Available online at www.sciencedirect.com
SCIENCE DIRECT®
Leukemia Research 29 (2005) 943–949
Leukemia Research
www.elsevier.com/locate/leukres

Retroviral expression screening of oncogenes in natural killer cell leukemia

Young Lim Choi [a, b], Ryozo Moriuchi [c], Mitsujiro Osawa [d], Atsushi Iwama [d], Hideki Makishima [e], Tomoaki Wada [a], Hiroyuki Kisanuki [a], Ruri Kaneda [a], Jun Ota [a, f], Koji Koinuma [a], Madoka Ishikawa [a], Shuji Takada [a], Yoshihiro Yamashita [a], Kazuo Oshimi [b], Hiroyuki Mano [a, f, *]

[a] *Division of Functional Genomics, Jichi Medical School, 3311-1 Yakushiji, Kawachigun, Tochigi 329-0498, Japan*
[b] *Division of Hematology, Department of Medicine, Juntendo University School of Medicine, Tokyo, Japan*
[c] *Department of Molecular Microbiology and Immunology, Nagasaki University Graduate School of Medicine, Nagasaki, Japan*
[d] *Center for Experimental Medicine, Institute of Medical Science, University of Tokyo, Tokyo, Japan*
[e] *Second Department of Internal Medicine, Shinshu University School of Medicine, Nagano, Japan*
[f] *CREST, Japan Science and Technology Agency, Saitama, Japan*

Received 17 October 2004; accepted 22 January 2005
Available online 24 February 2005

Abstract

Aggressive natural killer cell leukemia (ANKL) is an intractable malignancy that is characterized by the outgrowth of NK cells. To identify transforming genes in ANKL, we constructed a retroviral cDNA expression library from an ANKL cell line KHYG-1. Infection of 3T3 cells with recombinant retroviruses yielded 33 transformed foci. Nucleotide sequencing of the DNA inserts recovered from these foci revealed that 31 of them encoded KRAS2 with a glycine-to-alanine mutation at codon 12. Mutation-specific PCR analysis indicated that the *KRAS* mutation was present only in KHYG-1 cells, not in another ANKL cell line or in clinical specimens ($n = 8$).

Keywords: Aggressive NK cell leukemia; cDNA expression library; Retrovirus; KRAS2 oncogene

1. Introduction

Outgrowth of $CD3^{-}CD16/CD56^{+}$ natural killer (NK) cells in peripheral blood is diagnosed as either chronic NK lymphocytosis (CNKL) or aggressive NK cell leukemia (ANKL) [1,2]. Whereas the former condition has an indolent clinical course with few symptoms, the latter is characterized by chemoresistance and multiorgan failure and has a poor outcome.

The Epstein–Barr virus (EBV) genome is frequently present episomally in ANKL cells [3], suggesting a role for EBV in disease pathogenesis. However, little is known of how infection with EBV might trigger clonal growth of NK cells. Inactivation of tumor suppressor genes has been associated with NK cell neoplasia. For instance, a homozygous deletion of the genes for p16INK4A, p15INK4B, or p14ARF has been detected [4]. Additionally, inactivating mutations of the FAS gene have been found in nasal NK/T cell lymphoma [5].

A few studies have identified a potential contribution of oncogenes to NK cell malignancy. Mutations that affect codons 13 or 22 of *KRAS2* were found in NK/T cell lymphoma [6] but not in ANKL [7]. Furthermore, although mutations in *KIT* were associated with NK/T cell lymphoma, transforming activity of the mutant KIT proteins was not demonstrated [8]. A role for oncogenes in ANKL has not been identified to date.

Functional screening based on transforming ability is one potential approach to the efficient isolation of tumor-promoting genes in ANKL. Focus formation assays with

* Corresponding author. Tel.: +81 285 58 7449; fax: +81 285 44 7322.
E-mail address: hmano@jichi.ac.jp (H. Mano).

0145-2126/$ – see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.leukres.2005.02.002

mouse 3T3 fibroblasts have indeed proved successful for the identification of oncogenes in human cancer [9]. In such screening assays, genomic DNA isolated from cancer specimens is used to transfect 3T3 cells and the formation of transformed cell foci is then evaluated. Expression of the exogenous genes in such experiments is driven by their own promoters or enhancers, however, so that oncogenes can exert transforming effects in 3T3 cells only if their regulatory regions are active in fibroblasts. Given the distinct developmental origins of NK cells and fibroblasts, expression of oncogenes associated with ANKL in 3T3 cells under these conditions is not guaranteed.

This problem might be expected to be overcome by the expression of test cDNAs under the control of an ectopic promoter in 3T3 cells. We have therefore constructed a retroviral cDNA expression library from the ANKL cell line KHYG-1 [10] and used this library to infect 3T3 cells. In preparation of the cDNA library, we took advantage of a polymerase chain reaction (PCR)-based system that preferentially amplifies full-length cDNAs. The resulting library was found to have sufficient complexity and to contain a high percentage of full-length cDNAs. Focus formation assays with 3T3 cells resulted in the identification of *KRAS2* as a transforming gene in KHYG-1 cells.

2. Materials and methods

2.1. Cell culture and clinical samples

KHYG-1 and NKL cells [11] were kindly provided by M. Yagita and Y. Yokota, respectively, and were cultured in RPMI 1640 medium (Invitrogen, Carlsbad, CA) supplemented with 10% fetal bovine serum (Invitrogen) and human interleukin-2 (20 U/mL) (Roche, St. Louis, MO). The BOSC23 packaging cell line for ecotropic retroviruses [12] and mouse 3T3 fibroblasts were maintained in Dulbecco's modified Eagle's medium (DMEM)–F12 (Invitrogen) supplemented with 10% fetal bovine serum and 2 mM L-glutamine.

Mononulear cells were isolated by Ficoll–Hypaque density gradient centrifugation from peripheral blood of the subjects with informed consent. The cells were incubated with anti-CD3 MicroBeads (Miltenyi Biotec, Auburn, CA), and loaded onto MIDI-MACS magnetic cell separation columns (Miltenyi Biotec) to remove $CD3^+$ cells. The flow-through was then mixed with anti-CD56 MicroBeads (Miltenyi Biotec), and was subjected to a MINI-MACS column for the "positive selection" of $CD56^+$ cells. Cells bound specifically to the column was then eluted according to the manufacturer's instructions.

2.2. Construction of a retrovirus library

Total RNA was extracted from KHYG-1 cells with the use of an RNeasy Mini column and RNase-free DNase (Qiagen, Valencia, CA), and first-strand cDNA was synthesized from the RNA with PowerScript reverse transcriptase, a SMART IIA oligonucleotide, and CDS primer IIA (Clontech, Palo Alto, CA). The resulting cDNA molecules were then amplified for 12 cycles with 5′-PCR primer IIA and a SMART PCR cDNA synthesis kit (Clontech), with the exception that LA Taq polymerase (Takara Bio, Shiga, Japan) was substituted for the Advantage 2 DNA polymerase provided with the kit. The PCR products were treated with proteinase K, rendered blunt-ended with T4 DNA polymerase, and ligated to a *Bst*XI adapter (Invitrogen). Unbound adapters were removed with a cDNA size fractionation column (Invitrogen), and the modified cDNAs were ligated into the pMX retroviral plasmid (kindly provided by T. Kitamura) [13] that had been digested with *Bst*XI. The pMX-cDNA plasmids were introduced into ElecroMax DH10B cells (Invitrogen) by electroporation.

2.3. Focus formation assay

BOSC23 cells (1.8×10^6) were seeded onto 6-cm culture plates, cultured for 1 day, and then transfected with a mixture comprising 2 μg of retroviral plasmids, 0.5 μg of pGP plasmid (Takara Bio), 0.5 μg of pE-eco plasmid (Takara Bio), and 18 μL of Lipofectamine reagent (Invitrogen). Two days after transfection, polybrene (Sigma, St. Louis, MO) was added at a concentration of 4 μg/mL to the culture supernatant, which was then used to infect 3T3 cells for 48 h. For the focus formation assay, the culture medium of 3T3 cells was changed to DMEM–high glucose (Invitrogen) supplemented with 5% calf serum and 2 mM L-glutamine. Transformed foci were isolated after 3 weeks of culture.

2.4. Recovery of cDNAs from 3T3 cells

Each 3T3 cell clone was harvested with a cloning syringe and cultured independently in a 10-cm culture plate. Genomic DNA was subsequently extracted from the cells and subjected to PCR with 5′-PCR primer IIA and LA Taq polymerase for 50 cycles of 98 °C for 20 s and 68 °C for 6 min. Amplified genomic fragments were purified by gel electrophoresis and ligated into the pT7Blue-2 vector (EMD Biosciences, San Diego, CA) for nucleotide sequencing.

2.5. Mutation-specific PCR for KRAS2

Detection of $KRAS2^{G12A}$ cDNA was performed as described previously [14] but with minor modifications. In brief, a 5′-region of *KRAS2* cDNA was amplified from oligo(dT)–primed cDNA by PCR with 5′-RAS primer (5′-ACTGAATATAAACTTGTGGTAGTTGGACCT-3′; the underlined cytosine was incorporated to generate a *Bst*NI site) and 3′-RAS primer A (5′-CTGTGTCGAGAATATCCAAGAGACA-3′). The PCR product was subjected to digestion with *Bst*NI (New England Biolabs, Beverly, MA) and then to a second PCR with 5′-RAS primer and 3′-RAS primer B (5′-CTGTGTCGAGAATCCAGGAGACA-3′; the under-

lined guanine was incorporated to generate a *Bst*NI site). The second PCR product was then also subjected to digestion with *Bst*NI, and the resulting DNA fragments were separated by agarose gel electrophoresis.

3. Results

3.1. Construction of a full-length cDNA expression library for KHYG-1 cells

Full-length cDNAs were selectively amplified from mRNA of KHYG-1 cells and ligated into the retroviral vector pMX. We obtained a total of 5.61×10^6 colony-forming units (cfu) of independent plasmid clones. To evaluate the quality of the library, we randomly selected 40 clones and examined the incorporated cDNAs. Thirty-nine of the 40 clones contained inserts with an average size of 2.03 kbp. The cDNA inserts from 20 out of these 39 clones were sequenced from both ends, and the determined sequences were used to screen, with the BLAT search program [15], the nucleotide sequence database assembled as of July 2003 by the Genome Bioinformatics Group of the University of California at Santa Cruz (http://genome.ucsc.edu/). Both ends of 14 of the 20 cDNAs could be matched to the mRNA sequences of known genes, and 13 of these cDNAs included complete open reading frames (data not shown). We therefore concluded that the retroviral cDNA expression library was of sufficient complexity and sufficiently enriched in full-length cDNAs for the present study.

3.2. Identification of $KRAS2^{G12A}$ in KHYG-1 cells

We generated a recombinant ecotropic retrovirus library by introducing 7.1×10^5 cfu of the generated plasmids into a packaging cell line. This library was then used to infect mouse 3T3 fibroblasts. After culture of the infected cells for 3 weeks, we detected 33 transformed foci (Fig. 1). Each focus was isolated, expanded independently, and subjected to extraction of genomic DNA for the recovery of retroviral inserts by PCR with the primer used originally to amplify the cDNAs during construction of the library. In most instances, a single major DNA fragment was recovered from each genome (Fig. 2A), suggestive of original infection of a single 3T3 cell with a single retrovirus.

The recovered cDNA fragments were sequenced from both ends for all 33 clones. Screening of the human genome sequence database with the insert sequences revealed that those from 31 of the 33 clones (#1–#29, #31, #33) matched, with >98% identity, the sequence of human *KRAS2* (GenBank accession number, NM_004985). The genome of 3T3 clone ID #30 yielded two PCR fragments (Fig. 2A); the larger (~1.4 kbp) and the smaller (~0.9 kbp) fragments were revealed to be derived from β-actin (*ACTB*; GenBank accession number, NM_001101) and profilin 1 (*PFN1*; GenBank accession number, NM_005022) genes, respectively. The final 3T3 clone (#32) yielded a major PCR fragment corresponding to the gene for isocitrate dehydrogenase 3 (NAD^+) β (*IDH3B*; GenBank accession number, NM_006899).

KRAS2 belongs to the *RAS* gene family and is involved in a wide variety of human cancers [16]. Given that point

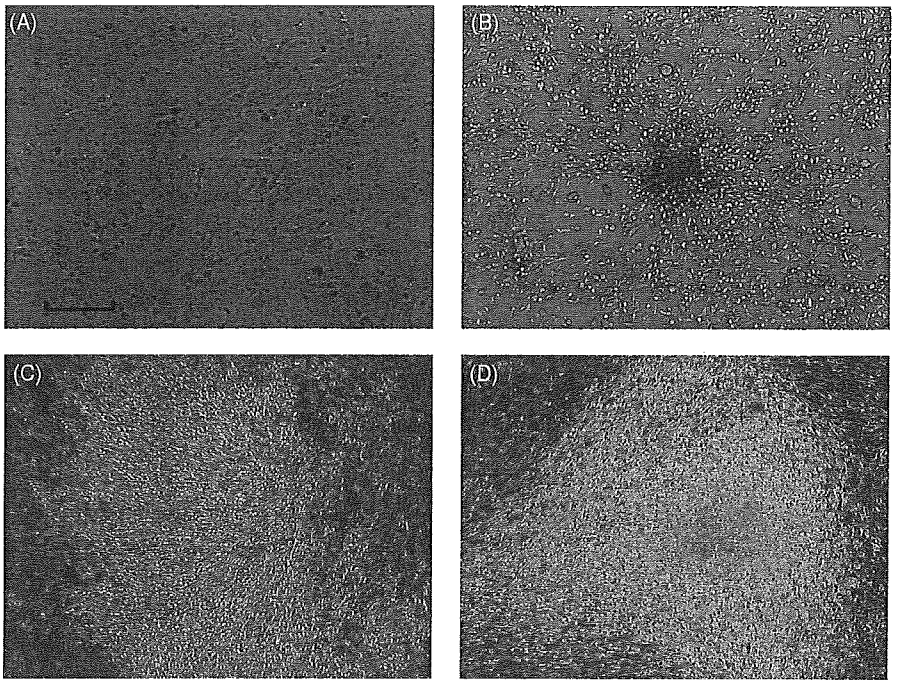

Fig. 1. Focus formation assay with a retroviral library derived from KHYG-1 cells. Mouse 3T3 cells were infected with the empty virus (A), a retrovirus expressing v-Ras as a positive control (B), or retroviruses from the KHYG-1 cell library (C and D). The cultures were photographed 3 weeks after infection. Scale bar, 100 μm.

