Prediction of Chemosensitivity for Patients with Acute Myeloid Leukemia, According to Expression Levels of 28 Genes Selected by Genome-wide Complementary DNA Microarray Analysis[1,2]

Jun-ichi Okutsu, Tatsuhiko Tsunoda, Yasuyuki Kaneta, Toyomasa Katagiri, Osamu Kitahara, Hitoshi Zembutsu, Rempei Yanagawa, Shuichi Miyawaki, Kazutaka Kuriyama, Nobuyuki Kubota, Yukihiro Kimura, Kohmei Kubo, Fumiharu Yagasaki, Toshio Higa, Hirokuni Taguchi, Tadasu Tobita, Hideki Akiyama, Akihiro Takeshita, Yan-Hua Wang, Toshiko Motoji, Ryuzo Ohno, and Yusuke Nakamura[3]

Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo 108-8639 [J-i. O., Y. Ka., T. K., O. K., H. Z., R. Y., Y. N.]; Laboratory of Medical Informatics, SNP Research Center, RIKEN (Institute of Physical and Chemical Research), Tokyo 108-8639 [T. Ts.]; Department of Medicine, Saiseikai Maebashi Hospital, Maebashi 371-0821 [S. M.]; Department of Hematology and Molecular Medicine, Atomic Bomb Disease Institute, Nagasaki University School of Medicine, Nagasaki 852-8102 [K. Kur.]; Tokai University School of Medicine, Isehara 259-1100 [N. K.]; First Department of Internal Medicine, Tokyo Medical University, Tokyo 160-0023 [Y. Ki.]; Aomori Prefectural Central Hospital, Aomori 030-0913 [K. Kub.]; First Department of Internal Medicine, Saitama-Medical School, Saitama 350-0495 [F. Y.]; Department of Internal Medicine, Sapporo Hokuyu Hospital, Shiroishi-ku, Sapporo 003-0006 [T. H.]; Third Department of Internal Medicine, Kochi Medical School, Kochi 783-0043 [H. T.]; Department of Hematology, Yaizu City Hospital, Yaizu [T. To.]; Tokyo Metropolitan Komagome Hospital, Tokyo 113-0024 [H. K.]; Department of Internal Medicine III, Hamamatsu University, School of Medicine, Hamamatsu 431-3192 [A. T.]; Department of Hematology, Tokyo Women's Medical University, Tokyo 162-0054 [Y-H. W., T. M.]; Aichi Cancer Center, Nagoya 464-0021 [R. O.], Japan

Abstract

To identify genes involved in the sensitivity of acute myeloid leukemia (AML) cells to chemotherapy, we monitored gene-expression profiles of cancer cells from 76 AML patients using a cDNA microarray consisting of 23,040 genes. We identified 63 genes that were commonly overexpressed and 372 genes suppressed in AML. Because these genes represent key molecules for disclosing the molecular mechanisms of AML, they may be potential targets for drug development. We also found 28 that revealed different expression levels between good and poor responders to chemotherapy and appeared to be associated with chemosensitivity. On that basis, we developed a "Drug Response Scoring" system that was correlated well with individual sensitivity to an anticancer drug regimen. Among the 44 cases with positive drug-response scores by our definition, 40 achieved complete remission after treatment, whereas the only 3 of the 20 cases with negative scores responded well to the treatment. An ability to predict chemosensitivity should eventually lead to achievement of our goal of "personalized therapy."

Introduction

AML[4] is a hematological malignancy of myeloid lineage, characterized by blast cells arrested at a certain stage of myeloid differentiation. Immature blast cells proliferate in bone marrow and then are released to the peripheral blood. Although recent studies have indicated that genetic and epigenetic alterations in several genes are involved in the etiology of AML as is the case with solid tumors, the molecular mechanism of this cancer is not fully understood. We consider a genome-wide analysis to be an essential step toward a better understanding of the molecular basis of AML, as well as a way to identify molecular targets for development of novel diagnostic and therapeutic methods.

The development of cDNA microarray or DNA-chip technology in recent years has made it possible to analyze the expression of thousands of genes simultaneously. Patterns of gene expression in yeast (1), cell lines (2), and cancer patients (3–8) have been investigated this way, and importantly, a large amount of information can be retrieved by adding statistical methods such as cluster analysis (9–11). For example, Golub *et al.* (12) reported that patients with AML could be distinguished from those with acute lymphoblastic leukemia by differential expression of certain genes. These techniques are the most powerful tool yet developed for clarifying the molecular mechanisms involved or associated with diseases and for investigating which genes are involved in signaling and metabolic pathways. To construct a genome-wide gene-expression database for cancer, we established a cDNA microarray system consisting of 23,040 genes (6). In the study reported here, we used the microarray to study AML.

Because of advances in therapeutic methods, the proportion of AML patients who can be induced into remission has

Received 5/6/02; revised 7/8/02; accepted 8/7/02.
[1] This work was supported in part by Research for the Future Program Grant 00L01402 from the Japan Society for the Promotion of Science.
[2] Supplementary data for this article are available at Molecular Cancer Therapeutics Online (http://mct.aacrjournals.org).
[3] To whom requests for reprints should be addressed, at Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan. Phone: 81-3-5449-5372; Fax: 81-3-5449-5433; E-mail: yusuke@ims.u-tokyo.ac.jp.

