Distinct patterns of mutations occurring in de novo AML versus AML arising in the setting of severe congenital neutropenia

Daniel C. Link,[1] Ghada Kunter,[1] Yumi Kasai,[2] Yu Zhao,[1] Tracie Miner,[2] Michael D. McLellan,[2] Rhonda E. Ries,[1] Deepak Kapur,[4] Rakesh Nagarajan,[3] David C. Dale,[4] Audrey Anna Bolyard,[4] Laurence A. Boxer,[5] Karl Welte,[6] Cornelia Zeidler,[6] Jean Donadieu,[7] Christine Bellanné-Chantelot,[8] James W. Vardiman,[9,11] Michael A. Caligiuri,[10,11] Clara D. Bloomfield,[10,11] John F. DiPersio,[1] Michael H. Tomasson,[1] Timothy A. Graubert,[1] Peter Westervelt,[1] Mark Watson,[3] William Shannon,[1] Jack Baty,[12] Elaine R. Mardis,[2,13] Richard K. Wilson,[2,13] and Timothy J. Ley[1]

[1]Department of Medicine, [2]Genome Sequencing Center, [3]Department of Pathology and Immunology, Washington University, St Louis, MO; [4]University of Washington, Seattle; [5]Division of Pediatric Hematology/Oncology, Department of Pediatrics, C. S. Mott Children's Hospital, University of Michigan, Ann Arbor; [6]Department of Pediatric Hematology/Oncology, Hannover Medical School, Hannover, Germany; [7]Service d'Hémato Oncologie Pédiatrique, Hopital Trousseau, Paris, France; [8]Hôpital Saint-Antoine, Laboratoire de Cytogenetique, Paris, France; [9]University of Chicago, IL; [10]Cancer and Leukemia Group B, Chicago, IL; [11]Ohio State University Comprehensive Cancer Center and James Cancer Hospital, Columbus, OH; Divisions of [12]Biostatistics and [13]Genetics, Washington University, St Louis, MO

Severe congenital neutropenia (SCN) is an inborn disorder of granulopoiesis. Like most other bone marrow failure syndromes, it is associated with a marked propensity to transform into a myelodysplastic syndrome (MDS) or acute leukemia, with a cumulative rate of transformation to MDS/leukemia that exceeds 20%. The genetic (and/or epigenetic) changes that contribute to malignant transformation in SCN are largely unknown. In this study, we performed mutational profiling of 14 genes previously implicated in leukemogenesis using 14 MDS/leukemia samples from patients with SCN. We used high-throughput exon-based resequencing of whole-genome–amplified genomic DNA with a semiautomated method to detect mutations. The sensitivity and specificity of the sequencing pipeline was validated by determining the frequency of mutations in these 14 genes using 188 de novo AML samples. As expected, mutations of tyrosine kinase genes (*FLT3, KIT,* and *JAK2*) were common in de novo AML, with a cumulative frequency of 30%. In contrast, no mutations in these genes were detected in the SCN samples; instead, mutations of *CSF3R,* encoding the G-CSF receptor, were common. These data support the hypothesis that mutations of *CSF3R* may provide the "activated tyrosine kinase signal" that is thought to be important for leukemogenesis. (Blood. 2007;110:1648-1655)

Introduction

Bone marrow failure syndromes are characterized by a deficiency of 1 or more hematopoietic lineage. A common feature of both congenital and acquired forms of bone marrow failure is a marked propensity to develop acute myeloid leukemia (AML) or a myelodysplastic syndrome (MDS). The cumulative risk of developing AML or MDS in patients with chronic bone marrow failure ranges from 2% to greater than 20%.[1-7] Distinct genetic changes are associated with AML arising in the setting of bone marrow failure. The most prevalent genetic changes are abnormalities of chromosome 7; while only 5% of de novo AML samples display a loss of some or all of chromosome 7, they are present in 35% to 68% of AML arising in the setting of bone marrow failure.[3,8-14] These data suggest that distinct genetic mechanisms mediate leukemogenesis in patients with bone marrow failure syndromes.

Severe congenital neutropenia (SCN) is a congenital bone marrow failure syndrome characterized by severe neutropenia present from birth, an arrest of myeloid differentiation at the promyelocyte/myelocyte stage, and frequent infections.[15,16] Treatment with granulocyte colony-stimulating factor (G-CSF) is effective in increasing neutrophil counts in most patients.[4,17] Like other bone marrow failure syndromes, SCN is associated with a marked propensity to develop MDS or AML. The French Neutropenia Registry reported a cumulative incidence of MDS or AML in patients with SCN of 10.8% at 20 years of age.[18] A recent update of the Severe Chronic Neutropenia International Registry (SCNIR) showed that the cumulative incidence of MDS or AML was 21% after 10 years of G-CSF therapy.[4] Moreover, no plateau in the incidence of AML or MDS was observed, suggesting that the cumulative risk of progression may be even higher.

The molecular mechanisms that mediate leukemic transformation in SCN are poorly understood. Transformation has been associated with acquired clonal cytogenetic changes, most commonly involving the partial or complete loss of chromosome 7, activating *RAS* mutations, or abnormalities of chromosome 21.[3,10] In addition, acquired mutations of the *CSF3R* gene encoding the G-CSF receptor (G-CSFR) are present in a subset of patients with SCN and are strongly associated with the development of AML or MDS.[19-23] However, the contribution of these *CSF3R* mutations to leukemogenesis remains unclear, since transgenic mice expressing mutant G-CSFR do not

Submitted March 21, 2007; accepted April 23, 2007. Prepublished online as *Blood* First Edition paper, May 9, 2007; DOI 10.1182/blood-2007-03-081216.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

develop AML or MDS,[24,25] despite displaying a hyperproliferative response to G-CSF.

In the present study, the frequency of mutations in 14 genes previously implicated in the pathogenesis of AML was determined in 14 SCN patients who had transformed to MDS or acute leukemia. A high-throughput sequencing pipeline was established using exon-based resequencing of whole-genome–amplified genomic DNA (since sample abundance was extremely low for these samples) and a semiautomated method to detect mutations. The sensitivity and specificity of the resequencing pipeline was validated by determining the frequency of mutations in these 14 genes in 188 de novo AML samples. As expected, *FLT3, NPM1,* and *CEBPA* were the most frequently mutated genes in the de novo AML samples, but no mutations in these genes were detected in the SCN samples. The pattern of gene mutations in SCN-MDS/AML has implications both for the design of targeted therapies and for our understanding of the mechanisms of leukemic transformation in SCN patients.

