Table 1. Clinical characteristics of 11 TAM patients

Patient no.	Gender	Period of gestation, wk	Weight at birth, g	PB at diagnosis of TAM				Cytogenetics	*GATA1* mutation	Treatment		Clinical outcome	Onset of ML-DS (m of age)	Follow-up interval
				WBC, $\times 10^3$/μL	Blast, %	Hb, g/dL	Plt, $\times 10^3$/μL	International System for Human Cytogenetic Nomenclature (2009)		Exchange transfusion	Low-dose Ara-C			
1	F	36	3050	159	91	9.1	247	47,XX,+21 [20]	c.38_39delAG	No	Yes	Alive	Yes (14)	27
2	M	37	1868	45.0	65	19.0	80	47,XY,+21 [20]	c.49C>T	No	Yes	Alive	No	24
3	F	39	3102	40.3	37	15.1	304	47,XX,+21 [20]	c.59_174del116	No	No	Alive	No	23
4	M	37	2780	15.6	24	17.0	50	47,XY,+21 [20]	c.163_169del	No	No	Alive	No	22
5	M	39	3052	60.9	48	20.7	258	47,XY,+21 [19]	c.37G>T	No	No	Alive	No	21
6	F	37	2050	13.6	12	20.9	291	47,XX,+21 [20]	N/A	No	No	Alive	No	19
7	M	38	2694	280	87	13.6	26	47,XY,+21 [20]	c.186C>G	Yes	Yes	DOD	No	1
8	M	35	2070	174	80	19.2	117	47,XY,+21 [20]	c.-19-1G>A, c.1A>G	Yes	No	Alive	No	17
9	M	39	3380	56.2	65	20.2	156	47,XY,+21 [20]	c.35C>G	No	Yes	Alive	No	14
10	M	36	2131	199	84	12.2	73	47,XY,+21 [20]	c.19_20insCCTGA	Yes	Yes	Alive	No	14
11	F	33	2032	254	90	14.3	178	47,XX,+21 [20]	c.-19-62_-56delinsA	Yes	Yes	Alive	No	10

Brackets under International System for Human Cytogenetic Nomenclature indicate the number of analyzed cells in metaphase.
Ara-C, cytosine arabinoside; DOD, died of disease; Hb, hemoglobin; N/A, not assessed; Plt, platelet; WBC, white blood cell.

expanded rapidly in the 3° recipients (Figure 2C). The colony-forming ability of the engrafted cells also increased in subsequent generations (Figure 2D). These cells could be grown by serial transplantation for >1 year and ≥8° recipients, indicating that some TAM clones had long-term self-renewal capacity, a characteristic of leukemia. Indeed, patient 1 developed ML-DS at the age of 1 year, whereas the other patients did not (Table 1).

TAM-NOG xenograft model recapitulates leukemic evolution from TAM

Additional chromosomal alterations are frequently observed in ML-DS in comparison with TAM, suggesting that these alterations in genomic structure could be related to the evolution of ML-DS from TAM.[2,11,12] Therefore, we first investigated whether the serially engrafted TAM-derived cells (from patient 1) had DNA copy number alterations (CNAs) using Affymetrix GeneChip Mapping 250K arrays. Primary samples from patient 1 had no CNAs other than the gain of chromosome 21. However, the TAM-derived cells in the 1° recipients showed heterozygous deletion of 16q22 and 16q24 (Figure 3). To determine whether these deletions were present in the same cell, we calculated the signal intensities of each deletion using array data. Nearly 100% of TAM-derived cells harbored each deletion, indicating that these 2 deletions exist in a single TAM-derived cell. Although 2° recipients showed the same CNAs, 3° recipients showed additional CNAs, namely the gain of the entire chromosome 1q (Figure 3; supplemental Figure 1A). Interestingly, the 1q gain was not detected in the 4° to 7° recipients, whereas deletions of 16q22 and 16q24 were present (Figure 3). In this series of transplantations, the original *GATA1* mutation found in the primary patient sample (patient 1) remained unchanged (supplemental Figure 1B).

Gain of 1q and deletions in 16q are recurrent chromosomal abnormalities in ML-DS.[11,12,31] The result of G-band karyotyping of TAM-derived cells in 3° recipients was 47,XX,+1, der (1;15)(q10;q10),+21 in 20/20 metaphase cells (supplemental Figure 1C), confirming genomic structural change, which is a hallmark of ML-DS. These data suggest that leukemic evolution of TAM-derived cells was observed in our NOG mouse model.

Genetically heterogeneous subclones with varying repopulating capacity expanded in the TAM-NOG xenograft model

To examine the kinetics of the leukemic evolution of TAM cells, another 2 sets of serial transplantations were performed using the preserved patient 1 sample (Figure 4A). Four of 5 mice in the second group (m2-1–m2-5) and 5 of 11 mice in the third group (m3-1–m3-11) harbored TAM cells from the patient. Of the total of 9 engrafted mice, 2 had the same CNAs detected in the first series of serial transplantations: deletion of 16q22 and 16q24 (m3-5 and m3-8; Figure 3). Moreover, 2 combinations of new CNAs were detected in the 1° recipients: deletion of 9q22 +12p12 (m3-4 and m3-7) and gain of 1q25.2-1q44 (m3-11). No CNAs other than the gain of chromosome 21 were detected in the other recipients (m2-1, m2-2, m2-4, and m2-5).

Each 1° engrafted mouse was subjected to 2° transplantation, and 5 of 9 series (m2-5, m3-4, m3-7, m3-8, and m3-11) successfully gave rise to the xenografts in the 2° recipients. It is noteworthy that the TAM-derived cells of the 2° recipients in 2 of the 3 analyzed series (m2-5 and m3-4) acquired additional CNAs, whereas the CNAs in 2 descendent 2° recipients of m3-8 remained unchanged. The additional CNA of gain of 1q was detected in the 2° recipients of m3-4, similar to that observed in the 3° recipient in Figure 3. Although

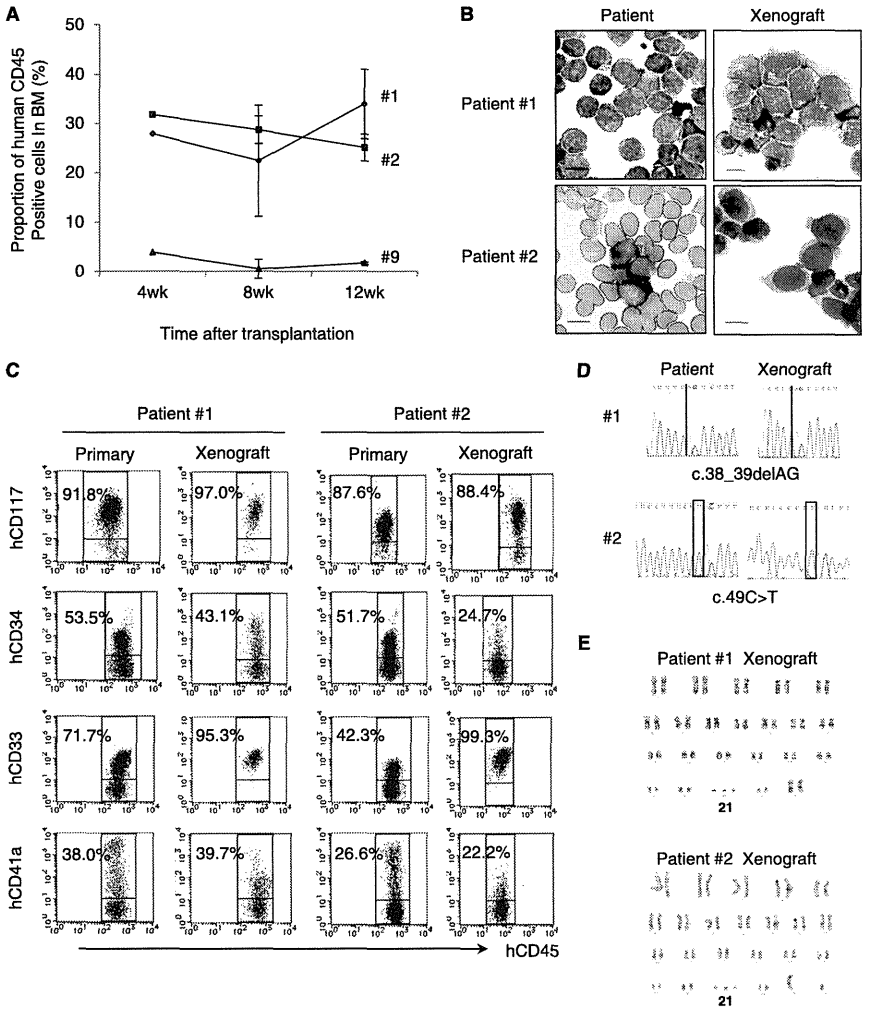

Figure 1. TAM cells engrafted in NOG mice. *(A) Proportion of human CD45$^+$ cells in the BM of NOG mice at 4, 8, and 12 weeks after transplantation (n = 3–5 per group).* (B) May-Giemsa staining of the BM smear of patients and cytospin preparation of human CD45$^+$ cells in the recipient NOG mice. Blast cells with cytoplasmic blebbing consistent with megakaryocytic differentiation were present in the BM of recipient mice. (C) Surface marker analysis of engrafted TAM cells. Human CD45$^+$ TAM-derived cells expressing hCD117, hCD34, hCD33, and hCD41a are detected in the recipient's BM. Blast cells were identified by CD45/SSC gating, and debris (low forward scatter) and dead cells (4′,6-diamidino-2-phenylindole positive) were excluded from the analysis. A representative result of >3 experiments is shown. (D) Genomic direct sequencing shows the presence of concordant GATA1 mutation in xenograft and original patients (1 and 2). (E) G-band karyotyping of TAM-derived cells in recipient murine BM shows no *additional chromosome abnormality apart from constitutional trisomy 21, consistent with the findings in the original patients. The GATA1 mutation and the karyotype of* engrafted cells from patient 9 were not assessed because of a low cell number.

gain of 1q was recurrently observed in this series, the duplicated regions were diverse: 1q25.2-1q44 (1°, m3-11), 1q21.3-1q44 (2°, m3-4), 1q31.2-1q44 (2°, m3-4), and the whole arm of chromosome 1q (3° in Figures 3 and 4A; supplemental Figure 2). In m2-5, a deletion of 3q24 appeared in the 2° and 3° recipients. These results demonstrated that TAM cells derived from patient 1

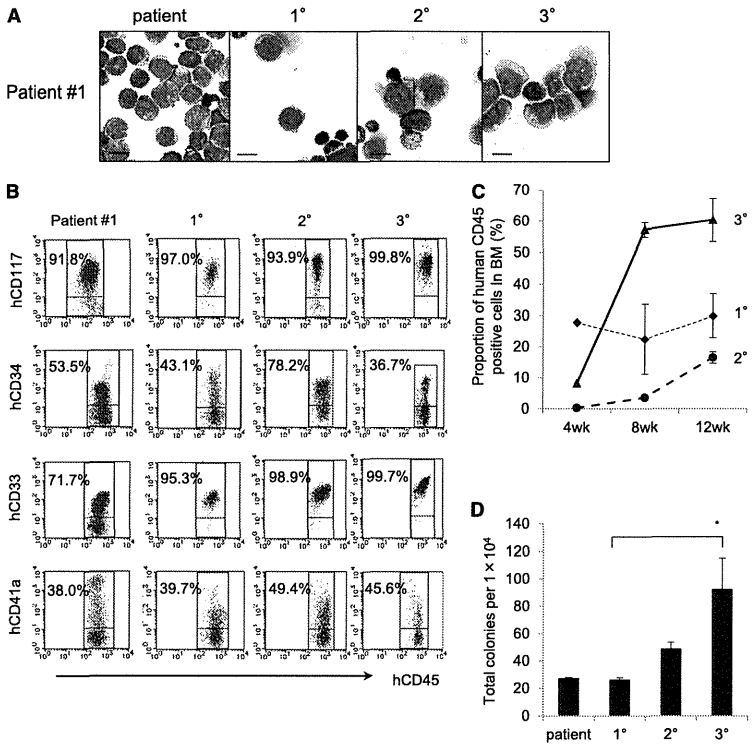

Figure 2. The NOG mouse model can support self-renewal of TAM-derived cells. (A) May-Giemsa staining of TAM-derived cells in recipients of patient 1. (B) Surface marker analysis of TAM-derived cells in recipients by flow cytometry. Viable cells were gated according to their forward scatter (FSC) and 4′,6-diamidino-2-phenylindole staining, blast cells were identified by CD45/SSC gating, and hCD45$^+$-gated cells were tested for the expression of hCD117, hCD34, hCD33, and hCD41a. (C) Proportion of hCD45$^+$ cells in BM of 1°, 2°, and 3° recipient mice at 4, 8, and 12 weeks after transplantation. (D) Colony assay of hCD45$^+$ cells in BM of 1°, 2°, and 3° recipient mice. hCD45$^+$ cells were seeded at 1.0×10^4 cells per 35-mm dish in triplicate, and the number of colonies in each dish was counted. Bars represent the standard deviation of the mean of 3 independent experiments. *Significant difference ($P < .05$).

acquired various CNAs and showed divergent repopulating capacity in our xenograft model.

TAM-NOG xenograft model revealed the presence of a minor clone with a distinct GATA1 mutation

ML-DS can arise from a minor TAM clone with a *GATA1* mutation that is distinct from that of the major TAM clone in a patient.[13] To determine whether the *GATA1* mutation in the primary patient's TAM cells was preserved in engrafted TAM-derived cells, *GATA1* mutation analysis was performed. TAM-derived cells in the series m3-4, m3-5, m3-7, and m3-8 had the same *GATA1* mutation (c.38_39delAG) as that of patient 1 (Figure 4A). Surprisingly, this mutation was not detected in TAM-derived cells in m2-1, m2-2, m2-5, and m3-11; instead, these samples showed a distinct *GATA1* mutation (c.1A>G) that was not detectable in the primary patient sample by direct sequencing. One of the 1° recipients (m2-4) showed both *GATA1* mutations. These results suggested that a minor clone with a distinct *GATA1* mutation (c.1A>G) was present in the primary patient sample and that this minor clone coexisted with, or predominated over, other clones in some 1° recipients. Therefore, a mutation-specific restriction enzyme digestion assay was performed using the primary sample from patient 1, which confirmed the presence of cells with the *GATA1* mutation (c.1A>G) as a minor clone (Figure 4B). Moreover, this minor clone propagated and acquired CNAs in NOG mice independently of the major clones (Figure 4A), further demonstrating the genetic heterogeneity of TAM cells. Interestingly, the major clone in the original patient 1 sample with a c.38_39delAG *GATA1* mutation and no CNAs did not become dominant in any of the recipients.

Minor subclone with additional CNAs was present in the primary TAM patient sample

TAM-derived cells in multiple 1° recipients derived from patient 1 had various CNAs including deletions of 16q22 and 16q24

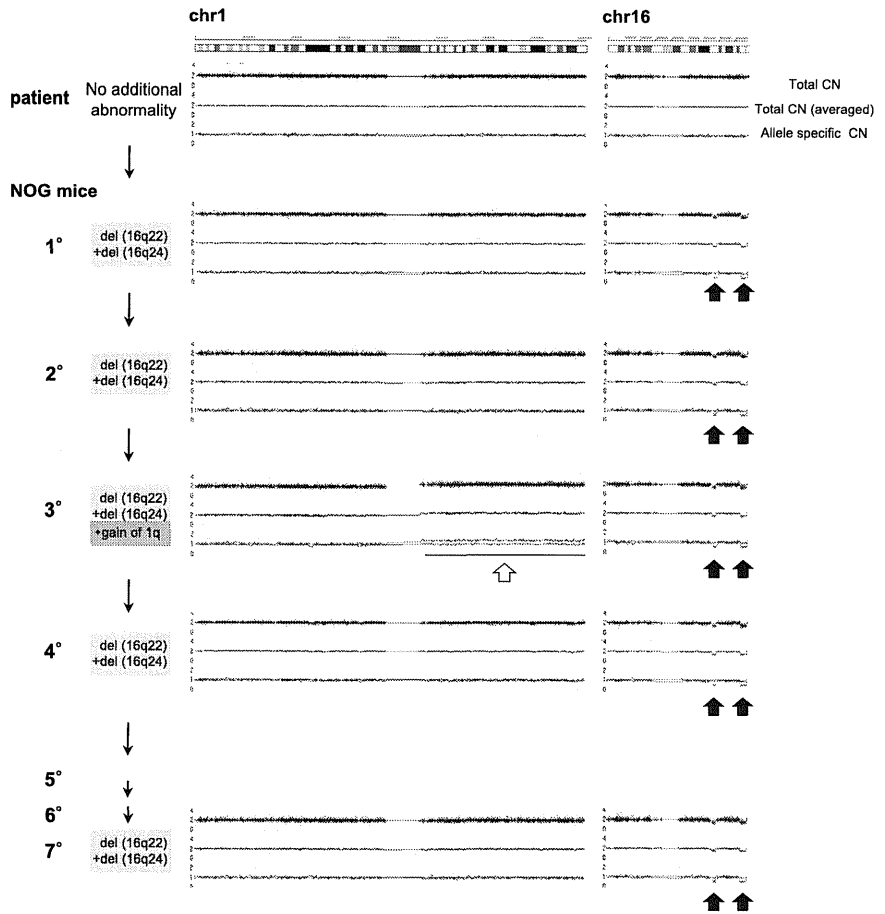

Figure 3. Sequential CNA analysis of TAM-derived cells in the recipients of patient 1. DNA obtained from the original patient sample and sorted hCD45$^+$ recipient BM cells were analyzed by Affymetrix GeneChip Mapping 250K arrays and compared with the PB sample of the original patient in complete remission phase. The primary sample of the patient in TAM phase (blast 92%) had no CNA. hCD45$^+$ BM cells of 1° to 7° recipients had a hemi-allelic deletion in regions 16q22 and 16q24 (black arrows). The 3° recipient had a gain of the entire arm of chromosome 1q (white arrow) in addition to deletion of 16q22 and 16q24. Arrowhead indicates abnormal CNA.

(Figure 4A). To determine whether these subclones were present at low levels in the primary sample of patient 1, specific PCR for the 16q22 deletion was performed using primer pairs designed to bookend the deletion site. CNA analysis and genome sequencing data in these deletion sites (16q22 and 16q24) revealed the presence of genomic breakage and inversion (Figure 5A; supplemental Figures 3 and 4; see supplemental Methods for details). A primer set was designed to detect the deduced breakpoint and used to perform PCR on TAM-derived cells from patient 1 in the recipients with 16q22 and 16q24 deletions. PCR using genomic DNA from TAM-derived cells in the 1° to 8° recipients of the first series of transplantations (Figure 3) produced a uniformly bright DNA fragment of the same size, consistent with the results of CNA profiling (Figure 5B). A faint fragment was detected by applying this PCR method to genomic DNA from the primary patient sample (patient 1), which was confirmed to contain the deletion breakpoint in 16q22 by Sanger sequencing. These results demonstrated that TAM cells with the 16q22 and 16q24 deletions already resided as a minor population in the original sample of patient 1. The frequency of the mutant cells was estimated to be ~1.0% to 0.2% of the patient's PBMCs by a serial dilution assay (supplemental Figure 5).

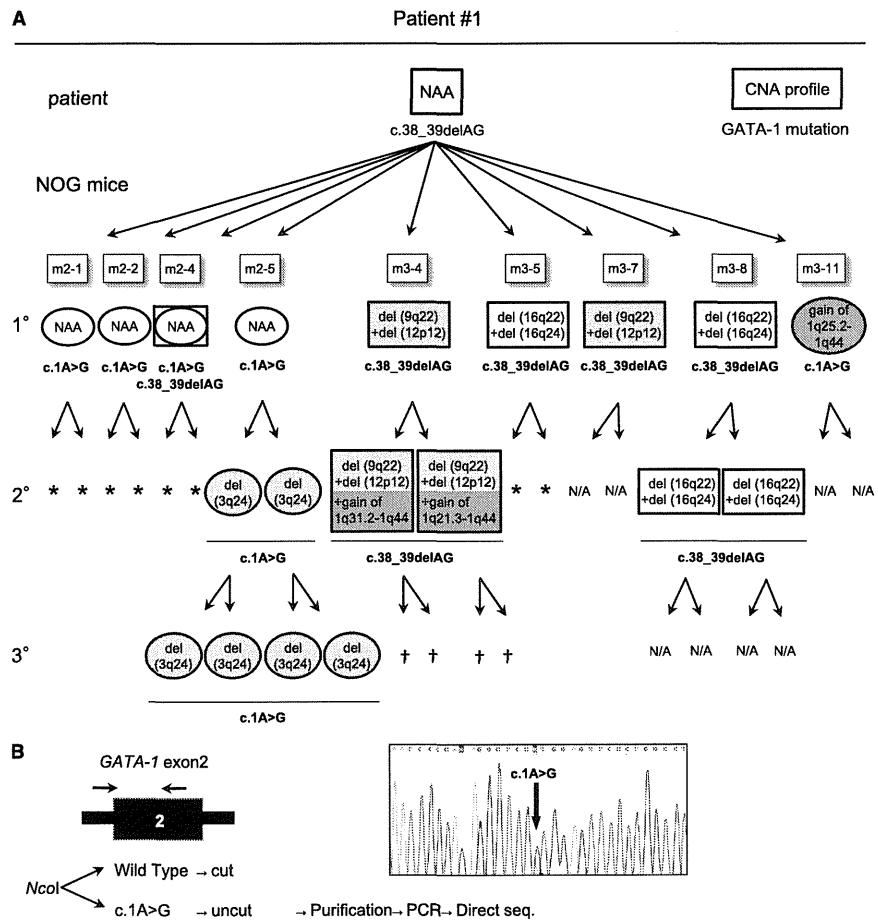

Figure 4. TAM-derived cells show genetic and functional diversity. (A) Summary of the serial transplantation of TAM cells of patient 1 and the results of CNA profiling and *GATA1* mutation analysis. The original patient sample had a single *GATA1* mutation, c.38G_39delAG, and no additional CNAs. Diverse subpopulations with or without additional CNAs expanded in each recipient. *GATA1* mutation analysis showed 2 distinct mutations in recipients: one identical to that of the original patient (c.38_39delAG) and a different mutation (c.1A>G). The mice harboring cells with the original mutation (c.38_39delAG) are shown in rectangles, and the mice with cells harboring the other mutation (c.1A>G) are shown in ovals, with a CNA profile note inside. The *GATA1* mutation is indicated below the symbol. NAA, no additional alteration; N/A, not assessed because of low blast cell count. †Death of recipient before analysis. *No engraftment. (B) Detection of a minor clone with the c.1A>G mutation in the original sample of patient 1. NcoI digestion of a DNA fragment obtained by PCR of *GATA1* exon 2 yielded 2 fragments in the wild type, whereas the mutant allele (c.1A>G) remained undigested. PCR of the undigested band and direct sequence analysis identified the same *GATA1* mutation (c.1A>G mutation) in the patient sample. Black arrow indicates the primer set.

