CANCER GENOMICS & PROTEOMICS *3*: 169-182 (2006)

Gene Expression Profiles of CD133-positive Fractions Predict the Survival of Individuals with Acute Myeloid Leukemia

YOSHIHIRO YAMASHITA[1], JUN OHASHI[2], YUJI HIRAI[3], YOUNG LIM CHOI[1], RURI KANEDA[1], SHIN-ICHIRO FUJIWARA[1,4], YUKIHIRO ARAI[5], MIYUKI AKUTSU[6], CHIZUKO TSUTSUMI[7], YASUSHI MIYAZAKI[7], KENSUKE USUKI[8], MASANAO TERAMURA[3], KINUKO MITANI[5], YASUHIKO KANO[6], MICHAEL C. O'NEILL[9], AKIO URABE[8], MASAO TOMONAGA[7], KEIYA OZAWA[4] and HIROYUKI MANO[1,10]

Divisions of [1]Functional Genomics and [4]Hematology, Jichi Medical University, Tochigi 329-0498; [2]Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Tokyo 113-0033; [3]Department of Hematology, Tokyo Women's Medical University, Tokyo 162-8666; [5]Department of Hematology, Dokkyo University School of Medicine, Tochigi 321-0293; [6]Tochigi Cancer Center, Tochigi 320-0834; [7]Department of Hematology and Molecular Medicine Unit, Nagasaki University, Nagasaki 852-8523; [8]Department of Hematology, Kanto Medical Center NTT EC, Tokyo 141-8625, Japan; [9]Department of Biological Sciences, University of Maryland, Baltimore, Maryland 21250, U.S.A.; [10]CREST, Japan Science and Technology Agency, Saitama 332-0012, Japan

Abstract. *Background: The current classification of acute myeloid leukemia (AML) is based predominantly on the cytogenetic abnormalities and morphology of the malignant blasts and is not always helpful for optimization of the treatment strategy. Gene expression profiles of AML blasts were obtained and a gene expression-based means of predicting the outcome of AML patients was developed. Materials and Methods: CD133-positive hematopoietic stem cell-like fractions were purified from the bone marrow of 99 individuals with AML-related disorders and the expression profiles of ~33,000 human transcripts in these cells were characterized with the use of DNA microarray analysis. Results: The comparison of the expression data between individuals with short- or long-term survival by application of Cox's proportional hazard model led to the identification of four genes, whose expression patterns discriminated between the two groups. The gene expression-based stratification (GES) system, based on a combination of the karyotype approach and the risk index calculated from the expression levels of the four outcome predictor genes, was developed to separate the patients into subgroups with distinct prognoses. Conclusion: DNA microarray analysis of purified fractions provides novel stratification schemes for AML based on the expression profiles of a handful of genes.*

Correspondence to: Hiroyuki Mano, Division of Functional Genomics, Jichi Medical School, 3311-1 Yakushiji, Kawachigun, Tochigi 329-0498, Japan. Tel: +81-285-58-7449, Fax: +81-285-44-7322, e-mail: hmano@jichi.ac.jp

Key Words: DNA microarray, prognosis prediction, gene expression profile.

Acute myeloid leukemia (AML) is characterized by clonal growth of immature leukemic blasts in the bone marrow (BM). Although current chemotherapeutic regimens induce an initial complete remission in >70% of affected individuals, the long-term survival of such patients remains poor (5-year survival rate of <30%) (1). Given that leukemic blasts of individuals with AML differ in their abilities to differentiate into cells of the granulocyte, monocyte, erythrocyte, or megakaryocyte lineages, the French-American-British Cooperative Group (FAB) established a classification scheme for AML (M0 to M7) based on blast morphology and differentiation commitment (2). Although some FAB subtypes have proved to be related to good or poor prognosis, the clinical relevance of this classification scheme is complicated by other clinical parameters. A preceding history of myelodysplastic syndrome (MDS) or anticancer treatment, for example, is an important indicator of poor prognosis (3).

One of the most robust predictors of AML prognosis is the blast karyotype (1, 4). A good prognosis is thus predicted from the presence in leukemic clones of t(8;21), t(15;17), or inv(16) chromosomal rearrangements ("favorable" karyotype), whereas –7/7q–, 11q23, or more complex (affecting three or more chromosomes) abnormalities are indicative of a poor outcome ("adverse" karyotype). Other karyotypes are classified as "intermediate." The World Health Organization (WHO) has proposed a classification of AML that separates

individuals with cytogenetic (or molecular) evidence of t(8;21), t(15;17), inv(16)/t(16;16), or 11q23 abnormalities into distinct subcategories (5). However, this classification is of little help for predicting the prognosis of AML patients with a normal karyotype, who constitute ~50% of the AML population. Given that AML patients with a normal karyotype are considered to be at intermediate risk, the corresponding leukemic blasts may harbor heterogeneous minor genomic rearrangements or mutations.

DNA microarray analysis has the potential to provide a stratification scheme for AML based on gene expression profiles and is able to predict the prognosis of each affected individual (6). To facilitate the development of such a genomics approach to the classification of human leukemias, a large-scale cell bank (the Blast Bank) was set up for the storage of CD133 (AC133)-positive hematopoietic stem cell (HSC)-like fractions purified from individuals with a wide range of leukemic disorders (7, 8). CD133 is specifically expressed on HSCs and hemangioblasts that are $CD34^+CD38^-$ (9, 10). There are at least two advantages to the use of such purified immature fractions for the characterization of AML. First, given that the proportion of leukemic blasts within BM varies substantially (20 to almost 100%) among patients and that leukemic blasts possess the ability to differentiate to various extents, a simple comparison of BM mononuclear cells (MNCs) among heterogeneous AML patients is likely to reveal a large number of changes in gene expression that reflect differences either in the percentage of blasts within BM, or in the differentiation ability of the blasts. Analysis of Blast Bank specimens should thus minimize such population-shift effects (7). Second, although leukemic cells in a given patient comprise a mixture of malignant clones at various levels of differentiation, they are thought to be generated from a small number of leukemic stem cells (LSCs), similar to the situation for normal hematopoiesis (11, 12). Characterization of the LSCs should provide insight into the molecular mechanisms of leukemogenesis. Given that such LSCs are exclusively $CD34^+CD38^-$, the Blast Bank may represent a diverse collection of LSC specimens. Cancer stem cells of malignant glioblastoma, but not their progenies, have also been shown to specifically express CD133 (13).

DNA microarray analysis was used here to characterize the expression profiles of ~44,000 probe sets in the $CD133^+$ fractions of 99 adults with AML-related disorders. Statistical analyses of the resulting large data set provided the basis for a new classification of AML that facilitates prediction of the long-term prognosis of affected individuals.

Materials and Methods

Purification of $CD133^+$ cells. Informed consent was obtained from each of the study subjects and the study was approved by the appropriate institutional review boards. Table I shows the clinical characteristics of the study subjects. MNCs were isolated by density gradient centrifugation from BM aspirates of each patient and were subjected to chromatography on a miniMACS column (Miltenyi Biotec, Auburn, CA, USA) with magnetic bead-conjugated monoclonal antibodies to CD133 (AC133 MicroBeads; Miltenyi Biotec) as described previously (7). In most instances, the $CD34^{hi}CD38^{lo}$ fraction constituted >90% of the eluate of the affinity column, as judged by flow cytometry.

Microarray analysis. Total RNA was extracted from the $CD133^+$ cell preparations with the use of an RNeasy Mini column and RNase-free DNase (Qiagen, Valencia, CA, USA). It was then subjected to two rounds of amplification of mRNA with T7 RNA polymerase (14); the high fidelity of the amplification step has been demonstrated previously (15). The resulting cRNAs were labeled with biotin and subjected to hybridization with GeneChip HGU133 A&B microarrays with the GeneChip system (Affymetrix, Santa Clara, CA, USA). The fidelity of the microarray data was confirmed by quantitative RT-PCR analysis.

