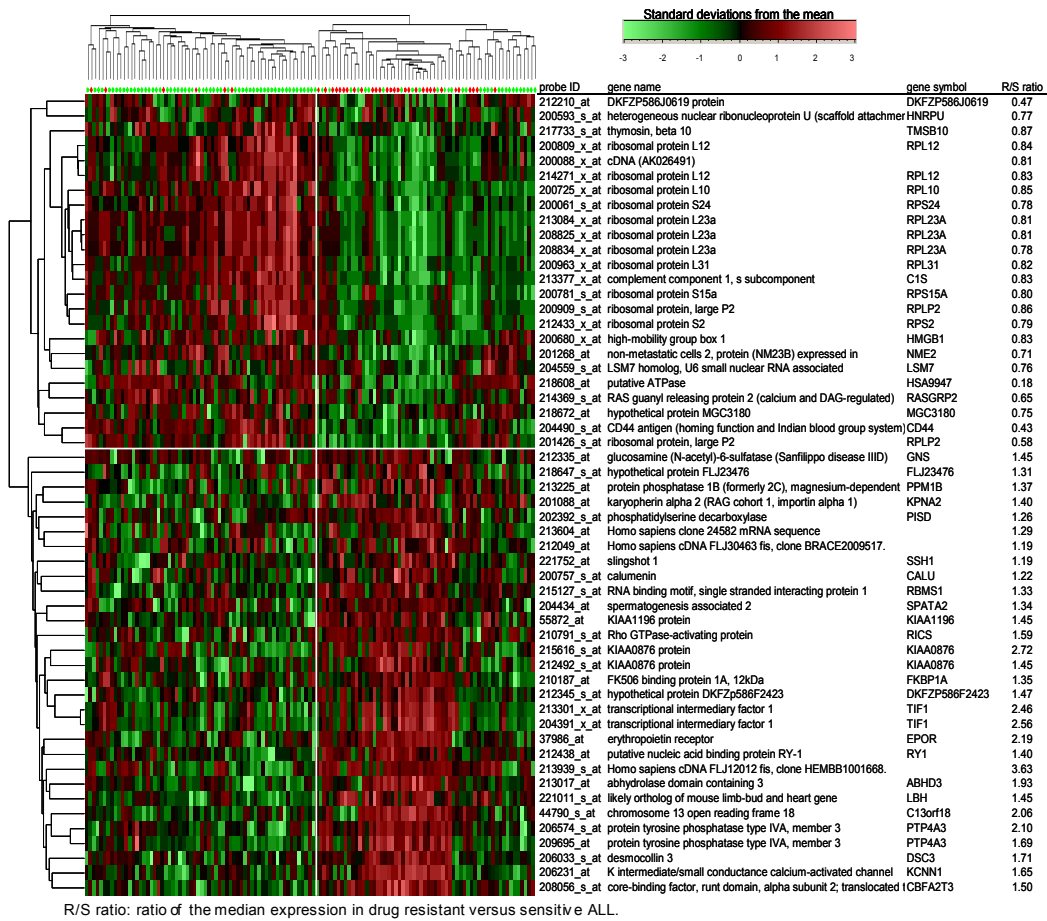
**Supplemental Figure 7: Supervised hierarchical clustering discriminating asparaginase resistant and sensitive B-lineage ALL.**