The same method was used to detect a subclone with a 3q24 deletion in the primary patient sample (m2-5; Figure 4A; supplemental Figure 6). At the site of the deletion, genomic breakage was confirmed, and the ends were bound by insertion of a G-nucleotide (Figure 5C; supplemental Figure 7). Consistent with the results of CNA profiling, PCR using DNA from engrafted cells in the 2° and 3° mice (m2-5; Figure 4A) produced a bright DNA fragment, which was confirmed to contain the deletion breakpoint in 3q24 by Sanger sequencing (Figure 5D). Engrafted cells from the BM of the 1° recipient produced a faint DNA fragment, although CNAs were not detected in these cells by array-based methods. We could not detect the corresponding DNA fragment in the primary sample of patient 1. These results suggest the subclone with the 3q24 deletion arose in the 1° recipient mouse as a minor population, emerged as a major population in the 2° recipient, and subsequently engrafted into the 3° recipients.

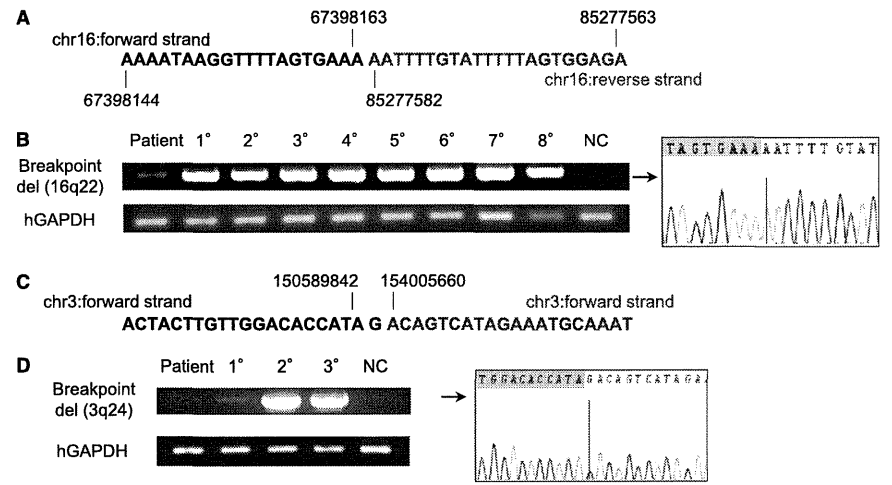

Figure 5. A minor subclone with additional CNAs was present in the primary TAM patient sample, whereas a new clone emerged in a 1° recipient. (A) Contig of del (16q22) breakpoint deduced by whole genome sequencing of the clone containing del(16q22) and del(16q24) in patient 1. Details are shown in supplemental Figure 6. (B) Breakpoint-specific PCR for the del(16q22) clone using genomic DNA from the original patient sample (1; PB in TAM phase), 1° to 8° xenografts (hCD45$^+$ BM cells; Figure 3), and NC (negative control; PBMCs from a healthy adult). Cells from 1° to 8° recipients showed a bright band. The original patient sample showed a faint band, and direct sequencing revealed the presence of the deduced breakpoint for del(16q22). Human glyceraldehyde-3-phosphate dehydrogenase (hGAPDH) was used as an internal control. (C) Contig of del(3q24) breakpoint deduced by whole-genome sequencing. Details are shown in supplemental Figure 9. (D) Breakpoint-specific PCR for the del(3q24) clone using genomic DNA from the original patient sample (1; PB in TAM phase), 1° to 3° xenografts (hCD45$^+$ BM cells, m2–5; Figure 4A), and NC. Cells from 2° and 3° recipients showed a bright band. No band was detected in the original patient sample, but a faint band was detected in the 1° recipient sample. hGAPDH was used as an internal control. Direct sequencing confirmed the presence of cells with the deduced breakpoint for del(3q24) in the 1° recipient.

However, because the sensitivity of the specific PCR targeting of the 3q24 deletion was ~0.1% as determined by the dilution assay (data not shown), it is also possible that this minor clone already existed in the primary patient sample at a frequency below the sensitivity limit. Collectively, our results provide evidence that subclones with additional genetic alterations already exist in the TAM phase and suggest that clonal selection occurs continuously in this xenograft model.

TAM cells derived from patients who did not develop ML-DS had limited self-renewal capacity and fewer additional CNAs than those from the patient who developed ML-DS

To assess whether TAM cells derived from the patients who did not develop ML-DS had similar self-renewal capacity and genetic instability to those from patient 1, CNA analysis of TAM-derived cells was performed by transplanting the preserved PBMC samples of patients 2 and 9. In patient 2, 4 1° transplantation attempts resulted in successful engraftment. The primary sample of patient 2 had no CNAs (Figure 6A). However, TAM-derived cells in 1 of the 1° recipients (m2-2) showed 7p and 7q deletions, suggesting that a subclone with these CNAs may exist in the primary patient sample. The other 2 1° recipients had no additional CNAs (m2-1 and m3-6). In patient 9, engraftment succeeded in 5 1° recipients, and no additional CNAs were detected in either primary patient sample or engrafted TAM-derived cells (Figure 6B). The engrafted cells in all of the recipient mice harbored the same *GATA1* mutation as that of the primary samples of patients 2 and 9. In these 2 cases, our xenograft assay did not detect potent TAM clones with self-renewal capacity in serial transplantation assays (Figure 6A-B; supplemental Table 1), which may reflect the favorable clinical outcome of these patients.

Taken together, the results show that only the TAM cells derived from patients who subsequently developed ML-DS had long-term self-renewal capacity with additional CNAs in our serial transplantation assay.

Discussion

New genomic technologies have led to a better understanding of the complex clonal architecture of leukemia and have shown that disease progression occurs through clonal evolution.[20-22,32] However, most studies have been based on the retrospective analysis of frank leukemia samples, and data on the evolutionary process occurring in the preleukemic phase are limited because primary preleukemia samples are rarely available and are difficult to maintain in vitro or in vivo. TAM is a unique hematologic condition associated with DS that is mostly self-limited but leads to ML-DS in 20% of cases after spontaneous remission. Therefore, TAM has been considered a preleukemic state and is a suitable pathological condition to analyze the evolutionary process of leukemia.

Because 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

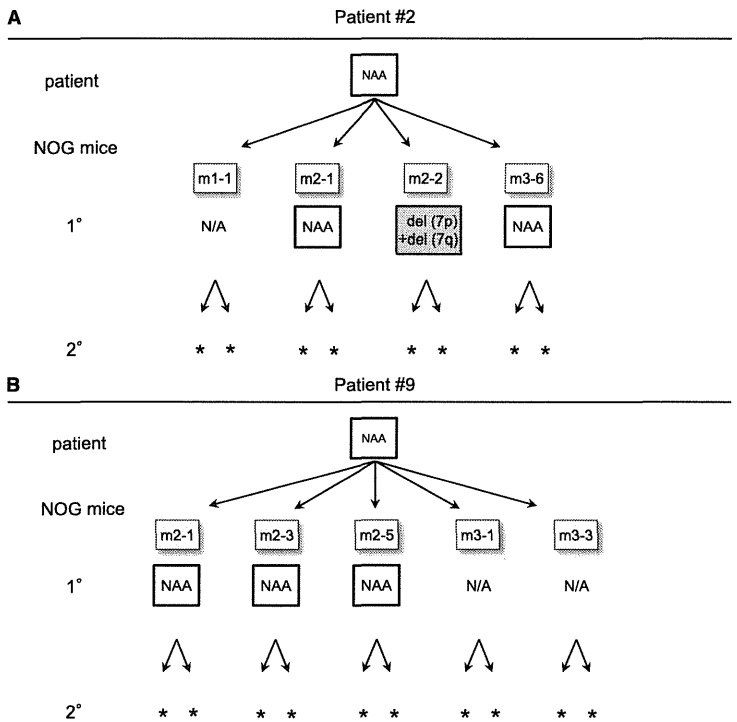

Figure 6. Serial transplantation and CNA profiling of TAM-derived cells from patients that did not develop ML-DS. (A) Serial transplantation assay using TAM cells from patient 2. Four attempts resulted in successful analysis in 1° recipients. CNAs with del(7p) and del(7q) were observed in 1 recipient (m2-2). No additional CNAs were observed in any other recipients. No engraftment was observed in 2° recipients. (B) Serial transplantation assay of TAM cells from patient 9. Five attempts resulted in successful analysis in 1° recipients. No additional CNAs were observed in any analyzed recipients, and no engraftment was observed in 2° recipients. NAA, no additional alteration; N/A, not assessed because of low cell count. *No engraftment.

leukemia,[19-22] we hypothesized that xenograft models of TAM cells would be an attractive method to investigate leukemogenesis. In this report, we demonstrated the long-term engraftment of primary TAM cells in NOG mice and showed that TAM cells from a patient that subsequently developed ML-DS had the potential to gain diverse additional genomic alterations and self-renewal capacity. Although we were unable to determine whether the clonal evolution of TAM cells observed in our model reflected the clinical phenotype of the original patient because of insufficient sample from the ML-DS phase, our model is likely to enable the prospective evaluation of leukemic evolution and can be a powerful tool to study the pathophysiology of leukemogenesis. Our model using NOG mice contrasts somewhat with the study by Chen et al,[27] who reported that TAM cells resided only in the BM after intra-BM infusion into NOD/SCID mice. We speculate that a severe and unique immunodeficient microenvironment may have contributed to the successful engraftment of TAM cells in NOG mice.

In the present study, primary TAM cell samples from 3 of 11 patients engrafted in NOG mice (Figure 1), but serial engraftment was successful only with cells obtained from the patient who developed ML-DS at the age of 1 year (Figures 2, 3, and 6). The results of extensive serial transplantation revealed the emergence of subclones with various additional CNAs characteristic of ML-DS (Figures 3 and 4). Furthermore, we showed that minor subclones with various CNAs and a distinct *GATA1* mutation were already present in the ML-DS patient during the early TAM phase (Figures 4 and 5), as previously described for polyclonality of TAM.[2,33] These findings suggest that several preleukemic clones with high leukemia-initiating potential may already reside as minor clones in TAM cells of patients fated to develop ML-DS and show high repopulating capacity in the special microenvironment of NOG mice. Our findings support the hypothesis that ML-DS develops from a pool of heterogeneous minor clones through clonal selection, illustrating the early evolutionary process of leukemia.[34]

Long-term engraftment of TAM-derived cells was observed for only a minority of TAM patients. This finding suggests that factors other than the properties of the TAM-derived cells, such as technical issues, affected engraftment efficiency. In this regard, increasing the number of transplanted cells resulted in a higher rate of engraftment in recipients of samples from patient 9 (supplemental Table 1).

However, there was no clear association between the percentage of TAM blast cells in transplanted samples and successful engraftment (supplemental Table 1). Likewise, frozen samples from 3 patient samples (patients 1, 2, and 9) were as efficient for engraftment as fresh samples from these patients. Therefore, although the number of injected TAM cells and technical issues may affect engraftment, we speculate that engraftment efficiency is an intrinsic property of each TAM-derived cell population.

In addition to trisomy 21, somatic *GATA1* mutation is considered an early essential event of TAM and ML-DS occurring in utero.[35,36] Interestingly, our TAM-NOG mice model enhanced the emergence of a minor clone with a distinct *GATA1* mutation that was not detectable in the original patient sample by conventional sequencing methods. In our model, a minor *GATA1* mutant clone expanded predominantly in some recipients and acquired CNAs independently of clones with the original *GATA1* mutation, raising the possibility that leukemic evolution occurred from this minor clone, similar to the clinical observation in our previous report.[13] In this scenario, a common founder clone of TAM/ML-DS may be established before the acquisition of the *GATA1* mutation, or TAM clones with distinct *GATA1* mutations may arise independently in the fetal period.

It has long been considered that the linear sequential acquisition of genetic alterations induces disease progression in TAM/ML-DS.[37] By contrast, recent studies using high-throughput genomic technology indicate that evolutionary trajectories are more complex and branching in other cancers and leukemias, as previously proposed by Nowell.[38] In this theory, genomic instability in founder cells gives rise to heterogeneous mutant subclones, and under selective pressure, some subclones expand to result in disease progression, whereas others become extinct or remain dormant. Thus, leukemic clones may evolve and emerge through the complex interaction of selectively advantageous "driver" mutations, additional advantageous "cooperating" mutations, neutral "passenger" mutations, and deleterious mutations.[32,38] It is clinically true that genomic alterations are more frequently observed in ML-DS than in TAM.[2,11,39] In this paper, we showed that diverse subclones with various CNAs can be generated in TAM, and these events occurred preferentially in a patient who later developed ML-DS. These findings suggest the presence of leukemic driver mutations in the early phase of TAM in this patient, which may have induced genomic instability. We were unable to find any candidate tumor-associated genes on the deletion sites (3q24, 9q22, 12p12, 16q22, and 16q24) of TAM-derived cells using the The National Center for Biotechnology Information database, suggesting that other genetic mutations and epigenetic events may contribute to the progression to ML-DS, including a few candidate mutations identified previously.[40-42] It is also noteworthy that subclones in each recipient mouse showed different repopulating capacities in this study. The dominant clones in each recipient were not always identical in the 1° generations, and the dominant clone in a certain recipient did not always propagate dominantly in the next generation recipients (Figure 4A). Differences between the recipient mice or technical problems may have caused variations in engraftment outcome, which is a potential weakness of this xenograft model; however, it is more likely that cooperating genetic event(s) important for leukemogenesis led to the cells of a specific TAM clone becoming the dominant population in each recipient. Such cooperating event(s) could have a considerable impact on a preleukemic TAM clone, and clonal selection might occur in a somewhat random manner. Thus, leukemic evolution may depend on random chance to an extent. Our TAM xenograft model may help demonstrate the branching architecture of clonal evolution in a preleukemic phase, which contrasts with a linear and deterministic pattern of evolution.[34,38] Further genomewide analysis is needed to elucidate the true driver or cooperating mutation(s) and unravel the evolutionary process of leukemia.

In conclusion, we established a xenograft model of TAM using highly immunodeficient NOG mice. Our model enabled the observation of clonal selection and expansion of minor mutant TAM clones and is likely to mimic the early phase of the leukemic evolutionary process, demonstrating the striking genetic heterogeneity and the propagating potential of minor clones in a preleukemic phase. Our xenograft model could be valuable tool for gaining insight into the leukemogenesis of ML-DS and for evaluating the prognosis of TAM patients.

Acknowledgments

The authors thank all the TAM patients and their families for their participation. The authors thank Drs Akira Niwa, Masashi Sanada, Hironao Numabe, Tomoki Kawai, Takahiro Yasumi, and Ryuta Nishikomori for technical advice.

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

Authorship

Contribution: S.S., I.K., T.M., H.F., and K.U. performed sample collection and processing; S.S., K.T., K.Y., and R.W. performed experiments; A.S.-O. performed microarray analysis (accession number GSE44739); Y.S. and S.M. provided expert statistical analysis; S.S., Y.O., and T.T. analyzed results and made the figures; M.I. and T.N. generated NOG mice; and S.S., K.W., H.H., S.A., E.I., S.O., and T.H. designed the research and wrote the paper.

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

Correspondence: Toshio Heike, Department of Pediatrics, Graduate School of Medicine, Kyoto University, 54 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto, 606-8507, Japan; e-mail: heike@kuhp.kyoto-u.ac.jp.

References

1. Pine SR, Guo Q, Yin C, Jayabose S, Druschel CM, Sandoval C. Incidence and clinical implications of GATA1 mutations in newborns with Down syndrome. *Blood.* 2007;110(6):2128-2131.
2. Massey GV, Zipursky A, Chang MN, et al; Children's Oncology Group (COG). A prospective study of the natural history of transient leukemia (TL) in neonates with Down syndrome (DS): Children's Oncology Group (COG) study POG-9481. *Blood.* 2006;107(12):4606-4613.
3. Zipursky A, Poon A, Doyle J. Leukemia in Down syndrome: a review. *Pediatr Hematol Oncol.* 1992;9(2):139-149.
4. Hitzler JK. Acute megakaryoblastic leukemia in Down syndrome. *Pediatr Blood Cancer.* 2007; 49(7 Suppl):1066-1069.
5. Hitzler JK, Cheung J, Li Y, Scherer SW, Zipursky A. GATA1 mutations in transient leukemia and acute megakaryoblastic leukemia of Down syndrome. *Blood.* 2003;101(11):4301-4304.
6. Wechsler J, Greene M, McDevitt MA, Anastasi J, Karp JE, Le Beau MM, Crispino JD. Acquired mutations in GATA1 in the megakaryoblastic leukemia of Down syndrome. *Nat Genet.* 2002; 32(1):148-152.

7. Groet J, McElwaine S, Spinelli M, et al. Acquired mutations in GATA1 in neonates with Down's syndrome with transient myeloid disorder. *Lancet.* 2003;361(9369):1617-1620.
8. Xu G, Nagano M, Kanezaki R, et al. Frequent mutations in the GATA-1 gene in the transient myeloproliferative disorder of Down syndrome. *Blood.* 2003;102(8):2960-2968.
9. Weiss MJ, Orkin SH. Transcription factor GATA-1 permits survival and maturation of erythroid precursors by preventing apoptosis. *Proc Natl Acad Sci USA.* 1995;92(21):9623-9627.
10. Ferreira R, Ohneda K, Yamamoto M, Philipsen S. GATA1 function, a paradigm for transcription factors in hematopoiesis. *Mol Cell Biol.* 2005; 25(4):1215-1227.
11. Hayashi Y, Eguchi M, Sugita K, et al. Cytogenetic findings and clinical features in acute leukemia and transient myeloproliferative disorder in Down's syndrome. *Blood.* 1988;72(1):15-23.
12. Forestier E, Izraeli S, Beverloo B, et al. Cytogenetic features of acute lymphoblastic and myeloid leukemias in pediatric patients with Down syndrome: an iBFM-SG study. *Blood.* 2008; 111(3):1575-1583.
13. Xu G, Kato K, Toki T, Takahashi Y, Terui K, Ito E. Development of acute megakaryoblastic leukemia from a minor clone in a Down syndrome patient with clinically overt transient myeloproliferative disorder. *J Pediatr Hematol Oncol.* 2006;28(10): 696-698.
14. Vyas P, Crispino JD. Molecular insights into Down syndrome-associated leukemia. *Curr Opin Pediatr.* 2007;19(1):9-14.
15. Kirsammer G, Jilani S, Liu H, Davis E, Gurbuxani S, Le Beau MM, Crispino JD. Highly penetrant myeloproliferative disease in the Ts65Dn mouse model of Down syndrome. *Blood.* 2008;111(2): 767-775.
16. Carmichael CL, Majewski IJ, Alexander WS, Metcalf D, Hilton DJ, Hewitt CA, Scott HS. Hematopoietic defects in the Ts1Cje mouse model of Down syndrome. *Blood.* 2009;113(9): 1929-1937.
17. Alford KA, Slender A, Vanes L, et al. Perturbed hematopoiesis in the Tc1 mouse model of Down syndrome. *Blood.* 2010;115(14):2928-2937.
18. Malinge S, Bliss-Moreau M, Kirsammer G, Diebold L, Chlon T, Gurbuxani S, Crispino JD. Increased dosage of the chromosome 21 ortholog Dyrk1a promotes megakaryoblastic leukemia in a murine model of Down syndrome. *J Clin Invest.* 2012;122(3):948-962.
19. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med.* 1997;3(7):730-737.
20. Clappier E, Gerby B, Sigaux F, et al. Clonal selection in xenografted human T cell acute lymphoblastic leukemia recapitulates gain of malignancy at relapse. *J Exp Med.* 2011;208(4): 653-661.
21. Notta F, Mullighan CG, Wang JCY, et al. Evolution of human BCR-ABL1 lymphoblastic leukaemia-initiating cells. *Nature.* 2011; 469(7330):362-367.
22. Anderson K, Lutz C, van Delft FW, et al. Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature.* 2011;469(7330):356-361.
23. Hiramatsu H, Nishikomori R, Heike T, Ito M, Kobayashi K, Katamura K, Nakahata T. Complete reconstitution of human lymphocytes from cord blood CD34+ cells using the NOD/SCID/gammacnull mice model. *Blood.* 2003;102(3):873-880.
24. Fujino H, Hiramatsu H, Tsuchiya A, et al. Human cord blood CD34+ cells develop into hepatocytes in the livers of NOD/SCID/gamma(c)null mice through cell fusion. *FASEB J.* 2007;21(13):3499-3510.
25. Kato M, Sanada M, Kato I, et al. Frequent inactivation of A20 in B-cell lymphomas. *Nature.* 2009;459(7247):712-716.
26. Kato I, Niwa A, Heike T, et al. Identification of hepatic niche harboring human acute lymphoblastic leukemic cells via the SDF-1/ CXCR4 axis. *PLoS ONE.* 2011;6(11):e27042.
27. Chen J, Li Y, Doedens M, Wang P, Shago M, Dick JE, Hitzler JK. Functional differences between myeloid leukemia-initiating and transient leukemia cells in Down's syndrome. *Leukemia.* 2010;24(5):1012-1017.
28. Ito M, Hiramatsu H, Kobayashi K, et al. NOD/ SCID/gamma(c)(null) mouse: an excellent recipient mouse model for engraftment of human cells. *Blood.* 2002;100(9):3175-3182.
29. Lacombe F, Durrieu F, Briais A, et al. Flow cytometry CD45 gating for immunophenotyping of acute myeloid leukemia. *Leukemia.* 1997;11(11): 1878-1886.
30. Suda J, Eguchi M, Akiyama Y, et al. Differentiation of blast cells from a Down's syndrome patient with transient myeloproliferative disorder. *Blood.* 1987;69(2):508-512.
31. Silva ML, do Socorro Pombo-de-Oliveira M, Raimondi SC, et al. Unbalanced chromosome 1 abnormalities leading to partial trisomy 1q in four infants with Down syndrome and acute megakaryocytic leukemia. *Mol Cytogenet.* 2009;2:7.
32. Welch JS, Ley TJ, Link DC, et al. The origin and evolution of mutations in acute myeloid leukemia. *Cell.* 2012;150(2):264-278.
33. Ahmed M, Sternberg A, Hall G, et al. Natural history of GATA1 mutations in Down syndrome. *Blood.* 2004;103(7):2480-2489.
34. Greaves M, Maley CC. Clonal evolution in cancer. *Nature.* 2012;481(7381):306-313.
35. Shimada A, Xu G, Toki T, Kimura H, Hayashi Y, Ito E. Fetal origin of the GATA1 mutation in identical twins with transient myeloproliferative disorder and acute megakaryoblastic leukemia accompanying Down syndrome. *Blood.* 2004; 103(1):366.
36. Hitzler JK, Zipursky A. Origins of leukaemia in children with Down syndrome. *Nat Rev Cancer.* 2005;5(1):11-20.
37. Khan I, Malinge S, Crispino J. Myeloid leukemia in Down syndrome. *Crit Rev Oncog.* 2011;16(1-2): 25-36.
38. Nowell PC. The clonal evolution of tumor cell populations. *Science.* 1976;194(4260):23-28.
39. Blink M, van den Heuvel-Eibrink MM, Aalbers AM, et al. High frequency of copy number alterations in myeloid leukaemia of Down syndrome. *Br J Haematol.* 2012;158(6):800-803.
40. Malkin D, Brown EJ, Zipursky A. The role of p53 in megakaryocyte differentiation and the megakaryocytic leukemias of Down syndrome. *Cancer Genet Cytogenet.* 2000;116(1):1-5.
41. Norton A, Fisher C, Liu H, 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.* 2007;110(3):1077-1079.
42. Malinge S, Ragu C, Della-Valle V, et al. Activating mutations in human acute megakaryoblastic leukemia. *Blood.* 2008;112(10):4220-4226.

nature genetics

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

Kenichi Yoshida[1,2,17], Tsutomu Toki[3,17], Yusuke Okuno[1,17], Rika Kanezaki[3], Yuichi Shiraishi[4], Aiko Sato-Otsubo[1,2], Masashi Sanada[1,2], Myoung-ja Park[5], Kiminori Terui[3], Hiromichi Suzuki[1,2], Ayana Kon[1,2], Yasunobu Nagata[1,2], Yusuke Sato[1,2], RuNan Wang[3], Norio Shiba[5], Kenichi Chiba[4], Hiroko Tanaka[6], Asahito Hama[7], Hideki Muramatsu[7], Daisuke Hasegawa[8], Kazuhiro Nakamura[9], Hirokazu Kanegane[10], Keiko Tsukamoto[11], Souichi Adachi[12], Kiyoshi Kawakami[13], Koji Kato[14], Ryosei Nishimura[15], Shai Izraeli[16], Yasuhide Hayashi[5], Satoru Miyano[4,6], Seiji Kojima[7], Etsuro Ito[3,18] & Seishi Ogawa[1,2,18]

Transient abnormal myelopoiesis (TAM) is a myeloid proliferation resembling acute megakaryoblastic leukemia (AMKL), mostly affecting perinatal infants with Down syndrome. Although self-limiting in a majority of cases, TAM may evolve as non-self-limiting AMKL after spontaneous remission (DS-AMKL). Pathogenesis of these Down syndrome–related myeloid disorders is poorly understood, except for GATA1 mutations found in most cases. Here we report genomic profiling of 41 TAM, 49 DS-AMKL and 19 non-DS-AMKL samples, including whole-genome and/or whole-exome sequencing of 15 TAM and 14 DS-AMKL samples. TAM appears to be caused by a single GATA1 mutation and constitutive trisomy 21. Subsequent AMKL evolves from a pre-existing TAM clone through the acquisition of additional mutations, with major mutational targets including multiple cohesin components (53%), CTCF (20%), and EZH2, KANSL1 and other epigenetic regulators (45%), as well as common signaling pathways, such as the JAK family kinases, MPL, SH2B3 (LNK) and multiple RAS pathway genes (47%).