[4] The abbreviations used are: AML, acute myeloid leukemia; aRNA, amplified RNA; RT-PCR, reverse transcription-PCR; Ara-C, cytosine-arabinoside; EST, expressed sequence tag; PCNA, proliferating cell nuclear antigen.

increased substantially in the last decade. However, no method yet exists for predicting the response of an individual patient to therapy with anticancer drugs. Some patients suffer from adverse effects without any positive results, thereby losing the chance of trying alternative chemotherapy if their physical condition has deteriorated too far. Hence, accurate prediction regarding the effectiveness of a specific therapy is of critical importance for cancer patients. Certain factors are known to be associated with chemosensitivity or prognosis (13–16), but the information from only one or a few of these factors have thus far failed to predict individual response; a larger body of information is needed to predict chemosensitivity more precisely.

We believed that data concerning expression of thousands of genes might provide a better understanding of the characteristics of AML cells. On the basis of this hypothesis, we performed cDNA microarray analysis of >20,000 genes and selected the ones that were differentially expressed between good responders and poor responders among AML patients treated with a standard drug regimen. Here we report that activity of genes selected in this way can explain the effectiveness of chemotherapy and suggest that such information may lead ultimately to our goal of "personalized therapy."

Materials and Methods

RNA Preparation and T7-based RNA Amplification. We prepared mononuclear cells (2×10^7 cells) using Ficoll (Amersham Biosciences, Buckinghamshire, United Kingdom) immediately after samples were transferred and extracted total RNA using TRIzol (Life Technologies, Inc., Grand Island, NY) according to the manufacturer's instructions. After treatment with DNase I (Nippon Gene, Tokyo, Japan), T7-based RNA amplification was carried out as described previously (5). Using 2 μg of total RNA as starting material, we performed two rounds of amplification and finally gained 40–100 μg of aRNA for each sample. For control samples, we also performed two rounds of T7-based RNA amplification to obtain sufficient volumes of aRNA. RNA amplified by this method accurately reflects the proportions in the original RNA source (17, 18), as we have confirmed by semiquantitative RT-PCR experiments (5). Data from the microarrays were consistent with results from RT-PCR, whether total RNA or aRNA was used as the template.

Preparation of the Microarray. To obtain cDNAs for spotting on the glass slides, we performed RT-PCR for each gene as described previously (6). The PCR products were spotted on type 7 glass slides (Amersham Biosciences) by a Microarray Spotter Generation III (Amersham Biosciences); 4,608 genes were spotted in duplicate on a single slide. We prepared five different sets of slides (total, 23,040 genes), on each of which the same 52 housekeeping genes and 2 negative-control genes were spotted as well.

Labeling, Hybridization, and Scanning. The cDNA probes were prepared from aRNA in the manner described previously (5). For the hybridization experiments, 12.5-μg aliquots of aRNA from healthy volunteers and AML patients were labeled with Cy3-dCTP and Cy5-dCTP (Amersham Biosciences), respectively. Hybridization and washing were performed according to protocols described previously (5), except that all processes were carried out with an Automated Slide Processor (Amersham Biosciences).

Quantification of Signals. We calculated the intensity of each hybridization signal photometrically using the ArrayVision computer program (Amersham Biosciences) and then averaged the intensities of duplicate spots. After subtraction of background, data were normalized so that the average of log (Cy5:Cy3) of all spots would be 0. To consider the reliability of signal intensities, cutoff values were automatically calculated according to fluctuation of the data (5). If both Cy3 and Cy5 signal intensities were lower than the cutoff values, expression of the corresponding gene in that sample was assessed as low or absent. For other genes, we calculated Cy5:Cy3 as a relative expression ratio. In cases where only the Cy3 signal was lower than the cutoff value, we replaced its signal with the cutoff value and calculated Cy5:cutoff as a relative expression ratio. If only the Cy5 signal was lower than the cutoff value, we calculated cutoff:Cy3.

Selection of Genes Associated with Chemosensitivity and Calculation of "Drug Response Score." We compared relative expression ratios of 32 cases belonging to group 1 with 12 cases of group 3. We calculated the difference between the median ratios of the two groups for each gene and selected genes showing differences >2-fold between the two groups. Then, for each selected gene we calculated *U* values by the Mann-Whitney test, *i.e.,* the number of samples that overlapped when arranged according to the order of their relative expression ratios. In cases of low or no expression, we assumed a relative expression ratio of 1.0 to calculate the *U* value. We considered genes with *P*s lower than 0.01, *i.e.,* with *U*s <95, to be differentially expressed between the two groups.

Using these data, we developed a "Drug Response Scoring" system. If genes were preferentially expressed in AML in group 1, we assigned "plus ten" to them. Genes that showed the opposite pattern were provided the sign of minus ten. For each sample, we calculated the sum of $\log_2$ (Cy5:Cy3), multiplied by each sign, and that became the drug response score. We performed cDNA microarray analysis and subsequent calculation of drug response scores for an additional 29 AML cases to further evaluate our scoring system. These test samples consist of 11 good responders, 9 intermediate responders, and 9 poor responders who had not been part of the original procedure for selecting discriminating genes.