Method for real-time RT-PCR. Portions of non-amplified cDNA were subjected to PCR with a QuantiTect SYBR Green PCR Kit (Qiagen). The amplification protocol comprised incubations at 94°C for 15 sec, 60°C for 30 sec and 72°C for 60 sec. Incorporation of the SYBR Green dye into the PCR products was monitored in real time with an ABI PRISM 7700 sequence detection system (PE Applied Biosystems, Foster City, CA, USA), thereby allowing determination of the threshold cycle (C_T) at which exponential amplification of products begins. The C_T values for cDNAs corresponding to *GAPDH* and the target genes were used to calculate the abundance of target gene mRNA relative to that of *GAPDH* mRNA. The primer sequences used for RT-PCR were as follows:

FGFR1:
Correlation coefficient between RT-PCR and microarray data=0.746.
Sense primer: 5'-CCACCAGAGTGATGTGTG GTCTTT–3'
Antisense primer: 5'-CATCATCATGTACAGCTC GTTGGT–3'

NRLN1
Correlation coefficient between RT-PCR and microarray data=0.638.
Sense primer: 5'-AATAGACGGAAAATGCTGCA AGGT–3'
Antisense primer: 5'-TGAGGTGGTCTCTCAGTCTCC AGT–3'

ZNF6
Correlation coefficient between RT-PCR and microarray data=0.779.
Sense primer: 5'-ATCTGGTGCAAAAACAGA AAGGTG–3'
Antisense primer: 5'-GGCGGGTTTATGCAGTATTT AACAG–3'

SCGF
Correlation coefficient between RT-PCR and microarray data=0.715.
Sense primer: 5'-TACTACGTCTGCGAGTTC CCCTTC–3'
Antisense primer: 5'-GCCCCTTCAAGGAAAGA CACTAAC–3'

HOXA9
Correlation coefficient between RT-PCR and microarray data=0.586.
Sense primer: 5'-CTCAGGTTGTTTATGAGG GGAAAA–3'
Antisense primer: 5'-ATGAATCTATGCATCCCC GAGAAC–3'

ANGPT1
Correlation coefficient between RT-PCR and microarray data=0.719.
Sense primer: 5'-GGCTGGGGAATGAGTTT ATTTTTG–3'
Antisense primer: 5'-AAATCAGCACCGTGTAAG ATCAGG–3'

241376_at (EST)
Correlation coefficient between RT-PCR and microarray data=0.672.
Sense primer: 5'-CAACTCGAAGCTCAAA TACCCTCA–3'
Antisense primer: 5'-ACCGTTTATACACCAACGG TCACA–3'

FLJ13063
Correlation coefficient between RT-PCR and microarray data=0.502.
Sense primer: 5'-AGAGTTCTGCTGTGT CCTCTG–3'
Antisense primer: 5'-CAGGACAGTGCTGAAC CAATG–3'

TSPAN-2
Correlation coefficient between RT-PCR and microarray data=0.479.
Sense primer: 5'-GCAGTTGAAAATTGTG GGAAAGAG–3'
Antisense primer: 5'-CCCACACACAACTAGGA GAAGATG–3'

KIAA0830
Correlation coefficient between RT-PCR and microarray data=0.575.
Sense primer: 5'-TCCAGAGGCATCTGAGG GAAGTAG–3'
Antisense primer: 5'-ATGGCCATGAAGTATGA GATGGTG–3'

DJ79P11.1
Correlation coefficient between RT-PCR and microarray data=0.625.
Sense primer: 5'-CCATCCTGCAGTATAG ATGGGACA–3'
Antisense primer: 5'-GATTCAGGGCATAAGGC AAAACTC–3'

POU4F1
Correlation coefficient between RT-PCR and microarray data=0.643.
Sense primer: 5'-ATGAACAAGCCTGAGC TCTTCAAC–3'
Antisense primer: 5'-GAGAATTTCATCCGCT TCTGCTTC–3'

IGHM
Correlation coefficient between RT-PCR and microarray data=0.593.
Sense primer: 5'-CAGAAGAACATCGGAG ACCAGAGA–3'
Antisense primer: 5'-AACAAGCGTATACACAG CAAAGCA–3'

GAPDH
Sense primer: 5'-GTCAGTGGTGGACCT GACCT–3'
Antisense primer: 5'-TGAGCTTGACAAAGTG GTCG–3'

Statistical analysis. The mean expression intensity of the internal positive control probe sets (http://www.affymetrix.com/support/technical/mask_files.affx) was set to 500 units (U) in each hybridization and the fluorescence intensity of each test probe set was normalized accordingly. All normalized array data are available at the Gene Expression Omnibus web site (http://www.ncbi.nlm.nih.gov/geo) under the accession number GSE1427. All expression values were transformed to logarithms prior to statistical analyses. Hierarchical clustering of the data set and Student's *t*-test were performed with GeneSpring 7.0 software (Silicon Genetics, Redwood, CA, USA). Principal component analysis (16) and survival analyses were performed with the SAS software package (ver. 8.0.2).

Results

Purification of CD133$^+$ HSC-like fractions. The number of CD133$^+$ cells isolated from BM MNCs varied markedly among FAB subtypes of AML. For individuals whose leukemic blasts had a low differentiation capacity (FAB subtypes M0 to M2), for example, the blasts purified by CD133-based affinity chromatography constituted ≥30% with MNCs. In contrast, the blasts purified from individuals with FAB subtypes M3 (characterized by prominent differentiation to the promyelocyte level) or M5 (differentiation to the promonocyte level) constituted <1% of MNCs. In both the latter instances, large promyelocytes (M3) or large promonocytes (M5) constituted a major proportion of the BM MNCs, whereas the corresponding column eluates contained only medium-sized, agranular blasts with a high nucleus-to-cytoplasm ratio (Figure 1A).

The affinity column thus appeared to select for a minor population of cells at a highly immature level of differentiation. To verify that such cells in the column eluates did indeed represent leukemic clones, fluorescence *in situ* hybridization analysis was performed with the cells purified from several individuals. In the case of FAB subtype M3, described above, a t(15;17) was detected in 75% of the purified blasts (data not shown). Similarly, in an AML case characterized by 5q– or a case of chronic myelogenous leukemia characterized by t(9;22), the corresponding abnormality was detected in >90 or 70.7% (8) of cells in the column eluates, respectively. These findings indicate that most of the cells in the column eluates were malignant clones, especially given the only moderate sensitivity of fluorescence *in situ* hybridization.

Gene expression profiles of the Blast Bank specimens. All the specimens were collected from the patients before initiation of chemotherapy. The expression profiles of >44,000 probe

Table I. *Clinical characteristic of the study subjects.*

Blast Bank ID	Age (yr)	Gender	Disease	FAB subtype	Karyotype
ID020	66	Female	AML evolved from MDS	M2	Normal
ID023	34	Female	*de novo* AML	M6	Others>=3
ID026	46	Male	*de novo* AML	M3	Others≤2
ID027	49	Male	AML evolved from MDS	M2	Normal
ID028	62	Male	AML evolved from MDS	M2	Normal
ID032	75	Male	RAEB		Others>=3
ID035	61	Male	*de novo* AML	M2	Others≤2
ID036	74	Male	*de novo* AML	M0	Normal
ID042	67	Female	AML evolved from MDS	M2	Others>=3
ID045	69	Male	RAEB		+8
ID046	84	Male	*de novo* AML	M2	–7
ID051	42	Male	RAEB		Others>=3
ID054	84	Male	AML evolved from MDS	M2	–7
ID062	72	Male	*de novo* AML	M0	Normal
ID063	67	Female	RAEB		Others>=3
ID066	73	Male	RAEB		t(8;21)
ID076	37	Male	*de novo* AML	M2	t(8;21)
ID077	55	Male	RAEB		+8
ID083	64	Female	*de novo* AML	M4	Normal
ID087	66	Male	*de novo* AML	M5	Normal
ID093	53	Female	*de novo* AML	M5	Others≤2
ID098	66	Male	AML evolved from MDS	M0	Normal
ID104	72	Male	*de novo* AML	M0	Others>=3
ID107	48	Female	*de novo* AML	M5	Others≤2
ID109	86	Male	*de novo* AML	M1	Normal
ID127	41	Male	*de novo* AML	M2	Normal
ID139	50	Male	*de novo* AML	M1	Others≤2
ID142	38	Male	*de novo* AML	M2	t(8;21)
ID148	74	Male	AML evolved from MDS	M2	Normal
ID154	49	Male	RAEB		Normal
ID174	51	Male	*de novo* AML	M0	Normal
ID180	47	Male	*de novo* AML	M4	Normal
ID183	50	Female	*de novo* AML	M2	t(8;21)
ID188	59	Male	*de novo* AML	M2	t(8;21)
ID195	61	Female	*de novo* AML	M2	t(8;21)
ID205	39	Male	*de novo* AML	M1	Normal
ID215	52	Male	RAEB		Others>=3
ID226	52	Male	*de novo* AML	M2	t(8;21)
ID227	29	Male	*de novo* AML	M2	Normal
ID234	68	Male	RAEB		Others≤2
ID239	48	Male	*de novo* AML	M2	t(8;21)
ID243	54	Female	RAEB		Others>=3
ID262	70	Female	*de novo* AML	M2	Others≤2
ID265	65	Male	*de novo* AML	M7	Normal
ID266	38	Male	*de novo* AML	M4	Normal
ID267	80	Male	*de novo* AML	M2	Normal
ID269	32	Female	*de novo* AML	M4	Others≤2
ID270	46	Female	*de novo* AML	M1	Others≤2
ID272	57	Female	*de novo* AML	M4	+8
ID274	67	Male	*de novo* AML	M1	–7
ID275	70	Male	AML evolved from MDS	M2	Others>=3
ID276	64	Male	de novo AML	M2	–7, >3
ID277	23	Female	*de novo* AML	M4	inv16
ID278	16	Female	de novo AML	M4	Normal
ID279	68	Male	*de novo* AML	M2	t(8;21)
ID280	68	Male	*de novo* AML	M1	–7
ID288	61	Male	de novo AML	M5	Normal
ID292	74	Male	*de novo* AML	M6	Normal
ID305	63	Male	*de novo* AML	M2	Normal
ID306	43	Male	*de novo* AML	M1	Normal
ID310	75	Male	*de novo* AML	M0	Normal
ID313	88	Female	AML evolved from MDS	M1	Others>=3
ID314	55	Female	*de novo* AML	M5	+8
ID315	45	Female	RAEB		Others>=3
ID316	30	Male	*de novo* AML	M1	Others>=3
ID317	36	Male	*de novo* AML	M2	t(8;21)
ID318	53	Male	*de novo* AML	M2	Normal
ID319	47	Male	*de novo* AML	M6	Others>=3
ID321	68	Female	*de novo* AML	M2	t(8;21)
ID325	49	Male	*de novo* AML	M2	Others>=3
ID326	68		*de novo* AML	M1	Normal
ID329	79	Female	*de novo* AML	M2	Normal
ID338	57	Male	RAEB		Normal
ID339	72	Female	*de novo* AML	M4	inv16
ID347	52	Male	*de novo* AML	M2	Others≤2
ID349	70	Male	*de novo* AML	M6	Others>=3
ID355	69	Male	*de novo* AML	M2	t(8;21)
ID362	59	Female	*de novo* AML	M2	Normal
ID363	67	Male	*de novo* AML	M0	Others>=3
ID375	32	Male	*de novo* AML	M2	Normal
ID376	23	Male	*de novo* AML	M0	Others≤2
ID378	28	Female	*de novo* AML	M2	t(8;21)
ID379	62	Male	*de novo* AML	M2	Normal
ID380	51	Female	*de novo* AML	M6	Others>=3
ID382	49	Female	AML evolved from MDS	M2	Others>=3
ID385	79	Male	RAEB		Normal
ID388	45	Male	*de novo* AML	M5	inv16
ID395	80	Male	AML evolved from MDS	M0	Normal
ID402	63	Male	*de novo* AML	M1	Others≤2
ID409	72	Female	*de novo* AML	M1	Others≤2
ID410	67	Female	RAEB		Normal
ID413	61	Female	*de novo* AML	M2	Normal
ID414	61	Male	*de novo* AML	M0	Normal
ID415	85	Female	*de novo* AML	M2	Others>=3
ID416	42	Female	*de novo* AML	M2	Normal
ID418	57	Female	*de novo* AML	M2	Normal
ID421	81	Female	RAEB		Normal
ID427	62	Male	*de novo* AML	M5	Normal
ID430	66	Female	RAEB		Others≤2

