Corrigendum: The landscape of somatic mutations in Down syndrome–related myeloid disorders

Kenichi Yoshida, Tsutomu Toki, Yusuke Okuno, Rika Kanezaki, Yuichi Shiraishi, Aiko Sato-Otsubo, Masashi Sanada, Myoung-ja Park, Kiminori Terui, Hiromichi Suzuki, Ayana Kon, Yasunobu Nagata, Yusuke Sato, RuNan Wang, Norio Shiba, Kenichi Chiba, Hiroko Tanaka, Asahito Hama, Hideki Muramatsu, Daisuke Hasegawa, Kazuhiro Nakamura, Hirokazu Kanegane, Keiko Tsukamoto, Souichi Adachi, Kiyoshi Kawakami, Koji Kato, Ryosei Nishimura, Shai Izraeli, Yasuhide Hayashi, Satoru Miyano, Seiji Kojima, Etsuro Ito & Seishi Ogawa
Nat. Genet. **45**, 1293–1299 (2013); **published online 22 September 2013; corrected after print 30 October 2013**

In the version of this article initially published, the discussion of cited reference 52 should also have noted that the work "reported accumulation of additional somatic mutations (including single cases of *SMC3* and *EZH2* mutation) during progression from TAM to DS-AMKL." The error has been corrected in the HTML and PDF versions of the article.

NUP98-NSD1 Gene Fusion and Its Related Gene Expression Signature Are Strongly Associated with a Poor Prognosis in Pediatric Acute Myeloid Leukemia

Norio Shiba,[1,2] Hitoshi Ichikawa,[3] Tomohiko Taki,[4] Myoung-Ja Park,[1] Aoi Jo,[3] Sachiyo Mitani,[3] Tohru Kobayashi,[2] Akira Shimada,[5] Manabu Sotomatsu,[1] Hirokazu Arakawa,[2] Souichi Adachi,[6] Akio Tawa,[7] Keizo Horibe,[8] Masahiro Tsuchida,[9] Ryoji Hanada,[10] Ichiro Tsukimoto,[11] and Yasuhide Hayashi[1]*

[1]Department of Hematology/Oncology, Gunma Children's Medical Center, Shibukawa, Japan
[2]Department of Pediatrics, Gunma University Graduate School of Medicine, Maebashi, Japan
[3]Division of Genetics, National Cancer Center Research Institute, Tokyo, Japan
[4]Department of Molecular Diagnostics and Therapeutics, Kyoto Prefectural University of Medicine Graduate School of Medical Science, Kyoto, Japan
[5]Department of Pediatrics, Okayama University Graduate School of Medicine, Okayama, Japan
[6]Department of Human Health Sciences, Kyoto University Graduate School of Medicine, Kyoto, Japan
[7]Department of Pediatrics, National Hospital Organization Osaka National Hospital, Osaka, Japan
[8]Clinical Research Center, National Hospital Organization Nagoya Medical Center, Nagoya, Japan
[9]Department of Pediatrics, Ibaraki Children's Hospital, Ibaraki, Japan
[10]Division of Hematology/Oncology, Saitama Children's Medical Center, Saitama, Japan
[11]Department of First Pediatrics, Toho University School of Medicine, Tokyo, Japan

The cryptic t(5;11)(q35;p15.5) creates a fusion gene between the *NUP98* and *NSD1* genes. To ascertain the significance of this gene fusion, we explored its frequency, clinical impact, and gene expression pattern using DNA microarray in pediatric acute myeloid leukemia (AML) patients. *NUP98-NSD1* fusion transcripts were detected in 6 (4.8%) of 124 pediatric AML patients. Supervised hierarchical clustering analyses using probe sets that were differentially expressed in these patients detected a characteristic gene expression pattern, including 18 *NUP98-NSD1*-negative patients (*NUP98-NSD1*-like patients). In total, a *NUP98-NSD1*-related gene expression signature (*NUP98-NSD1* signature) was found in 19% (24/124) and in 58% (15/26) of cytogenetically normal cases. Their 4-year overall survival (OS) and event-free survival (EFS) were poor (33.3% in *NUP98-NSD1*-positive and 38.9% in *NUP98-NSD1*-like patients) compared with 100 *NUP98-NSD1* signature-negative patients (4-year OS: 86.0%, 4-year EFS: 72.0%). Interestingly, t(7;11)(p15;p15)/*NUP98-HOXA13*, t(6;11)(q27;q23)/MLL-*MLLT4* and t(6;9)(p22;q34)/*DEK-NUP214*, which are known as poor prognostic markers, were found in *NUP98-NSD1*-like patients. Furthermore, another type of *NUP98-NSD1* fusion transcript was identified by additional RT-PCR analyses using other primers in a *NUP98-NSD1*-like patient, revealing the significance of this signature to detect *NUP98-NSD1* gene fusions and to identify a new poor prognostic subgroup in AML. © 2013 Wiley Periodicals, Inc.

INTRODUCTION

Acute myeloid leukemia (AML) is a complex disease caused by mutations, epigenetic modifications, and deregulated expression of genes, leading to increased proliferation and decreased differentiation of hematopoietic progenitor cells (Frohling et al., 2005; Marcucci et al., 2011; Pui et al., 2011). Several important molecular markers have been discovered in AML that have not only helped to characterize better patients, but also to improve risk stratification (Marcucci et al., 2011; Pui et al., 2011). However, in a subset of AML patients, no prognosis-associated cytogenetic aberrations or mutations are known (Frohling et al., 2005; Marcucci et al., 2011; Pui et al., 2011). In

Additional Supporting Information may be found in the online version of this article.

Supported by: a grant for Cancer Research, a grant for Research on Children and Families, and Research on Intractable Diseases, Health and Labor Sciences Research Grants from the Ministry of Health, Labor, and Welfare of Japan, Grants-in-Aid for Scientific Research (B, C) and Exploratory Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan, the Program for Promotion of Fundamental Studies in Health Sciences of the National Institute of Biomedical Innovation (NiBio) of Japan, and a Research grant for Gunma Prefectural Hospitals.

*Correspondence to: Yasuhide Hayashi, Department of Hematology/Oncology, Gunma Children's Medical Center, 779, Shimohakoda, Hokkitsu, Shibukawa, 377-8577 Gunma, Japan. E-mail: hayashiy-tky@umin.ac.jp

Received 21 December 2012; Accepted 15 March 2013

DOI 10.1002/gcc.22064

Published online 30 April 2013 in Wiley Online Library (wileyonlinelibrary.com).

hematological malignancies, 11p15 translocations involving the nucleoporin 98-kDa (NUP98) protein gene are relatively rare. This notwithstanding, more than 20 different chromosomal rearrangements have been identified (Romana et al., 2006). These translocations fuse *NUP98* with respective partner genes, including many homeobox genes and nuclear non-homeobox genes (Arai et al., 1997; Taketani et al., 2002a,b,c). Although the *NUP98* fusion genes are rare, they have provided valuable information regarding the role of homeobox proteins in leukemogenesis (Nakamura et al., 1996; Nakamura, 2005).

The cryptic t(5;11)(q35;p15.5), which is frequently accompanied by deletion of the long arm of chromosome 5, del(5q), creates a fusion gene between *NUP98* and the nuclear receptor-binding SET-domain protein 1 (*NSD1*) gene (Jaju et al., 2001). This fusion gene has mainly been identified in pediatric AML, by the use of fluorescence in situ hybridization (FISH) with subtelomeric probes (Jaju et al., 2001; Brown et al., 2002; Panarello et al., 2002; Cerveira et al., 2003), it is rare in adult AML (Brown et al., 2002; Casas et al., 2003; Nebral et al., 2005; Walter et al., 2009). Recently, the *NUP98-NSD1* gene fusion, identified by high-resolution genome-wide copy number analysis and reverse transcription (RT)-PCR in pediatric and adult AML patients, was associated with poor prognosis (Hollink et al., 2011). To increase our understanding of this gene fusion, we explored the frequency, clinical significance, and gene expression pattern of *NUP98-NSD1* using DNA microarray in pediatric AML patients.

MATERIALS AND METHODS

Patients and Samples

From January 2000 to December 2002, 318 patients were diagnosed with de novo AML. The diagnosis of AML was based on the French-American-British (FAB) classification, and cytogenetic analysis was performed using conventional G-banding. Of these patients, samples from 124 patients with known mutation status and gene expression profiling data were available, including 10 patients with FAB-M3 and six patients with Down syndrome who were treated on different treatment protocols (Kudo et al., 2007; Tsukimoto et al., 2009; Imaizumi et al., 2011). Age and initial white blood cell (WBC) count were higher, patients with t(8;21) were more frequent, and M7 patients with Down syndrome were fewer in the present study cohort than in the non-analyzed patients (Supporting Information Table S1). There were no significant differences in survival between these two groups [4-year overall survival (OS): 76.6% versus 79.9% and 4-year event-free survival (EFS): 65.3% versus 71.1%]. Informed consent was obtained from the patients or the patients' parents, according to guidelines based on the tenets of the revised Helsinki protocol. The institutional review boards of Gunma Children's Medical Center and National Cancer Center approved this project.

Reverse Transcription-PCR and Sequence Analysis

Total RNA extracted from leukemic cells at diagnosis was reverse transcribed to cDNA with a cDNA Synthesis Kit (GE Healthcare, Tokyo, Japan). PCR was performed with AmpliTaq Gold DNA polymerase (Applied Biosystems, Branchburg, NJ), using a DNA thermal cycler (Applied Biosystems). For the detection of *NUP98-NSD1* and the reciprocal fusion transcript, *NUP98*−5F and *NSD1*−1R, and *NSD1*−2F and *NUP98*−6R were used, respectively (Supporting Information Table S2) (Jaju et al., 2001; Brown et al., 2002). PCR conditions were as follows: initial denaturation at 94°C (9 min), 40 cycles of 96°C (45 sec), 58°C (45 sec), and 72°C (1 min), followed by final elongation at 72°C (7 min). For sequencing, PCR products were amplified using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) under the following conditions: 95°C (2 min) followed by 25 cycles of 95°C (10 sec), 50°C (5 sec), and 60°C (4 min). Direct sequencing was performed using an ABI PRISM 310 Genetic Analyzer (Applied Biosystems). To confirm and identify the gene fusions precisely, we used various primer sets for RT-PCR followed by direct sequencing (Supporting Information Table S2). Mutations of *NPM1* were also examined as previously reported (Döhner et al., 2005). Mutation analyses of the *DNMT3A*, *FLT3*, *MLL*, *KIT*, *NRAS*, *KRAS*, and *WT1* genes have been reported previously (Shimada et al., 2006, 2008; Sano et al., 2012; Shiba et al., 2012).

Microarray Analysis

Gene expression profiling data for the 124 patients (Gene Expression Omnibus accession number, GSE35784) were obtained and analyzed as follows. Total RNA was re-purified using the RNeasy MinElute cleanup kit (Qiagen, Hilden,

Germany), and the integrity of the purified RNA was confirmed using a 2100 Bioanalyzer and an RNA 6000 Nano LabChip kit (Agilent Technologies, Santa Clara, CA). The DNA microarray used was a Human Genome U133 plus 2.0 array (Affymetrix, Santa Clara, CA). Target cRNA was prepared from 20 ng purified RNA with a two-cycle cDNA synthesis kit and 3′-amplification reagents for IVT labeling (Affymetrix). Hybridization to the microarrays, washing, and staining with the antibody amplification procedure and scanning were performed according to the manufacturer's instructions. Using the GeneChip Operating Software version 1.4 (Affymetrix), the scanned image data were processed and the expression value (signal) of each probe set was calculated. The signal values were normalized so that the mean in each experiment was set at 100 to adjust for minor differences between the experiments. Statistical analyses and fold change calculations were performed using expression values that were log-transformed after addition of 10 to reduce adverse effects caused by noise at low expression levels (Ichikawa et al., 2006; Jo et al., 2009). To identify differentially expressed probe sets between *NUP98-NSD1*-positive and -negative patients, *P*-values in Student's *t*-test and fold change values were used. Hierarchical clustering analysis was performed using the Cluster and Tree View software (Eisen et al., 1998). For this analysis, log-transformed expression values were normalized for each probe set by subtracting the mean, and uncentered correlation metric and complete linkage clustering methods were used. Unsupervised clustering analyses was performed and visualized as previously described (Ichikawa et al., 2006).

Statistical Analysis

All analyses were carried out using the SPSS statistical package program (version 18.0J; SPSS Tokyo, Japan). Survival distributions were assessed using the Kaplan-Meier method and the differences were compared using the log-rank test. EFS and OS were defined as the times from diagnosis to event (relapse or death of any cause) and from diagnosis to death from any cause, respectively. Statistical analyses were performed using Fisher's exact test for categorical variables and Mann-Whitney's U test for continuous variables. For all analyses, the *P*-values were two-tailed, and $P < 0.05$ was considered significant. To identify independent predictors of poor prognosis, multivariate Cox regression analysis was performed using factors that were significant in frequencies between *NUP98-NSD1* signature positive and negative cases.

RESULTS

Detection of *NUP98-NSD1*-Positive Patients

We identified the *NUP98-NSD1* fusion transcript in 6 (4.8%) out of 124 Japanese pediatric AML patients. The reciprocal fusion transcript *NSD1-NUP98* was detected in five of the patients (Supporting Information Fig. S1A). Sequence analysis of the PCR products confirmed that *NUP98* and *NSD1* were fused in-frame, joining *NUP98* exon 12 with *NSD1* exon 6 (Supporting Information Fig. S1B) (Jaju et al., 2001; Brown et al., 2002; Panarello et al., 2002; Casas et al., 2003; Cerveira et al., 2003; Hollink et al., 2011). The reciprocal fusion products were also fused in-frame, joining *NSD1* exon 5 with *NUP98* exon 13 (Supporting Information Fig. S1C).

