Wnt3a stimulates maturation of impaired neutrophils developed from severe congenital neutropenia patient-derived pluripotent stem cells

Takafumi Hiramoto[a,b], Yasuhiro Ebihara[b,c,1], Yoko Mizoguchi[d], Kazuhiro Nakamura[d], Kiyoshi Yamaguchi[e], Kazuko Ueno[f], Naoki Nariai[f], Shinji Mochizuki[b,c], Shohei Yamamoto[b,c], Masao Nagasaki[f], Yoichi Furukawa[e], Kenzaburo Tani[a], Hiromitsu Nakauchi[g], Masao Kobayashi[d], and Kohichiro Tsuji[b,c]

[a]Division of Molecular and Clinical Genomics, Medical Institute of Bioregulation, Kyushu University, Higashi-ku, Fukuoka 812-8582, Japan; [c]Department of Pediatric Hematology/Oncology, Research Hospital, Divisions of [b]Stem Cell Processing and [g]Stem Cell Therapy, Center for Stem Cell Biology and Regenerative Medicine, and [e]Division of Clinical Genome Research, Advanced Clinical Research Center, Institute of Medical Science, University of Tokyo, Minato-ku, Tokyo 108-8639, Japan; [d]Pediatrics, Hiroshima University Graduate School of Biomedical and Health Sciences, Minami-ku, Hiroshima 734-8551, Japan; and [f]Department of Integrative Genomics, Tohoku Medical Megabank Organization, Tohoku University, Aramaki, Aoba-ku, Sendai 980-8573, Japan

Edited by George Q. Daley, Children's Hospital Boston, Boston, MA, and accepted by the Editorial Board January 4, 2013 (received for review October 1, 2012)

The derivation of induced pluripotent stem (iPS) cells from individuals of genetic disorders offers new opportunities for basic research into these diseases and the development of therapeutic compounds. Severe congenital neutropenia (SCN) is a serious disorder characterized by severe neutropenia at birth. SCN is associated with heterozygous mutations in the neutrophil elastase [elastase, neutrophil-expressed (ELANE)] gene, but the mechanisms that disrupt neutrophil development have not yet been clarified because of the current lack of an appropriate disease model. Here, we generated iPS cells from an individual with SCN (SCN-iPS cells). Granulopoiesis from SCN-iPS cells revealed neutrophil maturation arrest and little sensitivity to granulocyte-colony stimulating factor, reflecting a disease status of SCN. Molecular analysis of the granulopoiesis from the SCN-iPS cells vs. control iPS cells showed reduced expression of genes related to the wingless-type mmtv integration site family, member 3a (Wnt3a)/β-catenin pathway [e.g., lymphoid enhancer-binding factor 1], whereas Wnt3a administration induced elevation lymphoid enhancer-binding factor 1-expression and the maturation of SCN-iPS cell-derived neutrophils. These results indicate that SCN-iPS cells provide a useful disease model for SCN, and the activation of the Wnt3a/β-catenin pathway may offer a novel therapy for SCN with ELANE mutation.

apoptosis | unfolded protein response | SCN disease model

Severe congenital neutropenia (SCN) is a heterogeneous bone marrow (BM) failure syndrome characterized by severe neutropenia at birth, leading to recurrent infections by bacteria or fungi (1). SCN patients reveal an arrest in neutrophil differentiation in the BM at the promyelocyte or myelocyte stage (1), as well as a propensity to develop myelodysplastic syndrome and acute myeloid leukemia (2). Current treatment by high-dose granulocyte-colony stimulating factor (G-CSF) administration induces an increase in the number of mature neutrophils in the peripheral blood of most SCN patients (3). Although this treatment is curative for the severe infections, there is a concern that high-dose G-CSF may increase the risk of hematologic malignancy in these individuals (4).

Several genetic mutations have been identified in SCN patients. Approximately 50% of autosomal-dominant SCN cases were shown to have various heterozygous mutations in the gene encoding neutrophil elastase [elastase, neutrophil-expressed (ELANE)] (5, 6), a monomeric, 218-amino acid (25 kDa) chymotryptic serine protease (7) that is synthesized during the early stages of primary granule production in promyelocytes (8, 9). However, the mechanism(s) causing impaired neutrophil maturation in SCN patients remains unclear due to the current lack of an appropriate disease model.

Results and Discussion

In the present study, we generated induced pluripotent stem (iPS) cells from the BM cells obtained from an SCN patient with a heterologous ELANE gene mutation (exon 5, 707 region, C194X) (SCN-iPS cells) to provide the basis for an SCN disease model. The patient who donated BM cells recurrently suffered from severe infections without exogenous G-CSF administration, but the G-CSF administration once a week prevented his repeated infection. The SCN-iPS cells continued to show embryonic stem cell morphology after >20 passages and also expressed pluripotent markers (Fig. S1*A*). The silencing of exogenous genes and the capability to differentiate into three germ layers by teratoma formation were confirmed for each of the three SCN-iPS cell clones (Fig. S1 *B* and *C*). Furthermore, the same ELANE gene mutation that was present in the patient persisted in the SCN-iPS cells (Fig. S1*D*). The SCN-iPS cells, as well as control iPS cells that were generated from healthy donors, had the normal karyotype (Fig. S1*E*) (10, 11) and no mutations in the mutation-sensitive region of the G-CSF receptor gene (12).

We first compared the hematopoietic differentiation from SCN-iPS cells with that from control iPS cells that were generated from healthy donors. SCN-iPS and control iPS cells were cocultured with a 15-Gy-irradiated murine stromal cell line (the AGM-S3 cell line), as reported (13). After 12 d, the cocultured cells were harvested, and the $CD34^+$ cells separated from these cells (SCN-iPS–$CD34^+$ and control iPS–$CD34^+$ cells, respectively) were cultured in a hematopoietic colony assay by using a cytokine mixture (*Materials and Methods*). The number and size of the erythroid (E) and mixed-lineage (Mix) colonies derived from SCN-iPS–$CD34^+$ cells (1×10^4 cells) were nearly identical to those of the corresponding colonies derived from control iPS–$CD34^+$ cells (E colonies: SCN-iPS cells, 11.0 ± 3.0, and control iPS cells, 11.4 ± 3.9; Mix colonies: SCN-iPS cells, 25.1 ± 7.2, and control iPS cells, 17.4 ± 4.0) (Fig. 1 *B* and *C* and Fig. S2 *A* and *B*). However, the number of myeloid colonies derived from SCN-iPS–$CD34^+$ vs. control iPS–$CD34^+$ cells was significantly lower (SCN-iPS cells, 47.4 ± 19.5; control iPS cells, 127.8 ± 17.9; $P < 0.01$), and the size of the colonies was also smaller (Fig. 1 *A* and *D*). In particular, only a few SCN-iPS cell-derived granulocyte (G) colonies—myeloid colonies consisting of only granulocytes—were detected (Fig. 1*A*). SCN-iPS cell-derived granulocyte–macrophage (GM) colonies—myeloid colonies consisting of macrophages/monocytes with/without granulocytes—contained a few immature myeloid cells in addition to macrophages/monocytes, whereas control iPS cell-derived GM colonies included a substantial number of mature, segmented, and band neutrophils (Fig. 1*D*).

We also found that Mix colonies derived from SCN-iPS cells, but not control iPS cells, contained immature myeloid cells and few mature neutrophils (Fig. S2 *C* and *D*). Next, we conducted a hematopoietic colony assay using various concentrations of G-CSF alone instead of the cytokine mixture to examine the G-CSF dose dependency of neutrophil differentiation from SCN-iPS and control iPS–$CD34^+$ cells. For all concentrations of G-CSF used (1–1,000 ng/mL), the SCN-iPS cell-derived myeloid colonies were significantly lower in number and smaller in size than the control iPS cell-derived myeloid colonies (Fig. 1*E*). Myeloid colony formation from control iPS cells reached a plateau at ~1–10 ng/mL G-CSF, whereas the number and size of those from SCN-iPS cells gradually increased with increasing concentrations of G-CSF. However, the values observed for SCN-iPS cells did not reach those for the control iPS cells, even at the highest dose of G-CSF used (1,000 ng/mL). Furthermore, large colonies consisting of >100 cells derived from SCN-iPS cells were only found with higher concentrations of G-CSF (Fig. 1*F*). Thus, granulopoiesis initiated from SCN-iPS cells was relatively insensitive to G-CSF, reflecting the inadequate in vivo response of neutrophils to G-CSF in SCN patients (14, 15). Therefore, these results support the applicability of the SCN-iPS cells established herein as a disease model for SCN.

To examine neutrophil development from SCN-iPS cells in more detail, SCN-iPS and control iPS–$CD34^+$ cells (1×10^4 cells each) were cocultured in suspension with AGM-S3 cells in the presence of neutrophil differentiation medium (*SI Materials and Methods*). The number of nonadherent cells derived from SCN-iPS–$CD34^+$ cells was lower than that from control iPS–$CD34^+$ cells on day 14 of culture (SCN-iPS cells, $9.77 \times 10^4 \pm 1.65 \times 10^4$ cells; control iPS cells, $52.48 \times 10^4 \pm 23.13 \times 10^4$ cells; $P < 0.05$) (Fig. 2*A*). The proportion of mature neutrophils among the nonadherent cells was also significantly lower for SCN-iPS cells relative to control iPS cells on day 14 (SPN-iPS cells, 15.53% ± 4.33%; control iPS cells, 71.285 ± 3.30%; $P < 0.05$) (Fig. 2 *B* and *C*), indicating that myeloid cells derived from SCN-iPS cells revealed the maturation arrest in the neutrophil development. We then examined a possibility that the maturation arrest in SCN-

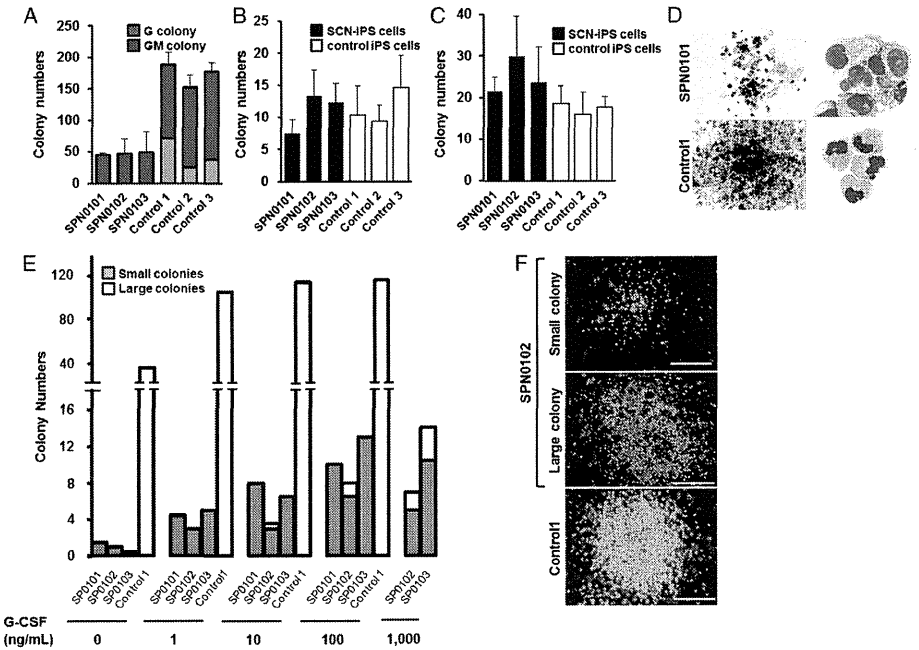


Fig. 1. Impaired neutrophil development from SCN-iPS cells. (*A–C*) A hematopoietic colony assay was performed by using 1×10^4 $CD34^+$ cells derived from three SCN-iPS cell clones (SPN0101, SPN0102, and SPN0103) and three control iPS cell clones (controls 1, 2, and 3) in the presence of a cytokine mixture. Colonies were sorted as myeloid (*A*), erythroid (*B*), and mixed-lineage (Mix) (*C*). Data are shown as mean ± SD. (*D*) Photographs of colonies (*Left*; 100x) and cells in a GM colony (*Right*; 400x; May–Grünwald–Giemsa staining). (*E*) A hematopoietic colony assay with dose escalation of G-CSF was performed by using 1×10^5 $CD34^+$ cells derived from SCN-iPS and control iPS cells. Filled and open bars indicate small colonies consisting of <100 cells and large colonies consisting of >100 cells, respectively. Data are shown as the average of three independent experiments. (*F*) Photographs of a small colony derived from SCN-iPS cells (SPN0102) in the presence of 10 ng/mL G-CSF, large colonies derived from SCN-iPS cells in the presence of 1,000 ng/mL G-CSF, and large colonies derived from control iPS cells (control 1) in the presence of 10 ng/mL G-CSF. (Scale bars, 200 μm.)

Author contributions: T.H., Y.E., K.Y., S.M., S.Y., Y.F., K. Tani, H.N., M.K., and K. Tsuji designed research; T.H., Y.M., K.N., and K.Y. performed research; T.H., Y.E., Y.M., K.N., K.Y., K.U., N.N., S.M., S.Y., M.N., and K. Tsuji analyzed data; and T.H., Y.E., and K. Tsuji wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. G.Q.D. is a guest editor invited by the Editorial Board.

[1]To whom correspondence should be addressed. E-mail: ebihara@ims.u-tokyo.ac.jp.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1217039110/-/DCSupplemental.

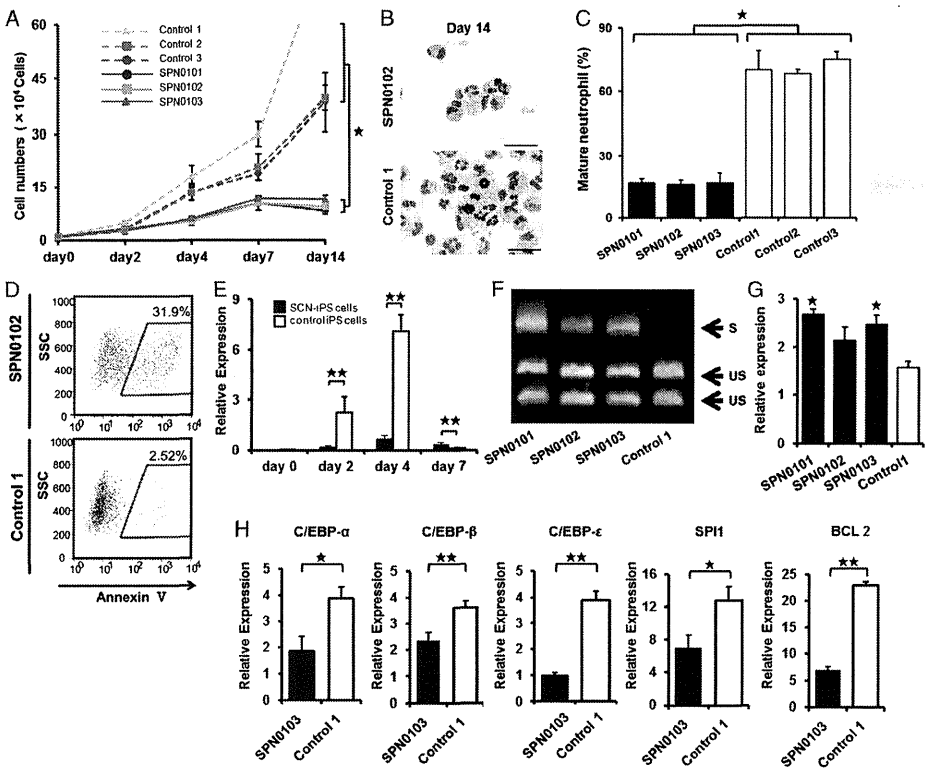


Fig. 2. Analysis of impaired neutrophil development from SCN-iPS cells. (*A*) Total number of nonadherent cells in the suspension culture of 1×10^4 CD34$^+$ cells derived from SCN-iPS and control iPS cells. Data are shown as mean ± SD. $^*P < 0.01$. (*B*) Photographs of nonadherent cells derived from SCN-iPS (SPN0103) and control iPS cells (control 1) on day 14 of culture (400×; May–Grünwald–Giemsa staining; scale bars, 50 μm.) (*C*) Filled and open bars show the proportion of mature neutrophils among the cells derived from SCN-iPS (filled bars) and control iPS (open bars) cells on day 14 of suspension culture. Data are shown as mean ± SD. $^*P < 0.05$. (*D*) Flow cytometric analysis of annexin V expression on cultured cells from SCN-iPS cells (SPN0102) or control iPS cells (control 1) on day 7. (*E*) Sequential qRT-PCR analysis of the relative expression of ELANE mRNA [ELANE/hypoxanthine–guanine phosphoribosyltransferase (HPRT) expression]. Data obtained from independent experiments using three SCN-iPS cell clones (SPN0101, SPN0102, and SPN0103) and three control iPS cell clones are shown as mean ± SD. $^{**}P < 0.01$. (*F* and *G*) CD34$^+$ cells derived from SCN-iPS or control iPS cells were cultured in neutrophil differentiation medium (see text). On day 7, nonadherent cells were collected and analyzed. (*F*) Representative gel showing spliced (S) and unspliced (US) XBP-1 bands on day 7. (*G*) qRT-PCR analysis of the relative mRNA expression (target/HPRT expression) of BiP on day 7. Data are shown as mean ± SD. $^*P < 0.05$; different from control 1). (*H*) qRT-PCR analysis of the relative mRNA expression (target / HPRT expression) of C/EBP-α, C/EBP-β, C/EBP-ε, SPI1, and BCL2 genes in non-adherent cells derived from SCN-iPS cells (filled bars, SPN0103) and control iPS cells (open bars, control 1) on day 2 of suspension culture. Data are shown as the mean ± the s.d. ($^{**}P < 0.01$, $^*P < 0.05$).

iPS cell-derived myeloid cells might be caused by their apoptosis. In flow cytometric analysis, SCN-iPS cell-derived myeloid cells contained a significantly higher proportion of annexin V-positive cells than control iPS-derived myeloid cells on day 7 of culture, suggesting that the maturation arrest in myeloid cells derived from SCN-iPS cells might be caused by their apoptosis (Fig. 2*D*).

We next examined ELANE mRNA expression levels in nonadherent cells derived from SCN-iPS vs. control iPS cells (Fig. 2*E*). ELANE expression was significantly lower in nonadherent cells derived from SCN-iPS vs. control iPS cells on days 2 and 4 of culture ($P < 0.01$), as reported (16, 17). However, the former was a little higher than the latter on day 7 ($P < 0.01$). This result may be explained by the existence of SCN-iPS cell-derived myeloid cells arrested at an early stage along the neutrophil differentiation pathway even on day 7 of culture. We also examined the expression of proteinase 3 and azurocidin, which comprise a family of closely related genes encoding neutrophil granule proteins along with ELANE, and found these genes were more highly expressed on day 4 (Fig. S3).

It has been reported that induction of the endoplasmic reticulum stress (ER) response and the unfolded protein response (UPR) has been advanced as a potential explanation for the molecular pathogenesis of SCN (18, 19). Thus, we examined activation of the UPR by X-box binding protein 1 (XBP-1) mRNA splicing on day 7. As shown in Fig. 2*F*, SPN-iPS cells induced XBP-1 mRNA splicing. We also found the up-regulation of BiP (also known as GRP78 or HSPA5) (Fig. 2*G*). These results suggested that ER stress response and UPR might be involved in the pathogenesis in SCN.

To examine further the differences in gene expression between the two cell types, a microarray analysis was carried out by using CD34$^+$ cells derived from SCN-iPS and control iPS cells (three clones of each) in suspension culture on day 2. At this early time point, differences in cell number and morphology were not yet readily discernible between SCN-iPS and control iPS cells, as shown in Fig. 2*A*. However, the microarray analysis revealed a differential expression of various genes between the two cell types. Transcription factor genes, which were related to neutrophil development [e.g., CCAAT/enhancer-binding protein (C/EBP)-α (20), C/EBP-β (21), C/EBP-ε (22), and SPI1 (also known as PU.1) (23)], were all down-regulated in SCN-iPS cells. B-cell chronic lymphocytic leukemia/lymphoma 2, which regulates cell death under ER stress through the core mitochondrial apoptosis pathway (24), was also down-regulated (Fig. 3*A*). These findings were confirmed by quantitative reverse-transcriptional PCR (qRT-PCR), as shown in Fig. 2*H*.

Notably, the down-regulation of the genes in SCN-iPS cells related to and regulated by the wingless-type mmtv integration site family, member 3a (Wnt3a)/β-catenin pathway [e.g., Wnt3a, lymphoid enhance-binding factor (LEF)-1, BIRC5 (also known as survivin), and cyclin D1] was also uncovered by microarray analysis and qRT-PCR (Fig. 3 *A–C* and Fig. S4). Therefore, we examined the effect of enhancement of Wnt3a/β-catenin signaling by exogenous Wnt3a addition on the neutrophil development of CD34$^+$ cells derived from SCN-iPS and control iPS cells. Although Wnt3a did not stimulate the survival, proliferation, and differentiation of CD34$^+$ cells derived from both iPS cells in the absence of cytokines stimulating myelopoiesis including G-CSF, the addition of Wnt3a to the neutrophil differentiation medium induced a dose-dependent increase in the percentage of mature neutrophils among the cultured cells, as shown in Fig. 3 *D* and *E*. Furthermore, when Wnt3a was added concurrently with 1,000ng/mL G-CSF, the proportion of mature neutrophils increased more than it did with Wnt3a or 1,000 ng/mL G-CSF alone, reaching a value comparable with that observed for control iPS cells (Fig. 4 *A* and *B*).

The reduced expression of LEF-1 (as regulated by the Wnt3a/β-catenin pathway) reportedly plays a critical role in the defective maturation of neutrophils in SCN patients (25). Therefore, we next examined LEF-1 mRNA expression in SCN-iPS–CD34$^+$ cells cultured in the presence of Wnt3a, G-CSF (1,000 ng/mL), or both. Wnt3a and G-CSF both enhanced LEF-1 mRNA expression, but the most significant increase was observed in the presence of Wnt3a plus G-CSF. LEF-1 expression in SCN-iPS–CD34$^+$ cells in response to Wnt3a plus G-CSF was almost the same as that in control iPS–CD34$^+$ cells (Fig. 4*C*). These results substantiate the importance of LEF-1 in neutrophil development and the pathogenesis of SCN, as shown (25). Moreover the

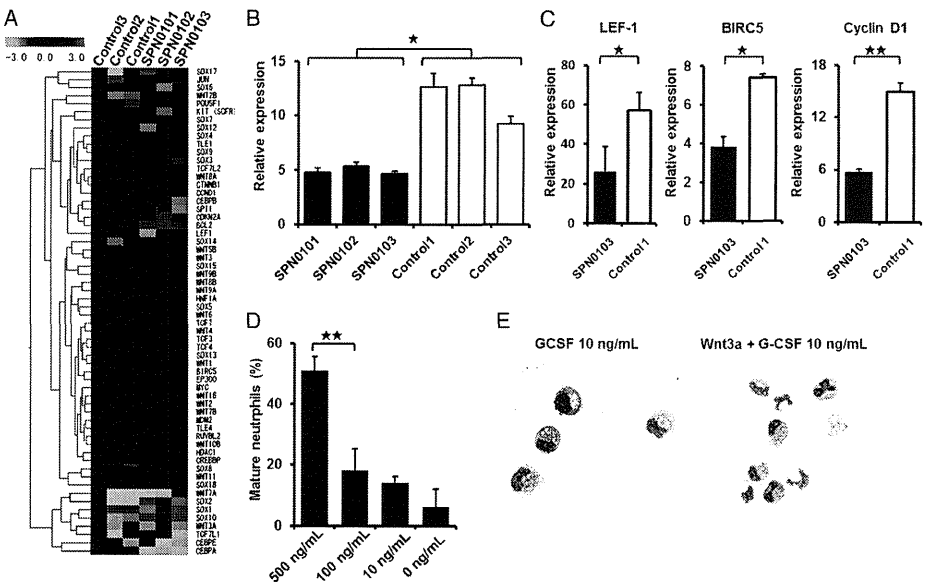


Fig. 3. Effects of Wnt3a on neutrophil development from SCN-iPS cells. (*A*) Heat map showing differential gene expression among SCN-iPS and control iPS cells on day 2. Red, high gene expression; blue, low gene expression compared with gene expression in control 3. (*B*) qRT-PCR analysis of the relative mRNA expression (target/HPRT expression) of Wnt3a on day 2. Filled and open bars indicate experiments using SCN-iPS cells (SPN0101, SPN0102, and SPN0103) and control iPS cells (controls 1, 2, and 3), respectively. Data are shown as mean ± SD. $^*P < 0.05$. (*C*) qRT-PCR analysis of the relative expression (target/HPRT expression) of genes regulated by the Wnt3a/β-catenin pathway (LEF-1, survivin, and cyclin D1) in SCN-iPS cells (filled bars, SPN0103) vs. control iPS cells (open bars, control 1) on day 2 of suspension culture. Data are shown as mean ± SD. $^{**}P < 0.01$; $^*P < 0.05$. (*D*) Proportion of mature neutrophils among the cells derived from SCN-iPS cells (SPN0102) on day 14 of suspension culture with dose escalation of Wnt3a. Data are shown as mean ± SD. $^{**}P < 0.01$. (*E*) Photographs of nonadherent cells on day 7 of suspension culture with or without Wnt3a (500 ng/mL) (400×; May–Grünwald–Giemsa staining).

