Acute Myeloid Leukemia • Research Paper

Combined genetic and transcriptional profiling of acute myeloid leukemia with normal and complex karyotypes

[haematologica]
2004;89:1072-1081

CHARLOTTA LINDVALL
KYLE FURGE
MAGNUS BJÖRKHOLM
XIANG GUO
BRIAN HAAB
ELISABETH BLENNOW
MAGNUS NORDENSKJÖLD
BIN TEAN TEH

A B S T R A C T

Background and Objectives. Acute myeloid leukemia (AML) is a heterogeneous group of diseases. Patients with a normal karyotype constitute the largest single group; multiple chromosome rearrangements involving three or more chromosomes occur in 5–10% of AML patients. The pathophysiologic mechanisms underlying both groups are largely unknown. In the current study, we have systematically combined transcriptional profiles with cytogenetic data from 15 AML patients with either normal or complex karyotypes.

Design and Methods. The expression profiles were investigated by unsupervised hierarchical clustering, supervised cluster analysis, and comparative genomic microarray analysis. In addition, the samples were analyzed by G-banding and/or spectral karyotyping and comparative genomic hybridization.

Results. Our results show that AML with complex karyotypes exhibit a gene expression profile that is specific to this group of patients. The differentially expressed genes included several located on 5q and 7q, as well as genes involved in controlling cell division. We also found that DNA gains and losses caused by multiple chromosome rearrangements result in altered gene expression in a gene-dosage-dependent manner.

Interpretation and Conclusions. These data provide insight into the mechanisms of multiple chromosome rearrangements and further demonstrate that the expression patterns of AML are strongly linked to the karyotypic status, even for the relatively undefined cytogenetic subgroup *AML with complex karyotype*.

Key words: AML, cytogenetics, gene expression profiling.

From the Department of Molecular Medicine (CL,EB MN) and Medicine (MB), Karolinska Hospital and Institutet, Stockholm, Sweden Bioinformatics Program (KF), Laboratory of DNA and Protein Microarray Technology (BH) and Laboratory of Cancer Genetics (BTT), Van Andel Research Institute, Grand Rapids, MI, USA.

Correspondence:
Charlotta Lindvall, MD, PhD, Laboratory of Cell Signaling and Carcinogenesis, Van Andel Research Institute, 333 Bostwick Ave NE, Grand Rapids, MI 49503, USA
Email: charlotta.lindvall@vai.org

Acute myeloid leukemia (AML) results from a clonal expansion of a malignant transformed progenitor cell in the bone marrow, blood, or other tissue. Acquired cytogenetic aberrations are detected in 55–75% of newly diagnosed patients with AML; the rest show no cytogenetic changes, and this masks any clues to their molecular pathogenesis.[1,2] Individuals with normal cytogenetics (CN) constitute the largest single group of AML patients. Many karyotypic abnormalities are associated with specific disease subtypes, characteristic morphologic and immunologic profiles, and distinct therapeutic and prognostic implications.[2,3] Direct involvement of many recurring translocations, inversions, and deletions in the leukemogenetic process is supported by molecular dissection and cloning of genes adjacent to translocation breakpoints. However, approximately 10% of AML with cytogenetic aberrations do not have leukemia-specific aberrations at diagnosis, but do have complex karyotypes with multiple chromosome rearrangements (MCR) involving three or more chromosomes.[2] AML patients with MCR respond poorly to antileukemic treatment, and it is likely that some of these rearrangements contribute to drug resistance and disease progression.[2,4] We and other groups have shown that the MCR seen in AML often result in loss of chromosome arms 5q, 7q, and 17p and gain of chromosome 8 and arm 11q.[5-7]

In the current study, we have combined genetic and transcriptional profiling systematically in order to characterize AML with normal and complex karyotypes. The pathophysiologic mechanisms underlying both groups are largely unknown. However, it is believed that normal hematopoiet-

ic differentiation, which is at least partly regulated at the expression level, is blocked. Therefore, gene expression profiling using cDNA microarrays should be a powerful tool for identifying these transcriptional changes. In fact, gene expression profiling has already successfully been used to identify gene expression patterns specific to AML and certain AML cytogenetic subgroups, but this paper is, to our knowledge, the first cDNA microarray report on AML with complex karyotypes.[8-11] We obtained the expression profiles of 15 AML samples — 9 with normal karyotypes and 6 with complex karyotypes. We also obtained expression profiles of seven mononuclear cell samples from healthy bone marrow donors as controls. We studied the data to: (i) investigate whether AML with complex karyotypes exhibits a specific gene expression pattern; (ii) identify genes that are differentially expressed in AML with complex karyotypes relative to AML with normal cytogenetics; and (iii) test whether MCR cause a change in gene expression in a gene-dosage-dependent manner. We also present a relatively new technique for analyzing microarray data, termed comparative genomic microarray analysis (CGMA). CGMA identifies chromosome regions that exhibit a disproportionate number of genes with induced or reduced gene expression.

Design and Methods

Bone marrow sample collection and processing

Mononuclear cells were isolated from diagnostic AML samples (bone marrow or peripheral blood) by using a Ficoll-Hypaque gradient and were then stored at –70°C. The diagnosis and subtyping were performed according to the FAB classification.[12] Mononuclear cells (MNC) isolated from seven healthy bone marrow donors were purchased from Poietics (BioWhittaker Inc., Walkersville, MD, USA). For the microarray experiment, total RNA was extracted by using Trizol (Invitrogen, Carlsbad, CA, USA). RNA isolated from pooled white blood cells obtained from six healthy blood donors was used as the common reference in all microarray experiments. This study was performed with ethical approval from the Karolinska hospital's ethics committee. A summary of the patients' characteristics is given in Table 1.