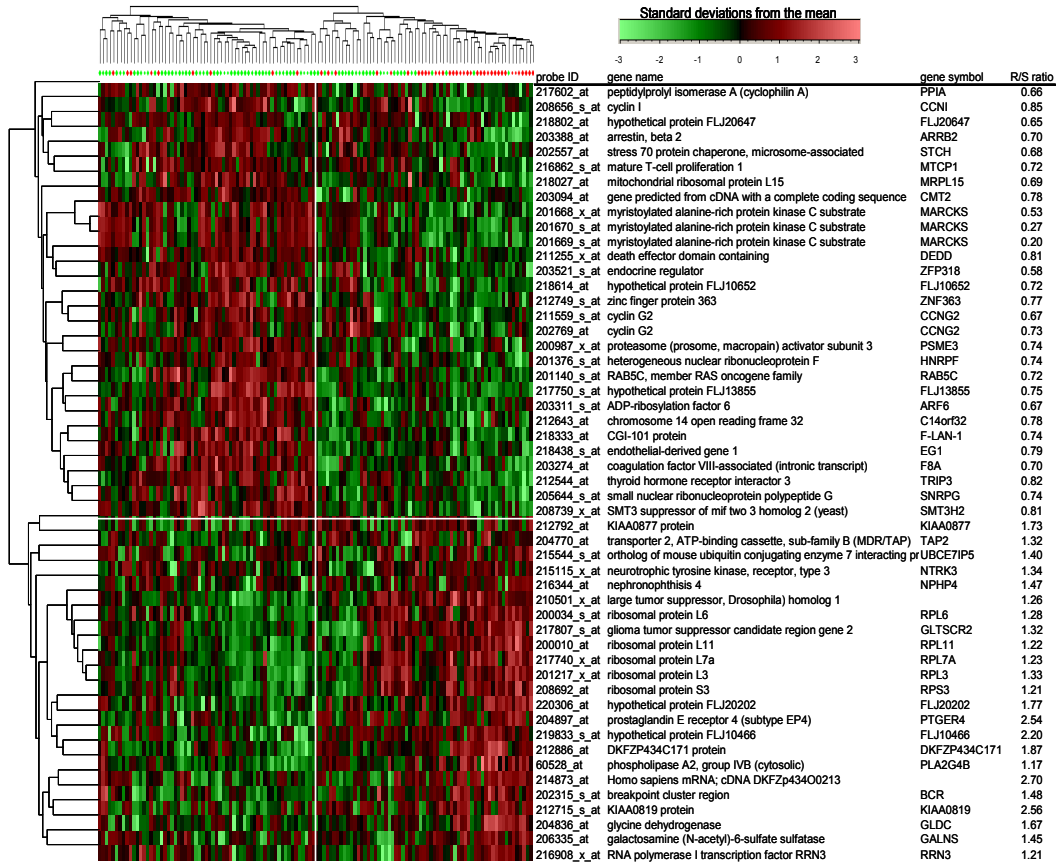
**Supplemental Figure 8: Supervised hierarchical clustering discriminating daunorubicin resistant and sensitive B-lineage ALL.**



**Supplemental Figure 9: Supervised hierarchical clustering discriminating prednisolone resistant and sensitive ALL (B- and T-lineage ALL).**

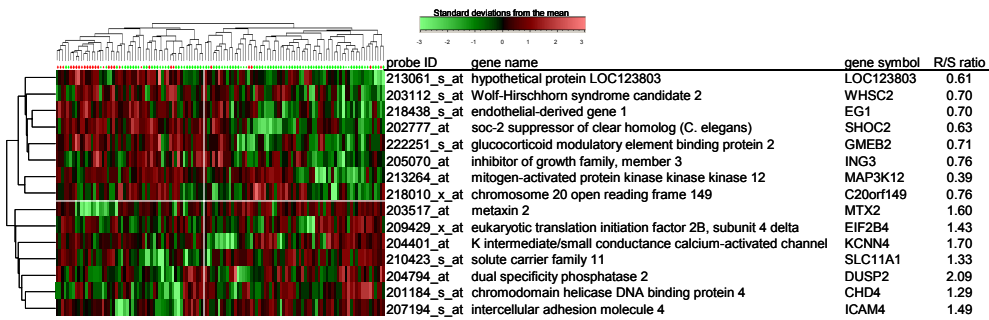


**Supplemental Figure 10: Supervised hierarchical clustering discriminating vincristine resistant and sensitive ALL (B- and T-lineage ALL).**