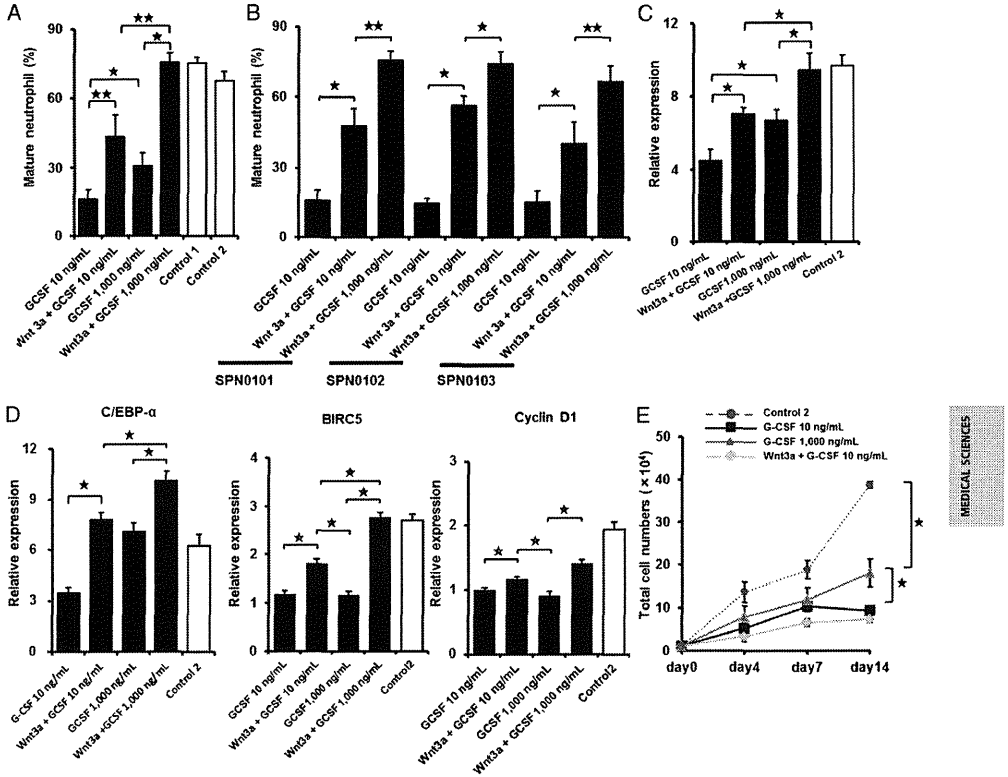


Fig. 4. Effects of Wnt3a in combination with high-dose G-CSF. (*A*) Filled and open bars show the proportion of mature neutrophils among the cells derived from SCN-iPS cells (SPN0101) on day 14 of suspension culture in the presence of neutrophil differentiation medium containing 10 ng/mL G-CSF (GCSF 10 ng/mL); 500 ng/mL Wnt3a and 10 ng/mL G-CSF (Wnt3a+GCSF 10 ng/mL); 1,000 ng/mL G-CSF (GCSF 1,000 ng/mL); or 500 ng/mL Wnt3a and 1,000 ng/mL G-CSF (Wnt3a + GCSF 1,000 ng/mL); and that from control iPS cells (controls 1 and 2) cultured in the neutrophil differentiation medium containing 10 ng/mL G-CSF, respectively. Data are shown as mean ± SD. $^{**}P < 0.01$; $^{*}P < 0.05$. (*B*) The proportion of mature neutrophils among the cells derived from three SCN-iPS cell clones (SPN0101, SPN0102, and SPN0103) on day 14 of suspension culture in the presence of neutrophil differentiation medium containing 10 ng/mL G-CSF (GCSF 10 ng/mL); 500 ng/mL Wnt3a and 10 ng/mL G-CSF (Wnt3a+GCSF 10 ng/mL); or 500 ng/mL Wnt3a and 1,000 ng/mL G-CSF (Wnt3a + GCSF 1,000 ng/mL). Data are shown as mean ± SD. $^{**}P < 0.01$; $^{*}P < 0.05$. (*C*) Filled and open bars show the relative expression (target/HPRT expression) of LEF-1 mRNA in SCN-iPS cells (SPN0101) on day 2 of suspension culture in the presence of differentiation medium containing the same combinations of Wnt3a and G-CSF as shown in *A* and that from control iPS cells (control 2), respectively. Data are shown as mean ± SD. $^{**}P < 0.01$; $^{*}P < 0.05$. (*D*) Filled and open bars show the relative expression (target/HPRT expression) of C/EBP-α, BIRC5, or cyclin D1 mRNA in SCN-iPS cells (SPN0101) on day 2 of suspension culture in the presence of differentiation medium containing the same combinations of Wnt3a and G-CSF as shown in *A* and that from control iPS cells (control 2), respectively. Data are shown as mean ± SD. $^{**}P < 0.01$; $^{*}P < 0.05$. (*E*) Total cell numbers of nonadherent cells in suspension cultures of 1×10^4 $CD34^+$ cells derived from control iPS cells (control 2; red broken line) and SCN-iPS cells (SPN0101) in the presence of neutrophil differentiation medium (black line) and those from SCN-iPS cells in the presence of neutrophil differentiation medium containing 500 ng/mL Wnt3a (yellow line) or 1,000 ng/mL G-CSF (black line). Data are shown as mean ± SD. $^{**}P < 0.05$.

administration of Wnt3a led to up-regulate C/EBP-α, cyclin D1, and BIRC5/survivin in addition to LEF-1 in the presence of G-CSF (Fig. 4*D*). These results suggested that the up-regulation of LEF-1 expression might promote granulopoiesis by increasing the expressions of cyclin D1, BIRC5/survivin, and C/EBP-α and its binding to LEF-1 in accordance with the previous report (25). Interestingly, Wnt3a did not stimulate the proliferation of myeloid cells, whereas 1,000 ng/mL G-CSF did to a certain extent (Fig. 4*E*). Hence, Wnt3a was capable of stimulating the maturation of impaired neutrophils in the presence of G-CSF, but not the proliferation of myeloid cells from SCN-iPS cells.

Importantly, aside from providing new insights into the mechanisms behind impaired neutrophil development in SCN patients, the present study demonstrates that agents activating the Wnt3a/β-catenin pathway are potential candidates for new drugs for SCN with mutations in the ELANE gene. Because endogenous G-CSF is readily increased in SCN patients (26), these activating agents may be viable alternatives to exogenous G-CSF treatment.

Materials and Methods

Additional information is available in *SI Materials and Methods*.

Generation of Human iPS Cells. BM fibroblasts from a patient with SCN and skin dermal fibroblasts from a healthy donor were acquired after obtaining informed consent after getting the approval by the Ethics Committee of the Institute of Medical Science, University of Tokyo, in accordance with the Declaration of Helsinki. The SCN patient presented with a heterozygous mutation in the ELANE gene in the 707 region of exon 5. SCN-iPS cells were established from the SCN-BM fibroblasts by transfection with the pMX retroviral vector, as described (10). This vector expressed the human transcription factors OCT3/4, SOX2, KLF4, and c-MYC. Control iPS cell clones, control 1 (TkDN4-M) and control 3 (201B7), were gifts from K. Eto and S. Yamanaka (Kyoto University, Kyoto), respectively (10, 11). Control 2 (SPH0101) was newly generated from another healthy donor's skin dermal fibroblasts by using the same methods.

Hematopoietic Colony Assay. A hematopoietic colony assay was performed in an aliquot of culture mixture, which contained 1.2% methylcellulose (*Shin-Etsu Chemical*), 30% (*vol/vol*) *FBS*, 1% (*vol/vol*) *deionized fraction* V BSA, 0.1 mM 2-mercaptoethanol (2-ME), α-minimum essential medium, and a cytokine mixture consisting of 100 ng/mL human stem cell factor (hSCF) (Wako), 100 ng/mL fusion protein 6 [FP6; a fusion protein of interleukin (IL)-6 and IL-6 receptor] (a gift from Tosoh), 10 ng/mL human IL-3 (hIL-3) (a gift from Kirin Brewery), 10 ng/mL human thrombopoietin (hTPO) (a gift from Kirin Brewery), 10 ng/mL human G-CSF (a gift from Chugai Pharmaceutical), and 5 U/mL human erythropoietin (a gift from Kirin Brewery). For dose escalation experiments, various concentrations (0, 1, 10, 100, and 1,000 ng/mL) of G-CSF were used instead of the cytokine mixture described above. Colony types were determined according to established criteria on day 14 of culture by in situ observations under an inverted microscope (IX70; Olympus) (27).

Suspension Culture and Neutrophil Differentiation Assay. $CD34^+$ cells (1×10^4 cells) were cocultured with irradiate confluent AGM-S3 cells in neutrophil differentiation medium containing Iscove's modified Dulbecco's medium, 10% FBS, 3 mM L-glutamine, 1×10^{-4} M 2-ME, 1×10^{-4} M nonessential amino acids solution, 100 ng/mL hSCF, 100 ng/mL FP6, 10 ng/mL hIL-3, 10 ng/mL hTPO, and 10 or 1,000 ng/mL human G-CSF. Wnt3a (10, 100, or 500 ng/mL) (R&D) was then added. The medium was replaced with an equivalent volume of fresh medium every 4 d. Living, nonadherent cells were counted following 0.4% trypan blue staining.

PCR primer. All primer sets used in this study are shown in Table S1.

Statistical Analysis. All data are presented as mean ± SD. $P < 0.05$ was considered significant. Statistical analyses were performed by using Prism software (*GraphPad*).

ACKNOWLEDGMENTS. We thank the individual with SCN who participated in this study; K. Eto for providing control iPS cells (control 1; TkDN4-M); S. Yamanaka for providing control iPS cells (control 3; 206B7); and E. Matsuzaka and S. Hanada for technical assistance. This work was supported by in part by Ministry of Education, Culture, Sports, Science, and Technology of Japan (MEXT) Grants-in-Aid (to Y.E.) and Project for Realization of Regenerative Medicine (MEXT) Grants-in-Aid (to K.Tsuji).

1. Zeidler C, Germeshausen M, Klein C, Welte K (2009) Clinical implications of *ELA2*-, *HAX1*-, and G-CSF-receptor (*CSF3R*) mutations in severe congenital neutropenia. *Br J Haematol* 144(4):459–467.
2. Freedman MH, et al. (2000) Myelodysplasia syndrome and acute myeloid leukemia in patients with congenital neutropenia receiving G-CSF therapy. *Blood* 96(2):429–436.
3. Dale DC, et al. (1993) A randomized controlled phase III trial of recombinant human granulocyte colony-stimulating factor (filgrastim) for treatment of severe chronic neutropenia. *Blood* 81(10):2496–2502.
4. Rosenberg PS, et al.; Severe Chronic Neutropenia International Registry (2006) The incidence of leukemia and mortality from sepsis in patients with severe congenital neutropenia receiving long-term G-CSF therapy. *Blood* 107(12):4628–4635.
5. Xia J, et al. (2009) Prevalence of mutations in ELANE, GFI1, HAX1, SBDS, WAS and G6PC3 in patients with severe congenital neutropenia. *Br J Haematol* 147(4):535–542.
6. Horwitz MS, et al. (2007) Neutrophil elastase in cyclic and severe congenital neutropenia. *Blood* 109(5):1817–1824.
7. Hajjar E, Broemstrup T, Kantari C, Witko-Sarsat V, Reuter N (2010) Structures of human proteinase 3 and neutrophil elastase—so similar yet so different. *FEBS J* 277(10): 2238–2254.
8. Fouret P, et al. (1989) Expression of the neutrophil elastase gene during human bone marrow cell differentiation. *J Exp Med* 169(3):833–845.
9. Pham CT (2006) Neutrophil serine proteases: Specific regulators of inflammation. *Nat Rev Immunol* 6(7):541–550.
10. Takayama N, et al. (2010) Transient activation of c-MYC expression is critical for efficient platelet generation from human induced pluripotent stem cells. *J Exp Med* 207(13):2817–2830.
11. Takahashi K, et al. (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131(5):861–872.
12. Germeshausen M, Ballmaier M, Welte K (2007) Incidence of CSF3R mutations in severe congenital neutropenia and relevance for leukemogenesis: Results of a long-term survey. *Blood* 109(1):93–99.
13. Ma F, et al. (2007) Novel method for efficient production of multipotential hematopoietic progenitors from human embryonic stem cells. *Int J Hematol* 85(5):371–379.
14. Konishi N, et al. (1999) Defective proliferation of primitive myeloid progenitor cells in patients with severe congenital neutropenia. *Blood* 94(12):4077–4083.
15. Nakamura K, et al. (2000) Abnormalities of primitive myeloid progenitor cells expressing granulocyte colony-stimulating factor receptor in patients with severe congenital neutropenia. *Blood* 96(13):4366–4369.
16. Skokowa J, Fobiwe JP, Dan L, Thakur BK, Welte K (2009) Neutrophil elastase is severely down-regulated in severe congenital neutropenia independent of ELA2 or HAX1 mutations but dependent on LEF-1. *Blood* 114(14):3044–3051.
17. Kawaguchi H, et al. (2003) Dysregulation of transcriptions in primary granule constituents during myeloid proliferation and differentiation in patients with severe congenital neutropenia. *J Leukoc Biol* 73(2):225–234.
18. Köllner I, et al. (2006) Mutations in neutrophil elastase causing congenital neutropenia lead to cytoplasmic protein accumulation and induction of the unfolded protein response. *Blood* 108(2):493–500.
19. Grenda DS, et al. (2007) Mutations of the ELA2 gene found in patients with severe congenital neutropenia induce the unfolded protein response and cellular apoptosis. *Blood* 110(13):4179–4187.
20. Pabst T, et al. (2001) AML1-ETO downregulates the granulocytic differentiation factor C/EBPalpha in t(8;21) myeloid leukemia. *Nat Med* 7(4):444–451.
21. Hirai H, et al. (2006) *C/EBPbeta is required for 'emergency' granulopoiesis. Nat Immunol* 7(7):732–739.
22. Bedi R, Du J, Sharma AK, Gomes I, Ackerman SJ (2009) Human C/EBP-ε activator and repressor isoforms differentially reprogram myeloid lineage commitment and differentiation. *Blood* 113(2):317–327.
23. Friedman AD (2007) Transcriptional control of granulocyte and monocyte development. *Oncogene* 26(47):6816–6828.
24. Hetz C (2012) The unfolded protein response: Controlling cell fate decisions under ER stress and beyond. *Nat Rev Mol Cell Biol* 13(2):89–102.
25. Skokowa J, et al. (2006) LEF-1 is crucial for neutrophil granulocytopoiesis and its expression is severely reduced in congenital neutropenia. *Nat Med* 12(10):1191–1197.
26. Mempel K, Pietsch T, Menzel T, Zeidler C, Welte K (1991) Increased serum levels of granulocyte colony-stimulating factor in patients with severe congenital neutropenia. *Blood* 77(9):1919–1922.
27. Nakahata T, Ogawa M (1982) Hemopoietic colony-forming cells in umbilical cord blood with extensive capability to generate mono- and multipotential hemopoietic progenitors. *J Clin Invest* 70(6):1324–1328.

Int J Hematol (2013) 97:20–29
DOI 10.1007/s12185-012-1249-9

Inherited bone marrow failure syndromes in 2012

Hirotoshi Sakaguchi · Koji Nakanishi · Seiji Kojima

Received: 11 October 2012 / Revised: 7 December 2012 / Accepted: 10 December 2012 / Published online: 28 December 2012

Abstract Inherited bone marrow failure syndromes (CBMFS) are a heterogeneous group of genetic disorders characterized by bone marrow failure, congenital anomalies, and an increased risk of malignant disease. The representative diseases with trilineage involvement are Fanconi anemia and dyskeratosis congenita, while the disease with the single lineage cytopenia is Diamond–Blackfan anemia. Recent advances in our understanding of these diseases have come from the identification of genetic lesions responsible for the disease and their pathways. Although recent studies have identified many causative genes, mutations of these genes have only been found in less than half of the patients. Next-generation sequencing technologies may reveal new causative genes in these patients. Also, induced pluripotent stem cells derived from patients with CBMFS will be useful to study the pathophysiology of the diseases. The only long-term curative treatment for bone marrow failure in patients with inherited bone marrow failure syndromes is allogeneic hematopoietic stem cell transplantation, although this procedure has a risk of severe adverse effects. Multicenter prospective studies are warranted to establish appropriate conditioning regimens aimed at reducing transplant-related mortality.

Keywords Inherited bone marrow failure syndrome · Fanconi anemia · Diamond–Blackfan anemia · Dyskeratosis congenita

H. Sakaguchi · K. Nakanishi · S. Kojima (✉)
Department of Pediatrics,
Nagoya University Graduate School of Medicine,
65 Tsurumai-cho, Showa-ku, Nagoya 466-8550, Japan
e-mail: kojimas@med.nagoya-u.ac.jp

Introduction

Inherited bone marrow failure syndromes (IBMFS) are a heterogeneous group of genetic disorders characterized by bone marrow failure, congenital anomalies, and increased risk of malignant disease. Such bone marrow failure may affect all three hematopoietic cell lineages or single cell lineages individually. Diseases characterized by trilineage involvement include Fanconi anemia and dyskeratosis congenita, while Diamond–Blackfan anemia results in single-lineage cytopenia. Recent advances in our understanding of these diseases have arisen from the identification of genetic lesions responsible for such diseases and their pathogenic pathways. These investigations have further clarified both normal and pathological hematopoiesis. In this current review, we describe recent insights into three IBMFS: Fanconi anemia, Diamond–Blackfan anemia, and dyskeratosis congenita.

Fanconi anemia

Fanconi anemia (FA) is a rare autosomal recessive disease characterized by congenital abnormalities, progressive bone marrow failure, and cancer susceptibility. FA, which has an incidence of less than 10 per million live births, is the most frequent inherited cause of aplastic anemia [1]. FA is a genetically heterogeneous disease defined by complementation groups. To date, 15 genes have been identified as playing a causative in FA and these genes, FANCA to FANCP, have been cloned [2]. Children with FA often develop aplastic anemia during the first decade of life, with death often resulting from complications of bone marrow failure, such as severe infection or bleeding. FA

patients also develop clonal chromosomal abnormalities in bone-marrow progenitor cells, such as monosomy 7, which are associated with myelodysplastic syndrome (MDS) and acute myeloblastic leukemia (AML) [3]. The gene FANCD1, which is responsible for complementation group FA-D1, is identical to the hereditary breast cancer susceptibility gene, BRCA2, and has been reported as affected in 3 % of patients with Fanconi anemia. As compared with children from other FA groups, more severe phenotypes are seen in FA-D1 patients, such as co-occurrence of multiple anomalies, development of multiple malignancies with earlier onset, and increased incidence of leukemia and solid tumors [4], including Wilms tumor, neuroblastoma, and brain tumors.

While the treatment of choice for FA patients remains allogeneic stem cell transplantation (SCT) from an HLA-matched sibling or unrelated donor, older patients may develop squamous-cell carcinomas (SCCs) of the head and neck or gynecological system. In particular, some studies have demonstrated there is a high incidence of SCC, such as esophageal cancer, in FA patients who have received SCT. The age-specific hazard of SCC has been shown to be 4.4-fold higher in patients who receive transplants and, in addition, SCCs occurred at significantly younger ages in the transplant group [5]. Thus, further investigations of the complete care of FA patients need to be undertaken.

Complementation groups and genes of FA

Since cells derived from FA patients are hypersensitive to DNA interstrand cross-linking (ICL) agents, such as diepoxybutane (DEB), mitomycin C (MMC), and cisplatin, it is expected that FA genes are involved in ICL repair. In 1993, FANCC was the first FA gene to be cloned by expression cloning [6, 7]. Subsequently, 15 other genes have been cloned (Table 1). At present, the FANC genes, which range from FANC"A" to FANC"P", and the FA pathway have been shown to resolve ICLs encountered during DNA replication. There are three primary groups of FA proteins, which include the FA core complex, the ID (FANCD2/I) complex, and the BRCA complex (Fig. 1). In these groups, there are eight FA proteins (FANCA/B/C/E/F/G/L/M) that form a multi-subunit ubiquitin E3 ligase complex, the FA core complex, which activates the monoubiquitination of the ID complex after genotoxic stress, such as ICL, or during the S phase (Fig. 1) [8, 9]. The monoubiquitinated ID complex forms foci on damaged DNA. FANCM is also a crucial gene, as it is a sensor for detecting stalled DNA replication. The BRCA complex, which is also referred to as homologous recombination (HR), consists of FANCD1, FANCJ, FANCN, and FANCO, and is located downstream of the DI complex on the ICL repair pathway.

Table 1 Genes mutated in patients with Fanconi anemia

	Other names	Chromosomal Locus	Population in FA patients (%)
FANCA		16q24.3	60–70
FANCB		Xp22.2	2
FANCC		9q22.3	14
FANCD1	BRCA2	13q13.1	3
FANCD2		3p25.3	3
FANCE		6p22-p21	3
FANCF		11p15	2
FANCG		9p13	10
FANCI		15q25-q26	1
FANCJ	BACH1/BRIP1	17q22	2
FANCL		2p16.1	0.20
FANCM		14q21.3	0.20
FANCN	PALB2	16p12.3-p12.2	0.70
FANCO	RAD51C	17q22	0.20
FANCP	SLX4	16p13.3	0.20

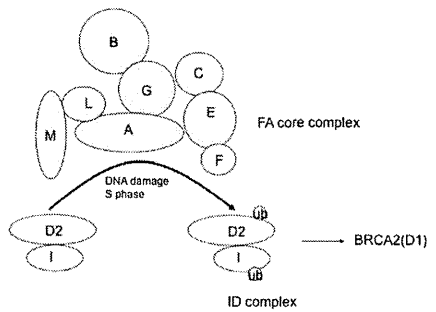


Fig. 1 Simplified scheme of the FA pathway. Depending on the FA core complex, FANCD2 and FANCI are monoubiquitinated after DNA damage or during the S-phase

Prognosis factors in FA patients

Although FANC gene knockout mice models have been established, they differ from the hematological phenotypes of human FA patients [10]. With the exception of the lethal phenotype of the BRCA2/FANCD1 knockout mouse, the hematological parameters of the other FA groups show only a slightly decreased platelet count and a slightly increased erythrocyte mean cell volume in mice at a young age, which did not progress to aplastic anemia or leukemia. However, both male and female mice showed hypogonadism and impaired fertility, which is consistent with the human FA patient phenotypes.

Recent studies have revealed a relationship between the acetaldehyde and FA pathways. Acetaldehyde is an organic chemical compound that is naturally present in coffee, bread, and ripe fruit, which is produced as product of a plant's normal metabolism. It is also produced by the oxidation of ethylene. In the liver, the enzyme alcohol dehydrogenase (ADH) oxidizes ethanol into acetaldehyde, which is then further oxidized into harmless acetic acid by acetaldehyde dehydrogenase (ALDH). These oxidation reactions are coupled with the reduction of NAD^+ to NADH. ALDH2, an isozyme of ALDH, contains the functional polymorphism, *ALDH2* Glu487Lys. An association between this polymorphism and squamous cell carcinomas, such as esophageal cancer in alcoholics, has been reported. A recent study reported that exposure of cells to acetaldehyde results in a concentration-dependent increase in FANCD2 monoubiquitination [11]. Acetaldehyde also stimulates BRCA1 phosphorylation at Ser1524 and increases the level of H2AX, a marker of homologous recombination. Both modifications occur in a dose-dependent manner.