cDNA microarray experiments

DNA microarray production, labeling, and hybridization were performed as described by Takahashi *et al.*[13] cDNA transcribed from the common reference RNA was labeled with Cy3 (green channel) and cDNA transcribed from each individual AML or MNC RNA sample was labeled with Cy5 (red channel). The relative expression of each gene could be determined as the ratio of the red/green signal. For example, if one gene was highly expressed in the AML-CN group but not expressed in the AML-MCR group, the array spot representing this gene lit up red on the AML-CN arrays and green on the AML-MCR arrays and would be included in the analyses. Those spots that were not significantly stronger than background in either the red or the green channel were excluded. Two samples, AML-MCR 1 and AML-MCR 5, were arrayed in duplicate. For the first microarray experiment, the two AML samples were labeled with Cy5 and the common reference with Cy3, and for the second microarray experiment the two AML samples were labeled with Cy3 and the reference with Cy5. This was done to determine the reproducibility of the arrays and to control for possible differential labeling efficiencies associated with the specific dyes used. In addition, to determine the amount of noise in our microarray data, one aliquot of cDNA transcribed from the common reference RNA was labled with Cy3 and one with Cy5, and then arrayed to each other.

Slides were scanned in a commercially available confocal fluorescent Scan Array Lite scanner equipped with lasers operating at 532 (Cy3; green channel) and 635 nm (Cy5; red channel) (GSI Lumonics, Billerica, CA, USA). Image files were analyzed by GenePix Pro 3 image analysis software (Axon Instruments, Union City, USA).

Unsupervised and supervised cluster analyses

Gene ratios were log-transformed and normalized prior to clustering (this file is available as Supplementary Data 1). Normalization was performed as described by Yang *et al.*[14] Unsupervised hierarchical cluster analysis was performed using the program CLUSTER and visualized using the program TREEVIEW (*http://rana.lbl.gov*). Only genes present in 80% of the samples, in total 2,529, were included in the analysis. In CLUSTER, genes were first median-centered, and then genes and subsequently arrays were normalized as recommended by the software manual. Correlation-based, uncentered, average linkage clustering was then applied to the genes and arrays.

To identify genes with a significantly different expression in the experimental groups, we analyzed the log-transformed and normalized data set (Supplementary Data 1) using the computer software CIT.[15,16] CIT uses a statistical discrimination metric and permutation analysis to identify clusters of genes or individual genes that differentiate between experimental groups. CIT is freely available on the Van Andel Research Institute bioinformatics core program's webpage (*http://www.vai.org/*).

Table 1. Patients' characteristics.

Patient	*AML subgroup*[1]	*Age*[2] */sex*	*Survival time*[3]	*Karyotype*[4]	*DNA gain or loss*[5]	*Comparative Genomic Microarray Analysis*[6]
AML-CN 1	M2	78/F	42	46,XX[27]	Balanced	-19q
AML-CN 1R	M2			46,XX[25]	Balanced	-17p,-19p,-22q
AML-CN 2	M2	80/F	3	46,XX[26]	Balanced	-1p,-1q,-15q,-22q
AML-CN 2R	M2			46,XX[27]	Balanced	Balanced
AML-CN 3	M1/M2	54/F	7	46,XX[27]	Balanced	-10q,-19q
AML-CN 3R	M1/M2			46,XX[28]	Balanced	-1p,-6q
AML-CN 4	M2	77/F	11	46,XX[29]	Balanced	-22q
AML-CN 4R	M2			46,XX[24]	Balanced	-1p,-1q,-8q,-11q,-17p, -19p,-19q -22q
AML-CN 5	M4	50/M	44+	46,XY[29]	Balanced	Balanced
AML-CN 6	M5B	43/M	0	46,XY[27]	Balanced	Balanced
AML-CN 7	M4	77/M	38+	46,XY[27]	Balanced	-19p
AML-CN 8	M2	61/M	1	46,XY[27]	Balanced	-19p
AML-CN 9	M5	27/M	34+	46,XY[27]	Balanced	Balanced
AML-MCRs 1	M0	80/M	4	42~43,XY,-3,der(5)t(5;19) (q11.2;?),del(6)(p22), der(9)t(6;9)(p22;q34),+11, der(11;15) (q10;q10),der(12)t(12;19)(p11;?),-17, -19,der(20)t(11;20)(q23;q13.3)[cp8]	Loss: 3,5q11.2-qter, 12p11-pter,17,19q10-qter Gain: 11q23-qter	-3p,-3q,5q,-17p, -17q,-19p,-19q
AML-MCRs 2	M1	74/F	3	48,X,t(X;1)(p11;p11),t(1;22)(q31;q12), der(3)t(3;11)(q12;q23),del(5)(q21q31), der(6)(17pter→17p?::15q?→15q?: :14q?→14q?::6p23→6qter),der(7)t(6;7)(p23;q22), +der(8)t(3;8)(q21;q24),-14,+15, ider(15)(q10)t(11;15)(q13;q?)x2 der(17)t(15;17)(q21;p11.2),+22[14]	Loss: 5q21-q31,7q22-qter Gain: 8pter-q24,11q21-qter	-5q,-7q
AML-MCRs 3	M2	77/F	6	46,XX,del(5)(q15q31),+11, +13,i(13)(q10),-16, der(17)t(16;17)(p11;p11),-18[10]	Loss: 5q15-q31,16p11-qter, 17p11-pter,18 Gain: 11,13	-5q,-16q, -17p,-18q +11p,+11q,+13q
AML-MCRs 4	M4	70/F	1	47~50,XX,+r(3)(p11q?)x2, +r(3)(p11q?)x2,+r(3)(p11q?), der(5)t(5;14)(q22;q22),-7, der(14)del(14)(q13q22)ins(14;7)(q12;?) [cp8]/46,XX[2]	Loss: 5q22-qter,7, 14q12-q22 Gain: 3q12-qter	-5q,-7p,-7q,-14q +3q
AML-MCRs 5	M5A	58/M	1	45,XY,ins(4;11)(q12;q?q?),del(5) (q13q33),der(12)t(11;12)(q23;p11.2) hsr(11)(q23),-16[9]	Loss: 5q13-q33, 12p11-pter,16 Gain:11q13-qter	-5q,-16p,-16q,-19p
AML-MCRs 6	M1	64/M	+8	46,XY,del(5)(q21q31), del(7)(q21),del(20)(q11.2q12)	Loss: 5q21-q31, 7q21-qter, 20q11.2-q12	-4p,-5q,-7q,-19p,-20q