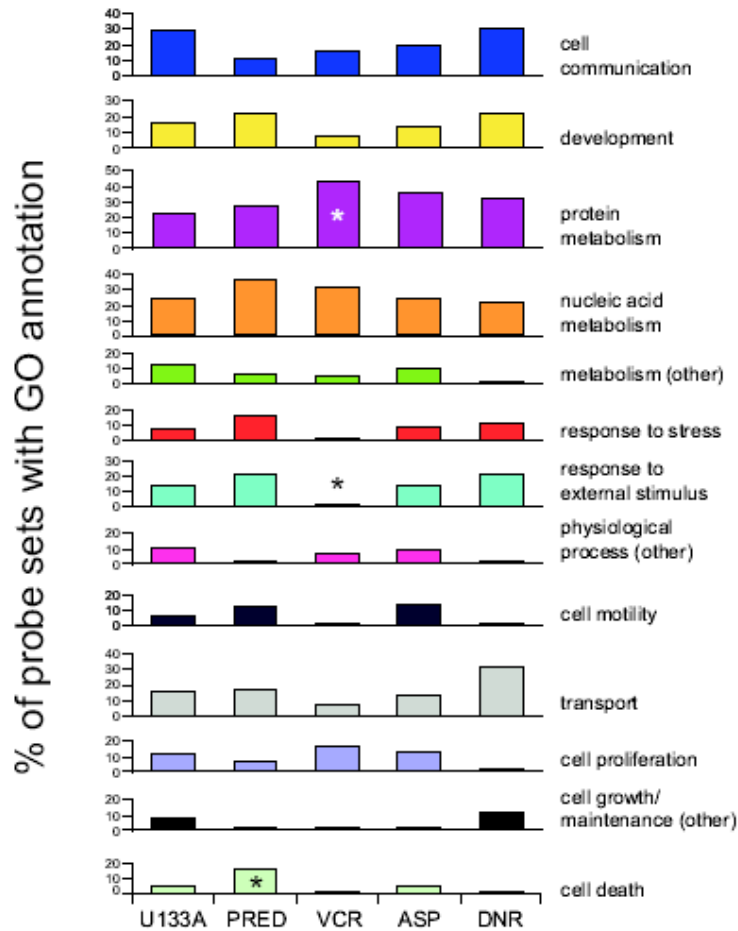
R/S ratio: ratio of the median expression in drug resistant versus sensitive ALL.

**Supplemental Figure 11: Supervised hierarchical clustering discriminating asparaginase resistant and sensitive ALL (B- and T-lineage ALL).**



R/S ratio: ratio of the median expression in drug resistant versus sensitive ALL.

**Supplemental Figure 12: Supervised hierarchical clustering discriminating daunorubicin resistant and sensitive ALL (B- and T-lineage ALL).**



**Supplemental Figure 13: Gene Ontology classification of genes discriminating drug resistance in patients with acute lymphoblastic leukemia (B- and T-lineage ALL).**

Functional classification in Gene Ontology (GO) of the probe sets discriminating drug resistance in ALL compared to the entire genome, as represented by all probe sets on the U133A GeneChip (22,283 probe sets; 12,983 with GO annotation) for four antileukemic agents. For prednisolone (PRED) 32 significant probe sets were identified, for vincristine (VCR) there were 54 significant probe sets, for asparaginase (ASP) 52 probe sets, and for daunorubicin (DNR) 15 probe sets discriminated resistant and sensitive ALL cells. There were 19, 33, 30 and 11 probe sets annotated in the GO database for PRED, VCR, ASP, and DNR, respectively. Functional categories that were proportionally over-represented in discriminating resistance drug compared to the entire genome, are indicated by an asterisk ( $P < 0.05$ , Fisher's exact test). Redundant gene symbols were counted as one.