Another report showed that ALDH2 is essential for the development of FANCD 2(−/−) embryos [12]. Nevertheless, mothers with AA enzyme (ALDH 2(+/−)) can support the development of double-mutant (ALDH2 (−/−)FANCD2(−/−)) mice. These embryos are unusually sensitive to ethanol exposure in utero, with ethanol consumption by postnatal double-deficient mice rapidly precipitating bone marrow failure. ALDH2 (−/−)FANCD2(−/−) mice also spontaneously develop acute leukemia.

This previous study also provided the first evidence of the factors responsible for driving the FA hematological phenotype in mice. DNA damage caused by acetaldehyde may contribute critically to the genesis of fetal alcohol syndrome in fetuses, as well as to abnormal development, bone marrow failure, and cancer predisposition in FA patients. This research group also focused on hematopoietic stem cells (HSCs) in another study [13]. They reported finding that some aged ALDH2(−/−)FANCD2(−/−) mutant mice that did not develop leukemia spontaneously developed aplastic anemia, with a concomitant accumulation of damaged DNA within the hematopoietic stem and progenitor cell (HSPC) pool. Only HSPCs and not the more mature blood precursors require Aldh2 for protection against acetaldehyde toxicity. There is more than a 600-fold reduction in the HSC pool of mice deficient in both FA pathway-mediated DNA repair and acetaldehyde detoxification. This study data indicated that the emergence of bone marrow failure in FA was probably due to aldehyde-mediated genotoxicity restricted to the HSPC pool.

All of the ALDH data suggest that ALDH2 polymorphism is critical to the prognosis of FA patients.

Intercrosslink repair

DNA ICLs are toxic to dividing cells, as they induce mutations, chromosomal rearrangements, and cell death. In order to survive, organisms have developed strategies for dealing with DNA damage. As such, specialized repair pathways have evolved for specific kinds of DNA damage, including double-strand break (DSB) and ICL. Inducers of ICLs are important drugs in cancer treatment and include the well-known chemotherapeutic agents mitomycin C, cisplatin, cyclophosphamide, and their respective derivatives. While cells derived from most individuals with FA are hypersensitive to ICLs, they are generally not hypersensitive to inducers of DSBs such as ionizing radiation, indicating that the ICL repair pathway is distinct from that of DSB.

Homologous recombination is a DNA repair pathway that utilizes strand exchange in a gene conversion reaction involving a single-strand and a DNA duplex. In mammalian cells, this is a major repair pathway for DNA damage such as DSBs. The strand exchange protein RAD51 and the products of the hereditary breast cancer susceptibility genes BRCA1 and BRCA2 [14, 15] are critical proteins in HR in mammalian cells.

In 2005, a cellular study in humans showed that mutation of either the FA core complex members or the FANCD2 monoubiquitination site resulted in HR defects [16]. These defects, however, are mild compared with those resulting from a BRCA2 deficiency. HR measurements in these previous studies were performed with the widely used DR-GFP reporter system, in which a DSB formed by I-SceI endonuclease results in green fluorescent protein-expressing (GFP+) cells repaired by HR (Fig. 2). Further studies have reported on the mechanisms of ICL repair, particularly in terms of the replication-coupled manner. A 2008 study using a cell-free system based on *Xenopus* egg extracts found that ICL is repaired in a replication-dependent manner [17]. Another study in the *Xenopus* egg showed that ubiquitinated FANCI-FANCD2 is essential for replication-dependent ICL repair and that it is able to control the incision step [18]. Development and use of a TR-GFP assay, a modified version of the DR-GFP HR assay system, demonstrated that ICL repair in mammalian cells is dependent on DNA replication. The TR-GFP assay uses a DNA template with a site-specific ICL at sequences that are complemented to triplex-forming oligonucleotide conjugated with psoralen (pso-TFO) [19]. The construct also contains an origin of replication from the Epstein–Barr virus (EBV), enabling replication in human cells. Their results showed that ICL-induced HR was substantially compromised in the absence of FA proteins, suggesting that the FA pathway is specifically involved in replication-coupled HR repair. Use of direct assays for ICL-induced HR in vivo, along with studies that

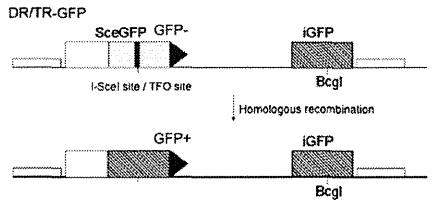


Fig. 2 DR-GFP assay. In the *DR-GFP* substrate, an I-SceI site is inserted into the GFP gene on sceGFP. GFP is inactivated by the stop-codon in the I-SceI site. To restore functionality of the *GFP* gene, the *iGFP* gene has .8 kb of sequence homology to direct the repair of an I-*SceI*-cleaved *SceGFP* gene

have demonstrated the involvement of the FA pathway overall, may facilitate delineation of the mechanisms and factors involved in this process.

Diamond–Blackfan anemia

Diamond–Blackfan anemia (DBA) is a rare congenital bone marrow failure syndrome characterized by severe normochromic macrocytic anemia and reticulocytopenia, with selective hypoplasia of erythroid precursors in the bone marrow. Up to 50 % of affected individuals have physical abnormalities including short stature, craniofacial dysmorphism, heart defects, and anomalies of the thumbs and genitourinary tract [20]. Increased risk of malignant disease, such as acute myeloid leukemia and osteogenic sarcoma, has also been reported to occur in this syndrome [21]. The incidence of DBA has been estimated to be 5–7 per million live births in Europe and North America. In a national study conducted in Japan between 2006 and 2010, 65 new DBA patients were registered. During this study period, the mean number of live births per year in Japan was reported to be 1.08 million, putting the incidence of DBA at 12 per each million live births, as most of these patients were diagnosed as DBA during infancy. The majority of these patients are sporadic, with the percentages of patients with autosomal dominant inheritance reported to be less than 10 %. Corticosteroids are recommended as a first line therapy, as these have been reported to improve erythropoiesis in approximately 80 % of DBA patients. In patients refractory to corticosteroids or who develop other forms of cytopenia, HSC transplantation has been suggested as a viable alternative [22].

Molecular pathogenesis

The first DBA gene (RPS19) was identified in 1999 and was found in approximately 25 % of the probands in western countries [23]. Since then, a total of nine genes encoding large (RPL) or small (RPS) ribosomal subunit proteins were found to be mutated in DBA patients, including RPL5 (6.6 %), RPL11 (4.8 %), RPL35A (3 %), RPS24 (2 %), RPS17 (1 %), RPS7 (1 %), RPS10 (6.4 %), and RPS26 (2.6 %) [24] (Table 2). Collectively, mutations in at least one of these nine genes have been detected in approximately 50–60 % of DBA patients. Of 68 Japanese been examined, mutations in RPS 19, RPL5, RPL11, RPS17, RPS26 were identified in 10 (14.7 %), six (8.8 %), three (4.4 %), one (1.5 %), one (1.5 %), and one (1.5 %), respectively. These mutations have subsequently been determined to occur in 32.4 % of Japanese patients [25, 26]. A low incidence of mutations in the RPS19 gene may account for the overall lower incidence of total mutations in the Japanese population.

As conventional gene sequencing cannot identify large gene deletions, there have been only a few reports of patients with allelic losses in the RPS19 and RPL35A genes. Kuramitsu et al. [26] investigated large deletions of the RP genes using gene copy number variation analysis based on a quantitative-PCR and a single-nucleotide polymorphism (SNP) array. This study used sequencing to screen for large gene deletion in 27 patients without gene mutations. The PCR-based gene copy number assay identified a large deletion in seven (25.9 %) of 27 patients. Of these, three patients had RPS17, two had RPS19, one had RPL5, while one had RPL35A deletions. The SNP array confirmed six of the seven large deletions. Based on these new methods, the frequency of RP gene abnormalities in the DBA patients increased to 42.6 %. All patients with large deletions in DBA genes exhibited malformation with growth retardation. However, half of the patients with a mutation due to sequencing had growth retardation, while all seven patients with a large deletion exhibited growth retardation. While four of seven patients responded to corticosteroids, there were no phenotypic

Table 2 Ribosomal protein gene mutations and deletions in 68 Japanese patients with Diamond–Blackfan anemia

Gene	Mutation	Deletion	Total (%)
RPS 19	10	2	12 (17.6 %)
RPL 5	6	1	7 (10.3 %)
RPS 17	3	0	3 (4.4 %)
RPL 11	3	1	4 (5.9 %)
RPS 10	1	0	1 (1.5 %)
RPS 26	1	0	1 (1.5 %)
RPS 35A	0	1	1 (1.5 %)
RPS 24	0	0	0
RPS 14	0	0	0
Total	22	7	29 (42.6 %)

differences noted between patients with and without large deletions, including response rate to corticosteroids and other malformations.

Farrar et al. [27] also identified RP gene deletions in nine (17 %) of 51 patients without any identifiable mutation by SNP array. Of these nine patients, three had RPS17, two had RPS26, two had RPS19, and two had RPL35A deletions. Clinically, five of the nine patients responded to corticosteroids. Two exhibited short stature. These two studies suggested that genomic deletions may be detected in 4–10 % of DBA patients, which is more common than has been previously suspected. Thus, in addition to conventional gene sequencing, molecular studies of suspected DBA cases should also include either a SNP array or PCR-based gene copy number assay.

Despite extensive sequencing of all the RP genes, at present mutations have only been found in approximately half of DBA patients examined, which raises the question whether other genes are responsible for DBA. Recent advance in genomic sequencing have made it possible to search for new candidate genes. Sankaran et al. [28] performed exome sequencing on two siblings without RP gene mutations. Both affected siblings satisfied the diagnostic criteria for DBA and both parents had normal blood values, suggesting X-linked or autosomal recessive inheritance. During sequencing, at least 10-fold coverage was obtained in more than 93 % of the target bases. After filtering, a total of 74 variants were identified as being shared by the three affected siblings. Of these 74 mutations, 31 were found in two affected siblings but not in the unaffected sibling. No variants were identified that would fit an autosomal recessive model of inheritance. Only the GATA1 gene showed appropriate segregation for an X-linked disease with full penetrance. The mutation in the GATA1 gene is a G-C transversion at position 48,649,736 on the X chromosome and results in a substitution of leucine for valine at amino acid 74 of the GATA1 protein. This mutation impaired production of the full-length form of the exon 2 protein. After screening 62 additional male DBA patients without known mutations for the GATA1 mutation, the study also identified one patient with a mutation in GATA1 at the exon 2-intron 2 junction. It was predicted that this would result in impaired splicing and a frameshift of the full-length GATA1 open reading frame. Overall, this study has opened new avenues for studying the molecular pathogenesis of DBA.

Role of p53 in the pathophysiology of DBA

Although current evidence suggests that impaired ribosomal biogenesis should affect all blood cell lineages, one question remains as to why it affects only the erythroid progenitors. Several animal models have demonstrated the role of p53 in the pathophysiology of DBA. The RPS19-deficient zebrafish model has been shown to have many features of DBA and is accompanied by the up-regulation of the p53 family [29]. Suppression of p53 in the RPS19-deficient zebrafish alleviated the phenotype and improved survival. RPS19 knockdown mouse fetal liver cells, which were created by retrovirus-infected siRNA, showed reduced proliferation but normal differentiation of erythroid cells and an increased level of p53 and p21 [30]. Dutt et al. [31] have examined the accumulation and activity of p53 in different hematopoietic lineages after a partial knockdown of the RPS19 gene in primary human bone marrow-derived CD34 cells. Their study showed that p53 accumulates selectively in erythroid progenitors, resulting in lineage-specific p53 target gene expression, cell cycle arrest, and apoptosis. While pifithrin-α has been shown to inhibit the activity of p53, nutlin-3 activates p53 through the inhibition of HDM2. In addition, nutlin-3 selectivity impairs erythropoiesis, whereas inhibition of p53 by pifithrin-α rescues the erythroid defect. To directly examine whether p53 accumulation is operative in patients with DBA, bone marrow biopsies from eight patients with DBA were stained with anti-human p53 antibody and shown to have strong nuclear staining in two patients and weak nuclear staining in six patients. The erythroid lineage has a low threshold for the induction of p53, which accounts for the selective impaired erythropoiesis in patients with DBA.

Alternative therapies for DBA

Although approximately 80 % of DBA patients initially respond to corticosteroid, half of the responders are steroid-dependent. Only 20 % of these patients achieve remission. Although historically many alternative drugs have been tried, there has been no agreement on a second-line therapy. L-leucine is an essential amino acid and is known to be an activator of mRNA and stimulate protein synthesis through the mammalian target of rapamycin (mTOR) pathway. L-leucine treatment of the RPS19-deficient zebrafish model results in a striking improvement of anemia and developmental defects. These findings were reproduced in primary human CD34 cells after knockdown of the RPS19 gene [32]. Therapeutic effect of L-leucine has also been confirmed in the mouse model for RPS19-deficient DBA and shown to be associated with reduced p53 activity in hematopoietic progenitors [33]. Recently, leucine has been used on an investigational basis in one patient with DBA and is reported to have achieved a remission [34].

These findings support commencement of a clinical trial with L-leucine as an alternative therapy for DBA.

Dyskeratosis congenita

Clinical features of patients with dyskeratosis congenita

Dyskeratosis congenita (DC) is a rare inherited disease characterized by the classical mucocutaneous triad of abnormal skin pigmentation, nail dystrophy, and mucosal leucoplakia in approximately 80–90 % of patients [34]. Patients with DC are unable to maintain the telomere complex that protects the chromosome ends and consequently have very short telomeres [35]. Shortened telomeres can cause a wide variety of clinical features across a phenotypic spectrum consisting not only of mucocutaneous abnormalities but also multisystem symptoms including bone marrow failure, pulmonary fibrosis, hepatic fibrosis, and predisposition to malignancy [36, 37]. Indeed, non-mucocutaneous features, such as bone marrow failure and pulmonary fibrosis, occasionally precede mucocutaneous abnormalities, making it difficult to diagnose patients with DC based on clinical features alone. The incidence of DC is estimated to be one per million live births.

The diagnostic criteria for DC proposed by Vulliamy [38] include one or more of the three classic mucocutaneous features combined with hypoplastic bone marrow and at least two other somatic features known to occur in DC. The primary causes of mortality in patients with DC are bone marrow failure syndrome (60–70 %), pulmonary complications (10–15 %), and malignancy (particularly MDS and AML) (10 %) [36, 37].

Genetic background of DC

DC is a genetically heterogeneous disorder, showing autosomal recessive, autosomal dominant, and X-linked inheritance. The *DKC1* gene on chromosome (chr) Xq28, which encodes dyskerin, was the first gene identified in the X-linked DC patients [39]. Dyskerin has a close association with the RNA component of telomerase (*TERC*), and mutations in dyskerin cause a reduction in accumulation of *TERC* and reduced telomere length [35]. In addition to its role in the biogenesis of telomerase RNA dyskerin is involved in ribosomal RNA biogenesis. Dyskerin catalyzes uridine to pseudouridine, which is a critical step for ribosomal RNA maturation and function. These findings imply that both telomere and ribosomal defects may occur in patients with *DKC1* mutations. Subsequently, heterozygous *TERC* mutations have also been found in autosomal dominant DC patients [40]. Genetic screening has identified mutations of other components of the telomerase complex, including *TERT* (chr 5p15) [41, 42], *NOP10* (chr 15q14-q15) [43], and *NHP2* (chr 5q35) [44] in patients with rare autosomal recessive DC. Mutations of *TERT* have also been reported in the autosomal dominant family [45]. Moreover, heterozygous mutations of *TINF2* (chr 14q12) that encode TIN2, which is the main component of shelterin and which protects telomeres, have been identified in <11 % of DC patients [46, 47].

More recently, mutations of *TCAB1* (chr 17p13) were identified in patients with DC as autosomal recessive forms [48]. Venteicher et al. [49] found that TCAB1 associates with TERT, dyskerin and TERC, and small Cajal body RNAs (scaRNAs) that are involved in modifying splicing RNAs to control telomerase trafficking. TCAB1 defects prevented TERC from associating with the Cajal bodies, which disrupted the telomerase–telomere association. A recent case report described biallelic mutations of the *CTC1* gene (chr 17p13) in a patient with DC [50]. This gene was originally described as causative gene of the Coats plus syndrome, which is a form of cerebroretinal microangiopathy with calcifications and cysts (CRMCC). The mutation frequencies of these new genes for DC remain unknown.

At present, eight of the mutated genes in DC have been shown to be associated with the telomerase holoenzyme (TERT, *TERC, DKC1, NOP10, NHP2, TCAB1,* and *CTC1)* or the shelterin complex (TINF2), accounting for approximately 50 % of DC patients. Mutations in telomerase and telomere components have also been identified in patients with aplastic anemia, pulmonary fibrosis, and liver diseases that did not have any mucocutaneous manifestations [45, 46, 51–59]. These findings suggest that defective telomere maintenance causes not only classical DC, but also a broad spectrum of diseases previously thought to be idiopathic and thus this has led to a new concept of diseases termed ''syndromes of telomere shortening''.

Cryptic DC patients in aplastic anemia

Patients with DC have been shown to have disease diversity in terms of age at onset, symptoms, and severity. This diversity occurs even among the patients with the same gene mutation. Bone marrow failure sometimes precedes mucocutaneous manifestations in patients with DC, and a substantial proportion of patients with aplastic anemia have shorter telomeres compared with normal individuals [60, 61]. These observations have prompted screening for gene mutations responsible for telomere maintenance in patients with aplastic anemia and other bone marrow failure syndromes. This screening identified mutations in *TERC* and *TERT* in 3 % of the aplastic anemia patients [54, 55]. Our research group conducted a study in Japanese children with aplastic anemia and identified two of 96 as having the *TERT* mutations, although none of the patients had a *TERC* mutation [53]. Patients with *TERC* or *TERT* mutations have been shown to have very short telomeres in their blood cells. Recently, Du et al. [52] found that 6 (5.5 %) of 109

pediatric patients with severe aplastic anemia had mutations of *TINF2*. In an unpublished study, our research group screened for mutations of *TINF2* and found that of the 96 pediatric patients with aplastic anemia that were examined, none exhibited any mutations of this gene.

Three methods are commonly used for measuring telomere length, including Southern blot, real-time polymerase chain reaction, and flow cytometry and fluorescence in situ hybridization (flow-FISH). Of these, the flow-FISH has been shown to be the most appropriate when undertaking "prospective" screening [62, 63]. As shown in Fig. 3, patients with DC and aplastic anemia with the *TERT* mutation were all found to have very short telomeres as compared with the idiopathic aplastic anemia patients and normal individuals. As a small subset of patients with apparently idiopathic aplastic anemia have been shown to carry telomere gene mutations, identification of such patients is critical for informing treatment decisions. Aplastic anemia patients should be routinely screened for telomere gene mutations prior to starting any treatment. However, because screening of gene mutations can be both laborious and time consuming, we have adopted the screening of telomere length in blood cells rather than screening of gene mutations.

It should be noted that short telomeres are not specific for patients with DC, as they are also seen in patients with other bone marrow failure syndromes. Although short telomeres have also been found in patients with other congenital bone marrow failure syndromes, such as Shwachman–Diamond syndrome and Fanconi anemia, telomere lengths in patients with DC have been demonstrated to be

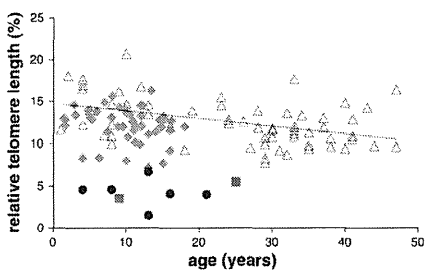


Fig. 3 Relative telomere length in peripheral blood lymphocytes from patients with dyskeratosis congenita (*filled circles*), patients with aplastic anemia harboring TERT mutations (*filled squares*), patients with idiopathic aplastic anemia (*filled argyles*) and normal individuals (*open triangles*). Telomere lengths were measured by flow cytometry-fluorescent in situ hybridization (flow-FISH). Relative telomere length was calculated as the ratio between the telomere signal of each sample and the telomere signal of the control cell line (cell line 1301). These data were provided by the Department of Pediatrics, Nagoya University Graduate School of Medicine

shorter than those in all other bone marrow failure syndromes. In fact, telomere length in most patients with DC is below the first percentile of telomere length found in healthy controls [64].

Family members of patients with DC should receive genetic counseling to rule out if they are silent carriers. In particular, genetic counseling is necessary during the proband search for a donor for HSC transplantation. Studies on telomere length analyses in families with DC have shown that mutated carriers with clinical signs of bone marrow failure have short telomeres. Even so, telomere length cannot predict the presence or absence of a mutation in family members with bone marrow failure. In addition, there have been rare cases that show normal telomere length, even though the subject harbors the same mutation as the proband. This suggests that mutation alone does not sufficiently explain the reduction of telomere length [51].

Clinical management for DC

Bone marrow failure and immune deficiency are the most common causes of death in up to 60–70 % of patients with DC. Androgen (e.g. oxymetholone) has been used to improve cytopenia in patients with DC since the 1960s. However, the mechanism of action of androgen has not been well understood until recently. Calado et al. [65] showed that in vitro exposure of normal peripheral blood cells to androgen produced higher TERT mRNA levels. When these patients were treated with cells from patients who had a heterozygous mutation of the telomerase, it was possible to restore their low baseline telomerase activity to normal levels. Thus, as telomere shortening is closely associated with malignant disease, androgen therapy might be able to prevent or postpone the development of various types of cancers. Erythropoietin and/or G-CSF combined with androgen has occasionally provided transient hematopoietic recovery to poor responders to androgen alone [66]. However, this combination should be used with caution, as severe splenic peliosis and fatal rupture have been reported in two patients with DC who received simultaneous administration of androgen and G-CSF [67].

Allogeneic HSC transplantation is the only curative treatment for bone marrow failure in patients with DC. However, the outcome in previous reports has been disappointing due to unacceptable transplant-related toxicities, including severe pulmonary/liver complications, especially in transplants from an alternative donor [68, 69]. To avoid these complications, non-myeloablative conditioning regimens have been recently used in several cases. Dietz et al. [70] reported encouraging results of six patients with DC who received a fludarabine-based non-myeloablative regimen. Of the four surviving patients, three were recipients of unrelated grafts. Non-myeloablative

transplants are expected to provide improvement in the short-term survival. At our institute, three patients with DC underwent allogeneic bone marrow transplantation following non-myeloablative conditioning from 2003 to 2009. Successful engraftment was achieved in all patients with only a few regimen-related toxicities, and at the present time all continue to survive without any symptoms [71]. However, due the late effects of conditioning agents and allogeneic immune responses within the recipient's organs, such as the lung and liver, longer-term follow-ups are necessary to definitively clarify the present results.

Conclusion

Although recent studies have identified many causative genes, mutations of these genes have only been found in half of the patients with DBA or DKC. Next-generation sequencing (or massive parallel sequencing) technologies have led to a tremendous revolution in genomics, with their effects currently becoming increasingly widespread. This new strategy may soon be able to reveal the remaining unknown causative genes in IBMFS.

Recently, Agarwal et al. [72] established induced pluripotent stem cells (iPSCs) derived from a patient with DC and showed that the reprogrammed DC cells overcame a critical limitation in TERC levels to restore the telomere maintenance and self-renewal. These findings indicate that drugs or gene therapy that upregulate TERC activity may show therapeutic potential in patients with DC. These same strategies may also be applicable for other IBMFS.

The only long-term curative treatment for bone marrow failure in patients with IBMFS is allogeneic HSC transplantation, although this procedure has a risk of severe adverse effects. Multicenter prospective studies are needed to establish appropriate conditioning regimens aimed at reducing transplant-related mortality. Future studies must aim to improve short-term outcomes, such as hematological recovery, and to decrease the incidence of late adverse effects.