Patients, materials, and methods

Human subjects

Ninety-four adult patients with de novo AML (Discovery set) were enrolled in a study at Washington University to identify genetic factors contributing to leukemia initiation and progression. Approval was obtained from the Washington University institutional review board for these studies. Informed consent was obtained in accordance with the Declaration of Helsinki. Eligibility criteria included age greater than 18, more than 30% myeloblasts in the bone marrow, and the absence of antecedent chemotherapy, radiation therapy, or history of myelodysplasia. Samples with 2 or fewer clonal cytogenetic abnormalities were given preference for inclusion in the study. Peripheral blood and bone marrow were obtained for analysis of tumor cells, and a 6-mm punch biopsy of skin was obtained for analysis of unaffected somatic cells. DNA was prepared from both tissues using standard protocols. Ninety-four de novo AML samples from patients with the same entry criteria were also obtained from Cancer and Leukemia Group B (CALGB); similar numbers of patients with the M0/1, M2, M3, and M4 French-American-British (FAB) subtype were selected for analysis.[26] The SCN bone marrow samples were obtained from the SCNIR and French Neutropenia Registry. The institutional review boards of all participating institutions approved this study.

Sequencing

Due to the limited amounts of starting material available for many samples, Phi29-based whole-genome amplification was performed on genomic DNA samples (both SCN and de novo AML samples) using the Qiagen Repli-G service (Qiagen, Valencia, CA).[27] Gene sequences were extracted from 2 public draft human genome databases, GenBank[28] and Ensembl,[29] and used as reference sequences for assembly and primer construction. The 500-bp 5′ flanking region and exonic regions, including the exon/intron boundaries, were amplified by polymerase chain reaction (PCR) to generate amplicons for resequencing. Universal tails were added to the 5′ ends of amplification primers to serve as the sequencing primer sites. The primers were either designed by ABI VariantSEQr (www.ncbi.nlm.nih.gov/projects/genome/probe/doc/ProjVariantSEQr.shtml) or with scripts that are based on the use of the Primer 3 program version 1.1.1.[30] Amplicons were purified by Exonuclease/SAP treatment and then directly sequenced using BigDye Terminator chemistry on an ABI3730 automated sequencer (Applied Biosystems, Foster City, CA).

Sequence analysis

The sequence traces were assembled and scanned for variations from the reference sequence using 2 parallel mutation detection pipelines: the mutation profiling pipeline and the PolyScan (Washington University, St Louis, MO) informatics suite.[31] The mutation profiling pipeline consists of the following steps: (1) analyze data using Phred version 0.20425.c to determine sequence quality;[32,33] (2) assemble sequence into contigs using Phrap version 0.990329 (http://www.phrap.org/phredphrapconsed.html); (3) align contigs to the reference sequence using Needle (Staden's Package version 1.6.0);[34] and (4) perform automated mutation detection using Informax Automated Sequence Assembly and Analysis Pipeline (ASAAP) version 1.2 (Invitrogen, Carlsbad, CA). To manually confirm the tagged sequence variations, we developed a new graphical user interface called Mutation Viewer (MV), where the mutations/polymorphisms are painted onto a DNA schematic showing protein motifs (eg, kinase domain). This program also prioritizes mutations based on their potential functional significance (conservative vs nonconservative substitutions), frequency, and whether they are known single-nucleotide polymorphisms (SNPs; derived by scanning the NCBI[28] and Ensembl[29] SNP databases). Finally, specimens from the same patient (eg, tumor versus germ line) were covisualized, allowing for the rapid identification of somatic mutations. All gene mutations reported in the literature (eg, *NRAS* G12D) or discovered in our initial analysis were considered "hot spots," and the relevant sequence tracings were manually reviewed for all other samples. Detailed protocols and primer sequences are available through the Washington University School of Medicine Genome Sequencing Center (GSC) website.[35] The primary sequence data are available at the Bioinformatics Core website.[36]

Sequence coverage for each amplicon was assessed as follows. Sequence traces for each amplicon were aligned with the reference sequence; alignments with a BLAST expect score of greater than 1×10^{-35} indicate poor sequence quality and were considered inadequate coverage. For amplicons in which more than 15% of samples failed, primers were redesigned and new amplicons were generated until adequate coverage for all amplicons was obtained.

Statistical analysis

Pairwise mutation analyses were done by permutation testing. Each pair of genes was tabulated to determine the frequency of the number of cases with both mutations present. Permutations were done by shuffling the mutation present/absent values within each gene independently and recalculating how often both genes were mutated within a single case. This null distribution was used to determine the significance value of the unpermuted number of double mutations. The Fisher exact test was used to test the independence of NPM1 mutations from mutations in the grouped tyrosine kinase pathway genes.

Results

Patient characteristics

A total of 188 de novo AML samples were analyzed (Table 1). For the Discovery set of 94 de novo AML samples obtained at Washington University, both skin ("germ line") and leukemic cell genomic DNA were obtained. This allowed us to determine whether an observed nucleotide change in a leukemic sample was somatically acquired. Of note, most of the de novo AML samples displayed normal or simple cytogenetic abnormalities (Table 1); abnormalities of chromosomes 5 or 7 were rare.

A total of 17 samples from 14 patients with SCN were analyzed (Table 2); 2 of the patients (nos. 14251 and 14252) have been described previously.[18] Nine patients had AML, 3 had MDS, 1 had acute lymphocytic leukemia (ALL), and 1 had a mixed AML/ALL picture. Most of the SCN samples were obtained from children (age range, 2.9-49.3 y; median, 10.8 y). Consistent with previous reports,[3,10] abnormalities of chromosome 7 were common, observed in 6 of the 13 samples where cytogenetics was available. The dose and duration of G-CSF therapy were variable.

Table 1. Clinical characteristics of the de novo AML samples

	Discovery	CALGB
Cytogenetic subgroup (%)		
Normal	35 (37.2)	41 (43.6)
t(15;17) only	12 (12.8)	20 (21.3)
t(8;21) only	1 (1.1)	3 (3.2)
inv (16) only	2 (2.1)	6 (6.4)
Trisomy 8 only	5 (5.3)	0 (0.0)
5q-/-5 only	1 (1.1)	0 (0.0)
7q-/-7 only	1 (1.1)	0 (0.0)
Complex karyotype, typical*	2 (2.1)	0 (0.0)
Complex karyotype, atypical*	10 (10.6)	0 (0.0)
Other	25 (26.6)	24 (25.5)
FAB subtype (%)		
M0	6 (6.4)	5 (5.3)
M1	18 (19.1)	18 (19.1)
M2	21 (22.3)	23 (24.5)
M3	17 (18.1)	23 (24.5)
M4	18 (19.1)	25 (26.6)
M5	9 (9.6)	0 (0.0)
M6	3 (3.2)	0 (0.0)
M7	2 (2.1)	0 (0.0)
Age, y (range)	52.8 (16-81)	40.3 (22-70)
Sex, no. (%)		
Male	56 (59.6)	54 (57.4)
Female	38 (40.4)	40 (42.6)
Ethnicity, no. (%)		
Asian	1 (1.1)	4 (4.2)
Black	9 (9.6)	6 (6.4)
White	83 (88.3)	70 (74.5)
Hispanic	1 (1.1)	11 (11.7)
Other	0 (0.0)	3 (3.2)