Detection of *NUP98-NSD1*-Like Patients by Gene Expression Analysis

To investigate the gene expression pattern characteristic of *NUP98-NSD1*-positive patients, we used gene expression profiling data on the 124 patients. In unsupervised hierarchical clustering analyses (Supporting Information Fig. S2), *NUP98-NSD1*-positive patients did not form a cluster, whereas patients with t(8;21)(q22;q22), t(15;17)(q22;q21), and inv(16)(p13;q22) formed unique clusters. Comparing the 6 *NUP98-NSD1*-positive patients with the other 118 patients resulted in the identification of 87 differentially expressed probe sets ($P < 0.001$, fold change ≥ 2) (Supporting Information Table S3). Fifty-one probe sets, including several *HOX* genes, were overexpressed, and 36 probe sets were underexpressed. Interestingly, in supervised hierarchical clustering analysis using those 87 probe sets, a relatively large cluster including 18 *NUP98-NSD1*-negative patients was found, not only a cluster of the 6 *NUP98-NSD1*-positive patients (Fig. 1). These 24 patients also formed a single cluster in similar supervised clustering analyses in which probe sets were selected by slightly different criteria (Supporting Information Fig. S3). The robustness in clustering suggests that these 24 patients have certain common features and that they constitute a distinct subtype. Thus, we termed this characteristic gene expression pattern the *NUP98-NSD1* signature. We also designated the 18 patients lacking the *NUP98-NSD1* gene fusion but displaying the *NUP98-NSD1* signature as *NUP98-NSD1*-like AML patients. Of the 24 patients displaying the *NUP98-NSD1* signature, a common overexpression of 3 *HOXB* genes (*HOXB3*,

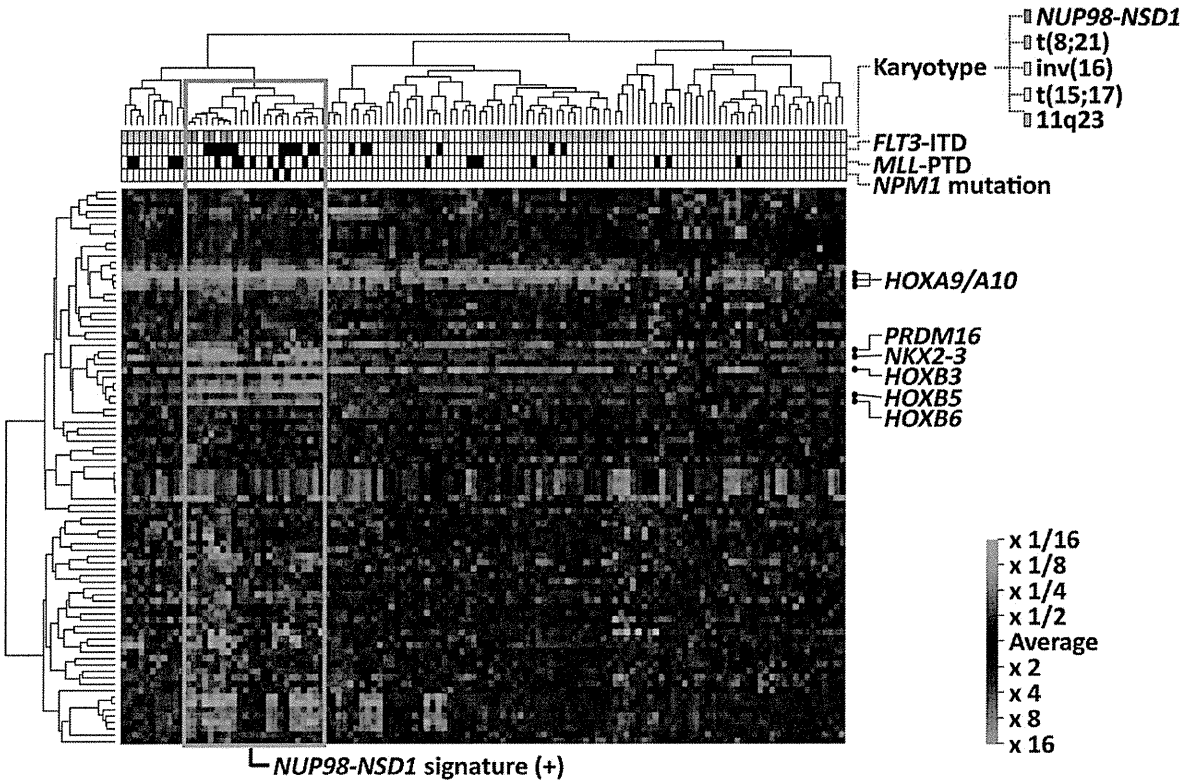

Figure 1. Supervised hierarchical clustering analysis of the 124 patients using differentially expressed genes between *NUP98-NSD1*-positive and -negative patients. Two-dimensional hierarchical clustering analysis of the 124 patients was performed using 87 probe sets that were differentially expressed between the *NUP98-NSD1*-positive and -negative patients (Table S3). Each column represents a patient and each row represents a probe set. The karyotype and the *FLT3*-ITD, *MLL*-PTD, and *NPM1* mutation status of each patient are indicated. Relative expression levels normalized to the average for each probe set are indicated by color, where red and green represent high and low expressions, respectively.

HOXB5, and *HOXB6*), another homeobox gene *NKX2–3*, a zinc finger transcription factor gene *PRDM16*, and a noncoding RNA gene *LOC404266* was particularly marked (Fig. 1). Although *HOXA9* and *HOXA10* were also overexpressed in these patients, their overexpression was shared by many patients with 11q23 abnormalities (Fig. 1).

Clinical and Molecular Characteristics of *NUP98-NSD1*-Positive and *NUP98-NSD1*-Like Patients

We compared the clinical and molecular features of the *NUP98-NSD1*-positive and *NUP98-NSD1*-like patients; there were no significant differences in terms of age, initial WBC count, gender, FAB subtypes, and cytogenetics (Table 1). Frequent gene mutations were also detected in both the *NUP98-NSD1*-positive and -like patients. All six *NUP98-NSD1*-positive patients and 16 *NUP98-NSD1*-like patients had one or several of the mutations analyzed (Table 2). *NUP98-NSD1*-positive patients frequently had class I aberrations, such as *FLT3*-ITD (4/6), and *NRAS* (1/6), *KRAS* (1/6), and *KIT* (1/6) mutations; however, no class II aberrations (*NPM1* or *MLL*-PTD) were detected in *NUP98-NSD1*-positive patients. On the other hand, *NUP98-NSD1*-like patients frequently had class II aberrations [*MLL*-PTD (7/18) and *NPM1* (3/18)] in addition to class I aberrations [*FLT3*-ITD (8/18), *KRAS* (3/18), and *KIT* (2/18)] (Table 2). Furthermore, we identified *DEK-NUP214*, *NUP98-HOXA13* and *MLL-MLLT4* fusion genes, which were generated by t(6;9)(p22;q34), t(1;7;11)(q32;p15;p15), and t(6;11)(q27;q23), respectively, in *NUP98-NSD1*-like patients (Table 2). In addition, we identified another type of *NUP98-NSD1* fusion transcript using another forward primer of *NUP98*-F', located on the 5′ side further from the known *NUP98* junction point (Supporting Information Table S2), in a *NUP98-NSD1*-like patient (Supporting Information Fig. S4).

Clinical and Molecular Characteristics of *NUP98-NSD1* Signature-Positive Patients

When compared with 100 *NUP98-NSD1* signature-negative patients, age (median 8.8 years vs.

TABLE 1. Clinical and Molecular Characteristics of the *NUP98-NSD1*-Positive and *NUP98-NSD1*-Like Patients

	NUP98-NSD1-positive	*NUP98-NSD1*-like	P-value
Total (n = 24)	6	18	
Age (range), y	7.2 (2–15)	9.3 (0–15)	0.28
Mean WBC ($\times 10^9$/l) (range)	115.9 (9.0–329.0)	60.5 (2.3–255.0)	0.31
Gender			0.34
Male	5 (83.3%)	9 (50.0%)	
Female	1 (16.7%)	9 (50.0%)	
FAB subtype			
M0	0 (0.0%)	1 (5.6%)	
M1	1 (16.7%)	4 (22.2%)	
M2	0 (0.0%)	2 (11.1%)	
M3	0 (0.0%)	0 (0.0%)	
M4	2 (33.3%)	6 (33.3%)	
M5	2 (33.3%)	4 (22.2%)	
M6	0 (0.0%)	0 (0.0%)	
M7	0 (0.0%)	1 (5.6%)	
Unclassified	1 (16.7%)	0 (0.0%)	
Cytogenetics			
Normal	4 (66.7%)	11 (61.1%)	1.00
t(8;21)(q22;q22)	0 (0.0%)	0 (0.0%)	1.00
t(15;17)(q22;q12)	0 (0.0%)	0 (0.0%)	1.00
inv(16)(p13q22)	0 (0.0%)	0 (0.0%)	1.00
abnormal 11q23	0 (0.0%)	2 (11.1%)	1.00
del(9q)	2 (33.3%)	0 (0.0%)	0.054
Others	0 (0.0%)	5 (27.8%)	0.28
Gene mutations			
FLT3-ITD	4 (66.7%)	8 (44.4%)	0.64
DNMT3A	0 (0.0%)	0 (0.0%)	1.00
NPM1	0 (0.0%)	3 (16.7%)	0.55
KIT	1 (16.7%)	2 (11.1%)	1.00
NRAS	1 (16.7%)	0 (0.0%)	0.25
KRAS	1 (16.7%)	3 (16.7%)	1.00
WT1	2 (33.3%)	3 (16.7%)	0.57
MLL-PTD	0 (0.0%)	7 (38.9%)	0.13

Abbreviations: y, years; WBC, white blood cell count; FAB,French-American-British subtype.

6.7 years; $P = 0.069$) and initial WBC count (median 74.4 $\times$ 10^9/ L vs. 49.7 $\times$ 10^9/ L; $P = 0.025$) were higher in the 24 *NUP98-NSD1* signature-positive patients. They frequently had a normal karyotype (62.5%) or del(9q) (8.3%), but did not have the favorable chromosomal translocations t(8;21), t(15;17), and inv(16) (Tables 2 and 3). The frequencies of the M4 and M5 subtypes (14/24; $P < 0.001$) were higher than in *NUP98-NSD1* signature-negative patients.

Prognosis of *NUP98-NSD1*-Positive and *NUP98-NSD1*-Like Patients

Both the 4-year OS and EFS were 33.3% in *NUP98-NSD1*-positive patients and 38.9% in *NUP98-NSD1*-like patients, which is significantly worse than for those with *NUP98-NSD1* signature-negative patients (86.0% in OS and 72.0% in EFS; Figs. 2A and 2B). Five of the six *NUP98-NSD1*-positive and 8 of the 18 *NUP98-NSD1*-like patients received allogeneic-stem cell transplantation (allo-SCT) (Table 2). Of the *NUP98-NSD1*-positive patients, only two of the three who underwent allo-SCT in first CR are still alive and, in the *NUP98-NSD1*-like patients, all 4 who received SCT in first CR and 3 of 10 who were treated with only chemotherapy are still alive without relapse. All 12 relapsed patients and two who did not achieve CR died (Table 2). Multivariate Cox regression analysis of OS was used to construct a model including the *NUP98-NSD1* signature, *FLT3*-ITD, *WT1*, *NPM1*, t(8;21), del(9q), and initial WBC, which were statistically significant in univariate analysis. In this model, the *NUP98-NSD1* signature and *FLT3*-ITD were independent poor prognostic factors (Table 4). Among the 24 patients displaying the *NUP98-NSD1* signature, the outcome of the 12 *FLT3*-ITD-positive patients was worse than that of the 12 *FLT3*-ITD-negative

TABLE 2. Individual Characteristics of the *NUP98-NSD1* Signature-Positive Patients

ID	Age (y)	Sex	FAB	Cytogenetic aberrations	*NUP98-NSD1*	Risk	CR	Relapse	SCT	Class I mutations	Class II mutations	Other mutations	Survival (m)
A106	5	F	M1	50,XX,+6,+8,del (9q?),+21,+22	+	Off study	+	–	CR 1	*FLT3*-ITD	–	–	63+
A188	10	M	U/C	Normal	+	High	+	+	CR 2	*FLT3*-ITD	–	*WT1*	21
A282	2	M	M5	46,XY,del(9)(q13q22)	+	Intermediate	+	–	CR 1	*FLT3*-ITD	–	–	43+
A325	5	M	M5	Normal	+	High	+	+	CR 2	*FLT3*-ITD	–	–	15
A333	15	M	M4	Normal	+	Intermediate	+	+	–	*NRAS*	–	–	31
A335	6	M	M4	Normal	+	Intermediate	+	–	CR 1	*KRAS*+*KIT*	–	*WT1*	7[a]
A044	6	M	M2	Normal	–	Intermediate	+	+	–	–	–	–	30
A059	12	F	M7	Normal	–	Intermediate	+	+	CR 2	*FLT3*-ITD	–	*WT1*	16
A089	9	F	M4	Normal	–	Off study	–	–	Non-CR	*FLT3*-ITD	*MLL*-PTD	–	25
A154	8	F	M0	Normal	–	High	+	+	CR 2	*FLT3*-ITD	–	–	20
A167	2	M	M5	Normal	–	High	+	–	CR 1	–	*MLL*-PTD	–	55+
A171	12	M	M1	46,XY,t(1;7;11) (q32;p15;p15)	–	Intermediate	+	–	CR 1	–	*MLL*-PTD	*WT1*	55+
A173	13	M	M1	Normal	–	Intermediate	+	–	CR 1	*FLT3*-ITD +*KIT*	*MLL*-PTD	*WT1*	54+
A199	10	F	M5	46,XX,add(10)(p11.2), del(11)(q13q23)	–	Intermediate	+	+	–	*KIT*+*KRAS*	–	–	41
A202	15	F	M4	Normal	–	Intermediate	+	–	–	–	*NPM1*	–	50+
A211	14	F	M1	47,XX,+8	+[b]	Intermediate	+	+	CR 2	*FLT3*-ITD	*MLL*-PTD	–	14
A234	13	M	M4	46,XY, t(6;11) (q27;q23)	–	Intermediate	+	+	–	*KRAS*	*MLL*-PTD	–	29
A243	5	F	M1	Normal	–	Intermediate	+	+	–	*FLT3*-ITD	–	–	8
A245	11	F	M4	Normal	–	Intermediate	+	–	–	*FLT3*-TK (D835)	*NPM1*	–	46+
A249	0	M	M4	46,XY,t(5;6) (q33;q22)	–	Low	+	–	–	–	–	–	46+
A259	5	M	M5	46,XY,t(6;9) (p23;q34)	–	Intermediate	+	–	CR 1	*FLT3*-TK (D835) + *KRAS*	–	–	48+
A297	6	M	M5	Normal	–	Off study	–	–	–	*FLT3*-ITD	*NPM1*	–	6
A299	13	F	M4	Normal	–	Intermediate	+	+	–	–	*MLL*-PTD	–	27
A355	13	M	M2	46,XY,t(15;17) (q13;q11)	–	Intermediate	+	+	–	*FLT3*-ITD	–	–	15

Abbreviations: y, years; F, female; M, male; FAB, French-American-British subtype; U/C, unclassified; CR, complete remission; SCT, stem cell transplantation; CR 1, first CR; CR 2, second CR; m, month; +, alive. The *NUP98-NSD1*-positive cases are indicated in bold type.