References

1. Joenje H, Patel KJ. The emerging genetic and molecular basis of Fanconi anaemia. Nat Rev Genet. 2001;2:446–57.
2. Kim H, D'Andrea AD. Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway. Genes Dev. 2012;26:1393–408.
3. D'Andrea AD, Grompe M. The Fanconi anaemia/BRCA pathway. Nat Rev Cancer. 2003;3:23–34.
4. Alter BP, Rosenberg PS, Brody LC. Clinical and molecular features associated with biallelic mutations in FANCD1/BRCA2. J Med Genet. 2007;44:1–9.
5. Rosenberg PS, Socie G, Alter BP, Gluckman E. Risk of head and neck squamous cell cancer and death in patients with Fanconi anemia who did and did not receive transplants. Blood. 2005;105:67–73.
6. Whitney MA, Saito H, Jakobs PM, Gibson RA, Moses RE, Grompe M. A common mutation in the FACC gene causes Fanconi anaemia in Ashkenazi Jews. Nat Genet. 1993;4:202–5.
7. Gibson RA, Hajianpour A, Murer-Orlando M, Buchwald M, Mathew CG. A nonsense mutation and exon skipping in the Fanconi anaemia group C gene. Hum Mol Genet. 1993;2:797–9.
8. Garcia-Higuera I, Taniguchi T, Ganesan S, Meyn MS, Timmers C, Hejna J, et al. Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. Mol Cell. 2001;7:249–62.
9. Smogorzewska A, Matsuoka S, Vinciguerra P, McDonald ER 3rd, Hurov KE, Luo J, et al. Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. Cell. 2007;129:289–301.
10. Carreau M. Not-so-novel phenotypes in the Fanconi anemia group D2 mouse model. Blood. 2004;103:2430.
11. Marietta C, Thompson LH, Lamerdin JE, Brooks PJ. Acetaldehyde stimulates FANCD2 monoubiquitination, H2AX phosphorylation, and BRCA1 phosphorylation in human cells in vitro: implications for alcohol-related carcinogenesis. Mutat Res. 2009;664:77–83.
12. Langevin F, Crossan GP, Rosado IV, Arends MJ, Patel KJ. Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. Nature. 2011;475:53–8.
13. Garaycoechea JI, Crossan GP, Langevin F, Daly M, Arends MJ, Patel KJ. Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function. Nature. 2012;489:571–5.
14. Moynahan ME, Chiu JW, Koller BH, Jasin M. Brca1 controls homology-directed DNA repair. Mol Cell. 1999;4:511–8.
15. Moynahan ME, Pierce AJ, Jasin M. BRCA2 is required for homology-directed repair of chromosomal breaks. Mol Cell. 2001;7:263–72.
16. Nakanishi K, Yang YG, Pierce AJ, Taniguchi T, Digweed M, D'Andrea AD, et al. Human Fanconi anemia monoubiquitination pathway promotes homologous DNA repair. Proc Natl Acad Sci USA. 2005;102:1110–5.
17. Raschle M, Knipscheer P, Enoiu M, Angelov T, Sun J, Griffith JD, et al. Mechanism of replication-coupled DNA interstrand crosslink repair. Cell. 2008;134:969–80.
18. Knipscheer P, Raschle M, Smogorzewska A, Enoiu M, Ho TV, Scharer OD, et al. The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. Science. 2009;2009(326):1698–701.
19. Nakanishi K, Cavallo F, Perrouault L, Giovannangeli C, Moynahan ME, Barchi M, et al. Homology-directed Fanconi anemia pathway cross-link repair is dependent on DNA replication. Nat Struct Mol Biol. 2011;18:500–3.
20. Alter BP, Young NS. The bone marrow failure syndromes. In: Nathan DG, Orkin HS, editors. Hematology of infancy and childhood, vol 1. Philadelphia: Saunders; 1998. p. 237–335.
21. Vlachos A, Rosenberg PS, Atsidaftos E, Alter BP, Lipton JM. Incidence of neoplasia in Diamond Blackfan anemia: a report from the Diamond Blackfan Anemia Registry. Blood. 2012;119:3815–9.
22. Mugishima H, Ohga S, Ohara A, Kojima S, Fujisawa K, For the Aplastic Anemia Committee of the Japanese Society of Pediatric Hematology. Hematopoietic stem cell transplantation for Diamond–Blackfan anemia: a report from the Aplastic Anemia Committee of the Japanese Society of Pediatric Hematology. Pediatr Transpl. 2007;11:601–7.
23. Draptchinskaia N, Gustavsson P, Andersson B, Pettersson M, Willig TN, Dianzani I, et al. The gene encoding ribosomal

protein S19 is mutated in Diamond–Blackfan anaemia. Nat Genet. 1999;21:169–75.
24. Boria I, Garelli E, Gazda HT, Aspesi A, Quarello P, Pavesi E, et al. The ribosomal basis of Diamond–Blackfan anemia: mutation and database update. Hum Mutat. 2010;31:1269–79.
25. Konno Y, Toki T, Tandai S, Xu G, Wang R, Terui K, et al. Mutations in the ribosomal protein genes in Japanese patients with Diamond–Blackfan anemia. Haematologica. 2010;95: 1293–9.
26. Kuramitsu M, Sato-Otsubo A, Morio T, Takagi M, Toki T, Terui K, et al. Extensive gene deletions in Japanese patients with Diamond–Blackfan anemia. Blood. 2012;119:2376–84.
27. Farrar JE, Vlachos A, Atsidaftos E, Carlson-Donohoe H, Markello TC, Arceci RJ, et al. Ribosomal protein gene deletions in Diamond–Blackfan anemia. Blood. 2011;118:6943–51.
28. Sankaran VG, Ghazvinian R, Do R, Thiru P, Vergilio JA, Beggs AH, et al. Exome sequencing identifies GATA1 mutations resulting in Diamond–Blackfan anemia. J Clin Invest. 2012;122: 2439–43.
29. Danilova N, Sakamoto KM, Lin S. Ribosomal protein S19 deficiency in zebrafish leads to developmental abnormalities and defective erythropoiesis through activation of p53 protein family. Blood. 2008;112:5228–37.
30. Sieff CA, Yang J, Merida-Long LB, Lodish HF. Pathogenesis of the erythroid failure in Diamond Blackfan anaemia. Br J Haematol. 2010;148:611–22.
31. Dutt S, Narla A, Lin K, Mullally A, Abayasekara N, Megerdichian C, et al. Haploinsufficiency for ribosomal protein genes causes selective activation of p53 in human erythroid progenitor cells. Blood. 2011;117:2567–76.
32. Payne EM, Virgilio M, Narla A, Sun H, Levine M, Paw BH, et al. L-leucine improves the anemia and developmental defects associated with Diamond–Blackfan anemia and del(5q) MDS by activating the mTOR pathway. Blood. 2012;120:2214–24.
33. Jaako P, Debnath S, Olsson K, Bryder D, Flygare J, Karlsson S. Dietary L-leucine improves the anemia in a mouse model for Diamond–Blackfan anemia. Blood. 2012;120:2225–8.
34. Kirwan M, Dokal I. Dyskeratosis congenita, stem cells and telomeres. Biochim Biophys Acta. 2009;1792:371–9.
35. Mitchell JR, Wood E, Collins K. A telomerase component is defective in the human disease dyskeratosis congenita. Nature. 1999;402:551–5.
36. Dokal I. Dyskeratosis congenita in all its forms. Br J Haematol. 2000;110:768–79.
37. Walne AJ, Dokal I. Advances in the understanding of dyskeratosis congenita. Br J Haematol. 2009;145:164–72.
38. Vulliamy TJ, Marrone A, Knight SW, Walne A, Mason PJ, Dokal I. Mutations in dyskeratosis congenita: their impact on telomere length and the diversity of clinical presentation. Blood. 2006;107: 2680–5.
39. Heiss NS, Knight SW, Vulliamy TJ, Klauck SM, Wiemann S, Mason PJ, Poustka A, Dokal I. X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. Nat Genet. 1998;19:32–8.
40. Vulliamy T, Marrone A, Goldman F, Dearlove A, Bessler M, Mason PJ, et al. The RNA component of telomerase is mutated in autosomal dominant dyskeratosis congenita. Nature. 2001;413: 432–5.
41. Armanios M, Chen JL, Chang YP, Brodsky RA, Hawkins A, Griffin CA, et al. Haploinsufficiency of telomerase reverse transcriptase leads to anticipation in autosomal dominant dyskeratosis congenita. Proc Natl Acad Sci USA. 2005;102:15960–4.
42. Marrone A, Walne A, Tamary H, Masunari Y, Kirwan M, Beswick R, et al. Telomerase reverse-transcriptase homozygous mutations in autosomal recessive dyskeratosis congenita and Hoyeraal–Hreidarsson syndrome. Blood. 2007;110:4198–205.
43. Walne AJ, Vulliamy T, Marrone A, Beswick R, Kirwan M, Masunari Y, et al. Genetic heterogeneity in autosomal recessive dyskeratosis congenita with one subtype due to mutations in the telomerase-associated protein NOP10. Hum Mol Genet. 2007;16: 1619–29.
44. Vulliamy T, Beswick R, Kirwan M, Marrone A, Digweed M, Walne A, et al. Mutations in the telomerase component NHP2 cause the premature ageing syndrome dyskeratosis congenita. Proc Natl Acad Sci USA. 2008;105:8073–8.
45. Vulliamy TJ, Walne A, Baskaradas A, Mason PJ, Marrone A, Dokal I. Mutations in the reverse transcriptase component of telomerase (TERT) in patients with bone marrow failure. Blood Cells Mol Dis. 2005;34:257–63.
46. Walne AJ, Dokal I. Dyskeratosis congenita: a historical perspective. Mech Ageing Dev. 2008;129:48–59.
47. Savage SA, Giri N, Baerlocher GM, Orr N, Lansdorp PM, Alter BP. TINF2, a component of the shelterin telomere protection complex, is mutated in dyskeratosis congenita. Am J Hum Genet. 2008;82:501–9.
48. Zhong F, Savage SA, Shkreli M, et al. Disruption of telomerase trafficking by TCAB1 mutation causes dyskeratosis congenita. Genes Dev. 2011;25:11–6.
49. Venteicher AS, Abreu EB, Meng Z, McCann KE, Terns RM, Veenstra TD, Terns MP, Artandi SE. A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis. Science. 2009;323:644–8.
50. Keller RB, Gagne KE, Usmani GN, Asdourian GK, Williams DA, Hofmann I, Agarwal S. CTC1 Mutations in a patient with dyskeratosis congenita. Pediatr Blood Cancer. 2012;59:311–4.
51. Du HY, Pumbo E, Ivanovich J, An P, Maziarz RT, Reiss UM, et al. TERC and TERT gene mutations in patients with bone marrow failure and the significance of telomere length measurements. Blood. 2009;113:309–16.
52. Du HY, Mason PJ, Bessler M, Wilson DB. TINF2 mutations in children with severe aplastic anemia. Pediatr Blood Cancer. 2009;52:687.
53. Liang J, Yagasaki H, Kamachi Y, Hama A, Matsumoto K, Kato K, et al. Mutations in telomerase catalytic protein in Japanese children with aplastic anemia. Haematologica. 2006;91:656–8.
54. Yamaguchi H, Calado RT, Ly H, Kajigaya S, Baerlocher GM, Chanock SJ, et al. Mutations in TERT, the gene for telomerase reverse transcriptase, in aplastic anemia. N Engl J Med. 2005; 352:1413–24.
55. Yamaguchi H, Baerlocher GM, Lansdorp PM, Chanock SJ, Nunez O, Sloand E, et al. Mutations of the human telomerase RNA gene (TERC) in aplastic anemia and myelodysplastic syndrome. Blood. 2003;102:916–8.
56. Vulliamy T, Marrone A, Dokal I, Mason PJ. Association between aplastic anaemia and mutations in telomerase RNA. Lancet. 2002;359:2168–70.
57. Tsakiri KD, Cronkhite JT, Kuan PJ, Xing C, Raghu G, Weissler JC, et al. Adult-onset pulmonary fibrosis caused by mutations in telomerase. Proc Natl Acad Sci USA. 2007;104:7552–7.
58. Armanios MY, Chen JJ, Cogan JD, Alder JK, Ingersoll RG, Markin C, et al. Telomerase mutations in families with idiopathic pulmonary fibrosis. N Engl J Med. 2007;356:1317–26.
59. Calado RT, Regal JA, Kleiner DE, Schrump DS, Peterson NR, Pons V, et al. A spectrum of severe familial liver disorders associate with telomerase mutations. PLoS ONE. 2009;4:e 7926.
60. Ball SE, Gibson FM, Rizzo S, Tooze JA, Marsh JC, Gordon-Smith EC. Progressive telomere shortening in aplastic anemia. Blood. 1998;91:3582–92.
61. Lee JJ, Kook H, Chung IJ, Na JA, Park MR, Hwang TJ, et al. Telomere length changes in patients with aplastic anaemia. Br J Haematol. 2001;112:1025–30.

62. Baerlocher GM, Vulto I, de Jong G, Lansdorp PM. Flow cytometry and FISH to measure the average length of telomeres (flow FISH). Nat Protoc. 2006;1:2365–76.
63. Canela A, Klatt P, Blasco MA. Telomere length analysis. Methods Mol Biol. 2007;371:45–72.
64. Alter BP, Baerlocher GM, Savage SA, Chanock SJ, Weksler BB, Willner JP, et al. Very short telomere length by flow fluorescence in situ hybridization identifies patients with dyskeratosis congenita. Blood. 2007;110:1439–47.
65. Calado RT, Yewdell WT, Wilkerson KL, Regal JA, Kajigaya S, Stratakis CA, et al. Sex hormones, acting on the TERT gene, increase telomerase activity in human primary hematopoietic cells. Blood. 2009;114:2236–43.
66. Alter BP, Gardner FH, Hall RE. Treatment of dyskeratosis congenita with granulocyte colony-stimulating factor and erythropoietin. Br J Haematol. 1997;97:309–11.
67. Giri N, Pitel PA, Green D, Alter BP. Splenic peliosis and rupture in patients with dyskeratosis congenita on androgens and granulocyte colony-stimulating factor. Br J Haematol. 2007;138:815–7.
68. Alter BP, Giri N, Savage SA, Rosenberg PS. Cancer in dyskeratosis congenita. Blood. 2009;113:6549–57.
69. de la Fuente J, Dokal I. Dyskeratosis congenita: advances in the understanding of the telomerase defect and the role of stem cell transplantation. Pediatr Transpl. 2007;11:584–94.
70. Dietz AC, Orchard PJ, Baker KS, Giller RH, Savage SA, Alter BP, et al. Disease-specific hematopoietic cell transplantation: nonmyeloablative conditioning regimen for dyskeratosis congenita. Bone Marrow Transpl. 2011;46:98–104.
71. Nishio N, Takahashi Y, Ohashi H, Doisaki S, Muramatsu H, Hama A, Shimada A, Yagasaki H, Kojima S. Reduced-intensity conditioning for alternative donor hematopoietic stem cell transplantation in patients with dyskeratosis congenita. Pediatr Transpl. 2011;15:161–6.
72. Agarwal S, Loh YH, McLoughlin EM, Huang J, Park IH, Miller JD, et al. Telomere elongation in induced pluripotent stem cells from dyskeratosis congenita patients. Nature. 2010;464:292–6.

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 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*

[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. (ogawa-hky@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

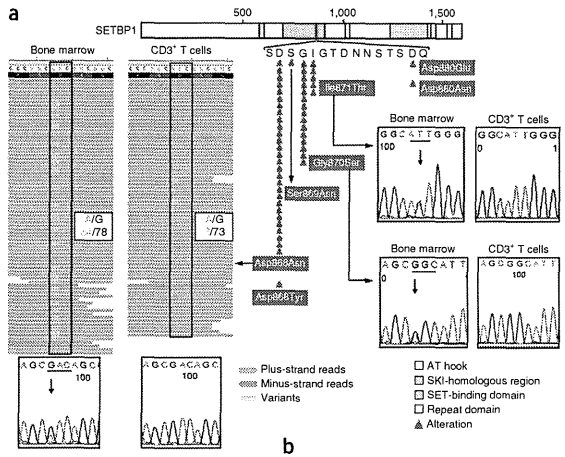


Figure 1 Somatic *SETBP1* mutations as detected by next-generation whole-exome sequencing and Sanger sequencing. (**a**) Distribution of *SETBP1* mutations detected in 52 of 727 myeloid neoplasms, all of which were located within the portion of the gene encoding the SKI-homologous domain. Bottom and 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.

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

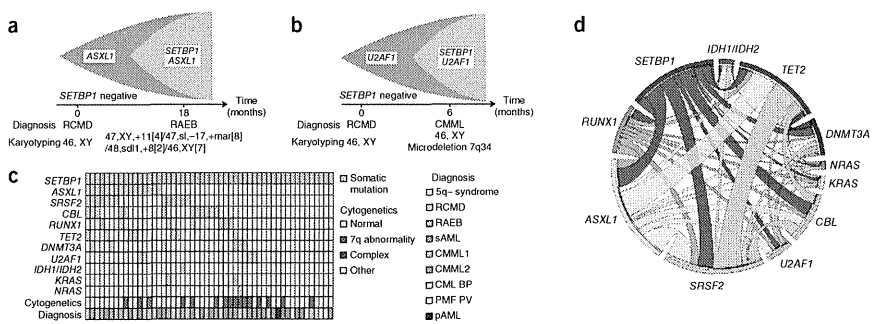
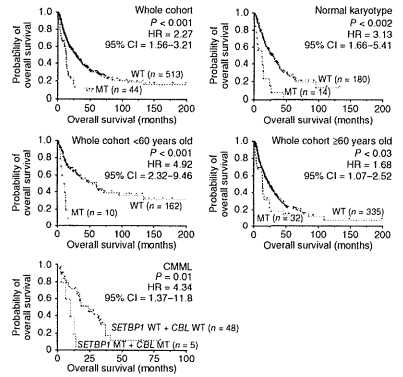


Figure 2 The relationship of *SETBP1* mutations with other common mutations. (**a**,**b**) Clonological profiles of gene mutations in two representative cases with MDS that transformed to RAEB (**a**) and CMML (**b**). Initially, hypocellular MDS (RCMD) was diagnosed on the basis of hypocellular bone 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.

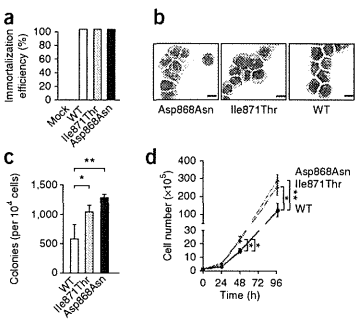


Figure 4 Immortalization of mouse myeloid progenitors by *SETBP1* mutations. (**a**) Efficiencies of empty pMYs retroviral vector (mock) or of pMYs constructs expressing wild-type Setbp1 (WT) or Setbp1 mutants (Asp868Asn and Ile871Thr) in immortalizing mouse myeloid progenitor cells in three independent 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.

URLs. The February 2009 human reference sequence (GRCh37) produced by the Genome Reference Consortium was used as the reference genome (UCSC Genome Browser; http://genome.ucsc.edu/cgi-bin/hgGateway). Basewise conservation scores were calculated using PhyloP in the UCSC Genome Browser. Expression array and methylation array data were extracted from Oncomine (https://www.oncomine.org/), BioGPS (http://biogps.org/) and The Cancer Genome Atlas (TCGA; http://cancergenome.nih.gov/) and were analyzed by Matlab software (http://www.mathworks.com/). Somatic mutation data were searched in the Catalogue of Somatic Mutations in Cancer (COSMIC) database on the Welcome Trust Sanger Institute website (http://www.sanger.ac.uk/genetics/CGP/cosmic/). Each potential mutation was compared against databases of known SNPs, including Entrez Gene (http://www.ncbi.nlm.nih.gov/gene) and the Ensembl Genome Browser (http://useast.ensembl.org/index.html). SAMtools (http://samtools.sourceforge.net/) and Integrative Genomics Viewer (http://www.broadinstitute.org/igv/) software were used. The Database of Genomic Variants is a publically available database of copy number variations (http://projects.tcag.ca/variation).

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Whole-exome sequencing results have been deposited in the Sequence Read Archive (SRA; BioProject accession PRJNA203580).

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We thank T. Yamaguchi (The University of Tokyo) for providing CS-Ubc lentivirus vector. This work was supported by US National Institutes of Health (NIH) grants RO1 HL-082983 (J.P.M.), U54 RR019391 (J.P.M.), K24 HL-077522 (J.P.M.) and RO1 CA-143193 (Y.D.), by a grant from the AA & MDS International Foundation, by the Robert Duggan Charitable Fund (J.P.M.), by a Scott Hamilton CARES grant (H. Makishima) and by Grants-in-Aid from the Ministry of Health, Labor and Welfare of Japan and KAKENHI (23249052, 22134006 and 21790907; S.O.), the project for the development of innovative research on cancer therapies (p-direct; S.O.), the Japan Society for the Promotion of Science (JSPS) through the Funding Program for World-Leading Innovative R&D on Science and Technology, initiated by the Council for Science and Technology Policy (CSTP; S.O.), NHRI-EX100-10003NI Taiwan (L.-Y.S.) and Uniformed Services University of the Health Sciences Pediatrics grant KM86GI (Y.D.). The results presented here are partly based on data generated by The Cancer Genome Atlas (TCGA) pilot project established by the National Cancer Institute and the National Human Genome Research Institute. Information about TCGA and the investigators and institutions that constitute the TCGA research network can be found at http://cancergenome.nih.gov/.