CALGB indicates Cancer and Leukemia Group B.
*Defined based on the criteria outlined by Bacher et al.[37]

Sequencing strategy

The genetic progression factors that contribute to leukemogenesis in the setting of SCN are largely unknown. We reasoned that genes that are frequently mutated in de novo AML might also contribute to leukemic transformation in SCN. A total of 14 genes were chosen for resequencing; they included tyrosine kinase genes and genes involved in the transmission of tyrosine kinase signals (*FLT3, KIT, CSF1R, JAK2*, *NRAS, KRAS,* and *PTPN11*), myeloid transcription factors (*RUNX1, CEBPA,* and *SPI1*), and *NPM1* and *TP53*. In addition, we sequenced the *ELA2* and *CSF3R* genes, since mutations of these genes have been implicated in leukemic transformation in SCN. The specific exons that were sequenced for each gene are shown in Table S1 (available on the *Blood* website; see the Supplemental Materials link at the top of the online article).

A high-throughput sequencing pipeline was employed using whole-genome–amplified genomic DNA and a semiautomated method to detect mutations. The amount of tumor and/or skin DNA available for many of these samples was very limited, necessitating whole-genome amplification to create adequate amounts of template for exonic amplification and resequencing. Since large-scale resequencing with whole-genome–amplified DNA has not previously been reported, it was first necessary to "credential" this material for analysis by assessing the frequency of known mutations in a large cohort of de novo AML patients.

A crucial factor in assessing mutation frequency is the quality and coverage of the sequences obtained. Therefore, we developed an automated method to rapidly assess sequence quality and coverage. Figure 1 contains a representative sequence coverage map for exon 2 of *NRAS*. High-quality double-stranded (green) or single-stranded (yellow) sequence was observed for nearly all of the samples. For the purposes of this study, a given sample had to meet 2 criteria to be considered to have adequate sequence coverage. First, for genes with mutational hot spots (eg, codons 12, 13, and 64 for *NRAS*), this region had to have at least single-stranded sequence coverage. Second, sequence gaps (if any) in the coding region had to be less than 10 nucleotides in length. Based on these criteria, only 5 of 210 tumor samples (both de novo AML and SCN samples) had inadequate sequence coverage for *NRAS* (Figure 1). In fact, the sequence quality and coverage was high for all analyzed genes, with an average of 95% plus or minus 4.3% of samples having adequate coverage (range, 86%-100%). In each case where inadequate sequence coverage was obtained, the sample was eliminated from the final analysis of the relevant gene mutation frequency.

Frequency of gene mutations in de novo AML

To assess the sensitivity and specificity of our sequencing pipeline, we first sequenced the 188 de novo AML samples, since the frequency of mutations in these genes had previously been established. Figure 2 summarizes the nucleotide deletions, insertions, and nonsynonymous single-nucleotide changes that were

Table 2. SCN patient characteristics

Patient ID	Age, y	Sex	Diagnosis	Cytogenetics	G-CSF dose, μg/kg/per day	G-CSF duration, y
12397	15	M	AML (M5)	45, XY, −7	1.9	6.3
17390	7.8	F	AML (M2)	46,XX,ins(1,?7)(q25;q31q36) t(8,21)(q22;qq22)[20]	3.8	3.7
12374	49.3	M	AML (M7)	45, XY, −7, del(6)(q21)	2.4	4.8
17392	5	F	AML (M2)	46, XY, inv (16), t(5:6)(q31:q27)	13.5	3.7
12377	20.3	F	MDS (RAEB)	46, XX, add(2)(q37), add 7(q22)	9	8
17393	18.3	M	MDS	NA	5	11.4
17394	13.4	M	AML (M1)	46, XY	3.7	13.5
12400	7.8	M	AML	46, XY	5.2	6.8
13462	16.8	F	AML (M5)	45, XX, −7/4n	2.6	8.5
13464	2.9	M	AML/B-ALL	46 XY,add (21)(q22); XY, +21	10-120	2.8
13476	7	F	AML (M0)	45,XX,−7[12]/46,XX[11]. nuc ish (D7S486x1)	0.72	6.8
13995	15.6	F	AML (M2)	del(10)(q32)	48	10.2
14251	11.8	F	ALL	48,XX,del(5)(q21q34),+21,+22[16]/46, XX[8]	26	6.8
14252	7.4	M	MDS	47,XY,−7,+21,+21[9]/46, XY[5]	18	4.4

All patients included in this table are white. The clinical status at the time the sample was obtained is shown, including the diagnosis, major cytogentic abnormality (if any), and duration and dose of G-CSF.
RAEB indicates refractory anemia with excess of blasts; NA, not available.