[1]According to the FAB classification; [2]Patient's age at the time of AML diagnosis (years); [3]Survival time (months); [4]Determined by G-banding for AML-CN 1–9 and by spectral karyotyping for AML-MCRs 1–6: [5]Determined by karyotyping for AML-CN 1–9 and by comparative genomic hybridization for AML-MCRs 1–6; [6] p = 0.01 and n = 15 (see "Design and Methods").

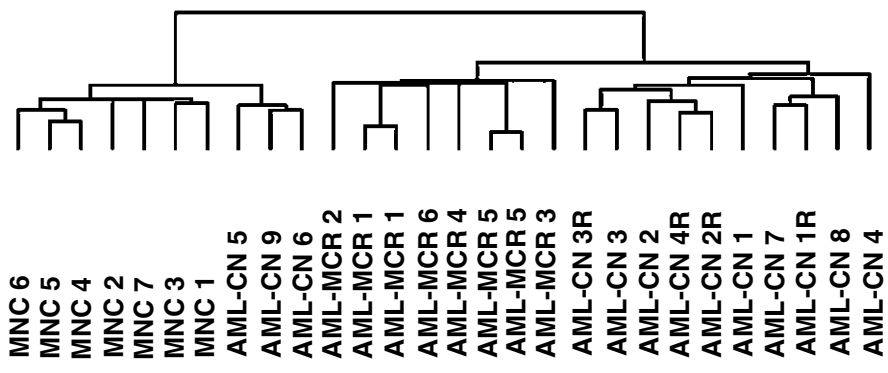

Figure 1. Unsupervised cluster analysis of 15 AML samples (6 with multiple chromosome rearrangements and 9 with normal cytogenetics) and 7 MNC samples from healthy bone marrow donors. The cluster dendrogram represents similarities in the expression of 2,529 cDNAs, which were selected because these genes were expressed in more than 80% of the samples.

Data transformation

For generation of transcriptional profiles using MNC from healthy bone marrow donors (MNC-N) as the common reference, the normalized log-transformed expression ratios of MNC-N/buffy coat (MNC-N/BC) were averaged and inverted (BC/MNC-N) by multiplication by –1. The log-transformed, normalized transcription profile of each AML (AML/BC) was multiplied by the transcription profile of the inverted MNC-N/BC, i.e., (BC/MNC-N) × (AML/BC) = AML/MNC-N.

Identification of chromosomal gains and losses by transcriptional profiling using comparative genomic microarray analysis (CGMA)

To identify regional gene expression biases, gene expression values were segregated into sets based on chromosomal arm mapping. A one sample Welsh's t-test was applied to the gene expression values in each set to determine whether a significant number of genes were up- or down-regulated ($p \leq 0.005$). The resulting t-statistics were plotted as a heatmap (Figure 3). Pink indicates genomic regions that show a significant number of up-regulated genes and blue indicates genomic regions containing a significant number of down-regulated genes. All data were analyzed using the base R environment[17] supplemented with the BioConductor packages (*http://www.bioconductor.org*). The CGMA computer program is freely available at on the Van Andel Research Institute bioinformatics core program's webpage (*http://www.vai.org*). The algorithm and applications of CGMA are also described in the listed references.[18-20]

Spectral karyotyping and comparative genomic hybridization

Spectral karyotyping (SKY) was performed as described by Lindvall *et al.*[5] For comparative genomic hybridization (CGH), DNA was isolated simultaneously with RNA by using Trizol (Invitrogen, Carlsbad, CA, USA). The CGH labeling and hybridization were performed as described by Kallioniemi *et al.*[21]

Results

We analyzed diagnostic leukemic samples from 15 patients with AML (Table 1) – 9 of whom had a normal karyotype (AML-CN), whereas 6 had a complex karyotype (AML-MCR). For four AML patients (AML-CN 1-4) we also obtained leukemic samples at the time of relapse. We analyzed the samples by karyotype (G-banding and/or spectral karyotyping), comparative genomic hybridization, and gene expression profiling. We also analyzed gene expression data of seven normal MNC samples obtained from healthy bone marrow donors as controls. The cDNA microarray we used contained 19,956 cDNA clones, representing approximately 16,000 genes or expressed sequence tags (EST); thus, many genes were represented by more than one cDNA on microarray. In order to determine the amount of noise in our microarray data, RNA from the same sample was reciprocally labeled and hybridized to the microarray. On this array, 99.98% of features did not show greater than 2-fold expression changes.

Gene expression profiling

As an introductory approach to analyzing the gene expression data, we used unsupervised hierarchical clustering analysis to group patients and genes on the basis of similarity of expression pattern. In this two-dimensional clustering analysis, patients with complex karyotypes formed a cluster separate from the normal MNC sample group and also separate from patients with normal cytogenetics, suggesting the existence of gene expression pattern specific to AML-MCR (Figure 1). One chromosomal abnormality, a deletion of the long arm of chromosome 5, was present in all AML-MCR samples (Table 1). To test whether this or any other chromosomal changes were the primary influence of

Table 2. Genes with lower expression in AML with complex karyotype compared to in AML with normal karyotype.