Results

Clinicopathological Features. In cooperation with the Japan Adult Leukemia Study Group (JALSG), we initially obtained samples from 56 AML patients with informed consent to find genes associated with chemosensitivity and later obtained 20 additional cases to verify our chemosensitivity prediction system. Clinicopathological features of the patients used in this study are summarized in Table 1. Samples had been taken at the time of diagnosis from patients before starting chemotherapy, and those in which the proportion of blast cells exceeded 70% were used for microarray analysis. The diagnosis was made according to the French-American-British (FAB) classification and samples examined in this study were classified from M0 to M5. FAB-M3 samples were

Table 1 Clinicopathological features of patients examined

ID. no.	Age	Sex	Response[a]	FAB	Sample	Blast cell (%)
092	51	F	1	M4	BM[b]	90
096	40	M	1	M1	BM	88
099	36	M	1	M2	BM	75
107	32	F	2	M1	BM	90
124	42	M	3	M1	BM	90
139	41	M	Others	M2	PB	82
142	61	M	1	M2	BM	70
144	44	F	1	M4	BM	94
159	64	F	1	M2	PB	76
167	42	M	3	M0	BM	84
170	15	M	2	M0	BM	98
174	16	F	1	M4	BM	90
175	26	M	1	M1	PB	95
178	48	M	1	M1	BM	80
181	63	M	3	M1	PB	90
184	62	F	1	M2	BM	90
185	51	M	1	M4	BM	77
187	34	M	1	M2	PB	70
199	30	F	3	M2	BM	85
202	17	F	3	M1	PB	90
204	56	M	3	M4	BM	80
205	46	M	1	M2	BM	81
212	52	F	Others	M1	BM	96
214	58	F	2	M2	BM	80
216	20	M	1	M5a	BM	88
219	63	M	1	M2	PB	74
222	41	F	1	M2	PB	70
223	34	M	Others	M5a	BM	91
234	63	F	1	M2	PB	92
240	38	M	1	M5a	PB	70
248	50	M	2	M4	PB	97
249	53	M	3	M4	BM	80
259	32	M	1	M2	PB	82
263	66	F	2	M4	BM	95
267	62	M	1	M1	BM	75
270	62	M	2	M2	BM	82
279	19	M	1	M2	BM	70
284	43	M	1	M2	BM	70
285	23	F	1	M4	PB	95
287	64	M	1	M1	BM	80
290	62	M	3	M5b	PB	90
297	22	F	1	M4	BM	70
298	38	F	3	M4	BM	90
299	33	F	1	M2	BM	70
300	28	M	3	M2	BM	98
303	47	F	1	M5a	BM	90
305	29	M	2	M1	BM	90
309	63	F	1	M1	BM	82
311	60	F	3	M2	BM	90
312	23	M	1	M5b	BM	95
317	46	F	3	M1	BM	80
324	46	F	1	M1	PB	93
326	45	M	2	M2	BM	85
328	49	F	2	M2	BM	70
334	47	M	1	M2	BM	81
336	58	M	1	M4	PB	90
339	43	F	1 (test)	M2	BM	80
344	17	M	1 (test)	M1	BM	70
347	59	F	1 (test)	M2	PB	85
349	28	F	1 (test)	M2	BM	70
356	62	F	1 (test)	M1	PB	99
362	53	M	1 (test)	M2	BM	80
363	40	M	3 (test)	M1	BM	96
367	56	M	1 (test)	M2	BM	73
368	33	F	1 (test)	M2	PB	90
380	55	M	1 (test)	M1	BM	90
381	39	F	1 (test)	M5b	PB	90
383	46	M	1 (test)	M0	PB	90
M5	35	F	3 (test)	M2	PB	78
M16	52	M	3 (test)	M1	PB	92
M19	46	M	3 (test)	M4	PB	74
M20	70	M	3 (test)	M1	PB	84
R34p	37	F	3 (test)	M2	PB	90
R36p	71	M	3 (test)	M4	PB	21[c]
R38p	43	F	3 (test)	M2	PB	46[c]
R44p	26	M	3 (test)	M1	PB	66[c]

[a] Response to induction therapy: 1, patients who achieved complete remission after one course; 2, patients who achieved complete remission after two courses; 3, patients who could not achieve complete remission even after two courses. (test), samples used for test cases in chemosensitivity analysis. Others, Patients who were not categorized into any of three groups.
[b] BM, bone marrow. PB, peripheral blood.
[c] Samples with higher percentage (>70%) of blast cells after preparation of mononuclear cells.