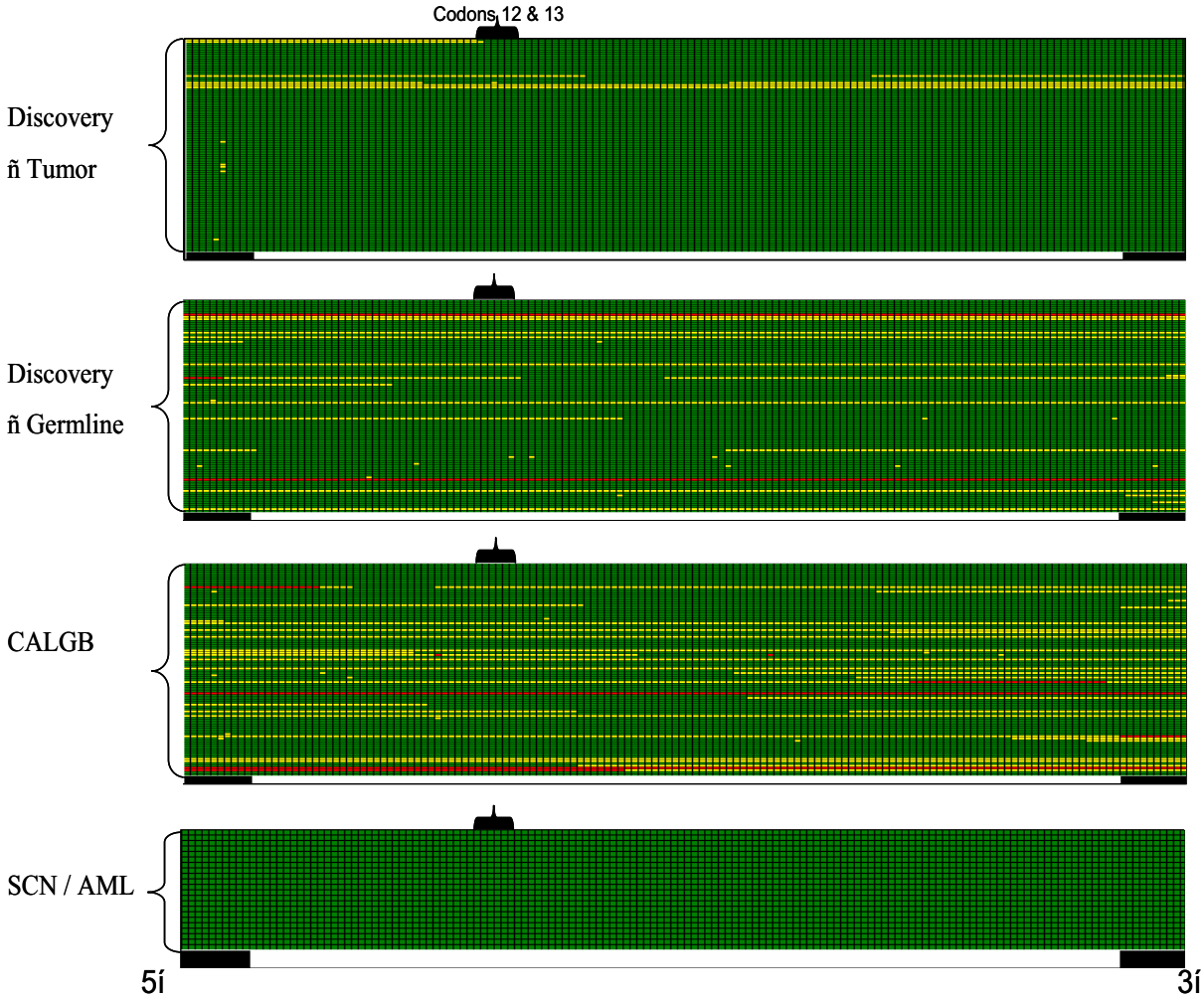

Figure 1. NRAS exon 2 sequence coverage maps. Sequence coverage for exon 2 of NRAS is shown for the Discovery set of 94 AML samples (both tumor and skin), the 94 CALGB AML samples, and the 22 SCN-AML samples; each row represents an individual sample. Each column represents a single nucleotide position starting from the 5′-intronic region (■), through exon 2 of NRAS (□) and 3′-intronic region (■). The position of codons 12 and 13 of NRAS is highlighted. High-quality double-stranded sequence is shown in green, single-stranded coverage in yellow, and no coverage in red.

detected in the Discovery set of 94 de novo AML samples; single-nucleotide changes present in the NCBI[28] or Ensembl[29] SNP databases were excluded from analysis. Consistent with previous reports, the 4 genes most often mutated were *FLT3,*[38] *NPM1,*[39] *NRAS*,[37,40] and *CEBPA*,[41] with frequencies in our Discovery set of 27.7%, 25.5%, 9.6%, and 5.3%, respectively (Table 3). As reported previously, somatic mutations of *KIT,*[42] *KRAS,*[40] *PTPN11,*[43] *JAK2,*[44] *RUNX1,*[45] *TP53,*[46] and *SPI1*[47,48] were observed but infrequent (< 5%). No somatic mutations of *CSF1R* (encoding c-fms), *ELA2,* or *CSF3R* were detected. Very similar data were obtained with the CALGB set of de novo AML samples (Figure S1). To conserve resources, *CSF1R* and *CSF3R* were not sequenced in this set. Collectively, the frequency of gene mutations in our de novo AML samples was comparable to that reported in the literature, validating our use of whole-genome–amplified DNA for the resequencing pipeline (Table 3).

The multiplexed gene sequencing approach allowed us to look for associations between gene mutations; 2 significant associations were observed (Figure 2). Although the numbers were small, a significant positive association between KIT mutations and RUNX1 mutations was observed ($P = .027$). As reported previously,[49] a trend toward a positive association between FLT3 mutations and NPM1 mutations was also observed ($P = .076$). Indeed, after grouping the tyrosine kinase pathway genes (*FLT3, KIT, CSF1R, NRAS, KRAS, PTPN11,* and *JAK2*) together, the association with mutations of NPM1 was significant ($P = .049$). A negative association of tyrosine kinase pathway gene mutations with each other was observed, but it was not statistically significant (ie, mutations of *FLT3, KIT, JAK2,* or *PTPN11* were never seen in the same sample; Figure 2). The same pattern does not appear to hold true for *RAS* mutations, since *NRAS* and *KRAS* mutations were seen in association with *FLT3* or *KIT* mutations.

Frequency of gene mutations in SCN AML/ALL/MDS

For the SCN samples, only tumor (bone marrow) samples were available, so we were unable to directly address whether the observed nucleotide changes were acquired. As expected, over 40% of patients (6 of 14) with SCN carried mutations (presumably germ line) in *ELA2* (Figure 3). Consistent with a previous report,[18] G185R *ELA2* mutations are frequent in patients with SCN who have progressed to leukemia or AML. *CSF3R* mutations were observed in 6 of 14 patients (42.8%). In 1 case (no. 12400), the *CSF3R* mutation was clearly acquired, since it was not present in a preleukemic sample (no. 12380) from the same patient. In contrast to the de novo AML samples, no