cDNA	*Gene Symbol*[1]	*Change*[2]	*Location*	*Gene Function*[1]
AA485373	VMP1	-3.9	17q23.2	
W49708	SRA1	-2.4	5q31.3	receptor activity
AA521411	CAMLG	-2.0	5q23	signal transduction
R76437	TBXAS1	-3.1	7q34-q35	blood coagulation, electron transport, oxidoreductase activity, prostaglandin biosynthesis
AA431196	SPEC1	-1.9	1q21.3	signal transduction
AA487914	HSD17B4	-1.8	5q21	metabolism, oxidoreductase activity, steroid biosynthesis
AA608575	PCCA	-1.9	13q32	metabolism
AA485427	CRIP2	-5.8	14q32.3	zinc ion binding
W47106	NID67	-2.0	5q33.1	
T52325	TES	-2.2	7q31.2	
W48726	PBX3	-4.6	9q33-q34	regulation of transcription, transcription factor activity
AA133191	QP-C	-2.7	5q31.1	electron transport, mitochondrial electron transport chain, oxidoreductase activity
W88497	LOC51064	-2.0	7q35	glutathione transferase activity, protein disulfide oxidoreductase activity
AA425782	KIAA0874	-1.7	18p11.22	
AA454740	PHF15	-2.0	5q31.2	regulation of transcription
R77144	TMPIT	-2.2	7q11.23	
AA406311	EVI2A	-1.9	17q11.2	cell growth and/or maintenance, transmembrane receptor activity
W46985	MMP24	-1.6	20q11.2	proteolysis and peptidolysis, enzyme activator activity
N80129	MT1X	-3.0	16q13	metal ion binding
AA789301	MRF-1	-2.1	2q11.2	DNA binding
T50041	DKFZP586A0522	-2.1	12q13.13	
N94357	SSH2	-2.3	17q11.2	protein amino acid dephosphorylation, protein phosphatase activity
R62412	C6orf89	-1.8	6p21.31	
H48472	SPEC2	-1.9	5q31.1	signal transduction, kinase activity
T50313	MAP4K1	-1.5	19q13.1 -q13.4	activation of JUNK, protein kinase activity
AA291773	TETRAN	-2.0	4p16.3	integral to membrane
AA452541	LYSAL1	-1.8	8p21.2	nucleoside diphosphatase activity, magnesium ion binding
N35086	FYN	-2.4	6q21	cell growth and/or maintenance, intracellular signaling cascade, protein kinase activity
AA456105	FLJ34969	-1.6	3p21.2	
AA053810	SMARCD3	-3.0	7q35-q36	chromatin modeling, regulation of transcription from Pol II promoter

[1]*According to NCBI (http://cgap.nci.nih.gov);* [2]*Log2 scale.*

the pattern of clustering, sets of genes that mapped to each chromosome were removed and the samples were re-clustered. In all cases, the pattern of clustering remained unchanged (*data not shown*). This suggests that features other than chromosome dosage determined the partitioning of the AML-MCR samples. Two of the AML-MCR samples (AML-MCR 1 and 5) were arrayed twice. The replicates clustered together, demonstrating high reproducibility of the microarray method.

To determine the homogeneity in gene expression within and between the AML-MCR, AML-CN, and MNC sample groups we computed the correlation coefficients (no correlation = 0, perfect correlation = 1). For the AML-MCR samples the correlations ranged from 0.5 to 0.9 with the average correlation being 0.63. Interestingly, the average correlation between this group of samples was higher than that for both the AML-CN 0.2-0.83 (0.55) and MNC 0.2-0.75 (0.5) sample groups. This suggests that while the AML-MCR samples contain heterogenous karyotypes, there is a great deal of similarity between the individual samples' gene expression profiles.

Differentially expressed genes in AML with complex karyotypes

A Student's t test was performed to identify genes associated with the AML-MCR cluster. *p* values were adjusted by permutation analysis.[16] In total, 169 genes were differentially expressed in AML-MCR relative to AML-CN ($p < 0.05$). Tables 2 and 3 summarize the 30 most significantly down- or up-regulated genes, respectively. Interestingly, 12 out of 30 (40%) genes

Table 3. Genes with higher expression in AML with complex karyotype compared to in AML with normal karyotype.

cDNA	*Gene Symbol*[1]	*Change*[2]	*Location*	*Gene Function*[1]
AA018659	KIAA0193	+6.0	7p14.3-p14.1	
N66064	CFP1	+1.7	10p11.22	cell cycle, cytokinesis
AA085978	ANAPC1	+1.8	2q12.1	
AA598621	SRPR	+2.7	11q24.3	cotranslational membrane targeting, signal recognition particle
AA007299	TRIO	+2.5	5p15.1-p14	protein kinase activity, receptor protein tyrosine phosphatase signaling pathway
AA453175	BIN1	+6.0	2q14	cell differentiation, cell proliferation, negative regulation of cell cycle, actin cytoskeleton
AA486761	KIAA1522	+2.0	1p34.3	
AA441935	ASCL1	+3.5	12q22-q23	cell differentiation, transcription factor activity
H89517	APLP2	+4.3	11q23-q25	G-protein coupled receptor protein signaling pathway
AA487921	KIAA0152	+1.6	12q24.31	integral to membrane
AA406601	ABLIM1	+5.8	10q25	cytoskeleton organization and biogenesis, actin cytoskeleton
W42849	APP	+9.7	21q21.2	
W40150	CSPG6	+2.9	10q25	DNA repair, cell cycle, chromosome segregation
AA701030	RTBDN	+2.6	19p12	
H72683	CFP1	+2.1	10p11.22	cell cycle, cytokinesis
W45165	CAPZB	+2.1	1p36.1	actin cytoskeleton organization and biogenesis
AA488413	MKNK2	+1.9	19p13.3	protein kinase activity
AA010247	FLJ10890	+2.2	11p11.2	
R39578	CTL2	+2.2	19p13.1	integral to membrane
H69583	BTG2	+6.5	1q32	DNA repair, negative regulation of cell proliferation, transcription factor activity
AA425401	STK24	+2.1	13q31.2-q32.3	protein kinase activity, signal transduction
AA701046	MLL	+2.5	11q23	cell growth and/or maintenance, transcription factor activity
AA598583	CYFIP1	+2.0	15q11	
AA644128	NASP	+1.5	1p34.1	DNA packaging
AA018591	SPTBN1	+2.7	2p21	cytoskeleton, actin binding
AA455056	MAPKAPK2	+1.7	1q32	MAPKKK cascade, protein kinase activity, signal transducer activity
W67174	ITGB1	+2.0	10p11.2	cell-matrix adhesion, integrin-mediated signaling pathway, receptor activity
AA775257	ITM2A	+4.0	Xq13.3-Xq21.2	integral to membrane
R44546	KIAA0153	+1.9	22q13.31	protein modification, tubulin-tyrosine ligase activity
AA437370	SLC35B2	+2.0	6p12.1-p11.2	electron transporter activity, copper ion binding