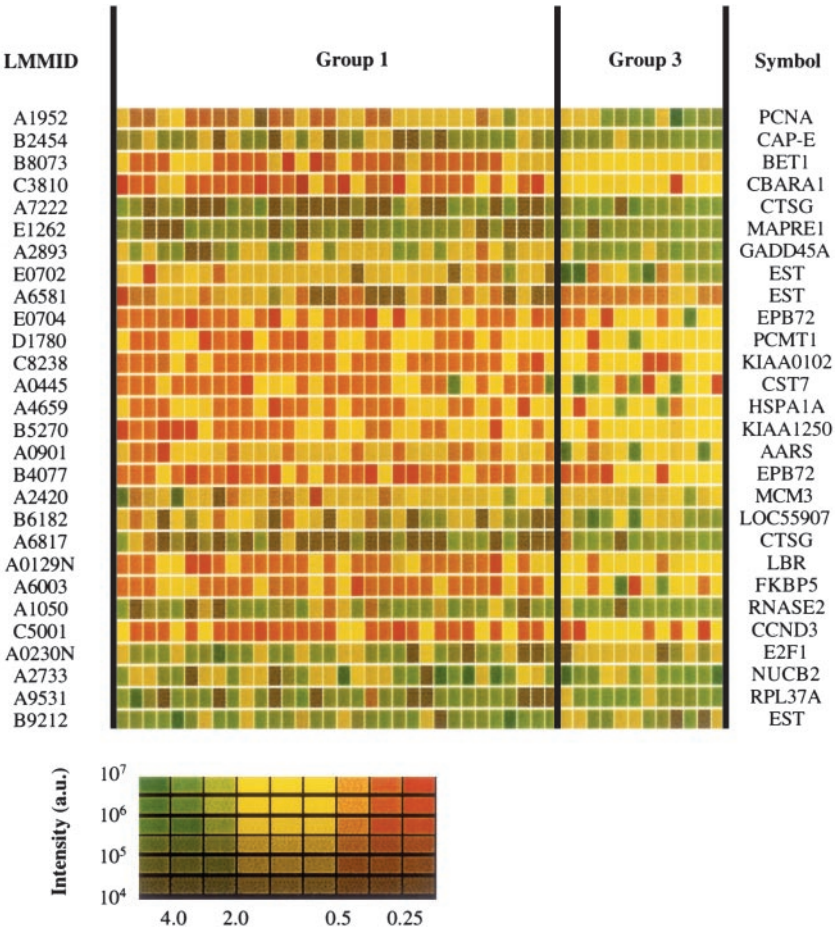


Fig. 1. Gene-expression patterns among 32 good responders and 12 poor responders. Expression data are shown by pseudocolors based on relative expression ratios (Cy5:Cy3) and signal intensities (average of Cy3 and Cy5 signals). *Brown tiles,* low or no expression.

not included in this study because most of those patients were treated by more effective chemotherapy with all-*trans* retinoic acid (ATRA). Among the 76 samples examined, 47 were from bone marrow and 29 from peripheral blood. All patients received the same induction therapy, consisting of a combination of 100 mg/m^2 of Ara-C for 7 days and 12 mg/m^2 of idarubicin for 3 days. According to their responses to the treatment, we categorized the patients into three groups: those who achieved complete remission after one course of induction therapy were classified as "good responders" (group 1); those who did not achieve complete remission even after two courses were classified as "poor responders" (group 3); and patients who showed complete remission after two courses were considered "intermediate-responders" (group 2). As controls, we used a mixture of mononuclear cells from peripheral blood from four healthy volunteers.

Overexpressed or Suppressed Genes. We screened for genes whose expression was commonly altered in AML cells. When relative expression ratios were >3 in more than half of the samples examined, we defined those genes as commonly overexpressed; those whose relative expression ratios were less than one-third in more than half of samples were defined as commonly suppressed genes. By those definitions, 63 genes were selected as commonly overexpressed in AML; 47 of them were of known function, and the other 16 were ESTs.[2] v-myb avian myeloblastosis viral oncogene homologue (*MYB*) was overexpressed in all 56 AML samples. On the other hand, a total of 372 genes, including 110 ESTs, were selected as commonly suppressed in AML.[2] Semaphorin 4D (*SEMA4D*), granzyme A (*GZMA*), T-cell receptor β locus (*TRB@*), transforming growth factor, β receptor II (*TGFBR2*), granzyme B (*GZMB*), protein tyrosine phosphatase, non-receptor type 12 (*PTPN12*), RNB6 (*RNB6*), and five ESTs were suppressed in all 56 cases.

Chemosensitivity Analysis. We then searched for genes that might be associated with response to chemotherapy of 53 samples as learning cases using the Mann-Whitney test (3 patients were excluded from this analysis because 2 died before starting the treatment and 1 was treated with another anticancer drug after one course of the treatment) and selected 28 genes whose expressions levels were significantly different between a group of 32 good responders (group 1) and a group of 12 poor responders (group 3; Fig. 1 and Table 2). PCNA, which promotes the cell cycle, was preferentially expressed in AML cells of the poor responders.

We established an algorithm to calculate "Drug Response Scores" using expression levels of the 28 selected genes to predict individual clinical responses to chemotherapy. We calculated the sum of the log expression ratio multiplied by the sign [+10 or −10] of each gene as the drug

Table 2 Genes associated with chemosensitivity

Genes preferentially expressed in good-responders and genes showing the opposite pattern are listed separately. *U* values of Mann-Whitney tests are listed in the first column.