[a]A335 died of severe GVHD and acute pneumonia.

[b]Another type of *NUP98-NSD1* fusion transcript was identified by additional RT-PCR using another forward primer.

TABLE 3. Clinical and Molecular Characteristics of the *NUP98-NSD1* Signature Positive and Negative Cases

	NUP98-NSD1 signature (+)	*NUP98-NSD1* signature (−)	*P*-value
Total ($n = 124$)	24	100	
Age (range), y	8.8 (0−15)	6.7 (0−15)	0.069
Mean WBC (-10^9/l) (range)	74.4 (2.3–329.0)	49.7 (1.0–440.0)	0.025
Gender			1.00
Male	14 (58.3%)	57 (57.0%)	
Female	10 (41.7%)	43 (43.0%)	
FAB subtype[a]			
M0	1 (4.2%)	3 (3.0%)	
M1	5 (20.8%)	14 (14.0%)	
M2	2 (8.3%)	39 (39.0%)	
M3	0 (0.0%)	10 (10.0%)	
M4	8 (33.3%)	9 (9.0%)	
M5	6 (25.0%)	11 (11.0%)	
M6	0 (0.0%)	1 (1.0%)	
M7	1(4.2%)	12 (12.0%)	
Unclassified	1 (4.2%)	1 (1.0%)	
Cytogenetics			
Normal	15 (62.5%)	11 (11.0%)	<0.001
t(8;21)(q22;q22)	0 (0.0%)	41 (41.0%)	<0.001
t(15;17)(q22;q12)	0 (0.0%)	10 (10.0%)	0.21
inv(16)(p13q22)	0 (0.0%)	6 (6.0%)	0.60
abnormal 11q23	2 (8.3%)	9 (9.0%)	1.00
del(9q)	2 (8.3%)	0 (0.0%)	0.036
Others	5 (20.8%)	23 (23.0%)	1.00
Gene mutations			
FLT3-ITD	12 (50.0%)	6 (6.0%)	<0.001
DNMT3A	0 (0.0%)	0 (0.0%)	1.00
NPM1	3 (12.5%)	0 (0.0%)	0.007
KIT	3 (12.5%)	12 (12.0%)	1.00
NRAS	1 (4.2%)	9 (9.0%)	0.69
KRAS	4 (16.7%)	9 (9.0%)	0.28
WT1	5 (20.8%)	6 (6.0%)	0.037
MLL-PTD	7 (29.2%)	14 (14.0%)	0.12

[a]*NUP98-NSD1* signature-positive patients were significantly associated with the M4 and M5 subtypes (14/24; $P < 0.001$) when compared with *NUP98-NSD1* signature-negative patients.

patients (4-year OS: 25% vs. 50%), although the difference was not significant ($P = 0.400$).

We further analyzed the prognostic significance of *NUP98-NSD1*-positive patients and *NUP98-NSD1*-like patients other than those with t(15;17) and Down syndrome because they represent distinct AML entities. All 10 patients with t(15;17) and all six patients with Down syndrome were *NUP98-NSD1* signature negative. The outcome of the six patients with *NUP98-NSD1* gene fusion was significantly worse than that of the *NUP98-NSD1* signature-negative patients in OS ($P < 0.001$; 4-year OS: 33.3% vs. 85.7%; Fig. 2C) and in EFS ($P = 0.022$; 4-year EFS: 33.3% versus 70.2%; Fig. 2D). Furthermore, the outcome of the 18 *NUP98-NSD1*-like AML patients was significantly worse than that of the *NUP98-NSD1* signature-negative patients in OS ($P < 0.001$; 4-year OS: 38.9% versus 85.7%; Fig. 2C) and in EFS ($P = 0.002$; 4-year EFS: 38.9% versus 70.2%; Fig. 2D).

DISCUSSION

In this study, we found 24 patients with *NUP98-NSD1*-related gene expression signature, including six with the *NUP98-NSD1* gene fusion (*NUP98-NSD1*-positive) and 18 without (*NUP98-NSD1*-like). This signature represented 19% (24/124) of all pediatric AML patients and 58% (15/26) of all cytogenetically normal cases (Fig. 3). Our results also revealed that the *NUP98-NSD1* signature, irrespective of the presence of the *NUP98-NSD1* fusion, is a novel poor prognostic factor in AML.

The relationship between *NUP98-NSD1*-positive AML and *NUP98-NSD1*-like AML resembles that of *BCR-ABL*-positive acute lymphoblastic

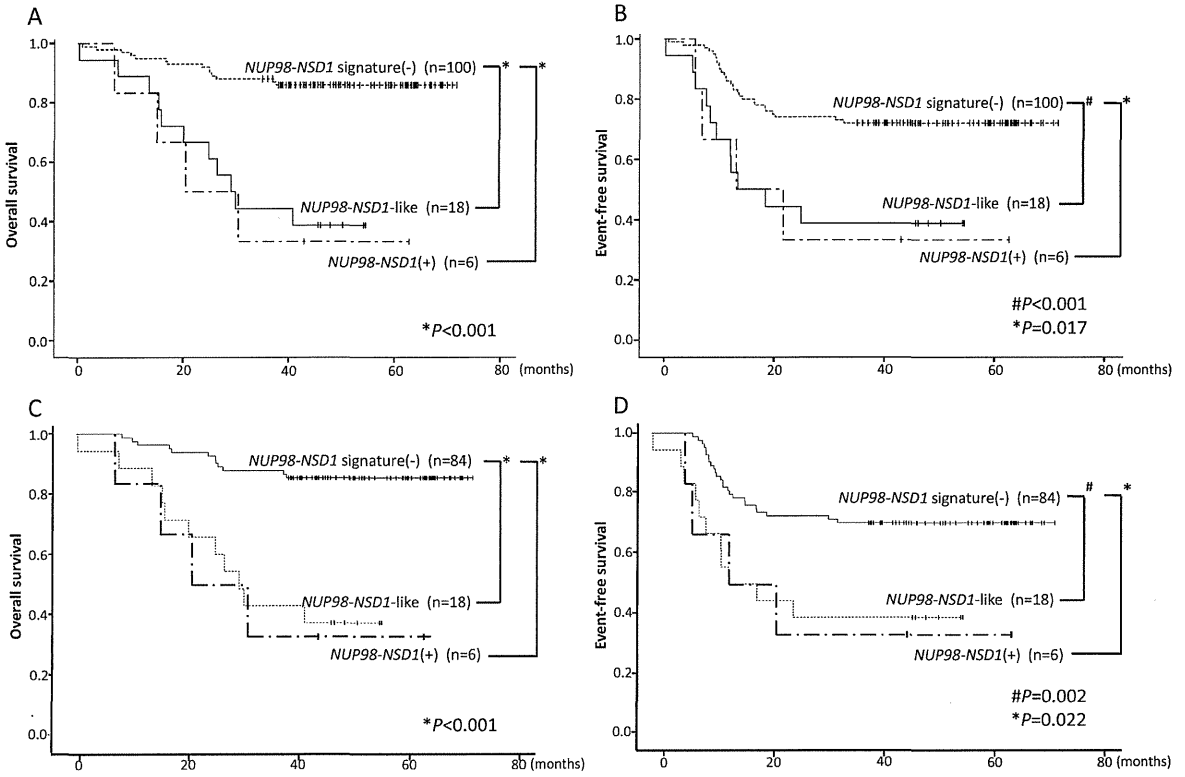

Figure 2. Survival based on *NUP98-NSD1* status by Kaplan-Meier method. Comparison of *NUP98-NSD1*-positive, *NUP98-NSD1*-like, and *NUP98-NSD1* signature-negative patients as regards OS (A) and EFS (B). Comparison of *NUP98-NSD1*-positive, *NUP98-NSD1*-like patients, and *NUP98-NSD1* signature-negative patients other than t(15;17) and Down syndrome, as regards OS (C) and EFS (D).

TABLE 4. Multivariate Analysis of OS of the Pediatric AML Patients

	P-value	SE	HR	Lower CI	Upper CI
NUP98-NSD1 signature	0.005	0.38	2.89	1.38	6.02
FLT3-ITD	0.005	0.40	3.06	1.41	6.63
WT1	0.20	0.54	0.50	0.18	1.43
NPM1	0.49	1.05	0.48	0.06	3.77
t(8;21)	0.18	0.46	0.54	0.22	1.32
del(9q)	0.97	371.69	<0.001	<0.001	–
Initial WBC	0.10	0.002	1.00	1.00	1.01

Abbreviations: SE, standard error; HR, hazard ratio; CI, confidence interval.

leukemia (ALL) and *BCR-ABL*-like ALL (Den Boer et al., 2009; Mullighan et al., 2009). In their gene expression analyses, a significant number of genetically unclassified B-cell precursor ALL patients clustered together with *BCR-ABL*-positive patients, and these *BCR-ABL*-like patients had a poor prognosis, similar to *BCR-ABL*-positive patients. Recently, it was reported that some *BCR-ABL*-like patients express tyrosine kinase/cytokine receptor gene-related fusion genes, such as *NUP214-ABL1*, *EBF1-PDGFRB*, *BCR-JAK2*, and *STRN3-JAK2* (Roberts et al., 2012). Thus, as potentially important fusions might be detected in *NUP98-NSD1*-like patients, we performed RT-PCR using various primer sets for detecting *NUP98-HOXA9*, *NUP98-HOXA11*, *NUP98-HOXA13*, *NUP98-TOP1*, *NUP98-PRRX1*, *NUP98-DDX10*, *NUP98-MLL*, *NUP98-NSD3*, *DEK-NUP214*, *MLL-MLLT4*, and other junction points of *NUP98-NSD1* (Supporting Information Table S2). As a result, t(6;9)/*DEK-NUP214*, t(1;7;11)/*NUP98-HOXA13*, and t(6;11)/*MLL-MLLT4*, which are well-known poor prognostic markers, were found in some *NUP98-NSD1*-like patients in our study (Table 2). Furthermore, another type of *NUP98-NSD1* fusion transcript was identified in a *NUP98-NSD1*-like patient (Supporting Information Fig. S4). This fusion transcript has previously been reported in an adult patient with refractory anemia with excess blasts (La Starza et al., 2004). It is likely that other *NUP98-NSD1*-like patients also have unknown fusion genes with the same functions as *NUP98-NSD1* gene fusion.

It has been reported that hematological malignancies with *NUP98*-fusion genes are strongly

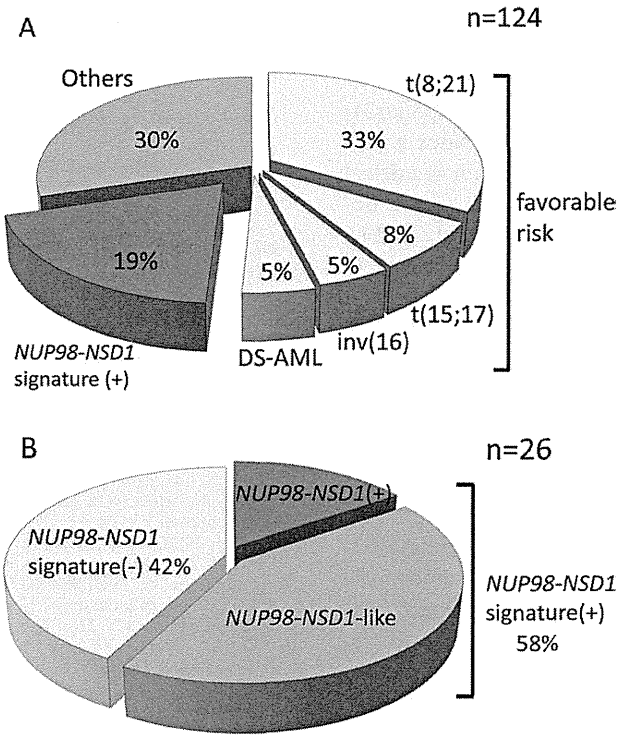

Figure 3. The *NUP98-NSD1* signature is frequent in pediatric AML. Frequencies of *NUP98-NSD1* signature-positive patients in the total pediatric AML cohort (A) and in cytogenetically normal AML (B).

associated with class I mutations (Chou et al., 2009; Taketani et al., 2010). Interestingly, *NUP98-NSD1*-positive patients had only class I aberrations, except for *WT1* mutations, although *NUP98-NSD1*-like patients frequently had Class II aberrations (*NPM1* mutations, *MLL*-PTD, and the *NUP98-HOXA13*, *MLL-MLLT4*, and *DEK-NUP214* fusion genes) in addition to class I aberrations. Because the *NUP98-NSD1* fusion is considered to act as a Class II aberration, no additional Class II aberrations might be necessary for leukemogenesis in *NUP98-NSD1*-positive patients.

The frequency of *FLT3*-ITD was higher than other mutations in both *NUP98-NSD1*-positive and *NUP98-NSD1*-like patients. Many studies have shown that *FLT3*-ITD has a negative impact on outcome in both adult and pediatric AML patients. In fact, in this study, *FLT3*-ITD was an independent poor prognostic factor in addition to the *NUP98-NSD1* signature (Table 4). This suggests that the *NUP98-NSD1* signature and *FLT3*-ITD play a key potential role in AML patients with poor prognosis.

Among the 100 *NUP98-NSD1* signature-negative patients, 23 patients relapsed and of those patients, nine died. On the other hand, of the 24 *NUP98-NSD1* signature-positive patients, 13 patients relapsed and all of those died. Except for 2 patients who did not achieve CR, 12 of the 22 *NUP98-NSD1* signature-positive patients received allo-SCT. Six of the seven patients (86%) who received allo-SCT in first CR were still alive without relapse; however, all five patients who received allo-SCT in second CR died. Only three patients were alive among the 10 patients who were treated with chemotherapy alone. Thus, allo-SCT is recommended in first CR of *NUP98-NSD1* signature-positive patients.