AML=acute myeloid leukemia; MDS=myelodysplastic syndrome; FAB=French-American-British Cooperative Group; RAEB=refractory anemia with excess of blasts.

sets (corresponding to ~33,000 human genes) were then determined for Blast Bank specimens derived from 83 individuals with AML and 16 individuals in the advanced stage [refractory anemia with excess of blasts (RAEB)] of MDS. The clinical characteristics of the study subjects are provided in Table I.

For analysis of the expression data, the criterion that the expression level of a given probe set should receive the "Present" call (from Microarray Suite 5.0 software) in at least 30% (n=30) of the samples was applied in order to exclude transcriptionally-silent genes from the analysis. A total of 11,595 probe sets passed this selection window. Unsupervised two-way hierarchical clustering analysis (17) was then applied to the 99 patients based on the expression profiles of these 11,595 probe sets, generating a dendrogram of the subjects (Figure 1B). Six out of the 16 patients with RAEB (MDS) clustered together in the dendrogram, whereas the remaining ten RAEB patients were intermixed with the AML cases. Given that eleven of our AML subjects had experienced a previous MDS phase, it was not surprising that the gene expression profiles did not clearly separate RAEB from heterogeneous AML.

Patients with a normal karyotype or "other" abnormalities (<3 or ≥3) were widely distributed throughout the dendrogram, indicative of the highly heterogeneous character of their blasts. Previous unsupervised clustering analysis of the gene expression profiles of BM MNCs from individuals with AML separated the patients into subgroups that were strongly related to the FAB subtype (18, 19). However, such a relationship was not apparent in our analysis of $CD133^+$ cells (data not shown), suggesting that changes in gene expression that accompany the differentiation of leukemic blasts within BM might greatly influence the overall gene expression profiles of MNCs.

Gene expression profiles linked to good prognosis. Among the 83 AML patients studied, 66 individuals were treated with standard chemotherapeutic regimens according to the protocols of the Japan Adult Leukaemia Study Group (JALSG). Kaplan-Meier analysis of these 66 cases revealed that, although the prognosis of individuals with the adverse karyotype was significantly worse than that of the other two karyotype groups (long-rank test, $p<0.001$), the prognosis of patients with the favorable karyotype did not differ from that of those with the intermediate one ($p=0.06$) (Figure 2). The poor prognosis of individuals with the adverse karyotype (5-year survival rate of ~10%) has been confirmed extensively (1, 4, 20), but the intermediate group apparently contains both patients with a curable disorder and those with an intractable one. It is, therefore, of clinical importance to be able to identify individuals with a *bona fide* good prognosis from among those with a favorable or intermediate karyotype.

We next compared the transcriptomes of individuals with these two karyotype designations. Among the 66 AML patients who underwent standard chemotherapy, the blasts of 16 and 39 individuals had the favorable and intermediate karyotypes, respectively. Probe sets that received the "Present" call in ≥10% (n=6) of these 55 cases were selected first. The selected 17,724 probe sets were then screened for those whose expression differed significantly between the two karyotype groups (Student's *t*-test, $p<0.001$), resulting in the identification of 378 such sets. Application of principal component analysis reduced the number of principal gene expression patterns from 378 to three. On the basis of the calculated coordinates of these three principal expression patterns, the 55 samples were projected into a virtual three-dimensional space (Figure 3A). The samples with favorable and intermediate karyotypes were not clearly separated from each other, however.

We, therefore, examined whether it was possible to identify a gene expression profile, or "molecular signature," that was directly linked to the prognosis of leukemic patients. Among the 55 patients with favorable or intermediate karyotypes who underwent standard chemotherapy, 14 individuals survived for ≥755 days (long-term survivors) and eight individuals died within 365 days (short-term survivors) after the first chemotherapy. The application of the Student's *t*-test to the gene expression data for these 22 cases identified five probe sets, the expression level of which contrasted the two groups ($p<0.001$) (Table II). Principal component analysis and three-dimensional projection indicated that the molecular signature was clearly different between the two patient classes (Figure 3B).

Gene expression-based stratification scheme for AML. We next attempted to develop a gene expression-based stratification (GES) system for AML. Given that the adverse karyotype is a reliable indicator of poor prognosis, we tried to construct a GES scheme to isolate those individuals with good prognosis from other patients. The 66 AML patients who underwent standard chemotherapy were divided into training (n=44) and test (n=22) sets (Figure 4A). The former contained 37 patients with favorable or intermediate karyotypes, seven and six of whom were long-term or short-term survivors, respectively. To identify a prognosis-related molecular signature, the Student's *t*-test was first applied to the expression data for these latter 13 cases to isolate probe sets whose expression differed significantly ($p<0.01$) between the two groups. The expression profiles of the resulting 38 probe sets were then subjected to Cox's proportional-hazard regression analysis (21) ($p<0.01$) for the 13 selected individuals. Four independent probe sets (outcome predictor genes) were finally isolated, the expression profiles of which are shown in a dendrogram in Figure 3B. The risk index (RI)