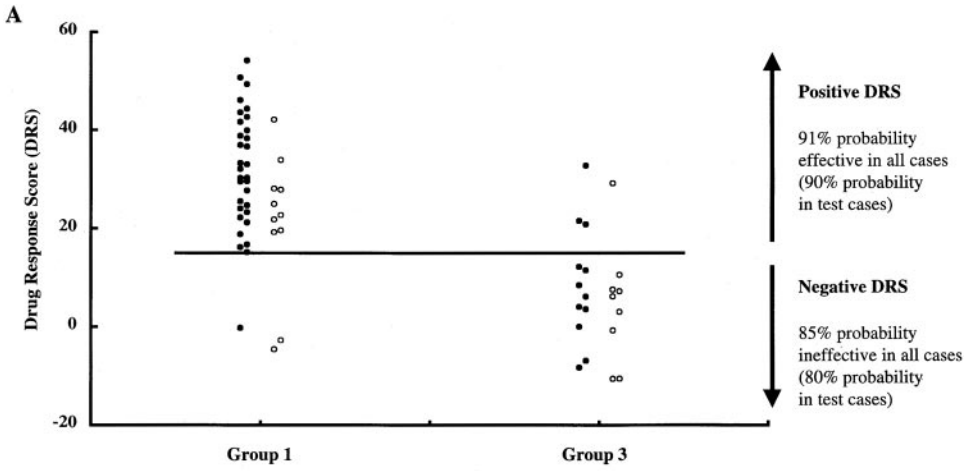
U value	LMMID	Accession number	UniGene ID	Cytoband	Symbol
Good responder < Poor responder					
53	A1952	M15796	Hs.78996	20pter–p12	PCNA
54	B2454	AA749076	Hs.119023	9q22.31–9q34.11	CAP-E
55	B8073	AA037028	Hs.23103	7q21.1–q22	BET1
60	C3810	NM_006077	Hs.61628	10	CBARA1
60.5	A7222	M16117	Hs.100764	14q11.2	CTSG
62.5	E1262	AL035071	Hs.234279	20q11.11.23	MAPRE1
65	A2893	M60974	Hs.80409	1p31.2–p31.1	GADD45A
66	E0702	BE891654	Hs.14846		
67	A6581	T81218	Hs.8003		
68	E0704	BE439695	Hs.160483	9q34.1	EPB72
70	D1780	D25545	Hs.79137	6q24–q25	PCMT1
70	C8238	D14658	Hs.77665	11cen1q13.5	KIAA0102
71	A0445	AF031824	Hs.143212	20p11.22–p11.21	CST7
72	A4659	M11717	Hs.8997	6p21.3	HSPA1A
75	B5270	AA744859	Hs.9873	2p24	KIAA1250
76	A0901	D32050	Hs.75102	16q22	AARS
77	B4077	M81635	Hs.160483	9q34.1	EPB72
78	A2420	D38073	Hs.179565	6p12	MCM3
80	B6182	AA417239	Hs.12492	12p12.32q14.3	LOC55907
83	A6817	AA355657	Hs.100764	14q11.2	CTSG
84	A0129N	AA099192	Hs.152931	1q42.1	LBR
86	A6003	AA678103	Hs.7557	6	FKBP5
88.5	A1050	M28129	Hs.728	14q24–q31	RNASE2
90	C5001	M92287	Hs.83173	6p21	CCND3
92	A0230N	S49592	Hs.96055	20q11.2	E2F1
92.5	A2733	X76732	Hs.3164	11p15.1–p14	NUCB2
94	A9531	AI357601	Hs.184109	2	RPL37A
Good responder > Poor responder					
90	B9212	AI281337	Hs.177788		

response score (see "Materials and Methods"). When we used the score of 15 as a borderline score, distribution of the drug response scores for the 44 patients belonging to group 1 and those to group 3 were clearly separated (Fig. 2) and then defined scores of over 15 as "positive," and those of 15 or lower as "negative." To clarify the system further, we calculated the drug response scores for the 9 original intermediate responders classified as group 2 as well as 20 additional test cases consisting of 11 good responders and 9 poor responders who had not been part of the original procedure for selecting discriminating genes. For the 9 patients belonging to group 2, 6 showed positive scores, and the remaining 3 showed negative scores (data not shown). Nine of the 10 patients with positive scores achieved complete remission after one course of the treatment, and 8 of the 10 patients with negative scores failed to achieve remission after two courses of the same chemotherapy. Hence, the predictive scores correctly reflected the clinical response of 17 (85%) of the 20 test cases. Among the total of 44 cases (including learning and test cases) with positive scores (over the borderline value), 40 achieved complete remission after one course of treatment; only 4 patients have thus far failed to show a positive response. On the other hand, 17 of the 20 cases with negative scores were unable to show any response.

Discussion

The pseudocolor views of the expression-profile images obtained from each of the 76 patients analyzed in this study indicated altered expression patterns between normal and cancer cells, but the expression profiles of AML cells were very similar to each other except for a small subset of genes that were likely to reflect differences in characteristics of individual cancers. Hence, analysis of expression profiles can extract a set of genes whose expression is commonly changed in AML and also those whose expression differs among groups with different clinicopathological features. In this study, we aimed to find genes whose expression levels could be correlated with clinical response to treatment with a combination of Ara-C and idarubicin, but we began by identifying numerous genes whose expression was commonly altered in AML cells.