TAM represents a transient proliferation of immature megakaryoblasts that occurs in 5–10% of perinatal infants with Down syndrome[1,2]. Although morphologically indistinguishable from AMKL, TAM is self-limiting in the majority of cases and usually terminates spontaneously within 3–4 months of birth[1]. Hepatic infiltration of myeloid cells is a common finding and can be severe enough to be fatal, owing to hepatic failure, with liver fibrosis occurring in 5–16% of cases[2–4]. Moreover, even when spontaneous remission is achieved, approximately 20–30% of surviving infants develop DS-AMKL years after remission, although some DS-AMKL cases have no documented history of TAM[4]. In contrast to non–Down syndrome–related AMKL (non-DS-AMKL), which generally shows poor prognosis, individuals with DS-AMKL typically have a favorable prognosis. In molecular pathogenesis of these Down syndrome–related myeloid disorders, *GATA1* mutations are detected in virtually all affected infants, suggesting their central role in Down syndrome–related myeloid proliferation[5,6]. However, it is still open to question whether a *GATA1* mutation is sufficient for the development of TAM in individuals with Down syndrome, what is the cellular origin of the subsequent AMKL, whether additional gene mutations are required for progression to AMKL, and, if so, what are their gene targets, although several genes have been reported to be mutated in occasional cases with DS-AMKL, including *JAK1*, *JAK2* and *JAK3* (refs. 7–10), *TP53* (refs. 10,11), *FLT3* (ref. 8) and *MPL*[12]. We reasoned that identifying a comprehensive registry of gene mutations and tracking them at a clonal level using massively parallel sequencing would provide vital information for addressing these questions.

RESULTS

Genomic landscape of Down syndrome–related myeloid neoplasms

We performed whole-genome sequencing of 4 trios consisting of samples from TAM, AMKL and complete remission phases (**Supplementary Figs. 1** and **2** and **Supplementary Table 1**). In total, we confirmed 411 single-nucleotide variants (SNVs) and 17 small nucleotide insertions and deletions (indels) by Sanger sequencing and/or deep resequencing (**Supplementary Fig. 1** and **Supplementary Table 2**). We detected only a few structural variants, including deletion, amplification and uniparental disomy, in the TAM and DS-AMKL genomes (**Fig. 1** and **Supplementary Fig. 3**). The mean number of validated somatic mutations in DS-AMKL samples (71 or 0.023 mutations/Mb) was twice the number observed in TAM samples (36 or 0.012 mutations/Mb) (**Supplementary Fig. 1a**). Mutation numbers in samples from both phases were substantially lower than in most other cancers (**Supplementary Fig. 4**), although differences in mutation rates could partly be affected by different definitions and algorithms for mutation calling. The spectrum of mutations was over-represented by C-to-T and G-to-A transitions in both TAM and DS-AMKL samples, resembling the mutational spectra in gastric and colorectal cancers[13] and in other blood cancers (**Supplementary Fig. 1b**)[14,15]. We unmasked the details of clonal evolution and expansion leading to AMKL through the use of deep sequencing of individual mutations detected by combined whole-genome and whole-exome sequencing (**Fig. 2** and **Supplementary Table 2**). Intratumoral heterogeneity was evident at initial diagnosis with TAM and in the AMKL phase in all cases (**Supplementary Fig. 5**). In UPN001, UPN002 and UPN004, AMKL evolved from one of the major subclones in the TAM phase with a shared *GATA1* mutation, as reported previously in relapsed acute myeloid leukemia (AML) in adults (**Fig. 2a,b,d**)[15]. In contrast, UPN003 showed a unique pattern of clonal evolution, in which AMKL originated from a minor subclone in the TAM phase that was totally unrelated to the predominant clone in terms of somatic mutations, with no mutation shared by both phases, and carried an independent *GATA1* mutation (**Fig. 2c**). In both scenarios, progression to AMKL seemed to be accompanied by many additional mutations, including common driver mutations that were absent in the original TAM population, indicating a multistep process of leukemogenesis.

Exome sequencing

We further investigated non-silent mutations by whole-exome sequencing of additional samples to generate a full registry of driver mutations that are relevant to the development of TAM and subsequent progression to AMKL (**Supplementary Fig. 6** and **Supplementary Table 1**). We detected *GATA1* mutations in all TAM and DS-AMKL cases, indicating sufficient sensitivity in our whole-exome analysis. In total, we confirmed 26 and 81 non-silent somatic mutations identified in the exome analysis of 15 TAM and 14 DS-AMKL samples, respectively, with 3 *GATA1* mutations common to both phases (**Supplementary Table 3**). The mean number of non-silent mutations was significantly higher in DS-AMKL samples (5.8; range of 1–11) than in TAM samples (1.7; range of 1–5) ($P = 0.0002$) (**Fig. 3a**). Of the 107 mutations, 84 were single-nucleotide substitutions that were mostly within coding sequences, except for 4 splice-site mutations. We also observed predominantly C-to-T and G-to-A transitions for non-silent substitutions (**Supplementary Fig. 7**). The remaining mutations were frameshift ($n = 21$) or non-frameshift ($n = 2$) indels, most frequently involving *GATA1* ($n = 13$). One individual with DS-AMKL (UPN004) had no SNVs or indels (**Fig. 3a**), but copy number analysis identified a large deletion at 16q involving the *CTCF* locus (**Supplementary Fig. 3**), suggesting that the alteration of *CTCF* could be a driver event in this case. Therefore, at least one additional genetic lesion other than *GATA1* mutation was detected in our whole-exome sequencing, despite the low frequency of leukemic cells appearing to show the morphology of immature megakaryoblasts (blast percentage) in many cases, which is a known characteristic of DS-AMKL samples[16,17]. Whole-exome sequencing results suggested the presence of intratumoral heterogeneity in the majority of DS-AMKL cases (**Fig. 3b**).

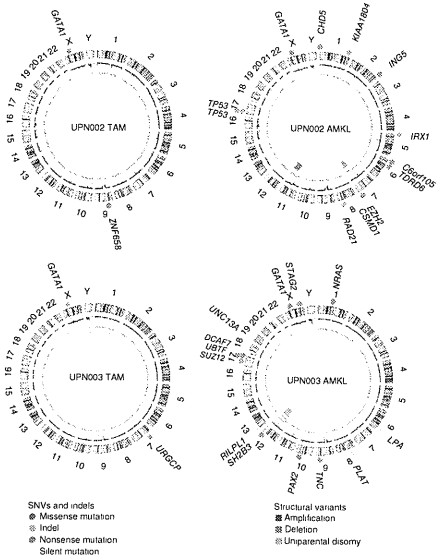

Figure 1 Representative Circos plots of paired TAM and DS-AMKL cases. Locations of somatic mutations, including of missense, frameshift, nonsense and silent mutations (colored circles), are indicated. Total (black) and allele-specific (red and green for alleles showing relatively larger and smaller copy numbers, respectively) genomic copy numbers, as well as somatic structural variants (colored bars), are indicated in the inner circle. Sample IDs are shown within each plot; plots were created with Circus[53].

Spectrum of recurrent mutations in DS-AMKL

Recurrently affected genes are of primary interest in identifying driver mutations. Whereas *GATA1* was the only recurrent mutational target in TAM samples, an additional eight genes were recurrently mutated in the DS-AMKL samples, including *RAD21*, *STAG2*, *NRAS*, *CTCF*, *DCAF7*, *EZH2*, *KANSL1* and *TP53* (**Table 1**). These genes are expressed in a wide variety of hematopoietic compartments, including in both myeloid and lymphoid cells, except for *EZH2*, whose expression is largely confined to CD34$^+$ cells[18] (**Supplementary Fig. 8**). We also found that these genes were expressed in DS-AMKL cells at similar levels to common hematopoietic genes[19], although we did not observe significant difference in their expression levels in DS-AMKL and non-DS-AMKL cells (**Supplementary Fig. 9**).

We then performed targeted deep sequencing of these 8 genes in an extended set of 109 samples (including 29 samples in 25 discovery cases) consisting of 41 TAM, 49 DS-AMKL and 19 non-DS-AMKL samples (**Supplementary Tables 1** and **4**). We also included additional genes in targeted sequencing that were either functionally related to the above eight genes or were mutated only in single cases but had been previously reported to be mutated in DS-AMKL (*JAK3*) or other myeloid neoplasms (*SH2B3*, *SUZ12*, *SRSF2* and *WT1*), together with other common mutational targets in adult myeloid malignancies

[1]Cancer Genomics Project, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. [2]Department of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto University, Kyoto, Japan. [3]Department of Pediatrics, Hirosaki University Graduate School of Medicine, Hirosaki, Japan. [4]Laboratory of DNA Information Analysis, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan. [5]Department of Hematology/Oncology, Gunma Children's Medical Center, Shibukawa, Japan. [6]Laboratory of Sequence Analysis, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan. [7]Department of Pediatrics, Nagoya University Graduate School of Medicine, Nagoya, Japan. [8]Department of Pediatrics, St. Luke's International Hospital, Tokyo, Japan. [9]Department of Pediatrics, Hiroshima University Graduate School of Biomedical Sciences, Hiroshima, Japan. [10]Department of Pediatrics, Graduate School of Medicine, University of Toyama, Toyama, Japan. [11]Division of Neonatology, National Center for Child Health and Development, Tokyo, Japan. [12]Human Health Sciences, Graduate School of Medicine, Kyoto University, Kyoto, Japan. [13]Department of Pediatrics, Kagoshima City Hospital, Kagoshima, Japan. [14]Department of Hematology and Oncology, Children's Medical Center, Japanese Red Cross Nagoya First Hospital, Nagoya, Japan. [15]Department of Pediatrics, School of Medicine, Institute of Medical, Pharmaceutical and Health Sciences, Kanazawa University, Kanazawa, Japan. [16]Functional Genomics, Cancer Research Center, Sheba Medical Center, Tel Hashomer and Tel Aviv University, Tel Aviv, Israel. [17]These authors contributed equally to this work. [18]These authors jointly directed this work. Correspondence should be addressed to S.O. (sogawa-tky@umin.ac.jp) or E.I. (eitoh@cc.hirosaki-u.ac.jp).

Received 3 May; accepted 19 August; published online 22 September 2013; corrected after print 30 October 2013; doi:10.1038/ng.2759

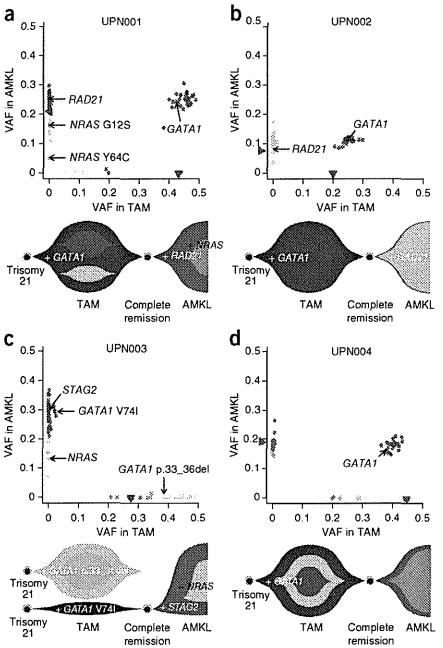

Figure 2 Clonal evolution of Down syndrome–related myeloid disorders. (*a–d*) *Observed VAFs of validated mutations listed in* **Supplementary Table 2** in both TAM and AMKL phases are shown in diagonal plots (top) for UPN001 (**a**), UPN002 (**b**), UPN003 (**c**) and UPN004 (**d**), where VAFs of genes on the X chromosome in male cases or in regions of uniparental disomy were halved. Half the value of the blast percentage, which corresponds to the allele frequency of a heterozygous mutation distributed in all tumor cells, is also shown by a red arrowhead, except for UPN003 AMKL, for which clinical data were not available. Driver mutations including in *GATA1, STAG2, RAD21* and *NRAS* are indicated by black arrows. Predicted chronological behaviors of different leukemia subclones are depicted below each diagonal plot. Distinct mutation clusters are indicated by color. In UPN001, UPN002 and UPN004, founding clones of TAM shown in blue became dominant in the AMKL samples, in which some subsequent subclones evolved through the serial acquisition of SNVs. In contrast, in UPN003, a subclone in the TAM phase (blue) and not the founding clone of TAM (aqua) became dominant in the AMKL sample. VAFs of some mutations were higher than for *GATA1* but seem to be actually equivalent to it given the error range of PCR-based deep sequencing.

(**Supplementary Fig. 10** and **Supplementary Tables 5** and **6**). We also analyzed by RT-PCR two recurrent fusion genes previously reported in non-DS-AMKL cases, *RBM15-MKL1* (*OTT-MAL*)[20,21] and *CBFA2T3-GLIS2* (refs. 22,23).

Mutations of cohesin and associated molecules

Major components of the cohesin complex, including *RAD21* and *STAG2*, were frequent targets of gene mutations in DS-AMKL (**Table 1**). Including an additional mutation in *NIPBL*, 8 of the 14 discovery DS-AMKL cases (57%) had a mutated cohesin or associated component (**Supplementary Table 3**). Cohesin is a multiprotein complex consisting of 4 core components, including the SMC1, SMC3, RAD21 and STAG proteins[24,25]. In concert with several functionally associated proteins, such as the NIPBL and ESCO proteins, cohesin is engaged in the cohesion of newly replicated sister chromatids by forming a ring-like structure[25], preventing their premature separation before late anaphase. Cohesin has also been implicated in post-replicative DNA repair and long-range regulation of gene expression[26–30]. Targeted deep sequencing confirmed recurrent mutations and deletions in all core cohesin components (*STAG2, RAD21, SMC3* and *SMC1A*) and in *NIPBL* in 26 of 49 DS-AMKL cases (53%) but in none of the 41 TAM cases, although 2 non-DS-AMKL cases (11%) had *STAG2* mutations (**Fig. 4a,b** and **Supplementary Tables 7** and **8**). Strikingly, all mutations and deletions in different cohesin components were completely mutually exclusive, suggesting that cohesin function was the common target of these mutations. All but one *STAG2* mutation (encoding a p.Arg370Gln substitution) was either a nonsense, frameshift or splice-site change (**Fig. 4a,b**, **Supplementary Figs. 11** and **12a**, and **Supplementary Table 7**). Similarly, 6 of 9 *RAD21* mutations were heterozygous nonsense or frameshift alterations. Four of the five mutations in *NIPBL, SMC1A* and *SMC3* were also nonsense or splice-site changes causing abnormal exon skipping (**Fig. 4a** and **Supplementary Table 7**). Thus, most of these mutations were thought to result in premature truncation, leading to loss of cohesin function. The leukemogenic mechanism of mutated cohesin components is still elusive; some studies have implicated aneuploidy caused by cohesin dysfunction in oncogenic actions[31]. However, DS-AMKL cases have been characterized by a largely normal karyotype[32]. We found no significant difference in the frequency of aneuploidy between cases with mutated and wild-type cohesin in the current DS-AMKL cohort. Many cases with mutated cohesin had completely normal karyotypes, except for constitutive trisomy 21, arguing against the hypothesis that aneuploidy has a major role in the pathogenesis of cohesin-mutated DS-AMKL (**Fig. 5a**).

CTCF mutations

Given the high frequency of cohesin mutations, new recurrent *CTCF* mutations were of particular interest because the functional interaction of cohesin and CTCF proteins has been of emerging interest in the long-range regulation of gene expression[26,30,33,34]. CTCF is a zinc-finger protein implicated in diverse regulatory functions, including transcriptional activation and/or repression, insulation, formation of chromatin barrier, imprinting and X-chromosome inactivation[35]. CTCF binds to target sequence elements and blocks the interaction of enhancers and promoters through DNA loop formation (insulator activity)[36], and several lines of evidence suggest that cohesin occupies CTCF-binding sites to contribute to the long-range regulation of gene expression by participating in the formation and stabilization of a repressive loop[26,37]. *CTCF* was mutated or deleted in ten DS-AMKL cases (20%), one TAM case (2%) and four non-DS-AMKL cases (21%), with seven mutations representing nonsense, frameshift or splice-site changes and an additional six alterations representing deletions resulting in the loss of protein function (**Fig. 4a,b**, **Supplementary Figs. 11** and **12b**, and **Supplementary Tables 7** and **8**). To our knowledge, this is the first report of frequent recurrent *CTCF* mutations in cancer, although rare mutations (occurring in approximately 2% of cases) have recently been reported in breast cancer sequencing[38].

Mutations in epigenetic regulators

EZH2, which encodes a catalytic subunit of the Polycomb repressive complex 2 (PRC2) that is responsible for di- and trimethylation of histone H3 lysine 27 (H3K27)[39], is another recurrent mutational target in DS-AMKL (**Table 1**). Inactivating mutations in *EZH2* have been reported in up to 13% of myelodysplastic syndromes and related chronic myeloid neoplasms[40]. Although rarely mutated in adult AML[41], *EZH2* represents one of the most frequently mutated and deleted genes in childhood AMKL, as we identified mutations or deletions in 16 of 49 DS-AMKL cases (33%) and in 3 of 19 non-DS-AMKL cases (16%) (**Fig. 4a,b**, **Supplementary Fig. 12c** and **Supplementary Tables 7** and **8**). No other PRC2 components were mutated, except for *SUZ12*, which was mutated in a single DS-AMKL case (**Fig. 4a** and **Supplementary Table 7**). Although frequent mutations in other epigenetic regulators, including in *TET2, IDH1* or *IDH2, DNMT3A* and *ASXL1*, are cardinal features of myeloid neoplasms in adults, we rarely found these mutations in DS-AMKL and non-DS-AMKL cases, only identifying occasional *DNMT3A* ($n = 1$), *ASXL1* ($n = 1$) and *BCOR* ($n = 2$) mutations in DS-AMKL (**Fig. 4a**).

KANSL1 (encoding KAT8 regulatory NSL complex subunit 1; also known as MSL1V1 or NSL1) represents a new recurrent mutational target in human cancer (**Table 1**), although haploinsufficiency of *KANSL1* through germline deletions or mutations has been implicated in a congenital disease known as 17q21.31 microdeletion syndrome (MIM 610443)[42,43]. We found heterozygous mutations in *KANSL1* in three DS-AMKL and three non-DS-AMKL cases, and most of these mutations were nonsense or frameshifts, leading to loss of protein function (**Fig. 4a** and **Supplementary Table 7**). KANSL1 protein is necessary and sufficient for the activity of the KAT8 (MOF) histone acetyltransferase complex, which is engaged in the acetylation of histone H4 lysine 16 (H4K16), leading to transcriptional activation. Loss of acetylation of H4K16 has been reported to be a common hallmark of human cancer, and other histone acetyltransferases for H4K16 have been reported to form recurrent fusion partners in leukemia, including MOZ and MORF[44], suggesting a role for compromised H4K16 acetylation by *KANSL1* mutations in leukemogenesis. Of interest, KANSL1 is also responsible for the acetylation of the TP53 tumor suppressor that is important for TP53-dependent transcriptional activation[45]. KAT8 also interacts with a histone H3 lysine 4 (H3K4) methyltransferase, MLL, and the interaction of MLL and KAT8 complexes facilitates the cooperative recruitment of both complexes to gene promoters and enhances transcription initiation at target genes[45]. Thus, impaired TP53 function and/or deregulated expression of MLL gene targets could also contribute to leukemogenesis by *KANSL1* mutations.

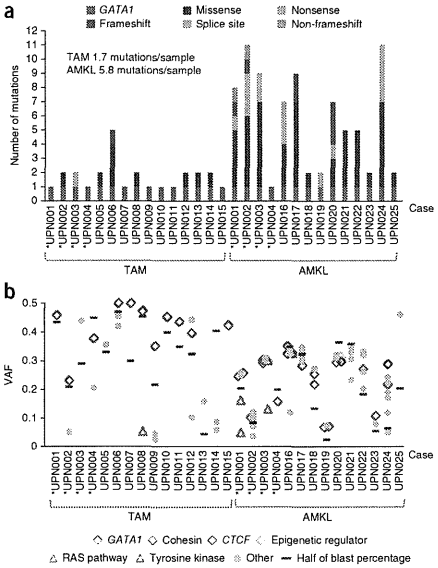

Figure 3 Somatic mutations detected by whole-exome sequencing of *Down syndrome–related myeloid disorders.* (**a**) *Number of validated* somatic mutations in 25 individuals with TAM and DS-AMKL identified by whole-exome sequencing. Paired samples are indicated by asterisks. The mutation rates per phase are given. (**b**) VAFs of individual mutations determined by deep sequencing, with VAFs adjusted for genomic copy numbers. Long indels of >3 bp were excluded from the analysis because their VAFs were difficult to accurately estimate. The VAF for each sample estimated on the basis of blast percentage is indicated by a purple horizontal bar.