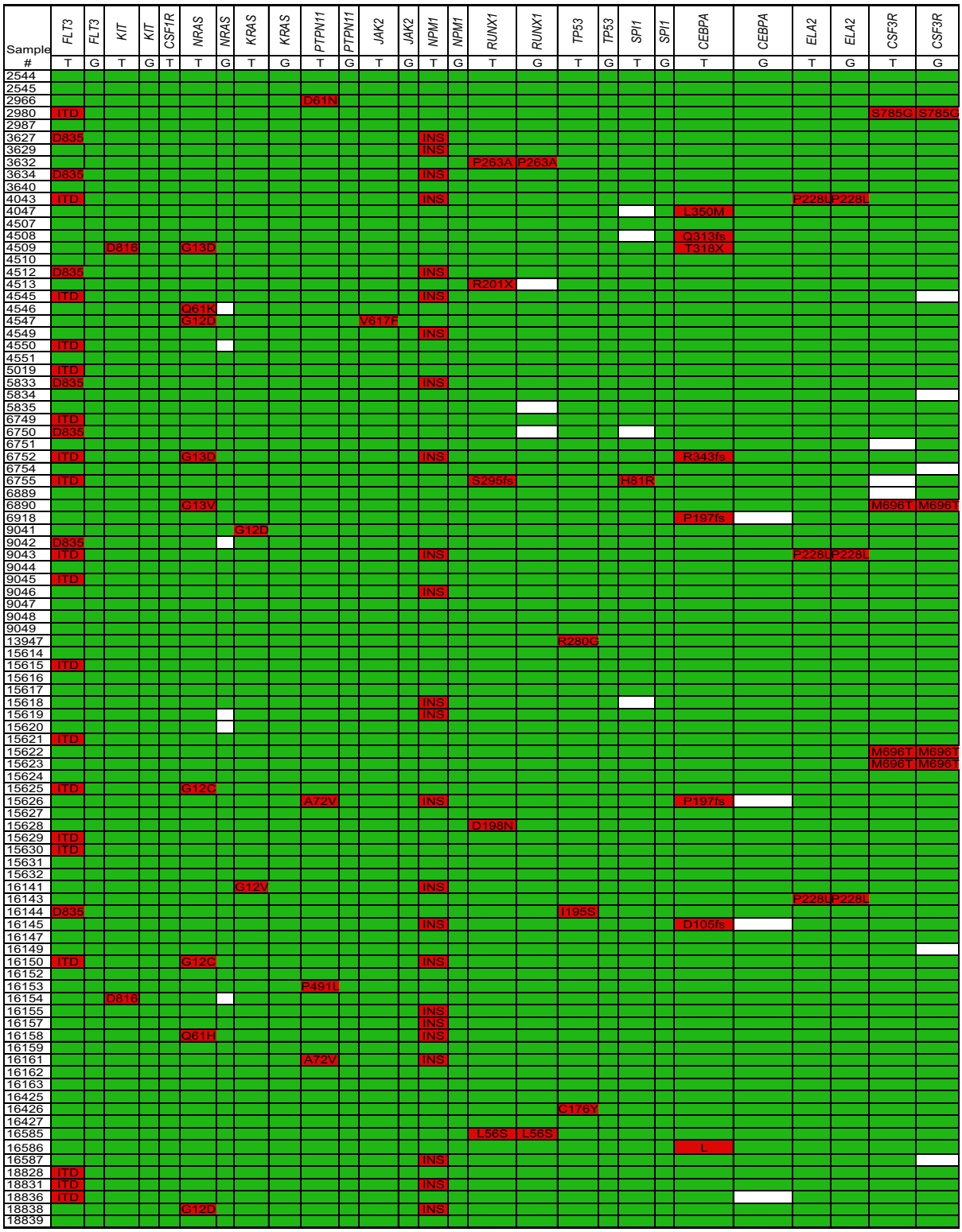

Figure 2. Nucleotide changes for the discovery set of de novo AML samples. Nonsynonymous nucleotide changes and gene deletions or insertions are highlighted in red; white boxes indicate missing sequence data. Known SNPs are excluded from this analysis. Both leukemia (T) and germline (G) samples for each patient were sequenced. ITD indicates internal tandem duplication (FLT3); INS, tetranucleotide insertion of exon 12 (NPM1).

mutations of *FLT3, NRAS, CEBPA,* or *NPM1* were detected. Nonsynonymous single-nucleotide changes of *RUNX1* were detected in 2 SCN samples. Although the R174Q mutation is associated with AML,[45] the significance of the A160T sequence change is less clear, since it has not previously been reported in AML or MDS. A single SCN-AML sample was found to have a homozygous T801I substitution in *KIT;* however, this substitution does not result in constitutive c-*kit* activation[50] and has not previously been detected in AML or MDS[51]; its significance is therefore unclear.

To compare mutation frequencies with de novo AML, we limited our analysis to the 10 SCN samples with AML and

Table 3. Mutation frequency in SCN-AML versus de novo AML

	SCN			Discovery		
Gene	Mutation	n	Frequency, %	Mutation	n	Frequency, %
FLT3	0	10	0.0	51	185	27.6
KIT	0†	10	0.0	4	188	2.1
*CSF1R**	0	9	0.0	0	94	0.0
NRAS	0	10	0.0	17	182	9.3
KRAS	0	10	0.0	3	187	1.6
PTPN11	0	10	0.0	5	183	2.7
JAK2	0	10	0.0	2	187	1.1
NPM1	0	10	0.0	43	180	23.9
RUNX1	0	10	0.0	9	182	4.9
TP53	0	10	0.0	3	183	1.6
SPI1	0	9	0.0	2	184	1.1
CEBPA	0	10	0.0	11	186	5.9
ELA2	3	10	30.0	0	188	0.0
*CSF3R**	3	10	30.0	0	91	0.0

Only those SCN samples with AML were included. The Discovery and CALBG de novo AML data sets were combined. Germ line nonsynonymous mutations and samples with inadequate sequence coverage were excluded.

*Not sequenced in the CALBG samples.

†The T801I KIT SNP was excluded because it has been shown to be a nonactivating mutation.

eliminated those samples with inadequate sequence coverage (Table 3). Compared with the de novo AML samples, there was a nonsignificant trend toward a reduced frequency of mutations of *FLT3* (27.6% vs 0%; $P < .07$) and *NPM1* (23.9% vs 0%; $P = .12$) in the SCN-AML samples. Consistent with previous reports, activating mutations of the G-CSFR were specifically associated with SCN-AML (0% vs 30%; $P < .01$).

Discussion

In this study, we used high-throughput exonic resequencing to determine the mutation frequency of 14 AML-associated genes in patients with SCN-associated AML. To validate our resequencing methodologies, we sequenced the regions containing these common mutations in 188 de novo AML samples and found that our pipeline detected all of these mutations at the expected frequencies. The pattern of mutations in the SCN-MDS/leukemia samples was distinct, since no mutations in *NPM1, CEBPA,* or *FLT3* were detected. As expected, we found that the *CSF3R* gene was frequently mutated in SCN-AML but not in de novo AML. These results therefore suggest that the genetic basis of SCN-AML is different from that of de novo AML. However, it is important to note that de novo AML samples analyzed in this study had certain favorable biologic features compared with the SCN samples, including the lack of antecedent myelodysplasia and relatively few cytogenetic abnormalities. These features make comparisons of the 2 groups difficult, but the new information regarding the SCN-AML patients is still informative regarding their pathophysiology.

Mutations of *CSF3R* are present in a substantial proportion of patients with SCN. In the largest published series, which included 121 patients with SCN, *CSF3R* mutations were detected in 41%.[23] These mutations are acquired, typically heterozygous, and nearly always introduce a premature stop codon resulting in the truncation of the distal cytoplasmic portion of the G-CSFR.[19-22] *CSF3R* mutations are strongly associated with the development of AML/MDS. Germeshausen et al[23] reported that the incidence of *CSF3R* mutations was 78% (18/23) in individuals with SCN and monosomy 7, MDS, or AML compared with 34% (43/125) in patients without MDS or AML. In the present study, *CSF3R* mutations were detected in 43% (6/14) of SCN patients with malignant transformation. The lower incidence of *CSF3R* mutations in our study is likely due to differences in the methods used to detect *CSF3R* mutations. In the study by Germeshausen et al,[23] sequencing was performed on PCR-amplified cDNA after reverse transcription of hematopoietic cell RNA; using this approach, they were able to detect *CSF3R* mutations in samples in which only 5% of the cells contained this mutation. In contrast, our method of direct amplification of tumor DNA to generate amplicons for resequencing has been shown to reliably detect mutations that are present in at least 33% of cells (this applies to all genes sequenced in this study).[26] Consistent with previous reports,[23,52,53] no *CSF3R* mutations were detected in the de novo AML samples, suggesting that *CSF3R* mutations are rare in de novo AML.