Many of the genes that were commonly overexpressed in AML cells had previously shown relationships to AML, underscoring the reliability of our microarray analysis. Myeloperoxidase (MPO) is a well-known AML marker, and almost all cases examined here showed elevated expression of that gene. Cell surface protein, neural cell adhesion molecule (*NCAM2*), is present in immature myeloid cells (19). Others have reported overexpression of Fms-related tyrosine kinase 3 (*FLT3*) in AML cells (20), and polymorphism of *FLT3* appears to have some relationship to leukocytosis and prog-



B

Cases belonging to Groups 1 and 3		Remission (+)	Remission (-)	Total
DRS	positive	40	4	44
DRS	negative	3	17	20
Total		43	21	64

$p=3X10^{-9}$ (Fisher's exact test)

Test Cases		Remission (+)	Remission (-)	Total
DRS	positive	9	1	10
DRS	negative	2	8	10
Total		11	9	20

p=0.002 (Fisher's exact test)

Fig. 2. Distribution of the drug response scores. *A,* drug response scores for individual patients. ●, scores for patients whose expression data were used for selecting discriminating genes (learning). ○, scores for additional test cases. *B,* correlation between the drug response score and the clinical responses.

nosis (21). v-myb avian myeloblastosis viral oncogene homologue (*MYB*) is a transcriptional activator that causes acute leukemia and transforms only hematopoietic cells (22). Together with C/EBP, *MYB* activates transcription of neutrophil elastase 2 (*ELA2*; Ref. 23), and along with other transcriptional cofactors *MYB* also regulates myeloperoxidase (24). Because expression of *MYB* in AML is consistent with a role in malignant proliferation, the overexpression of *MYB* in AML in this study demonstrated that *MYB* might play a role in this process. Apart from these known genes, 16 ESTs that were also extracted in our experiments may prove to be novel AML markers and/or good candidates as molecular targets for drug development.

A great number of genes were selected in our experiments as commonly suppressed in AML. Most of them were known to be expressed specifically in lymphocytes, *e.g.,* immunoglobulins and the MHC. Moreover, interleukin 4 receptor (*IL4R*; Ref. 25) and interleukin 7 receptor (*IL7R*; Ref. 26), whose ligands are required for lymphocyte differentiation, were also suppressed. The majority of the samples examined here corresponded to blast cells of myeloid lineage; lymphocytes accounted for only a small proportion of the cells present in the samples, although the peripheral white cell samples used as controls contained cells of both myeloid and lymphoid lineages. In addition to the known lymphocyte-specific genes, many interesting genes including apoptosis-inducing non-caspase proteases, *e.g.,* calpain 2 (*CAPN2*), granzyme A (*GZMA*), and granzyme B (*GZMB*), were selected (27). From among the 10 caspases spotted on slides, *CASP1, CASP4,* and *CASP5,* all cytokine-processing enzymes (28, 29) were suppressed in our AML samples.

We also retrieved a set of genes whose expression levels correlated with a specific clinical parameter, *i.e.,* response to treatment with a combination of Ara-C and idarubicin. Despite recent advances in chemotherapy, 20–30% of patients who receive chemotherapy still show no response to the drugs and suffer from adverse effects such as myelosuppression and gastrointestinal and/or cardiac toxicity. Often, such patients cannot be treated with another drug at that point because their physical condition has deteriorated too far. To avoid ineffective and potentially damaging treatment, it is urgent for clinicians to have a way to predict sensitivity to chemotherapy before treatment is undertaken (13–16). Many investigators have approached this dilemma by measuring a single or a few factors, but those efforts have failed thus far to establish a method that is widely acceptable in clinical practice because the features of cancer cells in individual patients vary too much; a larger set of factors is required to encompass these differences. Hence, we selected 28 genes on the basis of expression data from an array of thousands of genes and attempted to predict chemosensitivity on the basis of the behavior of those 28 genes. Of 28 genes, *PCNA* was significantly expressed in poor responders but not in good responders. Del Giglio *et al.* (30) reported that high levels of *PCNA* were associated with poor response to Ara-C induction therapy. Moreover, *GADD45A,* reported recently to be involved in DNA repair via FOXO3A (31), is highly expressed in AML cells from poor responders. High expression levels of *GADD45A* in AML cells might be protective against DNA-damaging chemicals such as anthracyclines. Regarding the cases classified into groups 1 and 3, 40 of the 44 cases (including learning and test cases) with positive scores achieved complete remission after treatment; only the 4 patients have thus far failed to show good responses. On the other hand, 17 of the 20 cases with negative scores were unable to show any response (Fig. 2). Moreover,

when we confirmed the Predictive Scores of original 44 cases by using cross validation leave-one-out method (12), 37 of 44 cases were accurately predicted. These data suggested that accuracy of our prediction system was ~85% both in original and test cases. Without this scoring system, AML patients have 20–30% probability for the failure of this protocol of chemotherapy, but the system we developed here can predict the risk of the failure at 80–90% probability. Because the patients with negative scores have a higher possibility to suffer from adverse effect without any good response and lose a chance to survive, the application of other protocols also increase a chance to have a better prognosis and a better quality of life. Although a larger-scale study will be required to further evaluate our scoring system and establish the method of predicting chemosensitivity more precisely, at least we have shown that our goal of "personalized medicine," an appropriate drug for each patient, may be achievable by selection of a set of genes by this kind of approach.