[1]According to NCBI (http://cgap.nci.nih.gov); [2]Log2 scale.

significantly down-regulated in AML-MCR were located on either 5q or 7q (Table 2). Moreover, the gene *PBX3* was highly expressed in AML-CN (i.e., down-regulated in AML-MCR). *PBX3*, a homeobox gene, has previously been shown to be induced in AML with normal cytogenetics.[11] The up-regulated genes included two genes known to be involved in DNA repair (*CASPG6, BTG2*) and two in chromosome segregation (*CFP1, CSPG6*). CFP1 was represented by two different cDNA on the microarray; both cDNA were identified as significant features (Table 3). The *MLL* gene, located on 11q23 and often amplified in AML with 11q23 gain,[22] was also more highly expressed in the AML-MCR group than in the AML-CN group. Interestingly, four of the 30 (13%) up-regulated genes are involved with the actin cytoskeleton. Certain actin-binding proteins have been shown to be critical for reliable chromosome segregation in mitosis.[23]

Combining transcriptional profiles with SKY and CGH

To test whether a gene-dose effect plays a role in the origin of AML-MCR, we identified all cDNA corresponding to mapped Unigene clusters and averaged the expression ratios of cDNA from the same Unigene cluster. This resulted in almost 10.000 mapped Unigene clusters covering the genome. SKY or G-banding was performed on all samples, and CGH was performed on all AML-MCR samples (Table 1). In the group of patients with complex karyotypes, the most common chromosome deletion was 5q, followed by 7q, whereas the most common chromosome gain was 11q. This is consistent with recently published cytogenetic reports of AML with MCR.[5-7]

We first plotted the AML-MCR vs. MNC-N expression profiles of the Unigene clusters mapped to the chromosomal regions showing DNA losses or gains by SKY and CGH (Figure 2). In these regions, 47–88% of genes in the deleted regions and 17–75% of genes in amplified regions exhibited a change in gene expression consistent with the chromosome aberrations. To make the predictions of whether regions of chromosome loss or gain are associated with a corresponding change in gene expression more objective, we used a computational approach to identify regional expression biases termed CGMA.[18,19] CGMA arranges gene expression data based on genomic mapping information and determines whether a genomic region (in this analysis, *genomic region* is defined by chromosomal arm boundaries) contains a significantly disproportionate number of genes with increased or reduced expression (see *Design and Methods*). We found that most of the chromosome aberrations identified by CGH were associated with aberrant gene expression profiles as determined by CGMA (Table 1 and Figure 3).

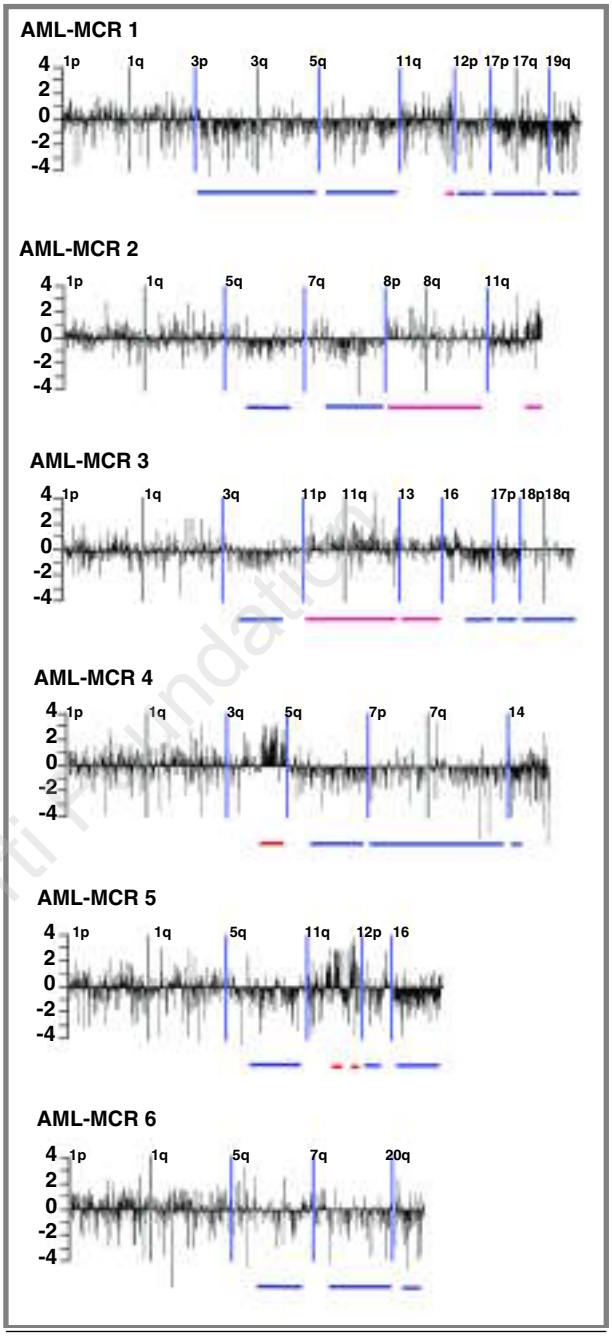

Figure 2. Illustration of the gene expression in regions with DNA gains and losses as determined by comparative genomic hybridization (CGH). The black bars show log-transformed AML-MCRs vs. MNC-N expression ratios of individual genes, the blue bars show chromosome boundaries, and the gray shows centromere location if both chromosome arms are represented. The horizontal blue and pink bars indicate DNA loss or gain, respectively. For each patient, the gene expression of chromosome 1 is shown as a reference, because none of the patients exhibited any aberrations on this chromosome by spectral karyotyping or CGH.