Table 1 Recurrently mutated genes other than *GATA1* in DS-AMKL samples in whole-exome sequencing

Gene	Mutation type	RefSeq	Amino acid change	Nucleotide change	Sample (UPN) number
CTCF	Splice site	NM_006565	p.Gly318_splice	c.953–2A>G	016
CTCF	Frameshift	NM_006565	p.Asn314fs	c.940_941insAC	020
DCAF7	Missense	NM_005828	p.Leu340Phe	c.1018C>T	001
DCAF7	Missense	NM_005828	p.Leu340Phe	c.1018C>T	003
EZH2	Frameshift	NM_004456	p.710_716del	c.2129_2148delATCACAGGA TAGGTATTTTT	001
EZH2	Missense	NM_004456	p.Arg25Gln	c.74G>A	002
KANSL1	Frameshift	NM_001193466	p.Arg720fs	c.2159_2160insCG	020
KANSL1	Nonsense	NM_001193466	p.Arg462*	c.1384C>T	024
NRAS	Missense	NM_002524	p.Gly12Ser	c.34G>A	001
NRAS	Missense	NM_002524	p.Tyr64Cys	c.191A>G	001
NRAS	Missense	NM_002524	p.Gly12Ala	c.35G>C	003
RAD21	Nonsense	NM_006265	p.Arg139*	c.415A>T	001
RAD21	Frameshift	NM_006265	p.374_375del	c.1120_1124delTCTTT	002
RAD21	Missense	NM_006265	p.Leu611Arg	c.1832T>G	018
RAD21	Nonsense	NM_006265	p.Arg65*	c.193C>T	024
STAG2	Nonsense	NM_001042750	p.Arg604*	c.1810C>T	003
STAG2	Nonsense	NM_001042750	p.Arg216*	c.646C>T	019
STAG2	Frameshift	NM_001042750	p.Asn863fs	c.2588_2589insT	020
TP53	Nonsense	NM_000546	p.Glu68*	c.202G>T	002
TP53	Non-frameshift	NM_000546	p.157_162del	c.469_486delGTCCGCGCCA TGGCCATC	002

Other mutations in DS-AMKL

RAS pathway mutations are common in hematopoietic malignancies and other human cancers but have not to our knowledge been described in DS-AMKL. In the current cohort, we identified RAS pathway

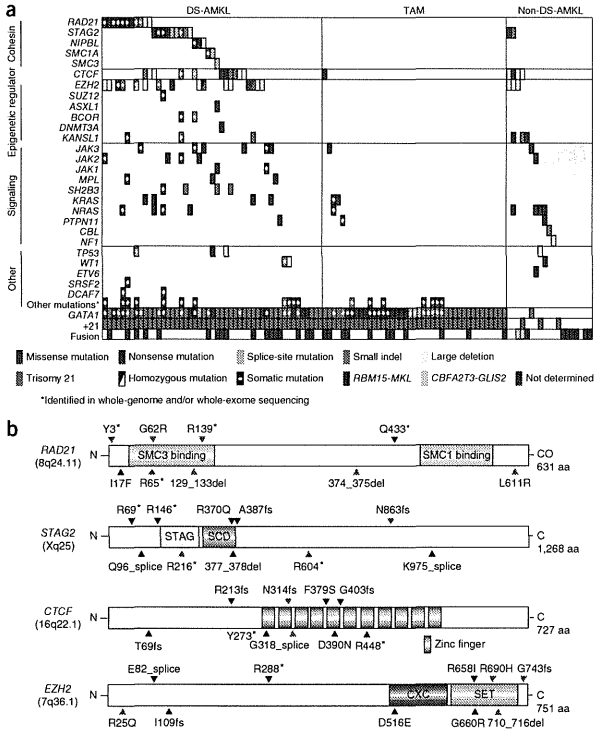

Figure 4 Driver mutations in Down syndrome–related myeloid disorders and non-DS-AMKL. (**a**) Driver mutations in 109 samples of 49 DS-AMKL, 41 TAM and 19 non-DS-AMKL cases. Types of mutations are distinguished by color. Each sample is also described in **Supplementary Table 12**. (**b**) Distribution of RAD21, STAG2, CTCF and EZH2 alterations. Alterations encoded by confirmed somatic mutations are indicated by red arrowheads.

mutations in the *NRAS*, *KRAS*, *PTPN11*, *NF1* and *CBL* genes in 8 DS-AMKL cases (16%) and 6 non-DS-AMKL cases (32%), but these mutations were rarely found in TAM cases (n = 3; 7%) (**Fig. 4a**). Tyrosine kinase and cytokine receptor mutations were also common in DS-AMKL. We found mutations in *JAK1*, *JAK2*, *JAK3*, *MPL* or *SH2B3* (*LNK*) in 17 DS-AMKL cases (35%) but rarely in TAM (n = 1) and non-DS-AMKL (n = 2) cases. We found no *FLT3* mutations in our cohort. The identified mutations were largely mutually exclusive. We found *JAK2* mutations in 4 DS-AMKL cases and 1 non-DS-AMKL case, including mutations encoding p.Val617Phe (n = 2), p.Leu611Ser (n = 1), p.Arg683Ser (n = 1) and p.Arg867Gln (n = 1); of these, *JAK2* mutations encoding p.Arg683Ser and p.Arg867Gln substitutions have been reported in acute lymphoblastic leukemia (ALL)[46,47] but not in myeloid malignancies[8,46]. Thus, we re-evaluated the diagnosis of AMKL in both UPN097 (p.Arg683Ser) and UPN023 (p.Arg867Gln), in whom the initial diagnosis of AMKL was strongly supported by typical surface marker expression of CD41, CD41b, CD117, CD13, CD33, CD34 and CD36 in UPN097 and of CD7, CD13, CD34, CD41a and CD42b in UPN023, together with characteristic cytomorphology. Similarly, the mutation encoding p.Leu611Ser was reported in both ALL[48] and polycythemia vera[49]. Thus, it seems that some *JAK2* mutations are involved in both myeloid and lymphoid leukemogenesis. As reported previously[10,11], *TP53* mutations were found in approximately 10% of DS-AMKL cases. Two identical somatic mutations found in the *DCAF7* gene (encoding p.Leu340Phe) might be interesting because the DCAF7 protein interacts with the DYRK1a kinase encoded within the Down syndrome critical region on chromosome 21 (ref. 50). DCAF7 has been shown to interact with DYRK1a through its N-terminal or C-terminal region, and the p.Leu340Phe substitution identified in our study was also located in the C-terminal domain. However, no additional mutation was detected in the extended cohort; therefore, the relevance of *DCAF7* remains to be determined.

Allelic burden of major recurrent mutations relative to *GATA1* mutations

We assessed intratumoral heterogeneity and the clonal origin of mutations by calculating the variant allele frequency (VAF) of each mutation relative to that of the *GATA1* mutation using deep sequencing. Mutations in cohesin components, *CTCF* and *EZH2* showed comparable VAFs to *GATA1* mutations (**Fig. 5b**), suggesting their role in the early stage of DS-AMKL development. In contrast, RAS pathway and other tyrosine kinases and cytokine receptor mutations showed significantly lower VAFs than corresponding *GATA1* mutations (P = 0.0001) (**Fig. 5b**), indicating that they are more likely to represent subclonal mutations, which were typically preceded by mutations in cohesin components, *CTCF* and *EZH2* and were involved in the evolution of multiple DS-AMKL subclones. Although RAS and JAK pathways activated by gene mutations represent potentially druggable targets and several promising compounds are currently available, this observation may largely preclude the efficient use of such compounds in eradicating founding DS-AMKL clones.

Distinct genetic features of Down syndrome– and non–Down syndrome–related AMKL

Despite their morphological similarities, both forms of AMKL in childhood are characterized by distinctive genetic features. According to the current study and a recent report of integrated analysis of non-DS-AMKL[22], *GATA1* mutations and trisomy 21 are less common in non-DS-AMKL than in DS-AMKL cases (**Fig. 4a** and **Supplementary Table 9**). In our series, DS-AMKL was characterized by high frequencies of mutations in the cohesin complex, *EZH2* and other epigenetic regulators, as well as in JAK family kinases, which were less

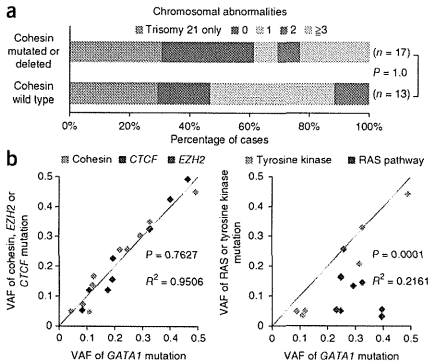

Figure 5 Relationship of cohesin mutations with karyotypes and comparison of mutation loads between major gene targets in DS-AMKL and *GATA1*. (**a**) The number of chromosomal abnormalities is compared between cases with and without cohesin mutations or deletions for DS-AMKL cases. Zero signifies chromosomal abnormalities without change in chromosome count, such as partial amplification or deletion of the chromosomal region or balanced translocation. (**b**) Diagonal plots of copy number–adjusted VAFs comparing coexisting *GATA1* and other pathway mutations, including cohesin, *CTCF*, *EZH2*, tyrosine kinase and the RAS pathway mutations, as indicated by color.

common mutational targets in non-DS-AMKL. Previous studies identified recurrent *CBFA2T3-GLIS2* and *RBM15-MKL* gene fusions in non-DS-AMKL, which were found in 27% and 15.2% of non-DS-AMKL cases, respectively[22,51], whereas these fusions were not detected in DS-AMKL cases in another report (n = 10 cases)[23]. Similarly, in the current cohort, RT-PCR analysis identified 2 *CBFA2T3-GLIS2* and 3 *RBM15-MKL* fusion genes in 19 non-DS-AMKL cases but not in TAM and DS-AMKL cases (**Fig. 4a** and **Supplementary Table 10**), illustrating the genetic differences between DS-AMKL and non-DS-AMKL. In addition, our RNA sequencing of the current cases (n = 17) (**Supplementary Table 11**) also showed no *CBFA2T3-GLIS2* and *RBM15-MKL* fusions.

DISCUSSION

Whole-genome and/or whole-exome analyses and follow-up targeted sequencing identified several new aspects of the pathogenesis of Down syndrome–related myeloid proliferation. First, the initial TAM phase was characterized by a paucity of somatic mutations. The mean number of non-silent mutations per sample (1.7; range of 1–5) was surprisingly small compared with that reported in other human cancers (**Supplementary Fig. 13**), in line with a recent report that identified 1.2 (range of 1–2) mutations per sample by whole-exome sequencing in 5 TAM samples[52]. In addition to reporting a low somatic mutation frequency in their initial TAM phase, Nikolaev *et al.*[52] also reported accumulation of somatic mutations (including single cases of *SMC3* and *EZH2* mutation) during progression from TAM to DS-AMKL. Excluding common *GATA1* mutations, we identified no other recurrent mutations, with only 0.7 non-silent mutations per case, indicating that TAM could be caused by a single acquired *GATA1* mutation in addition to constitutive trisomy 21.

Intratumoral heterogeneity was evident not only in the DS-AMKL phase but also at the initial diagnosis of TAM, and subsequent DS-AMKL originated from one of the multiple subclones present in the TAM phase, usually representing the progeny of the largest subpopulation. In most cases, the DS-AMKL clone was accompanied by newly acquired driver mutations not shared by the original TAM population, generating a unique landscape of gene mutations in DS-AMKL, which was characterized by high mutational frequencies in cohesin or *CTCF* (65%), other epigenetic regulators (45%), and RAS or signal-transducing molecules (47%) (**Fig. 4a**). Tumor recurrence or evolution has not to our knowledge been characterized by the distinct gene mutations in greater detail than in the present study. In total, 44 of the 49 DS-AMKL cases had additional mutations beyond those in *GATA1* (**Fig. 4a**), even though there was a clear limitation on capturing mutations using the targeted sequencing approach.

The very high frequency of cohesin (53%) and *EZH2* (33%) mutations and deletions in DS-AMKL but not in TAM or non-DS-AMKL cases was noteworthy because the reported mutation rates of cohesin and *EZH2* in adult AML and other human cancers remain approximately 10% (refs. 14,40,41), underscoring a major role for these mutations in the pathogenesis of DS-AMKL. The leukemogenic mechanism of mutated cohesin remains elusive, and frequent *CTCF* mutations also need further evaluation to characterize their possible cooperative role with cohesin mutations[26,30,33,34]. To our knowledge, *KANSL1* mutations have not been reported previously and represent a new recurrent mutational target in human cancer, although their functional impact on AMKL development remains unknown. Evaluation of the allelic burden of these mutations by deep sequencing disclosed a clonal hierarchy among different driver mutations in which clonal mutations in cohesin, *CTCF* and epigenetic regulators frequently preceded subclonal mutations in RAS and signal transduction molecules.

In conclusion, Down syndrome–related myeloid proliferation is shaped by multiple rounds of acquisition of new mutations and clonal selection, which are initiated by a *GATA1* mutation in the TAM phase and further driven by mutation in cohesin or *CTCF*, *EZH2* or other epigenetic regulators, and RAS or signal-transducing molecules, leading to AMKL. DS-AMKL and non-DS-AMKL showed similar phenotypes but had distinct genetic features, which may underlie their different clinical characteristics.

URLs. European Genome-phenome Archive (EGA), https://www.ebi.ac.uk/ega/; EBCall, https://github.com/friend1ws/EBCall; Catalogue of Somatic Mutations in Cancer (COSMIC), http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/; PubMed, http://www.ncbi.nlm.nih.gov/pubmed/; UCSC Genome Browser, http://genome.ucsc.edu/; Integrative Genomics Viewer, http://www.broadinstitute.org/igv/; DNACopy, http://biostatistics.oxfordjournals.org/content/5/4/557.full.pdf; Genomon-fusion (in Japanese), http://genomon.hgc.jp/rna/.

METHODS

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

Accession codes. Sequencing data have been deposited in the European Genome-phenome Archive (EGA) under accession EGAS00001000546.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank Y. Mori, M. Nakamura, O. Hagiwara and N. Mizota for their technical assistance. This work was supported by the Research on Measures for Intractable

Diseases Project and Health and Labor Sciences Research grants (Research on Intractable Diseases) from the Ministry of Health, Labour and Welfare, by Grants-in-Aid from the Ministry of Health, Labor and Welfare of Japan and KAKENHI (22134006, 23249052, 23118501, 23390266 and 25461579) and by the Japan Society for the Promotion of Science (JSPS) through the Funding Program for World-Leading Innovative Research and Development on Science and Technology (FIRST Program), initiated by the Council for Science and Technology Policy (CSTP) and research grants from the Japan Science and Technology Agency CREST.

AUTHOR CONTRIBUTIONS

Y.O., Y. Shiraishi, A.S.-O., K.C., H.T. and S.M. performed bioinformatics analyses of the resequencing data. M.S., A.S.-O., Y. Sato, A.H. and H.M. performed microarray experiments and analyses. R.K. and A.H. performed RT-PCR analyses. M.P., K. Terui, R.W., D.H., K.N., H.K., K. Tsukamoto, S.A., K. Kawakami, K. Kato, R.N., S.I., Y.H., S.K. and E.I. collected specimens and were involved in planning the project. K.Y., T.T., H.S., Y.N. and N.S. processed and analyzed genetic materials, prepared the library and performed sequencing. K.Y., T.T., Y.O., A.K. and S.O. generated figures and tables. E.I. and S.O. led the entire project. K.Y. and S.O. wrote the manuscript. All authors participated in discussions and interpretation of the data and results.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Khan, I., Malinge, S. & Crispino, J. Myeloid leukemia in Down syndrome. *Crit. Rev. Oncog.* **16**, 25–36 (2011).
2. Massey, G.V. *et al.* A prospective study of the natural history of transient leukemia (TL) in neonates with Down syndrome (DS): Children's Oncology Group (COG) study POG-9481. *Blood* **107**, 4606–4613 (2006).
3. Muramatsu, H. *et al.* Risk factors for early death in neonates with Down syndrome and transient leukaemia. *Br. J. Haematol.* **142**, 610–615 (2008).
4. Klusmann, J.H. *et al.* Treatment and prognostic impact of transient leukemia in neonates with Down syndrome. *Blood* **111**, 2991–2998 (2008).
5. Xu, G. *et al.* Frequent mutations in the *GATA-1* gene in the transient myeloproliferative disorder of Down syndrome. *Blood* **102**, 2960–2968 (2003).
6. Wechsler, J. *et al.* Acquired mutations in *GATA1* in the megakaryoblastic leukemia of Down syndrome. *Nat. Genet.* **32**, 148–152 (2002).
7. Walters, D.K. *et al.* Activating alleles of *JAK3* in acute megakaryoblastic leukemia. *Cancer Cell* **10**, 65–75 (2006).
8. Malinge, S. *et al.* Activating mutations in human acute megakaryoblastic leukemia. *Blood* **112**, 4220–4226 (2008).
9. Blink, M. *et al.* Frequency and prognostic implications of *JAK 1–3* aberrations in Down syndrome acute lymphoblastic and myeloid leukemia. *Leukemia* **25**, 1365–1368 (2011).
10. Hama, A. *et al.* Molecular lesions in childhood and adult acute megakaryoblastic leukaemia. *Br. J. Haematol.* **156**, 316–325 (2012).
11. Malkin, D., Brown, E.J. & Zipursky, A. The role of p53 in megakaryocyte differentiation and the megakaryocytic leukemias of Down syndrome. *Cancer Genet. Cytogenet.* **116**, 1–5 (2000).
12. Hussein, K. *et al.* MPL^{W515L} mutation in acute megakaryoblastic leukaemia. *Leukemia* **23**, 852–855 (2009).
13. Greenman, C. *et al.* Patterns of somatic mutation in human cancer genomes. *Nature* **446**, 153–158 (2007).
14. Welch, J.S. *et al.* The origin and evolution of mutations in acute myeloid leukemia. *Cell* **150**, 264–278 (2012).
15. Ding, L. *et al.* Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* **481**, 506–510 (2012).
16. Creutzig, U. *et al.* Diagnosis and management of acute myeloid leukemia in children and adolescents: recommendations from an international expert panel. *Blood* **120**, 3187–3205 (2012).
17. Swerdlow, S.H., Jaffe, E.S. & International Agency for Research on Cancer & World Health Organization *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues* (International Agency for Research on Cancer, Lyon, France, 2008).
18. Wu, C. *et al.* BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol.* **10**, R130 (2009).
19. Bourquin, J.P. *et al.* Identification of distinct molecular phenotypes in acute megakaryoblastic leukemia by gene expression profiling. *Proc. Natl. Acad. Sci. USA* **103**, 3339–3344 (2006).
20. Mercher, T. *et al.* Involvement of a human gene related to the *Drosophila spen* gene in the recurrent t(1;22) translocation of acute megakaryocytic leukemia. *Proc. Natl. Acad. Sci. USA* **98**, 5776–5779 (2001).
21. Ma, Z. *et al.* Fusion of two novel genes, *RBM15* and *MKL1*, in the t(1;22)(p13;q13) of acute megakaryoblastic leukemia. *Nat. Genet.* **28**, 220–221 (2001).
22. Gruber, T.A. *et al.* An inv(16)(p13.3q24.3)-encoded CBFA2T3-GLIS2 fusion protein defines an aggressive subtype of pediatric acute megakaryoblastic leukemia. *Cancer Cell* **22**, 683–697 (2012).
23. Thiollier, C. *et al.* Characterization of novel genomic alterations and therapeutic approaches using acute megakaryoblastic leukemia xenograft models. *J. Exp. Med.* **209**, 2017–2031 (2012).
24. Gruber, S., Haering, C.H. & Nasmyth, K. Chromosomal cohesin forms a ring. *Cell* **112**, 765–777 (2003).
25. Nasmyth, K. & Haering, C.H. Cohesin: its roles and mechanisms. *Annu. Rev. Genet.* **43**, 525–558 (2009).
26. Wendt, K.S. *et al.* Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**, 796–801 (2008).
27. Ström, L. *et al.* Postreplicative formation of cohesion is required for repair and induced by a single DNA break. *Science* **317**, 242–245 (2007).
28. Watrin, E. & Peters, J.M. The cohesin complex is required for the DNA damage-induced G2/M checkpoint in mammalian cells. *EMBO J.* **28**, 2625–2635 (2009).
29. Dorsett, D. *et al.* Effects of sister chromatid cohesion proteins on *cut* gene expression during wing development in *Drosophila*. *Development* **132**, 4743–4753 (2005).
30. Parelho, V. *et al.* Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132**, 422–433 (2008).
31. Solomon, D.A. *et al.* Mutational inactivation of *STAG2* causes aneuploidy in human cancer. *Science* **333**, 1039–1043 (2011).
32. Forestier, E. *et al.* Cytogenetic features of acute lymphoblastic and myeloid leukemias in pediatric patients with Down syndrome: an iBFM-SG study. *Blood* **111**, 1575–1583 (2008).
33. Rubio, E.D. *et al.* CTCF physically links cohesin to chromatin. *Proc. Natl. Acad. Sci. USA* **105**, 8309–8314 (2008).
34. Stedman, W. *et al.* Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. *EMBO J.* **27**, 654–666 (2008).
35. Ohlsson, R., Bartkuhn, M. & Renkawitz, R. CTCF shapes chromatin by multiple mechanisms: the impact of 20 years of CTCF research on understanding the workings of chromatin. *Chromosoma* **119**, 351–360 (2010).
36. Phillips, J.E. & Corces, V.G. CTCF: master weaver of the genome. *Cell* **137**, 1194–1211 (2009).
37. Wendt, K.S. & Peters, J.M. How cohesin and CTCF cooperate in regulating gene expression. *Chromosome Res.* **17**, 201–214 (2009).
38. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
39. Cao, R. *et al.* Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**, 1039–1043 (2002).
40. Ernst, T. *et al.* Inactivating mutations of the histone methyltransferase gene *EZH2* in myeloid disorders. *Nat. Genet.* **42**, 722–726 (2010).
41. Patel, J.P. *et al.* Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. *N. Engl. J. Med.* **366**, 1079–1089 (2012).
42. Koolen, D.A. *et al.* Mutations in the chromatin modifier gene *KANSL1* cause the 17q21.31 microdeletion syndrome. *Nat. Genet.* **44**, 639–641 (2012).
43. Zollino, M. *et al.* Mutations in *KANSL1* cause the 17q21.31 microdeletion syndrome phenotype. *Nat. Genet.* **44**, 636–638 (2012).
44. Yang, X.J. The diverse superfamily of lysine acetyltransferases and their roles in leukemia and other diseases. *Nucleic Acids Res.* **32**, 959–976 (2004).
45. Li, X., Wu, L., Corsa, C.A., Kunkel, S. & Dou, Y. Two mammalian MOF complexes regulate transcription activation by distinct mechanisms. *Mol. Cell* **36**, 290–301 (2009).
46. Bercovich, D. *et al.* Mutations of *JAK2* in acute lymphoblastic leukaemias associated with Down's syndrome. *Lancet* **372**, 1484–1492 (2008).
47. Mullighan, C.G. *et al.* JAK mutations in high-risk childhood acute lymphoblastic leukemia. *Proc. Natl. Acad. Sci. USA* **106**, 9414–9418 (2009).
48. Kratz, C.P. *et al.* Mutational screen reveals a novel JAK2 mutation, L611S, in a child with acute lymphoblastic leukemia. *Leukemia* **20**, 381–383 (2006).
49. Nussenzveig, R.H. *et al.* Detection of *JAK2* mutations in paraffin marrow biopsies by high resolution melting analysis: identification of L611S alone and in *cis* with V617F in polycythemia vera. *Leuk. Lymphoma* **53**, 2479–2486 (2012).
50. Miyata, Y. & Nishida, E. DYRK1A binds to an evolutionarily conserved WD40-repeat protein WDR68 and induces its nuclear translocation. *Biochim. Biophys. Acta* **1813**, 1728–1739 (2011).
51. de Rooij, J.D. *et al.* *NUP98*/*JARID1A* is a novel recurrent abnormality in pediatric acute megakaryoblastic leukemia with a distinct *HOX* gene expression pattern. *Leukemia* doi:10.1038/leu.2013.87 (27 March 2013).
52. Nikolaev, S.I. *et al.* Exome sequencing identifies putative drivers of progression of transient myeloproliferative disorder to AMKL in infants with Down Syndrome. *Blood* **122**, 554–561 (2013).
53. Krzywinski, M. *et al.* Circos: an information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645 (2009).