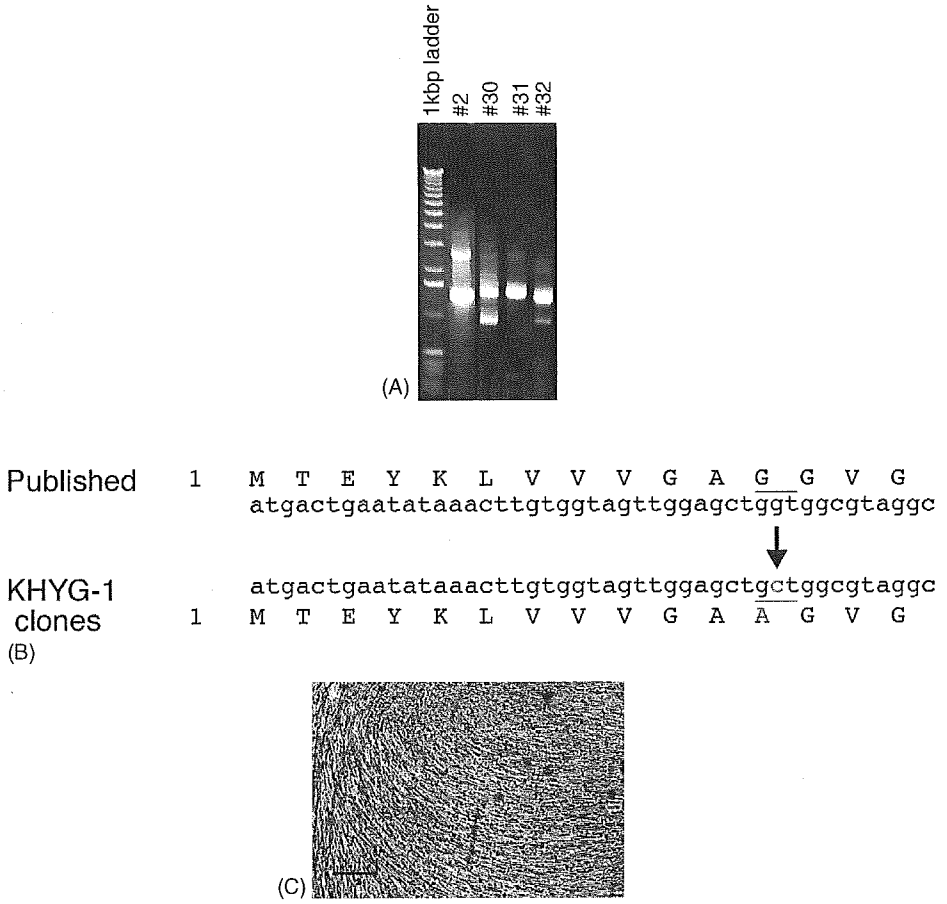

Fig. 2. Identification of a *KRAS* mutant gene with transforming activity: (A) Genomic DNA isolated from transformed 3T3 cell foci (clones #2, #30, #31, and #32) was subjected to PCR for amplification of the DNA inserts. The left lane contains DNA size markers (1-kbp DNA ladder; Invitrogen); (B) The nucleotide sequences of the DNA inserts derived from 31 of the 33 transformed foci matched that of *KRAS* with a single nucleotide substitution (G to C) in the second position of codon 12 that results in a change in the encoded amino acid from glycine to alanine; (C) A recombinant retrovirus encoding KRAS2^{G12A} was used to infect 3T3 cells. The cells were photographed after culture for 2 weeks. Scale bar, 50 μm.

mutations in *KRAS2* confer oncogenic activity, we compared the nucleotide sequences of the *KRAS2* cDNAs derived from KHYG-1 cells with the published sequence of the wild-type gene. Although oncogenic mutations have previously been shown to affect codons 12, 13, and 59 of *KRAS2*, all of the KHYG-1 cell cDNAs harbored the same mutation: the GGT sequence of codon 12 was changed to GCT, resulting in the substitution of an alanine residue for the glycine normally present at this position (Fig. 2B). To verify the transforming ability of *KRAS2*^{G12A}, we inserted the corresponding cDNA into the pMX retroviral vector and generated recombinant retroviruses for infection of 3T3 cells. The recipient 3T3 cells indeed underwent transformation (Fig. 2C), confirming that *KRAS2*^{G12A} possesses oncogenic activity.

3.3. Screening for *KRAS2*^{G12A} in NK cell leukemia

To determine whether *KRAS2*^{G12A} is frequently associated with NK cell leukemia, we applied a slightly modified version of a rapid screening method previously described by Kahn et al. [14]. *KRAS2* cDNA was first amplified by PCR with 5′-RAS primer and 3′-RAS primer A (Fig. 3A). The 3′ end of 5′-RAS primer corresponds to codon 11 of *KRAS2* but contains a cytosine substitution in the first position of codon 11, which generates a *Bst*NI recognition site [CC(T/A)GG] that includes the first and second positions of codon 12. The presence of a point mutation at the first or second position of codon 12 would therefore prevent digestion of the PCR product by *Bst*NI.

After *Bst*NI digestion, the PCR product was subjected to a second PCR with 5′-RAS primer and 3′-RAS primer B. Given that *Bst*NI digestion removes the binding site for 5′-RAS primer, only *KRAS2* cDNA with a mutation at the first or second position of codon 12 should yield a second PCR product. Even if *Bst*NI digestion of the initial PCR product was not complete and the second PCR amplified a trace amount of wild-type *KRAS2* cDNA, a second *Bst*NI digestion further discriminates between the wild-type and mutant genes. The 3′-RAS primer B thus contains a guanine substitution that generates a *Bst*NI site. The second PCR product of wild-type *KRAS2* cDNA would thus contain two *Bst*NI sites, whereas that of mutant *KRAS2* contains only one.

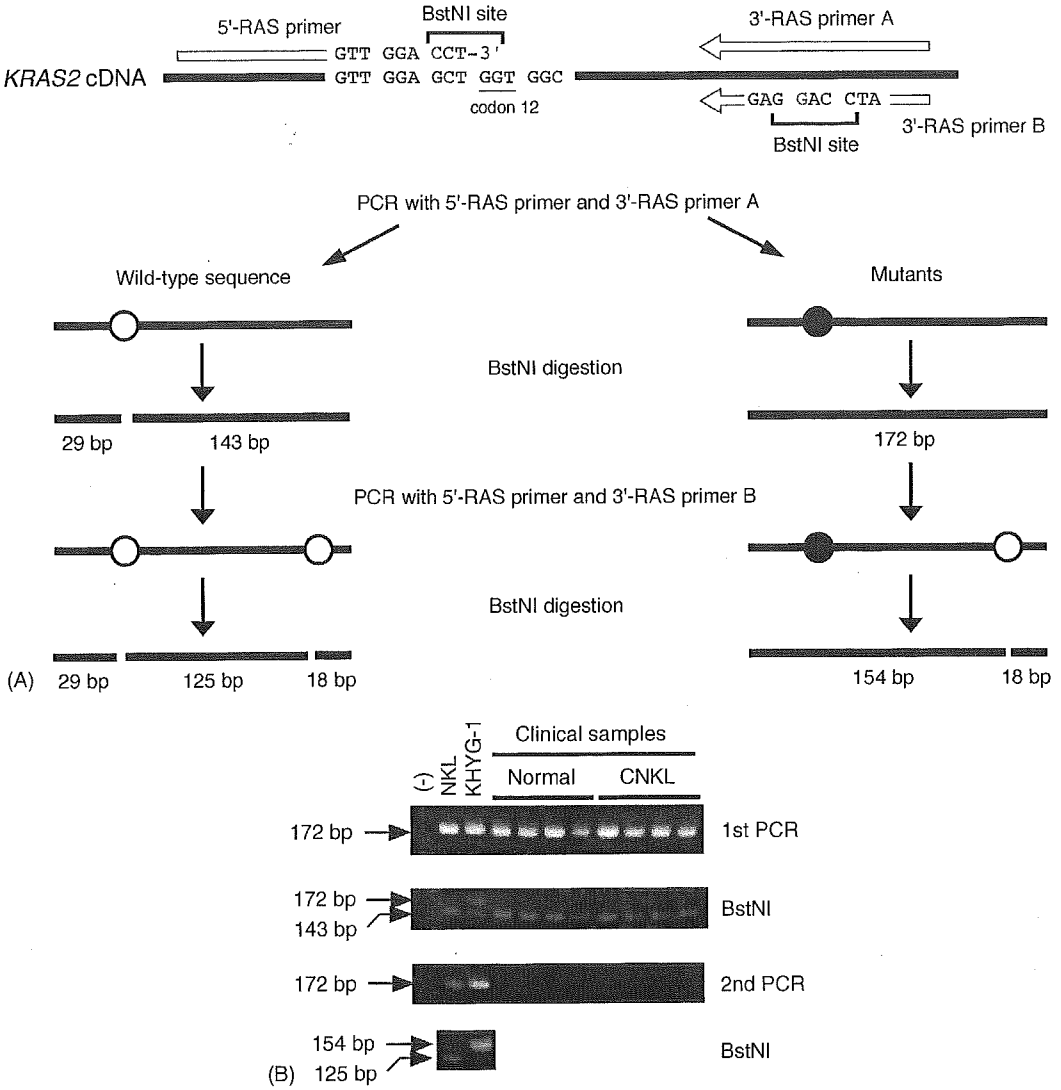

Fig. 3. Mutation-specific PCR analysis of NK cell leukemia cell lines and clinical specimens: (A) *KRAS2* cDNA was amplified with 5′-RAS primer and 3′-RAS primer A. The PCR product was subjected to digestion with *Bst*NI and then to a second PCR with 5′-RAS primer and 3′-RAS primer B. The second PCR product was also subjected to digestion with *Bst*NI. The nucleotides shown in red were incorporated into the primers to generate a *Bst*NI site. Open circles indicate *Bst*NI sites; closed circles indicate corresponding mutant sequences that are not susceptible to *Bst*NI; (B) cDNA isolated from the NKL and KHYG-1 cell lines as well as from $CD3^{-}CD56^{+}$ NK cell fractions derived from healthy volunteers (Normal) or individuals with CNKL was subjected to mutation-specific PCR analysis. A reaction without input DNA was also performed as a negative control (–). The size of DNA fragments is indicated on the left.

With this approach, we analyzed cDNA prepared from two ANKL cell lines (NKL and KHYG-1) and from $CD3^{-}CD56^{+}$ NK cell fractions purified from the peripheral blood of healthy individuals ($n=4$) and patients with CNKL ($n=4$). The first PCR step yielded a single DNA fragment of 172 bp from all samples. Furthermore, only the PCR product from KHYG-1 cells was refractory to *Bst*NI digestion, indicating that only KHYG-1 cells harbor a codon 12 mutation of *KRAS2*. The presence of the 143-bp band may indicate that KHYG-1 cells are heterozygous for the *KRAS2* mutation. The second PCR generated a 172-bp DNA fragment only with the NKL and KHYG-1 cell samples. Whereas this fragment derived from NKL cells was completely digested by *Bst*NI to generate a 125-bp band, *Bst*NI digestion of the fragment derived from KHYG-1 cells generated a band of 154 bp. Of all the samples analyzed, therefore, mutation of the first or second position of codon 12 of *KRAS2* was detected only in KHYG-1 cells.

4. Discussion

We have constructed a retroviral cDNA expression library for an ANKL cell line. Given that >97% (39/40) of the viral plasmids contained cDNA inserts and that the overall clone number was $>5 \times 10^6$, our library likely represents most of

the transcriptome of KHYG-1 cells. The high probability that the incorporated cDNAs are full length is also an important advantage for functional screening.

In our screening, most of the 3T3 transformants were found to have incorporated a single cDNA fragment corresponding to *KRAS2*^{G12A}, with only two transformants found to contain other cDNAs. One of these two cDNA inserts was derived from the gene for PFN1, a protein that binds to unpolymerized actin [17]. Homozygous deletion of *Pfn1* results in embryonic death in mice, with the encoded protein apparently being indispensable for cell growth or differentiation during embryonic development [18]. The other cDNA insert isolated from 3T3 transformants contained the entire open reading frame for IDH3B, which catalyzes the oxidative decarboxylation of isocitrate and is a key enzyme in the tricarboxylic acid cycle [19]. Neither *PFN1* nor *IDH3B* has previously been shown to possess oncogenic activity. It is currently under examination whether a long terminal repeat (LTR)-driven overexpression of *PFN1* or *IDH3B* leads focus formation in 3T3 cells.

Comparative genomic hybridization analysis identified a wide variety of genetic alterations at a high frequency in ANKL cells [20], suggesting that leukemogenesis in ANKL is associated with multiple steps of oncogene activation. An analysis of patients with NK cell neoplasia failed to detect any changes in the genes for members of the RAS and MYC families of proteins [7], however. This previous study did find a frequent increase in the abundance of the cell cycle regulator MDM2.

In contrast, we have detected a transforming *KRAS2* mutant gene in an ANKL cell line. Given that the mutation in codon 12 of this gene was detected by two different approaches (retroviral screening of PCR-amplified cDNAs and mutation-specific PCR), we conclude that it was not an artifact of PCR. KRAS2 is a GTP-binding protein with a relative molecular mass of ~21 kDa. Together with HRAS and NRAS, it plays an important role in cell growth and differentiation. Many mitogenic signals promote the loading of KRAS2 with GTP, which in turn triggers various downstream signaling events including activation of the mitogen-activated protein kinase (MAPK) pathway.

Activating mutations of *KRAS2* have been identified in a wide range of human cancers. Mutations of codon 12, for example, are associated with acute lymphoblastic leukemia [21], lung cancer [22], and pancreatic cancer [23]. No such mutations have previously been detected in association with ANKL, however. Although we have now identified a *KRAS2* mutation affecting codon 12 in the ANKL cell line KHYG-1, we did not detect this mutation in another ANKL cell line (NKL) or in $CD3^{-}CD56^{+}$ NK cell fractions isolated from healthy volunteers or from individuals with CNKL. Mutation of *KRAS2*, at least of codon 12 of this gene, might therefore be an infrequent event in NK cell neoplasia. Indeed, it remains possible that the detected *KRAS2* mutation is specific to the KHYG-1 cell line. Nevertheless, our identification of an activating *KRAS2* mutation in KHYG-1 cells might provide insight into the role of the RAS-MAPK signaling pathway in ANKL carcinogenesis. Furthermore, given the high quality of our retroviral expression library, the strategy adopted in the present study also might prove useful for the functional screening of genes associated with various clinical characteristics of ANKL, such as chemoresistance.