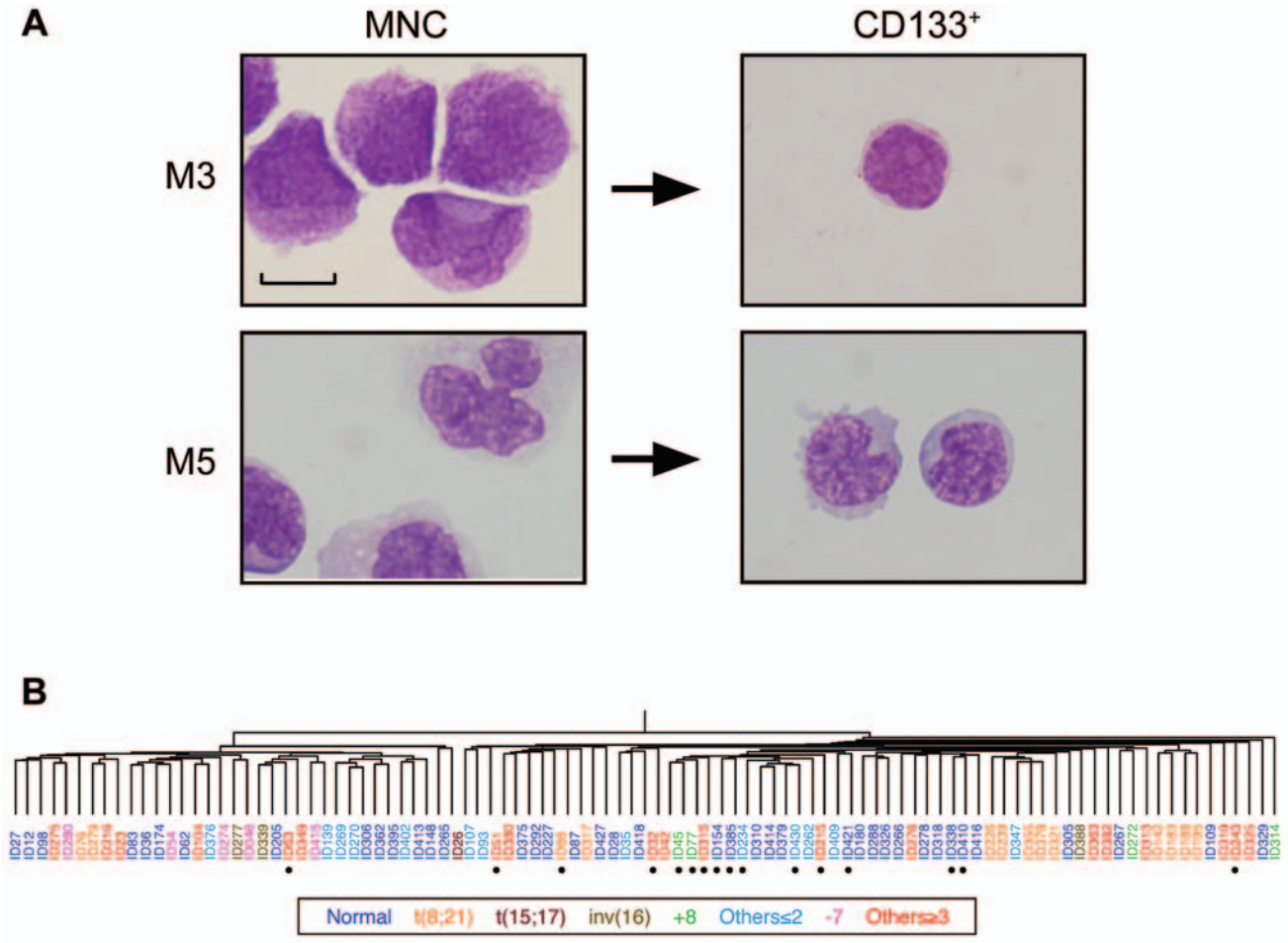

Figure 1. *Clustering analysis of purified CD133+ fractions from individuals with acute myeloid leukemia (AML)-related disorders. A) The CD133+ fractions (right panels) were purified from bone marrow mononuclear cells (MNCs) (left panels) of AML patients of FAB subtypes M3 (upper panels) or M5 (lower panels). The cells were stained with Wright-Giemsa solution. Scale bar, 10 μm. B) Unsupervised clustering of the study subjects based on the similarity in expression profiles of 11,595 probe sets. The karyotype of each patient is colored differentially. Individuals with refractory anemia with excess blasts are indicated by dots.*

(22) was calculated for each patient based on the expression intensity and the parameter estimate (Table III) for each of these four probe sets.

A GES system for AML, based on a combination of this RI-based classification scheme and the karyotype-based scheme, is thus proposed (Figure 5A). AML patients with an adverse karyotype are classified as GES class III, whereas the other patients are classified as either GES class I (RI < –9.36) or GES class II (RI ≥ –9.36) on the basis of the calculated RI for the four outcome predictor genes. To examine whether the GES system is able to select patients with good prognosis, Kaplan-Meier analysis was performed with the 44 individuals in the training set classified according either to the karyotype-based scheme (Figure 5B) or to our GES scheme (Figure 5C). The prognosis of the favorable group of the karyotype-based classification did not

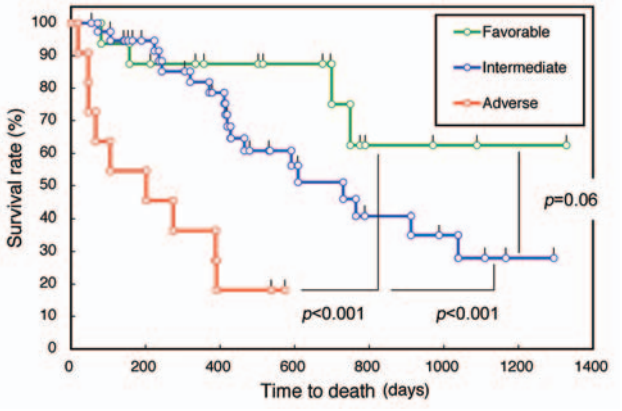

Figure 2. *Long-term survival according to the keryotype-based stratification of the 66 acute myeloid leukemia patients treated with standard chemotherapy.*

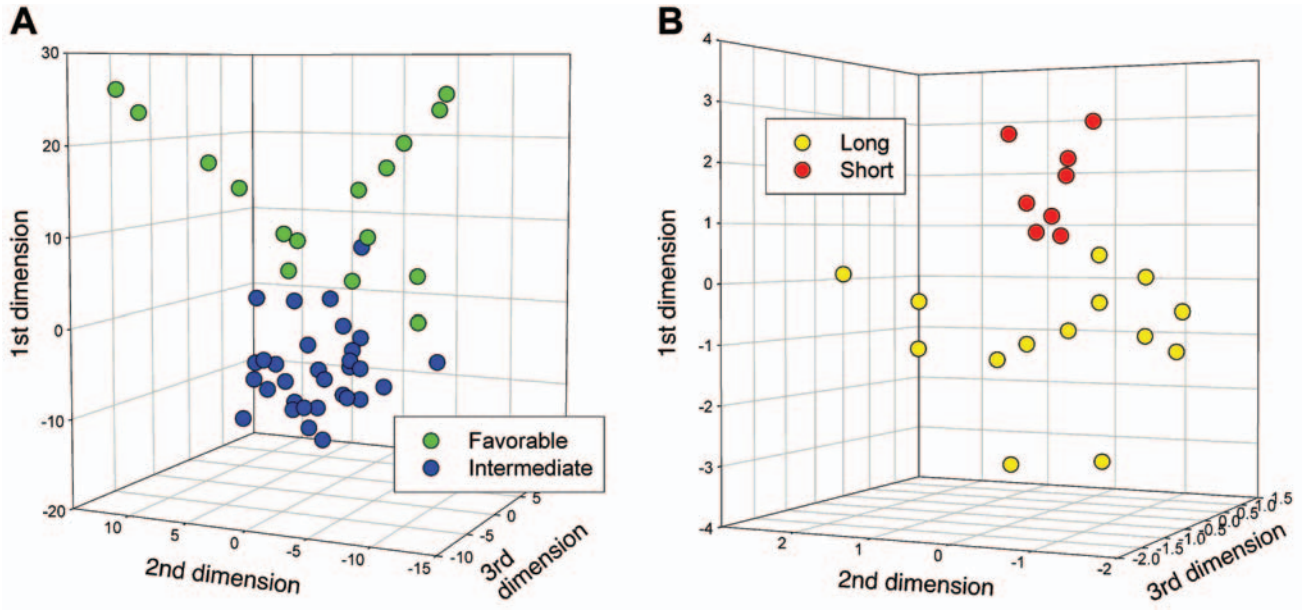

Figure 3. *Comparison of gene expression profiles by principal component analysis. A) Three principal components were derived from the gene expression patterns of 378 probe sets that contrasted acute myeloid leukemia (AML) patients with favorable or intermediate karyotypes who underwent standard chemotherapy. Samples were projected into a virtual space based on the coordinates of the three components. B) Sample projection was similarly performed for five probe sets that contrasted the long-term and short-term survivors among the patients in A.*

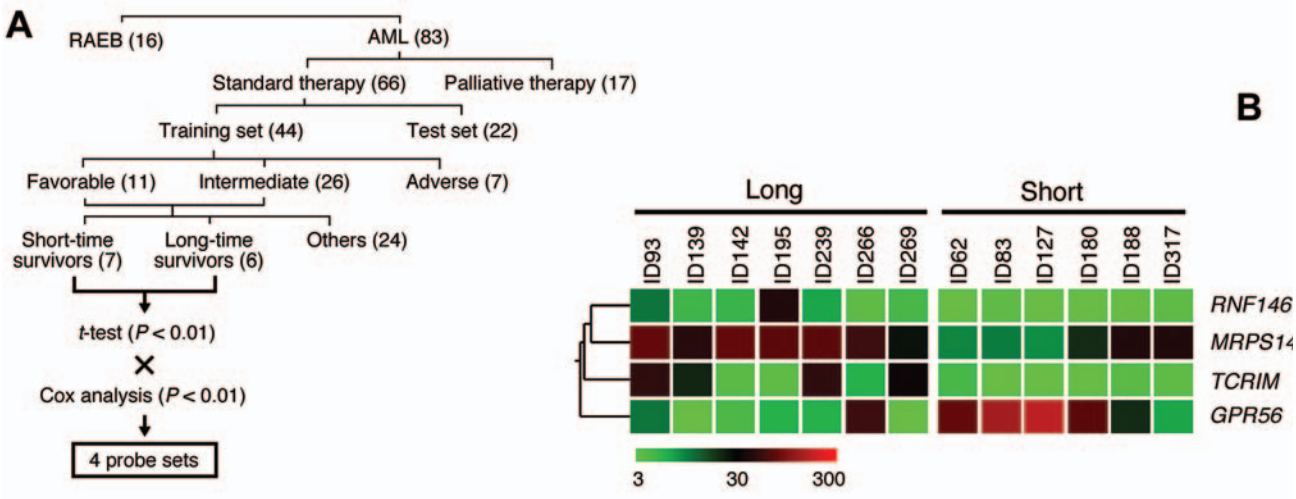

Figure 4. *Isolation of probe sets linked to survival time. A) Approach adopted to identify four probe sets whose expression contrasted the long-term and short-term survivors in a training set of acute myeloid leukemia (AML) patients. The numbers in parentheses indicate the corresponding numbers of subjects. RAEB; refractory anemia with excess of blasts. B) Gene tree for the expression levels (color coded as indicated by the scale at the bottom) of the four human genes identified in A for CD133$^+$ cells derived from seven long-term and six short-term survivors. Each row corresponds to a single gene (symbol shown at right) and each column to a different patient (Blast Bank ID shown at top).*