ONLINE METHODS

Subjects and samples. Genomic DNA from 84 individuals with Down syndrome–related myeloid disorders (41 samples from the TAM phase and 49 from the AMKL phase) and 19 with non-DS-AMKL were analyzed by whole-genome and/or-whole-exome and/or targeted deep sequencing. In six cases with Down syndrome–related myeloid disorders, samples were collected from both the TAM and AMKL phases. RNA sequencing was also performed for 12 of the 49 DS-AMKL cases and for 5 additional DS-AMKL cases. RNA samples were also available for RT-PCR analysis from 30 cases with TAM, 32 cases with DS-AMKL and 15 cases with non-DS-AMKL. Written informed consent was obtained from each subject's parents before sample collection (**Supplementary Note**). This study was approved by the Ethics Committees of the University of Tokyo according to the Helsinki convention. *GATA1* mutations were detected by Sanger sequencing of all TAM and DS-AMKL samples according to the previously described procedure[5]. Detailed information on subjects and samples is provided in **Supplementary Tables 1**, **4**, **11** and **12**. Tumor DNA was extracted from bone marrow– or peripheral blood–derived mononuclear cells at diagnosis. Genomic DNA samples from peripheral blood from subjects in remission or from nail tissues at diagnosis were used as germline controls. Genomic DNA was extracted using a QIAamp DNA Blood Mini kit and a QIAamp DNA Investigator kit (Qiagen). Total RNA was extracted using the RNeasy kit (Qiagen) with RNase-free DNase (Qiagen).

Whole-genome sequencing. DNA samples were processed for whole-exome sequencing using NEBNext DNA sample Prep Reagent (New England Biolabs) according to the modified Illumina protocol. Sequence data were generated on the Illumina HiSeq 2000 platform in 100-bp paired-end reads. Data processing and variant calling were performed as described previously[54]. All candidate variants were validated by deep sequencing.

Validation and quantitative measurements of the frequencies of mutant alleles by deep sequencing. Individual mutation sites were amplified by genomic PCR using primers tagged with NotI cleavage sites and subjected to high-throughput sequencing as described previously[55], except that target DNA was not pooled. Deep sequencing was performed using the MiSeq or HiSeq 2000 platform. Data processing was performed according to the previously described method with minor modifications[55]. Briefly, each read was aligned to a set of PCR-amplified target sequences using BLAT[56], and dichotomic variant alleles were differentially enumerated. For indels, individual reads were first aligned to each of the wild-type and indel sequences and then assigned to the one to which better alignment was obtained in terms of the number of matched bases. Each SNV and indel whose VAF in the tumor sample was equal to or greater than 2.0% and significantly higher than the frequency in the germline sample was adopted as a somatic mutation. The error size for estimated VAFs was evaluated by assuming binomial distributions in deep sequencing, which were confirmed by observed allele frequencies at heterozygous SNPs in normal DNA samples (**Supplementary Fig. 14a**), in which the variance (σ^2) ranged from $4.0–11.0 \times 10^{-4}$ (**Supplementary Fig. 14b**).

Clustering analysis of mutations. To identify the chronological behavior of the structure of the tumor subpopulation for the TAM and AMKL phases, somatic mutations detected in both phases by whole-genome sequencing were clustered according to their VAFs as measured by deep sequencing. Copy number–adjusted deep sequencing data, in which the VAFs of genes on the X chromosome in male cases or in regions of uniparental disomy were halved, were subjected to unsupervised clustering. Six mutations located in amplified or deleted genomic regions were excluded from the analysis. Long indels of >3 bp, except for those affecting key genes such as *GATA1* and *RAD21*, and mutations in repetitive regions were excluded from the analysis because their VAFs could tend to be underestimated.

All validated mutations were grouped into three categories according to the following criteria: (i) mutations found only in TAM (VAF in AMKL < 0.02), (ii) mutations found only in AMKL (VAF in TAM < 0.02) and (iii) mutations found in both TAM and AMKL (VAF in TAM > 0.02 and VAF in AMKL > 0.02). Clustering of mutations in each category was performed using Mclust, provided as an R package, on the basis of the VAFs of the mutations in the TAM and AMKL phases, where one-dimensional clustering of mutations in categories (i) and (ii) was performed on the basis of the homoscedastic model and two-dimensional clustering was performed for mutations in category (iii) on the basis of the ellipsoidal model. The most appropriate number of clusters was determined by using the Bayesian information criterion (BIC) score. Singleton points identified by this algorithm were regarded as outliers. Clonal subpopulations within tumors were also evaluated by kernel density analysis (**Supplementary Fig. 5**), where we drew kernel density estimate plots for the VAFs of validated variants using the density function in R.

Whole-exome sequencing and detection of somatic mutations. Exome capture was performed using SureSelect Human All Exon V3 or V4 (Agilent Technologies) or the TruSeq Exome Enrichment kit (Illumina). Enriched exome fragments were then subjected to massively parallel sequencing using the Genome Analyzer IIx or HiSeq 2000 platform (Illumina). Candidate somatic mutations were detected using our in-house pipeline EBCall (Empirical Bayesian mutation Calling; see URLs)[57]. All candidates were validated by Sanger sequencing or independent deep sequencing.

PCR-based targeted deep sequencing. Deep sequencing of *DCAF7*, *EED*, *JAK1*, *JAK3*, *KANSL1*, *SH2B3*, and *SUZ12* was performed using the primers tagged with NotI cleavage sites whose sequences are listed in **Supplementary Table 6**. Data processing and variant calling were performed as described previously[58]. All candidate variants were validated by Sanger sequencing or independent deep sequencing using non-amplified DNA.

Targeted deep sequencing. In total, 39 gene targets were exhaustively examined for mutations in all 109 cases using deep sequencing (**Supplementary Table 5**). Genomic DNA (1–1.5 μg) from bone marrow–derived mononuclear cells or peripheral blood was enriched for target exons using a SureSelect custom kit (Agilent Technologies) designed to capture all of the coding exons from the 39 target genes, and high-throughput sequencing was performed on the enriched targets using the HiSeq 2000 platform with a standard 100-bp paired-end read protocol. Sequencing reads were aligned to hg19 using Burrows-Wheeler Aligner (BWA) version 0.5.8 with default parameters. The allele frequencies of SNVs and indels were calculated at each genomic position by enumerating the relevant reads with SAMtools[59]. Initially, all variants showing VAF > 0.02 were extracted and annotated using ANNOVAR[60] for further consideration if they were found in >6 reads out of >10 total reads and appeared in both plus- and minus-strand reads. For the cases for which no germline DNA was available, relevant somatic mutations were called by eliminating the following entries, unless they were registered in the Catalogue of Somatic Mutations in Cancer (COSMIC) v60 (ref. 61) or reported as somatic mutations in PubMed: (i) synonymous variants and those having ambiguous (unknown) annotations, (ii) known SNPs in public and private databases, including dbSNP131, the 1000 Genomes Project as of 23 November 2010 and our in-house database, (iii) sequencing or mapping errors, (iv) all missense SNVs with allele frequencies of 0.45–0.55 and (v) variants localized to duplicated regions found in SegDups of the UCSC Genome Browser. To eliminate sequencing errors in category (iii), we excluded all variants found in 31 normal Japanese samples at, on average, allele frequency > 0.25. Mapping errors were removed by visual inspection with the Integrative Genomics Viewer browser[62]. All candidate variants were validated by Sanger sequencing or independent deep sequencing.

Calculation of copy numbers for target exons. Letting $d_j^{i,s}$ be the sequencing depth at the ith nucleotide of the jth exon in sample s, the standardized depth of the jth exon is calculated as

$$D_j^s = k_s \sum_i d_i^{j,s}$$

where k_s is determined to satisfy

$$k_0 = \sum_j D_j^s$$

for a fixed constant k_0 (for example, $k_0 = 1$). The correlation coefficient ($R = R^{s,t}$) between two vectors D_i^s and D_i^t was calculated, where D_i^s and D_i^t represent the depth for a given sample (sample s) and each of the 443

samples (sample t), analyzed for other projects, with completely normal copy numbers in array–comparative genomic hybridization (aCGH; $t = 1, 2, 3, \ldots, 443$), respectively, through which a total of m_0 (= 12) control samples showing the largest R values were selected (T_m; $m = 1, 2, 3, \ldots, m_0$) and used for copy number calculation. The copy number of the ith target exon of sample s (Cn_i^s) was calculated as

$$Cn_i^s = D_i^s / \hat{D}_i^s$$

where $\hat{D}_i^s$ was calculated by averaging m_0 samples by

$$\hat{D}_i^s = \sum_{m=1}^{m_0} D_i^{T_m} / m_0$$

Copy numbers were calculated for exons with mean depth of >500. Circular binary segmentation was also used to identify discrete copy number segments using DNACopy (see URLs); segmented copy number ($\widehat{Cn}_i^s$) was defined for the ith exon of sample s. The distribution of $\widehat{Cn}_i^s$ was calculated for all samples, and exons showing $|\widehat{Cn}_i^s - E(\widehat{Cn}_i^s)| > 4$ s.d. were considered to have copy number losses or gains.

Screening for *CBFA2T3-GLIS2* and *RBM15-MKL1* fusion genes. *CBFA2T3-GLIS2* and *RBM15-MKL1* fusion genes were screened by RT-PCR[22,63]. Primer sequences are given in **Supplementary Table 13**. PCR amplification was performed by 40 cycles at 94 °C for 2 min, 60 °C for 30 s and 68 °C for 1 min, followed by denaturation at 94 °C for 2 min and extension at 68 °C for 7 min.

SNP array analyses. All tumor samples subjected to whole-exome sequencing were also analyzed for copy number alterations using SNP arrays (Affymetrix GeneChip Human Mapping 250K NspI Array or Genome-Wide Human SNP Array 6.0) as described previously[10,64,65].

RT-PCR analysis of *STAG2* and *CTCF* transcripts. To confirm abnormal splicing of *CTCF* in UPN016 and UPN071 and that of *STAG2* in UPN067, RT-PCR were performed using cDNA derived from each subject, with cDNA from CMK11-5 (DS-AMKL–derived cell line with no known mutations in both genes) used as a control (**Supplementary Fig. 11**). Primer sequences are given in **Supplementary Table 14**. Total RNA (1 μg) was subjected to reverse transcription using M-MLV reverse transcriptase (Invitrogen) according to the manufacturer's instructions. Electrophoresis was performed using Experion (Bio-Rad).

RNA sequencing. Detailed information on samples is provided in **Supplementary Table 11**. Library preparation and sequencing were performed as described previously[54]. Fusion transcripts were detected using Genomon-fusion.

Gene expression analysis of recurrently mutated genes. Expression data for the recurrently mutated genes in whole-exome sequencing were retrieved from the BioGPS database[18] for normal hematopoietic cells, including whole bone marrow, $CD33^+$ myeloid cells, $CD34^+$ cells, $CD19^+$ B cells and $CD4^+$ T cells, and from published data[19] and our RNA sequencing data for DS-AMKL samples.

Statistical analysis. The number of non-silent mutations identified by whole-exome sequencing in TAM and DS-AMKL samples (**Fig. 2a**) and the number of chromosome abnormalities in DS-AMKL cases with and without cohesin mutations or deletions (**Fig. 5a**) were compared using the Mann-Whitney U test. The difference in VAF between two mutations (**Fig. 5b**) was tested by Wilcoxon signed-rank test.

54. Sato, Y. *et al.* Integrated molecular analysis of clear-cell renal cell carcinoma. *Nat. Genet.* **45**, 860–867 (2013).
55. Yoshida, K. *et al.* Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **478**, 64–69 (2011).
56. Kent, W.J. BLAT—the BLAST-like alignment tool. *Genome Res.* **12**, 656–664 (2002).
57. Shiraishi, Y. *et al.* An empirical Bayesian framework for somatic mutation detection from cancer genome sequencing data. *Nucleic Acids Res.* **41**, e89 (2013).
58. Sakaguchi, H. *et al.* Exome sequencing identifies secondary mutations of *SETBP1* and *JAK3* in juvenile myelomonocytic leukemia. *Nat. Genet.* **45**, 937–941 (2013).
59. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
60. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
61. Forbes, S.A. *et al.* COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* **39**, D945–D950 (2011).
62. Robinson, J.T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
63. Torres, L. *et al.* Acute megakaryoblastic leukemia with a four-way variant translocation originating the *RBM15-MKL1* fusion gene. *Pediatr. Blood Cancer* **56**, 846–849 (2011).
64. 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).
65. Yamamoto, G. *et al.* Highly sensitive method for genomewide detection of allelic composition in nonpaired, primary tumor specimens by use of Affymetrix single-nucleotide-polymorphism genotyping microarrays. *Am. J. Hum. Genet.* **81**, 114–126 (2007).

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

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

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

Brief Report

MYELOID NEOPLASIA

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

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

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

Key Points

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

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

Introduction

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

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

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

Study design

This study was approved by the Ethics Committee of the Hirosaki University Graduate School of Medicine. All animal experiments were approved by the Institutional Animal Experiment Committee of Tohoku University. All clinical samples were obtained with informed consent from the parents of all patients with TAM in accordance with the Declaration of Helsinki. Additional information can be found in the supplemental text on the *Blood* website.

Results and discussion

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

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

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

The online version of this article contains a data supplement.

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

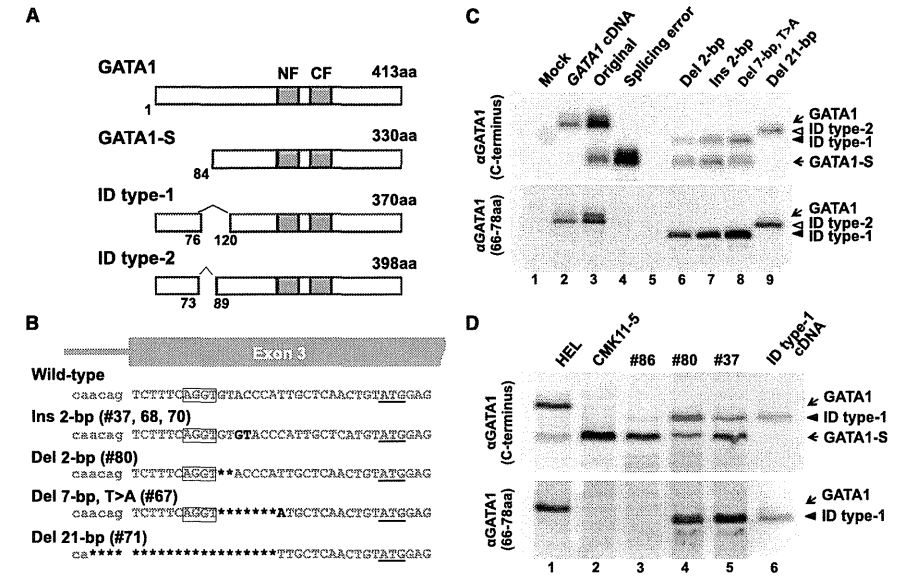

Figure 1. GATA1 mutant proteins with internal deletions. (A) A schema of mutant GATA1 proteins observed in patients with TAM. The amino acid sequence of GATA1-ID proteins was deduced from the sequence of *GATA1* cDNA obtained from patients with TAM. Dark boxes indicate *N*-finger (NF) and C-finger (CF)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

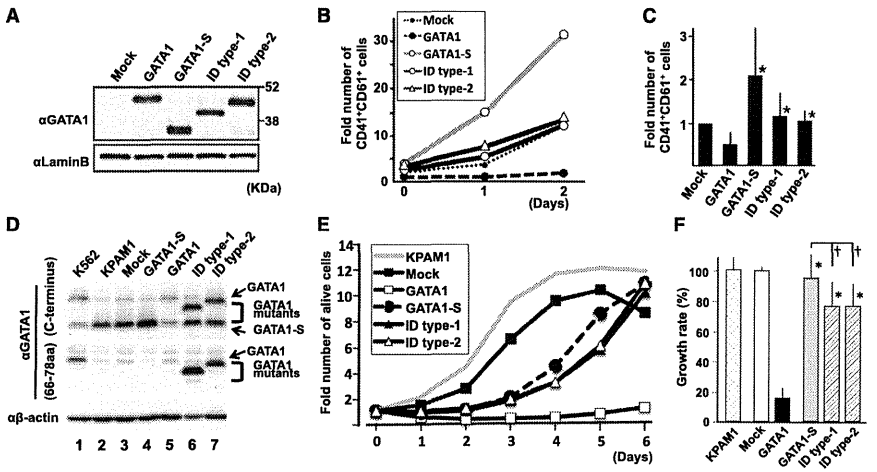

Figure 2. GATA1 ID proteins showed restricted antiproliferative activity. (A) Expression of GATA1 and GATA1 mutant proteins in cultured megakaryocytes at day 0, using an antibody against the C terminus of GATA1. The amount of protein loaded was quantified using an anti-Lamin B antibody on the same membrane.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

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

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

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

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

Our newly identified GATA1-ID mutants have highlighted a much narrower set of sequences responsible for the pathogenesis of TAM than has previously been suggested by the loss of the *N*-terminal sequence, as in GATA1-S. The missing region identified by the GATA1-ID proteins contains a consensus motif (LxCxE, amino acids 81-85) essential for the interaction with pRb,[17] which is also lost in GATA1-S. Interaction with hypophosphorylated pRb-E2F complex has been reported to be important for GATA1 to support the normal proliferation and differentiation of erythroid progenitors.[17] Consistent with this notion, GATA1-S failed to repress E2F activation, which was followed by activation of mTOR signaling in the GATA1-S fetal megakaryocytes and DS-AMKL cells.[18] Because the protein levels of cyclin D1 and $p27^{Kip}$ are reciprocally regulated by the mTOR pathway, and thereby cause pRb to be phosphorylated,[19] cell-cycle progression in response to the mTOR pathway may be potentiated by the enfeebled function of LxCxE motif of GATA1-S. Thus, we are one step closer to a molecular understanding of GATA1-related leukemias.

Acknowledgments

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

Authorship

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

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

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

References

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

Congenital amegakaryocytic thrombocytopenia iPS cells exhibit defective MPL-mediated signaling

Shinji Hirata,[1] Naoya Takayama,[1] Ryoko Jono-Ohnishi,[1] Hiroshi Endo,[1,2] Sou Nakamura,[1] Takeaki Dohda,[1] Masanori Nishi,[3] Yuhei Hamazaki,[3] Ei-ichi Ishii,[4] Shin Kaneko,[1,2] Makoto Otsu,[2] Hiromitsu Nakauchi,[2] Shinji Kunishima,[5] and Koji Eto[1,2]

[1]Clinical Application Department, Center for iPS Cell Research and Application, Kyoto University, Kyoto, Japan. [2]Laboratory of Stem Cell Therapy, Center for Stem Cell Biology and Regenerative Medicine, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan. [3]Department of Pediatrics, University of Saga School of Medicine, Saga, Japan. [4]Department of Pediatrics, University of Ehime School of Medicine, Tohno, Japan. [5]Department of Advanced Diagnosis, Clinical Research Center, National Hospital Organization Nagoya Medical Center, Nagoya, Japan.

Congenital amegakaryocytic thrombocytopenia (CAMT) is caused by the loss of thrombopoietin receptor–mediated (MPL-mediated) signaling, which causes severe pancytopenia leading to bone marrow failure with onset of thrombocytopenia and anemia prior to leukopenia. Because *Mpl*$^{-/-}$ mice do not exhibit the human disease phenotype, we used an in vitro disease tracing system with induced pluripotent stem cells (iPSCs) derived from a CAMT patient (CAMT iPSCs) and normal iPSCs to investigate the role of MPL signaling in hematopoiesis. We found that MPL signaling is essential for maintenance of the CD34$^+$ multipotent hematopoietic progenitor (MPP) population and development of the CD41$^+$GPA$^+$ megakaryocyte-erythrocyte progenitor (MEP) population, and its role in the fate decision leading differentiation toward megakaryopoiesis or erythropoiesis differs considerably between normal and CAMT cells. Surprisingly, complimentary transduction of *MPL* into normal or CAMT iPSCs using a retroviral vector showed that MPL overexpression promoted erythropoiesis in normal CD34$^+$ hematopoietic progenitor cells (HPCs), but impaired erythropoiesis and increased aberrant megakaryocyte production in CAMT iPSC–derived CD34$^+$ HPCs, reflecting a difference in the expression of the transcription factor *FLI1*. These results demonstrate that impaired transcriptional regulation of the MPL signaling that normally governs megakaryopoiesis and erythropoiesis underlies CAMT.

Introduction

It has been well documented that thrombopoietin (TPO) plays an essential role in the self-renewal of HSCs (1, 2) and in the production of megakaryocytes (MKs) and platelets (3, 4). TPO acts via defined signaling pathways that include JAK-STAT, MAPK-ERK1/2, and PI3K-v-akt murine thymoma viral oncogene homolog 1 (PI3K-AKT) (5, 6). Congenital amegakaryocytic thrombocytopenia (CAMT) is a genetic disorder caused by the loss of function or deletion of myeloproliferative leukemia virus oncogene (*MPL*), the gene encoding the TPO receptor (7, 8). CAMT presents at birth with severe thrombocytopenia and absent bone marrow MKs and develops into bone marrow failure/aplastic anemia during the childhood years or even earlier. The disease is fatal unless successfully treated with HSC transplantation, which indicates that TPO/MPL signaling is indispensable for hematopoietic homeostasis in humans. Notably, most CAMT patients show a reduction in platelet and erythrocyte counts prior to the decrease in leukocyte counts (9), and repetitive transfusion of erythrocyte and/or platelets is usually necessary prior to curative bone marrow transplantation. These clinical features imply that each hematopoietic lineage has its own distinct dependency on MPL signaling. On the other hand, the precise roles of MPL at defined differentiation steps during normal hematopoiesis and the effects of its loss in CAMT patients remain unclear, in large part because of the difficulty of obtaining patients' HSCs for in vitro analysis. Although mouse *Tpo*$^{-/-}$ or *Mpl*$^{-/-}$ models show sustained thrombocytopenia with smaller numbers of MKs and smaller myeloid and erythrocyte progenitor pools in the bone marrow (4, 10), they do not fully recapitulate the phenotype manifested in CAMT patients. For example, *Mpl*$^{-/-}$ mice have normal levels of erythrocytes and leukocytes in their peripheral blood throughout life, and live to an old age without developing bone marrow failure/aplastic anemia.