AUTHOR CONTRIBUTIONS

H. Makishima and K.Y. designed research, performed research, collected data, performed statistical analysis and wrote the manuscript. Y.O., N.N., K.P.N., B.P., K.O.G., B.A.V., A.J., I.G.-S., Y. Shiraishi, Y.N., M.S., M.T., K.C., H.T., H. Muramatsu, H.S., S.M. and L.-Y.S. performed research and analyzed data. K.G. and H. Mori collected data. M.A.S., R.L.P., M.A.M., S.K. and Y. Sauntharajah designed research, analyzed and interpreted data, and wrote the manuscript. Y.D., S.O. and J.P.M. designed research, contributed analytical tools, collected data, analyzed and interpreted data, and wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Piazza, R. *et al.* Recurrent *SETBP1* mutations in atypical chronic myeloid leukemia. *Nat. Genet.* **45**, 18–24 (2013).
2. Hoischen, A. *et al. De novo* mutations of *SETBP1* cause Schinzel-Giedion syndrome. *Nat. Genet.* **42**, 483–485 (2010).
3. Damm, F. *et al. SETBP1* mutations in 658 patients with myelodysplastic syndromes, chronic myelomonocytic leukemia and secondary acute myeloid leukemias. *Leukemia* **27**, 1401–1403 (2013).
4. Laborde, R.R. *et al. SETBP1* mutations in 415 patients with primary myelofibrosis or chronic myelomonocytic leukemia (CMML): independent prognostic impact in CMML. *Leukemia* published online; doi:10.1038/leu.2013.97 (5 April 2013).
5. Thol, F. *et al. SETBP1* mutation analysis in 944 patients with MDS and AML. *Leukemia* published online; doi:10.1038/leu.2013.145 (7 May 2013).
6. Osato, M. *et al.* Biallelic and heterozygous point mutations in the runt domain of the AML1/PEBP2αB gene associated with myeloblastic leukemias. *Blood* **93**, 1817–1824 (1999).
7. Levine, R.L. *et al.* Activating mutation in the tyrosine kinase *JAK2* in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. *Cancer Cell* **7**, 387–397 (2005).
8. Farr, C.J., Saiki, R.K., Erlich, H.A., McCormick, F. & Marshall, C.J. Analysis of *RAS* gene mutations in acute myeloid leukemia by polymerase chain reaction and oligonucleotide probes. *Proc. Natl. Acad. Sci. USA* **85**, 1629–1633 (1988).
9. Lyons, J., Janssen, J.W., Bartram, C., Layton, M. & Mufti, G.J. Mutation of Ki-*ras* and N-*ras* oncogenes in myelodysplastic syndromes. *Blood* **71**, 1707–1712 (1988).
10. Sanada, M. *et al.* Gain-of-function of mutated *C-CBL* tumour suppressor in myeloid neoplasms. *Nature* **460**, 904–908 (2009).
11. Delhommeau, F. *et al.* Mutation in *TET2* in myeloid cancers. *N. Engl. J. Med.* **360**, 2289–2301 (2009).
12. Ernst, T. *et al.* Inactivating mutations of the histone methyltransferase gene *EZH2* in myeloid disorders. *Nat. Genet.* **42**, 722–726 (2010).
13. Ley, T.J. *et al. DNMT3A* mutations in acute myeloid leukemia. *N. Engl. J. Med.* **363**, 2424–2433 (2010).
14. Mardis, E.R. *et al.* Recurring mutations found by sequencing an acute myeloid leukemia genome. *N. Engl. J. Med.* **361**, 1058–1066 (2009).
15. Yoshida, K. *et al.* Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **478**, 64–69 (2011).
16. Papaemmanuil, E. *et al.* Somatic *SF3B1* mutation in myelodysplasia with ring sideroblasts. *N. Engl. J. Med.* **365**, 1384–1395 (2011).
17. Graubert, T.A. *et al.* Recurrent mutations in the *U2AF1* splicing factor in myelodysplastic syndromes. *Nat. Genet.* **44**, 53–57 (2012).
18. Walter, M.J. *et al.* Clonal architecture of secondary acute myeloid leukemia. *N. Engl. J. Med.* **366**, 1090–1098 (2012).
19. Walter, M.J. *et al.* Clonal diversity of recurrently mutated genes in myelodysplastic syndromes. *Leukemia* **27**, 1275–1282 (2013).
20. Minakuchi, M. *et al.* Identification and characterization of SEB, a novel protein that binds to the acute undifferentiated leukemia-associated protein SET. *Eur. J. Biochem.* **268**, 1340–1351 (2001).
21. Cristóbal, I. *et al.* SETBP1 overexpression is a novel leukemogenic mechanism that predicts adverse outcome in elderly patients with acute myeloid leukemia. *Blood* **115**, 615–625 (2010).
22. Ott, M.G. *et al.* Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of *MDS1-EVI1*, *PRDM16* or *SETBP1*. *Nat. Med.* **12**, 401–409 (2006).
23. Schinzel, A. & Giedion, A. A syndrome of severe midface retraction, multiple skull anomalies, clubfeet, and cardiac and renal malformations in sibs. *Am. J. Med. Genet.* **1**, 361–375 (1978).
24. Rodríguez, J.I., Jimenez-Heffernan, J.A. & Leal, J. Schinzel-Giedion syndrome: autopsy report and additional clinical manifestations. *Am. J. Med. Genet.* **53**, 374–377 (1994).
25. Pardanani, A. *et al. CSF3R* T618I is a highly prevalent and specific mutation in chronic neutrophilic leukemia. *Leukemia* published online; doi:10.1038/leu.2013.122 (22 April 2013).
26. Meggendorfer, M. *et al. SETBP1* mutations occur in 9% of MDS/MPN and in 4% of MPN cases and are strongly associated with atypical CML, monosomy 7, isochromosome i(17)(q10), *ASXL1* and *CBL* mutations. *Leukemia* published online; doi:10.1038/leu.2012.133 (30 April 2013).
27. Makishima, H. *et al. CBL mutation-related patterns of phosphorylation and* sensitivity to tyrosine kinase inhibitors. *Leukemia* **26**, 1547–1554 (2012).
28. Sakaguchi, H. *et al.* Exome sequencing identifies secondary mutations of *SETBP1* and *JAK3* in juvenile myelomonocytic leukemia. *Nat. Genet.* published online; doi:10.1038/ng.2693 (7 July 2013).
29. Goyama, S. *et al.* Evi-1 is a critical regulator for hematopoietic stem cells and transformed leukemic cells. *Cell Stem Cell* **3**, 207–220 (2008).
30. Greenberg, P. *et al.* International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood* **89**, 2079–2088 (1997).
31. Oakley, K. *et al. Setbp1* promotes the self-renewal of murine myeloid progenitors via activation of *Hoxa9* and *Hoxa10*. *Blood* **119**, 6099–6108 (2012).
32. Cohen, S.B., Zheng, G., Heyman, H.C. & Stavnezer, E. Heterodimers of the SnoN and Ski oncoproteins form preferentially over homodimers and are more potent transforming agents. *Nucleic Acids Res.* **27**, 1006–1014 (1999).
33. Cristóbal, I. *et al.* PP2A impaired activity is a common event in acute myeloid leukemia and its activation by forskolin has a potent anti-leukemic effect. *Leukemia* **25**, 606–614 (2011).

ONLINE METHODS

Subject population. Bone marrow aspirates or blood samples were collected from 727 individuals with various myeloid malignancies seen at the Cleveland Clinic, the University of Tokyo, the University of California, Los Angeles, the Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Chang Gung University and Showa University (**Supplementary Table 6**). Informed consent for sample collection was obtained according to protocols approved by the institutional review board at each participating institute and in accordance with the Declaration of Helsinki. Diagnosis was confirmed and assigned according to World Health Organization (WHO) classification criteria[34]. Prognostic risk assessment was assigned according to the International Scoring Criteria for individuals with MDS and chronic myelomonocytic leukemia with a white cell count of <12,000 cells/μl[30]. For the purpose of this study, low-risk MDS was defined as having <5% myeloblasts. Individuals with ≥5% myeloblasts constituted those with higher risk disease. Serial samples were obtained for 12 individuals with *SETBP1* mutations. As a source of germline controls, immunoselected $CD3^+$ T lymphocytes were used in an additional nine cases. Cytogenetic analysis was performed according to standard banding techniques on the basis of 20 metaphases, if available. Clinical parameters studied included age, sex, overall survival, bone marrow blast counts and metaphase cytogenetics.

Cytogenetics and SNP arrays. Technical details regarding sample processing for SNP array assays were previously described[35,36]. The Gene Chip Mapping 250K Assay kit and the Genome-Wide Human SNP Array 6.0 (Affymetrix) were used. A stringent algorithm was applied for the identification of lesions using SNP arrays. Individuals with lesions identified by SNP array concordant with those identified in metaphase cytogenetics or typical lesions known to be recurrent required no further analysis. Changes reported in our internal or publicly available (Database of Genomic Variants; see URLs) copy number variation (CNV) databases were considered non-somatic and were excluded. Results were analyzed using CNAG (v3.0)[37] or Genotyping Console (Affymetrix). All other lesions were confirmed as somatic or germline by analysis of CD3-sorted cells[38].

Whole-exome sequencing. Whole-exome sequencing was performed as previously reported[15]. Briefly, tumor DNA was extracted from bone marrow or peripheral blood mononuclear cells from affected individuals. For germline controls, DNA was obtained from paired $CD3^+$ T cells. Whole-exome capture was accomplished using liquid-phase hybridization of sonicated genomic DNA with mean length of 150–200 bp to the bait cRNA library synthesized on magnetic beads (SureSelect, Agilent Technologies) according to the manufacturer's protocol. The SureSelect Human All Exon 50Mb kit was used for 20 cases (**Supplementary Table 1**). Captured targets were subjected to massive sequencing using the Illumina HiSeq 2000 platform with the paired-end 75- to 108-bp read option, according to the manufacturer's instructions. Raw sequence data generated from HiSeq 2000 sequencers were processed through the in-house pipeline constructed for the whole-exome analysis of paired cancer genomes at the Human Genome Center, Institute of Medical Science, University of Tokyo, which is summarized in a previous report[15]. Data processing is divided into two steps: (i) generation of a BAM file (using SAMtools) for paired normal and tumor samples for each case and (ii) detection of somatic SNVs and indels by comparing normal and tumor BAM files. Alignment of sequencing reads on the hg19 reference genome was visualized using Integrative Genomics Viewer (IGV) software[39].

For all candidate somatic mutations, the accuracy of the prediction of these SNVs and indels by whole-exome sequencing was tested by validation of 65 genes (80 events) by Sanger sequencing and targeted deep sequencing. Prediction had a true positive rate of 47% (39% for missense mutation, 75% for nonsense mutations and 75% for indels). It is of note that prediction of known somatic mutations (for example, in *TET2* ($n = 9$), *CBL* ($n = 2$), *SETBP1* ($n = 2$) and *ASXL1* ($n = 2$)) showed accuracy of 100% (**Supplementary Tables 2–4**).

Targeted deep sequencing. To detect allelic frequencies for mutations or SNPs, we applied deep sequencing to targeted exons as previously described[15]. Briefly, we screened for possible mutations of *SETBP1* and other genes that were concomitantly mutated in the cases with *SETBP1* mutation (*U2AF1*, *DNMT3A*, *NRAS*, *ASXL1*, *SRSF2*, *CBL*, *IDH1*, *IDH2*, *SRSF2*, *TET2*, *PTPN11* and *RUNX1*). Each targeted exon was amplified with NotI linker attached to each primer as previously described[15]. After digestion with NotI, amplicons were ligated with T4 DNA ligase and sonicated into fragments that were on average up to 200 bp in size using Covaris. Sequencing libraries were generated according to an Illumina paired-end library protocol and were subjected to deep sequencing on the Illumina Genome Analyzer IIx or HiSeq 2000 sequencers according to the standard protocol.

Sanger sequencing and allele-specific PCR. Exons of selected genes were amplified and underwent direct genomic sequencing by standard techniques on the ABI 3730xl DNA analyzer (Applied Biosystems) as previously described[40–42]. Coding and sequenced exons are shown in **Supplementary Table 8**. All mutations were detected by bidirectional sequencing and were scored as pathogenic if not present in non-clonal paired DNA from CD3-selected cells. When a mutant allele with small burden was not confirmed by Sanger sequencing, cloning and sequencing of individual colonies (TOPO TA cloning, Invitrogen) was performed for validation. The allelic presence of p.Asp868Asn and p.Gly870Ser alterations was determined by allele-specific PCR. Primer sequences for *SETBP1* sequencing and *SETBP1* allele-specific PCR are provided in **Supplementary Table 14**.

Quantitative RT-PCR using TaqMan probes. Total RNA was extracted from bone marrow mononuclear cells and cell lines. cDNA was synthesized from 500 ng of total RNA using the iScript cDNA synthesis kit (Bio-Rad). Quantitative gene expression levels were detected using RT-PCR with the ABI PRISM 7500 Fast Sequence Detection System and FAM dye–labeled TaqMan MGB probes (Applied Biosystems). TaqMan probes for all genes analyzed were gene expression assay products purchased from Applied Biosystems (*SETBP1*, Hs00210209_m1; *HOXA9*, Hs00365956_m1; *HOXA10*, Hs00172012_m1; *GAPDH*, Hs99999905_m1). Expression levels of target genes were normalized to *GAPDH* mRNA levels.

Retrovirus generation. pMYs-*Setbp1* retrovirus expressing 3× Flag–tagged wild-type Setbp1 protein and green fluorescent protein (GFP) marker was described previously[31]. Point mutations of *Setbp1* (encoding p.Asp868Asn and p.Ile871Thr alterations) were generated using the same construct and the QuickChange II Site-Directed Mutagenesis kit (Agilent Technologies). Virus was produced by transient transfection of Plat-E cells (Cell Biolabs) using FuGene 6 (Roche). Viral titers were calculated by infecting NIH3T3 cells with serially diluted viral stock and counting GFP-positive colonies 48 h after infection.

Immortalization of myeloid progenitors. Immortalization of myeloid progenitors was performed as described according to protocols approved by the Institutional Animal Care and Use Committee of the Uniformed Services University of the Health Sciences[31]. Briefly, whole-bone marrow cells harvested from three young C57BL/6 mice were first cultured in StemSpan medium (Stemcell Technologies) with 10 ng/ml mouse SCF, 20 ng/ml mouse TPO, 20 ng/ml mouse IGF-2 (all from R&D Systems) and 10 ng/ml human FGF-1 (Invitrogen) for 6 d to expand primitive stem and progenitor cells. Myeloid differentiation was subsequently induced by growing the expanded cells in IMDM supplemented with 20% heat-inactivated horse serum with 100 ng/ml mouse SCF (PeproTech) and 10 ng/ml mouse IL-3 for 4 d. Resulting cells (5×10^5) were infected with retrovirus (1×10^5 colony-forming units (CFUs)) on plates coated with Retronectin (Takara) for 48 h. Infected cells were then continuously passaged at a 1:10 ratio every 3 d for 4 weeks to test whether transduction caused immortalization of the myeloid progenitors. In the absence of immortalization, transduced cultures generally ceased expanding in 2 weeks.

Methylation analysis. The DNA methylation status of bisulfite-treated genomic DNA was probed at 27,578 CpG dinucleotides using the Illumina Infinium HumanMethylation 27k BeadChip assay as previously described[43]. Briefly, methylation status was calculated from the ratio of methylation-specific and demethylation-specific fluorophores (β value) using the BeadStudio Methylation Module (Illumina).

Resistance of SETBP1 protein degradation associated with *SETBP1* mutation. Full-length wild-type human *SETBP1* cDNA encoding 3× HA–tagged protein was cloned from peripheral blood mononuclear cells. Mutagenesis of *SETBP1* (to introduce mutations encoding the p.Asp868Asn and p.Ile871Thr alterations) was performed using the PrimeSTAR kit (Takara Bio). Wild-type and mutant cDNA constructs were cloned into the CS-Ubc lentivirus vector (a kind gift of T. Yamaguchi). Vectors were cotransfected with packaging vector and with vectors expressing VSV-G and Rev into 293T cells, and lentiviral particles were harvested. Protein blotting experiments on whole lysates from Jurkat cell line stably transduced with viruses expressing wild-type and mutant SETBP1 were carried out with antibodies for HA at a 1:2,000 dilution (MMS-101R, Covance) and actin at a 1:1,000 dilution (sc-1616, Santa Cruz Biotechnology). Both cell lines were obtained from ATCC. For proteasomal inhibition, cell lines were treated with 0.5 μM lactacystin (Peptide Institute) and 0.25 μM bafilomycin A1 (Wako Junyaku) for 2 h.

Statistical analysis. The Kaplan-Meier method was used to analyze survival outcomes (overall survival) by the log-rank test. Pairwise comparisons were performed by Wilcoxon test for continuous variables and by two-sided Fisher's exact test for categorical variables. Paired data were analyzed by Wilcoxon signed-rank test. For multivariate analyses, a Cox proportional hazards model was conducted for overall survival. Variables considered for model inclusion were IPSS risk group, age, sex and gene mutation status. Variables with $P < 0.05$ in univariate analyses were included in the model. Statistical analyses were performed with JMP9 software (SAS). Significance was determined at a two-sided α level of 0.05, except for P values in multiple comparisons, in which Bonferroni correction was applied.

34. Shaffer, L.G. & Tommerup, N. *ISCN 2009. An International System for Human Cytogenetics Nomenclature* (Karger, Basel, Switzerland, 2009).
35. Maciejewski, J.P., Tiu, R.V. & O'Keefe, C. Application of array-based whole genome scanning technologies as a cytogenetic tool in haematological malignancies. *Br. J. Haematol.* **146**, 479–488 (2009).
36. Gondek, L.P. *et al.* Chromosomal lesions and uniparental disomy detected by SNP arrays in MDS, MDS/MPD, and MDS-derived AML. *Blood* **111**, 1534–1542 (2008).
37. Nannya, Y. *et al.* A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. *Cancer Res.* **65**, 6071–6079 (2005).
38. Tiu, R.V. *et al.* New lesions detected by single nucleotide polymorphism array–based chromosomal analysis have important clinical impact in acute myeloid leukemia. *J. Clin. Oncol.* **27**, 5219–5226 (2009).
39. Robinson, J.T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
40. Dunbar, A.J. *et al.* 250K single nucleotide polymorphism array karyotyping identifies acquired uniparental disomy and homozygous mutations, including novel missense substitutions of *c-Cbl*, in myeloid malignancies. *Cancer Res.* **68**, 10349–10357 (2008).
41. Jankowska, A.M. *et al.* Loss of heterozygosity 4q24 and *TET2* mutations associated with myelodysplastic/myeloproliferative neoplasms. *Blood* **113**, 6403–6410 (2009).
42. Makishima, H. *et al. CBL, CBLB, TET2, ASXL1,* and *IDH1/2* mutations and additional chromosomal aberrations constitute molecular events in chronic myelogenous leukemia. *Blood* **117**, e198–e206 (2011).
43. Ko, M. *et al.* Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant *TET2*. *Nature* **468**, 839–843 (2010).

doi:10.1038/ng.2696

nature genetics

Exome sequencing identifies secondary mutations of *SETBP1* and *JAK3* in juvenile myelomonocytic leukemia

Hirotoshi Sakaguchi[1,8], Yusuke Okuno[2,8], Hideki Muramatsu[1,8], Kenichi Yoshida[2,8], Yuichi Shiraishi[3], Mariko Takahashi[2], Ayana Kon[2], Masashi Sanada[2,4], Kenichi Chiba[3], Hiroko Tanaka[5], Hideki Makishima[6], Xinan Wang[1], Yinyan Xu[1], Sayoko Doisaki[1], Asahito Hama[1], Koji Nakanishi[1], Yoshiyuki Takahashi[1], Nao Yoshida[7], Jaroslaw P Maciejewski[6], Satoru Miyano[3,5], Seishi Ogawa[2,4,9] & Seiji Kojima[1,9]

Juvenile myelomonocytic leukemia (JMML) is an intractable pediatric leukemia with poor prognosis[1] whose molecular pathogenesis is poorly understood, except for somatic or germline mutations of RAS pathway genes, including *PTPN11*, *NF1*, *NRAS*, *KRAS* and *CBL*, in the majority of cases[2–4]. To obtain a complete registry of gene mutations in JMML, whole-exome sequencing was performed for paired tumor-normal DNA from 13 individuals with JMML (cases), which was followed by deep sequencing of 8 target genes in 92 tumor samples. JMML was characterized by a paucity of gene mutations (0.85 non-silent mutations per sample) with somatic or germline RAS pathway involvement in 82 cases (89%). The *SETBP1* and *JAK3* genes were among common targets for secondary mutations. Mutations in the latter were often subclonal and may be involved in the progression rather than the initiation of leukemia, and these mutations associated with poor clinical outcome. Our findings provide new insights into the pathogenesis and progression of JMML.

JMML is a rare myelodysplastic/myeloproliferative neoplasm unique to childhood, characterized by excessive proliferation of myelomonocytic cells and hypersensitivity to granulocyte-macrophage colony-stimulating factor[1]. A cardinal genetic feature of JMML is frequent somatic and/or germline mutation of RAS pathway genes, such as *NF1*, *NRAS*, *KRAS*, *PTPN11* and *CBL*, which are mutated in more than 70% of JMML cases in a mutually exclusive manner[2–4]. However, it is still open to question whether RAS pathway mutations are sufficient for the development of JMML or if secondary mutations have a role in the development and progression of this cancer. To address these issues and to better define the molecular pathogenesis of JMML, we performed whole-exome sequencing of paired tumor-normal DNA from 13 cases (**Supplementary Table 1**). We obtained mean coverage in exome sequencing of 137× for tumor samples and 143× for normal samples (**Supplementary Fig. 1**). A Monte-Carlo simulation indicated that the study detected 88% of the existing somatic mutations (Online Methods and **Supplementary Fig. 2**).

Sanger sequencing of 25 candidate non-silent somatic nucleotide alterations confirmed 1 nonsense and 10 missense mutations (**Table 1** and **Supplementary Fig. 3**), with the low true positive rate consistent with the very low numbers of somatic mutations in JMML. Of the 11 somatic mutations, 6 involved known RAS pathway genes. In addition, non-overlapping RAS pathway mutations (6 somatic and 6 germline) were confirmed in 11 of the 13 discovery cases (86%; **Table 1**). For the remaining two cases that lacked documented RAS pathway mutations, we intensively searched for possible germline mutations that could be relevant to the development of JMML. In total, 179 and 167 candidate germline mutations were detected in subjects 77 and 92, respectively, but these mutations did not affect known RAS pathway genes or other cancer-related genes, including the ones registered in the pathway databases (Online Methods). A frameshift deletion in *KMT2D* (also known as *MLL2*; encoding p.Val1670fs) was found in subject 92, who had been diagnosed as having Noonan syndrome on the basis of typical features such as hypertelorism, webbed neck and congenital heart disease (**Supplementary Fig. 3**) but lacked the distinctive facial appearance of Kabuki syndrome, which was shown to be caused by germline *KMT2D* mutations[5].

Five of the 11 somatic mutations were non–RAS pathway mutations, involving *SETBP1* (3 p.Asp868Asn alterations), *JAK3* (1 p.Arg657Gln alteration) and *SH3BP1* (1 p.Ser277Leu alteration), which had not been reported in JMML cases. SETBP1 was originally isolated as a 170-kDa nuclear protein that interacts with SET, a small protein inhibitor of the putative tumor suppressors PP2A and NM23-H1 (ref. 6). Several lines of recent evidence suggest that SETBP1 has a role in leukemogenesis (**Supplementary Fig. 4**)[7–11]. *SETBP1* participates in translocations that result in an aberrant fusion gene (*NUP98-SETBP1*) and overexpression of *SETBP1* in T cell acute lymphoblastic leukemia (T-ALL) and acute myeloid leukemia (AML), respectively[12,13]. *SETBP1* is one of the downstream targets induced by the Evi-1 oncoprotein[14] and, together with *EVI1* and its homolog *PRDM16* (also known as *MEL1*), was reported to be activated through retrovirus integration. SETBP1 is also known to augment the recovery of granulopoiesis after gene therapies for chronic granulomatous disease[15]. *SETBP1* overexpression is found in more than 27% of adult AML cases and is associated with poor survival[13]. The discovery of recurrent hotspot mutations of *SETBP1* provides unequivocal evidence for the leukemogenic role of deregulated SETBP1 function. Notably, the *SETBP1* mutation encoding p.Asp868Asn was identical to one of the *de novo* mutations reported to be causative in Schinzel-Giedion syndrome (SGS; MIM 269150), which is a highly recognizable congenital disease characterized by severe mental retardation, distinctive facial features and

[1]Department of Pediatrics, Nagoya University Graduate School of Medicine, Nagoya, Japan. [2]Cancer Genomics Project, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. [3]Laboratory of DNA Information Analysis, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan. [4]Department of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto University, Kyoto, Japan. [5]Laboratory of Sequence Analysis, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan. [6]Department of Translational Hematology and Oncology Research, Taussig Cancer Institute, Cleveland Clinic, Cleveland, Ohio, USA. [7]Department of Hematology and Oncology, Children's Medical Center, Japanese Red Cross Nagoya First Hospital, Nagoya, Japan. [8]These authors contributed equally to this work. [9]These authors jointly directed this work. Correspondence should be addressed to S.O. (sogawa-tky@umin.ac.jp) or S.K. (kojimas@med.nagoya-u.ac.jp).

Received 6 November 2012; accepted 17 June 2013; published online 7 July 2013; doi:10.1038/ng.2698

Table 1 List of gene mutations identified by whole-exome sequencing

	RAS pathway mutations								Other somatic mutations			
	Somatic				Germline							
Subject number	Gene	Change at DNA level	Change at protein level	VAF[a]	Gene	Change at DNA level	Change at protein level	VAF[a]	Gene	Change at DNA level	Change at protein level	VAF[a]
11[b]	*NF1*	c.4537C>T	p.Arg1513*	40.1/24.2	*NF1*	c.5927delG	p.Trp1976fs	44.0/47.1	*SETBP1*	c.2602G>A	p.Asp868Asn	32.6/27.0
63	*KRAS*	c.38G>A	p.Gly13Asp	44.3/0.0	–	–	–	–	–	–	–	–
72	*PTPN11*	c.172A>T	p.Asn58Tyr	48.2/5.7	–	–	–	–	*SETBP1*	c.2602G>A	p.Asp868Asn	45.9/2.5
									JAK3	c.1970G>A	p.Arg657Gln	30.5/2.2
									SH3BP1	c.830C>T	p.Ser277Leu	47.8/5.1
77	–	–	–	–	–	–	–	–	*SETBP1*	c.2602G>A	p.Asp868Asn	33.4/2.1
78	*NRAS*	c.35G>C	p.Gly12Ala	45.5/9.5	–	–	–	–	–	–	–	–
82	–	–	–	–	*CBL*	c.1217del22	p.Thr406fs	34.7/38.9	–	–	–	–
83	–	–	–	–	*NF1*	c.4970A>G	p.Tyr1657Cys	50.0/51.0	–	–	–	–
84	–	–	–	–	*CBL*	c.1096–110del643	p.Glu366_Phe488del	NA/NA	–	–	–	–
85	*PTPN11*	c.226G>A	p.Glu76Lys	47.5/4.4	–	–	–	–	–	–	–	–
86	*KRAS*	c.38G>A	p.Gly13Asp	38.9/3.1	–	–	–	–	–	–	–	–
89[c]	–	–	–	–	*PTPN11*	c.1502T>G	p.Ser502Ala	50.0/49.9	–	–	–	–
91[c]	–	–	–	–	*PTPN11*	c.218C>T	p.Thr73Ile	49.0/48.0	–	–	–	–
92[c]	–	–	–	–	–	–	–	–	–	–	–	–

NA, not available.
[a]Variant allele frequency (VAF) in tumor/reference samples, where the reference was CD3+ T cells, except for subject 63, for whom umbilical cord was used as the reference. [b]Substantial contamination of tumor cell components in the CD3+ T cell reference. [c]Noonan syndrome–associated myeloproliferative disorder.