**Supplemental Table 7: Genes previously linked to drug resistance or prognosis in ALL.**

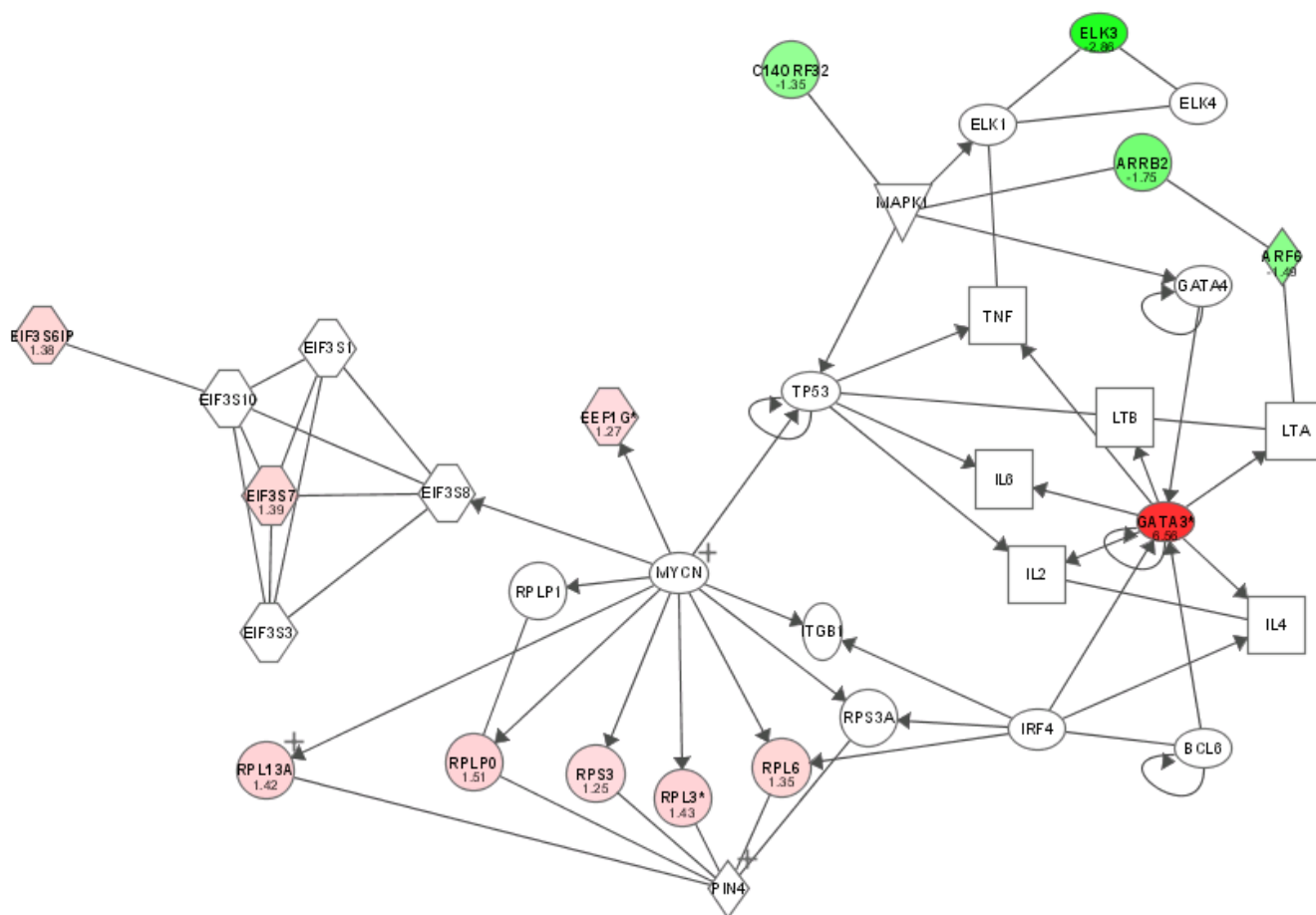
Gene expression ratios of resistant (R) and sensitive (S) samples are presented. These genes were not significant at the level required for inclusion in our model (Supplemental Table 4), but those indicated in red were significant at the \*P<0.05 or \*\*P<0.001 level.