Acknowledgments

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 grants from Research Foundation for Community Medicine of Japan, Sankyo Foundation of Life Science, Takeda Science Foundation, and Mitsubishi Pharma Research Foundation. Y.-L.C conducted most of the experiments. R.M., M.O., A.I. and T.W. helped to establish a retroviral expression library. Hideki Makishima, J.O. and K.O. collected the ANKL specimens, and conducted the mutation-specific PCR method. H.K., R.K., K.K., M.I., S.T. and Y.Y. helped the 3T3 focus formation screening, and provided suggestions on molecular biology. Hiroyuki Mano designed this project with Y.-L.C., and was responsible for all aspects of this project.

References

[1] Harris NL, Jaffe ES, Diebold J, Flandrin G, Muller-Hermelink HK, Vardiman J, et al. World Health Organization classification of neoplastic diseases of the hematopoietic and lymphoid tissues: report of the Clinical Advisory Committee meeting, Airlie House, Virginia, November. J Clin Oncol 1999;17:3835–49.

[2] Oshimi K. Leukemia and lymphoma of natural killer lineage cells. Int J Hematol 2003;78:18–23.

[3] Gelb AB, van de Rijn M, Regula Jr DP, Cornbleet JP, Kamel OW, Horoupian DS, et al. Epstein–Barr virus-associated natural killer-large granular lymphocyte leukemia. Hum Pathol 1994;25:953–60.

[4] Sakajiri S, Kawamata N, Egashira M, Mori K, Oshimi K. Molecular analysis of tumor suppressor genes, Rb, p53, p16INK4A, p15INK4B and p14ARF in natural killer cell neoplasms. Jpn J Cancer Res 2001;92:1048–56.

[5] Takakuwa T, Dong Z, Nakatsuka S, Kojya S, Harabuchi Y, Yang WI, et al. Frequent mutations of Fas gene in nasal NK/T cell lymphoma. Oncogene 2002;21:4702–5.

[6] Hoshida Y, Hongyo T, Nakatsuka S, Nishiu M, Takakuwa T, Tomita Y, et al. Gene mutations in lymphoproliferative disorders of T and NK/T cell phenotypes developing in renal transplant patients. Lab Invest 2002;82:257–64.

[7] Sugimoto KJ, Kawamata N, Sakajiri S, Oshimi K. Molecular analysis of oncogenes, ras family genes (N-ras, K-ras, H-ras), myc family genes (c-myc, N-myc) and mdm2 in natural killer cell neoplasms. Jpn J Cancer Res 2002;93:1270–7.

[8] Hongyo T, Li T, Syaifudin M, Baskar R, Ikeda H, Kanakura Y, et al. Specific c-kit mutations in sinonasal natural killer/T-cell lymphoma in China and Japan. Cancer Res 2000;60:2345–7.

[9] Aaronson SA. Growth factors and cancer. Science 1991;254:1146–53.

[10] Yagita M, Huang CL, Umehara H, Matsuo Y, Tabata R, Miyake M, et al. A novel natural killer cell line (KHYG-1) from a patient with aggressive natural killer cell leukemia carrying a p53 point mutation. Leukemia 2000;14:922–30.
[11] Robertson MJ, Cochran KJ, Cameron C, Le JM, Tantravahi R, Ritz J, et al. Characterization of a cell line, NKL, derived from an aggressive human natural killer cell leukemia. Exp Hematol 1996;24:406–15.
[12] Pear WS, Nolan GP, Scott ML, Baltimore D. Production of high-titer helper-free retroviruses by transient transfection. Proc Natl Acad Sci USA 1993;90:8392–6.
[13] Onishi M, Kinoshita S, Morikawa Y, Shibuya A, Phillips J, Lanier LL, et al. Applications of retrovirus-mediated expression cloning. Exp Hematol 1996;24:324–9.
[14] Kahn SM, Jiang W, Culbertson TA, Weinstein IB, Williams GM, Tomita N, et al. Rapid and sensitive nonradioactive detection of mutant K-ras genes via 'enriched' PCR amplification. Oncogene 1991;6:1079–83.
[15] Kent WJ. BLAT—the BLAST-like alignment tool. Genome Res 2002;12:656–64.
[16] Ayllon V, Rebollo A. Ras-induced cellular events. Mol Membr Biol 2000;17:65–73.
[17] Goldschmidt-Clermont PJ, Janmey PA, Profilin. A weak CAP for actin and RAS. Cell 1991;66:419–21.
[18] Witke W, Sutherland JD, Sharpe A, Arai M, Kwiatkowski DJ. Profilin I is essential for cell survival and cell division in early mouse development. Proc Natl Acad Sci USA 2001;98:3832–6.
[19] Kim YO, Park SH, Kang YJ, Koh HJ, Kim SH, Park SY, et al. Assignment of mitochondrial NAD(+)-specific isocitrate dehydrogenase beta subunit gene (IDH3B) to human chromosome band 20p13 by in situ hybridization and radiation hybrid mapping. Cytogenet Cell Genet 1999;86:240–1.
[20] Siu LL, Wong KF, Chan JK, Kwong YL. Comparative genomic hybridization analysis of natural killer cell lymphoma/leukaemia. Recognition of consistent patterns of genetic alterations. Am J Pathol 1999;155:1419–25.
[21] Perentesis JP, Bhatia S, Boyle E, Shao Y, Shu XO, Steinbuch M, et al. RAS oncogene mutations and outcome of therapy for childhood acute lymphoblastic leukemia. Leukemia 2004;18:685–92.
[22] Santos E, Martin-Zanca D, Reddy EP, Pierotti MA, Della Porta G, Barbacid M. Malignant activation of a K-ras oncogene in lung carcinoma but not in normal tissue of the same patient. Science 1984;223:661–4.
[23] Motojima K, Urano T, Nagata Y, Shiku H, Tsurifune T, Kanematsu T. Detection of point mutations in the Kirsten-ras oncogene provides evidence for the multicentricity of pancreatic carcinoma. Ann Surg 1993;217:138–43.

Progress in Hematology

International Journal of
HEMATOLOGY

Stratification of Acute Myeloid Leukemia Based on Gene Expression Profiles

Hiroyuki Mano

Division of Functional Genomics, Jichi Medical School, Tochigi, Japan

Received July 28, 2004; accepted August 10, 2004

Abstract

Acute myeloid leukemia (AML) is characterized by clonal growth of immature leukemic blasts and develops either de novo or secondarily to anticancer treatment or to other hematologic disorders. Given that the current classification of AML, which is based on blast karyotype and morphology, is not sufficiently robust to predict the prognosis of each affected individual, new stratification schemes that are of better prognostic value are needed. Global profiling of gene expression in AML blasts has the potential both to identify a small number of genes whose expression is associated with clinical outcome and to provide insight into the molecular pathogenesis of this condition. Emerging genomics tools, especially DNA microarray analysis, have been applied in attempts to isolate new molecular markers for the differential diagnosis of AML and to identify genes that contribute to leukemogenesis. Progress in bioinformatics has also yielded means with which to classify patients according to clinical parameters such as long-term prognosis. The application of such analysis to large sets of gene expression data has begun to provide the basis for a new AML classification that is more powerful with regard to prediction of prognosis.
Int J Hematol. 2004;80:389-394. doi: 10.1532/IJH97.04111

Key words: DNA microarray; CD133; CD34; Transcriptome

1. Introduction

The human genome project was launched in 1991 as a joint program by the United States, the United Kingdom, Japan, France, Germany, and China. Ten years later, the first draft sequence of the entire human genome, including the nucleotide sequence of approximately 90% of euchromatin with 99.9% accuracy, was completed by the international team [1]. The total number of protein-coding genes was estimated to be approximately 31,000, only twice that for a nematode [2]. Completion of the human sequencing project was announced in 2003, with the determined nucleotide sequence encompassing >99% of euchromatin with an accuracy of 99.99% (http://www.ncbi.nlm.nih.gov/genome/guide/human). Annotation of the final sequence is still in progress but is expected to be completed soon. The postgenome era is therefore about to begin.

Compilation of the catalog of human genes has spurred the development of new technologies to investigate and characterize changes—at the gene, messenger RNA (mRNA), or protein level—in the entire gene set simultaneously. Among such genomics approaches, DNA microarray analysis has probably been the most successful to date. A DNA microarray resembles a microscope slide and contains tens of thousands of genomic fragments, complementary DNAs (cDNAs), or oligonucleotides present in individual spots. Hybridization of such arrays with labeled cDNAs derived from a sample mRNA population allows measurement of the amounts of the original mRNAs for all genes represented on the array [3]. The high density of DNA fragments achievable on currently available microarrays makes it possible to quantitate the mRNA levels for all human genes with 1 array experiment. Such profiling of gene expression in a given cell or tissue type promises to provide a new dimension to our understanding of biology.

Microarray analysis has been applied, for instance, to characterize the differentiation [4] and the inflammatory response [5] of human cells. Gene expression profiling has also helped to develop new classification systems for human diseases that had previously been categorized on the basis of pathology or cell morphology. Microarray analysis of human specimens of prostate cancer, for example, has resulted in the identification of new biomarkers that predict poor prognosis [6]. Characteristic patterns of gene expression, or "molecular signatures," are similarly expected to provide a basis for the subdivision of patients with the same clinical diagnosis into groups with distinct prognoses [7].

Acute myeloid leukemia (AML) is characterized by the clonal growth of immature leukemic blasts in bone marrow

Correspondence and reprint requests: Hiroyuki Mano, MD, PhD, Division of Functional Genomics, Jichi Medical School, 3311-1 Yakushiji, Kawachigun, Tochigi 329-0498, Japan; 81-285-58-7449; fax: 81-285-44-7322 (e-mail: hmano@jichi.ac.jp).

(BM) and can develop de novo or from either myelodysplastic syndrome (MDS) or anticancer treatment [8]. One of the most robust predictors of AML prognosis is blast karyotype [9,10]. A good prognosis is thus predicted from the presence in leukemic clones of t(8;21), t(15;17), or inv(16) chromosomal rearrangements, whereas –7/7q–, 11q23, or more complex abnormalities are indicative of a poor outcome. Such stratification is not informative, however, for predicting the prognoses of patients with a normal karyotype, who constitute approximately 50% of the AML population.

A clinical record of a preceding MDS phase is also an indicator of poor prognosis for individuals with AML. Therapy-related acute leukemia (TRL) can develop after the administration of alkylating agents, topoisomerase inhibitors, or radiotherapy. The clinical outcome of TRL is generally worse than that of de novo AML [11], and a subset of individuals with TRL also exhibits multilineage dysplasia of blood cells. Prediction of the outcome of and optimization of the treatment for each AML patient would thus be facilitated by the ability to differentiate de novo AML from MDS-related AML and TRL. However, dysplastic changes (in particular, dyserythropoiesis) in differentiated blood cells are also found not infrequently in the BM of healthy elderly individuals [12]. The differential diagnosis among AML-related disorders is therefore not always an easy task in the clinical setting, especially if a prior record of hematopoietic parameters is not available.

The application of DNA microarray analysis to AML has the potential (1) to identify molecular markers for the differential diagnosis of AML-related disorders, (2) to provide a basis for the subclassification of such disorders, and (3) to yield insight into the molecular pathogenesis of AML. In this article, I review progress related to the first 2 of these goals.

2. Identification of Molecular Markers for the Differential Diagnosis of AML

2.1. Karyotype

Given that the current classification of AML relies on blast karyotype, it would be informative to determine whether karyotype is related to the gene expression profile of blasts. In other words, is DNA microarray analysis able to substitute for conventional karyotyping?

Schoch et al performed microarray analysis with BM mononuclear cells (MNCs) isolated from individuals with AML and compared the data among the patients with t(8;21), t(15:17), or inv(16) chromosomal anomalies [13]. Each of these 3 AML subgroups was found to possess a distinct molecular signature, and it was possible to predict the karyotype correctly on the basis of the expression level of specific genes. The leukemic blasts of these subgroups of AML manifest distinct differentiation abilities, however. Blasts with t(8;21) remain as immature myeloblasts, those with t(15;17) differentiate into promyelocytes, and those with inv(16) differentiate into cells of the monocytic lineage. The overall gene expression profiles of these 3 types of blasts therefore might be substantially affected by the mRNA repertoires of the differentiated cells present within BM. It remains to be determined whether such "karyotype-specific" molecular signatures are indeed dependent on karyotype or are related to French-American-British (FAB) subtype (differentiation ability).

The gene expression profiles of BM MNCs derived from a large number of pediatric AML patients were examined by Yagi et al [14]. Clustering of these patients according to the expression pattern of the entire gene set resulted in their separation into FAB subtype–matched groups, indicative of a prominent influence of differentiated cells within BM on the gene expression profile. It may nevertheless prove possible to capture bona fide karyotype-dependent genes from large data sets with the use of sophisticated bioinformatics approaches and then to use the expression profiles of these genes for "pseudokaryotyping."

Virtaneva et al purified $CD34^+$ progenitor cells from the BM of individuals with AML and compared the gene expression profiles of the patients with a normal karyotype and those with trisomy 8 [15]. They also compared such AML blasts with $CD34^+$ fractions isolated from the BM of healthy volunteers. The use of $CD34^+$ (immature) cells for microarray analysis would be expected to reduce the influence of differentiated cells within BM on the overall pattern of gene expression. These researchers found that the AML blasts differed markedly from normal $CD34^+$ cells in terms of the gene expression profile. However, the blasts with trisomy 8 did not appear to differ substantially from those with a normal karyotype. It is possible that blasts with a normal karyotype or those with trisomy 8 are too diverse to allow identification of distinguishing gene markers.

2.2. De Novo AML versus MDS-Related AML

Although dysplasia is the diagnostic hallmark of MDS, such abnormal cell morphology is also associated with other conditions, and the subjective assessment of the extent of dysplasia suffers from the risk of variability from physician to physician. It is therefore desirable to identify molecular markers that are able to distinguish de novo AML from MDS-associated AML. It is also important to clarify whether de novo AML and MDS-associated AML are indeed distinct clinical entities or whether they overlap to some degree.

Although DNA microarray analysis is a promising tool for the identification of such molecular markers able to differentiate de novo AML from MDS-related AML, a simple comparison of BM MNCs for these 2 conditions is likely to be problematic. The cellular composition of BM MNCs differs markedly among individuals. Differences in the gene expression profile between BM MNCs from a given pair of individuals may thus reflect these differences in cell composition [16]. The elimination of such pseudopositive and pseudonegative data necessitates the purification of background-matched cell fractions from the clinical specimens before microarray analysis.

Given that de novo AML and MDS both result from the transformation of hematopoietic stem cell (HSC) clones, HSCs would be expected to be an appropriate target for purification and gene expression analysis. With the use of an affinity purification procedure based on the HSC-specific surface protein CD133, also known as AC133 [17], we have purified $CD133^+$ HSC-like fractions from individuals with

various hematopoietic disorders, and we have stored these fractions in a cell depository referred to as the *Blast Bank*. With such background-matched purified samples, we attempted to identify differences in gene expression profiles between de novo AML and MDS-associated AML [18]. To minimize further the influence of differentiation commitment of blasts toward certain lineages, we used only Blast Bank samples with the same phenotype, the M2 subtype according to the FAB classification. We thus characterized the expression profiles of >12,000 genes in $CD133^+$ Blast Bank samples from 10 patients with de novo AML of the M2 subtype as well as from 10 individuals with MDS-related AML of the same FAB subtype.