Disease-specific human induced pluripotent stem cells (iPSCs) are an attractive tool for elucidating the pathogenesis of hematological diseases (11–15), for validating gene therapy models (13, 15–17), and for drug screening. Of importance in the present study is that MKs and erythrocytes generated in vitro from disease-specific iPSCs are an effective tool for studying the mechanism of not only thrombopoiesis (11), but also erythropoiesis (18–20).

Here, we established iPSCs derived from a patient diagnosed with CAMT and treated with curative allogenic stem cell transplantation (referred to herein as CAMT iPSCs) (7, 21). In several established CAMT iPSCs, the *MPL* mutations responsible for the complete loss of MPL expression were carried over. Using CAMT iPSCs and an in vitro disease tracking system we established previously (22–24), we determined the precise link between MPL signaling and development of a common MK/erythrocyte progenitor (MEP) and elucidated the pathogenesis of CAMT by recapitulating the clinical manifestations of the disease.

Results

Disease-specific iPSCs from a CAMT patient failed to generate MKs and platelets. A candidate patient was treated with bone marrow transplantation at 12 years of age (7, 21) after being diagnosed with CAMT. We used skin fibroblasts from the patient to create iPSCs with normal karyotypes using the previously established method with G glycoprotein of the vesicular stomatitis virus (VSV-G) pseudotyped retroviruses (23, 25) harboring 4 (*OCT3/4, SOX2, KLF4,* and *c-MYC*) or 3 (*OCT3/4, SOX2,* and *KLF4*) reprogramming factors (Supplemental Figure 1A; supplemental material available online with this article; doi:10.1172/JCI64721DS1). The resultant CAMT iPSCs exhibited mutations corresponding to the original donor skin, including compound heterozygous point mutations in the *MPL* locus: a C-to-T transition at the cDNA nucleotide position 556 in exon 4, and a single nucleotide deletion of thymine at position 1,499 in exon 10 (Figure 1A and ref. 7). The following parameters were taken as evidence of the pluripotency of CAMT iPSCs: alkaline phosphatase staining; immunostaining for SSEA-4, TRA1-60, and TRA1-81 (Supplemental Figure 1B); gene expression (data not shown); and the capacity for teratoma formation in NOD/SCID mice (Supplemental Figure 1C). We also confirmed that the exogenous reprogramming factors were all silenced in the established iPSCs (data not shown).

To explore the hematopoietic differentiation potential of CAMT iPSCs, we evaluated 3 CAMT iPSC clones and compared them with normal iPSCs (clone TkDA3-4; see Methods) previously established from age-matched dermal fibroblasts using 4 reprogramming factors (23). Using our recently established in vitro differentiation system (22–24), we confirmed that all of the CAMT iPSC clones generated few MKs or platelets, even in the presence of 100 ng/ml TPO, 50 ng/ml stem cell factor (SCF), and 25 U/ml heparin (Figure 1B).

CAMT patients are indeed thrombocytopenic at diagnosis: their platelet counts range 20,000–50,000 platelets/mm^3, equivalent to 5%–10% of that in healthy individuals. Conversely, platelet numbers from CAMT iPSCs in this study were less than 1% of that obtained with normal iPSCs (0.51% ± 0.29%, 0.62% ± 0.42%, and

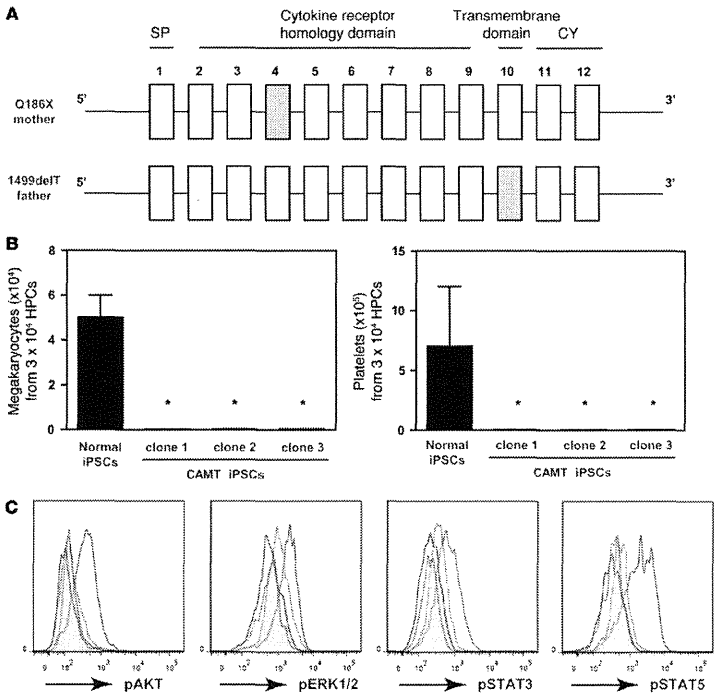

Figure 1

Disease-specific iPSCs recapitulate the disease phenotype manifested in a patient with CAMT. (**A**) Maternal mutation Q186X (C-to-T transition at the cDNA nucleotide position 556) in exon 4 and paternal mutation 1,499delT (single nucleotide deletion of thymine at position 1,499) in exon 10 in a patient with CAMT. SP, signal peptide; CY, cytoplasmic domain. (**B**) Generation of MKs and platelets from normal or CAMT iPSC–derived HPCs cultured on C3H10T1/2 feeder cells for 10 days in the presence of SCF (50 ng/ml), TPO (100 ng/ml), and heparin (25 U/ml). All CAMT iPSC clones generated few MKs or platelets. (**C**) Flow cytometric analysis of MPL-mediated downstream signaling in normal iPSC– (red lines) or CAMT iPSC–derived (blue lines) CD34$^+$ HPCs stimulated with 100 ng/ml TPO for 10 minutes (solid lines) or with vehicle control (dotted lines). No response was observed with CAMT iPSCs. *$P < 0.05$.

Conflict of interest: The authors have declared that no conflict of interest exists.
Citation for this article: *J Clin Invest.* 2013;123(9):3802–3814. doi:10.1172/JCI64721.

Downloaded on January 6, 2014. The Journal of Clinical Investigation. More information at www.jci.org/articles/view/64721

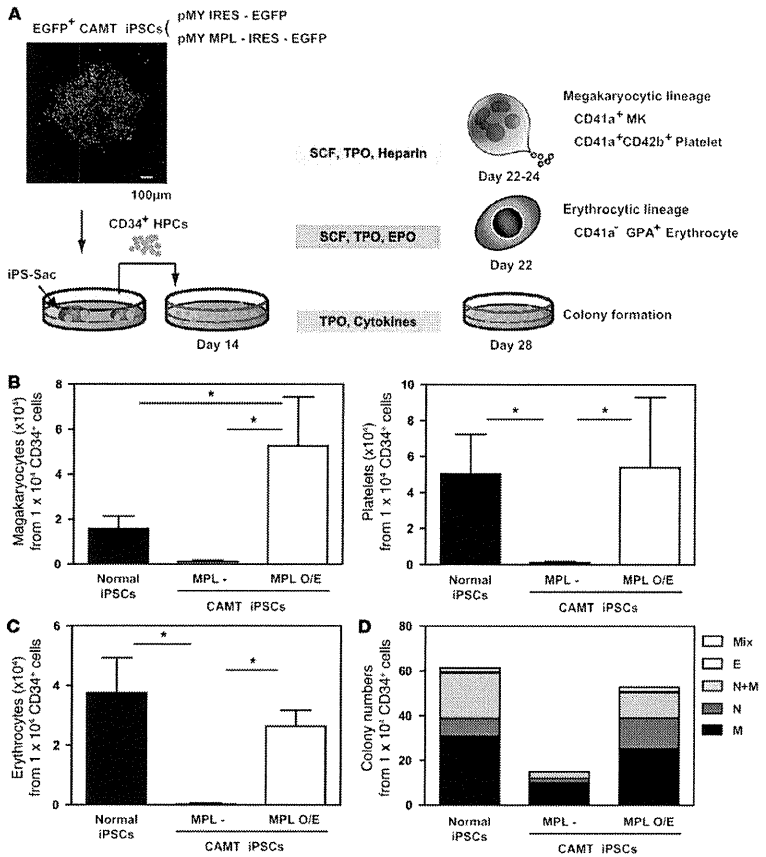

Figure 2

Disease phenotype of CAMT iPSCs was seemingly restored by MPL supplementation. (**A**) In vitro differentiation protocol using CAMT iPSCs transduced with *MPL*. CAMT iPSCs were transduced using a pMY retroviral vector harboring *EGFP* and *MPL* or *EGFP* alone and then selected as EGFP$^+$ populations. Using the indicated cytokines, CAMT iPSC–derived EGFP$^+$ CD34$^+$ HPCs overexpressing MPL were subjected to in vitro differentiation toward the MK or erythrocyte lineage or colony formation. Scale bar: 100 μm. (**B** and **C**) Numbers of MKs and platelets (**B**) cultured in SCF, TPO, and heparin for 8–10 days and erythrocytes (**C**) cultured with SCF (50 ng/ml), TPO (10 ng/ml), and EPO (6 U/ml) for 8 days. Cells were cultured without or with supplemental MPL to induce overexpression (O/E). (**D**) Colony-forming potential for myeloid lineage in normal iPSC–derived CD34$^+$ HPCs, and CAMT iPSC–derived EGFP$^+$ CD34$^+$ HPCs with or without MPL supplementation, in MethoCult H4434 semisolid medium containing TPO (50 ng/ml), SCF, EPO, IL-3, and GM-CSF. M, macrophage; N, neutrophil; N+M, neutrophil and macrophage; E, erythrocyte. *$P < 0.05$.

0.56% ± 0.21%; Figure 1B). However, when we reduced the TPO concentration to a more physiological level (0.1–1 ng/ml) (8), platelet numbers from CAMT iPSCs reached 5%–10% of those obtained with normal iPSCs (Supplemental Figure 2); i.e., they approximated the relative numbers obtained in vivo. This finding suggested that CAMT iPSCs recapitulate the thrombocytopenia seen in the CAMT patient, yielding a useful CAMT disease model.

These results were further confirmed by the failure of TPO stimulation to induce phosphorylation of mediators downstream of MPL (i.e., pAKT, pERK1/2, pSTAT3, and pSTAT5) in CD34$^+$ HPCs (Figure 1C) and by the absence of *MPL* mRNA and MPL protein expression (Supplemental Figure 3, A–C), which corresponded to the complete absence of *MPL* expression in bone marrow cells from the patient (26).

Downloaded on January 6, 2014. The Journal of Clinical Investigation. More information at www.jci.org/articles/view/64721

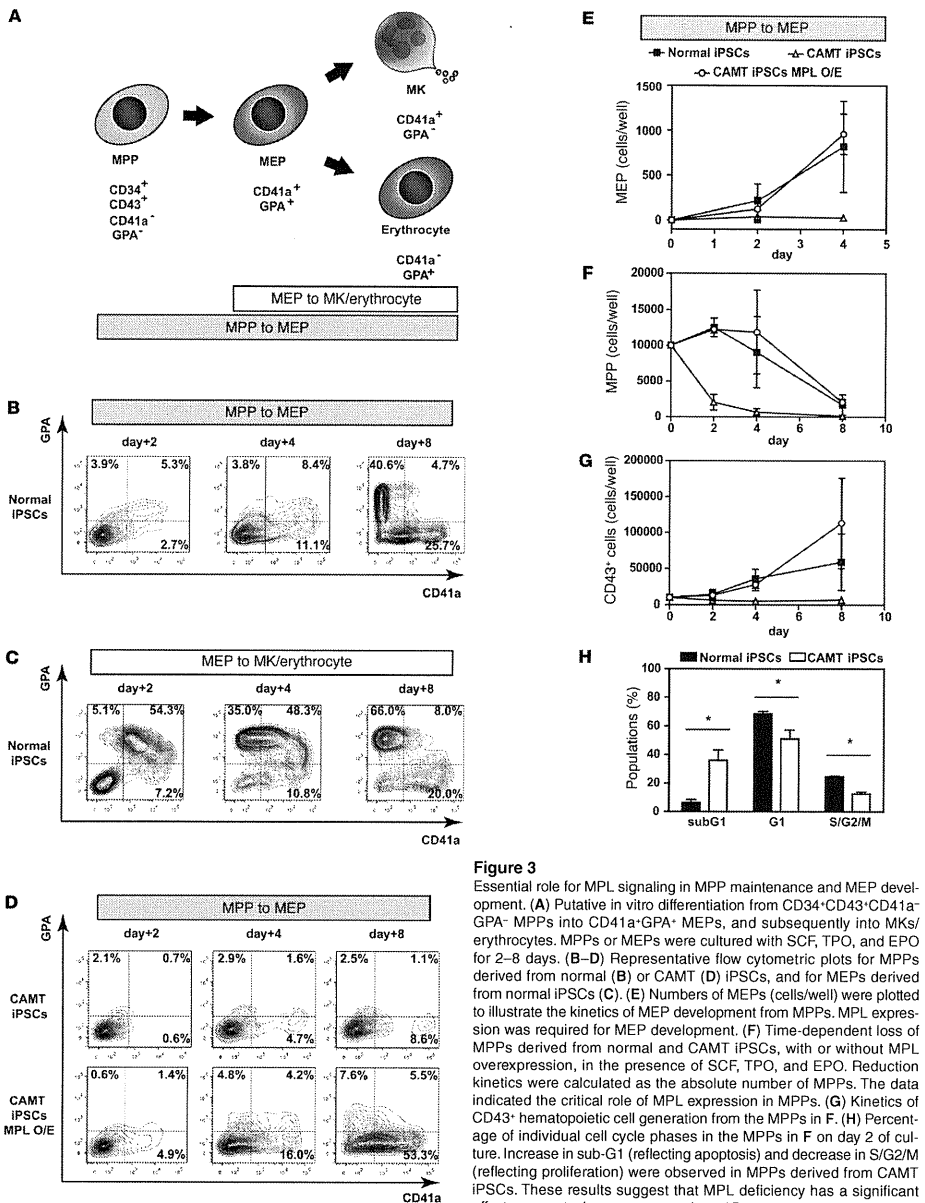

Figure 3

Essential role for MPL signaling in MPP maintenance and MEP development. (**A**) Putative in vitro differentiation from CD34$^+$CD43$^+$CD41a$^-$GPA$^-$ MPPs into CD41a$^+$GPA$^+$ MEPs, and subsequently into MKs/erythrocytes. MPPs or MEPs were cultured with SCF, TPO, and EPO for 2–8 days. (**B**–**D**) Representative flow cytometric plots for MPPs derived from normal (**B**) or CAMT (**D**) iPSCs, and for MEPs derived from normal iPSCs (**C**). (**E**) Numbers of MEPs (cells/well) were plotted to illustrate the kinetics of MEP development from MPPs. MPL expression was required for MEP development. (**F**) Time-dependent loss of MPPs derived from normal and CAMT iPSCs, with or without MPL overexpression, in the presence of SCF, TPO, and EPO. Reduction kinetics were calculated as the absolute number of MPPs. The data indicated the critical role of MPL expression in MPPs. (**G**) Kinetics of CD43$^+$ hematopoietic cell generation from the MPPs in **F**. (**H**) Percentage of individual cell cycle phases in the MPPs in **F** on day 2 of culture. Increase in sub-G1 (reflecting apoptosis) and decrease in S/G2/M (reflecting proliferation) were observed in MPPs derived from CAMT iPSCs. These results suggest that MPL deficiency has a significant effect on apoptosis among progenitors. *$P < 0.05$.

Retroviral complementation of WT MPL restored hematopoiesis to CAMT iPSCs. To determine whether complementation of WT MPL in CAMT iPSCs could restore blood generation, CAMT iPSCs were transduced with *MPL* and *EGFP* using retroviral vectors and then purified by sorting out EGFP$^+$ cells using the differentiation protocol depicted in Figure 2A. CAMT iPSCs transduced with vehicle vector showed severely impaired generation of MKs, platelets, and erythrocytes, although they retained reduced granulocyte and macrophage differentiation potential in colony formation assays (Figure 2, B–D). Conversely, overexpression of MPL in CAMT iPSCs restored the differentiation potential of all myeloid cell lineages to levels comparable to those seen in normal iPSCs (Figure 2, B–D).

TPO/MPL signaling was indispensable for MPP maintenance and transition to common MEPs. Earlier studies demonstrated that, in an in vitro hematopoietic differentiation system using human ES cells (ESCs), CD34$^+$CD43$^+$CD41a$^-$GPA$^-$ and CD41a$^+$GPA$^+$ populations represent multipotent hematopoietic progenitors (MPPs) (27) and MEPs (28), respectively (Figure 3A). We evaluated whether human iPSCs exhibit traceable differentiation steps, similar to those seen with normal dermal fibroblast-derived iPSCs, in our differentiation system. CD34$^+$CD43$^+$CD41a$^-$GPA$^-$ MPPs sorted on day 14 of culture differentiated into CD41a$^+$GPA$^+$ MEPs on additional day +4, which in turn differentiated into CD41a$^+$GPA$^-$ MKs (25.7%) and CD41a$^-$GPA$^+$ erythrocytes (40.6%) on additional day +8 (Figure 3B). We then further confirmed that the selected CD41a$^+$GPA$^+$ MEPs could be differentiated into both CD41a$^+$GPA$^-$ MKs (20.0%) and CD41a$^-$GPA$^+$ erythrocytes (66%) on additional day +8 (Figure 3C). In our culture system, therefore, CD34$^+$CD43$^+$CD41a$^-$GPA$^-$ MPPs derived from normal iPSCs had the potential to generate CD41a$^+$GPA$^+$ MEPs that preferentially differentiated into erythrocytes rather than MKs (Figure 3, B and C).

CAMT iPSCs transduced with vehicle vector showed severely defective transition from MPPs to MEPs, which was corrected in CAMT iPSCs overexpressing MPL (Figure 3, D and E). In addition, whereas the number of MPPs derived from normal or CAMT iPSCs overexpressing MPL was maintained at least until day 4, the number of cells derived from untreated CAMT iPSCs declined rapidly, even in the presence of SCF and TPO (Figure 3F). Consequently, fewer CD43$^+$ pure hematopoietic cells were derived from CAMT iPSCs (Figure 3G). In addition, serial replating assays revealed that CD34$^+$ HPCs derived from CAMT iPSCs generated a few colonies exclusively in the first replating trial, whereas CD34$^+$ HPCs from normal iPSCs or from ESCs produced much larger numbers of colonies, even after the second replating trial (Supplemental Figure 4), which implies impairment of self-replication by CAMT iPSC-derived MPPs. The smaller numbers of MPPs could potentially be accounted for by increased apoptosis (sub-G1 population) and/or decreased proliferation (S/G2/M populations) (Figure 3H and Supplemental Figure 5); however, this was more likely due to the loss of MPL. Collectively, these results indicate that MPL signaling is essential for the transition of MPPs to MEPs as well as for MPP maintenance, and that MPL also contributes to their differentiation into granulocytes and macrophages (Figure 2D). Notably, this differentiation profile was consistent with the clinical course in CAMT patients, who typically show early onset of severe thrombocytopenia and anemia prior to leukopenia (9) and HSC exhaustion.

Supplementing CD34$^+$ HPCs derived from CAMT iPSCs with normal levels of MPL contributed to lineage commitments to both MKs and erythrocytes, but higher levels blocked erythropoiesis. Retroviral gene transduction is a powerful tool for studying the actions by a gene of interest. We noted that the expression levels obtained after MPL supplementation in CAMT iPSCs paralleled fluorescence intensity after retroviral transduction-mediated EGFP expression (Supplemental Figure 6, A and B). On day 22 of culture, normal iPSC-derived MPPs could be differentiated into both erythrocytes and MKs in the presence of SCF, TPO, and EPO. Under these conditions, they preferentially differentiated into erythrocytes (Figure 4, A and B). Flow cytometric analysis showed that a CAMT iPSC clone overexpressing MPL produced 2 distinct populations exhibiting different levels of EGFP expression (EGFP^{hi} and EGFP^{lo}; indicative of high and low MPL expression, respectively) on day 22 of culture (Figure 4A). The EGFP^{lo} population showed nearly equal levels of erythrocyte and MK differentiation, similar to normal iPSCs. However, the EGFP^{hi} population showed apparently increased MK differentiation, with few erythrocytes (Figure 4, A and B). These results were identical to results obtained in a prospective analysis in which CD34$^+$ HPCs derived from 7 individual CAMT iPSC clones were EGFP^{hi} or EGFP^{lo} on day 14, faithfully reflecting MPL expression level – i.e., EGFP^{hi} and EGFP^{lo} populations equalized at 10 times normal and at normal levels of MPL, respectively (Supplemental Figure 6), and were respectively subjected to the differentiation protocol (Figure 4, C and D). We also confirmed that the intensity of EGFP expression, as detected by flow cytometry, corresponded well to the level of pSTAT5 (Figure 4E). Based on our results thus far, we concluded that the intensity of MPL signaling determines the fate of erythrocyte and MK differentiation at the MPP stage and that excessive MPL signaling may block erythrocyte differentiation in CAMT iPSCs.

Excessive TPO/MPL signaling in MKs induces aberrant megakaryopoiesis, leading to generation of CD42b$^-$ platelets. As shown in Figure 2B, 3 times as many MKs were generated from CAMT iPSCs overexpressing MPL than from normal iPSCs. The level of *MPL* expression in CD34$^+$ HPCs derived from CAMT iPSCs overexpressing MPL was greater than in those from normal iPSCs (Supplemental Figure 6B); however, the numbers of CD42b$^+$ platelets were similar, which suggests that excessive MPL signaling might adversely influence megakaryopoiesis. Flow cytometric analysis on day 24 of culture revealed that CAMT iPSCs overexpressing MPL generated greater numbers of MKs showing incomplete maturation – i.e., with both CD41a$^+$CD42b$^-$CD42a$^-$ and CD41a$^+$CD42b$^-$CD42a$^+$ populations and low ploidy (Figure 5, A–C). Moreover, the dysregulated MKs appeared to release CD41a$^+$CD42b$^-$CD42a$^-$ platelets (Figure 5, D and E), even in the presence of GM-6001 (Supplemental Figure 7), which we previously showed to induce retention of CD42b on platelets by preventing its metalloproteinase-catalyzed shedding (29).