Category	Ratio R/S				Ref.	Category	Ratio R/S				Ref.
	PRED	VCR	ASP	DNR			PRED	VCR	ASP	DNR	
<b>Drug efflux pumps:</b>						<b>Apoptosis:</b>					
MRP1	1.3	1.01	1.03	1.02	<sup>4</sup>	BAD	0.75	0.97	1.09	0.93	<sup>5,6</sup>
MDR1	1.11	0.99	1.60	0.90	<sup>4</sup>	BAK	0.99	1.10	1.06	1.10	<sup>5,6</sup>
LRP	1.08	0.87	1.06	1.35	<sup>4</sup>	BAX	0.96	0.95	1.12	0.85	<sup>5,6</sup>
BCRP	1.15	0.94	1.01	1.17	<sup>7</sup>	BCL2	1.34	1.01	1.13	1.16	<sup>5,6</sup>
<b>Drug metabolism:</b>						BCL2:BAX	0.72	0.94	0.99	0.74	<sup>5</sup>
GSTA1	1.25	0.83	1.14	0.95	<sup>8</sup>	BCL2L1	0.86	1.28	1.04	0.89	<sup>5,6</sup>
GSTM1	1.12	1.04	1.06	1.14	<sup>8</sup>	BAG1	1.20	0.82	1.26*	1.13	<sup>6</sup>
GSTP1	0.96	0.79*	0.94	0.94	<sup>8,10</sup>	CASP3	1.42	1.30*	0.92	0.98	<sup>9</sup>
MT1E	1.21	0.94	1.24*	1.27*	<sup>11</sup>	CD95	0.98	0.96	0.9	0.93	<sup>9</sup>
MT1F	1.50	0.81	1.68*	1.60*	<sup>11</sup>	XIAP	0.74	0.88	0.99	0.94	<sup>12</sup>
NM23A	0.82	0.78	1.07	1.25	<sup>13</sup>	<b>Transcription factors:</b>					
<b>Intracellular drug targets:</b>						NKFB1	1.22	0.94	0.95	1.13	<sup>14</sup>
GR	0.97	1.18*	0.97	0.97	<sup>15</sup>	RELA	0.93	0.91	1.19	1.00	<sup>14</sup>
MAP4	0.93	0.82	0.96	0.82	<sup>16</sup>	FOS	1.23	1.01	1.14	1.50	<sup>10</sup>
TUBB1	1.01	1.04	1.20	1.03	<sup>16</sup>	JUN	1.45	0.95	1.35	1.46	<sup>10</sup>
TUBB4	0.96	0.85	0.91	0.90	<sup>16</sup>	WT1	1.29	0.93	2.27*	1.77	<sup>17</sup>
TOP2A	0.84	1.01	1.08	0.94	<sup>18</sup>	<b>Miscellaneous:</b>					
TOP2B	0.71*	0.95	1.18	0.74	<sup>18</sup>	PARP-1	0.97	1.01	1.24	1.21*	<sup>19</sup>
<b>Cell cycle:</b>						HSP27	2.11	1.28	0.79	1.69	<sup>20</sup>
MDM2	0.97	1.23	0.96	1.07	<sup>14</sup>	HSP70	2.16	1.39	0.74	1.18	<sup>20</sup>
P15INK4B	1.11	0.78	1.09	1.32	<sup>22</sup>	HSP90	1.15	0.99	1.04	1.06	<sup>21</sup>
P16INK4A	0.88	1.07	0.98	0.88	<sup>22</sup>	ATM	0.98	1.00	1.24	1.23	<sup>23</sup>
P21	0.90	0.98	0.89*	0.93	<sup>25</sup>	ASNS	0.56	1.48	2.89**	2.08	<sup>24</sup>
TP53	0.78	0.82	1.06	1.13	<sup>27</sup>	GCS	124	1.12	1.81**	1.48	<sup>26</sup>
RB1	1.02	0.98	1.03	0.92	<sup>29</sup>	VEGF	1.27	1.16	0.92	1.17	<sup>28</sup>
CCND1	1.36	1.49	0.8	0.98	<sup>30</sup>						

Red indicates \*P<0.05, \*\*P<0.001 for Wilcoxon rank sum test and t-test.

For example, we found that the gene encoding asparagine synthetase (ASNS) was significantly over-expressed in ALL that was resistant to asparaginase, consistent with previously reported differences in the NCI panel of 60 human cancer cell lines.<sup>31,32</sup> However,

ASNS ( $P < 0.001$ , Wilcoxon rank sum test;  $P < 0.001$ , t-test) was not among the 54 most discriminating genes for asparaginase sensitivity, as defined by  $P < 0.0001$  (Supplemental Table 4).



**Supplemental Figure 14: Pathway analysis of genes associated with asparaginase resistance in B-lineage ALL.**

The 54 probe sets that were discriminative between asparaginase resistant and sensitive B-lineage ALL served as the input for the pathway analysis using Ingenuity software (<http://www.ingenuity.com>). The probe sets were mapped to the corresponding gene within the Ingenuity Pathway Knowledge Base. The pathway shown here is the biological network that included the highest number of genes queried. Highlighted genes are present in the set

of 54 probe sets that were associated with asparaginase resistance in B-lineage ALL. The figure indicates the ratio of expression between resistant and sensitive cases (R/S ratio); green color represents underexpression and red color represents overexpression in asparaginase resistant ALL.

**Supplemental Table 8: Multivariate analysis of gene expression and known prognostic factors (age, WBC count) to discriminate in vitro drug resistance.**

Multiple logistic regression was used with gene expression score (described in Supplemental Table 6), age and WBC count as continuous variables for the analysis in B-lineage ALL. An odds ratio >1 indicates increased probability of in vitro resistance.

	<b>predictor</b>	<b>odds ratio (95% C.I.)</b>	<b>P-value</b>
<b>PRED</b>	gene expression score	57.8 (4.9-681.6)	0.001
	age [years]	1.17 (1.01-1.36)	0.038
	WBC count [ $10^9/L$ ]	1.0 (0.98-1.01)	0.427
<b>VCR</b>	gene expression score	13.5 (1.97-92.6)	0.008
	age [years]	1.0 (0.89-1.13)	0.995
	WBC count [ $10^9/L$ ]	0.99 (0.99-1.0)	0.084
<b>ASP</b>	gene expression score	22.1 (4.8-102.6)	<0.001
	age [years]	1.13 (1.01-1.26)	0.033
	WBC count [ $10^9/L$ ]	1.0 (0.99-1.01)	0.792
<b>DNR</b>	gene expression score	151.2 (4.1-5533.6)	0.006
	age [years]	1.04 (0.93-1.17)	0.455
	WBC count [ $10^9/L$ ]	1.0 (0.99-1.0)	0.273