In the *NUP98-NSD1* signature-positive patients, the *HOXA9*, *HOXA10*, *HOXB3*, *HOXB5*, and *HOXB6* genes were up-regulated (Fig. 1). *NUP98* is frequently fused to homeobox genes, and some *NUP98*-homeobox and *NUP98*-non-homeobox fusion genes were revealed to activate *HOXA* cluster genes in hematopoietic cells (Gough et al., 2011). Overexpression of some *HOX* genes is known to enhance the self-renewal of hematopoietic stem and progenitor cells and to perturb differentiation (Grier et al., 2005). It is expected that the aberrant expression of *HOX* genes plays an important role in the leukemogenesis of AML displaying the *NUP98-NSD1* signature. In addition,

two transcription factor genes, *PRDM16* and *NKX2–3*, were markedly up-regulated (Fig. 1). *PRDM16* (also known as *MEL1*) was originally isolated as a translocated gene in t(1;3)(p36;q21) AML (Mochizuki et al., 2000) and encodes a zinc finger protein with a PR domain, which is critical for the establishment and maintenance of the hematopoietic stem cell pool (Agulio et al., 2011). *NKX2–3* is an NKX family homeobox gene (Pabst et al., 1999), whose involvement in leukemogenesis has not been reported; however, its highly homologous paralog *NKX2–5* is rearranged and ectopically expressed in T-cell ALL with t(5;14)(q35;q32) and t(5;14)(q35;q11.2) (Nagel et al., 2003; Przybylski et al., 2006). Taken together, these data suggest that both *PRDM16* and *NKX2–3* play an important role in leukemogenesis.

Our results indicate that the *NUP98-NSD1*-related gene expression signature is associated with a poor outcome in addition to the *NUP98-NSD1* gene fusion in pediatric AML. Most of the patients displaying the *NUP98-NSD1* signature were classified into an intermediate risk group, but their unfavorable outcome suggests that a high-risk group is a more suitable stratification. Although further investigations are necessary, we believe that our work contributes to improving the risk stratification of pediatric AML.

ACKNOWLEDGMENTS

The authors would like to thank the following Japanese childhood AML Cooperative Study Group institutions and investigators that participated in the study: Akira Morimoto, Department of Pediatrics, Kyoto Prefectural University of Medicine; Ryoji Kobayashi, Department of Pediatrics, Hokkaido University School of Medicine; Hiromasa Yabe, Department of Pediatrics, Tokai University School of Medicine; Kazuko Hamamoto, Department of Pediatrics, Hiroshima Red Cross Hospital; Shigeru Tsuchiya, Department of Pediatric Oncology, Institute of Development, Aging and Cancer, Tohoku University; Yuichi Akiyama, Department of Pediatrics, National Hospital Organization Kyoto Medical Center; Hisato Kigasawa, Department of Hematology, Kanagawa Children's Medical Center; Akira Ohara, Department of First Pediatrics, Toho University School of Medicine; Hideki Nakayama, Department of Pediatrics, Hamanomachi Hospital; Kazuko Kudo, Department of Pediatrics, Nagoya University Graduate School of Medicine; and Masue Imaizumi, Department of Hematology/Oncology, Miyagi Prefectural Children's Hospital.

REFERENCES

Agulio F, Avagyan S, Labar A, Sevilla A, Lee DF, Kumar P, Lemischka IR, Zhou BY, Snoeck HW. 2011. Prdm16 is a physiologic regulator of hematopoietic stem cells. Blood 117:5057–5066.

Arai Y, Hosoda F, Kobayashi H, Kawakami K, Takahara J. 1997. The inv(11)(p15q22) chromosome translocation of de novo and therapy-related myeloid malignancies results in fusion of the nucleoporin gene, NUP98, with the putative RNA helicase case, DDX10. Blood 89:3936–3944.

Brown J, Jawad M, Twigg SR, Saracoglu K, Sauerbrey A, Thomas AE, Eils R, Harbott J, Kearney L. 2002. A cryptic t(5;11)(q35;p15.5) in 2 children with acute myeloid leukemaia with apparently normal karyotypes, identified by a multiplex fluorescence in situ hybridization telomere assay. Blood 99:2526–2531.

Casas S, Aventín A, Nomdedéu J, Sierra J. 2003. Cryptic t(5;11)(q35;p15.5) in adult de novo acute myelocytic leukemia with normal karyotype. Cancer Genet Cytogenet 145:183.

Cerveira N, Correia C, Dória S, Bizarro S, Rocha P, Gomes P, Torres L, Norton L, Borges BS, Castedo S, Teixeira MR. 2003. Frequency of NUP98-NSD1 fusion transcript in childhood acute myeloid leukaemia. Leukemia 17:2244–2247.

Chou WC, Chen CY, Hou HA, Lin LI, Tang JL, Yao M, Tsay W, Ko BS, Wu SJ, Huang SY, Hsu SC, Chen YC, Huang YN, Tseng MH, Huang CF, Tien HF. 2009. Acute myeloid leukemia bearing t(7;11)(p15;p15) is a distinct cytogenetic entity with poor outcome and a distinct mutation profile: comparative analysis of 493 adult patients. Leukemia 23:1303–1310.

Den Boer ML, van Slegtenhorst M, De Menezes RX, Cheok MH, Buijs-Gladdines JG, Peters ST, Van Zutven LJ, Beverloo HB, Van der Spek PJ, Escherich G, Horstmann MA, Janka-Schaub GE, Kamps WA, Evans WE, Pieters R. 2009. A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide classification study. Lancet Oncol 10:125–134.

Döhner K, Schlenk RF, Habdank M, Scholl C, Rücker FG, Corbacioglu A, Bullinger L, Fröhling S, Döhner H. 2005. Mutant nucleophosmin (NPM1) predicts favorable prognosis in younger adults with acute myeloid leukemia and normal cytogenetics: interaction with other gene mutations. Blood 106:3740–3746.

Eisen MB, Spellman PT, Brown PO, Botstein D. 1998. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 95:14863–14868.

Frohling S, Scholl C, Gilliland DG, Levine RL. 2005. Genetics of myeloid malignancies: Pathogenetic and clinical implications. J Clin Oncol 23:6285–6295.

Gough SM, Slape CI, Aplan PD. 2011. NUP98 gene fusions and hematopoietic malignancies: Common themes and new biological insights. Blood 118:6247–6257.

Grier DG, Thompson A, Kwasniewska A, McGonigle GJ, Halliday HL, Lappin TR. 2005. The pathophysiology of HOX genes and their role in cancer. J Pathol 205:154–171.

Hollink IH, van den Heuvel-Eibrink MM, Arentsen-Peters ST, Pratcorona M, Abbas S, Kuipers JE, van Galen JF, Beverloo HB, Sonneveld E, Kaspers GJ, Trka J, Baruchel A, Zimmermann M, Creutzig U, Reinhardt D, Pieters R, Valk PJ, Zwaan CM. 2011. NUP98/NSD1 characterizes a novel poor prognostic group in acute myeloid leukemia with a distinct HOX gene expression pattern. Blood 118:3645–3656.

Ichikawa H, Tanabe K, Mizushima H, Hayashi Y, Mizutani S, Ishii E, Hongo T, Kikuchi A, Satake M. 2006. Common gene expression signatures in t(8;21)- and inv(16)-acute myeloid leukemia. Br J Haematol 135:336–347.

Imaizumi M, Tawa A, Hanada R, Tsuchida M, Tabuchi K, Kigasawa H, Kobayashi R, Morimoto A, Nakayama H, Hamamoto K, Kudo K, Yabe H, Horibe K, Tsuchiya S, Tsukimoto I. 2011. Prospective study of a therapeutic regimen with all-trans retinoic acid and anthracyclines in combination of cytarabine in children with acute promyelocytic leukaemia: the Japanese childhood acute myeloid leukaemia cooperative study. Br J Haematol 152:89–98.

Jaju RJ, Fidler C, Haas OA, Strickson AJ, Watkins F, Clark K, Cross NC, Cheng JF, Aplan PD, Kearney L, Boultwood J,

Wainscoat JS. 2001. A novel gene, NSD1, is fused to NUP98 in the t(5;11)(q35;p15.5) in de novo childhood acute myeloid leukemia. Blood 98:1264–1267.

Jo A, Tsukimoto I, Ishii E, Asou N, Mitani S, Shimada A, Igarashi T, Hayashi Y, Ichikawa H. 2009. Age-associated difference in gene expression of paediatric acute myelomonocytic lineage leukaemia (FAB M4 and M5 subtypes) and its correlation with prognosis. Br J Haematol 144:917–929.

Kudo K, Kojima S, Tabuchi K, Yabe H, Tawa A, Imaizumi M, Hanada R, Hamamoto K, Kobayashi R, Morimoto A, Nakayama H, Tsuchida M, Horibe K, Kigasawa H, Tsukimoto I. 2007. Prospective study of a pirarubicin, intermediate-dose cytarabine, and etoposide regimen in children with Down syndrome and acute myeloid leukemia: The Japanese Childhood AML Cooperative Study Group. J Clin Oncol 25:5442–5447.

La Starza R, Gorello P, Rosati R, Riezzo A, Veronese A, Ferrazzi E, Martelli MF, Negrini M, Mecucci C. 2004. Cryptic insertion producing two NUP98/NSD1 chimeric transcripts in adult refractory anemia with an excess of blasts. Genes Chromosomes Cancer 41:395–399.

Marcucci G, Haferlach T, Döhner H. 2011. Molecular genetics of adult acute myeloid leukemia: Prognostic and therapeutic implications. J Clin Oncol 29:475–486.

Mochizuki N, Shimizu S, Nagasawa T, Tanaka H, Taniwaki M, Yokota J, Morishita K. 2000. A novel gene, MEL1, mapped to 1p36.3 is highly homologous to the MDS1/EVI1 gene and is transcriptionally activated in t(1;3)(p36;q21)-positive leukemia cells. Blood 96:3209–3214.

Mullighan CG, Su X, Zhang J, Radtke I, Phillips LA, Miller CB, Ma J, Liu W, Cheng C, Schulman BA, Harvey RC, Chen IM, Clifford RJ, Carroll WL, Reaman G, Bowman WP, Devidas M, Gerhard DS, Yang W, Relling MV, Shurtleff SA, Campana D, Borowitz MJ, Pui CH, Smith M, Hunger SP, Willman CL, Downing JR. 2009. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. N Engl J Med 360:470–480.

Nagel S, Kaufmann M, Drexler HG, MacLeod RA. 2003. The cardiac homeobox gene NKX2–5 is deregulated by juxtaposition with BCL11B in pediatric T-ALL cell lines via a novel t(5;14)(q35.1;q32.2). Cancer Res 63:5329–5334.

Nakamura T. 2005. NUP98 fusion in human leukemia: Dysregulation of the nuclear pore and homeodomain proteins. Int J Hematol 82:21–27.

Nakamura T, Largaespada DA, Lee MP, Johnson LA, Ohyashiki K, Toyama K, Chen SJ, Willman CL, Chen IM, Feinberg AP, Jenkins NA, Copeland NG, Shaughnessy JD Jr. 1996. Fusion of the nucleoporin gene NUP98 to HOXA9 by the chromosome translocation t(7;11)(p15;p15) in human myeloid leukaemia. Nature Genet 12:154–158.

Nebral K, König M, Schmidt HH, Lutz D, Sperr WR, Kalwak K, Brugger S, Dworzak MN, Haas OA, Strehl S. 2005. Screening for NUP98 rearrangements in hematopoietic malignancies by fluorescence in situ hybridization. Haematologica 90:746–752.

Pabst O, Zweigerdt R, Arnold HH. 1999. Targeted disruption of the homeobox transcription factor Nkx2–3 in mice results in postnatal lethality and abnormal development of small intestine and spleen. Development 126:2215–2225.

Panarello C, Rosanda C, Morerio C. 2002. Cryptic translocation t(5;11)(q35;p15.5) with involvement of the NSD1 and NUP98 genes without 5q deletion in childhood acute myeloid leukemia. Genes Chromosomes Cancer 35:277–281.

Przybylski GK, Dik WA, Grabarczyk P, Wanzeck J, Chudobska P, Jankowski K, von Bergh A, van Dongen JJ, Schmidt CA, Langerak AW. 2006. The effect of a novel recombination between the homeobox gene NKX2–5 and the TRD locus in T-cell acute lymphoblastic leukemia on activation of the NKX2–5 gene. Haematologica 91:317–321.

Pui CH, Carroll WL, Meshinchi S, Arceci RJ. 2011. Biology, risk stratification, and therapy of pediatric acute leukemias: An update. J Clin Oncol 29:551–565.

Roberts KG, Morin RD, Zhang J, Hirst M, Zhao Y, Su X, Chen SC, Payne-Turner D, Churchman ML, Harvey RC, Chen X, Kasap C, Yan C, Becksfort J, Finney RP, Teachey DT, Maude SL, Tse K, Moore R, Jones S, Mungall K, Birol I, Edmonson MN, Hu Y, Buetow KE, Chen IM, Carroll WL, Wei L, Ma J, Kleppe M, Levine RL, Garcia-Manero G, Larsen E, Shah NP, Devidas M, Reaman G, Smith M, Paugh SW, Evans WE, Grupp SA, Jeha S, Pui CH, Gerhard DS, Downing JR, Willman CL, Loh M, Hunger SP, Marra MA, Mullighan CG. 2012. Genetic alterations activating kinase and cytokine receptor signaling in high-risk acute lymphoblastic leukemia. Cancer Cell 22:153–166.

Romana SP, Radford-Weiss I, Ben Abdelali R, Schluth C, Petit A, Dastugue N, Talmant P, Bilhou-Nabera C, Mugneret F, Lafage-Pochitaloff M, Mozziconacci MJ, Andrieu J, Lai JL, Terre C, Rack K, Cornillet-Lefebvre P, Luquet I, Nadal N, Nguyen-Khac F, Perot C, Van den Akker J, Fert-Ferrer S, Cabrol C, Charrin C, Tigaud I, Poirel H, Vekemans M, Bernard OA, Berger R. 2006. NUP98 rearrangements in hematopoietic malignancies: a study of the Groupe Francophone de Cytogénétique Hématologique. Leukemia 20:696–706.

Sano H, Shimada A, Taki T, Murata C, Park MJ, Sotomatsu M, Tabuchi K, Tawa A, Kobayashi R, Horibe K, Tsuchida M, Hanada R, Tsukimoto I, Hayashi Y. 2012. RAS mutations are frequent in FAB type M4 and M5 of acute myeloid leukemia, and related to late relapse: A study of the Japanese Childhood AML Cooperative Study Group. Int J Hematol 95:509–515.

Shiba N, Taki T, Park MJ, Shimada A, Sotomatsu M, Adachi S, Tawa A, Horibe K, Tsuchida M, Hanada R, Tsukimoto I, Arakawa H, Hayashi Y. 2012. DNMT3A mutations are rare in childhood acute myeloid leukaemia, myelodysplastic syndromes and juvenile myelomonocytic leukaemia. Br J Haematol 156:413–414.