Acknowledgments

This study was performed with thanks to all doctors participating in the Japan Adult Leukemia Study Group, and we also thank Hideaki Ogasawara, Hiroko Bando, Noriko Nemoto, and Noriko Sudo for the fabrication of the cDNA microarray.

References

1. Spellman, P. T., Sherlock, G., Zhang, M. Q., Iyer, V. R., Anders, K., Eisen, M. B., Brown, P. O., Botstein, D., and Futcher, B. Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. Mol. Biol. Cell, *9:* 3273–3297, 1998.

2. Scherf, U., Ross, D. T., Waltham, M., Smith, L. H., Lee, J. K., Tanabe, L., Kohn, K. W., Reinhold, W. C., Myers, T. G., Andrews, D. T., Scudiero, D. A., Eisen, M. B., Sausville, E. A., Pommier, Y., Botstein, D., Brown, P. O., and Weinstein, J. N. A gene expression database for the molecular pharmacology of cancer. Nat. Genet., *24:* 236–244, 2000.

3. Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., Yu, X., Powell, J. I., Yang, L., Marti, G. E., Moore, T., Hudson, J., Jr., Lu, L., Lewis, D. B., Tibshirani, R., Sherlock, G., Chan, W. C., Greiner, T. C., Weisenburger, D. D., Armitage, J. O., Warnke, R., Levy, R., Wilson, W., Grever, M. R., Byrd, J. C., Botstein, D., Brown P. O., and Staudt, L. M. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature (Lond.), *403*: 503–511, 2000.

4. Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., Fluge, O., Pergamenschikov, A., Williams, C., Zhu, S. X., Lonning, P. E., Borresen-Dale, A. L., Brown, P. O., and Botstein, D. Molecular portraits of human breast tumours. Nature (Lond.), *406:* 747–752, 2000.

5. Ono, K., Tanaka, T., Tsunoda, T., Kitahara, O., Kihara, C., Okamoto, A., Ochiai, K., Takagi, T., and Nakamura, Y. Identification by cDNA microarray of genes involved in ovarian carcinogenesis. Cancer Res., *60:* 5007–5011, 2000.

6. Okabe, H., Satoh, S., Kato, T., Kitahara, O., Yanagawa, R., Yamaoka, Y., Tsunoda, T., Furukawa, Y., and Nakamura, Y. Genome-wide analysis of gene expression in human hepatocellular carcinomas using cDNA microarray: identification of genes involved in viral carcinogenesis and tumor progression. Cancer Res., *61:* 2129–2137, 2001.

7. Kitahara, O., Furukawa, Y., Tanaka, T., Kihara, C., Ono, K., Yanagawa, R., Nita, M. E., Takagi, T., Nakamura, Y., and Tsunoda, T. Alterations of gene expression during colorectal carcinogenesis revealed by cDNA microarrays after laser-capture microdissection of tumor tissues and normal epithelia. Cancer Res., *61:* 3544–3549, 2001.

8. Kihara, C., Tsunoda, T., Tanaka, T., Yamana, H., Furukawa, Y., Ono, K., Kitahara, O., Zembutsu, H., Yanagawa, R., Hirata, K., Takagi, T., and Nakamura, Y. Prediction of sensitivity of esophageal tumors to adjuvant chemotherapy by cDNA microarray analysis of gene-expression profiles. Cancer Res., *61:* 6474–6479, 2001.

9. Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. USA, *95:* 14863–14868, 1998.

10. Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., Lander, E. S., and Golub, T. R. Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. Proc. Natl. Acad. Sci. USA, *96*: 2907–2912, 1999.

11. Brown, M. P., Grundy, W. N., Lin, D., Cristianini, N., Sugnet, C. W., Furey, T. S., Ares, M., Jr., and Haussler, D. Knowledge-based analysis of microarray gene expression data by using support vector machines. Proc. Natl. Acad. Sci. USA, *97:* 262–267, 2000.

12. Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Bloomfield, C. D., and Lander, E. S. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science (Wash. DC), *286:* 531–537, 1999.

13. Broxterman, H. J., Sonneveld, P., van Putten, W. J., Lankelma, J., Eekman, C. A., Ossenkoppele, G. J., Pinedo, H. M., Lowenberg, B., and Schuurhuis, G. J. P-glycoprotein in primary acute myeloid leukemia and treatment outcome of idarubicin/cytosine arabinoside-based induction therapy. Leukemia (Baltimore), *14:* 1018–1024, 2000.