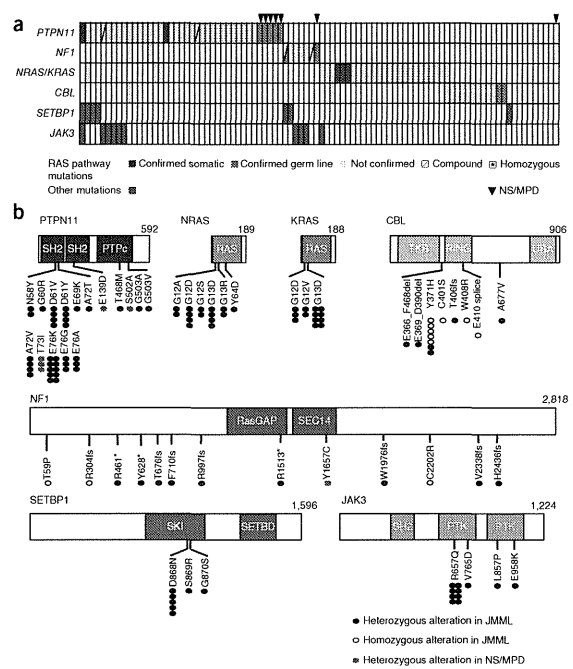


Figure 1 Mutation profiles of 92 JMML cases. (**a**) The mutation status of RAS pathway genes and 2 newly identified gene targets in a cohort of 92 JMML cases is summarized. NS/MPD, Noonan syndrome–associated myeloproliferative disorder. (**b**) The distribution of alterations is shown for each protein. SH2, Src homology 2 domain; PTPc, protein tyrosine phosphatase, catalytic domain; RAS, Ras GTPase family domain; TKB, tyrosine kinase–binding domain; RING, RING-finger domain; UBA, ubiquitin-associated domain; RasGAP, a region of similarity with the catalytic domain of the mammalian p120RasGAP protein in neurofibromin; SEC14, Sec14p-like lipid-binding domain; SKI, v-ski sarcoma viral oncogene homolog domain; SETBD, SET-binding domain; PTK, pseudokinase domain of the protein tyrosine kinases.

Table 2 Subject characteristics

Characteristic	Total cohort (n = 92)	Secondary mutations		P value
		Yes (n = 16)	No (n = 76)	
Sex (male/female)	61/31	12/4	49/27	NS
Median age at diagnosis in months (range)	19 (1–160)	38 (2–160)	13 (1–79)	<0.001
Diagnosis				
JMML	85	16	69	
NS/MPD	7	0	7	
Genetic mutations in RAS pathway				
PTPN11	39	9	30	NS
NF1	9	5	4	0.001
RAS (*NRAS* or *KRAS*)	28 (15/13)	2 (1/1)	26 (14/12)	0.08
CBL	14	0	14	0.06
Without RAS pathway mutation	10	1	9	NS
Secondary genetic mutations				
SETBP1	7	7	0	
JAK3	10	10	0	
Cytogenetics				
Normal karyotype	77	12	65	NS
Monosomy 7	8	1	7	NS
Trisomy 8	4	2	2	NS
Other abnormalities	3	1	2	NS
WBC count at diagnosis $\times 10^9$/l, median (range)	30.0 (1.0–563)	29.6 (5.6–563)	30.0 (1.0–131)	NS
Monocyte count at diagnosis $\times 10^9$/l, median (range)	4.6 (0.2–31.6)	3.1 (0.5–15.2)	4.9 (0.2–31.6)	NS
Percent HbF at diagnosis, median (range)	21 (0–68)	26 (9–55)	16 (0–68)	NS
PLT at diagnosis $\times 10^9$/l, median (range)	61.0 (1.4–483)	47.5 (1.4–175)	65.0 (5.0–483)	NS
HSCT (+/–)	56/36	16/0	40/36	
Alive/deceased	62/30	7/9	55/21	
Percent probability of 5-year overall survival (95% CI)	60 (46–71)	33 (10–59)	65 (49–77)	0.10
Percent probability of 5-year transplantation-free survival (95% CI)	15 (6–27)	0 (0–0)	18 (8–33)	0.007

JMML, juvenile myelomonocytic leukemia; NS/MPD, Noonan syndrome–associated myeloproliferative disorder; WBC, white blood cell; HbF, hemoglobin F; HSCT, hematopoietic stem cell transplantation; NS, not significant. We compared the difference between the subjects with and without secondary mutation, and P values were calculated by two-sided Fisher's exact test or Mann-Whitney U test.

multiple congenital malformations. Individuals with SGS with this mutation have a higher than normal prevalence of tumors, including of neuroepithelial neoplasia[16], although development of myeloid malignancies has not been reported so far.

To further validate our findings, we screened the entire cohort of 92 JMML cases for gene mutations in the newly identified 3 genes together with known RAS pathway targets using deep sequencing[17] (**Supplementary Fig. 5**).

RAS pathway mutations were found in 82 of 92 cases (89%) in a mutually exclusive manner, with *PTPN11* mutations predominant, followed by *NRAS*, *KRAS*, *CBL* and *NF1* mutations (**Fig. 1a** and **Table 2**). In accordance with previous reports, most of the *CBL* (8/14) and *NF1* (4/9) mutations were biallelic (**Fig. 1a,b** and **Supplementary Table 2**)[2,3,18], whereas the majority of mutations in *PTPN11*, *NRAS* and *KRAS* were heterozygous[4]. The individuals without RAS pathway mutations (n = 10) were vigorously investigated by whole-genome sequencing of tumor-normal paired samples (n = 2; **Supplementary Fig. 6**) or by whole-exome sequencing of only tumor samples (n = 8; **Supplementary Fig. 7**). As anticipated, we found no known RAS pathway mutations.

On the other hand, 18 mutations were found in *SETBP1* (n = 7) or *JAK3* (n = 11) in 16 cases (**Fig. 1a,b**, **Table 2** and **Supplementary Table 2**), with these mutations more frequent in cases with mutated *PTPN11* (and possibly *NF1*) than in cases with mutated *NRAS*, *KRAS*

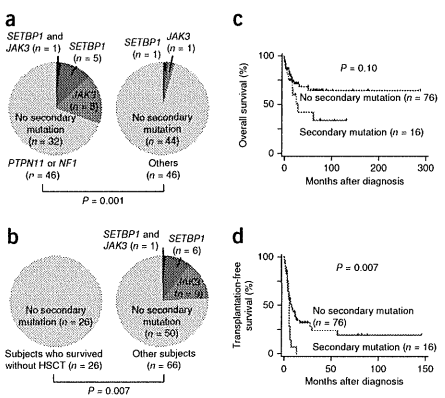


Figure 2 Clinical features of JMML cases with or without secondary mutations. (**a**,**b**) Frequency of secondary mutations in individuals with JMML depending on the type of RAS pathway mutations (left, *PTPN11* or *NF1*; right, other or no mutations) (**a**) and the status of HSCT (**b**). P values were calculated by two-sided Fisher's exact test. (**c**,**d**) The impact of secondary mutations on overall (**c**) and transplantation-free (**d**) survival is shown in Kaplan-Meier survival curves, where statistical significance was tested by log-rank test.

or *CBL* (**Fig. 2a**). Mutations in *SH3BP1*, encoding SH3 domain-binding protein 1, were not recurrent. All *SETBP1* mutations were heterozygous and occurred within the portion of the gene encoding the SKI domain, with six identical to the *de novo* recurrent mutations reported in SGS and five identical to the mutation encoding the p.Asp868Asn alteration (**Fig. 1b**). RT-PCR analysis showed that the wild-type and mutant alleles of *SETBP1* were equally expressed (**Supplementary Fig. 8**). Similarly, 8 of the 11 *JAK3* mutations in 10 cases were the well-described activating mutation (encoding a p.Arg657Gln alteration) found in various hematological malignancies, including Down syndrome–associated acute megakaryoblastic leukemia[19–23], ALL[24,25] and natural killer (NK)/T cell lymphoma[26], and the remaining 3 were also within the portions of the gene encoding the pseudokinase or kinase domain, suggestive of gain of function.

Deep sequencing of the relevant mutant alleles enabled an accurate estimation of allele frequencies for individual mutations (**Supplementary Fig. 9**). *SETBP1* and *JAK3* mutations showed lower allele frequencies (but not with statistical significance for *SETBP1*) than did the corresponding RAS pathway mutations (**Supplementary Fig. 10a**), indicating that the former mutations represent secondary genetic hits that contributed to clonal evolution after the main tumor population was established (**Supplementary Fig. 10b**). Individuals with secondary mutations had shorter lengths of survival compared to those without mutations: 5-year overall survival (hazards ratio (HR) = 1.90, 95% CI = 0.87–4.19). In addition, none of the individuals with JMML who survived without hematopoietic stem cell transplantation (HSCT; n = 26) harbored any of the secondary mutations, and individuals with secondary mutations showed significantly inferior 5-year transplant-free survival (HR = 2.18, 95% CI = 1.18–4.02) (**Fig. 2b–d** and **Table 2**).

JMML is characterized by a paucity of gene mutations. The average number of mutations per sample (0.85; range of 0–4) was unexpectedly low compared to those reported in other human cancers (**Supplementary Fig. 11**); excluding common RAS pathway mutations, only 5 mutations were detected in 3 of the 13 discovery cases. This small number of mutations is only comparable to the figure reported for retinoblastoma (mean of 3.3 per case; range of 0–5) (ref. 27) and is in stark contrast to the abundance of gene mutations in chronic myelomonocytic leukemia (CMML) in adult cases, where the mean number of non-silent mutations was 12.4 per sample, of which 3.1 represented known driver changes (ref. 17 and K.Y., M.S., Y.S., D. Nowak, Y. Nagata *et al.*, unpublished data), underscoring the distinct pathogenesis in these two neoplasms that show indistinguishable morphology. The impact of germline events is underscored by the fact that 6 of the 13 discovery cases harbored germline RAS pathway mutations and an additional case without known RAS pathway mutations showed constitutive abnormalities similar to Noonan syndrome. Despite the central role of RAS pathway mutations, a small subset of cases had no documented RAS pathway mutations, even after whole-exome analysis in the two RAS pathway mutation–negative cases, raising the possibility that the latter cases represent a genetically distinct myeloproliferative neoplasm in childhood.

Another key finding in the current study is the discovery of secondary mutations that involve *SETBP1* and *JAK3*. Detected only in a subpopulation of leukemic cells, most of these mutations are thought to be involved in the progression rather than the establishment of JMML and were associated with poor clinical outcome. *SETBP1* is a newly identified proto-oncogene, and identical mutations in this gene have recently been reported in 15–25% of adult cases with atypical chronic myeloid leukemia (CML)[10], CMML and secondary AML[28]. Affecting one of three highly conserved amino acid positions, *SETBP1* mutations have been shown to abolish the binding of an E3 ubiquitin ligase (β-TrCP1) to SETBP1, which prevents ubiquitination and subsequent degradation, leading to gain of function through the consequent increase in SETBP1 protein amounts[10,28]. Although the precise leukemogenic mechanisms of *SETBP1* mutations are still unclear, we have shown that mutant *SETBP1* alleles confer self-renewal capability to myeloid progenitors *in vitro*, and *SETBP1* mutations in adult leukemia were associated with increases in *HOXA9* and *HOXA10* expression[28]. Recurrent *JAK3* mutations in JMML are also noteworthy. The JAK-STAT pathway is a key component of normal hematopoiesis[29]. As in other hematopoietic malignancies[20], the p.Arg657Gln alteration represents the most frequent change in JMML. This alteration confers interleukin (IL)-3 independence to Ba/F3 cells and induces STAT5 phosphorylation[20]. Targeting the JAK-STAT pathway with a pan-JAK inhibitor such as CP-690550 (ref. 30) could be a promising therapeutic possibility for patients with *JAK3*-mutated JMML.

In conclusion, our whole-exome sequencing analysis identified the spectrum of gene mutations in JMML. Together with the high frequency of RAS pathway mutations, the paucity of non–RAS pathway mutations is a prominent feature of JMML. Mutations of *SETBP1* and *JAK3* were common recurrent secondary events presumed to be involved in tumor progression and were associated with poor clinical outcomes. Our findings provide an important clue to understanding the pathogenesis of JMML that may help in the development of novel diagnostics and therapeutics for this leukemia.

URLs. Genomon, http://genomon.hgc.jp/exome/en/; BioCarta, http://www.biocarta.com/; dbSNP131, http://www.ncbi.nlm.nih.gov/projects/SNP/; RefSeq database, http://www.ncbi.nlm.nih.gov/RefSeq/.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession code. We deposited whole-genome and whole-exome sequence data in the European Genome-phenome Archive under accession EGAS00001000521.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We thank the subjects and their parents for participating in this study. This work was supported by the Research on Measures for Intractable Diseases Project from the Ministry of Health, Labor and Welfare, by Grants-in-Aid from the Ministry of Health, Labor and Welfare of Japan and KAKENHI (23249052, 22134006 and 21790907), by the Project for the Development of Innovative Research on Cancer Therapeutics (P-DIRECT) and by the Japan Society for the Promotion of Science through the Funding Program for World-Leading Innovative R&D on Science and Technology.

AUTHOR CONTRIBUTIONS

H.S., Y.O., H. Muramatsu, K.Y., M.T., A.K. and M.S. designed and performed the research, analyzed the data and wrote the manuscript. Y.S., K.C., H.T. and S.M. performed bioinformatics analyses of the resequencing data. X.W. and Y.X. performed Sanger sequencing. S.D., A.H., K.N., Y.T. and N.Y. collected specimens and performed the research. H. Makishima and J.P.M. designed the research and analyzed the data. S.O. and S.K. led the entire project and wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Pinkel, D. *et al.* Differentiating juvenile myelomonocytic leukemia from infectious disease. *Blood* **91**, 365–367 (1998).
2. Loh, M.L. *et al.* Mutations in *CBL* occur frequently in juvenile myelomonocytic leukemia. *Blood* **114**, 1859–1863 (2009).
3. Muramatsu, H. *et al.* Mutations of an E3 ubiquitin ligase c-*Cbl* but not *TET2* mutations are pathogenic in juvenile myelomonocytic leukemia. *Blood* **115**, 1969–1975 (2010).
4. Pérez, B. *et al.* Genetic typing of *CBL*, *ASXL1*, *RUNX1*, *TET2* and *JAK2* in juvenile myelomonocytic leukaemia reveals a genetic profile distinct from chronic myelomonocytic leukaemia. *Br. J. Haematol.* **151**, 460–468 (2010).
5. Ng, S.B. *et al.* Exome sequencing identifies *MLL2* mutations as a cause of Kabuki syndrome. *Nat. Genet.* **42**, 790–793 (2010).
6. Minakuchi, M. *et al.* Identification and characterization of SEB, a novel protein that binds to the acute undifferentiated leukemia–associated protein SET. *Eur. J. Biochem.* **268**, 1340–1351 (2001).
7. Damm, F. *et al. SETBP1* mutations in 658 patients with myelodysplastic syndromes, chronic myelomonocytic leukemia and secondary acute myeloid leukemias. *Leukemia* **27**, 401–403 (2013).
8. Laborde, R.R. *et al. SETBP1* mutations in 415 patients with primary myelofibrosis or chronic myelomonocytic leukemia: independent prognostic impact in CMML. *Leukemia* published online; doi:10.1038/leu.2013.97 (5 April 2013).
9. Meggendorfer, M. *et al. SETBP1* mutations occur in 9% of MDS/MPN and in 4% of MPN cases and are strongly associated with atypical CML, monosomy 7, isochromosome i(17)(q10), *ASXL1* and *CBL* mutations. *Leukemia* published online; doi:10.1038/leu.2013.133 (30 April 2013).
10. Piazza, R. *et al.* Recurrent *SETBP1* mutations in atypical chronic myeloid leukemia. *Nat. Genet.* **45**, 18–24 (2013).
11. Thol, F. *et al. SETBP1* mutation analysis in 944 patients with MDS and AML. *Leukemia* published online; doi:10.1038/leu.2013.145 (7 May 2013).
12. Panagopoulos, I. *et al.* Fusion of *NUP98* and the SET binding protein 1 (*SETBP1*) gene in a paediatric acute T cell lymphoblastic leukaemia with t(11;18)(p15;q12). *Br. J. Haematol.* **136**, 294–296 (2007).
13. Cristóbal, I. *et al.* SETBP1 overexpression is a novel leukemogenic mechanism that predicts adverse outcome in elderly patients with acute myeloid leukemia. *Blood* **115**, 615–625 (2010).
14. Goyama, S. *et al.* Evi-1 is a critical regulator for hematopoietic stem cells and transformed leukemic cells. *Cell Stem Cell* **3**, 207–220 (2008).
15. Ott, M.G. *et al.* Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of *MDS1-EVI1*, *PRDM16* or *SETBP1*. *Nat. Med.* **12**, 401–409 (2006).
16. Hoischen, A. *et al. De novo* mutations of *SETBP1* cause Schinzel-Giedion syndrome. *Nat. Genet.* **42**, 483–485 (2010).
17. Yoshida, K. *et al.* Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **478**, 64–69 (2011).
18. Flotho, C. *et al.* Genome-wide single-nucleotide polymorphism analysis in juvenile myelomonocytic leukemia identifies uniparental disomy surrounding the *NF1* locus in cases associated with neurofibromatosis but not in cases with mutant *RAS* or *PTPN11*. *Oncogene* **26**, 5816–5821 (2007).
19. Walters, D.K. *et al.* Activating alleles of *JAK3* in acute megakaryoblastic leukemia. *Cancer Cell* **10**, 65–75 (2006).
20. Sato, T. *et al.* Functional analysis of *JAK3* mutations in transient myeloproliferative disorder and acute megakaryoblastic leukaemia accompanying Down syndrome. *Br. J. Haematol.* **141**, 681–688 (2008).
21. De Vita, S. *et al.* Loss-of-function *JAK3* mutations in TMD and AMKL of Down syndrome. *Br. J. Haematol.* **137**, 337–341 (2007).
22. Norton, A. *et al.* Analysis of *JAK3*, *JAK2*, and *C-MPL* mutations in transient myeloproliferative disorder and myeloid leukemia of Down syndrome blasts in children with Down syndrome. *Blood* **110**, 1077–1079 (2007).
23. Kiyoi, H., Yamaji, S., Kojima, S. & Naoe, T. *JAK3* mutations occur in acute megakaryoblastic leukemia both in Down syndrome children and non–Down syndrome adults. *Leukemia* **21**, 574–576 (2007).
24. Elliott, N.E. *et al.* FERM domain mutations induce gain of function in JAK3 in adult T-cell leukemia/lymphoma. *Blood* **118**, 3911–3921 (2011).
25. Zhang, J. *et al.* The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature* **481**, 157–163 (2012).
26. Koo, G.C. *et al.* Janus kinase 3–activating mutations identified in natural killer/T-cell Lymphoma. *Cancer Discov.* **2**, 591–597 (2012).
27. Zhang, J. *et al.* A novel retinoblastoma therapy from genomic and epigenetic analyses. *Nature* **481**, 329–334 (2012).
28. Makishima, H. *et al.* Somatic *SETBP1* mutations in myeloid malignancies. *Nat. Genet.* published online; doi:10.1038/ng.2696 (7 July 2013).
29. Crozatier, M. & Meister, M. *Drosophila* haematopoiesis. *Cell. Microbiol.* **9**, 1117–1126 (2007).
30. Changelian, P.S. *et al.* Prevention of organ allograft rejection by a specific Janus kinase 3 inhibitor. *Science* **302**, 875–878 (2003).

© 2013 Nature America, Inc. All rights reserved.

ONLINE METHODS

Subjects. We studied 92 children (61 boys and 31 girls) with JMML, including 7 individuals with NS/MPD, who were diagnosed as having JMML in institutions throughout Japan. Written informed consent was obtained from subjects' parents before sample collection. This study was approved by the ethics committees of the Nagoya University Graduate School of Medicine and the University of Tokyo in accordance with the Declaration of Helsinki. Diagnosis with JMML was made on the basis of internationally accepted criteria[1]. Characteristics of the 92 JMML cases are summarized in **Table 2**. The median age at diagnosis was 16 months (range of 1–160 months). Karyotypic abnormalities were detected in 16 subjects, including in 8 with monosomy 7. Fifty-six of the 92 subjects (61%) received allogeneic HSCT.

Sample preparation. Genomic DNA was extracted using the QIAamp DNA Blood Mini kit and the QIAamp DNA Investigator kit (Qiagen) according to the manufacturer's instructions. The T Cell Activation/Expansion kit, human (Miltenyi Biotec) was used for the expansion of $CD3^+$ T cells from subjects' peripheral blood or bone marrow mononuclear cells[3].

Whole-exome sequencing. Exome capture from paired tumor-reference DNA was performed using SureSelect Human All Exon V3 (Agilent Technologies), covering 50 Mb of coding exons, according to the manufacturer's protocol. Enriched exome fragments were subjected to massively parallel sequencing using the HiSeq 2000 platform (Illumina). Candidate somatic mutations were detected through our in-house pipeline (Genomon) as previously described[17].

Detection of mutations from whole-exome sequencing data. Detection of candidate somatic mutations was performed according to previously described algorithms with minor modifications[17]. Briefly, the number of reads containing single-nucleotide variations (SNVs) and indels in both tumor and reference samples was determined using SAMtools[31], and the null hypothesis of equal allele frequencies in tumor and reference samples was tested using the *two-tailed Fisher's exact test*. A variant was adopted as a candidate somatic mutation if it had $P < 0.01$, if it was observed in bidirectional reads (in both plus and minus strands of the reference sequence) and if its allele frequency was less than 0.25 in the corresponding reference sample. For the detection of germline mutations in RAS pathway genes, SNVs and indels having allele frequencies of more than 0.25 (SNVs) and 0.10 (indels) were interrogated for 46 genes, which consisted of known JMML-related RAS pathway genes and genes registered in the pathway databases ('Ras signaling pathway' in BioCarta and 'signaling to RAS' in Reactome[32]). For variant calls in tumor samples for which the paired normal reference was not available, candidate variants in the RAS pathway were detected at an allele frequency of >0.10. Finally, the list of candidate somatic and/or germline mutations was generated by excluding synonymous SNVs and other variants registered in either dbSNP131 or an in-house SNP database constructed from 180 individual samples. All candidates were validated by Sanger sequencing as previously described.

Estimation of tumor content. The tumor content of bone marrow specimens was estimated from the allele frequency of the somatic mutations identified by deep sequencing. For homozygous mutations, as indicated by an allele frequency of >0.75, the tumor content (F_{tumor}) was calculated from the observed frequency (F_{observed}) of the mutation according to the following equation: $F_{\text{tumor}} = 2 \times F_{\text{observed}} - 1$. For heterozygous mutations, the tumor content was calculated by doubling the allele frequency.

Power analysis of whole-exome sequencing. The power of detecting somatic mutations at each nucleotide position in whole-exome sequencing was estimated by Monte-Carlo simulation ($n = 1{,}000$) on the basis of the observed mean depth of coverage for each exon in germline and tumor samples and the observed tumor content for each sample, which were estimated using the allele frequencies of the observed mutations. For the samples with no observed somatic mutations, the average tumor content of the informative samples was employed. Simulations were performed across a total of 192,424 exons.

Copy number analysis in whole-exome sequencing data. To detect copy number lesions at a single-exon level, the mean coverage of each exon normalized by the mean depth of coverage of the entire sample was compared with that of 12 unrelated normal DNA samples. Exons showing normalized coverage greater than 3 s.d. from the mean coverage of the reference samples were called as candidates for copy number alterations. All candidate exons of RAS pathway genes were visually inspected using the Integrative Genomics Viewer[33] and were validated by Sanger sequencing of corresponding putative breakpoint-containing fragments.