**Supplemental Table 9: Association of the LC50 scores with gene expression scores and other known prognostic variables.**

The relation of the combined LC50 score (by MTT assay) to disease-free survival was assessed in a univariate analysis. This revealed that in this cohort of patients (COALL/DCOG), PVAD LC50 score was not statistically significant at the P<0.05 level (intermediate LC50 score: hazard ratio =1.7, P=0.27; high LC50 score: hazard ratio =2.12, P=0.14), whereas the combined gene expression (GE) score was significantly associated

with disease-free survival (intermediate GE score hazard ratio =2.34, P=0.07; high GE score hazard ratio =3.8, P=0.0023). The combined LC50 score and the combined gene expression score are significantly correlated with each other (R=0.69, P-value<0.0001, Pearson correlation), and thus can confound each other in a multivariate analysis.

The table below summarizes results of a multivariate analysis in which LC50 values, gene expression scores, and other known prognostic variables (patient age, WBC at diagnosis) were all included in the model to assess their relation to treatment outcome (disease free survival). As shown, WBC (>100x10<sup>9</sup>/L) and gene expression score were statistically related to treatment outcome, whereas patient age is marginally significant, and LC50 scores are not.

<b>Variable</b>	<b>Hazard Ratio (95% C.I.)</b>	<b>P-value</b>
<b>Age (years)</b>	1.08 (1.0, 1.17)	0.054
<b>WBC (10<sup>9</sup>/L)</b>	1.48 (1.15, 1.91)	0.0024
<b>Combined drug resistance LC50 score<sup>‡</sup></b>	0.91 (0.75, 1.10)	0.310
<b>Combined drug resistance gene expression score<sup>‡</sup></b>	1.78 (1.10, 2.80)	0.016

<sup>‡</sup>for prednisolone, vincristine, asparaginase, daunorubicin

In a multivariate analysis in which the four LC50 scores for each patient were included in the analysis along with drug resistance GE score, the GE score was marginally significant (intermediate GE score hazard ratio =2.37, P=0.098; high GE score hazard ratio =2.85, P=0.072).

**Supplemental Table 10: Association of the drug resistance gene expression scores and minimal residual disease (MRD).**

To test the relation between the combined drug resistance gene expression score (GE score) and the presence (≥0.01 percent) or absence (≤0.01 percent) of minimal residual disease (MRD) on day 19 of treatment and on day 43 (end of induction), we used MRD data that

were available for a subset of the SJCRH cohort (independent test set). MRD results were available on 33 patients on day 19 and 44 patients on day 43. Despite the small number of patients for whom both MRD and gene expression were available in our cohort, there was an association between the GE score and MRD on day 19 (**a**), with a resistant (high) GE score associated with presence of MRD (P=0.10, Wicoxon rank sum test). In contrast, there was no evidence of an association between the GE score and MRD on day 43 (**b**) (P=0.52, Wicoxon rank sum test). The predominant components of treatment were prednisolone, vincristine, asparaginase and daunorubicin (PVAD) during the first 19 days, whereas etoposide and cytarabine were also given between days 19 and 43, consistent with the GE score being an important predictor of disease response to these four agents (i.e., day 19 MRD). Over the entire course of treatment (2.5 to 3 years) PVAD constituted a substantially greater proportion of treatment than other agents such as etoposide and cytarabine, consistent with GE score being significantly related to long term outcome, as observed in the training set and the independent test set.

**a. MRD day19**

	Combined drug resistance gene expression score <sup>‡</sup>			Total
	Low (<4.7)	Intermediate (4.7-5.58)	High (>5.58)	
MRD day 19 ≤0.01%	5	11	2	18
MRD day 19 ≥0.01%	2	7	6	15
<b>Total</b>	7	18	8	33

**b. MRD day43**

	Combined drug resistance gene expression score <sup>‡</sup>			Total
	Low (<4.7)	Intermediate (4.7-5.58)	High (>5.58)	
MRD day 43 ≤0.01%	7	18	4	29
MRD day 43 ≥0.01%	4	7	4	15
<b>Total</b>	11	25	8	44