differ significantly (log-rank test, $p=0.10$) from that of the intermediate group. In contrast, the GES system clearly separated the patients into three groups with distinct prognoses. Importantly, only GES class I contained the long-term survivors.

The efficacy of the GES system was then examined with the 22 individuals in the test set, who were also classified by either the karyotype-based (Figure 5D) or GES (Figure 5E) schemes. Again the former scheme failed to separate the patients into prognosis-related

Table II. *Expersssion intensities of five probe sets linked to survival of 14 long-term and eight short-term survivors of acute myeloid leukemia.*

Affymetrix designation	Gene symbol	GenBank accession no.	Long-term survivors												
			ID026	ID035	ID087	ID093	ID139	ID142	ID174	ID195	ID226	ID227	ID239	ID266	ID269	ID270
217147_s_at	TRIM	AJ240085	22.2	22.4	5.3	42.9	26.5	4.9	0.9	3.6	14.2	6.4	43.0	11.0	30.5	27.9
203098_at	CDYL	AL050164	6.9	15.5	8.4	18.8	49.9	10.2	20.5	6.0	42.5	11.7	60.8	20.3	23.5	30.8
226333_at	IL6R	AV700030	16.3	34.0	68.4	39.2	20.2	52.3	60.5	80.3	54.2	60.6	49.4	107.4	50.6	24.4
243023_at		N34402	1.7	10.2	6.9	4.5	31.3	20.8	16.7	7.1	2.0	8.7	2.9	3.0	29.3	2.4
200663_at	CD63	NM_001780	170.0	73.0	132.4	181.5	74.4	97.7	60.4	94.2	112.9	112.8	102.1	68.2	82.0	76.6

Affymetrix designation	Gene symbol	GenBank accession no.	Short-term survivors								
			ID027	ID062	ID083	ID127	ID180	ID188	ID288	ID317	
217147_s_at	TRIM	AJ240085	1.3	7.3	1.1	2.9	2.7	5.0	1.5	4.1	
203098_at	CDYL	AL050164	48.5	60.1	35.1	30.7	44.3	67.4	33.2	47.0	
226333_at	IL6R	AV700030	177.2	146.2	127.5	72.5	120.7	85.7	61.8	70.1	
243023_at		N34402	28.8	33.5	33.4	23.4	24.5	10.8	19.7	19.5	
200663_at	CD63	NM_001780	57.9	65.9	67.7	52.7	61.8	49.4	62.5	76.6	

groups. In contrast, the GES scheme efficiently isolated the long-term survivors.

Discussion

In the present study, we attempted to develop a new classification scheme for AML based on the gene expression profiles of purified highly immature leukemic cells. Microarray analysis has proved effective for the prediction of prognosis in other hematological malignancies including diffuse large B cell lymphoma (DLBCL) (23-25) and acute lymphoblastic leukemia (ALL) (26, 27). The clinical specimens for these latter two disorders, however, appear to be more homogeneous than are those for AML. Lymphoma clones constitute most of the cell population within the affected lymph nodes of individuals with DLBCL and most ALL clones remain at an early precursor stage of the B or T cell lineage. The clinical specimens derived from individuals with DLBCL or ALL are thus probably representative of the corresponding malignant clones.

Caution is warranted, however, in the interpretation of microarray data for certain blood cell disorders. In Hodgkin's lymphoma, for example, the malignant cells constitute only a small proportion of cells within lymph nodes, most of which are inflammatory cells and normal lymphocytes. Hodgkin's lymphoma cells have thus been isolated with the use of laser-capture microdissection before microarray analysis (28, 29). Similar caution is required with AML, given the heterogeneity in both the proportion and differentiation ability of the blasts within BM. Indeed, our data indicated that so-called "leukemic blasts" are highly heterogeneous with regard to their differentiation ability, even within the BM of a single patient (Figure 1A), consistent with the notion of the presence of multiple types of LSC (12).

In the present study, statistical analysis of the gene expression profiles of $CD133^+$ HSC-like fractions resulted in the isolation of a very small number of outcome-predictor genes, including those for G protein-coupled receptor 56 (GPR56; GenBank accession number, NM_005682), ring finger protein 146 (RNF146; NM_030963), mitochondrial ribosomal protein S14 (MRPS14; NM_022100) and T cell receptor-interacting protein (TCRIM; NM_016388). Their predicted amino acid sequences indicate that GPR56 belongs to the family of cell surface proteins with seven transmembrane domains and that RNF146 participates in protein ubiquitination. MRPS14 is a component of the 28S subunit of mitochondrial ribosomes and contributes to protein synthesis in this organelle (30). TCRIM is a transmembrane protein that undergoes phosphorylation by Src family kinases (31). TCRIM potentially functions as a scaffold protein that recruits a variety of signaling proteins *via* its Src homology 2 (SH2) domain.

Table III. *The outcome predictor genes for acute myeloid leukemia (AML).*

Affymetrix designation	Gene symbol	GenBank accession no.	Parameter estimate	*P* value
203801_at	MRPS14	NM_022100	–2.30663	0.003
212070_at	GPR56	NM_005682	0.69679	0.0045
217147_s_at	TCRIM	NM_016388	–0.79050	0.0041
244517_x_at	RNF146	NM_030963	–0.97626	0.0018