Selection with a Student *t* test ($P < .01$) and an effect size of ≥5 units for discrimination between the 2 clinical conditions led to the identification of 57 "diagnosis-associated genes," the expression profiles of which are shown in a "gene tree" format in Figure 1A. In this format, genes with similar expression patterns across the samples are clustered near each other. Patients were also clustered in this tree (2-way clustering) on the basis of the similarity of the expression pattern of the 57 genes (dendrogram at the top). All subjects were clustered into 2 major groups, 1 composed mostly of de novo AML patients and the other containing predominantly patients with MDS-associated AML. Each of these 2 main branches contained misclassified samples, however, indicating that simple clustering was not sufficiently powerful to differentiate the 2 clinical conditions completely. Furthermore, this analysis might not adequately address whether de novo AML and MDS-associated AML should be treated as distinct entities, at least from the point of view of the gene expression profile.

Decomposition of the multidimensionality of gene expression profiles by the application of principal components analysis [19] or correspondence analysis [20], for example, is often informative for such purposes. Application of the latter method to the data set of the 57 genes reduced the number of dimensions from 57 to 3. On the basis of the calculated 3-dimensional (3D) coordinates for each sample, the specimens were then projected into a virtual 3D space (Figure 1B). Most of the de novo AML samples were positioned in a region of this space that was distinct from that occupied by the MDS-associated AML specimens. However, 2 of the former samples were localized within the MDS region. These results suggest that de novo AML and MDS-associated AML are distinct disorders but that the current clinical diagnostic system is not efficient enough to separate them completely.

Instead of extracting a molecular signature from the expression profile of multiple genes, an alternative approach is to attempt to identify individual gene markers specific to either de novo AML or MDS-associated AML. We used the Blast Bank array data to identify MDS-specific markers, defined as genes that are silent in blasts from all de novo AML patients but are active in those from at least some patients with MDS-related AML [16]. This approach resulted in the identification of a single gene, *DLK*, that matched these criteria. *DLK* is expressed in immature cells [21] and is implicated in the maintenance of the undifferentiated state [22]. Selective expression of *DLK* in MDS blasts may thus contribute to the pathogenesis of MDS. Increased expression

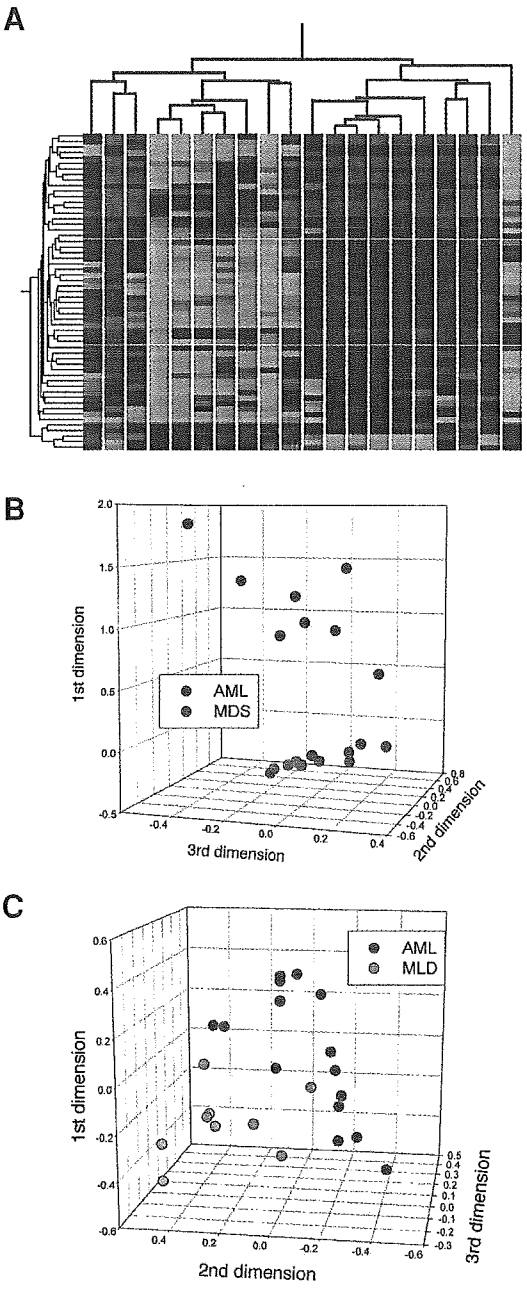

Figure 1. Correspondence analysis and 3-dimensional (3D) projection for differential diagnosis of acute myeloid leukemia (AML). A, Two-way hierarchical clustering for 57 disease-associated genes and 20 patients. Each row corresponds to a single gene and each column to $CD133^+$ cells from a patient with de novo AML (blue) or myelodysplastic syndrome (MDS)-related AML (red). The expression level of each gene is color coded, with a high level in red and a low level in green. B, Projection of the 20 specimens from (A) into a virtual space with 3 dimensions identified by correspondence analysis of the 57 genes. Patients with de novo AML (AML) were separated from those with MDS-related leukemia (MDS). C, Projection of specimens from patients with de novo AML without dysplasia (AML) or de novo AML with multilineage dysplasia (MLD) into a 3-dimensional space based on correspondence analysis of differences in gene expression. The dendrogram in (A) and 3-dimensional projections in (B) and (C) were constructed from data in [18] and [28].

of *DLK* in blasts from individuals with MDS has also been demonstrated in other studies [18,23].

Cell fractions other than CD133$^+$ cells are also potential targets for microarray analysis. Pellagatti et al chose peripheral blood neutrophils to investigate the gene expression profiles of MDS and MDS-associated AML [24]. They identified genes whose expression was dependent on MDS subtype. Given the high activity of ribonuclease in neutrophils, however, it is important to confirm the reproducibility of expression data obtained with these cells.

2.3. Dysplasia

In addition to the dysplastic blood cells associated with MDS, prominent dysplasia is apparent in certain individuals with de novo AML for whom the possibility of a prior MDS phase can be excluded [25,26]. Such de novo AML with dysplasia has a poor outcome with conventional chemotherapy [27], as does MDS-related leukemia. In the revised classification of AML by the World Health Organization [8], an entity of AML with multilineage dysplasia (AML-MLD) has been proposed; this entity probably includes both de novo AML with dysplasia and AML secondary to MDS. Whether such an amalgamation has clinical relevance awaits further studies.

To clarify directly whether de novo AML-MLD is indeed a clinical entity distinct from de novo AML without dysplasia, we searched for differences between the transcriptomes of CD133$^+$ cells derived from individuals with diagnoses of these 2 conditions [28]. We attempted to construct a 3D view of the samples with the coordinates calculated from correspondence analysis of genes found to be associated with dysplasia (Figure 1C). Most cases of AML-MLD were separated from those of AML without dysplasia in the 3D space. In contrast to the prominent separation power of the first dimension in Figure 1B, both the first and second dimensions substantially contributed to separation of the samples in Figure 1C. These data indicate that de novo AML without dysplasia can be differentiated from de novo AML-MLD on the basis of gene expression profiles.

3. Stratification of AML

3.1. Analysis of BM MNCs

Given that the current classification of AML is not sufficiently powerful to predict the prognosis of each affected individual, it is hoped that DNA microarray analysis will provide a better stratification scheme to separate AML patients into prognosis-dependent subgroups. To this end, Bullinger et al isolated MNCs from either BM or peripheral blood of 116 adult AML patients and examined the gene expression profiles of these cells with DNA microarrays harboring >39,000 cDNAs [29]. From the data set, they then screened "class predictor" genes, whose expression was correlated with the duration of patient survival. Class prediction based on such genes separated the patients into 2 classes, and long-term survival differed significantly between these 2 classes for both the training set ($P < .001$, log-rank test) and the test set ($P = .006$). This expression profile–based classification also separated AML patients with normal karyotype into 2 classes with distinct prognoses ($P = .046$). Although the number of genes used for the class prediction was relatively large (n = 133), these data demonstrated that DNA microarray analysis is able to predict the prognosis of AML patients in a manner independent of karyotype.

Similarly, Valk et al measured the expression levels of approximately 13,000 genes in BM MNCs isolated from 285 patients with AML [30]. Unsupervised clustering based on the gene expression profiles separated the patients into 16 subgroups. The prognoses of patients differed among these clusters, but whether the ability to predict prognosis was independent of karyotype was not addressed.

The gene expression profiles of BM MNCs from 54 pediatric AML patients were compared for those individuals who maintained complete remission (CR) for >3 years and those who failed to enter initial CR [14]. Thirty-five genes were identified as associated with prognosis and were used to separate the individuals into 2 groups. The difference in survival between the 2 groups was again statistically significant ($P = .03$), although whether this approach was independent of the current classification system was not addressed.

3.2. Analysis of Purified Fractions

As described in section 2.2, microarray analysis of purified fractions is likely to be more accurate than is that of BM MNCs for the extraction of prognosis-associated molecular signatures. We have analyzed the expression intensities of approximately 33,000 genes (likely representing almost the entire human genome) in CD133$^+$ HSC-like fractions isolated from 66 patients with AML who received standard chemotherapy. Of these patients, 51 individuals entered initial CR, whereas the remaining 15 failed to do so. Comparison of the data set for these 2 classes resulted in the identification of a small number of outcome-related genes (Yamashita et al, unpublished data). Principal components were extracted from the gene expression patterns, and the patients were projected into a virtual 3D space based on the coordinates obtained by correspondence analysis. Individuals who entered CR were separated from those who did not (Figure 2A), indicating that the gene expression profile of leukemic blasts indeed differs between AML patients who respond to chemotherapy and those who do not. In this analysis, the separation of the 2 groups of patients was achieved mostly in the first dimension. We therefore separated the patients into 2 subgroups according to whether the coordinate in the first dimension was <-0.3 or ≥ -0.3. The individuals in the latter group lived significantly longer than did those in the former (Figure 2B). These data support the feasibility of a novel stratification scheme for AML based on the gene expression profile.

4. Conclusion

Newly developed genomics tools allow global assessment of mRNA levels, DNA copy number, or genomic polymorphisms. Among these tools, DNA microarray analysis has proved highly successful in studies of human specimens, not only from individuals with hematopoietic disorders but also

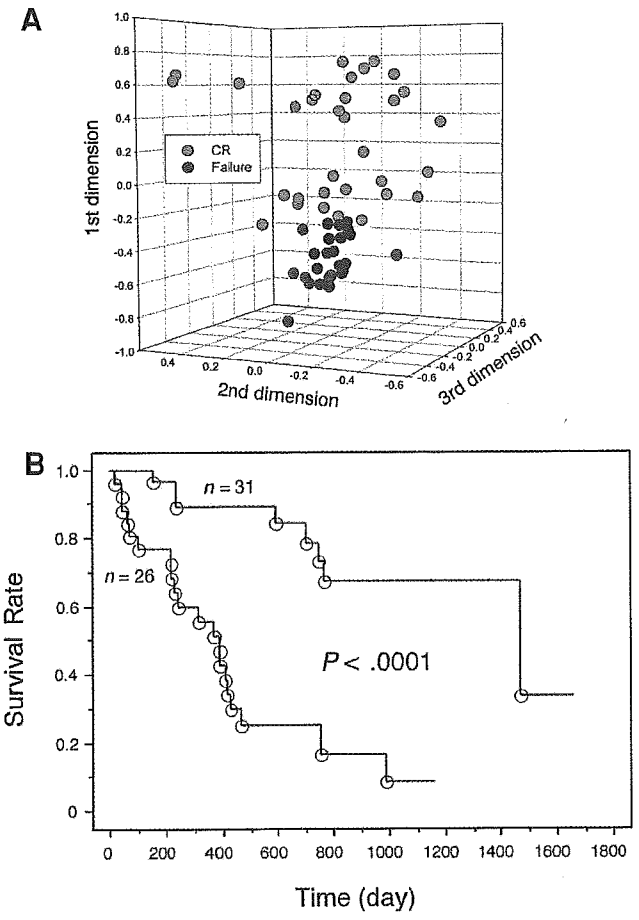

Figure 2. Stratification of acute myeloid leukemia (AML) on the basis of microarray analysis of gene expression. A, Projection of specimens from patients with AML who underwent initial chemotherapy-induced complete remission (CR) and those who did not (Failure) into a 3-dimensional space based on correspondence analysis of differences in gene expression. B, Kaplan-Meier analysis of the subjects in (A), revealing significantly different prognoses for the 2 classes with a coordinate in the first dimension of <–0.3 or ≥–0.3.

from those with solid tumors or nonmalignant degenerative diseases. Insight into many more human disorders is likely soon to be provided by such global profiling of gene expression. It is important to bear in mind, however, that DNA microarray data are prone to contamination with pseudopositive or pseudonegative results. The efficient characterization of AML thus appears to require optimal purification of target cell populations. Provided that experiments are designed carefully, DNA microarray analysis is likely to shed new light on the molecular pathogenesis of AML.

Acknowledgments

I thank colleagues in the Division of Functional Genomics, Jichi Medical School, for their enthusiasm in DNA microarray analyses, as well as the many physicians and patients who have participated in the Blast Bank project.