Shimada A, Taki T, Tabuchi K, Tawa A, Horibe K, Tsuchida M, Hanada R, Tsukimoto I, Hayashi Y. 2006. KIT mutations, and not FLT3 internal tandem duplication, are strongly associated with a poor prognosis in pediatric acute myeloid leukemia with t(8;21): a study of the Japanese Childhood AML Cooperative Study Group. Blood 107:1806–1809.

Shimada A, Taki T, Tabuchi K, Taketani T, Hanada R, Tawa A, Tsuchida M, Horibe K, Tsukimoto I, Hayashi Y. 2008. Tandem duplications of MLL and FLT3 are correlated with poor prognoses in pediatric acute myeloid leukemia: a study of the Japanese childhood AML Cooperative Study Group. Pediatr Blood Cancer 50:264–269.

Taketani T, Taki T, Shibuya N, Ito E, Kitazawa J, Terui K, Hayashi Y. 2002a The HOXD11 gene is fused to the NUP98 gene in acute myeloid leukemia with t(2;11)(q31;p15). Cancer Res 62:33–37.

Taketani T, Taki T, Ono R, Kobayashi Y, Ida K, Hayashi Y. 2002b. The chromosome translocation t(7;11)(p15;p15) in acute myeloid leukemia results in fusion of the NUP98 gene with a HOXA cluster gene, HOXA13, but not HOXA9. Genes Chromosomes Cancer 34:437–443.

Taketani T, Taki T, Shibuya N, Kikuchi A, Hanada R, Hayashi Y. 2002c Novel NUP98-HOXC11 fusion gene resulted from a chromosomal break within exon 1 of HOXC11 in acute myeloid leukemia with t(11;12)(p15;q13). Cancer Res 62: 4571–4574.

Taketani T, Taki T, Nakamura T, Kobayashi Y, Ito E, Fukuda S, Yamaguchi S, Hayashi Y. 2010. High frequencies of simultaneous FLT3-ITD, WT1 and KIT mutations in hematological malignancies with NUP98-fusion genes. Leukemia 24: 1975–1977.

Tsukimoto I, Tawa A, Horibe K, Tabuchi K, Kigasawa H, Tsuchida M, Yabe H, Nakayama H, Kudo K, Kobayashi R, Hamamoto K, Imaizumi M, Morimoto A, Tsuchiya S, Hanada R. 2009. Risk-stratified therapy and the intensive use of cytarabine improves the outcome in childhood acute myeloid leukemia: The AML99 trial from the Japanese Childhood AML Cooperative Study Group. J Clin Oncol 27:4007–4013.

Walter MJ, Payton JE, Ries RE, Shannon WD, Deshmukh H, Zhao Y, Baty J, Heath S, Westervelt P, Watson MA, Tomasson MH, Nagarajan R, O'Gara BP, Bloomfield CD, Mrózek K, Selzer RR, Richmond TA, Kitzman J, Geoghegan J, Eis PS, Maupin R, Fulton RS, McLellan M, Wilson RK, Mardis ER, Link DC, Graubert TA, DiPersio JF, Ley TJ. 2009. Acquired copy number alterations in adult acute myeloid leukemia genomes. Proc Natl Acad Sci USA 106:12950–12955.

Brief Report

MYELOID NEOPLASIA

Naturally occurring oncogenic GATA1 mutants with internal deletions in transient abnormal myelopoiesis in Down syndrome

Tsutomu Toki,[1] Rika Kanezaki,[1] Eri Kobayashi,[2,3] Hiroshi Kaneko,[3] Mikiko Suzuki,[3] RuNan Wang,[1] Kiminori Terui,[1] Hirokazu Kanegane,[4] Miho Maeda,[5] Mikiya Endo,[6] Tatsuki Mizuochi,[7] Souichi Adachi,[8] Yasuhide Hayashi,[9] Masayuki Yamamoto,[2] Ritsuko Shimizu,[3] and Etsuro Ito[1]

[1]Department of Pediatrics, Hirosaki University Graduate School of Medicine, Hirosaki, Japan; [2]Department of Medical Biochemistry and [3]Department of Molecular Hematology, Tohoku University Graduate School of Medicine, Sendai, Japan; [4]Department of Pediatrics, Graduate School of Medicine, University of Toyama, Toyama, Japan; [5]Department of Pediatrics, Nippon Medical School, Tokyo, Japan; [6]Department of Pediatrics, Iwate Medical University, Morioka, Japan; [7]Department of Pediatrics and Child Health, Kurume University School of Medicine, Kurume, Japan; [8]Human Health Sciences, Kyoto University Graduate School of Medicine, Kyoto, Japan; and [9]Department of Hematology/Oncology, Gunma Children's Medical Center, Gunma, Japan

Key Points

- Naturally occurring oncogenic GATA1 mutants with internal deletions contribute to transient abnormal myelopoiesis in Down syndrome.

Children with Down syndrome have an increased incidence of transient abnormal myelopoiesis (TAM) and acute megakaryoblastic leukemia. The majority of these cases harbor somatic mutations in the *GATA1* gene, which results in the loss of full-length GATA1. Only a truncated isoform of GATA1 that lacks the *N*-terminal 83 amino acids (GATA1-S) remains. We found through genetic studies of 106 patients with TAM that internally deleted GATA1 proteins (GATA1-IDs) lacking amino acid residues 77-119 or 74-88 (created by splicing mutations) contributed to the genesis of TAM in 6 patients. Analyses of GATA1-deficient embryonic megakaryocytic progenitors revealed that the GATA1 function in growth restriction was disrupted in GATA1-IDs. In contrast, GATA1-S promoted megakaryocyte proliferation more profoundly than that induced by GATA1 deficiency. These results indicate that the internally deleted regions play important roles in megakaryocyte proliferation and that perturbation of this mechanism is involved in the pathogenesis of TAM. (*Blood.* 2013;121(16):3181-3184)

Introduction

Children with Down syndrome (DS) are known to have a high risk of developing transient abnormal myelopoiesis (TAM) and subsequent acute megakaryoblastic leukemia (DS-AMKL).[1-4] Blast cells in the majority of patients with TAM and DS-AMKL have mutations in the second exon of the *GATA1* gene.[5,6] The mutations turn off the production of full-length GATA1. Instead, *N*-terminally truncated GATA1 protein (GATA1-S) was translated from the second methionine at codon 84, which is identical to the truncated GATA1 isoform found in the healthy human.[7] In contrast, only a few patients with AMKL have been reported to harbor 21-disomy blasts with the GATA1 mutation.[8,9] Therefore, GATA1-S is believed to be a prerequisite for the pathogenesis of TAM and DS-AMKL in children with DS, and unrestricted proliferation of megakaryocytic progenitors in DS-AMKL is thought to be provoked by a mechanism involving GATA1-S. However, the molecular mechanism of how GATA1-S contributes to the genesis of TAM and DS-AMKL remains elusive.

GATA1 regulates the proliferation of immature megakaryocytic progenitors. Indeed, active proliferation of immature megakaryocytic progenitors derived from GATA1-deficient mouse embryos is restricted by introduction of wild-type GATA1, but not by GATA-S.[10] GATA1-deficient mice rescued with transgenic expression of GATA1-S (or GATA1-ΔNT) are found to exhibit hyper-megakaryopoiesis in a limited embryonic and postnatal period, resembling the phenotype in human TAM cases.[11] In contrast, another report indicates that targeting mice expressing GATA1 protein with a deletion of 64 *N*-terminal amino acids, but retaining the 65th to 83rd amino acid residues intact, has demonstrated that the embryos display a transient megakaryocytic phenotype only during the early embryonic stage, not in the late-embryonic and postnatal stages.[12] We surmise that this difference simply may be a result of missing the region corresponding to the 65th to 83rd amino acids.

Here, we have identified novel GATA1 mutants with internal deletions (IDs) of either amino acid residues 77-119 or 74-88 (GATA1-IDs) in 6 patients. We found that the GATA1-IDs lost their activity in the regulation of megakaryocyte growth. These results demonstrate that disruption of ID regions is implicated in the pathogenesis of TAM.

Study design

This study was approved by the Ethics Committee of the Hirosaki University Graduate School of Medicine. All animal experiments

Submitted January 20, 2012; accepted February 7, 2013. Prepublished online as *Blood* First Edition paper, February 25, 2013; DOI 10.1182/blood-2012-01-405746.

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.

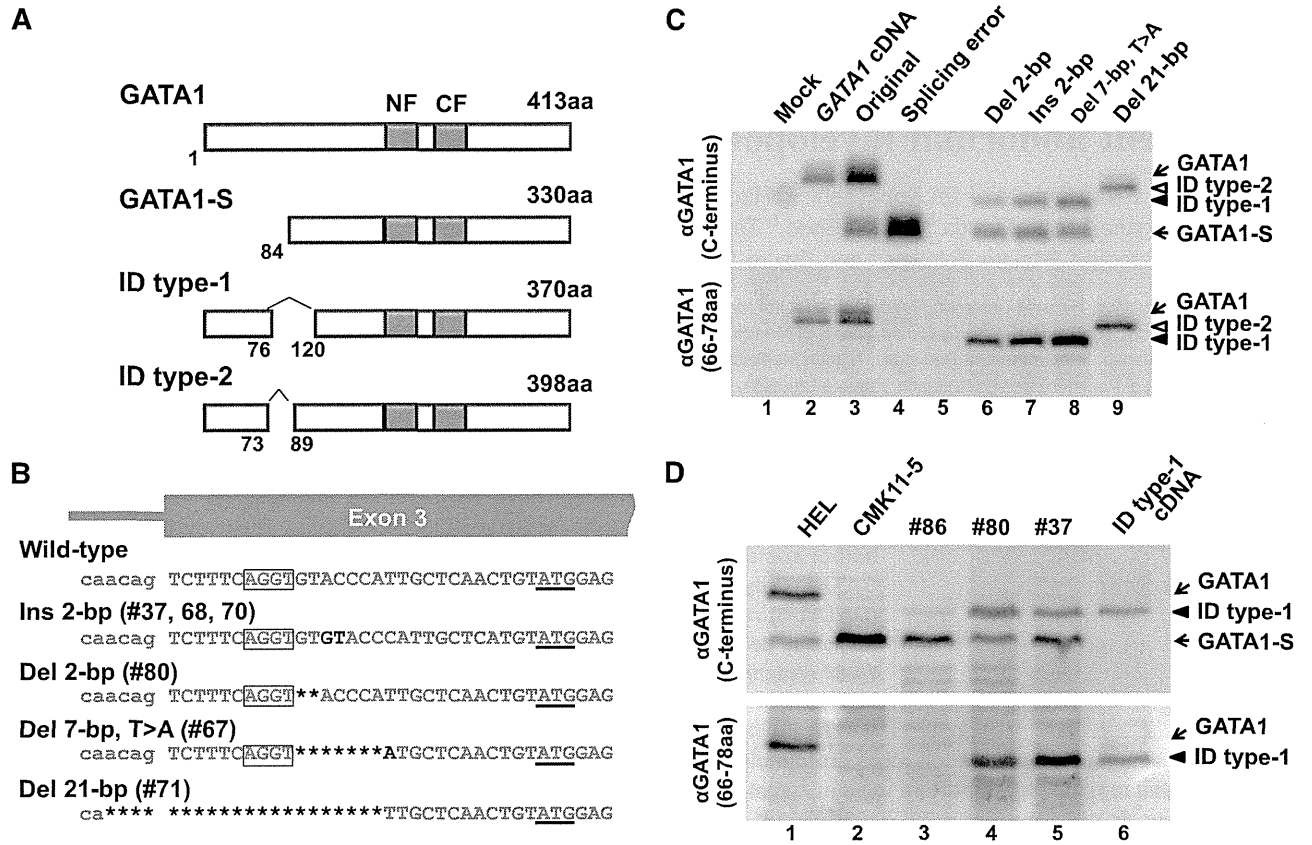

Figure 1. GATA1 mutant proteins with internal deletions. (A) A schema of mutant GATA1 proteins observed in patients with TAM. The amino acid sequence of GATA1-ID proteins was deduced from the sequence of *GATA1* cDNA obtained from patients with TAM. Dark boxes indicate *N*-finger (NF) and C-finger (CF) domains. ID indicates internal deletion. (B) Somatic mutations of the *GATA1* gene found in ID type 1 and type 2 patients. Missing, inserted, or substituted nucleotides are highlighted with dark color. A second translation initiation codon located in the third exon is underlined. The AGGT sequence functioning as an alternative splice donor site in mutant *GATA1* genes of ID type 1 patients is circled. Note that a mutant *GATA1* gene found in TAM patient 71f (ID type 2) lost a splice acceptor site in exon 3 because of the 21-nucleotide deletion. (C) Expression of GATA1 proteins in cells transfected with minigenes using anti-GATA1 antibodies recognizing the C terminus (upper) and residues between the 66th and 78th amino acids (lower) of the GATA1 protein. GATA1-ID proteins are recognized by the antibody against amino acid residues 66-78 of GATA1, whereas GATA1-S is not (lanes 6-9). Cells transfected with mock pcDNA3.1 (lane 1), pcDNA3.1-*GATA1* cDNA (lane 2), original minigene (lane 3), and *GATA1* minigene harboring a splicing error mutant in the 3′ boundary of intron 1[13] (lane 4) are used as positive and negative controls for GATA1 and GATA1-S, respectively. (D) GATA1 ID type 1 protein and GATA1-S are detected in the TAM blast cells from patients 80 (lane 4) and 37 (lane 5), whereas only GATA1-S is expressed in the blast cells from patient 86 harboring a conventional type of *GATA1* gene mutation in TAM cases (lane 3). Note that relatively abundant GATA1-S is recognized in patient 37 because of the intermixing of genetically distinct clone of cells expressing only GATA1-S (supplemental Table 1). Human erythroleukemia cells (HEL, lane 1) were used as a control for GATA1 and GATA1-S. DS-AMKL cells (CMK11-5, lane 2) and BHK-21 cells transfected with cDNA encoding GATA1 ID type 1 protein (lane 6) were used as controls for GATA1-S and GATA1 ID type 1, respectively.

were approved by the Institutional Animal Experiment Committee of Tohoku University. All clinical samples were obtained with informed consent from the parents of all patients with TAM in accordance with the Declaration of Helsinki. Additional information can be found in the supplemental text on the *Blood* website.