<sup>‡</sup>for prednisolone, vincristine, asparaginase, daunorubicin

When MRD is included in a multivariate survival analysis with the GE score (n=33), there is an association of a poor outcome in patients with a resistant (high) GE score, after adjusting for MRD status, but the sample size limits the power of such an analysis in our patient cohorts. (a) Day 19 MRD negative status ( $\leq 0.01$  percent) was associated with a lower probability of relapse (hazard ratio of 0.014,  $P=0.09$ ), and a drug resistant (high) GE score was associated with a higher risk of relapse (hazard ratio = 2.69,  $P=0.13$ ), after adjusting for MRD. (b) Using day 43 MRD data (N=44), the hazard ratio for MRD negative patients was 0.04 ( $P=0.001$ ), and for patients with a resistant (high) GE score the hazard ratio was 2.38 ( $P=0.08$ ) after adjusting for MRD status.

**a.**

Variable	No. of patients	Hazard Ratio (95% C.I.)	P-value
<b>Combined drug resistance gene expression score<sup>‡</sup></b>	33	2.69 (0.75, 9.36)	0.13
<b>MRD day19</b>			
MRD day19 $\geq 0.01\%$	15	1.0*	
MRD day19 $\leq 0.01\%$	18	0.14 (0.014, 1.34)	0.09

**b.**

Variable	No. of patients	Hazard Ratio (95% C.I.)	P-value
<b>Combined drug resistance gene expression score<sup>‡</sup></b>	44	2.38 (0.88, 6.44)	0.08
<b>MRD day43</b>			
MRD day43 $\geq 0.01\%$	15	1.0*	
MRD day43 $\leq 0.01\%$	29	0.04 (0.006, 0.30)	0.001

\*Reference subgroup

<sup>‡</sup>for prednisolone, vincristine, asparaginase, daunorubicin

It is also important to note that even if MRD were to overwhelm the prognostic significance of the gene expression (GE) score, this would not obviate the importance of the association between gene expression and drug resistance or disease outcome. The primary reason is that MRD is a marker of poor response, but it tells one nothing about why a given patient has had a poor response. The gene expression signature provides insights into the basis of drug resistance and poor response, thereby offering strategies whereby more effective therapy can be pursued.

## References

1. 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.
2. Kaspers GJ, Veerman AJ, Pieters R, et al. In vitro cellular drug resistance and prognosis in newly diagnosed childhood acute lymphoblastic leukemia. *Blood*. 1997;90:2723-2729.
3. 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;21:3262-3268.
4. Den Boer ML, Pieters R, Kazemier KM, et al. Relationship between major vault protein/lung resistance protein, multidrug resistance-associated protein, P-glycoprotein expression, and drug resistance in childhood leukemia. *Blood*. 1998;91:2092-2098.
5. Salomons GS, Smets LA, Verwijs-Janssen M, et al. Bcl-2 family members in childhood acute lymphoblastic leukemia: relationships with features at presentation, in vitro and in vivo drug response and long-term clinical outcome. *Leukemia*. 1999;13:1574-1580.
6. Kitada S, Andersen J, Akar S, et al. Expression of apoptosis-regulating proteins in chronic lymphocytic leukemia: correlations with In vitro and In vivo chemoresponses. *Blood*. 1998;91:3379-3389.
7. van den Heuvel-Eibrink MM, Wiemer EA, Prins A, et al. Increased expression of the breast cancer resistance protein (BCRP) in relapsed or refractory acute myeloid leukemia (AML). *Leukemia*. 2002;16:833-839.