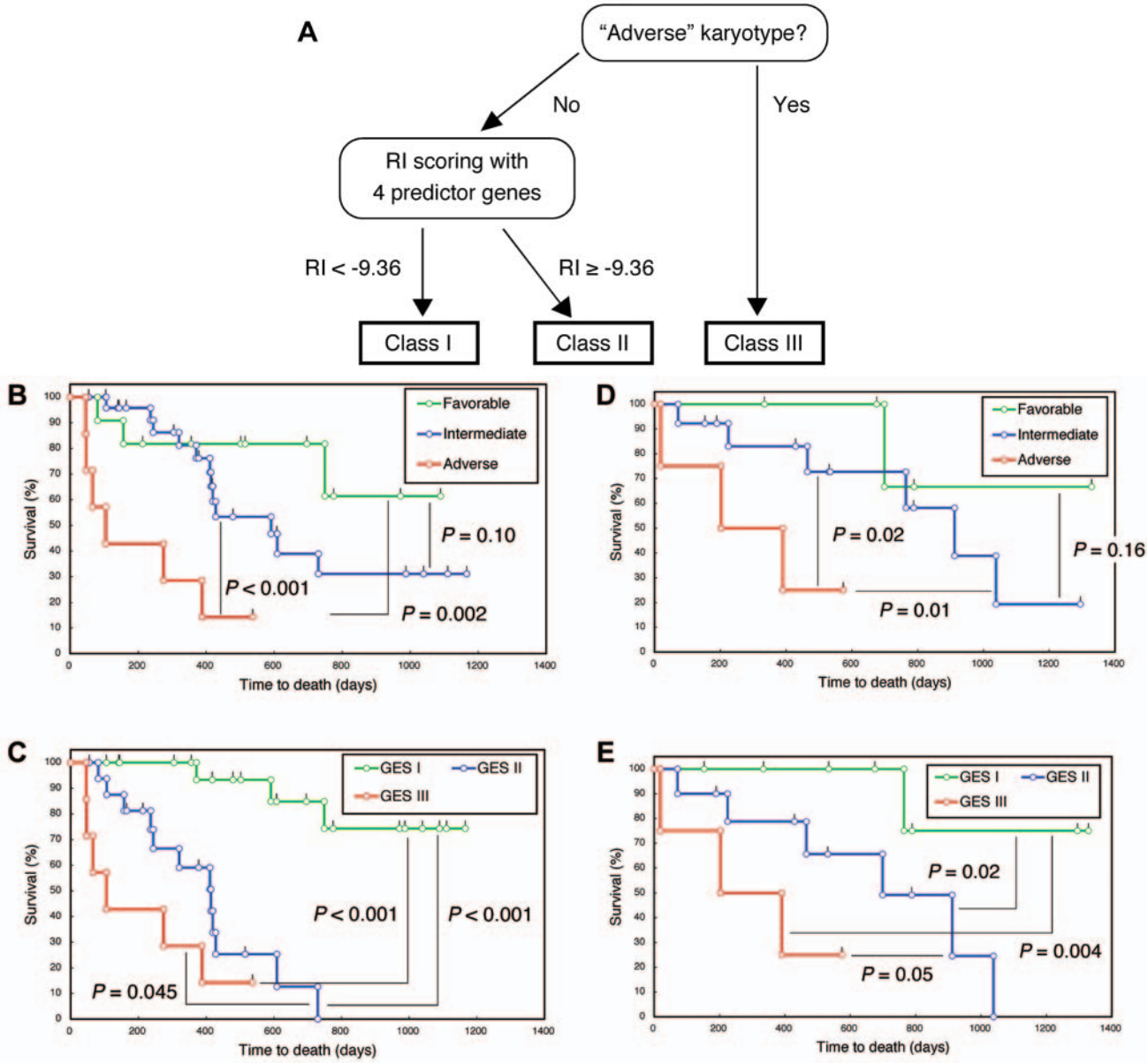

Figure 5. *The gene expression-based stratification (GES) system for acute myeloid leukemia (AML). A) Flow chart for GES classification of AML patients. B–E) Kaplan-Meier analysis of the 44 AML patients in the training set (B, C) and the 22 patients in the test set (D, E) classified according to karyotype-based stratification (B, D) or to the GES system (C, E).*

Table IV. *Expression intensities of 23 pobe sets linked to prognosis in the acute myeloid leukemia patients treated with standard chemotherapy.*

Affymetrix designation	Gene symbol	Genbank accession no.	Good prognosis																					
			ID026	ID035	ID076	ID087	ID093	ID107	ID139	ID142	ID174	ID183	ID195	ID226	ID227	ID239	ID265	ID266	ID269	ID270	ID277	ID278	ID279	ID314
37384_at	PPM1F	D13640	180.9	301.1	129.8	254.6	304.9	210.7	163.2	158.4	171.1	46	448.7	177.6	251.7	129.5	107.4	230.3	136.7	112.9	90	333.5	130	111.8
243579_at	MSI2	BF029215	84.5	40	21.9	170.4	39.7	32.3	30.3	11.4	59	21.3	97.1	47.5	62.5	42.8	35.3	72.5	18	12.5	15.9	208.2	20.2	47.2
228988_at	ZNF6	AU157017	10.6	16.2	7.7	28.8	55.5	57.2	2.4	4.9	138.3	5.5	8.7	20.3	22.1	18.2	4.9	25.2	4.5	1.9	5	57.7	4.6	5.8
228708_at		BF438386	23.2	10.5	54.5	44.7	16.6	20.1	39.6	10	38.8	57.9	12.9	22.5	59.2	97.7	89.4	49.8	60.7	63.3	112.4	171	32	3.2
225651_at	FLJ25157	BF431962	79.7	167	135	64.5	83.6	44.7	89.7	99.1	20.4	83.2	71.9	113	48.7	143	416.2	50.3	79.3	85.5	69.3	22.9	187.2	21.6
225351_at	HT011	AK027029	66.1	62.1	174.7	67.1	73.4	42.2	90.6	193.2	40.5	183.6	164.2	311.1	46.8	407.5	494.6	42.3	302	31	87.2	47	163.3	32.3
224516_s_at	HSPC195	BC006428	52.5	212.2	30.4	115.7	78.7	46.4	197.1	97.5	134	123	164.3	145	115.7	271.1	59.6	150.2	84.2	137.1	295.4	275	123.7	32.4
224367_at	DJ79P11.1	AF251053	68.5	44.3	9.2	18.1	29.4	158.7	5.8	9.9	188.8	26.4	15.8	27.7	12.4	10.8	37.9	97	4.6	7.2	14.4	177.4	5.4	14.2
219498_s_at	BCL11A	NM_1186	55.2	62.4	110.5	26.9	44	51	75.5	35.5	118.9	123.4	51.9	70.1	45.1	100.5	296.5	63.9	58.4	27.2	84.4	98.8	31.1	18.5
217975_at	LOC51186	NM_016303	58.4	32.6	6.1	3.3	27.9	40.9	5.5	16.1	54.4	20.9	22.5	18.9	31.2	6.7	244	55.5	7.3	6.6	9.7	208.2	6.5	86.9
215111_s_at	TSC22	AK027071	76.5	64	112.5	66.7	241.3	99.9	120.1	37.5	268.1	64.2	83.9	254.7	112.2	288	533.7	103.7	144.7	89.9	573.8	294.8	115.1	3.8
214651_s_at	HOXA9	U41813	14.7	14.6	7.5	55.6	185.4	163.1	7.8	11.9	134.3	3	4.4	19.2	61	8.1	190.7	114.6	0.6	2.3	118.6	401.7	0.9	19.9
212827_at	IGHM	X17115	753.6	151.3	35.5	222.3	136.6	142.3	47.2	57.6	214	36.7	62.8	23.8	138.8	24.8	69.9	49.1	21.6	20.8	205.3	185.4	55.4	48.7
211709_s_at	SCGF	BC005810	300	683.9	1184.4	268.5	552.3	3194.6	365.7	2041.3	245.1	2181.9	2493.3	2796.9	315.3	2800.3	373.8	432	788.9	517.7	643.6	2391.3	1150.7	4488.4
211341_at	POU4F1	L20433	59.6	612.2	233.8	28.9	16.6	5.5	29.9	813	11.4	607.5	1075	560.6	41	418.5	13.7	5.1	143	112.2	15.6	20.4	179.4	34.2
209905_at	HOXA9	AI246769	72.8	5.2	1.3	19	15.5	306.4	1.3	3.4	139.6	4.8	6	2.8	85.1	1.6	204.2	193.6	0.9	0.5	198.4	797.2	0.9	40
206478_at	KIAA0125	NM_014792	19	21.7	12.1	45.9	34	10.5	41.1	34.5	303.6	28.1	26.1	35.4	42.8	16.2	146.4	42	38.7	67.3	97.1	144.4	5.9	8.6
205609_at	ANGPT1	NM_001146	50	807.3	22.7	43.4	66.3	70.9	354.2	19.8	130.9	41.4	104.4	21.6	277.3	25.5	82.1	637.4	18	347.6	688.9	357	47	2.5
205608_s_at	ANGPT1	U83508	54.9	141.2	19.7	135.2	86.1	30.3	116.7	101.4	56.7	50.8	107.7	100	198.4	27.3	55.8	228.8	9	65.2	92.2	349.1	18	94.7
204949_at	ICAM3	NM_002162	36.9	49.7	84.3	65.1	52.5	27.9	163.9	24.3	369.8	58.3	70.8	135.4	82.2	98.6	51.4	158.3	159.5	181	186.2	341.7	169.1	34.7
204000_at	GNB5	NM_016194	11.9	37.8	13.6	19.8	63.7	9.9	87	19.3	27.1	31.6	41	39.6	21	2.7	25.1	7.4	62.9	55.1	29.4	69.4	9.3	82.6
203063_at	PPM1F	NM_014634	107.6	147.5	127.5	151.8	164.8	129.1	138.7	186.2	173.1	120	190.3	225.1	233.4	86.3	56.2	167.8	117.3	97.1	78.4	239	65.7	86.2
201315_x_at	IFITM2	NM_006435	217.8	156.4	339.9	286.1	437.7	201.2	55.2	64.7	268.1	67.3	139.9	218.8	132.7	279.4	41.7	121	84.5	93.9	49.3	148.2	320.1	134.4