Results and discussion

Between 2003 and 2010, we screened *GATA1* mutations by direct sequencing, using cDNAs prepared from TAM blasts provided by 106 patients with DS on request from referring hospitals. Acquired *GATA1* mutations were detected in 99 (93.4%) patients (supplemental Table 1). The majority of the mutations resulted in the GATA1-S mutant protein, which lacks the entire *N*-terminal transactivation domain. Importantly, we found new mutations harboring IDs of 43 and 15 amino acids in 5 patients (patients 37, 67, 68, 70, and 80) and in 1 patient (patient 71), respectively. We refer to these mutants as GATA1-ID type 1 and GATA1-ID type 2, respectively (Figure 1A). Clinical features in patients with TAM who have GATA1-ID mutations were shown in supplemental Table 2. All of these patients showed high white blood cell counts in the peripheral blood, which is known to be a risk factor for early death.[13]

We determined the genomic DNA sequences of these cases. As shown in Figure 1B, the mutations in GATA1-ID type 1 were located in a site immediately 3′ of the consensus motif for a splice donor site AGGT[14] (Ins 2-bp in patients 37, 68, and 70; Del 2-bp in patient 80; and Del 7-bp T>A in patient 67), whereas 21 bp containing a splice acceptor site in front of exon 3 was deleted in GATA1-ID type 2 (Del 21-bp). To verify the transcripts achieved through the putative splice donor site created by mutations in GATA1-ID type 1, we introduced identified mutations into *GATA1* minigene expression vectors[13] and transduced them into hamster fibroblast cell line BHK-21. We found 3 variant transcripts in the cases of GATA1-ID type 1 mutations (supplementary Figure 1A-B): a full-length transcript with deletion or insertion of nucleotides [Ex-2 (+) (PTC)], a short transcript lacking exon 2 by alternative splice variant skipping of exon 2 for GATA1-S [Ex-2 (−)], and an aberrant transcript in which 129 nucleotides were spliced out from exon 3 (Del 129-bp). In contrast, 2 disparate transcripts with deletions of 45 or 137 nucleotides were created by

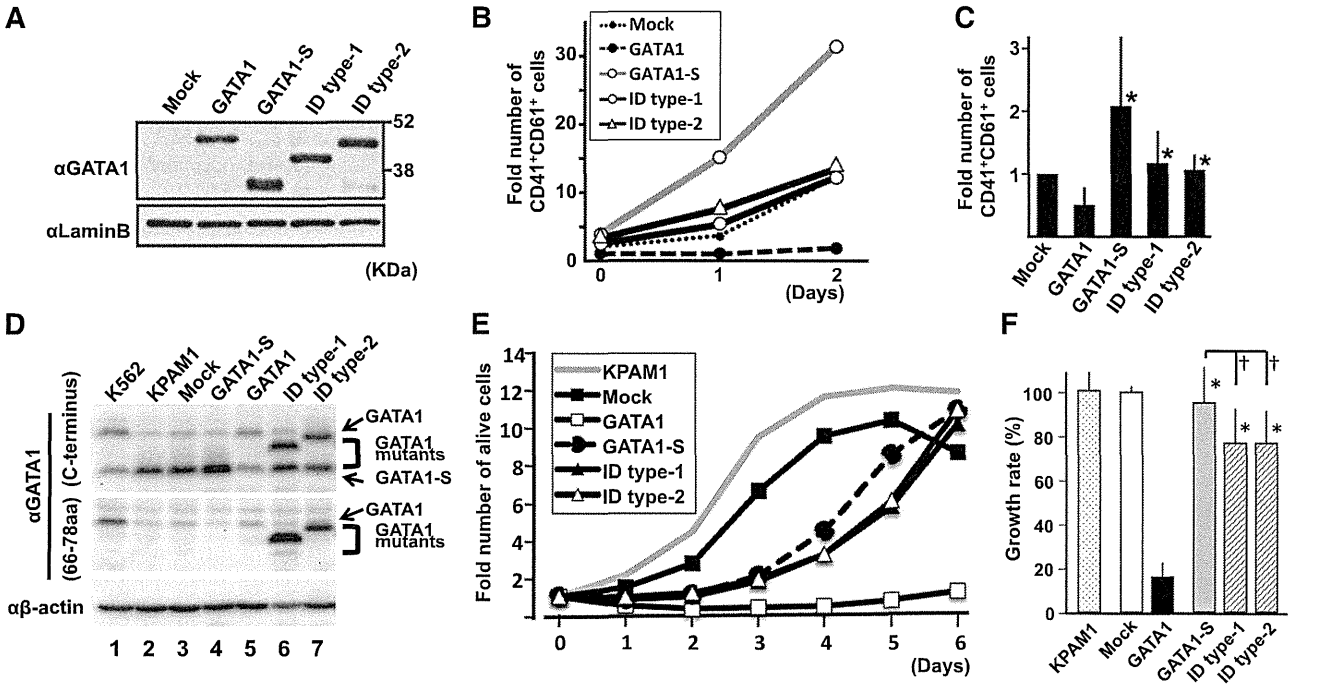

Figure 2. GATA1 ID proteins showed restricted antiproliferative activity. (A) Expression of GATA1 and GATA1 mutant proteins in cultured megakaryocytes at day 0, using an antibody against the C terminus of GATA1. The amount of protein loaded was quantified using an anti-Lamin B antibody on the same membrane. (B) Time-course change in the number of $CD41^+CD61^+$ cells. The value in the mock case at day 0 is set to 1. The result is representative of 4 independent experiments. (C) Comparison of the number of $CD41^+CD61^+$ cells at day 2. The value in the mock case is set to 1 in every experiment. The mean values and standard deviations from 4 independent experiments are presented. Asterisks indicate a significant difference compared with wild-type GATA1 ($P < .05$). (D) Immunoblot analysis of ectopic expression of GATA1 proteins in KPAM1 cells using anti-GATA1 antibodies against C terminus (upper) and residues between amino acids 66 and 78 (middle). The loading volume was quantified using anti-β-actin antibody (lower). (E) Growth curves of KPAM1 cells after ectopic expression of GATA1 proteins. Average values obtained from 6 wells are shown. The value at day zero is set to 1 for each. The growth curve of the original KPAM1 cells was analyzed as a control. Representative data from 3 independent experiments are shown. (F) Relative growth rate of KPAM1 cells at 5 days after ectopic expression of GATA1 mutant proteins. The average value of growth rate in the mock case is set to 100% in every experiment. The mean values and standard deviations from 18 wells obtained in 3 independent experiments (6 wells in each) are presented. Asterisks and daggers indicate significant differences compared with wild-type GATA1 and GATA-S, respectively ($P < .01$).

mutation in GATA1-ID type 2, using alternative acceptor sites in exon 3.

To examine whether the GATA1-ID proteins were produced from the mutant alleles, we performed immunoblotting analysis with 2 distinct antibodies recognizing the C terminus and amino acids 66-78, respectively. We detected GATA1-ID type 1 protein in addition to GATA1-S in the cells transfected with the minigenes harboring Ins 2-bp, Del 2-bp, or Del 7-bp T>A mutations, whereas only GATA1-ID type 2 protein was expressed on transfection of the minigene with a Del 21-bp mutation (Figure 1C). Consistent with the minigene results, a significant amount of GATA1-ID type 1 protein and GATA1-S had accumulated in patients 80 and 37, whereas only GATA1-S was detected in the TAM blasts of patient 86, who had only a short transcript skipping exon 2 because of a point mutation in the exon 2–intron 2 boundary (Figure 1D). Thus, splicing errors were occurred in GATA1-ID type 1 and type 2 patients, leading to the production of GATA1-ID proteins.

We next examined how GATA1-ID proteins affect the proliferation of embryonic megakaryocytic progenitors. We retrovirally transduced GATA1-S and GATA1-ID mutants into lineage-negative cells derived from megakaryocyte-specific *Gata1*-deficient ($Gata1^{\Delta neo\Delta HS}$) embryos[15] and induced differentiation toward the megakaryocytic lineage. The number of $CD41^+CD61^+$ megakaryocytes was significantly higher in cases transduced with GATA1-ID proteins than with wild-type GATA1, despite almost equivalent expression levels of GATA1 proteins (Figure 2A-C). GATA1-S-transduced cells unexpectedly acquired a hyperproliferative potential compared with mock cells, probably because of an unknown function that resides in the GATA1 *N*-terminal region (Figure 2B-C).

We next analyzed cell proliferation using the DS-AMKL cell line KPAM1, in which GATA1-S was predominantly expressed with a very low level of full-length GATA1 (Figure 2D).[16] On transduction with full-length GATA1 retrovirus, proliferation of KPAM1 cells was markedly reduced. In contrast, GATA1-ID type 1 and type 2 moderately restricted the proliferation of KPAM1 cells, but the restriction activity was significantly stronger than that of GATA1-S (Figure 2E-F). These results thus demonstrate that the ID regions indeed contribute to the regulation of AMKL cell proliferation.

Our newly identified GATA1-ID mutants have highlighted a much narrower set of sequences responsible for the pathogenesis of TAM than has previously been suggested by the loss of the *N*-terminal sequence, as in GATA1-S. The missing region identified by the GATA1-ID proteins contains a consensus motif (LxCxE, amino acids 81-85) essential for the interaction with pRb,[17] which is also lost in GATA1-S. Interaction with hypophosphorylated pRb-E2F complex has been reported to be important for GATA1 to support the normal proliferation and differentiation of erythroid progenitors.[17] Consistent with this notion, GATA1-S failed to repress E2F activation, which was followed by activation of mTOR signaling in the GATA1-S fetal megakaryocytes and DS-AMKL cells.[18] Because the protein levels of cyclin D1 and $p27^{Kip}$ are reciprocally regulated by the

mTOR pathway, and thereby cause pRb to be phosphorylated,[19] cell-cycle progression in response to the mTOR pathway may be potentiated by the enfeebled function of LxCxE motif of GATA1-S. Thus, we are one step closer to a molecular understanding of GATA1-related leukemias.

Acknowledgments

This work was supported in part by grants-in-aid for scientific research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (R.S., T.T., M.Y., and E.I.), sciences research grants from the Ministry of Health, Labour and Welfare of Japan (E.I.), the Asahi Glass Foundation (R.S.), the Mitsubishi Foundation (R.S. and M.Y.) and the Takeda Foundation (M.Y.).

Authorship

Contribution: T.T., R.K., E.K., H. Kaneko, R.W., and K.T. contributed to the experiments; T.T., M.S., R.S., M.Y., and E.I. contributed to the study design, funding, project conception, and manuscript writing; and H. Kanegane, M.M., M.E., T.M., S.A., and Y.H. contributed to the clinical sample collection and phenotype analyses.

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

Correspondence: Masayuki Yamamoto, Department of Medical Biochemistry, Tohoku University Graduate School of Medicine, 2-1 Seiryo-cho, Aoba-ku, Sendai 980-8575, Japan; e-mail: masiyamamoto@med.tohoku.ac.jp; and Etsuro Ito, Department of Pediatrics, Hirosaki University Graduate School of Medicine, 5 Zaifu-cho, Hirosaki, 036-8562, Japan; e-mail: eturou@cc.hirosaki-u.ac.jp.

References

1. Zipursky A, Poon A, Doyle J. Leukemia in Down syndrome: a review. *Pediatr Hematol Oncol.* 1992;9(2):139-149.
2. Hasle H, Niemeyer CM, Chessells JM, Baumann I, Bennett JM, Kerndrup G, Head DR. A pediatric approach to the WHO classification of myelodysplastic and myeloproliferative diseases. *Leukemia.* 2003;17(2):277-282.
3. Hitzler JK. Acute megakaryoblastic leukemia in Down syndrome. *Pediatr Blood Cancer.* 2007;49 (7 Suppl):1066-1069.
4. Malinge S, Izraeli S, Crispino JD. Insights into the manifestations, outcomes, and mechanisms of leukemogenesis in Down syndrome. *Blood.* 2009; 113(12):2619-2628.
5. Wechsler J, Greene M, McDevitt MA, Anastasi J, Karp JE, Le Beau MM, Crispino JD. Acquired mutations in GATA1 in the megakaryoblastic leukemia of Down syndrome. *Nat Genet.* 2002; 32(1):148-152.
6. Xu G, Nagano M, Kanezaki R, et al. Frequent mutations in the GATA-1 gene in the transient myeloproliferative disorder of Down syndrome. *Blood.* 2003;102(8):2960-2968.
7. Rainis L, Bercovich D, Strehl S, et al. Mutations in exon 2 of GATA1 are early events in megakaryocytic malignancies associated with trisomy 21. *Blood.* 2003;102(3):981-986.
8. Harigae H, Xu G, Sugawara T, Ishikawa I, Toki T, Ito E. The GATA1 mutation in an adult patient with acute megakaryoblastic leukemia not accompanying Down syndrome. *Blood.* 2004; 103(8):3242-3243.
9. Hama A, Yagasaki H, Takahashi Y, et al. Acute megakaryoblastic leukaemia (AMKL) in children: a comparison of AMKL with and without Down syndrome. *Br J Haematol.* 2008;140(5):552-561.
10. Muntean AG, Crispino JD. Differential requirements for the activation domain and FOG-interaction surface of GATA-1 in megakaryocyte gene expression and development. *Blood.* 2005; 106(4):1223-1231.
11. Shimizu R, Kobayashi E, Engel JD, Yamamoto M. Induction of hyperproliferative fetal megakaryopoiesis by an N-terminally truncated GATA1 mutant. *Genes Cells.* 2009;14(9): 1119-1131.
12. Li Z, Godinho FJ, Klusmann JH, Garriga-Canut M, Yu C, Orkin SH. Developmental stage-selective effect of somatically mutated leukemogenic transcription factor GATA1. *Nat Genet.* 2005; 37(6):613-619.
13. Kanezaki R, Toki T, Terui K, et al. Down syndrome and GATA1 mutations in transient abnormal myeloproliferative disorder: mutation classes correlate with progression to myeloid leukemia. *Blood.* 2010;116(22):4631-4638.
14. Shapiro MB, Senapathy P. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res.* 1987;15(17): 7155-7174.
15. Shivdasani RA, Fujiwara Y, McDevitt MA, Orkin SH. A lineage-selective knockout establishes the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet development. *EMBO J.* 1997;16(13):3965-3973.
16. Toki T, Kanezaki R, Adachi S, et al. The key role of stem cell factor/KIT signaling in the proliferation of blast cells from Down syndrome-related leukemia. *Leukemia.* 2009;23(1):95-103.
17. Kadri Z, Shimizu R, Ohneda O, et al. Direct binding of pRb/E2F-2 to GATA-1 regulates maturation and terminal cell division during erythropoiesis. *PLoS Biol.* 2009;7(6):e1000123.
18. Klusmann JH, Godinho FJ, Heitmann K, et al. Developmental stage-specific interplay of GATA1 and IGF signaling in fetal megakaryopoiesis and leukemogenesis. *Genes Dev.* 2010;24(15): 1659-1672.
19. Gera JF, Mellinghoff IK, Shi Y, et al. AKT activity determines sensitivity to mammalian target of rapamycin (mTOR) inhibitors by regulating cyclin D1 and c-myc expression. *J Biol Chem.* 2004; 279(4):2737-2746.