Targeted deep sequencing. Deep sequencing of the targeted genes was performed essentially as described in the 'deep sequencing of pooled target exons' section in ref. 17, except that target DNA was not pooled. Briefly, all exons of *PTPN11*, *NF1*, *KRAS*, *NRAS*, *CBL*, *SETBP1*, *JAK3* and *SH3BP1* were PCR amplified with Quick Taq HS DyeMix (TOYOBO) and the PrimeSTAR GXL DNA Polymerase kit (Takara Bio) using primers including the NotI restriction site (**Supplementary Table 3**). The PCR products from an individual sample were combined and purified with the QIAquick PCR Purification kit (Qiagen) for subsequent digestion with NotI (Fermentas). Digested PCR product was purified, concatenated with T4 DNA ligase (Takara Bio) and sonicated to generate fragments with an average size of 150 bp using Covaris. Fragments were processed for sequencing according to a modified Illumina paired-end library protocol, and sequences were read by a HiSeq 2000 instrument using a 100-bp paired-end read protocol.

Variant calls in targeted deep sequencing. Data processing and variant calling were performed with modifications to the protocol described in a previous publication[17]. Each read was aligned to the set of targeted sequences from PCR amplification, with BLAT[34] instead of Burrows-Wheeler Aligner (BWA)[35] used with the -fine option. Mapping information in the .psl format was converted to the .sam format with paired-read information. Of the successfully mapped reads, reads were excluded from further analysis if they mapped to multiple sites, mapped with more than four mismatched bases or had more than ten soft-clipped bases. Next, the Estimation_CRME script was run to eliminate strand-specific errors and exclude PCR-derived errors. A strand-specific mismatch ratio was calculated for each nucleotide variant for both strands using the bases from read cycles 11 to 50 on the next-generation sequencer. By excluding the top five cycles showing the highest mismatch rates, strand-specific mismatch rates were recalculated, and the smaller value between both strands was adopted as a nominal mismatch ratio for that variant. After excluding variants found in dbSNP131 or the in-house SNP database, non-silent variants having a mismatch ratio of greater than 0.05 were called as candidates, unless they had median values of the mismatch ratio at the relevant nucleotide positions in the 92 samples of greater than 0.01, as such variants were likely to be caused by systematic PCR problems. Finally, candidates with mismatch ratios of >0.15 were further validated by Sanger sequencing.

Annotation of the detected mutations. Detected mutations were annotated using ANNOVAR[36]. The positions of the mutations were based on the following RefSeq transcript sequences: NM_002834.3 for *PTPN11*, NM_000267.3 for *NF1*, NM_002524.4 for *NRAS*, NM_004985.3 for *KRAS*, NM_005188.3 for *CBL*, NM_015559.2 for *SETBP1* and NM_000215.3 for *JAK3*. The effect of the mutations on protein function was assessed by SIFT[37], PolyPhen-2 (ref. 38) and MutationTaster[39].

Whole-genome sequencing. Paired tumor-reference DNA samples were sequenced with the HiSeq 2000 platform according to the manufacturer's instructions to obtain 30× read coverage for reference samples and 40× coverage for tumor samples. Obtained FASTQ sequences were aligned to the human reference genome (hg19) using BWA[35] 0.5.8 with default parameters. Alignment of pairs of sequences, at least one of which was not mapped or was considered to have possible mapping problems (with mapping quality of less than 40, insertions or deletions, soft-clipped sequence of more than 10% of the length of the original sequence, irregular paired-read orientation or mate distance of greater than 2,000 bp), was attempted with BLAT[34] using default parameters, except for stepSize = 5 and repMatch = 2,253. Mapping statistics were calculated by counting the bases at each genomic position with SAMtools[31]. For variant calling, variant and reference bases with base quality of >30 were counted in both germline and tumor samples, and the Fisher's

exact test was applied. Variants with P of <0.01 were called. Variants having allele frequency of >0.25 in the germline sample were excluded. Variants found in 12 unrelated germline samples with an allele frequency of >0.01 on average were also excluded owing to the high probability that they represented false positive calls. Copy number estimation was performed by calculating the averaged ratio of read depths in germline and tumor samples in 10,000-base bins. An allele-specific copy number plot was generated by measuring the allele frequency of the tumor sample at the positions in which more than 25% of the allele mismatch was observed in germline samples. For the detection of chromosomal structural variations, soft-clipped sequences that could be mapped to a unique genomic position were selected. Structural variation candidates that had more than four supporting read pairs in total and at least one read pair from each side of the breakpoint were called. Contig sequences were generated by assembling the reads within 200 bp of the breakpoint with CAP3 (ref. 40), and structural variations having the contig sequence that could be aligned to the alternate assembly of the hg19 genome with more than 93% identity were excluded as false positives. Structural variations with read depth of greater than 150 on at least one side of the breakpoint were considered to be mapped to a repeat element and were also excluded. For detection of viruses, unmapped sequences were aligned to the collection of all viral genomes in the RefSeq database using BLAT. A virus was considered to be detected if its genome was covered by mean read coverage of >1.

cDNA sequencing. Total RNA was extracted using the RNeasy Mini kit (Qiagen) and was reverse transcribed with the ThermoScript RT-PCR system (Life Technologies). Target sequences were PCR amplified with the PrimeSTAR GXL DNA Polymerase kit using the primers listed in **Supplementary Table 3** and were sequenced.

Statistical analysis. For comparison of the frequency of mutations or other clinical features between disease groups, categorical variables were analyzed using the Fisher's exact test, and continuous variables were tested using the Mann-Whitney U test. Overall survival and transplantation-free survival were estimated by the Kaplan-Meier method. Hazard ratios for survival with 95% CIs were estimated according to the Cox proportional hazards model, and difference in survival was tested by log-rank test. STATA version 12.0 (StataCorp) was used for all statistical calculations.

31. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
32. Matthews, L. *et al.* Reactome knowledgebase of human biological pathways and processes. *Nucleic Acids Res.* **37**, D619–D622 (2009).
33. Thorvaldsdóttir, H., Robinson, J.T. & Mesirov, J.P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
34. Kent, W.J. BLAT—the BLAST-like alignment tool. *Genome Res.* **12**, 656–664 (2002).
35. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
36. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
37. Kumar, P., Henikoff, S. & Ng, P.C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **4**, 1073–1081 (2009).
38. Adzhubei, I.A. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
39. Schwarz, J.M., Rödelsperger, C., Schuelke, M. & Seelow, D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat. Methods* **7**, 575–576 (2010).
40. Huang, X. & Madan, A. CAP3: A DNA sequence assembly program. *Genome Res.* **9**, 868–877 (1999).

doi:10.1038/ng.2698

To the editor:

Rabbit antithymocyte globulin and cyclosporine as first-line therapy for children with acquired aplastic anemia

Horse antithymocyte globulin (hATG) and cyclosporine have been used as standard therapy for children with acquired aplastic anemia (AA) for whom an HLA-matched family donor is unavailable. However, in 2009, hATG (lymphoglobulin; Genzyme) was withdrawn and replaced by rabbit ATG (rATG; thymoglobulin; Genzyme) in Japan. Many other countries in Europe and Asia are facing the same situation.[1] Marsh et al recently reported outcomes for 35 adult patients with AA who were treated with rATG and cyclosporine as a first-line therapy.[2] Although the hematologic response rate was 40% at 6 months, several patients subsequently achieved late responses. The best response rate was 60% compared with 67% in a matched-pair control group of 105 patients treated with hATG. The overall and transplantation-free survival rates appeared to be significantly inferior with rATG compared with hATG at 68% versus 86% ($P = .009$) and 52% versus 76% ($P = .002$), respectively. These results are comparable to those from a prospective randomized study reported by Scheinberg et al comparing hATG and rATG.[3] Both studies showed the superiority of hATG over rATG.[2,3]

We recently analyzed outcomes for 40 Japanese children (median age, 9 years; range, 1-15) with AA treated using rATG and cyclosporine. The median interval from diagnosis to treatment was 22 days (range, 1-203). The numbers of patients with very severe, severe, and nonsevere disease were 14, 10, and 16, respectively. The ATG dose was 3.5 mg/kg/day for 5 days. The median follow-up time for all patients was 22 months (range, 6-38). At 3 months, no patients had achieved a complete response (CR) and partial response (PR) was seen in only 8 patients (20.0%). At 6 months, the numbers of patients with CR and PR were 2 (5.0%) and 17 (42.5%), respectively. After 6 months, 5 patients with PR at 6 months had achieved CR and 4 patients with no response at 6 months had achieved PR, offering a total best response rate of 57.5%. Two patients relapsed at 16 and 19 months without receiving any second-line treatments. Two patients with no response received a second course of rATG at 13 and 17 months, but neither responded. Sixteen patients underwent hematopoietic stem cell transplantation (HSCT) from alternative donors (HLA-matched unrelated donors, n = 13; HLA-mismatched family donors, n = 3). Two deaths occurred after rATG therapy, but no patients died after HSCT. Causes of death were intracranial hemorrhage at 6 months and acute respiratory distress syndrome at 17 months. The overall 2-year survival rate was 93.8% and the 2-year transplantation-free survival rate was 50.3% (Figure 1).

In our previous prospective studies with hATG, the response rates after 6 months were 68% and 70%, respectively, with no increases in response rates observed after 6 months.[4,5] Our results support the notion that rATG is inferior to hATG for the treatment of AA in children. First-line HSCT from an alternative donor may be justified, considering the excellent outcomes in children who received salvage therapies using alternative donor HSCT.

Yoshiyuki Takahashi
Department of Pediatrics, Nagoya Graduate School of Medicine, Nagoya, Japan

Hideki Muramatsu
Department of Pediatrics, Nagoya Graduate School of Medicine, Nagoya, Japan

Naoki Sakata
Department of Pediatrics, Kinki University School of Medicine, Osaka, Japan

Nobuyuki Hyakuna
Center of Bone Marrow Transplantation, Ryukyu University Hospital, Okinawa, Japan

Kazuko Hamamoto
Department of Pediatrics, Hiroshima Red Cross Hospital, Hiroshima, Japan

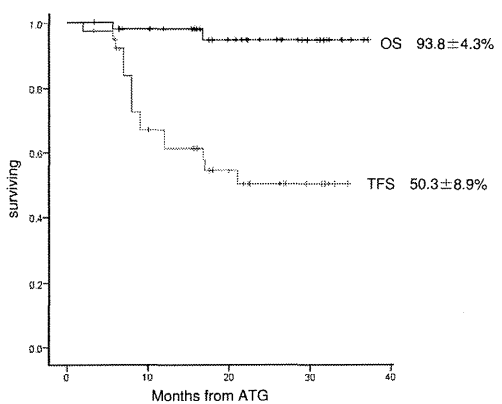


Figure 1. Kaplan-Meier estimates of overall survival (OS) and transplantation-free survival (TFS) in 40 Japanese children with AA. Survival was investigated using Kaplan-Meier methods. OS for all patients with AA after rATG and cyclosporine as first-line therapy included patients who later received HSCT for nonresponse to rATG. In the analysis of TFS for all patients treated with rATG and CSA, transplantation was considered an event.

Ryoji Kobayashi
Department of Pediatrics, Sapporo Hokuyu Hospital, Sapporo, Japan

Etsuro Ito
Department of Pediatrics, Hirosaki University School of Medicine, Hirosaki, Japan

Hiroshi Yagasaki
Department of Pediatrics, School of Medicine, Nihon University, Tokyo, Japan

Akira Ohara
Division of Blood Transfusion, Toho University Omori Hospital, Tokyo, Japan

Akira Kikuchi
Department of Pediatrics, Teikyo University School of Medicine, Tokyo, Japan

Akira Morimoto
Department of Pediatrics, Jichi Medical University School of Medicine, Tochigi, Japan

Hiromasa Yabe
Department of Cell Transplantation and Regenerative Medicine, Tokai University School of Medicine, Isehara, Japan

Kazuko Kudo
Division of Hematology and Oncology, Shizuoka Children's Hospital, Shizuoka, Japan

Ken-ichiro Watanabe
Department of Pediatrics, Graduate School of Medicine, Kyoto University, Kyoto, Japan

Shouichi Ohga
Department of Perinatal and Pediatric Medicine, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

Seiji Kojima
Department of Pediatrics, Nagoya Graduate School of Medicine, Nagoya, Japan

on behalf of the Japan Childhood Aplastic Anemia Study Group

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

Correspondence: Dr Seiji Kojima, Nagoya Graduate School of Medicine, Tsurumai-cho 65, Showa-ku, Nagoya, Ai, Japan 466-8550; e-mail: kojimas@med.nagoya-u.ac.jp.

References

1. Dufour C, Bacigalupo A, Oneto R, et al. Rabbit ATG for aplastic anaemia treatment: a backward step? *Lancet*. 2011;378(9806):1831-1833.
2. Marsh JC, Bacigalupo A, Schrezenmeier H, et al. Prospective study of rabbit antithymocyte globulin and cyclosporine for aplastic anemia from the EBMT Severe Aplastic Anaemia Working Party. *Blood*. 2012;119(23):5391-5396.
3. Scheinberg P, Nunez O, Weinstein B, et al. Horse versus rabbit antithymocyte globulin in acquired aplastic anemia. *N Engl J Med*. 2011;365(5):430-438.
4. Kojima S, Hibi S, Kosaka Y, et al. Immunosuppressive therapy using antithymocyte globulin, cyclosporine, and danazol with or without human granulocyte colony-stimulating factor in children with acquired aplastic anemia. *Blood*. 2000;96(6):2049-2054.
5. Kosaka Y, Yagasaki H, Sano K, et al. Prospective multicenter trial comparing repeated immunosuppressive therapy with stem-cell transplantation from an alternative donor as second-line treatment for children with severe and very severe aplastic anemia. *Blood*. 2008;111(3):1054-1059.

To the editor:

Peripheral blood stem cells versus bone marrow in pediatric unrelated donor stem cell transplantation

The relative benefits and risks of peripheral blood stem cells (PBSCs) versus bone marrow (BM) for allogeneic hematopoietic stem cell transplantation (SCT) are still a matter of highly controversial debates.[1-3] The first randomized study comparing the 2 stem cell sources in unrelated donor SCT recently documented comparable overall and event-free survival, but indicated a higher risk for chronic graft-versus-host disease (GVHD) with PBSCs.[4] Only a few pediatric patients were included in this study even though the long-term sequelae of chronic GVHD are of particular concern in this patient group.

We retrospectively compared the long-term outcome of contemporaneous unrelated donor SCT in 220 children transplanted with BM (n = 102) or PBSCs (n = 118) for hematologic malignancies and reported to the German/Austrian pediatric registry for SCT. All patients had received myeloablative conditioning followed by unmanipulated SCT from HLA-matched unrelated donors. The PBSC and BM groups were comparable with regard to patient and donor age, sex, cytomegalovirus (CMV) serostatus, disease status at transplantation, GVHD prophylaxis, growth factor use, and degree of HLA matching. The groups differed with regard to disease category with slightly more myelodysplastic syndrome patients ($P = .02$) and a higher CD34-cell dose ($P = .001$) in the PBSC group.

Neutrophil and platelet engraftment were achieved significantly faster after PBSC than BM transplantation (Figure 1A-B). In this entirely pediatric cohort, the incidence of clinically relevant grade II-IV acute GVHD (Figure 1C) did not differ. Most importantly, the incidence of chronic GVHD (PBSCs vs BM: 35% vs 33%, respectively; $P = .9$) and extensive chronic GVHD (Figure 1D) proved low and was virtually identical in the 2 groups. With a median follow-up time of 3 years, overall survival (PBSCs vs BM: 50% ± 5% vs 46% ± 6%, respectively; $P = .63$) and event-free survival (PBSCs vs BM: 45% ± 5% vs 44% ± 6%, respectively; $P = .59$) were comparable (Figure 1E-F). In multivariable analysis, taking into account all parameters with $P < .2$ in univariate analysis, the only significant independent risk factor for treatment failure was advanced disease status at the time of transplantation (relative risk = 2.4, 95% confidence interval, 1.5-3.8; $P = .001$). In contrast, stem cell source (PBSCs vs BM) had no effect (relative risk = 1.1, 95% confidence interval, 0.7-1.6; $P = .8$).

Our registry-based analysis provides evidence that in pediatric recipients of HLA-matched unrelated-donor transplantation with consistent antithymocyte globulin (ATG) use during conditioning, transplantation with PBSCs and BM results in comparable clinical outcomes without detectable differences in the risk of acute or, more importantly, chronic GVHD. Consistent with a recent study underscoring the role of ATG for the prevention of acute and chronic GVHD,[5] the use of ATG in 96% of our transplantation procedures compared with only 27% in the above-mentioned randomized study by Anasetti et al[4] might be one of the key factors responsible for the overall low and comparable incidence of

RED CELLS, IRON, AND ERYTHROPOIESIS

Variant ALDH2 is associated with accelerated progression of bone marrow failure in Japanese Fanconi anemia patients

Asuka Hira,[1] Hiromasa Yabe,[2] Kenichi Yoshida,[3] Yusuke Okuno,[3] Yuichi Shiraishi,[4] Kenichi Chiba,[4] Hiroko Tanaka,[5] Satoru Miyano,[4,5] Jun Nakamura,[6] Seiji Kojima,[7] Seishi Ogawa,[3,8] Keitaro Matsuo,[9] Minoru Takata,[1] and Miharu Yabe[2]

[1]Laboratory of DNA Damage Signaling, Department of Late Effects Studies, Radiation Biology Center, Kyoto University, Kyoto, Japan; [2]Department of Cell Transplantation and Regenerative Medicine, Tokai University School of Medicine, Isehara, Japan; [3]Cancer Genomics Project, Graduate School of Medicine, [4]Laboratory of DNA Information Analysis, and [5]Laboratory of Sequence Analysis, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan; [6]Department of Environmental Sciences and Engineering, University of North Carolina at Chapel Hill, Chapel Hill, NC; [7]Department of Pediatrics, Nagoya University Graduate School of Medicine, Nagoya, Japan; [8]Department of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto University, Kyoto, Japan; and [9]Department of Preventive Medicine, Kyushu University Faculty of Medical Sciences, Fukuoka, Japan

Key Points

- We found the defective ALDH2 variant is associated with accelerated progression of BMF in Japanese FA patients.
- The data support the view that aldehydes are an important source of genotoxicity in the human hematopoietic system.

Fanconi anemia (FA) is a severe hereditary disorder with defective DNA damage response and repair. It is characterized by phenotypes including progressive bone marrow failure (BMF), developmental abnormalities, and increased occurrence of leukemia and cancer. Recent studies in mice have suggested that the FA proteins might counteract aldehyde-induced genotoxicity in hematopoietic stem cells. Nearly half of the Japanese population carries a dominant-negative allele (rs671) of the aldehyde-catalyzing enzyme *ALDH2* (acetaldehyde dehydrogenase 2), providing an opportunity to test this hypothesis in humans. We examined 64 Japanese FA patients, and found that the *ALDH2* variant is associated with accelerated progression of BMF, while birth weight or the number of physical abnormalities was not affected. Moreover, malformations at some specific anatomic locations were observed more frequently in *ALDH2*-deficient patients. Our current data indicate that the level of ALDH2 activity impacts pathogenesis in FA, suggesting the possibility of a novel therapeutic approach. (*Blood*. 2013;122(18):3206-3209)

Introduction

Fanconi anemia (FA) is a genomic instability disorder with phenotypes including progressive bone marrow failure (BMF), developmental abnormalities, and increased occurrence of leukemia and cancer.[1] To date, 16 genes have been implicated in FA, and their products form a common DNA repair network ("FA pathway").[2,3] Because FA cells are hypersensitive to DNA interstrand crosslinks (ICLs), the FA pathway has been considered to be involved in the repair of ICLs.[2,3] However, it remains unclear what type of endogenous DNA damage is repaired through the FA pathway. Recent studies have suggested that FA cells are also sensitive to aldehydes,[4] which may create DNA adducts including ICLs or DNA-protein crosslinks. Furthermore, double knockout mice deficient in *Fancd2* and *Aldh2*, but neither of the single mutant mice, display an accelerated development of leukemia and BMF.[5,6] On the other hand, Fanc-deficient mice in general do not fully recapitulate the human FA phenotype, including overt BMF.[7] Thus, the role of aldehydes in the pathogenesis of human FA is still uncertain.

ALDH2 deficiency resulting from a Glu504Lys substitution (rs671, hereinafter referred to as the A allele) is highly prevalent in East Asian populations. The A allele (Lys504) acts as a dominant negative, since the variant form can suppress the activity of the Glu504 form (G allele) in GA heterozygotes by the formation of heterotetramers.[8] Individuals with the A variant experience flushing when drinking alcohol, and have an elevated risk of esophageal cancer with habitual drinking.[9] Because the frequency of the A allele is close to 50% in the Japanese population at large, some Japanese FA patients are expected to be deficient in *ALDH2*. We thus set out to determine the *ALDH2* status in a collection of Japanese FA patients.

Study design

The onset of BMF was defined according to the criteria used in the International Fanconi Anemia Registry (IFAR) study.[10] Criteria for diagnosis of aplastic anemia and other conditions are described in supplemental Methods (available on the *Blood* Web site). We observed physical abnormalities characteristic of FA, including skin abnormalities (hyperpigmentation and café au lait spots), low birth weight, growth defects, and malformations affecting skeletal systems and deep organs. Extensive malformation was defined as the involvement of at least 3 sites including at least 1 deep organ.[11] Mutation analysis of *FANCA/FANCC/FANCG* genes,[12] *ALDH2* genotyping,[13] multiplex ligation-mediated probe amplification (MLPA) test for *FANCA* (Falco), and whole-exome sequencing (WES)[14] were done as previously described. Details are provided as supplemental Methods. Development of BMF or acute myeloid leukemia (AML)/myelodysplasia (MDS) was analyzed by the Kaplan-Meier method or the cumulative incidence method,[15,16] respectively, since competing events (eg, death and stem cell transplantation [SCT]) existed in AML/MDS but not in BMF. This study was approved by the Research Ethics Committee of the Tokai University Hospital and Kyoto University. We obtained family informed consent from all subjects involved in this work in accordance with the Declaration of Helsinki.

Submitted June 13, 2013; accepted August 27, 2013. Prepublished online as *Blood* First Edition paper, September 13, 2013; DOI 10.1182/blood-2013-06-507962.

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.

Table 1. Summary of genotypes and clinical characteristics of the patients studied

	Total	ALDH2 genotype		
		GG	GA	AA
No. of cases	64	36	25	3
Mutated FA gene*				
FANCA	39	26†	11	2
FANCG	15	7	8	—
FANCI	2	—	2	—
FANCM	1	—	1	—
FANCP	2	—	1	1
Unknown	5	3	2	—
Disease				
Aplastic anemia	2	2	—	—
Severe aplastic anemia	40	21	19	—
MDS/AML	22	13	6	3‡
Tongue cancer	2	1	1	—
Median months of onset (range)				
BMF	52 (0-297)	72 (27-297)	28 (7-87)	0 (0-7)
MDS/AML	118 (4-384)	156 (61-384)	85 (41-192)	4 (4-12)
No. of cases with SCT (%)	58 (91)	33 (92)	23 (92)	2 (67)
Median months at SCT (range)	118 (12-448)	130 (52-448)	86 (28-248)	25 (13-36)

–, no case was found.

*Mutations found in the patients were listed in supplemental Table 1. Some of them were presumptive because their functional significance has not been determined.

†Somatic mosaicism due to reversion was confirmed in 2 cases and suspected in 1 case.

‡In these cases, onset of severe aplastic anemia and MDS was essentially simultaneous.

Results and discussion

All of the patients in this study (n = 64; supplemental Table 1) were referred to the Tokai University Hospital because of pancytopenia, in some cases with MDS or leukemia. The clinical diagnosis of FA was made based on clinical presentation and diepoxybutane (DEB)-induced chromosome fragility tests in peripheral blood lymphocytes,[17] except for 3 cases in which the DEB test was negative due to *FANCA* reversion mosaicism (supplemental Tables 1-2). Most of the patients underwent allogeneic SCT, indicating that our patients probably represent an FA population with relatively severe hematologic symptoms.