We also previously showed that excessive c-MYC activation in MKs blocks MK maturation and leads to CD42b^{lo} platelet generation (23). We therefore tested whether overexpression of MPL in human ESCs also blocks MK maturation in association with reduced platelet generation. Indeed, ESCs overexpressing MPL did generate the immature type of CD41a$^+$CD42b$^-$CD42a$^-$ MKs with lower ploidy and CD41a$^+$CD42b$^-$CD42a$^-$ platelets (Figure 5, F and G). We therefore concluded that excessive MPL signaling induces dysregulation of thrombopoiesis, which might be related to the MPL signaling level, as evidenced by the increased TPO sensitivity of CAMT iPSCs overexpressing MPL (Supplemental Figure 8). In addition, flow cytometric analysis suggested that the augmented MPL signaling might be associated with increased pAKT, pSTAT3, or pSTAT5A, but not pERK1/2, in CD41a$^+$ MKs (Supplemental Figure 9), which indicates that excessive signaling impairs normal MK development.

Downloaded on January 6, 2014. The Journal of Clinical Investigation. More information at www.jci.org/articles/view/64721

research article

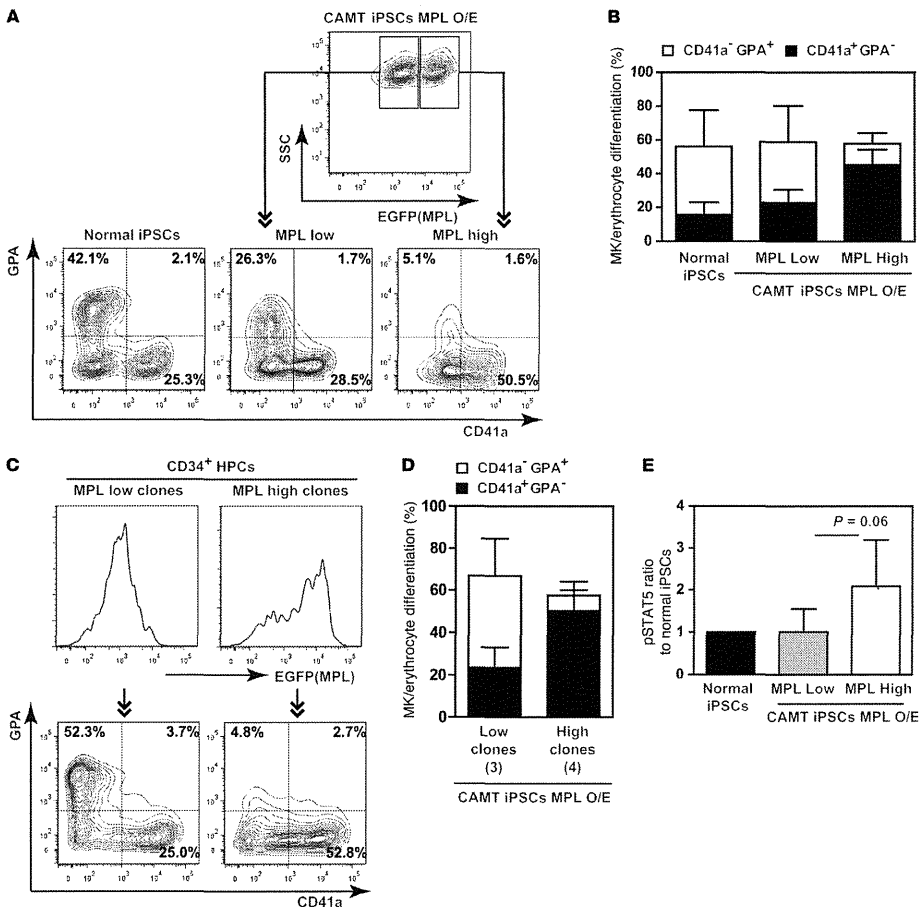

Figure 4

MPL expression levels may determine MK and erythrocyte specification. (**A**) Representative flow cytometric plots of MK and erythrocyte differentiation on day 22. EGFP^{hi} and EGFP^{lo} populations of CAMT iPSCs overexpressing MPL were retrospectively reanalyzed for CD41a and GPA expression. (**B**) Percent CD41a$^-$GPA$^+$ erythrocytes and CD41a$^+$GPA$^-$ MKs from the indicated iPSCs. High MPL expression led to preferential differentiation into MKs. (**C**) EGFP^{hi} and EGFP^{lo} CD34$^+$ HPC populations derived from several CAMT iPSC clones with MPL overexpression are shown as histograms and flow cytometric plots. Representative plots after induction of differentiation are shown for each clone (high and low MPL intensity). (**D**) Percent MK/erythrocyte differentiation. The number of clones is indicated in parentheses. MPL expression level affected the MK/erythrocyte lineage commitment of CD34$^+$ HPCs. (**E**) CD34$^+$ HPCs were incubated with TPO and analyzed for pSTAT5. Results are presented as fluorescence intensity relative to that of normal iPSCs (assigned as 1). The population with higher MPL overexpression showed a markedly augmented TPO-mediated response ($P = 0.06$).

Loss of MPL signaling in CAMT iPSCs impairs FLI1-mediated lineage determination toward erythrocytes or MKs. As shown in Figure 4, the strong intensity of TPO/MPL signaling influenced dominant differentiation into MKs from CAMT iPSC-derived CD34$^+$ HPCs at the expense of erythropoiesis. However, in normal hematopoiesis, some reports showed that TPO/MPL signaling contributed to erythropoiesis (30, 31). To assess the relationship between MPL signal intensity and erythropoiesis in normal cells, the differentiation

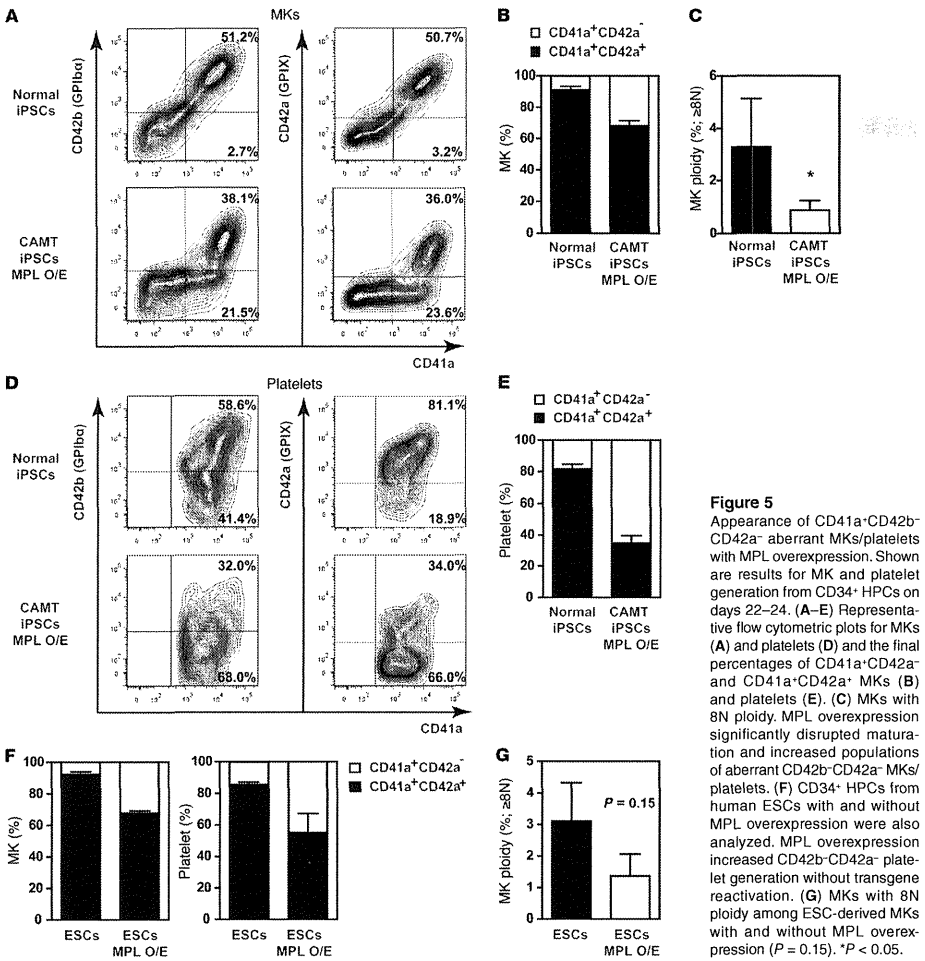

Figure 5
Appearance of CD41a+CD42b-CD42a- aberrant MKs/platelets with MPL overexpression. Shown are results for MK and platelet generation from CD34+ HPCs on days 22–24. (**A**–**E**) Representative flow cytometric plots for MKs (**A**) and platelets (**D**) and the final percentages of CD41a+CD42a- and CD41a+CD42a+ MKs (**B**) and platelets (**E**). (**C**) MKs with 8N ploidy. MPL overexpression significantly disrupted maturation and increased populations of aberrant CD42b-CD42a- MKs/platelets. (**F**) CD34+ HPCs from human ESCs with and without MPL overexpression were also analyzed. MPL overexpression increased CD42b-CD42a- platelet generation without transgene reactivation. (**G**) MKs with 8N ploidy among ESC-derived MKs with and without MPL overexpression ($P = 0.15$). *$P < 0.05$.

of normal iPSC- or ESC-derived CD34+ HPCs into MKs or erythrocytes in the presence of SCF, TPO, and EPO was manipulated by varying the level of exogenous MPL expression. Consistent with an earlier report (31), normal CD34+ HPCs strongly expressing MPL showed erythrocyte-biased differentiation that was dependent on TPO/MPL signaling (Figure 6A and Supplemental Figure 10). This suggests that the program of differentiation toward MKs or erythrocytes distinctly differs between normal and CAMT iPSC-derived CD34+ HPCs and that the lineage balance in normal iPSC-derived HPCs may correspond to normal erythropoiesis in vivo and in vitro (28, 32, 33). Although the underlying mechanism governing lineage commitment toward MKs or erythrocytes is not fully understood, it appears that Friend leukemia virus integration 1 (FLI1) and Kruppel-like factor 1 (KLF1) are potential fate-determining transcriptional factors at the MEP stage and that they mutually antagonize one another (34, 35). Therefore, to evaluate the fate determination program in normal and CAMT iPSC-derived CD34+ HPCs, we used quantitative PCR to assess *FLI1* and *KLF1* expression. Consistent with the mutually antagonistic relationship between FLI1 and KLF1 as well as with earlier reports (28, 34), higher *KLF1* and lower *FLI1* expression were observed with erythrocyte-biased differentiation in normal iPSC-derived cells, whereas lower *KLF1* and higher

Downloaded on January 6, 2014. The Journal of Clinical Investigation. More information at www.jci.org/articles/view/64721

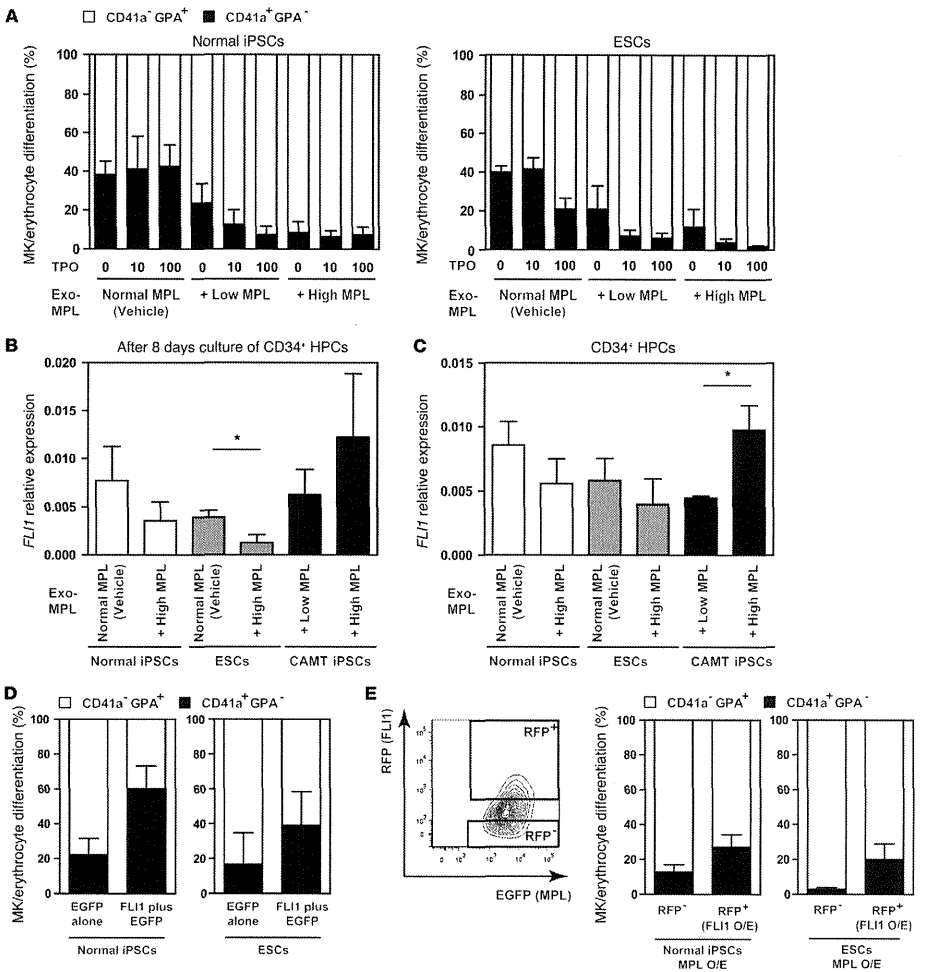

Figure 6
FLI1-mediated MK/erythrocyte differentiation differed between normal and CAMT iPSCs. Shown are results for MK and erythrocyte generation from CD34+ HPCs on day 22. (**A**) Percent CD41+GPA- MKs and CD41-GPA+ erythrocytes derived from normal iPSCs (left) or ESCs (right) and transduced with vehicle or MPL expression vector in the presence of 0, 10, or 100 ng/ml TPO with SCF and EPO. Exogenous MPL expression was assessed based on EGFP fluorescence intensity. Erythropoiesis was enhanced in a TPO/MPL signaling–dependent manner. (**B** and **C**) *FLI1* expression in CD34+ HPCs before (**C**) and after (**B**) differentiation. EGFP^{hi} CD34+ HPCs derived from normal iPSCs and ESCs with or without MPL overexpression, or EGFP^{lo} or EGFP^{hi} CD34+ HPCs from CAMT iPSCs overexpressing MPL, were sorted and cultivated for an additional 8 days. (**D**) Percent MK and erythrocyte differentiation from CD34+ HPCs derived from normal iPSCs or ESCs overexpressing EGFP alone (vehicle) or FLI1 plus EGFP. (**E**) Percent MK and erythrocyte differentiation from normal iPSCs or ESCs overexpressing MPL-EGFP and FLI1-RFP. CD34+ HPCs derived from normal iPSCs or ESCs overexpressing MPL were transfected with FLI1-RFP or vector and cultivated for an additional 8 days. Exogenous FLI1 expression was assessed based on RFP expression (shown in the contour plot). FLI1 overexpression attenuated erythrocyte-biased differentiation in normal iPSCs and ESCs. *$P < 0.05$.

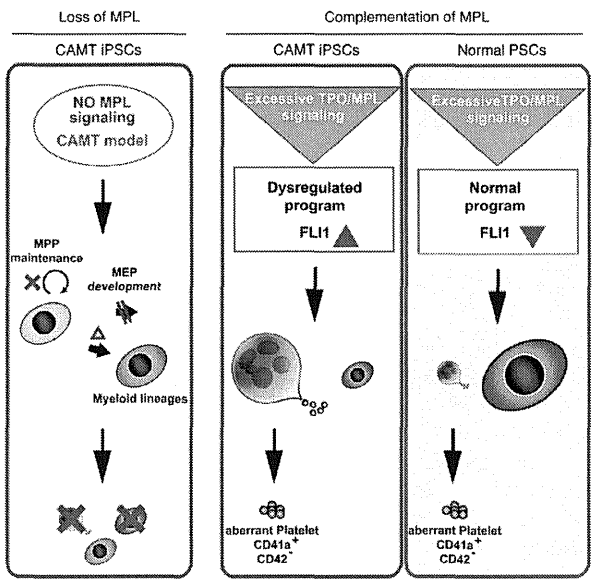

Figure 7
Model of human megakaryopoiesis and erythropoiesis regulated by MPL signaling. CAMT iPSC behavior showed that MPL signaling is essential for MPP maintenance and the development of MEPs from MPPs. In addition, MPL contributes to further myeloid development, but it is not indispensable, which indicates that the dependency on MPL signaling varies among lineages. MPL complementation experiments yielded 2 important findings. First, CAMT iPSCs exhibit dysregulated differentiation toward megakaryopoiesis and erythropoiesis, which is characterized by MK-biased differentiation due to intrinsically dominant FLI1 expression accelerated by ectopic TPO/MPL signaling. Conversely, FLI1 expression is diminished and erythropoiesis is dominant in normal PSCs. Second, excessive TPO/MPL signaling impairs MK maturation and facilitates generation of aberrant CD41a$^+$ and CD42-null platelets from both normal and CAMT iPSCs.

FLI1 were observed with MK-biased differentiation in CAMT iPSC-derived cells (Figure 6B and Supplemental Figure 11A). It is noteworthy that the *FLI1* expression profile at the CD34$^+$ progenitor stage was completely reversed between normal and CAMT iPSC-derived HPCs: *FLI1* expression was reduced in normal iPSC-derived CD34$^+$ HPCs but markedly elevated in CAMT iPSC-derived CD34$^+$ HPCs in a manner dependent on exogenous MPL expression, and this profile persisted even after differentiation (Figure 6, B and C). On the other hand, KLF1 expression behavior appeared to show a similar pattern, but only after differentiation (Supplemental Figure 11, A and B). This suggests that FLI1 is primarily involved in the fate decision from pluripotent stem cell–derived (PSC-derived) MEPs, which may subsequently affect KLF1 function.

We next confirmed that overexpression of FLI1 plus EGFP in CD34$^+$ HPCs derived from normal iPSCs facilitated MK-biased differentiation of the CD34$^+$ HPCs compared with HPCs overexpressing EGFP alone (Figure 6D). More importantly, lentiviral transduction of the same normal CD34$^+$ HPC population overexpressing MPL plus EGFP with red fluorescent protein–labeled (RFP-labeled) FLI1 diminished erythrocyte-biased differentiation (Figure 6E). Furthermore, *FLI1* mRNA levels were well correlated with RFP levels in PSCs, even in the presence of elevated MPL, whereas *KLF1* levels were inversely related to RFP levels (Supplemental Figure 12). These results indicate that CD34$^+$ HPCs derived from CAMT iPSCs exhibit MK-biased differentiation due to impairment of *FLI1* transcriptional factor. Hypermethylation of the *FLI1* promoter reportedly suppresses *FLI1* expression, leading to reduced collagen synthesis in fibroblasts derived from patients with scleroderma fibrosis (36). However, bisulfite sequence analysis of the *FLI1* promoter region revealed that most CpGs in CD34$^+$ HPCs derived from normal and CAMT iPSCs overexpressing MPL were unmethylated and did not significantly differ (data not shown). We therefore conclude that the opposite behavior of CAMT versus normal iPSC-derived HPCs is associated with hyperactivation of *FLI1*, independent of its DNA methylation status.

Discussion

Disease-specific iPSCs are a powerful tool for replacing invaluable cellular resources that are difficult or impossible to obtain from patients, enabling one to recapitulate a disease phenotype in stepwise fashion during differentiation in vitro (11). For example, the difficulty in obtaining living neurons from patients limits understanding of the pathogenesis of degenerative nerve diseases. On the other hand, in vitro neuronal differentiation systems using patient-derived iPSCs have contributed to clarifying the developmental pathogenesis of some diseases (37–40). Similarly, the rarity of MKs within bone marrow is closely associated with our inability to study normal and pathological megakaryopoiesis and platelet generation directly, especially in patients with severely reduced megakaryopoiesis. In that context, we used an in vitro disease tracing system with CAMT iPSCs to make 2 novel findings about the effect of MPL signaling on human hematopoiesis (Figure 7). First, a loss-of-function study revealed the diversity of MPL signaling required among hematopoietic lineages and recapitulated the clinical course of CAMT. Second, a gene complementation (gain-of-function) study using 2 normal iPSC clones and 4 CAMT iPSC clones overexpressing MPL showed that the program governing differentiation toward MKs or erythrocytes was impaired in CAMT iPSC-derived HPCs, and that the effect might be determined primarily by FLI1 expression downstream of TPO/MPL signaling. Another important finding of this study was that complimentary transduction of *MPL* into CAMT iPSCs using a retroviral vector enabled expression of different levels of MPL and revealed the critical importance of MPL-mediated signaling. Taken together, these findings indicate that the strength of the intracellular signaling of TPO/MPL to downstream mediators may determine lineage-dependent fate and behavior and that this important process is completely dysregulated by loss of MPL in CAMT patients.

Patients with CAMT exhibit severe thrombocytopenia, which most often begins at birth and worsens into bone marrow failure with aplastic anemia during childhood or sooner. These features ultimately lead to death unless the patient is successfully treated with HSC transplantation. King et al. provided detailed descriptions of the clinical symptoms and disease courses in 20 CAMT patients (9). These patients were divided into 2 groups, CAMT-I and CAMT-II, based on the time course of the platelet counts during their first year of life, which correlated well with the behavior of the pancytopenia and the type of MPL mutations (9). CAMT-I patients exhibited biallelic nonsense or frameshift mutation of *MPL*, leading to complete null expression of the mRNA. These patients had very low platelet counts and developed severe bone marrow failure during childhood. CAMT-II patients exhibited single amino acid substitutions in the intracellular domain of MPL but retained some residual MPL function, which appeared to delay the onset of bone marrow failure compared with CAMT-I patients. The most critical manifestation in both types of this disease was severe thrombocytopenia and anemia, which were evident prior to leukopenia (9). These findings are consistent with the notion that individual hematopoietic lineages have different dependencies on MPL signaling; however, the precise link between MPL and lineage commitment and blood cell generation at defined differentiation steps remains unclear, in large part because of the difficulty of examining HSCs obtained directly from CAMT patients. The *Mpl*$^{-/-}$ mouse model, which has normal erythrocyte and leukocyte counts throughout life despite sustained thrombocytopenia, does not recapitulate the human phenotype, illustrating the difference in the MPL signal dependencies of mice and humans. We therefore created human iPSCs using fibroblasts from a CAMT-I patient (7, 21) with compound heterozygous nonsense and frameshift mutations in the *MPL* locus (Supplemental Figure 3). These mutations generated a premature stop codon, leading to rapid degradation of the mutated mRNA through nonsense-mediated mRNA decay (41). In fact, none of the truncated MPL proteins or mRNAs could be detected in CD34$^+$ HPCs derived from CAMT iPSCs (Supplemental Figure 3, A–C).