References

1. The Genome International Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature*. 2001;409: 860-921.
2. The *C. elegans* Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science*. 1998;282:2012-2018.
3. Cheung VG, Morley M, Aguilar F, Massimi A, Kucherlapati R, Childs G. Making and reading microarrays. *Nat Genet*. 1999;21: 15-19.
4. Khan J, Bittner ML, Saal LH, et al. cDNA microarrays detect activation of a myogenic transcription program by the *PAX3-FKHR* fusion gene. *Proc Natl Acad Sci U S A*. 1999;96:13264-13269.
5. Heller RA, Schena M, Chai A, et al. Discovery and analysis of inflammatory disease-related genes using cDNA microarrays. *Proc Natl Acad Sci U S A*. 1997;94:2150-2155.
6. Dhanasekaran SM, Barrette TR, Ghosh D, et al. Delineation of prognostic biomarkers in prostate cancer. *Nature*. 2001;412: 822-826.
7. Beer DG, Kardia SL, Huang CC, et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med*. 2002;8:816-824.
8. Jaffe ES, Harris NL, Vardman JW, eds. *Tumours of Haematopoietic and Lymphoid Tissues*. Lyon, France: IARC Press; 2001.
9. Grimwade D, Walker H, Oliver F, 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; 92:2322-2333.
10. Byrd JC, Mrozek K, Dodge RK, et al. 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*. 2002;100:4325-4336.
11. Smith SM, Le Beau MM, Huo D, et al. Clinical-cytogenetic associations in 306 patients with therapy-related myelodysplasia and myeloid leukemia: the University of Chicago series. *Blood*. 2003; 102:43-52.
12. Bain BJ. The bone marrow aspirate of healthy subjects. *Br J Haematol*. 1996;94:206-209.
13. Schoch C, Kohlmann A, Schnittger S, et al. Acute myeloid leukemias with reciprocal rearrangements can be distinguished by specific gene expression profiles. *Proc Natl Acad Sci U S A*. 2002;99: 10008-10013.
14. Yagi T, Morimoto A, Eguchi M, et al. Identification of a gene expression signature associated with pediatric AML prognosis. *Blood*. 2003;102:1849-1856.
15. Virtaneva K, Wright FA, Tanner SM, et al. Expression profiling reveals fundamental biological differences in acute myeloid leukemia with isolated trisomy 8 and normal cytogenetics. *Proc Natl Acad Sci U S A*. 2001;98:1124-1129.
16. Miyazato A, Ueno S, Ohmine K, et al. Identification of myelodysplastic syndrome-specific genes by DNA microarray analysis with purified hematopoietic stem cell fraction. *Blood*. 2001;98:422-427.
17. Hin AH, Miraglia S, Zanjani ED, et al. AC133, a novel marker for human hematopoietic stem and progenitor cells. *Blood*. 1997;90: 5002-5012.
18. Oshima Y, Ueda M, Yamashita Y, et al. DNA microarray analysis of hematopoietic stem cell-like fractions from individuals with the M2 subtype of acute myeloid leukemia. *Leukemia*. 2003;17: 1990-1997.
19. Lefkovits I, Kuhn L, Valiron O, Merle A, Kettman J. Toward an objective classification of cells in the immune system. *Proc Natl Acad Sci U S A*. 1988;85:3565-3569.
20. Fellenberg K, Hauser NC, Brors B, Neutzner A, Hoheisel JD, Vingron M. Correspondence analysis applied to microarray data. *Proc Natl Acad Sci U S A*. 2001;98:10781-10786.
21. Moore KA, Pytowski B, Witte L, Hicklin D, Lemischka IR.

Hematopoietic activity of a stromal cell transmembrane protein containing epidermal growth factor-like repeat motifs. *Proc Natl Acad Sci U S A*. 1997;94:4011-4016.
22. Smas CM, Sul HS. Pref-1, a protein containing EGF-like repeats, inhibits adipocyte differentiation. *Cell*. 1993;73:725-734.
23. Hofmann WK, de Vos S, Komor M, Hoelzer D, Wachsman W, Koeffler HP. Characterization of gene expression of $CD34^+$ cells from normal and myelodysplastic bone marrow. *Blood*. 2002;100: 3553-3560.
24. Pellagatti A, Esoof N, Watkins F, et al. Gene expression profiling in the myelodysplastic syndromes using cDNA microarray technology. *Br J Haematol*. 2004;125:576-583.
25. Brito-Babapulle F, Catovsky D, Galton DAG. Clinical and laboratory features of de novo acute myeloid leukemia with trilineage myelodysplasia. *Br J Haematol*. 1987;66:445-450.
26. Jinnai I, Tomonaga M, Kuriyama K, et al. Dysmegakaryocytopoiesis in acute leukaemias: its predominance in myelomonocytic (M4) leukaemia and implication for poor response to chemotherapy. *Br J Haematol*. 1987;66:467-472.
27. Kuriyama K, Tomonaga M, Matsuo T, et al. Poor response to intensive chemotherapy in de novo acute myeloid leukaemia with trilineage myelodysplasia: Japan Adult Leukaemia Study Group (JALSG). *Br J Haematol*. 1994;86:767-773.
28. Tsutsumi C, Ueda M, Miyazaki Y, et al. DNA microarray analysis of dysplastic morphology associated with acute myeloid leukemia. *Exp Hematol*. 2004;32:828-835.
29. Bullinger L, Dohner K, Bair E, et al. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *N Engl J Med*. 2004;350:1605-1616.
30. Valk PJ, Verhaak RG, Beijen MA, et al. Prognostically useful gene-expression profiles in acute myeloid leukemia. *N Engl J Med*. 2004; 350:1617-1628.

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 280, No. 13, Issue of April 1, pp. 12621–12629, 2005
Printed in U.S.A.

Signal Transducers and Activators of Transcription 3 Augments the Transcriptional Activity of CCAAT/Enhancer-binding Protein α in Granulocyte Colony-stimulating Factor Signaling Pathway*

Received for publication, July 26, 2004, and in revised form, January 3, 2005
Published, JBC Papers in Press, January 21, 2005, DOI 10.1074/jbc.M408442200

Akihiko Numata‡, Kazuya Shimoda‡§, Kenjiro Kamezaki‡, Takashi Haro‡, Haruko Kakumitsu‡, Koutarou Shide‡, Kouji Kato‡, Toshihiro Miyamoto‡, Yoshihiro Yamashita¶, Yasuo Oshima‖, Hideaki Nakajima, Atsushi Iwama‡‡, Kenichi Aoki‡, Ken Takase‡, Hisashi Gondo‡, Hiroyuki Mano¶, and Mine Harada‡**

*From the ‡Medicine and Biosystemic Science, Kyushu University Graduate School of Medical Sciences, 3-1-1 Maidashi, Higashi-ku, Fukuoka, Fukuoka, 812-8582, the ¶Division of Functional Genomics, Jichi Medical School, 3311-1 Yakushiji, Kawaguchi-gun, Tochigi, 329-0498, the ‖Department of Clinical Pharmacology, Jichi Medical School, 3311-1 Yakushiji, Kawaguchi-gun, Tochigi, 329-0498, the **Department of Hematopoietic Factors, Institute of Medical Science, University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo, 108-8639, and the ‡‡Laboratory of Stem Cell Therapy, Center for Experimental Medicine, Institute of Medical Science, University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo, 108-8639, Japan*

The Janus kinase (Jak)-Stat pathway plays an essential role in cytokine signaling. Granulocyte colony-stimulating factor (G-CSF) promotes granulopoiesis and granulocytic differentiation, and Stat3 is the principle Stat protein activated by G-CSF. Upon treatment with G-CSF, the interleukin-3-dependent cell line 32D clone 3(32Dcl3) differentiates into neutrophils, and 32Dcl3 cells expressing dominant-negative Stat3 (32Dcl3/DNStat3) proliferate in G-CSF without differentiation. Gene expression profile and quantitative PCR analysis of G-CSF-stimulated cell lines revealed that the expression of *C/EBPα* was up-regulated by the activation of Stat3. In addition, activated Stat3 bound to CCAAT/enhancer-binding protein (C/EBP)α, leading to the enhancement of the transcription activity of C/EBPα. Conditional expression of C/EBPα in 32Dcl3/DNStat3 cells after G-CSF stimulation abolishes the G-CSF-dependent cell proliferation and induces granulocytic differentiation. Although granulocyte-specific genes, such as the G-CSF receptor, lysozyme M, and neutrophil gelatinase-associated lipocalin precursor (*NGAL*) are regulated by Stat3, only NGAL was induced by the restoration of C/EBPα after stimulation with G-CSF in 32Dcl3/DNStat3 cells. These results show that one of the major roles of Stat3 in the G-CSF signaling pathway is to augment the function of C/EBPα, which is essential for myeloid differentiation. Additionally, cooperation of C/EBPα with other Stat3-activated proteins are required for the induction of some G-CSF responsive genes including lysozyme M and the G-CSF receptor.

The proliferation and differentiation of hematopoietic progenitor cells are regulated by cytokines (1). Among these, granulocyte colony-stimulating factor (G-CSF)[1] specifically stimulates cells that are committed to the myeloid lineage (2). The importance of G-CSF to the regulation of granulopoiesis has been confirmed by the observation of severe neutropenia in mice carrying homozygous deletions of their G-CSF or G-CSF receptor genes (3, 4). Cytokines activate several intracellular signaling pathways, and the Janus kinase (Jak) signal transducers and activators of transcription (Stat) pathway is essential for cytokine function (5, 6). The binding of G-CSF to cell surface G-CSF receptors activates Jak1, Jak2, and Tyk2 followed by the activation of Stat1, Stat3, and Stat5 (7–9). Stat3 is the principle protein activated by G-CSF (8, 10). Phosphorylated Stats translocate from the cytoplasm into the nucleus and induce transcription of their target genes within a short period of time. 32Dcl3 cells differentiate to neutrophils following treatment with G-CSF. In contrast to their parental cells, 32Dcl3 cells expressing dominant-negative Stat3 (32Dcl3/DNStat3) proliferate in the presence of G-CSF, but they maintain immature morphologic characteristics without evidence of differentiation (11). Additionally, transgenic mice with a targeted mutation of their G-CSF receptor that abolishes G-CSF-dependent Stat3 activation show severe neutropenia with an accumulation of immature myeloid precursors in their bone marrows (12). To clarify the role of Stat3 in the G-CSF signaling pathway, we wished to identify target genes of Stat3.

We found that the levels of CCAAT/enhancer-binding protein (C/EBP) α mRNA were up-regulated following G-CSF stimulation in 32Dcl3 but were unchanged in 32Dcl3/DNStat3. In addition, the activation of Stat3 augmented the function of C/EBPα, which is the essential transcriptional factor for myeloid differentiation. G-CSF-induced granulocytic differentiation was restored in 32Dcl3/DNStat3 cells by the conditional expression of C/EBPα. These results show that one of the major

* This work was supported by a grant of the Japan Leukemia Research Foundation (2002) and Grants-in-aid for Scientific Research 11307015 and 15390302 from the Ministry of Education, Science, Sports, and Culture in Japan. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed. Tel.: 81-92-642-5230; Fax: 81-92-642-5247; E-mail: kshimoda@intmed1.med.kyushu-u.ac.jp.

[1] The abbreviations used are: G-CSF, granulocyte colony-stimulating factor; IL, interleukin; C/EBPα, CCAAT/enhancer-binding protein; NGAL, neutrophil gelatinase-associated lipocalin precursor; Jak, Janus kinase; Stat, signal transducers and activators of transcription; DNStat3, dominant-negative Stat3; IRES, internal ribosome entry site; GFP, green fluorescent protein; ER, endoplasmic reticulum; IFN, interferon; PBS, phosphate-buffered saline; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; FACS, fluorescence-activated cell sorter; LUC, luciferase; ERK, extracellular signal-regulated kinase; MAP, mitogen-activated protein; TK, thymidine kinase; 4-HT, 4-hydroxytamoxifen.

roles of Stat3 in the G-CSF signaling pathway is to enhance the function of C/EBPα.

MATERIALS AND METHODS

Cell Culture, Expression Plasmid, and Cytokines—32D clone 3 (32Dcl3) and 32Dcl3/DNStat3 cells (DNStat3 deletes the transactivation domain of Stat3) were cultured in RPMI 1640 supplemented with 10% heat-inactivated fetal bovine serum (ICN, Osaka, Japan), penicillin/streptomycin (Invitrogen), recombinant murine interleukin-3 (IL-3) (Kirin Brewery, Takasaki, Japan), and recombinant human G-CSF (Chugai Pharmaceutical, Tokyo, Japan). 293T cells were cultured in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum, penicillin/streptomycin, and L-glutamine.

For the construction of pTag2A-G-CSF receptor, the human G-CSF receptor cDNA (13) (pHQ3, kindly provided by S. Nagata and R. Fukunaga) was excited from the pBluescipt vector and inserted into the FLAG-tagged mammalian expression plasmid pCMV-Tag2A (Clontech). pcDNA3-rat C/EBPα was described before (14). Stat3 cDNA was amplified by PCR and inserted into pCMV-HA vector (Clontech). Stat3c cDNA was elicited from RCMV-Stat3c (15), kindly given from Dr. Darnell, and inserted into pcDNA3.1 (Clontech). For the construction of pMY-IRES-GFP/C/EBPα-ER, full-length human C/EBPα cDNA was fused in-frame with ligand-binding domain (amino acids 281–599) of mouse estrogen receptor harboring a mutation (G525R) that confers selective responsiveness to 4-hydroxytamoxifen (4-HT). A reporter construct of a minimal TK promoter with CEBP-binding sites (p(C/EBP)2TK) was described previously (14).

Murine recombinant leukemia inhibitory factor, natural IFN-α, and recombinant IFN-γ were purchased from Sigma, HyCult Biotechnology (Uden, The Netherlands), and Peprotech (Rocky Hill, NJ), respectively. For Western blotting, 32Dcl3 cells or 32Dcl3/DNStat3 cells were deprived of IL-3 for 12 h. Then cells were stimulated with G-CSF (10 ng/ml), IL-3 (10 ng/ml), leukemia inhibitory factor (10 ng/ml), IFN-α (1,000 units/ml), or IFN-γ (1,000 units/ml) for 30 min.

Microarray Analysis—32D cl3 and 32Dcl3/DNStat3 cells maintained in IL-3 were washed twice with PBS and starved of cytokine in RPMI 1640 containing 10% fetal bovine serum for 8 h and then stimulated with 10 ng/ml G-CSF. Total RNA was extracted, by the acid guanidinium method, from 32Dcl3 and 32Dcl3/DNStat3 cells before or after the stimulation for 2 h with G-CSF. Double-stranded cDNA synthesized from the total RNA (20 μg/sample) was then used to prepare biotin-labeled cRNA for the hybridization with GeneChip MGU74Avs2 microarrays (Affymetrix, Santa Clara, CA) harboring oligonucleotides corresponding to ~6000 known genes as well as ~6000 expressed sequence tag sequences. Hybridization, washing, and detection of signals on the arrays were performed with the GeneChip system (Affymetrix).