To determine which FA gene was mutated in each of these patients, we applied combinations of polymerase chain reaction–based methods (n = 26), the MLPA test for *FANCA* mutations (n = 44), and WES (n = 29). In our WES analysis, >90% of the 50-Mb target sequences were analyzed by >10 independent reads (data not shown). Fifty-nine patients were found to have a mutation in FA genes in at least 1 allele, but 5 of them were mutation-free in the known 16 FA genes, even after WES (Table 1; supplemental Table 1). These unclassified cases might be caused by large deletions or intronic mutations that are difficult to detect with these methods,[18] or possibly mutations in a novel FA gene.

We determined the *ALDH2* genotype in our series of 64 patients (Table 1; supplemental Table 1). The distribution of the *ALDH2* variant alleles appeared not significantly different from the reported allele frequencies in the healthy Japanese population.[13] The occurrence of leukemia and/or MDS was also not significantly different between patients with GA and GG genotypes. Strikingly, however, we found that progression of BMF was accelerated in heterozygous carriers of the variant A allele compared with homozygous GG patients (Figure 1A-B). Moreover, the 3 individuals carrying AA alleles developed MDS with BMF at a very young age (Figure 1A-B). None of these 3 patients belonged to FA-D1 or FA-N, the FA subgroups with severe symptoms.[19,20] Patient number 3 had biallelic frameshift mutations (S115AfsX11) in *FANCP/SLX4*. By contrast, of the FA-P patients that have previously been reported, none have displayed particularly severe symptoms.[21-23]

FA is a heterogeneous disorder, and our cohort of patients is quite heterogeneous in terms of complementation groups and types of mutations (Table 1). To reduce some of the variability, we selected only the *FANCA* patients having nonsense, frameshift, or large deletion mutations identified at both alleles, (n = 12; supplemental Table 1), and repeated the analysis. A patient with probable *FANCA* reversion (patient number 55) was excluded. In this subset of patients, a highly significant statistical difference was reproduced in BMF progression (Figure 1C) but not in AML/MDS development (data not shown).

We could not detect any significant difference in terms of the percentage of birth weight (Figure 1D) or number of physical abnormalities (Figure 1E) that correlated with the *ALDH2* genotypes. However, a significant difference was observed in the incidence of each class of malformations in the case of radial, cardiovascular, skeletal, or kidney anomalies, and in the incidence of extensive malformation (Figure 1F).

In conclusion, our current data indicate that endogenous aldehydes are an important source of genotoxicity in the human hematopoietic system, and the FA pathway counteracts them. If the FA pathway is compromised, hematopoietic stem cells (HSCs) likely accumulate aldehyde-induced DNA damage, resulting in BMF due to p53/p21-mediated cell death or senescence.[6,24] Consistent with this model, a recent study showed that the HSCs in *aldh2/fancd2* double knockout mice accumulate more DNA damage than HSCs in either of the single knockout mice.[6] Because some *ALDH2*-proficient FA patients developed BMF early, other modifier genes or environmental factors might affect levels of aldehydes or other genotoxic substances. Interestingly, our data predict that Japanese FA patients in general develop BMF at an earlier age compared with patients of other ethnic origins. We need to establish a Japanese FA registry similar to IFAR to test whether this is true or not. Finally, it seems worth considering *ALDH2* agonists such as Alda-1 as protective drugs against BMF in FA patients. Alda-1 can stimulate the enzymatic activity of both the normal and variant ALDH2,[25] suggesting that Alda-1 or a similar drug could be beneficial even for *ALDH2*-proficient FA cases.

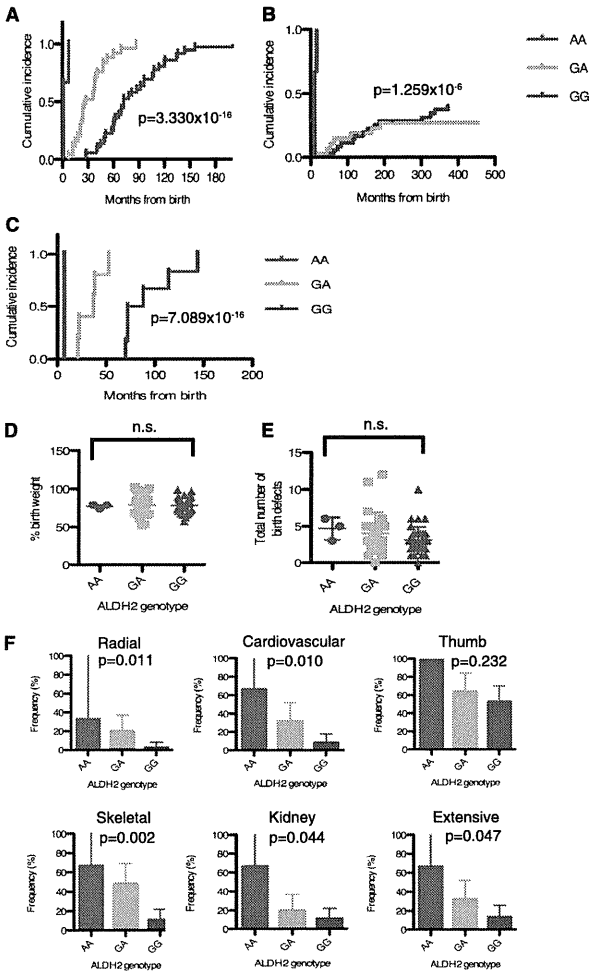


Figure 1. Effects of the ALDH2 deficiency on Japanese FA patients. (A-B) Cumulative incidence of BMF (A) or MDS/AML (B) were analyzed in 64 FA subjects. Numbers of AA, GA, and GG patients were 3, 25, and 36, respectively. (C) Cumulative incidence of BMF was analyzed in patients with confirmed biallelic *FANCA* mutations having protein truncations and/or large deletions (n = 12). Numbers of AA, GA, and GG patients were 1, 5, and 6, respectively. *P* values shown were calculated by the Gray test. In panel A, *P* values between genotypes were 8.625×10^{-7} (GG vs GA), 2.107×10^{-10} (GG vs AA), 1.259×10^{-6} (GA vs AA), respectively. In (B), the difference between GG and GA subjects was not significant ($P = .4564479$), whereas other statistical comparisons were highly significant (GG vs AA, 2.911×10^{-10}; GA vs AA, 8.813×10^{-8}). In panel C, the *P* values between GG and GA, GG and AA, or GA and AA were calculated as 0.001228433, 0.01430588, 0.02534732, respectively. (D) Percentage of birth weight or (E) total number of physical abnormalities (shown in supplemental Table 1) in 64 FA patients with 3 *ALDH2* genotypes. Birth weight was normalized to mean weight at gestational age in Japan. Mean and SEM are indicated. Birth weight records were missing for 3 patients (supplemental Table 1). There was no significant difference between the *ALDH2* genotypes (Kruskal-Wallis test). (F) Frequency (percentage) of cardiovascular, radial, thumb, skeletal, kidney, and extensive malformations in each *ALDH2* genotype. *P* values were calculated by the Cochran-Armitage test for trend, which detects statistical significance of effects across the genotypes. The error bars represent 95% confidence intervals.

Acknowledgments

The authors thank the individual patients and families in the study who made this work possible, Dr K. J. Patel (University of Cambridge) for communicating unpublished results, Dr James Hejna (Graduate School of Biostudies, Kyoto University) for critical reading of the manuscript and English editing, Mr Naoya Suzuki and Drs Akira Niwa and Megumu Saito (CiRA, Kyoto University) for discussion, and Ms Fumiko Tsuchida, Emi Uchida, Sumiyo Ariga, Chinatsu Ohki, and Mao Hisano for expert technical assistance.

This work was supported by grants from the Ministry of Health, Labor and Welfare, and grants from the Ministry of Education, Culture, Sports, Science and Technology (MEXT).

Authorship

Contribution: M.Y. and H.Y. examined DEB-induced chromosome aberrations, carried out MLPA testing, and analyzed clinical records; K.Y., Y.O., Y.S., K.C., H.T., S.M., S.K., and S.O. performed WES and analyzed sequence data; A.H. validated exome data and carried out genotyping; A.H., M.Y., H.Y., K.M., J.N., and M.T. analyzed data; and M.Y., M.T., and K.M. wrote the paper.

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

Correspondence: Minoru Takata, Laboratory of DNA Damage Signaling, Department of Late Effects Studies, Radiation Biology Center, Kyoto University, Kyoto 606-8501, Japan; e-mail: mtakata@house.rbc.kyoto-u.ac.jp; and Miharu Yabe, Department of Cell Transplantation and Regenerative Medicine, Tokai University School of Medicine, Isehara 259-1193, Japan; e-mail: miharu@is.icc.u-tokai.ac.jp.

References

1. Auerbach AD. Fanconi anemia and its diagnosis. *Mutat Res.* 2009;668(1-2):4-10.
2. Kim H, D'Andrea AD. Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway. *Genes Dev.* 2012;26(13):1393-1408.
3. Kottemann MC, Smogorzewska A. Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. *Nature.* 2013;493(7432):356-363.
4. Ridpath JR, Nakamura A, Tano K, et al. Cells deficient in the FANC/BRCA pathway are hypersensitive to plasma levels of formaldehyde. *Cancer Res.* 2007;67(23):11117-11122.
5. Langevin F, Crossan GP, Rosado IV, Arends MJ, Patel KJ. Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. *Nature.* 2011;475(7354):53-58.
6. Garaycoechea JI, Crossan GP, Langevin F, Daly M, Arends MJ, Patel KJ. Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function. *Nature.* 2012;489(7417):571-575.
7. Parmar K, D'Andrea A, Niedernhofer LJ. Mouse models of Fanconi anemia. *Mutat Res.* 2009; 668(1-2):133-140.
8. Crabb DW, Edenberg HJ, Bosron WF, Li TK. Genotypes for aldehyde dehydrogenase deficiency and alcohol sensitivity. The inactive ALDH2(2) allele is dominant. *J Clin Invest.* 1989; 83(1):314-316.
9. Matsuo K, Hamajima N, Shinoda M, et al. Gene-environment interaction between an aldehyde dehydrogenase-2 (ALDH2) polymorphism and alcohol consumption for the risk of esophageal cancer. *Carcinogenesis.* 2001;22(6):913-916.
10. Butturini A, Gale RP, Verlander PC, Adler-Brecher B, Gillio AP, Auerbach AD. Hematologic abnormalities in Fanconi anemia: an International Fanconi Anemia Registry study. *Blood.* 1994; 84(5):1650-1655.
11. Guardiola P, Pasquini R, Dokal I, et al. Outcome of 69 allogeneic stem cell transplantations for Fanconi anemia using HLA-matched unrelated donors: a study on behalf of the European Group for Blood and Marrow Transplantation. *Blood.* 2000;95(2):422-429.
12. Tachibana A, Kato T, Ejima Y, et al. The FANCA gene in Japanese Fanconi anemia: reports of eight novel mutations and analysis of sequence variability. *Hum Mutat.* 1999;13(3):237-244.
13. Matsuo K, Wakai K, Hirose K, Ito H, Saito T, Tajima K. Alcohol dehydrogenase 2 His47Arg polymorphism influences drinking habit independently of aldehyde dehydrogenase 2 Glu487Lys polymorphism: analysis of 2,299 Japanese subjects. *Cancer Epidemiol Biomarkers Prev.* 2006;15(5):1009-1013.
14. Kunishima S, Okuno Y, Yoshida K, et al. ACTN1 mutations cause congenital macrothrombocytopenia. *Am J Hum Genet.* 2013;92(3):431-438.
15. Gray RJ. A class of K-sample tests for comparing the cumulative incidence of a competing risk. *Ann Stat.* 1988;16(3):1141-1154.
16. Klein JP, Rizzo JD, Zhang MJ, Keiding N. Statistical methods for the analysis and presentation of the results of bone marrow transplants. Part I: unadjusted analysis. *Bone Marrow Transplant.* 2001;28(10):909-915.
17. Auerbach AD, Rogatko A, Schroeder-Kurth TM. International Fanconi Anemia Registry: relation of clinical symptoms to diepoxybutane sensitivity. *Blood.* 1989;73(2):391-396.
18. Chandrasekharappa SC, Lach FP, Kimble DC, et al; NISC Comparative Sequencing Program. Massively parallel sequencing, aCGH, and RNA-Seq technologies provide a comprehensive molecular diagnosis of Fanconi anemia. *Blood.* 2013;121(22):e138-e148.
19. Wagner JE, Tolar J, Levran O, et al. Germline mutations in BRCA2: shared genetic susceptibility to breast cancer, early onset leukemia, and Fanconi anemia. *Blood.* 2004;103(8):3226-3229.
20. Reid S, Schindler D, Hanenberg H, et al. Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer. *Nat Genet.* 2007;39(2):162-164.
21. Kim Y, Lach FP, Desetty R, Hanenberg H, Auerbach AD, Smogorzewska A. Mutations of the SLX4 gene in Fanconi anemia. *Nat Genet.* 2011; 43(2):142-146.
22. Stoepker C, Hain K, Schuster B, et al. SLX4, a coordinator of structure-specific endonucleases, is mutated in a new Fanconi anemia subtype. *Nat Genet.* 2011;43(2):138-141.
23. Schuster B, Knies K, Stoepker C, et al. Whole exome sequencing reveals uncommon mutations in the recently identified Fanconi anemia gene SLX4/FANCP. *Hum Mutat.* 2013;34(1):93-96.
24. Ceccaldi R, Parmar K, Mouly E, et al. Bone marrow failure in Fanconi anemia is triggered by an exacerbated p53/p21 DNA damage response that impairs hematopoietic stem and progenitor cells. *Cell Stem Cell.* 2012;11(1):36-49.
25. Chen CH, Budas GR, Churchill EN, Disatnik MH, Hurley TD, Mochly-Rosen D. Activation of aldehyde dehydrogenase-2 reduces ischemic damage to the heart. *Science.* 2008;321(5895): 1493-1495.

Regular Article

MYELOID NEOPLASIA

Clonal selection in xenografted TAM recapitulates the evolutionary process of myeloid leukemia in Down syndrome

Satoshi Saida,[1] Ken-ichiro Watanabe,[1] Aiko Sato-Otsubo,[2] Kiminori Terui,[3] Kenichi Yoshida,[2] Yusuke Okuno,[2] Tsutomu Toki,[3] RuNan Wang,[3] Yuichi Shiraishi,[4] Satoru Miyano,[4] Itaru Kato,[1] Tatsuya Morishima,[1] Hisanori Fujino,[1] Katsutsugu Umeda,[1] Hidefumi Hiramatsu,[1] Souichi Adachi,[5] Etsuro Ito,[3] Seishi Ogawa,[2] Mamoru Ito,[6] Tatsutoshi Nakahata,[7] and Toshio Heike[1]

[1]Department of Pediatrics, Graduate School of Medicine, Kyoto University, Kyoto, Japan; [2]Cancer Genomics Project, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; [3]Department of Pediatrics, Graduate School of Medicine, Hirosaki University, Hirosaki, Japan; [4]Laboratory of DNA Information Analysis, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan; [5]Department of Human Health Sciences, Graduate School of Medicine, Kyoto University, Kyoto, Japan; [6]Laboratory Animal Research Department, Central Institute for Experimental Animals, Kawasaki, Japan; and [7]Department of Clinical Application, Center for iPS Cell Research and Application, Kyoto University, Kyoto, Japan

Key Points

- Genetically heterogeneous subclones with varying leukemia-initiating potential exist in neonatal transient abnormal myelopoiesis.
- This novel xenograft model of transient abnormal myelopoiesis may provide unique insight into the evolutionary process of leukemia.

Transient abnormal myelopoiesis (TAM) is a clonal preleukemic disorder that progresses to myeloid leukemia of Down syndrome (ML-DS) through the accumulation of genetic alterations. To investigate the mechanism of leukemogenesis in this disorder, a xenograft model of TAM was established using NOD/Shi-*scid*, interleukin (IL)-2Rγ^{null} mice. Serial engraftment after transplantation of cells from a TAM patient who developed ML-DS a year later demonstrated their self-renewal capacity. A *GATA1* mutation and no copy number alterations (CNAs) were detected in the primary patient sample by conventional genomic sequencing and CNA profiling. However, in serial transplantations, engrafted TAM-derived cells showed the emergence of divergent subclones with another *GATA1* mutation and various CNAs, including a 16q deletion and 1q gain, which are clinically associated with ML-DS. Detailed genomic analysis identified minor subclones with a 16q deletion or this distinct *GATA1* mutation in the primary patient sample. These results suggest that genetically heterogeneous subclones with varying leukemia-initiating potential already exist in the neonatal TAM phase, and ML-DS may develop from a pool of such minor clones through clonal selection. Our xenograft model of TAM may provide unique insight into the evolutionary process of leukemia.
(*Blood*. 2013;121(21):4377-4387)

Introduction

Neonates with Down syndrome (DS) are at high risk of developing a unique hematologic disorder referred to as transient abnormal myelopoiesis (TAM), transient myeloproliferative disorder, or transient leukemia. In most cases, TAM resolves spontaneously within 3 months.[1,2] However, after spontaneous remission, 20% of TAM patients develop myelodysplastic syndrome and acute megakaryocytic leukemia referred to as myeloid leukemia of DS (ML-DS) within 4 years.[3,4] Blast cells in most patients with TAM and ML-DS have mutations in exon 2 of the gene coding for the transcription factor *GATA1*,[5-8] which is essential for the normal development of erythroid and megakaryocytic cells.[9,10] Although blast cells in most TAM and ML-DS patients share the identical *GATA1* mutation, recurrent additional cytogenetic abnormalities are commonly observed during disease progression.[2,5,11,12] In fact, a ML-DS case derived from a minor clone with a distinct *GATA1* mutation in the TAM phase was previously reported by our group.[13] These clinical findings suggest that although most TAM cells disappear in the early neonatal phase, a few clones persist during apparent remission to develop ML-DS later. Because only one fifth of TAM cases progress to ML-DS, additional genetic events besides *GATA1* mutation are likely to be involved in the progression of TAM to ML-DS.[14] As mentioned above, the development of ML-DS is significantly correlated with karyotypic abnormalities such as duplication (dup)(1q), deletion (del)(6q), del(7p), dup(7q), +8, +11, and del(16q),[2,11,12] which are rarely observed in the TAM phase. These clinical findings have led many physicians to consider TAM as preleukemia and the progression of TAM to ML-DS as an attractive model to investigate multistep leukemogenesis.

Animal models have contributed to our understanding of the pathogenesis of TAM/ML-DS and other leukemias.[15-21] Mice models in which primary human leukemic cells were transplanted into immunodeficient hosts provided significant clues to advance our understanding of the pathogenesis of human leukemia.[19-22] However, xenograft models of primary patient samples from the preleukemic phase have been rarely reported, and the TAM xenograft model would be an attractive method to investigate leukemogenesis.

We previously described the development of novel immunodeficient NOD/Shi-*scid*, interleukin (IL)-2Rγ^{null} (NOG) mice with a superior capacity for the engraftment of human hematopoietic and neoplastic cells.[23-26] In contrast to a previous study in which TAM cells showed a limited ability to expand in immunodeficient mice,[27] we established a xenograft model where TAM cells were transplanted into NOG mice to recapitulate the pathophysiology of TAM/ML-DS. This xenograft model in combination with high-throughput genomic technology was used to show that genetically heterogeneous minor subclones with leukemia-initiating potential already exist in the neonatal TAM phase and could serve as initiating clones evolving to ML-DS in a patient. Our TAM xenograft model may be of value to gain insight into the evolutionary process of leukemia.

Submitted December 18, 2012; accepted February 4, 2013. Prepublished online as *Blood* First Edition paper, March 12, 2013; DOI 10.1182/blood-2012-12-474387.

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.

Materials and methods

Patients and sample collection

Peripheral blood (PB) samples were obtained from patients diagnosed with TAM associated with DS in acute and complete remission phases. Mononuclear cells were separated by Ficoll-Hypaque (Pharmacia, Uppsala, Sweden) density gradient centrifugation, as previously described.[23] Informed consent was obtained from the patients' parents in accordance with the Declaration of Helsinki, and the research was approved by the institutional ethics committee of Kyoto University Hospital.

Mice

NOG mice were developed at the Central Institute of Experimental Animals (Kawasaki, Japan) as previously described[28] and were maintained in our pathogen-free facility and cared for in accordance with the institutional guidelines for animal welfare.

Primary and serial xenogeneic transplantation into NOG mice

Xenotransplantation and analysis of TAM cells were performed using a previously reported method with some modifications.[26] In brief, PB mononuclear cells (PBMCs) obtained from TAM patients ($1–3 \times 10^6$ cells) were injected into 2.4 Gy–irradiated 8- to 12-week-old NOG mice through the tail vein. To screen for the proliferation of TAM-derived cells, bone marrow (BM) cells were aspirated from the tibia every 4 weeks. Engraftment was defined as >1% of cells staining positive for human CD7 (hCD7), hCD33, hCD41a, hCD45, and hCD117 at 12 weeks after transplantation. For serial transplantation, recipient BM cells were collected 12 to 18 weeks after transplantation; the equivalent of 1×10^6 hCD45$^+$ cells was intravenously transplanted into new mice. For a detailed determination of chromosomal and genetic alterations in TAM-derived cells, serial transplantation experiments using preserved PBMC samples were performed.

Flow cytometric analysis of transplanted TAM-derived cells

For analysis of TAM-derived cells in murine BM, mice were euthanized, and the BM was removed and mechanically dispersed. Mononuclear cells were purified from the BM and stained with antibodies. Dead cells were excluded according to 4′,6-diamidino-2-phenylindole staining. Blast cells were identified by classical CD45/SSC blast gating.[29] See supplemental Methods on the *Blood* Web site for details.

Human cell sorting

Human cell isolation was performed according to a previously described method with some modifications.[23,24] See supplemental Methods for details.

Colony assay

Leukemic colony formation was assessed according to a previously described method with some modifications.[30] See supplemental Methods for details.

GATA1 genomic sequencing analysis

The *GATA1* gene was amplified using polymerase chain reaction (PCR) as previously described[8] and sequenced by an ABI 3130xl Genetic analyzer (Applied Biosystems, Foster City, CA).

DNA copy number analysis

DNA copy number analysis was performed using GeneChip Human Mapping 250K Nsp arrays (Affymetrix, Inc., Santa Clara, CA) according to the manufacturer's standard protocols. Genomic copy numbers including allele-specific copy numbers were calculated using CNAG/AsCNAR software (http://www.genome.umin.jp), and genomic DNA obtained from PB of patients in the remission phase was used as a control. Copy number abnormalities and other allelic imbalances were detected using a hidden Marcov model–based algorithm.

Statistical analysis

Data are presented as the mean ± standard deviation. The 2-sided P value was determined by testing the null hypothesis that the 2 population medians are equal. P values <0.05 were considered to be significant.

Results

Establishment of a TAM xenograft model using NOG mice

To determine whether NOG mice provide a TAM xenograft model, Ficoll-purified PB samples from 11 TAM patients were transplanted into irradiated NOG mice. Patient characteristics are shown in Table 1. Patients' ages at sample collection, percentage of blast cells, number of cells injected, and number of engrafted recipients for each PB sample are shown in supplemental Table 1. Of 11 patient samples, 3 (patients 1, 2, and 9) were engrafted successfully in the recipient mice. Engraftment was maintained ≥12 weeks in all cases (Figure 1A). The spleen and liver of the recipients were also infiltrated with hCD45$^+$ blast cells (data not shown). These TAM-derived cells were morphologically similar to the primary TAM cells obtained from the patients (Figure 1B). Flow cytometric analysis of surface antigens detected the expression of CD117, CD34, CD33, and CD41a on hCD45$^+$ cells, which was consistent with the pattern observed in primary cells of TAM patients (Figure 1C). The presence of the same *GATA1* mutation was confirmed in the primary TAM cells and the engrafted cells in NOG mice (Figure 1D; supplemental Table 1). Chromosomal analysis of engrafted cells showed no abnormalities other than trisomy 21 (Figure 1E). These TAM-derived cells were detectable in the recipient's BM for >24 weeks (data not shown).

NOG mice can support self-renewal of TAM-derived cells

To examine the self-renewal capacity of TAM-derived cells, we performed serial transplantation of engrafted cells in the BM of recipient mice. Only the TAM-derived cells from patient 1 were successfully engrafted into the secondary (2°) and tertiary (3°) recipients. The morphology and surface antigen expression of these engrafted cells remained unchanged throughout the serial transplantation (Figure 2A-B). Interestingly, the TAM-derived cells