continued

Table IV. *continued*

Affymetrix designation	Gene symbol	Genbank accession no.	Poor prognosis																					
			ID027	ID042	ID054	ID062	ID083	ID127	ID188	ID288	ID313	ID317	ID325	ID349	ID380	ID388	ID402	ID409	ID413	ID414	ID415	ID416	ID418	ID427
37384_at	PPM1F	D13640	222.4	415.7	235.3	199.8	165.9	204	275.8	274.9	339.4	394.5	518.6	281.9	268.6	317.5	160.1	444	240.2	134.5	193	180	276.7	191.9
243579_at	MSI2	BF029215	38.5	118.9	62.3	90.9	19.5	49.2	156.3	293.6	80.7	17.3	41.5	159.3	222.8	70.7	44.5	168.9	34.9	123.8	125.1	170.6	456.3	100
228988_at	ZNF6	AU157017	148.9	141.8	199	11.7	17.6	116.9	15.8	218.3	55.8	12.5	26.6	300.3	68	25.9	3.2	267.4	5.5	139	286.9	60	18.8	41.4
228708_at		BF438386	127.2	196.3	430.7	37.8	58.7	137.7	137	118.6	66.2	57.7	44.8	82.6	28	99.3	18.5	358.5	12.5	117.1	18.4	149.4	33.4	129.7
225651_at	FLJ25157	BF431962	62.9	47.7	11.8	86	18.7	60.5	26.1	19.1	56.5	93.7	23.8	39	97.8	92.2	30.4	56.7	45.7	33.7	54.7	25.9	88.7	74.8
225351_at	HT011	AK027029	64.6	55.2	26.1	122.5	48	61.5	106.2	21	93.5	91.7	46.7	22.8	68.6	96.5	58	74.6	27	43.8	63.6	20.7	28	72.6
224516_s_at	HSPC195	BC006428	295.5	158.1	466.3	174.2	208.9	241.3	62.4	199.8	168.4	131.8	107	127.4	223	387.3	138.4	166.9	110.7	165.6	172.2	175.6	207.8	146.2
224367_at	DJ79P11.1	AF251053	188.6	343.3	173.7	8.1	93.5	345	56.7	202.8	45.3	5	24	475.1	307.3	38.6	3.5	363.7	12	259.9	273.9	136.9	109.7	68.8
219498_s_at	BCL11A	NM_018014	106.4	233.2	108.4	237.3	135.8	66.6	90.8	80.4	182.9	83.7	54.5	169	101.5	238.5	71.7	229.2	66.3	155.1	114.6	87.3	201.6	54.2
217975_at	LOC51186	NM_016303	208	191.6	111.8	10.1	64.9	157.5	6.5	88.4	303.8	5.9	36.8	350	173.9	124.6	7.7	75.5	38.8	163.6	291	102.1	227.1	160.1
215111_s_at	TSC22	AK027071	366.5	709.6	696.9	562.8	281.6	255.7	143.3	198.4	282.8	115.8	114.9	904	195.1	669.3	50.8	316.5	68.2	233.1	289	143.8	258.5	215.9
214651_s_at	HOXA9	U41813	293.6	379.5	166.6	125.8	121.9	121.7	34.3	326.8	180.5	14.7	247.4	231.1	158.6	157.3	1.6	389	4.9	159.5	215.6	108.1	65.5	324
212827_at	IGHM	X17115	179.6	337.3	272	278.3	204.9	128.4	40.8	303.7	200.1	25.7	609.8	371.1	1002.5	471.9	19.6	120.2	62.4	393	224.2	456.1	124.2	295.4
211709_s_at	SCGF	BC005810	91.2	459.8	47.1	857.1	494	313.7	1762.7	298.9	1004.1	899.4	19.2	225.2	796.7	2637.8	66.2	361.6	866.5	90.2	69	453.1	183.7	679.3
211341_at	POU4F1	L20433	2	1.8	0.8	12.3	4	7.2	1089.2	9.8	22.5	790.7	12	16.3	7.5	49.2	39.3	95.6	2.5	13.1	13.4	17.5	21.3	55.8
209905_at	HOXA9	AI246769	428	382.5	253.7	99.4	152.4	148.7	9.2	125.7	282.8	9.6	77.4	502.6	129.7	96.6	1.4	356.1	0.7	262.6	415.7	202.3	121.2	477.9
206478_at	KIAA0125	NM_014792	517.3	198.9	584.7	127.7	317.1	39.8	27.4	31.2	94.3	43.3	39.7	185.6	90.9	84.5	33.7	144.8	16.7	99.1	79.5	49.5	40.1	105.4
205609_at	ANGPT1	NM_001146	1554.3	1007.2	524.9	27.5	264.3	847.5	16.3	308.5	651.8	93.5	8.7	223.6	882.5	1067.1	306.5	362.8	123.5	261	606.1	236.6	988.8	317.3
205608_s_at	ANGPT1	U83508	158.9	253.2	218.4	84.6	131.4	28.4	38	447.3	157.9	95.5	116.6	143.1	616	304.3	107.5	328.3	52.9	141	265.2	247.5	269	415.4
204949_at	ICAM3	NM_002162	104.8	209.3	156.5	119.2	271.2	316.1	38.4	217.9	358.6	109.4	136.5	382.6	487.1	305.4	265.4	291.4	233.4	245	245.5	118.9	44.7	295.6
204000_at	GNB5	NM_016194	15.4	91.7	57.4	26.4	58.7	140.5	36.3	40.6	83.2	90.9	61.7	48.3	35.6	129.1	30	20	151.1	33.3	76.8	75.7	654.9	119.8
203063_at	PPM1F	NM_014634	187.6	318.6	216.8	132.2	166.5	126.5	189.9	278.2	263.3	128.3	439.2	240.2	163	280.5	131.1	265.3	175.6	129.1	144.6	152.2	142.9	246.5
201315_x_at	IFITM2	NM_006435	274.8	291.4	464.2	492.5	374.1	390.3	127.1	115.3	299.1	245.8	670	379.9	176.6	337	219.7	177	235.3	158.9	187.5	68.6	185.7	170.9

ANOVA ($p<0.01$) and effect size selection (≥50 U) identified 31 probe sets, expression of which differed between individuals who failed to enter initial complete remission after the standard chemotherapy (poor prognosis) and those who remained at complete remission for >1 year after the standard chemotherapy (good prognosis). The Cox proportional hazard model was applied to such probe sets to isolate 23 probe sets whose expression levels correlated ($p<0.05$) with survival time. The expression intensities of these 23 probe sets are shown.

Recent microarray analyses of BM MNCs from AML patients identified a cluster of ~100 genes whose expression patterns discriminated among AML subtypes (19) and 133 genes whose expression patterns were predictive of clinical outcome (32). Both *GPR56* and *TCRIM* were among the former group of genes.

Given that our data set was obtained with purified HSC-like fractions, it should prove informative with regard to characterization, through various approaches, of undifferentiated leukemic clones (probably including LSCs). For example, comparison between the individuals with good and poor prognosis among the 66 AML patients who underwent standard chemotherapy revealed preferential expression of *ANGPT1* in the latter group (Table IV); this gene encodes an angiogenic factor (angiopoietin 1) and is frequently overexpressed in a wide variety of human cancers (33, 34). An increased level of expression of *TEK*, which encodes a receptor for ANGPT1, was also detected in the blasts of ~10% of all 99 study subjects, some of whom overexpressed both *TEK* and *ANGPT1* (data not shown). These data suggest that an autocrine loop consisting of ANGPT1 and TEK might contribute to the malignant transformation in AML.

In contrast to the requirement for quantitation of the expression of >100 genes in the previously described approaches to prognosis prediction with BM MNCs from AML patients (19, 32), our GES system relies on determination of the expression levels of only four genes. Analysis of such a small number of genes is within the scope of an assay based on simple methodology, such as multiplex PCR. Although the GES system requires purification of $CD133^+$ cells, a combination of karyotyping and multiplex PCR is relatively straightforward even in current clinical settings. A large prospective study is now needed to verify whether individuals with AML of GES class I should be treated by standard chemotherapies, and those with AML of GES class II or III should receive more aggressive treatments such as BM transplantation.

Acknowledgements

We thank P. Simon for helpful discussion and suggestions. This work was supported, in part, by grants for Research on Human Genome and Tissue Engineering and for Third-Term Comprehensive Control Research for Cancer from the Ministry of Health, Labor, and Welfare of Japan, as well as by a grant for Scientific Research on Priority Areas "Applied Genomics" from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

1 Byrd JC, Mrozek K, Dodge RK, Carroll AJ, Edwards CG, Arthur DC, Pettenati MJ, Patil SR, Rao KW, Watson MS, Koduru PR, Moore JO, Stone RM, Mayer RJ, Feldman EJ, Davey FR, Schiffer CA, Larson RA and Bloomfield CD: Pretreatment cytogenetic abnormalities are predictive of induction success, cumulative incidence of relapse, and overall survival in adult patients with *de novo* acute myeloid leukemia: results from Cancer and Leukemia Group B (CALGB 8461). Blood *100*: 4325-4336, 2002.

2 Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DA, Gralnick HR and Sultan C: Proposed revised criteria for the classification of acute myeloid leukemia. A report of the French-American-British Cooperative Group. Ann Intern Med *103*: 620-625, 1985.

3 Goasguen JE, Matsuo T, Cox C and Bennett JM: Evaluation of the dysmyelopoiesis in 336 patients with *de novo* acute myeloid leukemia: major importance of dysgranulopoiesis for remission and survival. Leukemia *6*: 520-525, 1992.

4 Grimwade D, Walker H, Oliver F, Wheatley K, Harrison C, Harrison G, Rees J, Hann I, Stevens R, Burnett A and Goldstone A: The importance of diagnostic cytogenetics on outcome in AML: analysis of 1,612 patients entered into the MRC AML 10 trial. The Medical Research Council Adult and Children's Leukaemia Working Parties. Blood *92*: 2322-2333, 1998.

5 Jaffe ES, Harris NL, Stein H and Vardiman JW: Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues. Lyon, IARC Press, 2001.