8. Den Boer ML, Pieters R, Kazemier KM, et al. Different expression of glutathione S-transferase alpha, mu and pi in childhood acute lymphoblastic and myeloid leukaemia. *Br J Haematol.* 1999;104:321-327.
9. Martinez-Lorenzo MJ, Gamen S, Etxeberria J, et al. Resistance to apoptosis correlates with a highly proliferative phenotype and loss of Fas and CPP32 (caspase-3) expression in human leukemia cells. *Int J Cancer.* 1998;75:473-481.
10. Volm M, Zintl F, Edler L, Sauerbrey A. Prognostic value of protein kinase C, proto-oncogene products and resistance-related proteins in newly diagnosed childhood acute lymphoblastic leukemia. *Med Pediatr Oncol.* 1997;28:117-126.
11. Sauerbrey A, Zintl F, Volm M. Expression of metallothionein in initial and relapsed childhood acute lymphoblastic leukemia. *Ann Hematol.* 1994;69:111-115.
12. Tamm I, Kornblau SM, Segall H, et al. Expression and prognostic significance of IAP-family genes in human cancers and myeloid leukemias. *Clin Cancer Res.* 2000;6:1796-1803.
13. Koomagi R, Sauerbrey A, Zintl F, Volm M. nm23-H1 protein expression in newly diagnosed and relapsed childhood acute lymphoblastic leukemia. *Anticancer Res.* 1998;18:4307-4309.
14. 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;99:3367-3375.
15. Wells RJ, Mascaro K, Young PC, Cleary RE, Bachner RL. Glucocorticoid receptors in the lymphoblasts of patients with glucocorticoid-resistant childhood acute lymphocytic leukemia. *Am J Pediatr Hematol Oncol.* 1981;3:259-264.
16. Kavallaris M, Tait AS, Walsh BJ, et al. Multiple microtubule alterations are associated with Vinca alkaloid resistance in human leukemia cells. *Cancer Res.* 2001;61:5803-5809.

17. Inoue K, Sugiyama H, Ogawa H, et al. WT1 as a new prognostic factor and a new marker for the detection of minimal residual disease in acute leukemia. *Blood*. 1994;84:3071-3079.
18. Brown GA, McPherson JP, Gu L, et al. Relationship of DNA topoisomerase II alpha and beta expression to cytotoxicity of antineoplastic agents in human acute lymphoblastic leukemia cell lines. *Cancer Res*. 1995;55:78-82.
19. Richardson DS, Allen PD, Kelsey SM, Newland AC. Effects of PARP inhibition on drug and Fas-induced apoptosis in leukaemic cells. *Adv Exp Med Biol*. 1999;457:267-279.
20. Stammer G, Volm M. Expression of heat shock proteins, glutathione peroxidase and catalase in childhood acute lymphoblastic leukemia and neuroblastoma. *Cancer Lett*. 1996;99:35-42.
21. Lauten M, Beger C, Gerdes K, et al. Expression of heat-shock protein 90 in glucocorticoid-sensitive and -resistant childhood acute lymphoblastic leukaemia. *Leukemia*. 2003;17:1551-1556.
22. 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;112:680-690.
23. Haidar MA, Kantarjian H, Manshour T, et al. ATM gene deletion in patients with adult acute lymphoblastic leukemia. *Cancer*. 2000;88:1057-1062.
24. Kiriya Y, Kubota M, Takimoto T, et al. Biochemical characterization of U937 cells resistant to L-asparaginase: the role of asparagine synthetase. *Leukemia*. 1989;3:294-297.
25. Baghdassarian N, Bertrand Y, Ffrench P, Duhaut P, Bryon PA, Ffrench M. Role of BCL-2 and cell cycle regulatory proteins for corticosteroid sensitivity assessment in childhood acute lymphoblastic leukaemia. *Br J Haematol*. 2000;109:109-116.

26. Itoh M, Kitano T, Watanabe M, et al. Possible role of ceramide as an indicator of chemoresistance: decrease of the ceramide content via activation of glucosylceramide synthase and sphingomyelin synthase in chemoresistant leukemia. *Clin Cancer Res.* 2003;9:415-423.
27. Lam V, McPherson JP, Salmena L, et al. p53 gene status and chemosensitivity of childhood acute lymphoblastic leukemia cells to adriamycin. *Leuk Res.* 1999;23:871-880.
28. Koomagi R, Zintl F, Sauerbrey A, Volm M. Vascular endothelial growth factor in newly diagnosed and recurrent childhood acute lymphoblastic leukemia as measured by real-time quantitative polymerase chain reaction. *Clin Cancer Res.* 2001;7:3381-3384.
29. Sauerbrey A, Stammeler G, Zintl F, Volm M. Expression and prognostic value of the retinoblastoma tumour suppressor gene (RB-1) in childhood acute lymphoblastic leukaemia. *Br J Haematol.* 1996;94:99-104.
30. Sauerbrey A, Hafer R, Zintl F, Volm M. Analysis of cyclin D1 in de novo and relapsed childhood acute lymphoblastic leukemia. *Anticancer Res.* 1999;19:645-649.
31. Scherf U, Ross DT, Waltham M et al. A gene expression database for the molecular pharmacology of cancer. *Nat Genet* 2000;24:236-244.
32. Weinstein JN, Myers TG, O'Connor PM et al. An information-intensive approach to the molecular pharmacology of cancer. *Science* 1997;275:343-349.