In total, 18 out of 24 (75%) DNA deletions and 3 out of 7 (43%) DNA amplifications were detectable by the

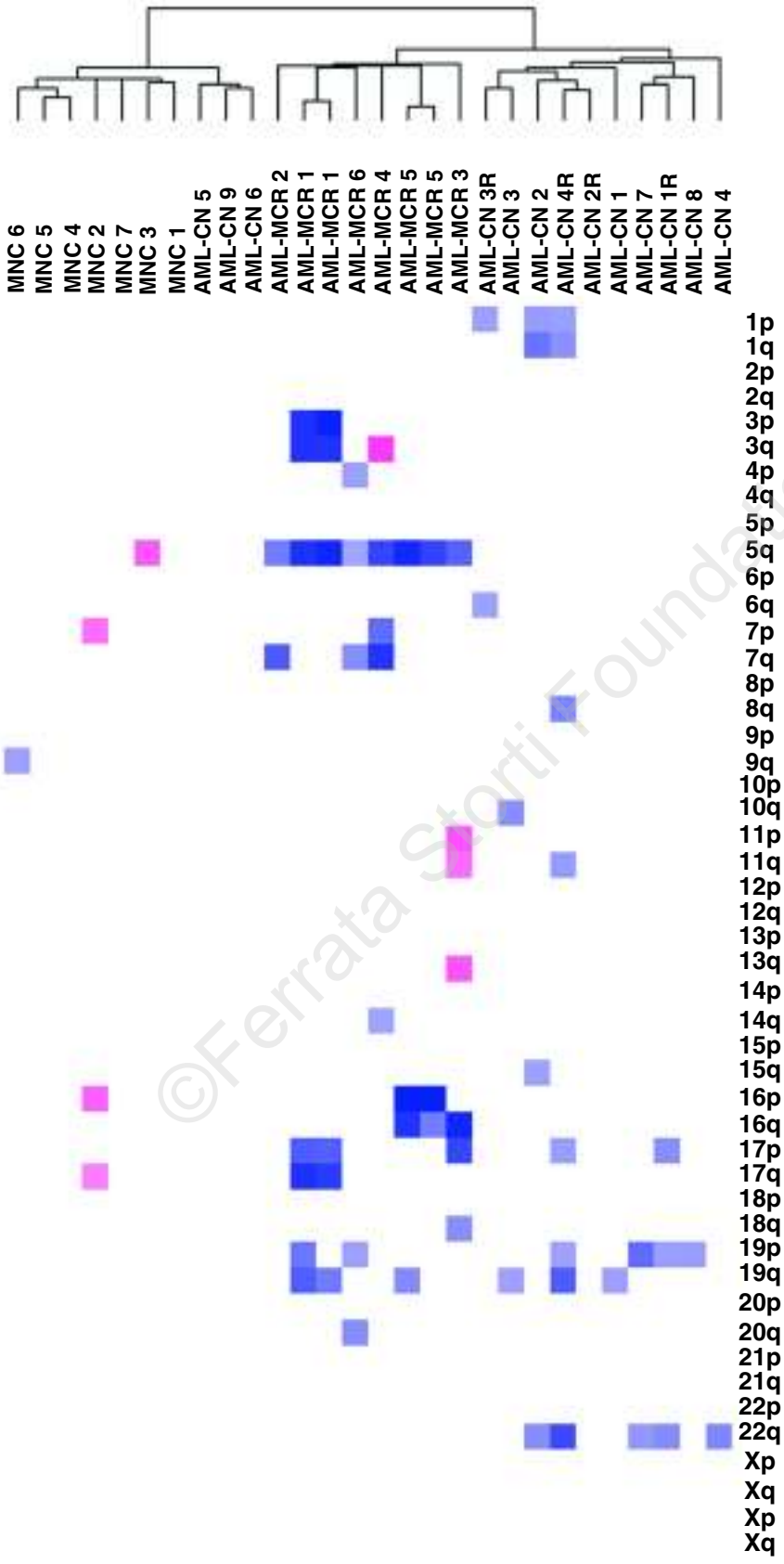

Figure 3. Illustration of comparative genomic microarray analysis (CGMA). Each column represents a sample, and each row represents a chromosome arm. Samples were sorted according to the clustering result (Figure 1). Genomic regions that showed a significant number of down-regulated and up-regulated genes are shown in blue and pink, respectively. The color intensity indicates the significance of the expression bias. The lowest-intensity color indicates a ζ-statistic = 3.48 ($p < 0.001$), and the most intense color indicate a ζ-statistic = 8.59.

use of CGMA. Altogether, CGMA identified 27 expression biases, 25 (93%) of these corresponding to regions with DNA gain or losses. The expression biases were always in the same direction as the DNA imbalance. We then performed CGMA on the AML-CN samples and found that 9 out of 13 samples exhibited one or more chromosome regions with low-expression biases (Table 1 and Figure 3). The most common low-expression bias in this group of patients involved chromosome 19, followed by 22q, present in 6 (46%) and 4 (31%) samples, respectively. No high-expression bias was found in the AML-CN group.