Quantitative Real-time Reverse Transcription-PCR Assay—32Dcl3 and 32Dcl3/DNStat3 cells maintained in IL-3, were washed twice with PBS and starved of cytokines for 8 h and then stimulated with 10 ng/ml G-CSF. Cells were harvested at the indicated times, and total RNA was isolated using Isogen (Nippon gene, Tokyo, Japan) according to the manufacturer's instructions. One microgram of extracted RNA was transcribed in a 20-μl cDNA synthesis reaction using an RNA PCR kit (AMV) (Takara, Tokyo, Japan). Real-time PCR for C/EBPα, G-CSF receptor, lysozyme M, neutrophil gelatinase-associated lipocalin precursor (NGAL), and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was performed by a TaqMan assay on an ABI 7000 system. PCR primers and probes were designed as follows: murine C/EBPα, sense, 5′-CCA TGT GGT AGG AGA CAG AGA CCT A-3′, and antisense, 5′-CTC TGG GAT GGA TCG ATT GTG-3′; probe FAM-5′-CGG CTG GCG ACA TAC AGT ACA CAC AAG-3′-TAMRA, and sense, 5′-CCA AGA AGT CGG TGG ACA AGA-3′, and antisense, 5′-CGG TCA TTG TCA CTG GTC AAC T-3′; probe FAM-5′-AGC ACC TTC TGT TGC GTC TCC ACG TT-3′-TAMRA; murine G-CSF receptor, sense, 5′-CTA AAC ATC TCC CTC CAT GAC TT-3′, and antisense, 5′-GGC CAT GAG GTA GAC ATG ATA CAA-3′; probe FAM-5′-CAT CTT CTC TGT CCC CAC CGA CCA A-3′-TAMRA; murine lysozyme M, sense, 5′-TGC CTG TGG GAT CAA TTG C-3′, and antisense, 5′-ATG CCA CCC ATG CTC GAA T-3′; probe 5′-FAM-CAG TGA TGT CAT CCT GCA GAC CA-TAMRA-3′; murine NGAL, sense, 5′-GGC CTC AAG GAC GAC AAC A-3′, and antisense, 5′-CAC CAC CCA TTC AGT TGT CAA T-3′; probe 5′-FAM-CAT CTT CTC TGT CCC CAC CGA CCA A-TAMRA-3′, and murine GAPDH sense, 5′-ACG GCA AAT TCA ACG GCA-3′, and antisense, 5′-AGA TGG TGA TGG GCT TCC-3′; probe 5′-FAM-AGG CCG AGA ATG GGA AGC TTG TCA TC-TAMRA-3′. PCR amplifications were performed in a 50-μl volume, containing 4 μl of cDNA template, 50 mM KCl, 10 mM Tris-HCl(pH 8.3), 10 mM EDTA, 60 nM, 200 μM dNTPs, 3 mM $MgCl_2$, 200 nM each primer, 0.625 units of AmpliTaqGold, and 0.25 units of AmpErase uracil *N*-glycosylase. Each amplification reaction also contained 100 nM appropriate detection probe. Each PCR amplification was performed in duplicate, using conditions of 50 °C for 2 min preceding 95 °C for 10 min followed by 40 cycles of amplification (95 °C for 15 s, 60 °C for 1 min). In each reaction, GAPDH was amplified as a housekeeping gene to calculate a standard curve and allow for the correction for variations in target sample quantities. Relative copy numbers were calculated for each sample from the standard curve after normalization to GAPDH by the instrument software.

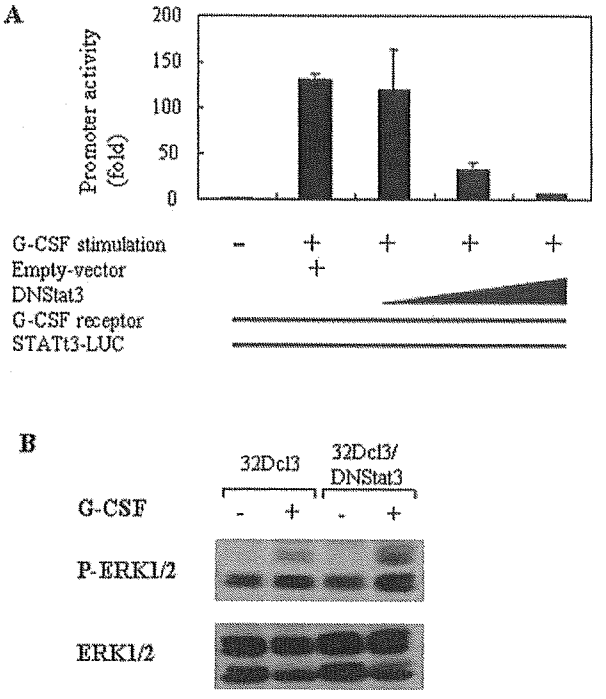

FIG. 1. **The effect of dominant-negative Stat3 on G-CSF signaling pathway.** *A*, transient transfection in 293T cells with a reporter construct with α2-macroglobulin promoter (*STAT3-LUC*), dominant-negative Stat3, and G-CSF receptor. Twelve hours after transfection, cells were stimulated with 10 ng/ml G-CSF. Promoter activity was measured as luciferase activity 36 h after transfection. The *vertical axis number* is the fold induction when compared with control. *B*, 32Dcl3 cells or 32Dcl3/DNStat3 cells were cultured with IL-3 and then deprived of IL-3 for 12 h. Cells were treated with the G-CSF for 30 min and lysed. Post-nuclear supernatants were resolved by 10% SDS-PAGE and transferred to nitrocellulose membranes. Membranes were probed using the indicated antibodies. *p-ERK1/2*, phosphorylated ERK1/2.

Conditional C/EBPα Expression—pMY-IRES-GFP/C/EBPα-ER was transfected into 32Dcl3 and 32Dcl3/DNStat3 cells by electroporation. 5×10^6 cells were transfected with 20 μg of expression vector, and GFP-positive cells were sorted by FACS Vantage (BD Biosciences). Expression of C/EBPα was determined by Western blotting analysis (see below).

Luciferase Assay—293T cells were transfected by the calcium phosphate precipitation method in 6-well plates, and luciferase activity was assayed using a luminometer Lumat LB9507 (Berthold Technologies, Bad Wildbad, Germany) according to the manufacturer's protocol. Each expression plasmid amount was 50–100 ng/well, and the same amount of empty expression vector was used as control, respectively. Results of reporter assays represent the average values for relative luciferase activity generated from five independent experiments.

Flow Cytometry—1×10^7 cells were incubated with 5 μl of recombinant phosphatidylethanolamine-conjugated rabbit anti-murine Gr1 monoclonal antibody (BD Biosciences) for 30 min at 4 °C, washed twice in PBS, and analyzed on a FACS Calibur (BD Biosciences).

Immunoprecipitation and Immunoblotting—Cells were lysed with lysis buffer, and lysates were immunoprecipitated with anti C/EBPβ (Santa Cruz Biotechnology, Santa Cruz, CA) as described previously (8). Total cell lysates or the immunoprecipitates were resolved by 10% SDS-PAGE and transferred to a nitrocellulose membrane. Membranes

TABLE I
Microarray analysis

32Dcl3 and 32Dcl3/DNStat3 cells were starved of cytokines for 8 h and then stimulated or left unstimulated with 10 ng/ml G-CSF. Total RNA was extracted from each fraction and was subjected to the hybridization with high-density oligonucleotide microarrays (MGU74Av2). Fold induction means a rate of increase in gene expression level by G-CSF stimulation. Candidate genes were identified as transcripts that were up-regulated in 32Dcl3 cells and down-regulated or unchanged in 32Dcl3/DNStat3 cells after G-CSF stimulation.

Gene product name	Abbreviation	Accession number	Fold induction	
			32Dcl3	32Dcl3/DNStat3
B-cell leukemia/lymphoma α	Bcl2	L31532	35.6	0.0629
CyclinE1	*Ccne1*	NM007633	29.7	0.690
Serotonin-gated ion channel	*5HT3*	M74425	27.2	0.592
KIF3B protein	*kif3b*	D26077	21.5	0.921
Protein kinase, serine/arginine-specific 1	*Srpk1*	AB012290	18.7	0.321
MAP kinase-interacting serine/threonine kinase 1	*Mknk1*	Y11091	15.7	0.845
Protein tyrosine phosphatase	*Ptpn13*	D83966	12.4	0.964
Transferrin receptor	*Trfr*	X57349	10.6	0.964
Lymphocyte antigen 57	*Ly57*	AF068182	9.62	0.968
Macrophage stimulating 1 receptor	*Mst1r*	X74736	8.83	0.762
Mitogen-activated protein kinase kinase 7	*MKK7*	AB005654	8.14	0.980
RAR-related orphan receptor alpha	*Rora*	U53228	7.94	0.861
Hemoglobin Y, β-like embryonic chain	*Hbb-y*	V00726	7.38	0.375
Runt related transcription factor 1	*Runx1*	NM009821	7.01	0.226
Microtubule-associated protein 6	*Mtap6*	Y14754	5.06	0.885
CCAAT/enhancer binding protein α	*C/EBPα*	M62362	2.05	0.840
Ecotoropic viral integration site 1	Evi1	M21829	1.55	0.239
Integrin alpha L	*Itgal*	M60778	1.35	0.567
Ninjurin 1	*Ninj1*	U91513	1.34	0.783
Interleukin 17 receptor	*IL17R*	U31993	1.24	0.449
Mucosal addressin	*MAdCAM*	D50434	1.14	0.527
Carbon catabolite repression 4 homolog	*Ccr4*	X16670	1.06	0.0768
Friend leukemia integration 1	*Fli1*	X59421	1.01	0.305

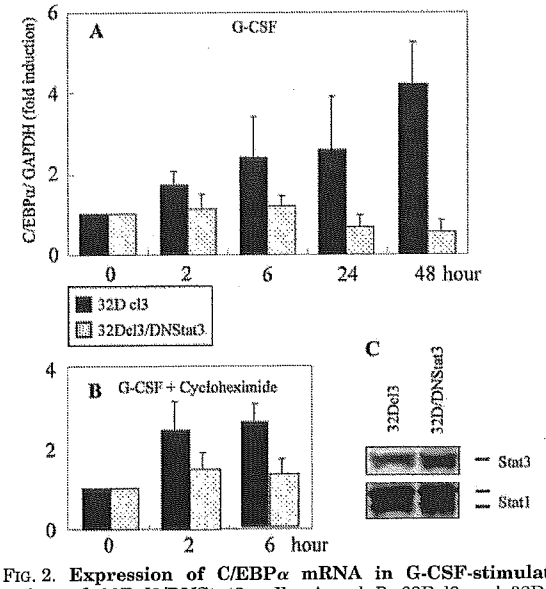

FIG. 2. **Expression of C/EBPα mRNA in G-CSF-stimulated 32Dcl3 and 32Dcl3/DNStat3 cells.** *A* and *B*, 32Dcl3 and 32Dcl3/DNStat3 cells maintained in IL-3 were washed twice with PBS and starved of cytokines for 8 h and stimulated with 10 ng/ml G-CSF (*A*) or 10 ng/ml G-CSF and 10 μg/ml cycloheximide (*B*). Total RNA was isolated from both cell lines at the indicated times and transcribed to cDNA, which was subjected to real-time PCR for murine C/EBPα. The *numbers* given on the *vertical axis* represent the fold induction of the ratios of GAPDH-normalized expression values when compared with those before G-CSF stimulation. Results are expressed as mean fold of two independent experiments.

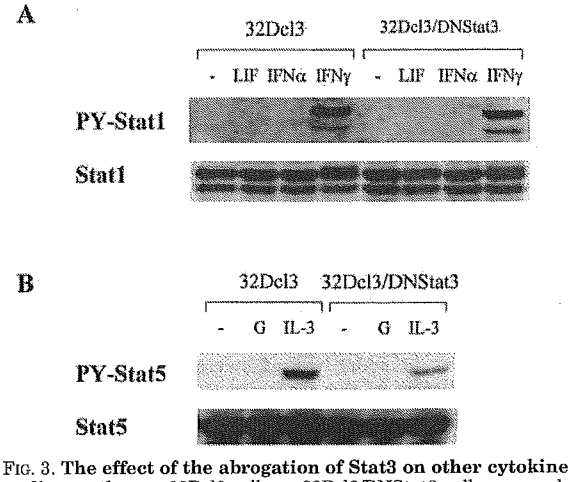

FIG. 3. **The effect of the abrogation of Stat3 on other cytokine signaling pathway.** 32Dcl3 cells or 32Dcl3/DNStat3 cells were cultured with IL-3 and then deprived of IL-3 for 12 h. Cells were treated with the indicated cytokines for 30 min and lysed. Post-nuclear supernatants were resolved by 10% SDS-PAGE and transferred to nitrocellulose membranes. Membranes were probed using the indicated antibodies. *LIF*, leukemia inhibitory factor.

were probed using the indicated antibodies followed by an IgG-horseradish peroxidase-conjugated secondary antibody (Amersham Biosciences) and visualized with the ECL detection system (Amersham Biosciences). Anti-phospho-ERK1/2 antibodies were purchased from Cell Signaling (Beverly, MA). Anti-phospho-Stat1 and -Stat5 antibodies were obtained from New England Biolabs (Beverly, MA), and anti-Stat1, -Stat3, and -C/EBPα antibodies were purchased from Santa Cruz Biotechnology. Membranes were probed using and visualizes with the ECL detection system (Amersham Biosciences).

Proliferation Assay—32D cl3 and 32Dcl3/DNStat3 cells maintained in IL-3 were washed twice with PBS and starved of cytokine for 8 h and then stimulated with 10 ng/ml G-CSF. The number of viable cells was determined by trypan blue dye exclusion using a hemocytometer. [^{3}H]Thymidine incorporation assays were also performed. Briefly, cells (1×10^5) in 100 μl of medium stimulated with murine IL-3 (1.0 ng/ml) or recombinant human G-CSF (10 ng/ml) were cultured for 48 h. During the final 4 h, [^{3}H]thymidine (1μCi/well) was added. Cells were then harvested by filtration, and radioactivity was counted by scintillation spectrophotometer.