6 Liu ET and Karuturi KR: Microarrays and clinical investigations. N Engl J Med *350*: 1595-1597, 2004.

7 Miyazato A, Ueno S, Ohmine K, Ueda M, Yoshida K, Yamashita Y, Kaneko T, Mori M, Kirito K, Toshima M, Nakamura Y, Saito K, Kano Y, Furusawa S, Ozawa K and Mano H: Identification of myelodysplastic syndrome-specific genes by DNA microarray analysis with purified hematopoietic stem cell fraction. Blood *98*: 422-427, 2001.

8 Ohmine K, Ota J, Ueda M, Ueno S-i, Yoshida K, Yamashita Y, Kirito K, Imagawa S, Nakamura Y, Saito K, Akutsu M, Mitani K, Kano Y, Komatsu N, Ozawa K and Mano H: Characterization of stage progression in chronic myeloid leukemia by DNA microarray with purified hematopoietic stem cells. Oncogene *20*: 8249-8257, 2001.

9 Hin AH, Miraglia S, Zanjani ED, Almeida-Porada G, Ogawa M, Leary AG, Olweus J, Kearney J and Buck DW: AC133, a novel marker for human hematopoietic stem and progenitor cells. Blood *90*: 5002-5012, 1997.

10 Peichev M, Naiyer AJ, Pereira D, Zhu Z, Lane WJ, Williams M, Oz MC, Hicklin DJ, Witte L, Moore MA and Rafii S: Expression of VEGFR-2 and AC133 by circulating human CD34(+) cells identifies a population of functional endothelial precursors. Blood *95*: 952-958, 2000.

11 Bonnet D and Dick JE: Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat Med *3*: 730-737, 1997.

12 Hope KJ, Jin L and Dick JE: Acute myeloid leukemia originates from a hierarchy of leukemic stem cell classes that differ in self-renewal capacity. Nat Immunol *5*: 738-743, 2004.

13 Singh SK, Hawkins C, Clarke ID, Squire JA, Bayani J, Hide T, Henkelman RM, Cusimano MD and Dirks PB: Identification of human brain tumour initiating cells. Nature *432*: 396-401, 2004.

14 Van Gelder RN, von Zastrow ME, Yool A, Dement WC, Barchas JD and Eberwine JH: Amplified RNA synthesized from limited quantities of heterogeneous cDNA. Proc Natl Acad Sci USA *87*: 1663-1667, 1990.

15 Oshima Y, Ueda M, Yamashita Y, Choi YL, Ota J, Ueno S, Ohki R, Koinuma K, Wada T, Ozawa K, Fujimura A and Mano H: DNA microarray analysis of hematopoietic stem cell-like fractions from individuals with the M2 subtype of acute myeloid leukemia. Leukemia *17*: 1990-1997, 2003.
16 Fellenberg K, Hauser NC, Brors B, Neutzner A, Hoheisel JD and Vingron M: Correspondence analysis applied to microarray data. Proc Natl Acad Sci USA *98*: 10781-10786, 2001.
17 Alon U, Barkai N, Notterman DA, Gish K, Ybarra S, Mack D and Levine AJ: Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. Proc Natl Acad Sci USA *96*: 6745-6750, 1999.
18 Ichikawa N, Kitano K, Ito T, Nakazawa T, Shimodaira S, Ishida F and Kiyosawa K: Abnormal proliferation of CD4-CD8+ gammadelta+ T cells with chromosome 6 anomaly: role of Fas ligand expression in spontaneous regression of the cells. Am J Hematol *60*: 305-308, 1999.
19 Valk PJ, Verhaak RG, Beijen MA, Erpelinck CA, Barjesteh van Waalwijk van Doorn-Khosrovani S, Boer JM, Beverloo HB, Moorhouse MJ, van der Spek PJ, Lowenberg B and Delwel R: Prognostically useful gene-expression profiles in acute myeloid leukemia. N Engl J Med *350*: 1617-1628, 2004.
20 Wheatley K, Burnett AK, Goldstone AH, Gray RG, Hann IM, Harrison CJ, Rees JK, Stevens RF and Walker H: A simple, robust, validated and highly predictive index for the determination of risk-directed therapy in acute myeloid leukaemia derived from the MRC AML 10 trial. United Kingdom Medical Research Council's Adult and Childhood Leukaemia Working Parties. Br J Haematol *107*: 69-79, 1999.
21 Cox DR: Regression models and life tables. J R Stat Soc *34*: 187-220, 1972.
22 Beer DG, Kardia SL, Huang CC, Giordano TJ, Levin AM, Misek DE, Lin L, Chen G, Gharib TG, Thomas DG, Lizyness ML, Kuick R, Hayasaka S, Taylor JM, Iannettoni MD, Orringer MB and Hanash S: Gene-expression profiles predict survival of patients with lung adenocarcinoma. Nat Med *8*: 816-824, 2002.
23 Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J Jr, Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Byrd JC, Botstein D, Brown PO and Staudt LM: Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature *403*: 503-511, 2000.
24 Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, Fisher RI, Gascoyne RD, Muller-Hermelink HK, Smeland EB, Giltnane JM, Hurt EM, Zhao H, Averett L, Yang L, Wilson WH, Jaffe ES, Simon R, Klausner RD, Powell J, Duffey PL, Longo DL, Greiner TC, Weisenburger DD, Sanger WG, Dave BJ, Lynch JC, Vose J, Armitage JO, Montserrat E, Lopez-Guillermo A, Grogan TM, Miller TP, LeBlanc M, Ott G, Kvaloy S, Delabie J, Holte H, Krajci P, Stokke T and Staudt LM: The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. N Engl J Med *346*: 1937-1947, 2002.
25 Shipp MA, Ross KN, Tamayo P, Weng AP, Kutok JL, Aguiar RC, Gaasenbeek M, Angelo M, Reich M, Pinkus GS, Ray TS, Koval MA, Last KW, Norton A, Lister TA, Mesirov J, Neuberg DS, Lander ES, Aster JC and Golub TR: Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. Nat Med *8*: 68-74, 2002.
26 Ferrando AA, Neuberg DS, Staunton J, Loh ML, Huard C, Raimondi SC, Behm FG, Pui CH, Downing JR, Gilliland DG, Lander ES, Golub TR and Look AT: Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. Cancer Cell *1*: 75-87, 2002.
27 Yeoh EJ, Ross ME, Shurtleff SA, Williams WK, Patel D, Mahfouz R, Behm FG, Raimondi SC, Relling MV, Patel A, Cheng C, Campana D, Wilkins D, Zhou X, Li J, Liu H, Pui CH, Evans WE, Naeve C, Wong L and Downing JR: Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. Cancer Cell *1*: 133-143, 2002.
28 Kapp U, Yeh WC, Patterson B, Elia AJ, Kagi D, Ho A, Hessel A, Tipsword M, Williams A, Mirtsos C, Itie A, Moyle M and Mak TW: Interleukin 13 is secreted by and stimulates the growth of Hodgkin and Reed-Sternberg cells. J Exp Med *189*: 1939-1946, 1999.
29 Skinnider BF, Elia AJ, Gascoyne RD, Trumper LH, von Bonin F, Kapp U, Patterson B, Snow BE and Mak TW: Interleukin 13 and interleukin 13 receptor are frequently expressed by Hodgkin and Reed-Sternberg cells of Hodgkin lymphoma. Blood *97*: 250-255, 2001.
30 Cavdar Koc E, Burkhart W, Blackburn K, Moseley A and Spremulli LL: The small subunit of the mammalian mitochondrial ribosome. Identification of the full complement of ribosomal proteins present. J Biol Chem *276*: 19363-19374, 2001.
31 Bruyns E, Marie-Cardine A, Kirchgessner H, Sagolla K, Shevchenko A, Mann M, Autschbach F, Bensussan A, Meuer S and Schraven B: T cell receptor (TCR) interacting molecule (TRIM), a novel disulfide-linked dimer associated with the TCR-CD3-zeta complex, recruits intracellular signaling proteins to the plasma membrane. J Exp Med *188*: 561-575, 1998.
32 Bullinger L, Dohner K, Bair E, Frohling S, Schlenk RF, Tibshirani R, Dohner H and Pollack JR: Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. N Engl J Med *350*: 1605-1616, 2004.
33 Shim WS, Teh M, Bapna A, Kim I, Koh GY, Mack PO and Ge R: Angiopoietin 1 promotes tumor angiogenesis and tumor vessel plasticity of human cervical cancer in mice. Exp Cell Res *279*: 299-309, 2002.
34 Giuliani N, Colla S, Lazzaretti M, Sala R, Roti G, Mancini C, Bonomini S, Lunghi P, Hojden M, Genestreti G, Svaldi M, Coser P, Fattori PP, Sammarelli G, Gazzola GC, Bataille R, Almici C, Caramatti C, Mangoni L and Rizzoli V: Proangiogenic properties of human myeloma cells: production of angiopoietin-1 and its potential relationship to myeloma-induced angiogenesis. Blood *102*: 638-645, 2003.

Received April 1, 2006
Accepted May 15, 2006