14. Kanda, Y., Hamaki, T., Yamamoto, R., Chizuka, A., Suguro, M., Matsuyama, T., Takezako, N., Miwa, A., Kami, M., Hirai, H., and Togawa, A. The clinical significance of CD34 expression in response to therapy of patients with acute myeloid leukemia: an overview of 2483 patients from 22 studies. Cancer (Phila.), *88:* 2529–2533, 2000.

15. Legrand, O., Perrot, J. Y., Baudard, M., Cordier, A., Lautier, R., Simonin, G., Zittoun, R., Casadevall, N., and Marie, J. P. The immunophenotype of 177 adults with acute myeloid leukemia: proposal of a prognostic score. Blood, *96:* 870–877, 2000.

16. Christiansen, D. H., Andersen, M. K., and Pedersen-Bjergaard, J. Mutations with loss of heterozygosity of p53 are common in therapy-related myelodysplasia and acute myeloid leukemia after exposure to alkylating agents and significantly associated with deletion or loss of 5q, a complex karyotype, and a poor prognosis. J. Clin. Oncol., *19:* 1405–1413, 2001.

17. Van Gelder, R. N., von Zastrow, M. E., Yool, A., Dement, W. C., Barchas, J. D., and Eberwine, J. H. Amplified RNA synthesized from limited quantities of heterogeneous cDNA. Proc. Natl. Acad. Sci. USA, *87:* 1663–1667, 1990.

18. Luo, L., Salunga, R. C., Guo, H., Bittner, A., Joy, K. C., Galindo, J. E., Xiao, H., Rogers, K. E., Wan, J. S., Jackson, M. R., and Erlander, M. G. Gene expression profiles of laser-captured adjacent neuronal subtypes. Nat. Med., *5:* 117–122, 1999.

19. Boyer, M. W., Waller, E. K., Bray, R. A., Unangst, T., Johnson, T. S., Phillips, C., Jurickova, I., Winton, E. F., and Yeager, A. M. Cytokine upregulation of the antigen presenting function of acute myeloid leukemia cells. Leukemia (Baltimore), *14:* 412–418, 2000.

20. Carow, C. E., Levenstein, M., Kaufmann, S. H., Chen, J., Amin, S., Rockwell, P., Witte, L., Borowitz, M. J., Civin, C. I., and Small, D. Expression of the hematopoietic growth factor receptor FLT3 (STK- 1/Flk2) in human leukemias. Blood, *87:* 1089–1096, 1996.

21. Yamamoto, Y., Kiyoi, H., Nakano, Y., Suzuki, R., Kodera, Y., Miyawaki, S., Asou, N., Kuriyama, K., Yagasaki, F., Shimazaki, C., Akiyama, H., Saito, K., Nishimura, M., Motoji, T., Shinagawa, K., Takeshita, A., Saito, H., Ueda, R., Ohno, R., and Naoe, T. Activating mutation of D835 within the activation loop of FLT3 in human hematologic malignancies. Blood, *97:* 2434–2439, 2001.

22. Lipsick, J. S., and Wang, D. M. Transformation by v-Myb. Oncogene, *18:* 3047–3055, 1999.

23. Nuchprayoon, I., Simkevich, C. P., Luo, M., Friedman, A. D., and Rosmarin, A. G. GABP cooperates with c-Myb and C/EBP to activate the neutrophil elastase promoter. Blood, *89:* 4546–4554, 1997.

24. Britos-Bray, M., and Friedman, A. D. Core binding factor cannot synergistically activate the myeloperoxidase proximal enhancer in immature myeloid cells without c-Myb. Mol. Cell. Biol., *17:* 5127–5135, 1997.

25. O'Garra, A., and Arai, N. The molecular basis of T helper 1 and T helper 2 cell differentiation. Trends Cell Biol., *10:* 542–550, 2000.

26. Kondo, M., Weissman, I. L., and Akashi, K. Identification of clonogenic common lymphoid progenitors in mouse bone marrow. Cell, *91:* 661–672, 1997.

27. Johnson, D. E. Noncaspase proteases in apoptosis. Leukemia (Baltimore), *14:* 1695–1703, 2000.

28. Wolf, B. B., and Green, D. R. Suicidal tendencies: apoptotic cell death by caspase family proteinases. J. Biol. Chem., *274:* 20049–20052, 1999.

29. Grutter, M. G. Caspases: key players in programmed cell death. Curr. Opin. Struct. Biol., *10:* 649–655, 2000.

30. Del Giglio, A., Drach, J., Kornblau, S. M., Patel, S., Novaes, M., Khetan, R., Sawaya, N., Dorlhiac-Llacer, P., Chamone, D. F., Andreeff, M., *et al.* Quantitative expression of proliferating cell nuclear antigen by Western blot (PCNAWB) in peripheral blasts correlates with remission induction in patients with acute myelogenous leukemia. Leuk. Lymphoma, *19:* 235–241, 1995.

31. Tran, H., Brunet, A., Grenier, J. M., Datta, S. R., Fornace, A. J., Jr., DiStefano, P. S., Chiang, L. W., and Greenberg, M. E. DNA repair pathway stimulated by the forkhead transcription factor FOXO3a through the Gadd45 protein. Science (Wash. DC), *296*: 530–534, 2002.