Discussion

We systematically combined genetic and transcriptional profiles in order to characterize AML with complex karyotypes. The aberrations seen in AML cases with multiple chromosome rearrangements have often been referred to as random, because the same chromosome rearrangement is rarely identified in more than one patient. However, in the past few years, new cytogenetic methods with significantly higher resolution than G-banding have been introduced and allow a more detailed characterization of MCR at the chromosome level. Spectral karyotyping, CGH, and M-FISH studies have shown that MCR result in loss of chromosome arms 5q, 7q, and 17p and gain of chromosome 8 and arm 11q.[5-7] The principal observations reported here are that these DNA gains and losses affect the gene expression in a gene-dosage dependent manner and that AML with complex karyotypes exhibits a specific gene expression profile.

In cancer research, an increasing number of gene expression profiling studies are identifying new potential diagnostic and prognostic variables, as well as providing clues for improving cancer therapy.[24,25] A common way to identify expression profiles associated with certain phenotypes is to use unsupervised cluster analysis. Unsupervised cluster analysis allows the gene expression patterns to drive the separation of samples into groups, without allowing the experimenter bias to influence the outcome.[26] In this study, all six AML samples with complex karyotypes clustered together even though they represented four different morphologic (FAB) subtypes (Figure 1 and Table 1). Furthermore, the average correlation of the AML-MCR gene expression profiles was higher than that of the normal controls (see *Results*). These findings suggest that there is a great deal of similarity between the expression profiles of AML with complex karyotypes. This is consistent with published studies of AML with primary reciprocal translocations [i.e., t(8;16), inv(16) and t(15;17)].[11,27] The AML samples in these studies also clustered based on cytogenetics rather than on FAB subgroup, suggesting that the expression patterns of AML are strongly linked to karyotypic status.

Furthermore, we used a computational approach, CGMA, to look for gene expression biases that correspond to chromosome regions. We found that the DNA imbalances identified by SKY and CGH could also be detected by CGMA (Table 1). This is consistent with a few other reports in which chromosome aberrations have been associated with changes in gene expression as seen by microarrays.[9,18,19,28-30]

Taken together, these findings suggest that CGMA predictions could be used as a first approximation of DNA copy number, for example when CGH data are not available for a particular cancer type but gene expression microarray data are. CGMA could also be used to confirm existing CGH data and to examine candidate genes whose expression changes most within a region of frequent DNA gain or loss. Additionally, Kyle *et al.* have recently shown that 82% of renal cell carcinoma samples could be correctly subgrouped solely based on their CGMA profiles.[20] It should be noted that in this and in a previous study[18] the consistency of CGMA differed for DNA deletions and amplifications, being higher for deletions (Table 1).

This may reflect a lack of cDNA clones in areas of genetic amplification, or that the size or copy number of the amplicons was too small to be detected by CGMA. It is also possible that this represents a form of gene silencing that has recently been reported for amplified genes in cancer.[31] Interestingly, when we applied CGMA to AML samples with normal cytogenetics, regional low-expression biases were also found (Table 1 and Figure 3). This finding could reflect limitations in the specificity of the CGMA methodology, but it is also possible that these regions harbor small deletions missed by karyotyping. At least two groups have reported on loss of heterozygosity (LOH) in AML.[32,33] Although cytogenetic information was available for most of the patients with regional allelic losses, few deletions were observed on these chromosome arms. Of the reported regions with LOH, 1q, 6q, 7q, 11q, 17p, and 19q showed repeated low-expression biases in our material.

In conclusion, we found that AML with a complex karyotype exhibits a specific expression profile and that multiple chromosome rearrangements alter the gene expression in a gene-dosage-dependent manner. The most differentially expressed genes included a number of genes involved in DNA repair, chromosome segregation and with the actin cytoskeleton. The consequences of regional low-expression biases in AML calls for further investigation, especially since Lu *et al.* recently showed that in Wilms' tumor the

expression pattern of chromosome arm 1q is a more accurate predictor of outcome than the copy number change found by CGH.[34]

CL planned the study together with EB, MN, MB and BTT. CL performed most of the experiments and wrote the manuscript. XG and BH contributed with technical assistance. KF performed most of the bioinformatic analyses. All authors approved the submitted manuscript. The authors reported no potential conflicts of interest

We thank Ramsi Haddad, Sigrid Sahlén, and Jacqueline Schoumans for technical assistance.

This research was supported by the Swedish Cancer Society and by the Van Andel Research Institute.

Manuscript received April 26, 2004. Accepted July 14. 2004.