nature genetics

Somatic *SETBP1* mutations in myeloid malignancies

Hideki Makishima[1], Kenichi Yoshida[2], Nhu Nguyen[3], Bartlomiej Przychodzen[1], Masashi Sanada[2,4], Yusuke Okuno[2,5], Kwok Peng Ng[1], Kristbjorn O Gudmundsson[3], Bandana A Vishwakarma[3], Andres Jerez[1], Ines Gomez-Segui[1], Mariko Takahashi[2], Yuichi Shiraishi[6], Yasunobu Nagata[2], Kathryn Guinta[1], Hiraku Mori[7], Mikkael A Sekeres[8], Kenichi Chiba[6], Hiroko Tanaka[9], Hideki Muramatsu[5], Hirotoshi Sakaguchi[5], Ronald L Paquette[10], Michael A McDevitt[11], Seiji Kojima[5], Yogen Saunthararajah[1], Satoru Miyano[6,9], Lee-Yung Shih[12], Yang Du[3,13], Seishi Ogawa[2,4,13] & Jaroslaw P Maciejewski[1,13]

Here we report whole-exome sequencing of individuals with various myeloid malignancies and identify recurrent somatic mutations in *SETBP1*, consistent with a recent report on atypical chronic myeloid leukemia (aCML)[1]. Closely positioned somatic *SETBP1* mutations encoding changes in Asp868, Ser869, Gly870, Ile871 and Asp880, which match germline mutations in Schinzel-Giedion syndrome (SGS)[2], were detected in 17% of secondary acute myeloid leukemias (sAML) and 15% of chronic myelomonocytic leukemia (CMML) cases. These results from deep sequencing demonstrate a higher mutational detection rate than reported with conventional sequencing methodology[3–5]. Mutant cases were associated with advanced age and monosomy 7/deletion 7q (–7/del(7q)) constituting poor prognostic factors. Analysis of serially collected samples indicated that *SETBP1* mutations were acquired during leukemic evolution. Transduction with mutant *Setbp1* led to the immortalization of mouse myeloid progenitors that showed enhanced proliferative capacity compared to cells transduced with wild-type *Setbp1*. Somatic mutations of *SETBP1* seem to cause gain of function, are associated with myeloid leukemic transformation and convey poor prognosis in myelodysplastic syndromes (MDS) and CMML.

During the past decade, substantial progress has been made in the understanding of the pathogenic gene mutations driving myeloid malignancies. Following the early identification of mutations in *RUNX1* (ref. 6), *JAK2* (ref. 7) and *RAS*[8,9], SNP array karyotyping led to the discovery of mutations in *CBL*[10], *TET2* (ref. 11) and *EZH2* (ref. 12). More recently, new sequencing technologies have enabled exhaustive screening of somatic mutations in myeloid malignancies, leading to the discovery of unexpected mutational targets, such as *DNMT3A*[13], *IDH1* (ref. 14) and spliceosomal genes[15–17]. Insights into the progression to sAML constitute an important goal of biomedical investigations, now augmented by the availability of next-generation sequencing technologies[18,19].

We performed whole-exome sequencing of 20 index cases with myeloid malignancies (**Supplementary Table 1**) and identified 38 non-silent somatic mutations that were subsequently confirmed by Sanger sequencing and targeted deep sequencing. We found that seven genes were recurrently mutated in multiple samples (**Supplementary Tables 2–4**). Of these, we identified a new recurrent somatic mutation in *SETBP1* (encoding a p.Asp868Asn alteration) in two cases with refractory anemia with excess blasts (RAEB) (**Fig. 1** and **Supplementary Tables 1–3** and **5**), which were confirmed using DNA from both tumor and $CD3^+$ T cells.

SETBP1 was initially identified as a 170-kDa nuclear protein that binds to SET[20,21] and is activated to support the recovery of granulopoiesis in chronic granulomatous disease[22]. Mutations in *SETBP1* are causative in SGS, a congenital disease characterized by a higher than normal prevalence of tumors, typically neuroepithelial neoplasia[23,24]. Notably, the mutations identified in our cohort exactly corresponded with the recurrent *de novo* germline mutations responsible for SGS, which prompted us to investigate *SETBP1* mutations in a large cohort of 727 cases with various myeloid malignancies (**Supplementary Table 6**).

SETBP1 mutations were found in 52 of 727 cases (7.2%). Consistent with recent reports[1,3–5,25,26], p.Asp868Asn (n = 28), p.Gly870Ser (n = 15) and p.Ile871Thr (n = 5) alterations were more frequent than p.Asp868Tyr, p.Ser869Asn, p.Asp880Asn and p.Asp880Glu alterations (n = 1 for each) (**Fig. 1** and **Supplementary Tables 1** and **7**). All these alterations were located in the SKI homology region, which is highly

[1]Department of Translational Hematology and Oncology Research, Taussig Cancer Institute, Cleveland Clinic, Cleveland, Ohio, USA. [2]Cancer Genomics Project, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. [3]Department of Pediatrics, Uniformed Services University of the Health Sciences, Bethesda, Maryland, USA. [4]Department of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto University, Kyoto, Japan. [5]Department of Pediatrics, Nagoya University Graduate School of Medicine, Nagoya, Japan. [6]Laboratory of DNA Information Analysis, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan. [7]Department of Hematology, Showa University, Tokyo, Japan. [8]Department of Hematologic Oncology and Blood Disorders, Taussig Cancer Institute, Cleveland Clinic, Cleveland, Ohio, USA. [9]Laboratory of Sequence Analysis, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan. [10]Department of Medicine, Hematology/Oncology, University of California, Los Angeles, Los Angeles, California, USA. [11]Department of Medicine and Oncology, Division of Hematology and Hematological Malignancy, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. [12]Department of Internal Medicine, Division of Hematology-Oncology, Chang Gung Memorial Hospital, Chang Gung University, Taipei, Taiwan. [13]These authors contributed equally to this work. Correspondence should be addressed to J.P.M. (maciejj@ccf.org), S.O. (sogawa-tky@umin.ac.jp) or Y.D. (yang.du@usuhs.edu).

Received 14 November 2012; accepted 13 June 2013; published online 7 July 2013; doi:10.1038/ng.2696

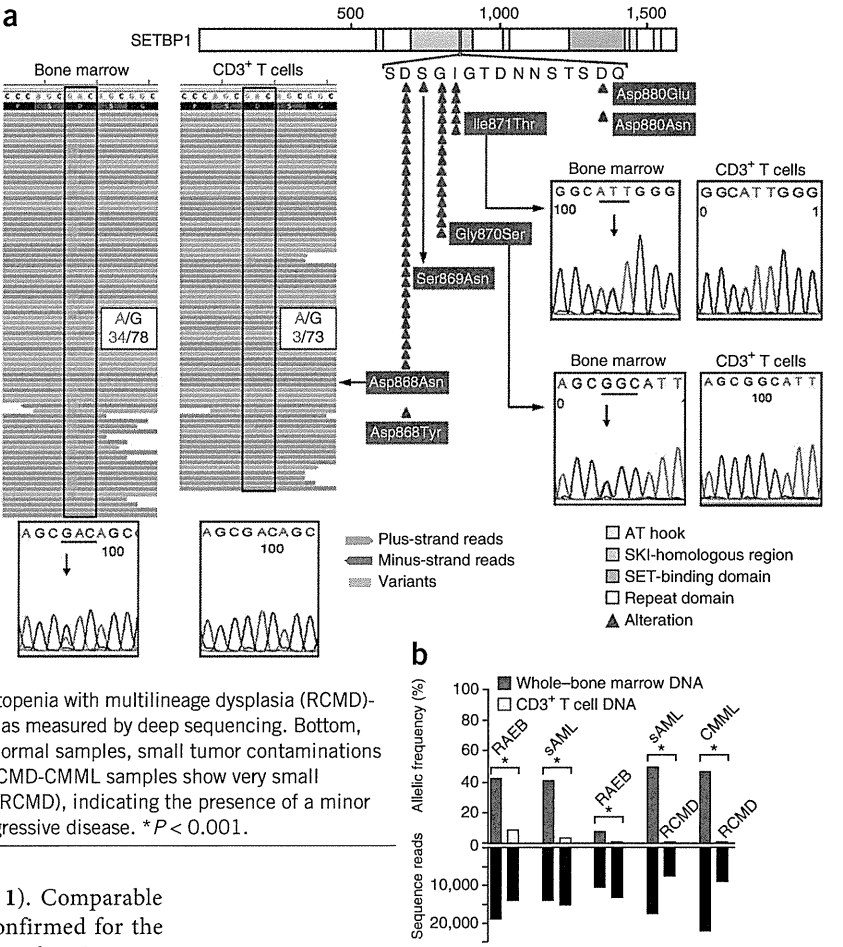

Figure 1 Somatic *SETBP1* mutations as detected by next-generation whole-exome sequencing and Sanger sequencing. (**a**) Distribution of *SETBP1* mutations detected in 52 of 727 myeloid neoplasms, all of which were located within the portion of the gene encoding the SKI-homologous domain. Bottom and right, representative mutations confirmed by Sanger sequencing (horizontal lines and vertical arrows indicate affected codons and nucleotides, respectively). Left, a somatic *SETBP1* mutation (encoding a p.Asp868Asn alteration) detected by whole-exome sequencing of paired tumor (bone marrow) and normal (CD3+ T cell) DNA from a case with RAEB (whole exome 4), where red and blue bars indicate positive and negative strands, respectively. Mutated nucleotides (c.2602G>A) are shown in green. Black rectangles highlight the codon affected by mutation. A small amount of tumor cell contamination caused occasional mutant reads in the CD3+ T cell sample, where the presence of multiple single-nucleotide variants (SNVs) of similar frequencies precluded the possibility of somatic mosaicism. (**b**) Top, allele frequencies in paired bone marrow and CD3+ T cell samples in two RAEB cases and one sAML case and in paired refractory cytopenia with multilineage dysplasia (RCMD)-sAML and RCMD-CMML samples from the same individual as measured by deep sequencing. Bottom, depth of coverage of independent reads. In paired tumor-normal samples, small tumor contaminations were detected in CD3+ T cells. Paired RCMD-sAML and RCMD-CMML samples show very small numbers of mutant reads in the initial MDS presentation (RCMD), indicating the presence of a minor *SETBP1*-mutated clone, which evolved later into more aggressive disease. $*P < 0.001$.

conserved between species (**Supplementary Fig. 1**). Comparable expression of mutant and wild-type alleles was confirmed for the p.Asp868Asn and p.Gly870Ser alterations by allele-specific PCR using genomic DNA and cDNA (**Supplementary Fig. 2**). *SETBP1* mutations were significantly associated with advanced age ($P = 0.01$) and −7/del(7q) ($P = 0.01$) and were frequently found in sAML (19 of 113 cases; 16.8%; $P < 0.001$) and CMML (22 of 152 cases; 14.5%; $P = 0.002$), whereas they were less frequent in primary AML (1 of 145 cases; <1%; $P = 0.002$) (**Table 1** and **Supplementary Fig. 3a**). The lack of apparent segmental allelic imbalance involving the *SETBP1* locus (18q12.3) in SNP array karyotyping in all mutated cases (**Supplementary Fig. 4**), together with no more than 50% mutant allele frequencies in deep sequencing and allele-specific PCR, suggested the presence of heterozygous mutations (**Fig. 1b** and **Supplementary Fig. 2**). Medical history and physical findings did not support clinical diagnosis with SGS in any of these cases, and formal confirmation of the somatic origin of all types of mutation found was carried out using germline DNA from CD3+ T cells and/or serial samples ($n = 21$).

Of the cases with *SETBP1* mutations, 12 had clinical material available to successfully analyze samples collected serially at multiple clinical time points. None of the 12 cases had *SETBP1* mutations at the time of initial presentation, indicating that the mutations were acquired only upon or during leukemic evolution (**Figs. 1** and **2**). Most of the *SETBP1* mutations (17 of 19) showed comparable or higher allele frequencies relative to other secondary events, suggesting a potential permissive role of *SETBP1* mutations (**Supplementary Fig. 5**). Such a secondary nature for *SETBP1* mutations was confirmed by mutational analysis of colonies derived from individual progenitor cells grown in methylcellulose culture (**Supplementary Fig. 6**).

To test potential associations with additional genetic defects, the frequencies of mutations in 13 common genes relevant to myeloid leukemogenesis were compared in the cases with *SETBP1* mutations and in cases with wild-type *SETBP1* (**Fig. 2c,d** and **Supplementary Table 8**). Only *CBL* mutations were significantly associated with *SETBP1* mutations ($P = 0.002$; **Supplementary Table 9**). Notably, mutations of *FLT3* and *NPM1* were not found in cases with *SETBP1* mutation. Coexisting *SETBP1* and *CBL* mutations were found in 12 cases, of which 6 were subjected to deep sequencing, and *CBL*-mutated clones were significantly smaller than *SETBP1*-mutated clones, suggesting that *CBL* mutations were acquired by a subclone with *SETBP1* mutation (**Supplementary Fig. 5**). The significant association of *CBL* and *SETBP1* mutations suggests their potential cooperation in leukemia progression. Although direct physical interaction between mutant Setbp1 and CBL proteins was not detected (**Supplementary Fig. 7**), it is possible that *CBL* mutations cooperate with *SETBP1* mutations indirectly by reducing the cytokine dependence of leukemia cells[10,27]. *SETBP1* mutations were also found in aCML[1] and juvenile chronic myelomonocytic leukemia[28], characterized by RAS pathway defects, including *CBL* mutations.