Accumulating data suggest that activating mutations of tyrosine kinase genes may play a key role in leukemogenesis. Indeed, we detected mutations of tyrosine kinase genes (*FLT3, KIT,* and *JAK2*) in 30% of de novo AML samples. There is evidence suggesting that the truncation mutations of *CSF3R* observed in SCN are likewise activating. Expression of mutant G-CSFR in myeloid cell lines results in enhanced proliferative and survival signals.[19,54,55] Moreover, transgenic mice carrying truncating mutations of *CSF3R* display a hyperproliferative response to G-CSF.[24,25] Since the G-CSFR and receptor tyrosine

Sample #	Diagnosis	*FLT3*	*KIT*	*CSF1R*	*NRAS*	*KRAS*	*PTPN11*	*JAK2*	*NPM1*	*RUNX1*	*TP53*	*SPI1*	*CEBPA*	*ELA2*	*CSF3R*
12371	Pre														
12397	AML														
17390	AML														
17392	AML														
12374	AML														
12380	Pre														
12400	AML														Y726X
17394	AML														
13476	AML														
13995	AML		T801I											G185R	Q717X
13462	AML													L123P	Q706X
13465	Pre													A28V	
13464	AML/ALL													A28V	
14252	MDS					G12D				R174Q				L92P	Q715X
17393	MDS														Q715X
12377	MDS														
14251	ALL									A160T				G185R	Q699X

Figure 3. Nucleotide changes in the SCN samples. Nonsynonymous nucleotide changes and gene deletions or insertions are highlighted in red; white boxes indicate missing sequence data. Known SNPs are excluded from this analysis. Pre- and post-leukemic samples from the same patient are outlined.

kinases share many signaling pathways, these observations suggest that the mutant G-CSFR present in many cases of SCN-AML may provide the "activated tyrosine kinase signal [pathway]." This hypothesis predicts that mutations of tyrosine kinase genes associated with de novo AML would not be required for disease progression. Consistent with this prediction, no activating mutations of the tyrosine kinase genes were observed in the SCN samples (0/14; $P = .007$ compared with de novo AML). The rarity of tyrosine kinase gene mutations in SCN is unlikely to be related to the age of the patients or the complexity of karyotypic abnormalities. The mutation frequency of *FLT3* and *KIT* in childhood AML is 19% and 3%, respectively[56,57]; similar to that reported for adult AML. Of particular relevance to SCN, the frequency of *FLT3* mutations in AML with karyotypic abnormalities of chromosome 7 is 5.8%.[38] In addition, since therapy-related MDS/AML shares many features with SCN-AML (including a high incidence of chromosome 7 abnormalities and a history of antecedent MDS), it is noteworthy that the reported frequency of *FLT3* mutations in therapy-related MDS/AML ranges from 4.3% to 7.8%.[58,59] Finally, a recent study reported that activating *KIT* mutations were present in 19% of patients with AML arising in the setting of MDS (MDS-AML), another condition associated with a high frequency of chromosome 7 abnormalities.[61] Collectively, these data suggest that the rarity of tyrosine kinase gene mutations is a unique finding in SCN-MDS/AML and support the hypothesis that *CSF3R* mutations provide the activated tyrosine kinase signal that is thought to be important for leukemogenesis.

In the present study, only a single *NRAS* or *KRAS* mutation was detected in the 14 SCN-AML/MDS samples. Similarly, a preliminary study reported a 7% frequency of *RAS* mutations (1/14 cases of SCN AML/MDS).[61] In contrast, an earlier study reported a *RAS* mutation frequency of 38% (5/13 cases of SCN AML/MDS).[10] Interestingly, most of the patients in this earlier study were older and had not been treated with G-CSF from infancy. Consequently, it is possible that long-term G-CSF therapy selects for *CSF3R* mutations, reducing the need for RAS mutations during disease progression. Consistent with this possibility, resequencing of 3 of the original 5 cases of SCN AML/MDS with *RAS* mutations revealed that none of these cases had mutations of *CSF3R*.[61] The mutual exclusivity of activating *RAS* and *CSF3R* mutations suggests that these genes may act on a shared pathway; however, it should be noted thatthe sole SCN sample in our study with a *KRAS* mutation also had a *CSF3R* mutation. In any case, combining data from all of these studies, the frequency of *RAS* mutations in SCN MDS/leukemia is approximately 15%, similar to that reported for de novo AML.

The frequency of *NPM1* mutations has been reported to be approximately 35% in adult de novo AML[39] and 6.5% in pediatric de novo AML.[63] In the present study, no *NPM1* mutations were detected in the 14 SCN samples versus 24% in the de novo AML samples ($P = .04$). Of note, *NPM1* mutations are more frequently detected in AML samples with a normal karyotype; the frequency of *NPM1* mutations in de novo AML with an abnormal karyotype is only 6.4%.[63] Since most cases of SCN-MDS/AML are associated with an abnormal karyotype, this may partially account for the reduced *NPM1* mutation frequency in these samples. Nonetheless, the data suggest that mutations of the *NPM1* are rare during leukemic progression in SCN patients.

There is accumulating evidence that the genetic changes that contribute to secondary leukemias, including therapy-related AML and AML arising in the setting of MDS (MDS-AML), are distinct from those that contribute to the pathogenesis of de novo AML. The present data suggest that the mutations that contribute to SCN-AML are more similar to those of the secondary AML syndromes. Similar to therapy-related AML and MDS-AML, mutations of *NPM1* and *CEBPA* are rare in SCN-MDS/AML, whereas abnormalities of chromosomes 5 and 7 are more common than in de novo AML. The one outlier is *CSF3R,* where mutations appear to be uniquely associated with SCN. Given the poor prognosis of SCN-AML and other secondary leukemia syndromes, it will be important to identify the genes that contribute to transformation under these circumstances. These data will improve our understanding of the mechanisms of leukemic transformation and may provide new targets for drug design.

Acknowledgments

This work was supported by National Institutes of Health (NIH) grant CA101937, by US Public Health Service grants CA31946 and CA101140, and by the Leukemia Clinical Research Foundation (CALGB).