References

1. Heim S & Mitelman F. Cancer Cytogenetics. Wiley-Liss, New York; 1995. p. 72-3.
2. Grimwade D, Walker H, Oliver F, Wheatley K, Harrison C, Harrison G, et al. 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 1998;92:2322-33.
3. Visani G, Bernasconi P, Boni M, Castoldi GL, Ciolli S, Clavio M, et al. The prognostic value of cytogenetics is reinforced by the kind of induction/consolidation therapy in influencing the outcome of acute myeloid leukemia-analysis of 848 patients. Leukemia 2001;15:903-9.
4. Schoch C, Haferlach T, Haase D, Fonatsch C, Loffler H, Schlegelberger B, et al. Patients with de novo acute myeloid leukaemia and complex karyotype aberrations show a poor prognosis despite intensive treatment: a study of 90 patients. Br J Haematol 2001; 112:118-26.
5. Lindvall C, Nordenskjold M, Porwit A, Bjorkholm M, Blennow E. Molecular cytogenetic characterization of acute myeloid leukemia and myelodysplastic syndromes with multiple chromosome rearrangements. Haematologica 2001;86:1158-64.
6. Van Limbergen H, Poppe B, Michaux L, Herens C, Brown J, Noens L, et al. Identification of cytogenetic subclasses and recurring chromosomal aberrations in AML and MDS with complex karyotypes using M-FISH. Genes Chromosomes Cancer 2002; 33:60-72.
7. Schoch C, Haferlach T, Bursch S, Gerstner D, Schnittger S, Dugas M, et al. Loss of genetic material is more common than gain in acute myeloid leukemia with complex aberrant karyotype: a detailed analysis of 125 cases using conventional chromosome analysis and fluorescence in situ hybridization including 24-color FISH. Genes Chromosomes Cancer 2002;35:20-9.
8. Golub TR, Slonim D K, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999;286:531-7.
9. Virtaneva K, Wright FA, Tanner SM, Yuan B, Lemon WJ, Caligiuri MA, et al. Expression profiling reveals fundamental biological differences in acute myeloid leukemia with isolated trisomy 8 and normal cytogenetics. Proc Natl Acad Sci USA 2001;98:1124-9.
10. Schoch C, Kohlmann A, Schnittger S, Brors B, Dugas M, Mergenthaler S, et al. Acute myeloid leukemias with reciprocal rearrangements can be distinguished by specific gene expression profiles. Proc Natl Acad Sci USA 2002;99:10008-13.
11. Debernardi S, Lillington DM, Chaplin T, Tomlinson S, Amess J, Rohatiner A, et al. Genome-wide analysis of acute myeloid leukemia with normal karyotype reveals a unique pattern of homeobox gene expression distinct from those with translocation-mediated fusion events. Genes Chromosomes Cancer 2003;37:149-58.
12. Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DA, Gralnick HR, et al. Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. Br J Haematol 1976; 33:451-8.
13. Takahashi M, Yang XJ, Lavery TT, Furge KA, Williams BO, Tretiakova M, et al. Gene expression profiling of favorable histology Wilms tumors and its correlation with clinical features. Cancer Res 2002; 62:6598-605.
14. Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, et al. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. Nucleic Acid Res 2002;30:15.
15. Takahashi M, Rhodes DR, Furge KA, Kanayama H, Kagawa S, Haab BB, et al. Gene expression profiling of clear cell renal cell carcinoma: gene identification and prognostic classification. Proc Natl Acad Sci USA 2001;98: 9754-9.
16. Rhodes DR, Miller JC, Haab BB, Furge KA. CIT:identification of differentially expressed clusters of genes from microarray data. Bioinformatics 2002;18:205-6.
17. Ihaka R, Gentleman RR. A language for data analysis and graphics. J Comput Graph Stat 1996; 5:299-314.
18. Haddad R, Furge KA, Miller JC, Haab BB, Schoumans J, Teh BT, et al. Genomic profiling and cDNA microarray analysis of human colon adenocarcinoma and associated intraperitoneal metastases reveals consistent cytogenetic and transcriptional aberrations associated with progression of multiple metastases. Applied Genomics Proteomics 2002;1:123-34.
19. Crawley JJ, Furge KA. Identification of frequent cytogenetic aberrations in hepatocellular carcinoma using gene-expression microarray data. Genome Biol 2002; 3: RESEARCH0075.
20. Furge KA, Lucas KA, Takahashi M, Sugimura J, Kort EJ, Kanayama HO, et al. Robust classification of renal cell carcinoma based on gene expression data and predicted cytogenetic profiles. Cancer Res 2004; 64:4117-21.
21. Kallioniemi OP, Kallioniemi A, Piper J, Isola J, Waldman FM, Gray JW, et al. Optimizing comparative genomic hybridization for analysis of DNA sequence copy number changes in solid tumors. Genes Chromosomes Cancer 1994;10:231-43.
22. Poppe B, Vandesompele J, Schoch C, Lindvall C, Mrozek K, Bloomfield CD, et al. Expression analyses identify MLL as a prominent target of 11q23 amplification and support an etiologic role for MLL gain of function in myeloid malignancies. Blood 2003;28:28.
23. Gerisch G, Faix J, Kohler J & Muller-Taubenberger A. Actin-binding proteins required for reliable chromosome segregation in mitosis. Cell Motil Cytoskeleton 2004; 57: 18-25.
24. Marx J. DNA arrays reveal cancer in its many forms. Science 2000;289:1670-2.
25. Bertucci F, Houlgatte R, Nguyen C, Viens P, Jordan BR, Birnbaum D. Gene expression profiling of cancer by use of DNA arrays: how far from the clinic? Lancet Oncol 2001; 2:674-82.
26. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 1998;95:14863-8.
27. Kohlmann A, Schoch C, Schnittger S, Dugas M, Hiddemann W, Kern W, et al. Molecular characterization of acute leukemias by use of microarray technology. Genes Chromosomes Cancer 2003;37:396-405.
28. Hughes TR, Roberts CJ, Dai H, Jones AR, Meyer MR, Slade D, et al. Widespread aneuploidy revealed by DNA microarray expression profiling. Nat Genet 2000;25:333-7.
29. Xu XR, Huang J, Xu ZG, Qian BZ, Zhu ZD, Yan Q, et al. Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver. Proc Natl Acad Sci USA 2001;98:15089-94.
30. Kauraniemi P, Barlund M, Monni O, Kallioniemi A. New amplified and highly expressed genes discovered in the ERBB2 amplicon in breast cancer by cDNA microarrays. Cancer Res 2001;61:8235-40.
31. Platzer P, Upender MB, Wilson K, Willis J, Lutterbaugh J, Nosrati A, et al. Silence of chromosomal amplifications in colon cancer. Cancer Res 2002;62:1134-8.
32. Mori N, Morosetti R, Hoflehner E, Lubbert M, Mizoguchi H, Koeffler HP. Allelic loss in the progression of myelodysplastic syndrome. Cancer Res 2000;60:3039-42.
33. Sweetser DA, Chen CS, Blomberg AA, Flowers DA, Galipeau PC, Barrett MT, et al. Loss of heterozygosity in childhood de novo acute myelogenous leukemia. Blood 2001; 98:1188-94.
34. Lu YJ, Hing S, Williams R, Pinkerton R, Shipley J, Pritchard-Jones K. Chromosome 1q expression profiling and relapse in Wilms' tumour. Lancet 2002;360:385-6.