Analysis of the expression patterns of *SETBP1* mRNA in normal hematopoietic tissues showed relatively low levels of this transcript in myeloid and/or monocytic cells as well as in CD34+ cells (**Supplementary Fig. 8**). In contrast, *SETBP1*-mutant cases showed significantly higher expression levels than samples with wild-type *SETBP1* ($P = 0.03$; **Supplementary Fig. 9**). When *SETBP1* expression was also evaluated using expression array data in the cases with different subtypes of myeloid neoplasm (**Supplementary Fig. 10**), *SETBP1* was found to be overexpressed in cases with non–core binding factor (CBF) primary AML, including MDS, whereas CBF leukemias showed normal levels of the corresponding mRNA. In particular, *SETBP1*

Table 1 Clinical characteristics of myeloid malignancies with or without *SETBP1* mutation

Characteristic	Wild-type *SETBP1*	Mutant *SETBP1*	*P*[a]
Number	675	52	
Age at study entry (years), mean ± s.d.	61 ± 15	67 ± 12	**0.01**[b]
Age range (years)	16–91	26–83	
Ancestry, number			0.27
Caucasian	222	29	
African American	10	0	
Asian	298	23	
Other	2	0	
Male sex, number	376	29	0.23
Increased (≥10%) bone marrow blasts, number	376	33	0.31
Diagnosis, number			
5q– syndrome	7	1	1.00
RCMD	52	2	1.00
RAEB	86	4	1.00
sAML	94	19	**<0.001**
CMML	130	22	**0.002**
CML BP	25	2	1.00
PMF	25	1	1.00
pAML	144	1	**0.002**
Cytogenetics, number			
Normal	208	17	1.00
–5,del(5q)	39	1	1.00
–7,del(7q)	72	15	**0.01**
–Y only	9	0	1.00
–20,del(20q)	18	1	1.00
+8	45	2	1.00
Complex (≥3)	69	2	1.00

CML BP, chronic myelogenous leukemia blast phase; PMF, primary myelofibrosis.
[a]A Fisher's exact test was used to determine *P* values, except where otherwise indicated. *P* values in multiple comparisons were evaluated by Bonferroni correction, and statistically significant *P* values are indicated with bold font. [b]A Wilcoxon test was used to calculate the *P* value.

expression was significantly higher in cases with loss of chromosome 7 ($P = 0.03$) and complex karyotype ($P < 0.001$) (**Supplementary Fig. 3**). Clustering analysis of gene expression profiles suggested that *SETBP1*-mutant cases had a similar expression pattern to that of cases with overexpression of wild-type *SETBP1*, including overexpression of *TCF4*, *BCL11A* and *DNTT* (**Supplementary Fig. 10** and **Supplementary Table 10**). Methylation array analysis showed that relative hypomethylation of the CpG site located in proximity to the *SETBP1* coding region was associated with higher expression and mutation of *SETBP1* (**Supplementary Fig. 11**). It remains unclear what factors drive the increase in *SETBP1* mRNA levels in these leukemias; however, these mechanisms may involve aberrant hypomethylation of the *SETBP1* promoter or activation of upstream regulators such as *MECOM*[22,29]).

Within the entire cohort, *SETBP1*-mutated cases were significantly associated with shorter overall survival time (hazards ratio (HR) = 2.27, 95% confidence interval (CI) = 1.56–3.21; $P < 0.001$), with this association especially prominent in the younger age group (<60 years; HR = 4.92, 95% CI = 2.32–9.46; $P < 0.001$). The presence of *SETBP1* mutations was also associated with compromised survival in the cohort with normal karyotype (HR = 3.13, 95% CI = 1.66–5.41; $P = 0.002$) (**Fig. 3**). Multivariate analysis confirmed that *SETBP1* mutation was an independent prognostic factor (HR = 2.90, 95% CI = 1.71–4.83; $P < 0.001$) together with male sex, advanced age and the presence of *ASXL1*, *CBL* and *DNMT3A* mutations. –7/del(7q) was associated with shorter length of survival in univariate analysis but did not remain an independent risk factor after multivariate analysis (**Supplementary Table 11**). The multivariate analysis in the subgroup of MDS and CMML cases (with white blood cell (WBC) counts of <12,000 cells/μl), in which the International Prognostic Scoring System (IPSS) score was applicable[30], also showed that *SETBP1* mutation was an independent prognostic factor (HR = 1.83, 95% CI = 1.04–3.12; $P = 0.04$), whereas the impact of the IPSS score

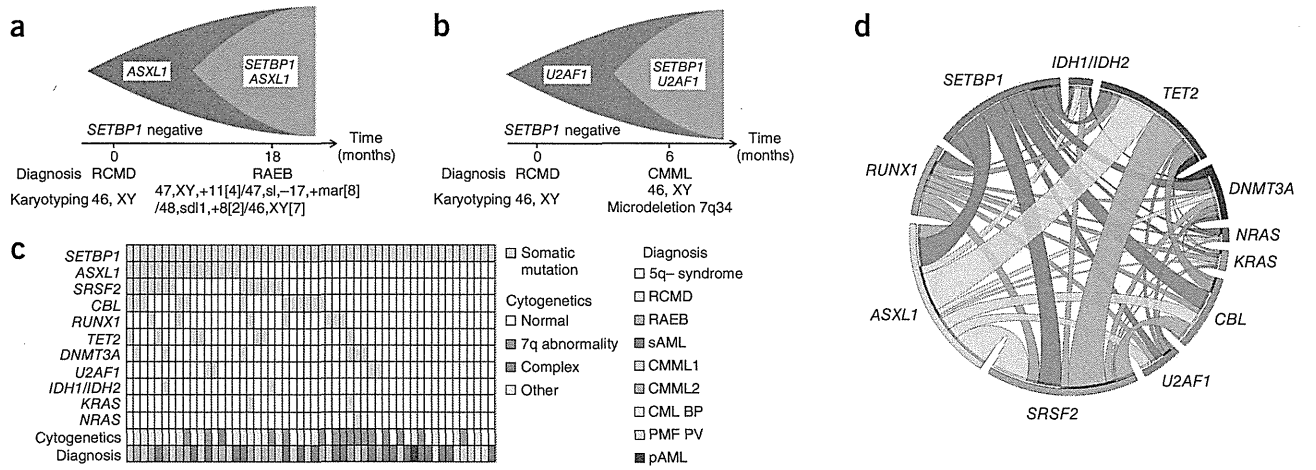

Figure 2 The relationship of *SETBP1* mutations with other common mutations. (**a**,**b**) Clonological profiles of gene mutations in two representative cases with MDS that transformed to RAEB (**a**) and CMML (**b**). Initially, hypocellular MDS (RCMD) was diagnosed on the basis of hypocellular bone marrow with normal karyotype in both cases. (**c**) Coexisting mutations in the *SETBP1*-mutated cohort are shown in a matrix: 36 of 52 cases (69%) were positive for other somatic concomitant mutations tested by Sanger sequencing. Sequenced genes are listed in **Supplementary Table 8**. CMML1 and CMML2 were discriminated by the number of blasts plus promonocytes in the peripheral blood and bone marrow. PV, polycythemia vera; pAML, primary AML. (**d**) Circos plots illustrating coexisting mutations in the selected 12 genes in the whole cohort. No mutations that occurred in a mutually exclusive manner were observed.

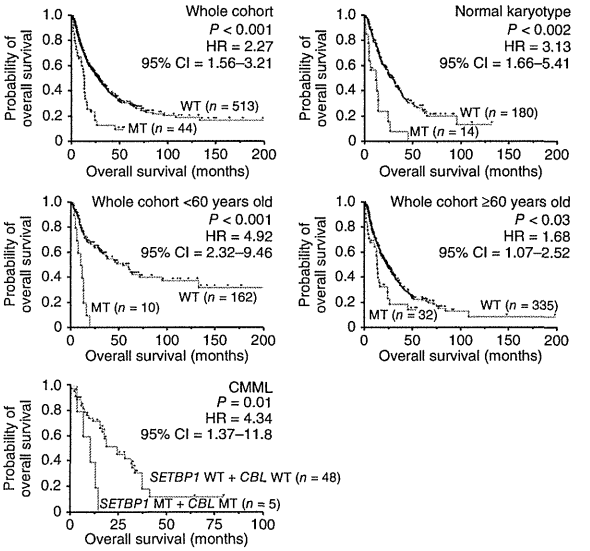

Figure 3 Impact of *SETBP1* mutations on clinical outcome. In the whole cohort, cases with *SETBP1* mutations (MT) had worse overall survival compared with those with wild-type *SETBP1* (WT). *SETBP1* mutations were poor prognostic factors for individuals with normal karyotype. *SETBP1* mutations were poor prognostic factors for individuals regardless of age (>60 and ≤60 years). In the CMML cohort, individuals with double mutations of *SETBP1* and *CBL* had worse prognosis than those with both wild-type genes. Data points indicate events and censors. The Kaplan-Meier method was used to analyze survival outcomes by the log-rank test.

dissipated after multivariate analysis (**Supplementary Tables 11** and **12**). Next, because comprehensive mutational screening identified a significant association between *SETBP1* and *CBL* mutations, we compared overall length of survival in cases with either of these mutations or with these mutations in combination (**Supplementary Figs. 12** and **13** and **Supplementary Table 13**). Overall length of survival was shorter in cases with mutation in both *SETBP1* and *CBL* compared to those with the wild-type forms of these genes, and the combination of these mutations was also unfavorable in an isolated CMML cohort in which either of these mutations alone did not affect survival (**Fig. 3** and **Supplementary Fig. 13**). However, no impact of these mutations was found in a sAML cohort, probably owing to the already very poor prognosis in this subset of individuals (**Supplementary Figs. 12** and **14**).

Previous studies demonstrated that overexpression of *Setbp1* can effectively immortalize mouse myeloid precursors[31]. Expression of Setbp1 mutants (either Asp868Asn or Ile871Thr) also caused efficient immortalization of mouse myeloid progenitors with similar phenotypes (**Fig. 4a,b** and **Supplementary Fig. 15**). Moreover, although having similar levels of Setbp1 protein expression as cells immortalized with wild-type *Setbp1*, cells immortalized with mutant *Setbp1* showed significantly more efficient colony formation and faster proliferation (**Fig. 4c,d** and **Supplementary Figs. 16** and **17**). This observation is consistent with the gain of leukemogenic function due to *SETBP1* mutation. As with overexpressed wild-type *Setbp1*, homeobox genes *Hoxa9* and *Hoxa10* represent critical targets of Setbp1 mutants, as cells immortalized by wild-type or mutant *Setbp1* expressed comparable levels of the corresponding mRNAs, and knockdown of either caused a marked reduction in colony-forming potential (**Supplementary Figs. 18** and **19**). In agreement with these findings, *SETBP1*-mutant leukemias ($n = 14$) showed significantly higher *HOXA9* and *HOXA10* expression levels compared to wild-type cases without *SETBP1* overexpression ($n = 9$; $P = 0.03$ and 0.03, respectively), supporting the notion that *HOXA9* and *HOXA10* are likely functional targets of mutated SETBP1 in myeloid neoplasms (**Supplementary Fig. 20**).

Multiple mechanisms could contribute to the enhanced oncogenic properties of *SETBP1* mutations. For instance, mutation could increase protein stability (**Supplementary Fig. 21**), resulting in greater protein amounts (analogous to upmodulation of *SETBP1* mRNA), in agreement with a previously reported observation[1]. However, we also showed that *SETBP1* mRNA overexpression *in vitro* was associated with the immortalization of progenitors and that there were primary cases of sAML with and without mutations of *SETBP1* and high levels of wild-type mRNA. Thus, although plausible, the mechanisms underlying increased *SETBP1* expression and its proto-oncogenic role may be more complicated. It is also possible that interaction of Ski and/or SnoN with SETBP1 through the SKI homology region could be affected by mutations, leading to transformation[20,32]. *SETBP1* was shown to regulate PP2A activity via binding to SET[20], and decreased PP2A activity has been described in AML[21,33]. In fact, we observed that mutant *Setbp1*–immortalized myeloid progenitors had increased tyrosine phosphorylation of Ppp2ca compared to myeloid progenitors immortalized with wild-type *Setbp1* (**Supplementary Fig. 22**), suggesting that *SETBP1* mutations could cause further PP2A inhibition.

In summary, recurrent somatic *SETBP1* mutations are new lesions that interact with previously defined pathways underlying poor prognosis and provide new insights into the process of leukemic evolution. The apparent association of *SETBP1* mutations with poor clinical outcome observed here provides an important focal point for future mechanistic studies as well as a goal for therapeutic targeting.

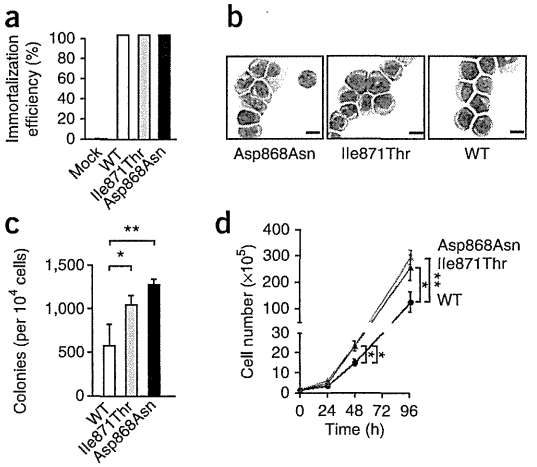

Figure 4 Immortalization of mouse myeloid progenitors by *SETBP1* mutations. (**a**) Efficiencies of empty pMYs retroviral vector (mock) or of pMYs constructs expressing wild-type Setbp1 (WT) or Setbp1 mutants (Asp868Asn and Ile871Thr) in immortalizing mouse myeloid progenitor cells in three independent experiments. (**b**) Wright-Giemsa staining of cells immortalized by transduction with retroviruses expressing the indicated wild-type or mutant Setbp1 proteins. Scale bars, 50 μm. (**c**) Mean and s.d. values for the colony-forming potentials of mouse myeloid progenitors immortalized by the expression of wild-type or mutant Setbp1 on methylcellulose medium in the presence of stem cell factor (SCF) (100 ng/ml) and interleukin (IL)-3 (20 ng/ml). Combined results from three independent myeloid progenitor populations immortalized by each retroviral construct are shown. (**d**) Expansion of myeloid progenitors immortalized by expression of wild-type or mutant Setbp1 in liquid medium with SCF and IL-3 over a 96-h period. Results from three independent populations immortalized by each retroviral construct are presented. Cell numbers were counted by trypan blue staining. Error bars, s.d. $*P < 0.05$, $**P < 0.005$; *t* tests were used for comparisons between strains.