RESULTS

G-CSF-induced Intracellular Signal Response in 32Dcl3/DNStat3 Cells—32Dcl3 cells differentiate into neutrophils following treatment with G-CSF, but 32Dcl3 cells expressing a

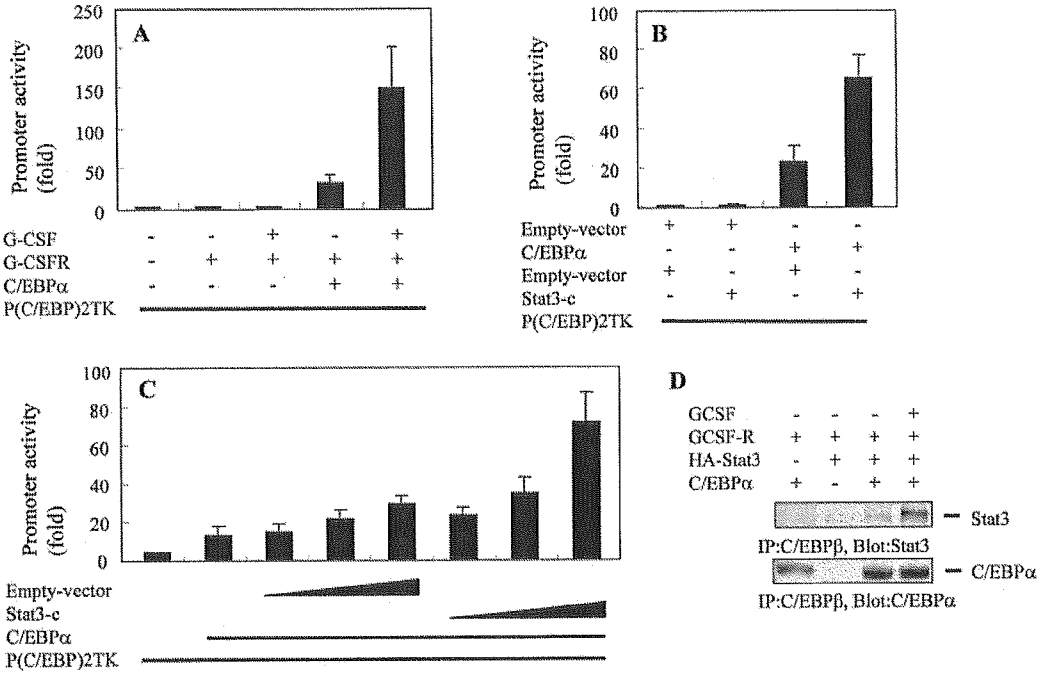

FIG. 4. **Activated Stat3 makes complex with C/EBPα, leading to the enhancement of C/EBPα-induced transcription.** *A*, transient transfection in 293T cells with a reporter construct of a minimal TK promoter with CEBP-binding sites (*p(C/EBP)2TK*), C/EBPα, and G-CSF receptor (*G-CSFR*). Twelve hours after transfection, cells were stimulated with 10 ng/ml G-CSF. Promoter activity was measured as luciferase activity 36 h after transfection. The *vertical axis number* is the fold induction when compared with control. *B* and *C*, transient transfection in 293T cells with a reporter construct of a minimal TK promoter with CEBP-binding sites (*p(C/EBP)2TK*), C/EBPα, Stat3c, and control vectors. Promoter activity was measured as luciferase activity 24 h after transfection. The *vertical axis number* is the fold induction when compared with control. *D*, transient transfection in 293T cells with a construct of G-CSF receptor, HA-Stat3, and C/EBPα and control vectors. After 24 h, cells were lysed and immunoprecipitated (*IP*) with anti C/EBPβ. Cells were stimulated with G-CSF during the final 9 h in the culture. The immunoprecipitates were resolved by 10% SDS-PAGE and transferred to a nitrocellulose membrane. Stat3 was detected by immunoblotting.

dominant-negative Stat3 (32Dcl3/DNStat3) proliferate following G-CSF treatment. These cells maintain immature morphologic characteristics without evidence of differentiation (11). First, we examined the effect of dominant-negative Stat3, carboxyl-truncated Stat3 that lacked 55 amino acids including the transactivation domain. We transfected reporter construct of STATt3-LUC, in which the α2-macroglobulin promoter (16) drives expression of the luciferase (LUC) reporter gene and G-CSF receptor, together with empty vector (pcDNA3) or DNStat3 to 293T cells. After 12 h of transfection, cells were stimulated with 10 ng/ml G-CSF. Cells were cultured for more 24 h, and luciferase assay was performed. As shown in Fig. 1*A*, G-CSF induced the transcriptional activity of Stat3 by 150-fold, and DNStat3 inhibited this G-CSF-induced Stat3 activation in a dose-dependent manner.

G-CSF mainly induces the phosphorylation of Stat3, but it also phosphorylates Stat1 and Stat5 in some cells among the Stats family (8) and induces the activation of MAP kinases. In both 32Dcl3 cells and 32Dcl3/DNStat3 cells, neither Stat1 nor Stat5 was phosphorylated in response to G-CSF (data not shown). As for the MAP kinase activation, the degree of the phosphorylation of ERK1/2 by G-CSF stimulation in 32Dcl3/DNStat3 cells was stronger than that in 32Dcl3 cells (Fig. 1*B*).

Identification of Genes Regulated by Stat3 in the G-CSF Signaling Pathway by Oligonucleotide Array Analysis—To identify Stat3-regulated genes involved in granulocytic differentiation, we compared gene expression change in both cell lines using microarray analysis. 32D cl3 and 32Dcl3/DNStat3 cells maintained in IL-3 were washed twice with PBS and starved in RPMI 1640 containing 10% fetal bovine serum lacking cytokine for 8 h and then stimulated with 10 ng/ml G-CSF. Total RNA was isolated from 32Dcl3 cells and 32Dcl3/DNStat3 cells treated with G-CSF after 0 and 2 h, transcribed to biotin-labeled cRNA, and hybridized to GeneChip MGU74Av2 arrays to compare the expression profile of ~12,000 murine genes. The fold induction in the expression level of each gene was calculated as the ratio of GAPDH-normalized fluorescence intensity value of G-CSF-stimulated cells when compared with those before G-CSF stimulation. As shown in Table I, we could identify a set of candidate genes for Stat3 targets, expression of which was up-regulated in 32D cl3 cells but down-regulated or unchanged in 32Dcl3/DNStat3 cells. Such Stat3-dependent expression profiles were confirmed in triplicate experiments.

C/EBPα Is a Target Gene for Stat3 in G-CSF Signaling Pathway—Among the identified genes, it was decided to focus further efforts on *C/EBPα*. C/EBPα has been shown to be critical for early granulocytic differentiation (17–19), and the factors regulating its activity are unclear. The expression of C/EBPα was examined by real-time quantitative reverse transcription-PCR. C/EBPα mRNA levels are rapidly up-regulated in 32Dcl3 cells, being elevated 2.39-fold after 6 h and 4.20-fold after 48 h (Fig. 2*A*). In contrast to 32Dcl3 cells, the C/EBPα mRNA levels were not changed in 32Dcl3/DNStat3 cells after G-CSF stimulation (Fig. 2*A*). A similar expression pattern was seen in separate experiments with independently designed primers and probes (data not shown). Levels of C/EBPα mRNA were unaffected by cycloheximide treatment (Fig. 2*B*). The expression level of the sum of Stat3 plus dominant-negative Stat3 in 32Dcl3/DNStat3 cells is a little larger than that of Stat3 in 32Dcl3 cells (Fig. 2*C*).

Activated Stat3 Binds to C/EBPα and Enhances the Transcription Activity of C/EBPα—We next examined the effect of

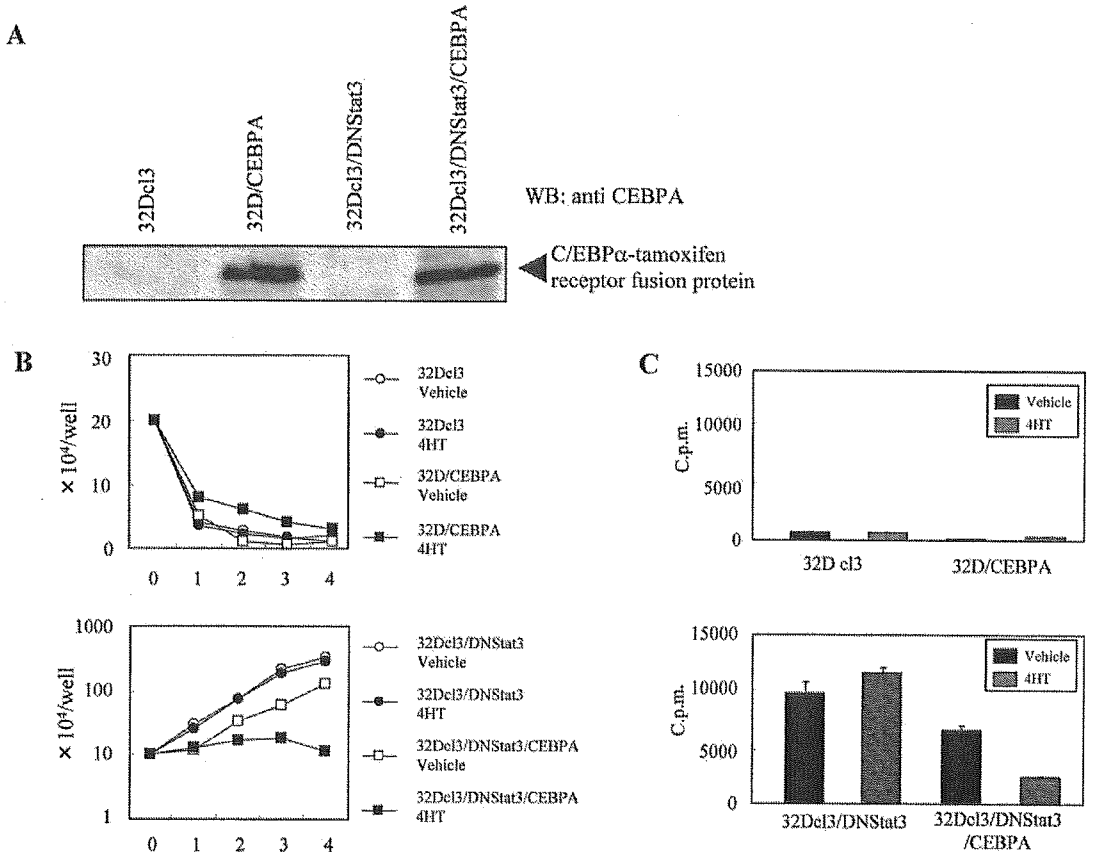

FIG. 5. **Proliferation of 32Dcl3 and 32Dcl3/DNStat3 by restoration of C/EBPα.** *A*, the expression vector pMY-IRES-GFP/C/EBPα-ER was transfected into 32Dcl3 and 32Dcl3/DNStat3 cells. The expression of C/EBPα-ER was examined by Western blotting (*WB*) using anti-C/EBPα polyclonal antiserum. *Lane 1*, 32Dcl3; *lane 2*, 32D/CEBPA; *lane 3*, 32Dcl3/DNStat3; *lane 4*, 32Dcl3/DNStat3/CEBPA. *B*, growth curve of 32Dcl3, 32Dcl3/CEBPA cells (*upper panel*), and 32Dcl3/DNStat3, 32Dcl3/DNStat3/CEBPA cells (*lower panel*). Cells maintained in IL-3 were washed twice with PBS and starved of cytokines for 8 h and stimulated with 10 ng/ml G-CSF plus 0.5 μM 4-HT or vehicle. Viable cells were counted daily by trypan blue dye exclusion method at the indicated times. The *numbers* given on the *vertical axis* represent the mean cell counts (×10⁴/well) of triplicate wells. Standard deviations (S.D.) were less than 15% of each mean. Three independent experiments were performed, and similar results were obtained. Data shown are representative of these results. *C*, ^{3}H incorporation assays in 32Dcl3, 32Dcl3/CEBPA (*upper panel*) and 32Dcl3/DNStat3 and 32Dcl3/DNStat3/CEBPA cells (*lower panel*). Cells maintained in IL-3 were washed twice with PBS and starved of cytokines for 8 h and stimulated with 10 ng/ml G-CSF plus 0.5 μM 4-HT or vehicle for 48 h. During the final 4 h, 1 μCi of [^{3}H]thymidine was added, cells were harvested by filtration, and radioactivity was counted by scintillation spectrophotometer. Results are expressed as mean cpm of triplicate wells ± S.D. Three independent experiments were performed, and similar results were obtained. Data shown are representative of these results.

Stat3 abrogation on the balance of intracellular signals in other cytokine pathways. Although Stat1 was not phosphorylated by leukemia inhibitory factor stimulation in neither 32Dcl3 cells nor 32Dcl3/DNStat3 cells, its activation in response to IFN-γ occurred at the same degree in both 32cl3 cells and 32Dcl3/DNStat3 cells (Fig. 3*A*). As for the Stat5 activation, the phosphorylation of Stat5 by IL-3 stimulation in 32Dcl3 cells was stronger than that in 32Dcl3/DNStat3 cells (Fig. 3*B*). These data indicated that there was the possibility that abrogation of Stat3 signaling can alter the balance of intracellular signals in other cytokine signaling pathways. The transcription of C/EBPα is regulated by C/EBPα itself (20, 21). Then we examined whether activated Stat3 in G-CSF signaling enhance C/EBPα activity or not.

We transfected a reporter construct of a minimal TK promoter with CEBP-binding sites (p(C/EBP)2TK), C/EBPα, and G-CSF receptor to 293T cells. After 12 h of transfection, cells were stimulated with 10 ng/ml G-CSF. Cells were cultured for more 24 h, and a luciferase assay was performed. C/EBPα up-regulated the C/EBPα-dependent gene expression, and the G-CSF stimulation enhanced this C/EBPα-dependent gene expression (Fig. 4*A*). Next we examined the effect of constitutive active Stat3 (Stat3C) on the augmentation of C/EBPα transcriptional activity instead of the G-CSF stimulation. We transfected reporter construct p(C/EBP)2TK, C/EBPα, and Stat3C to 293T cells. After 24 h of transfection, luciferase assay was performed. Stat3C augmented the C/EBPα-dependent gene expression, although Stat3C alone had no influence on the luciferase activity (Fig. 4, *B* and *C*).

As p(C/EBP)2TK contains only a C/EBPα-binding site and does not contain a Stat3-binding sequence, the possibility that Stat3C makes a complex with C/EBPα and augments the function of C/EBPα is raised. Then we transfected C/EBPα, Stat3, and G-CSF receptor to 293T cells and stimulated cells with G-CSF for 6 h. There is no detectable level of endogenous C/EBPα or C/EBPβ protein in 293T cells. Cells were lysed and immunoprecipitated with C/EBPβ antibody (this antibody cross-reacts with C/EBPα). As shown in Fig. 4*D*, immunoprecipitants with anti-C/EBPβ contain Stat3. In addition, the complex formation between C/EBPα and Stat3 is augmented by G-CSF stimulation, indicating that activated Stat3 makes the complex with C/EBPα.