Authorship

Contribution: G.K., Y.K., Y.Z., T.M., M.D.M., R.E.R., and D.K. generated and/or analyzed the data. W.S. and J.B. performed the statistical analyses. D.C.D., A.A.B., L.A.B., K.W., C.Z., J.D., C.B.-C., J.W.V., M.A.C., and C.D.B. provided crucial reagents. D.C.L., R.N., J.F.D., M.H.T., T.A.G., P.W., M.W., E.R.M., R.K.W., and T.J.L. designed the experiment. D.C.L. and T.J.L. wrote the paper.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Daniel C. Link, Division of Oncology, Department of Medicine, 660 S Euclid Avenue, Campus Box 8007, St Louis, MO 63110; e-mail: dlink@im.wustl.edu.

References

1. Alter B. Inherited bone marrow failure syndromes. In: Nathan DG, Orkin SH, Look AT, Ginsburg D, eds. Hematology of Infancy and Childhood, 6th ed. Philadelphia, PA: WB Saunders; 2003:280-365.
2. Rosenberg PS, Greene MH, Alter BP. Cancer incidence in persons with Fanconi anemia. Blood. 2003;101:822-826.
3. Freedman MH, Bonilla MA, Fier C, et al. Myelodysplasia syndrome and acute myeloid leukemia in patients with congenital neutropenia receiving G-CSF therapy. Blood. 2000;96:429-436.
4. Rosenberg PS, Alter BP, Bolyard AA, et al. The incidence of leukemia and mortality from sepsis in patients with severe congenital neutropenia receiving long-term G-CSF therapy. Blood. 2006; 107:4628-4635.
5. Smith OP, Hann IM, Chessells JM, Reeves BR, Milla P. Haematological abnormalities in Shwachman-Diamond syndrome. Br J Haematol. 1996; 94:279-284.
6. Socie G, Rosenfeld S, Frickhofen N, Gluckman E, Tichelli A. Late clonal diseases of treated aplastic anemia. Semin Hematol. 2000;37:91-101.
7. Germing U, Strupp C, Kuendgen A, et al. Prospective validation of the WHO proposals for the classification of myelodysplastic syndromes. Haematologica. 2006;91:1596-1604.

8. Auerbach AD, Allen RG. Leukemia and preleukemia in Fanconi anemia patients: a review of the literature and report of the International Fanconi Anemia Registry. Cancer Genet Cytogenet. 1991;51:1-12.
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. Kalra R, Dale D, Freedman M, et al. Monosomy 7 and activating RAS mutations accompany malignant transformation in patients with congenital neutropenia. Blood. 1995;86:4579-4586.
11. Larson RA. Myeloid leukemia after cytotoxic therapy and other hematotoxins. J Toxicol Environ Health A. 2000;61:381-386.
12. Maciejewski JP, Risitano A, Sloand EM, Nunez O, Young NS. Distinct clinical outcomes for cytogenetic abnormalities evolving from aplastic anemia. Blood. 2002;99:3129-3135.
13. Sole F, Espinet B, Sanz GF, et al. Incidence, characterization and prognostic significance of chromosomal abnormalities in 640 patients with primary myelodysplastic syndromes: Grupo Cooperativo Espanol de Citogenetica Hematologica. Br J Haematol. 2000;108:346-356.
14. Spirito FR, Crescenzi B, Matteucci C, Martelli MF, Mecucci C. Cytogenetic characterization of acute myeloid leukemia in Shwachman's syndrome: a case report. Haematologica. 2000;85:1207-1210.
15. Kostmann R. Infantile genetic agranulocytosis: a new recessive lethal disease in man. Acta Paediatr. 1956;105:1-78.
16. Welte K, Zeidler C, Dale DC. Severe congenital neutropenia. Semin Hematol. 2006;43:189-195.
17. Dale DC, Bonilla MA, Davis MW, et al. A randomized controlled phase III trial of recombinant human granulocyte colony-stimulating factor (filgrastim) for treatment of severe chronic neutropenia. Blood. 1993;81:2496-2502.
18. Donadieu J, Leblanc T, Bader Meunier B, et al. Analysis of risk factors for myelodysplasias, leukemias and death from infection among patients with congenital neutropenia: experience of the French Severe Chronic Neutropenia Study Group. Haematologica. 2005;90:45-53.
19. Dong F, Brynes RK, Tidow N, Welte K, Lowenberg B, Touw IP. Mutations in the gene for the granulocyte colony-stimulating-factor receptor in patients with acute myeloid leukemia preceded by severe congenital neutropenia [see comments]. N Engl J Med. 1995;333:487-493.
20. Dong F, Dale DC, Bonilla MA, et al. Mutations in the granulocyte colony-stimulating factor receptor gene in patients with severe congenital neutropenia. Leukemia. 1997;11:120-125.
21. Dong F, Hoefsloot LH, Schelen AM, et al. Identification of a nonsense mutation in the granulocyte-colony-stimulating factor receptor in severe congenital neutropenia. Proc Natl Acad Sci U S A. 1994;91:4480-4484.
22. Tidow N, Pilz C, Teichmann B, et al. Clinical relevance of point mutations in the cytoplasmic domain of the granulocyte colony-stimulating factor receptor gene in patients with severe congenital neutropenia [see comments]. Blood. 1997;89:2369-2375.
23. Germeshausen M, Ballmaier M, Welte K. Incidence of CSF3R mutations in severe congenital neutropenia and relevance for leukemogenesis: results of a long-term survey. Blood. 2007;109:93-99.
24. Hermans MH, Ward AC, Antonissen C, Karis A, Lowenberg B, Touw IP. Perturbed granulopoiesis in mice with a targeted mutation in the granulocyte colony-stimulating factor receptor gene associated with severe chronic neutropenia. Blood. 1998;92:32-39.
25. McLemore ML, Poursine-Laurent J, Link DC. Increased granulocyte colony-stimulating factor responsiveness but normal resting granulopoiesis in mice carrying a targeted granulocyte colony-stimulating factor receptor mutation derived from a patient with severe congenital neutropenia. J Clin Invest. 1998;102:483-492.
26. Ley TJ, Minx PJ, Walter MJ, et al. A pilot study of high-throughput, sequence-based mutational profiling of primary human acute myeloid leukemia cell genomes. Proc Natl Acad Sci U S A. 2003; 100:14275-14280.
27. Hosono S, Faruqi AF, Dean FB, et al. Unbiased whole-genome amplification directly from clinical samples. Genome Res. 2003;13:954-964.
28. National Center for Biotechnology Information. GenBank. www.ncbi.nlm.nih.gov. Accessed June 20, 2007.
29. European Bioinformatics Institute. Ensembl. www.ensembl.org. Accessed June 20, 2007.
30. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol. 2000;132:365-386.
31. Chen K, McLellan MD, Ding L, et al. Polyscan: an automatic Indel and SNP detection approach to the analysis of human re-sequencing data. Genome Res. 2007;17:659-666.
32. Ewing B, Green P. Base-calling of automated sequencer traces using phred, II: error probabilities. Genome Res. 1998;8:186-194.
33. Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred, I: accuracy assessment. Genome Res. 1998;8:175-185.
34. Needleman SB, Wunsch CD. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol. 1970;48:443-453.
35. Washington University School of Medicine Genome Sequencing Center. http://genome.wustl.edu/platforms.cgi?id = 7. Accessed June 20, 2007.
36. Siteman Cancer Center Bioinformatics Core. http://bioinformatics.wustl.edu/webTools/logOff.do. Accessed June 20, 2007.
37. Bacher U, Haferlach T, Schoch C, Kern W, Schnittger S. Implications of NRAS mutations in AML: a study of 2502 patients. Blood. 2006;107: 3847-3853.
38. Thiede C, Steudel C, Mohr B, et al. Analysis of FLT3-activating mutations in 979 patients with acute myelogenous leukemia: association with FAB subtypes and identification of subgroups with poor prognosis. Blood. 2002;99:4326-4335.
39. Thiede C, Koch S, Creutzig E, et al. Prevalence and prognostic impact of NPM1 mutations in 1485 adult patients with acute myeloid leukemia (AML). Blood. 2006;107:4011-4020.
40. Bowen DT, Frew ME, Hills R, et al. RAS mutation in acute myeloid leukemia is associated with distinct cytogenetic subgroups but does not influence outcome in patients younger than 60 years. Blood. 2005;106:2113-2119.
41. Leroy H, Roumier C, Huyghe P, Biggio V, Fenaux P, Preudhomme C. CEBPA point mutations in hematological malignancies. Leukemia. 2005;19: 329-334.
42. Beghini A, Peterlongo P, Ripamonti CB, et al. C-kit mutations in core binding factor leukemias. Blood. 2000;95:726-727.
43. Tartaglia M, Niemeyer CM, Fragale A, et al. Somatic mutations in PTPN11 in juvenile myelomonocytic leukemia, myelodysplastic syndromes and acute myeloid leukemia. Nat Genet. 2003;34:148-150.
44. Steensma DP, McClure RF, Karp JE, et al. JAK2 V617F is a rare finding in de novo acute myeloid leukemia, but STAT3 activation is common and remains unexplained. Leukemia. 2006;20:971-978.
45. Preudhomme C, Warot-Loze D, Roumier C, et al. High incidence of biallelic point mutations in the Runt domain of the AML1/PEBP2 alpha B gene in Mo acute myeloid leukemia and in myeloid malignancies with acquired trisomy 21. Blood. 2000; 96:2862-2869.
46. Stirewalt DL, Kopecky KJ, Meshinchi S, et al. FLT3, RAS, and TP53 mutations in elderly patients with acute myeloid leukemia [erratum appears in Blood. 2001;98:924]. Blood. 2001;97:3589-3595.
47. Mueller BU, Pabst T, Osato M, et al. Heterozygous PU. 1 mutations are associated with acute myeloid leukemia. Blood. 2002;100:998-1007.
48. Dohner K, Tobis K, Bischof T, et al. Mutation analysis of the transcription factor PU. 1 in younger adults (16 to 60 years) with acute myeloid leukemia: a study of the AML Study Group Ulm (AMLSG ULM). Blood. 2003;102:3850-3851.
49. Verhaak RGW, Goudswaard CS, van Putten W, et al. Mutations in nucleophosmin (NPM1) in acute myeloid leukemia (AML): association with other gene abnormalities and previously established gene expression signatures and their favorable prognostic significance. Blood. 2005;106: 3747-3754.
50. Kemmer K, Corless CL, Fletcher JA, et al. KIT mutations are common in testicular seminomas. Am J Pathol. 2004;164:305-313.
51. Wellcome Trust Sanger Institute. COSMIC: Catalogue Of Somatic Mutations In Cancer. http://www.sanger.ac.uk/perl/genetics/CGP/cosmic. Accessed June 20, 2007.
52. Bernard T, Gale RE, Linch DC. Analysis of granulocyte colony stimulating factor receptor isoforms, polymorphisms and mutations in normal haemopoietic cells and acute myeloid leukaemia blasts. Br J Haematol. 1996;93:527-533.
53. Carapeti M, Soede-Bobok A, Hochhaus A, et al. Rarity of dominant-negative mutations of the G-CSF receptor in patients with blast crisis of chronic myeloid leukemia or de novo acute leukemia. Leukemia. 1997;11:1005-1008.
54. Hermans MH, Antonissen C, Ward AC, Mayen AE, Ploemacher RE, Touw IP. Sustained receptor activation and hyperproliferation in response to granulocyte colony-stimulating factor (G-CSF) in mice with a severe congenital neutropenia/acute myeloid leukemia-derived mutation in the G-CSF receptor gene. J Exp Med. 1999;189:683-692.
55. Hunter MG, Avalos BR. Granulocyte colony-stimulating factor receptor mutations in severe congenital neutropenia transforming to acute myelogenous leukemia confer resistance to apoptosis and enhance cell survival. Blood. 2000;95:2132-2137.
56. Meshinchi S, Alonzo TA, Stirewalt DL, et al. Clinical implications of FLT3 mutations in pediatric AML. Blood. 2006;108:3654-3661.
57. Meshinchi S, Stirewalt DL, Alonzo TA, et al. Activating mutations of RTK/ras signal transduction pathway in pediatric acute myeloid leukemia. Blood. 2003;102:1474-1479.
58. Christiansen DH, Andersen MK, Desta F, Pedersen-Bjergaard J. Mutations of genes in the receptor tyrosine kinase (RTK)/RAS-BRAF signal transduction pathway in therapy-related myelodysplasia and acute myeloid leukemia. Leukemia. 2005;19:2232-2240.
59. Side LE, Curtiss NP, Teel K, et al. RAS, FLT3, and TP53 mutations in therapy-related myeloid malignancies with abnormalities of chromosomes 5 and 7. Genes Chromosomes Cancer. 2004;39:217-223.
60. Lorenzo F, Nishii K, Monma F, Kuwagata S, Usui E, Shiku H. Mutational analysis of the KIT gene in myelodysplastic syndrome (MDS) and MDS-derived leukemia. Leuk Res. 2006;30:1235-1239.
61. Germeshausen M, Kratz C, Ballmaier M, Shannon KM, Welte K. RAS and G-CSF receptor mutations are mutually exclusive in leukemogenesis in severe congential neutropenia [abstract]. Blood. 2005;106:860a. Abstract 3073.
62. Cazzaniga G, Dell'Oro MG, Mecucci C, et al. Nucleophosmin mutations in childhood acute myelogenous leukemia with normal karyotype. Blood. 2005;106:1419-1422.
63. Falini B, Mecucci C, Tiacci E, et al. Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. N Engl J Med. 2005; 352:254-266.