C/EBPα Restores G-CSF-induced Granulocytic Differentiation in 32Dcl3/DNStat3 Cells—To analyze the role of Stat3-

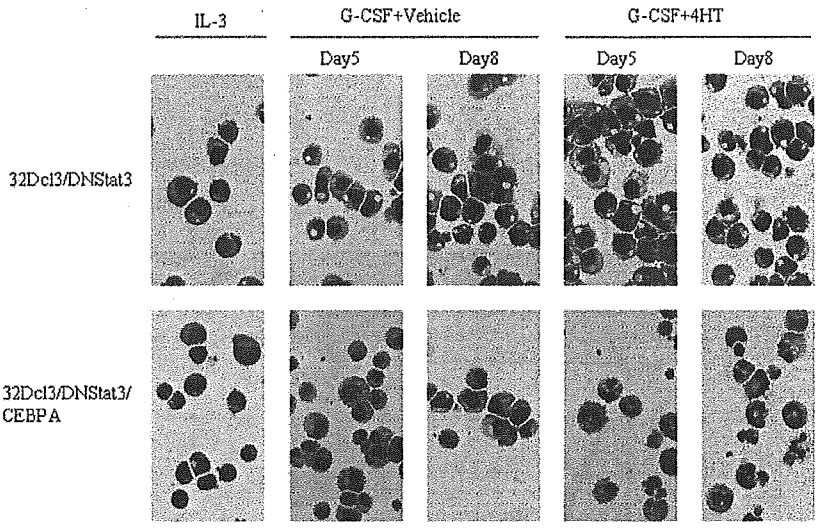

FIG. 6. **Morphologic features of 32Dcl3/DNStat3 and 32Dcl3/DNStat3/CEBPA cells.** Granulocytic differentiation of 32Dcl3/DNStat3 cells after induction of C/EBPα is shown. Cells were maintained in IL-3 and washed twice with PBS and then starved of cytokines for 8 h and stimulated with 10 ng/ml G-CSF plus 0.5 μM 4-HT or vehicle for 5 or 8 days. The cells were cytospun and stained with May-Grunwald and Giemsa stain (original magnification, ×400).

TABLE II
Differential count of 32Dcl3/DNStat3 and 32Dcl3/DNStat3/CEBPA cells

32Dcl3/DNStat3 and32Dcl3/DNStat3/CEBPA cells were maintained in IL-3 and starved of cytokines for 8 h and stimulated with 10 ng/ml G-CSF plus 0.5 μM 4-HT or vehicle for 5 days. Differential count was performed by May-Grunwald and Giemsa stain. Values are the mean ± S.D. percent of cells from three independent experiments. Myelocyte includes promyelocytes, myelocytes, and metamyelocytes. Band(seg) includes band and segmented neutrophils.

Cells	G-CSF+Vehicle	G-CSF+4HT
32Dcl3/DNStat3		
Myeloblasts	98.0 ± 0	99.3 ± 0.47
Myelocytes	1.3 ± 0.94	0.67 ± 0.47
Band(seg)	0.67 ± 0.94	0 ± 0
32Dcl3/DNStat3/CEBPA		
Myeloblasts	90.7 ± 3.3	3.0 ± 2.8
Myelocytes	5.0 ± 0.82	54.3 ± 2.4
Band(seg)	4.3 ± 2.6	42.7 ± 0.47

regulated C/EBPα function in the G-CSF signaling pathway, we transfected a C/EBPα-tamoxifen receptor fusion protein (C/EBPα-ER) into 32Dcl3 and 32Dcl3/DNStat3 cells (32Dcl3/CEBPA cells, 32Dcl3/DNStat3/CEBPA cells, respectively). The expression of C/EBPα-ER in these cells was verified by Western blotting (Fig. 5*A*). C/EBPα-ER localizes to the cytoplasm and is in an inactive form in the absence of tamoxifen. Upon treatment with tamoxifen, it translocates from cytoplasm to nucleus and becomes active. 32Dcl3, 32Dcl3/CEBPA, 32Dcl3/DNStat3, and 32Dcl3/DNStat3/CEBPA cells were cultured with G-CSF in the presence or absence of tamoxifen, and cell proliferation was examined by both counting viable cells and [^{3}H]thymidine incorporation. 32Dcl3/DNStat3 proliferated in response to G-CSF, and proliferation was not affected by the presence of tamoxifen. Conversely, G-CSF-induced proliferation of 32Dcl3/DNStat3/CEBPA cells in the presence of tamoxifen was dramatically reduced (Fig. 5, *B* and *C*).

32Dcl3/DNStat3 cells maintain morphologically immature characteristics and proliferate without granulocytic differentiation after G-CSF stimulation. We examined the morphological changes in 32Dcl3 and 32Dcl3/DNStat3 cells induced by G-CSF after translocation of C/EBPα from the cytoplasm to the nucleus. When tamoxifen was added to medium containing G-CSF, 32Dcl3/DNStat3/CEBPA cells rapidly began to differentiate into granulocytes, and 5 days later, about 40% of the cells were morphologically similar to mature neutrophils. In contrast, 32Dcl3/DNStat3/CEBPA cells cultured in G-CSF-containing medium without tamoxifen appeared immature with blast-like morphologic features (Fig. 6, Table II). To quantitatively analyze the difference in granulocyte maturation in 32Dcl3/DNStat3/CEBPA cells stimulated by G-CSF in the presence of tamoxifen, the mature granulocyte marker Gr-1 was monitored by FACS analysis. 32Dcl3 cells differentiate into Gr-1-positive neutrophils in response to G-CSF (Fig. 7*A*). As shown in Fig. 7*D*, Gr-1-positive cells were increased by the addition of tamoxifen in 32Dcl3/DNStat3/CEBPA cells treated with G-CSF, although low levels were detected in the absence of tamoxifen.

C/EBPα Up-regulates Genes That Are Related to Granulocytic Differentiation—In a conditional expression system, induction of C/EBPα leads to expression of granulocyte-specific genes, such as neutrophil primary granule genes (lysozyme M, *NGAL*) and the G-CSF receptor gene (17). In 32Dcl3/DNStat3 cells, the expression of these genes following G-CSF stimulation was inhibited (Fig. 8, *A*, *C*, and *E*). Interestingly, only *NGAL* was up-regulated by G-CSF in 32Dcl3/DNStat3/CEBPA cells following the restoration of C/EBPα (Fig. 8*B*). Conversely, the expression of lysozyme M and the G-CSF receptor were not changed by the restoration of C/EBPα (Fig. 8, *D* and *F*). These data suggest that regulatory factors in addition to C/EBPα may be involved in the induction of expression of granulocyte-specific genes by G-CSF.

DISCUSSION

G-CSF plays a pivotal role in granulopoiesis and granulocytic differentiation. The binding of G-CSF to its receptor leads to the activation of the Jak-Stat pathway, phosphatidylinositol-3 kinase pathway, and Ras-MAP kinase cascade (22). In the Jak-Stat pathway, G-CSF activates Jak1, Jak2, and Tyk2 followed by the activation of Stat1, Stat3, and Stat5 (7, 8).

Dominant-negative Stat3 inhibits G-CSF-induced transcriptional activity of Stat3 (Fig. 1*A*), as does G-CSF-induced granulocytic differentiation *in vitro* (11). Also, more transgenic mice with a targeted mutation of their G-CSF receptor that abolishes G-CSF-dependent Stat3 activation show severe neutropenia with an accumulation of immature myeloid precursors in their bone marrows (12). Consequently, Stat3 is thought to play an essential role in G-CSF-induced granulocytic differentiation.

32Dcl3 cells differentiate into neutrophils after treatment with G-CSF, and 32Dcl3/DNStat3 cells (32Dcl3 cells expressing dominant-negative Stat3) proliferate in G-CSF without differentiation. The degree of the phosphorylation of ERK1/2 by

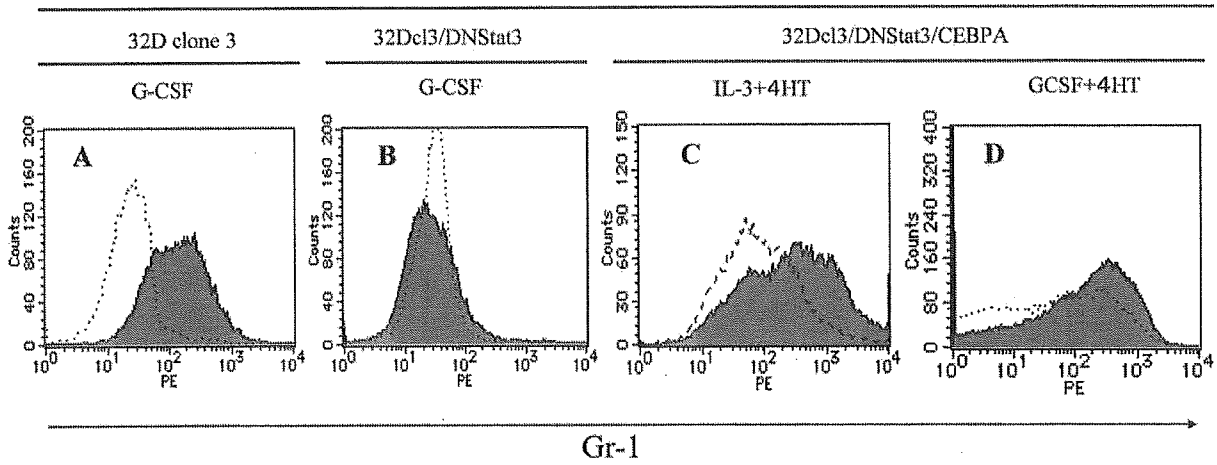

FIG. 7. **The expression of Gr-1 on 32Dcl3, 32Dcl3/DNStat3, and 32Dcl3/DNStat3/CEBPA cells.** 32Dcl3 (*A*) and 32Dcl3/DNStat3 cells (*B*) maintained in IL-3 (*broken line*) were starved of cytokines for 8 h and stimulated with 10 ng/ml G-CSF for 5 days (*solid line*). 32Dcl3/DNStat3/CEBPA (*C*) cells maintained in IL-3 were starved of cytokine for 8 h and stimulated with 1.0 ng/ml IL-3 (*C*) or 10 ng/ml G-CSF (*D*) plus 0.5 μM 4-HT (*solid line*) or vehicle (*broken line*) for 5 days.

G-CSF stimulation in 32Dcl3/DNStat3 cells was stronger than that in 32Dcl3 cells (Fig. 1*B*). We reported that Stat3 null bone marrow cells displayed a significant activation of ERK1/2 after G-CSF stimulation than wild-type bone marrow cells did using Stat3 conditional deficient mice (23). Then the augmented phosphorylation of ERK1/2 in response to G-CSF in 32Dcl3/DNStat3 cells might be caused by the functional abrogation of Stat3 in 32Dcl3/DNStat3 cells.

We compared gene profiles between two cell lines, 32Dcl3 and 32Dcl3/DNStat3 cells, to identify target genes of Stat3 in G-CSF signaling. We found that C/EBPα mRNA levels are rapidly up-regulated in 32Dcl3 cells following G-CSF treatment; these levels are increased 2.39-fold after 6 h and 4.20-fold after 48 h of treatment. In contrast to 32Dcl3 cells, C/EBPα mRNA levels are not changed in 32Dcl3/DNStat3 cells after G-CSF stimulation (Fig. 2*A*). The observation that cycloheximide does not inhibit G-CSF-induced increases in C/EBPα transcript levels (Fig. 2*B*) suggests that C/EBPα is induced by G-CSF directly downstream of Stat3. Dahl *et al.* (24) also reported that G-CSF induced the expression of C/EBPα in IL-3-dependent progenitors. *SOCS3* is one of the major target genes of Stat3. We previously reported that the expression level of SOCS3 protein in Stat3-deficient bone marrow cells is a trace, and it is not augmented by G-CSF stimulation (23). Contrary to this suppression of SOCS3 in Stat3-deficient cells, the induction of SOCS3 by G-CSF is not abolished in 32Dcl3/DNStat3 cells (data not shown).

The phosphorylation of ERK1/2 by G-CSF is stronger and the phosphorylation of Stat5 by IL-3 is weaker in 32Dcl3/DNStat3 cells when compared with those in 32D/Cl3 cells, although Stat1 phosphorylation by IFN-γ was not changed between these two cells (Figs. 1*B* and 3). Then there is the possibility that the transfection of dominant-negative Stat3 affects other signaling pathways in 32Dcl3/DNStat3 cells, resulting in the change of C/EBPα regulation. To clarify whether Stat3 directly up-regulates C/EBPα in the G-CSF signaling pathway in 32Dcl3 cells or not, we examined the effect of Stat3C on the transcription of C/EBPα. C/EBPα up-regulated the C/EBPα-dependent gene expression, and the G-CSF stimulation enhanced this C/EBPα-dependent gene expression (Fig. 4*A*). Strikingly, Stat3C augmented the C/EBPα-dependent gene expression as G-CSF stimulation did (Fig. 4, *B* and *C*). This means that G-CSF-induced up-regulation of C/EBPα-dependent gene expression is, at least partly, due to the activation of Stat3.

Two possibilities arise for the mechanism of the induction of C/EBPα transcription by activated Stat3 in the G-CSF signaling pathway. One is that activated Stat3 binds to the promoter region of *C/EBPα* and induces the transcription of C/EBPα. Analysis of the reported murine *C/EBPα* promoter sequence (20) identified no Stat-responsive elements (TTN5AA) (25, 26), but we found six Stat-responsive elements between 6 and 4 kb upstream of the *C/EBPα* transcription initiation site. Activated Stat3 might bind these Stat-responsive elements between 6 and 4 kb upstream of the C/EBPα transcription initiation site. The other possibility is that activated Stat3 might form the complex with C/EBPα and augment the transcriptional activity of C/EBPα because C/EBPα itself is the only protein reported to activate the murine *C/EBPα* promoter (20, 21). When a minimal TK promoter with CEBP-binding sites (p(C/EBP)2TK) together with C/EBPα was transfected to 293T cells, C/EBPα up-regulated C/EBPα-dependent gene expression. Activated Stat3 (Stat3C) enhanced this C/EBPα-dependent gene expression in collaboration with C/EBPα, although only Stat3C has no transcriptional activity on p(C/EBP)2TK (Fig. 4, *B* and *C*). In addition, the stimulation of G-CSF allows Stat3 to make the complex with C/EBPα (Fig. 4*D*). Then activated Stat3 by G-CSF makes the complex with C/EBPα and augments the transcriptional activity of C/EBPα. This is one of the reasons why induction of C/EBPα transcript through Stat3 activation by G-CSF occurred in 32Dcl3 cells. Several reports have described factors that repress *C/EBPα* promoter activity, such as SP1 (27), AP2A (28), or MYC (29). We show here for the first time that Stat3 augments the *C/EBPα* promoter activity.

Intracellular transcript levels of several genes were changed following G-CSF treatment downstream of Stat3 activation (Table I). To better identify the role of C/EBPα in Stat3-mediated signaling in G-CSF-induced granulocyte differentiation, C/EBPα-ER (C/EBPα-tamoxifen receptor fusion protein) was stably expressed in 32Dcl3 and 32Dcl3/DNStat3 cells. C/EBPα-ER translocates from the cytoplasm to the nucleus and becomes activated upon treatment with tamoxifen. Strikingly, transfection of C/EBPα-ER into 32Dcl3/DNStat3 cells abolished proliferation and induced myeloid differentiation by G-CSF without Stat3 activation (Figs. 5, *B* and *C*, and 6). These data indicate that C/EBPα activation induced by G-CSF through Stat3 plays an essential role in stopping the cell proliferation and inducing the differentiation to the myeloid lineage.