In vitro hematopoietic differentiation assays to track the progression of blood generation revealed that the CD34$^+$CD43$^+$CD41a$^-$GPA$^-$ MPPs derived from iPSC clones created from healthy human specimens can be differentiated into CD41a$^+$GPA$^+$ cells, which contain a MEP population (28) able to give rise to a proper balance of CD41a$^+$GPA$^-$ MKs and CD41a$^-$GPA$^+$ erythrocytes (Figure 3, B and C). In contrast, maintenance of the MPP population and the transition from MPPs to MEPs were completely blocked in CAMT iPSCs (Figure 3, D–F), although some potential for differentiation into the granulocyte and macrophage lineages was retained (Figure 2D). Unlike MPPs derived from normal iPSCs, those from CAMT iPSCs produced smaller numbers of CD43$^+$ cells (Figure 3G), consistent with the pancytopenia seen in CAMT patients and indicative of the role of human MPL signaling in MPPs. Moreover, administration of a nonpeptide TPO mimetic, eltrombopag, improved pancytopenia in patients with aplastic anemia (42), suggesting the potential importance of MPL signaling in human HSCs/HPCs. This is consistent with our novel finding that MPL was indispensable for normal erythropoiesis. Thus, by using CAMT iPSCs, we were able to recapitulate the clinical course of CAMT patients, including the distinct features of their blood cell generation and maintenance during childhood (9).

It was previously suggested that in order to keep platelet production in the normal range, MPL signaling must fall within a window that is presumably regulated by either definite serum level of TPO (43, 44), level of MPL expression and its subsequent dimerization and internalization (45), or various inhibitory signaling molecules (5). One possible mechanism for the dysregulation of CAMT iPSC-derived HPCs is a deficiency in MPL signaling, which could also account for the childhood thrombocytopenia observed in patients. Interestingly, experiments involving MPL expression in normal and CAMT iPSCs strongly implied that regulation of the transcriptional factor FLI1, downstream of TPO/MPL at the CD34$^+$ HPC level, is impaired in CAMT. For dominant production of erythrocytes (rather than MKs) through normal hematopoiesis in humans, erythrocytes must outnumber platelets 20:1. In that case, erythrocytes would likely outnumber MKs 40,000:1, as platelet/MK generation is estimated to be about 2,000:1 (46). Under the conditions of normal human cells, stronger TPO/MPL signaling may suppress FLI1 activity (Figure 6, B and C); determining the mechanism underlying the aberrant FLI1 expression in CAMT iPSCs will require further study.

Our present study also showed that excessive MPL signaling in both normal and CAMT iPSCs led to deleterious megakaryopoiesis and production of CD41a$^+$CD42b$^-$CD42a$^-$ MKs and platelets (Figure 5). Furthermore, we confirmed that these aberrant CD42-null platelets were dysfunctional in terms of inside-out signaling (data not shown). Similar aberrant hematopoiesis with greater numbers of immature MKs and reduced erythropoiesis seen with CAMT iPSCs overexpressing MPL is observed clinically in myeloproliferative disease (MPD) patients carrying the MPLW515 mutant, a constitutively active, TPO-hypersensitive mutant that transduces a signal stronger than that of WT MPL upon TPO stimulation (47–49). The degree of similarity between HPCs from MPLW515 patients and the CAMT iPSCs in this study remains uncertain, but both cells are expected to exhibit hematopoiesis that differs from normal.

MKs derived from both normal and CAMT iPSC-derived HPCs overexpressing MPL in the presence of TPO showed augmented pSTAT3, pSTAT5A, and pAKT, but not pERK. However, we were unable to identify the specific downstream signaling molecule(s) responsible for the aberrant megakaryopoiesis through single transduction of a constitutively active form of AKT (50) or STAT5A (Stat5 1*6) (51, 52) into CD34$^+$ HPCs derived from human ESCs (data not shown). This implies that enhanced signaling by STAT3 and/or other mediators is involved in the blockade of MK maturation and that MPL expression level governs MK maturation.

CAMT patients require HSC transplantation to survive, but a suitable donor is not always available. iPSC technology may enable development of new gene correction/administration therapies using the patients' own HSCs. In that regard, induction of adult-type HSCs from human PSCs was recently accomplished in vivo through teratoma formation in NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ (NSG) and NOD/SCID mice (53, 54).

Downloaded on January 6, 2014. The Journal of Clinical Investigation. More information at www.jci.org/articles/view/64721

research article

In summary, we used CAMT iPSCs to demonstrate that MPL signaling is indispensable for maintenance of MPPs and for transition from MPPs to MEPs during early hematopoiesis. Its absence led to deficiencies in both erythropoiesis and megakaryopoiesis, although some development of myeloid cells was retained (Figure 7). This constellation of effects recapitulates the clinical course seen in CAMT patients. We were also able to show the true pathogenesis of CAMT, whereby lineage commitment by MEPs toward erythropoiesis is the result of a dysregulated program of activated modification of FLI1 gene expression. These findings also provide a rationale for the use of MPL-stimulating drugs (i.e., TPO mimetics) in the treatment of anemia as well as other ailments, such as thrombocytopenia and stem cell suppression.

Methods

Further information can be found in Supplemental Methods.

Cells, reagents, and viral vectors. pMX and pMY retroviral vectors were a gift from T. Kitamura (The University of Tokyo, Tokyo, Japan). pGCDNsam IRES EGFP retroviral vector was from M. Onodera (National Children's Research Center, Tokyo, Japan). The CSII lentiviral vector was from T. Yamaguchi (The University of Tokyo, Tokyo, Japan). pMX mRFP1 retroviral vector was from A. Hotta (CiRA, Kyoto University, Kyoto, Japan). The human ESC line was from N. Nakatsuji and H. Suemori (Institute for Frontier Medical Sciences, Kyoto University, Kyoto, Japan). A normal human iPSC clone, TkDA3-4, was used as a reference (23). The UT-7/TPO cell line was from N. Komatsu (Juntendo University, Tokyo, Japan).

CAMT fibroblasts were cultivated in DMEM supplemented with 10% FBS, 2 mM L-glutamine, 100 U/ml penicillin, and 0.1 mg/ml streptomycin (all from Invitrogen). Disease-specific iPSCs derived from dermal fibroblasts from a CAMT patient (CAMT iPSCs) were established using retroviral vectors harboring 4 reprogramming factors (*OCT3/4*, *SOX2*, *KLF4*, and *c-MYC*; clones 1 and 2) or 3 reprogramming factors (*OCT3/4*, *SOX2*, and *KLF4*; clone 3), as previously described (23). All PSCs were maintained as described previously (23). The mouse C3H10T1/2 cell line was purchased from the RIKEN Bio-Resource Center and was cultured as described previously (23). Retroviral supernatants for establishing iPSCs were obtained from a 293 GPG system (provided by R.C. Mulligan, Children's Hospital Boston, Boston, Massachusetts, USA) (25). Human VEGF, TPO, and SCF were from R&D Systems. Human EPO was from Kyowa Hakko Kirin Co. Heparin was from Ajinomoto Pharmaceuticals Co. The metalloproteinase inhibitor GM-6001 was from Cosmobio Co. The following antibodies were used (from BD, unless indicated otherwise): allophycocyanin- (APC-) or PE/Cy7-conjugated anti-CD34, APC-conjugated anti-CD41a (integrin αIIb; HIP8 clone), eFluor 450–conjugated anti-CD42a (GPIX) (eBioscience), PE-conjugated anti-CD42b (GPIbα), PE-conjugated anti-CD43 (eBioscience), Alexa Fluor 405–conjugated anti–glycophorin A (GPA) (Biolegend), PE-conjugated anti-pAKT, anti-pERK1/2, anti-pSTAT3, and anti-pSTAT5.

Hematopoietic differentiation of human ESCs and iPSCs. Differentiation of human ESCs and iPSCs into hematopoietic cells was performed as described previously (22–24). In brief, small clumps of human PSCs (<100 cells treated with PBS containing 0.25% trypsin (Invitrogen), 1 mM $CaCl_2$ (Sigma-Aldrich), and 20% KSR (Invitrogen) were transferred onto mitomycin-treated or irradiated C3H10T1/2 cells and cocultured in hematopoietic cell differentiation medium, which was replaced every 3 days. On days 14–15 of culture, the HPCs within ESC- and iPSC-Sacs were collected and then sorted into $CD34^+$ HPCs, $CD34^+CD43^+CD41a^-GPA^-$ MPPs, and $CD41a^+GPA^+$ MEPs using an Aria flow cytometer (BD). They were then transferred onto freshly mitomycin-treated or irradiated feeder cells and further cultivated in differentiation medium supplemented with TPO and combinations of other cytokines/mediators/inhibitors (50 ng/ml SCF, 25 U/ml heparin sodium, 6 U/ml EPO, and 50 μM GM-6001), as described previously (22–24, 29). The medium was refreshed every 3 days, and nonadherent cells were collected and analyzed daily from day 16 to day 24.

Retroviral and lentiviral vectors and infection. Full-length *MPL* was subcloned into the retroviral vector pMY IRES EGFP (55). Full-length *FLI1* was subcloned into the retroviral vector pGCDNsam IRES EGFP or lentiviral vector pCSII harboring 2A mRFP1. Viral supernatant was generated as previously described (56). CAMT iPSCs, normal iPSCs, or ESCs were transduced with either vehicle vector or pMY-MPL-EGFP, after which $EGFP^+$ iPSCs were sorted using a flow cytometer. $CD34^+$ HPCs derived from normal iPSCs, with or without MPL overexpression, were also transduced with vehicle, pGCDNsam FLI1-EGFP, or pCS2 FLI1-RFP. This was followed by 8 days of culture in the presence of SCF, TPO, and EPO.

Hematopoietic colony-forming cell assay. Hematopoietic colony-forming cell assays were performed as described previously (23). In brief, $CD34^+$ HPCs from within iPSC-Sacs were cultivated for 14 days in MethoCult H4434 semisolid medium (Stem Cell Technologies Inc.) supplemented with 50 ng/ml human TPO. The colonies were then collected, stained with Hemacolor (Merck), and observed under a microscope.

Flow cytometric analysis of hematopoietic cells. Nonadherent cells on days 16–24 of culture were prepared in PBS containing 3% FBS (staining medium) and stained with combinations of antibodies for 30 minutes on ice. Samples were then washed with staining medium and analyzed by flow cytometry. Dead cells were identified using PI and excluded. Precise numbers of the cells were estimated using True Count Beads (BD Biosciences).

Flow cytometric analysis of polyploidy in MKs. Nonadherent cells on days 22–24 were prepared in staining medium and stained with 25 μg/ml Hoechst 33342 (Sigma-Aldrich) for 45 minutes at 37°C, followed by anti-human CD41a-APC and CD42b-PE for 30 minutes. Samples were then washed with cold staining medium and analyzed by flow cytometry in the presence of PI.

Flow cytometric analysis of platelets. Washed platelets were prepared as described previously (22). The resultant platelet pellets were resuspended with staining medium and stained with anti-human CD41a-APC, CD42a-Pacific blue, and CD42b-PE for 30 minutes at room temperature. The platelets were then diluted in 200 μl staining medium and analyzed by flow cytometry. Platelet numbers were estimated using True Count Beads.

Flow cytometric analysis of the cell cycle. Nonadherent cells from among MPPs on days 14 + 2 of culture were prepared in staining medium and stained with Vybrant DyeCycle Violet Stain (Invitrogen) for 30 minutes at 37°C. Samples were then stained with PI and analyzed by flow cytometry.

Phosphorylation of downstream molecules of MPL. $CD34^+$ HPCs derived from human iPSCs were starved overnight in differentiation medium containing 0.5% FBS, stained with CD34-APC for 30 minutes at 37°C, and then stimulated for 10 minutes with 100 ng/ml TPO. Immediately thereafter, they were fixed with Lyse/Fix buffer (BD) for 10 minutes at 37°C, permeabilized with Perm buffer III (BD) for 30 minutes on ice, and immunostained for pAKT, pERK1/2, pSTAT3, and pSTAT5 using PE-conjugated antibodies at room temperature for 1 hour. Samples were then washed with staining medium and analyzed by flow cytometry. Similarly, MKs in differentiation medium supplemented with SCF, TPO, and heparin on days 22–24 were stained with anti-CD41a–APC for 30 minutes at 37°C. They were then fixed, permeabilized, stained for downstream mediators, and analyzed as the HPCs were.

Quantitative RT-PCR. After sorting $CD34^+$ HPCs using flow cytometry, RNA was extracted and cDNA was synthesized as described previously (22). Real-time PCR was performed using an EagleTaq Master Mix with ROX and Universal Probe Library kit (Roche Applied Science) according to the manufacturer's instructions. Signals were detected using an ABI7900HT Real-Time PCR System (Applied Biosystems). Primer sets for *GAPDH*, *MPL*, and *KLF1* were determined using the Universal Probe Library Set for humans (http://www.roche-applied-science.com/sis/rtpcr/upl/index.jsp?id=UP030000). Primer sets for *FLI1* (Hs00956711_m1) were purchased from Applied Biosystems. The following primer sequences were used: *GAPDH* forward, 5′-AACAGCCTCAAGATCATCAGC-3′; *GAPDH* reverse, 5′-TTGGCAGGTTTTTCTAGACGG-3′; *MPL* primer 1 forward, 5′-CAGCGAGTCCTCTTTGTGG-3′; *MPL* primer 1 reverse, 5′-CCCAGCTGATCTGAAGTTCC-3′; *MPL* primer 2 forward, 5′-AGGCCATCAGGACTGGAA-3′; *MPL* primer 2 reverse, 5′-CAGCTGTAAACGGTAGCGAGA-3′; *MPL* primer 3 forward, 5′-GGTGACCGCTCTGCATCTA-3′; *MPL* primer 3 reverse, 5′-CAGGGCATGCCTCAGTCT-3′; *KLF1* forward, 5′-ACACCAAGAGCTCCCACCT-3′; *KLF1* reverse, 5′-GTAGTGGCGGGTCAGCTC-3′.

Statistics. All data are presented as mean ± SD. The statistical significance of the observed differences was determined using 1-way ANOVA for multiple comparisons and 2-tailed Student's *t* tests for pairwise comparisons. A *P* value less than 0.05 was considered significant.

Study approval. The human ESC clone Kyoto hESC-3 (KhES-3) was obtained from the Institute for Frontier Medical Sciences of Kyoto University after approval for human ESC use was granted by the Minister of Education, Culture, Sports, Science, and Technology of Japan (MEXT). Dermal fibroblasts derived from a CAMT patient were obtained at University of Saga School of Medicine (Saga, Japan) under informed consent. The Review Boards for ethics at the Institute of Medical Science, The University of Tokyo, and University of Saga School of Medicine approved this research protocol, including the informed consent. The entire study was conducted in accordance with the Declaration of Helsinki.

Acknowledgments

The authors thank N. Nakatsuji, H. Suemori, T. Kitamura, A. Hotta, M. Onodera, T. Yamaguchi, and N. Komatsu for providing reagents and cells. The authors also thank A. Watanabe for performing bisulfite sequence analysis and providing useful suggestions. This work was supported by Project of realization of regenerative medicine (phase II) from MEXT (to K. Eto and H. Nakauchi) and a Grant-in-aid (Kaken) for Young Scientist from MEXT (to N. Takayama). This research was also supported in part by the Japan Society for the Promotion of Science (JSPS) through its "Funding Program for World-Leading Innovative R&D on Science and Technology" (FIRST Program) to N. Takayama, S. Nakamura, and T. Dohda.

Received for publication May 9, 2012, and accepted in revised form May 30, 2013.

Address correspondence to: Koji Eto or Naoya Takayama, Clinical Application Department, Center for iPS Cell Research and Application, Kyoto University, 53 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan. Phone: 81.75.366.7075; Fax: 81.75.366.7095; E-mail: kojieto@cira.kyoto-u.ac.jp (K. Eto), naoya.takayama@cira.kyoto-u.ac.jp (N. Takayama).

1. Yoshihara H, et al. Thrombopoietin/MPL signaling regulates hematopoietic stem cell quiescence and interaction with the osteoblastic niche. *Cell Stem Cell.* 2007;1(6):685–697.
2. Qian H, et al. Critical role of thrombopoietin in maintaining adult quiescent hematopoietic stem cells. *Cell Stem Cell.* 2007;1(6):671–684.
3. de Sauvage FJ, et al. Physiological regulation of early and late stages of megakaryocytopoiesis by thrombopoietin. *J Exp Med.* 1996;183(2):651–656.
4. Alexander WS, Roberts AW, Nicola NA, Li R, Metcalf D. Deficiencies in progenitor cells of multiple hematopoietic lineages and defective megakaryocytopoiesis in mice lacking the thrombopoietic receptor c-Mpl. *Blood.* 1996;87(6):2162–2170.
5. Kaushansky K. The molecular mechanisms that control thrombopoiesis. *J Clin Invest.* 2005;115(12):3339–3347.
6. Drachman JG, Griffin JD, Kaushansky K. The c-Mpl ligand (thrombopoietin) stimulates tyrosine phosphorylation of Jak2, Shc, and c-Mpl. *J Biol Chem.* 1995;270(10):4979–4982.
7. Ihara K, et al. Identification of mutations in the c-mpl gene in congenital amegakaryocytic thrombocytopenia. *Proc Natl Acad Sci U S A.* 1999;96(6):3132–3136.
8. Ballmaier M, et al. c-mpl mutations are the cause of congenital amegakaryocytic thrombocytopenia. *Blood.* 2001;97(1):139–146.
9. King S, Germeshausen M, Strauss G, Welte K, Ballmaier M. Congenital amegakaryocytic thrombocytopenia: a retrospective clinical analysis of 20 patients. *Br J Haematol.* 2005;131(5):636–644.
10. Carver-Moore K, et al. Low levels of erythroid and myeloid progenitors in thrombopoietin-and c-mpl-deficient mice. *Blood.* 1996;88(3):803–808.
11. Takayama N, Eto K. Pluripotent stem cells reveal the developmental biology of human megakaryocytes provide a source of platelets for clinical application. *Cell Mol Life Sci.* 2012;69(20):3419–3428.
12. Park IH, et al. Disease-specific induced pluripotent stem cells. *Cell.* 2008;134(5):877–886.
13. Raya A, et al. Disease-corrected haematopoietic progenitors from Fanconi anaemia induced pluripotent stem cells. *Nature.* 2009;460(7251):53–59.
14. Ye Z, et al. Human-induced pluripotent stem cells from blood cells of healthy donors and patients with acquired blood disorders. *Blood.* 2009;114(27):5473–5480.
15. Zou J, Mali P, Huang X, Dowey SN, Cheng L. Site-specific gene correction of a point mutation in human iPS cells derived from an adult patient with sickle cell disease. *Blood.* 2011;118(17):4599–4608.
16. Zou J, et al. Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. *Cell Stem Cell.* 2009;5(1):97–110.
17. Sun N, et al. Patient-specific induced pluripotent stem cells as a model for familial dilated cardiomyopathy. *Sci Transl Med.* 2012;4(130):130ra47.
18. Lu SJ, et al. Biologic properties and enucleation of red blood cells from human embryonic stem cells. *Blood.* 2008;112(12):4475–4484.
19. Kobari L, et al. Human induced pluripotent stem cells can reach complete terminal maturation: in vivo and in vitro evidence in the erythropoietic differentiation model. *Haematologica.* 2012;97(12):1795–1803.
20. Chang CJ, et al. Production of embryonic and fetal-like red blood cells from human induced pluripotent stem cells. *PloS One.* 2011;6(10):e25761.
21. Muraoka K, et al. Successful bone marrow transplantation in a patient with c-mpl-mutated congenital amegakaryocytic thrombocytopenia from a carrier donor. *Pediatr Transplant.* 2005;9(1):101–103.
22. Takayama N, et al. Generation of functional platelets from human embryonic stem cells in vitro via ES-sacs, VEGF-promoted structures that concentrate hematopoietic progenitors. *Blood.* 2008;111(11):5298–5306.
23. Takayama N, et al. Transient activation of c-MYC expression is critical for efficient platelet generation from human induced pluripotent stem cells. *J Exp Med.* 2010;207(13):2817–2830.
24. Takayama N, Eto K. In vitro generation of megakaryocytes and platelets from human embryonic stem cells and induced pluripotent stem cells. *Methods Mol Biol.* 2012;788:205–217.
25. Ory DS, Neugeboren BA, Mulligan RC. A stable human-derived packaging cell line for production of high titer retrovirus/vesicular stomatitis virus G pseudotypes. *Proc Natl Acad Sci U S A.* 1996;93(21):11400–11406.
26. Muraoka K, et al. Defective response to thrombopoietin and impaired expression of c-mpl mRNA of bone marrow cells in congenital amegakaryocytic thrombocytopenia. *Br J Haematol.* 1997;96(2):287–292.
27. Vodyanik MA, Thomson JA, Slukvin II. Leukosialin (CD43) defines hematopoietic progenitors in human embryonic stem cell differentiation cultures. *Blood.* 2006;108(6):2095–2105.
28. Klimchenko O, et al. A common bipotent progenitor generates the erythroid and megakaryocyte lineages in embryonic stem cell-derived primitive hematopoiesis. *Blood.* 2009;114(8):1506–1517.
29. Nishikii H, et al. Metalloproteinase regulation improves in vitro generation of efficacious platelets from mouse embryonic stem cells. *J Exp Med.* 2008;205(8):1917–1927.
30. Kobayashi M, Laver JH, Kato T, Miyazaki H, Ogawa M. Recombinant human thrombopoietin (Mpl ligand) enhances proliferation of erythroid progenitors. *Blood.* 1995;86(7):2494–2499.
31. Parekh C, et al. Novel pathways to erythropoiesis induced by dimerization of intracellular C-Mpl in human hematopoietic progenitors. *Stem Cells.* 2012;30(4):697–708.
32. Lu J, et al. MicroRNA-mediated control of cell fate in megakaryocyte-erythrocyte progenitors. *Dev Cell.* 2008;14(6):843–853.
33. Bouilloux F, et al. EKLF restricts megakaryocytic differentiation at the benefit of erythrocytic differentiation. *Blood.* 2008;112(3):576–584.
34. Dore LC, Crispino JD. Transcription factor networks in erythroid cell and megakaryocyte development. *Blood.* 2011;118(2):231–239.
35. Starck J, et al. Functional cross-antagonism between transcription factors FLI-1 and EKLF. *Mol Cell Biol.* 2003;23(4):1390–1402.
36. Wang Y, Fan PS, Kahaleh B. Association between enhanced type I collagen expression and epigenetic repression of the FLI1 gene in scleroderma fibroblasts. *Arthritis Rheum.* 2006;54(7):2271–2279.
37. Israel MA, et al. Probing sporadic and familial Alzheimer's disease using induced pluripotent stem cells. *Nature.* 2012;482(7384):216–220.
38. Dimos JT, et al. Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. *Science.* 2008;321(5893):1218–1221.
39. Bilican B, et al. Mutant induced pluripotent stem cell lines recapitulate aspects of TDP-43 proteinopathies and reveal cell-specific vulnerability. *Proc*

Downloaded on January 6, 2014. The Journal of Clinical Investigation. More information at www.jci.org/articles/view/64721

research article

Natl Acad Sci U S A. 2012;109(15):5803–5808.
40. Yagi T, et al. Modeling familial Alzheimer's disease with induced pluripotent stem cells. *Hum Mol Genet.* 2011;20(23):4530–4539.
41. Isken O, Maquat LE. The multiple lives of NMD factors: balancing roles in gene and genome regulation. *Nat Rev Genet.* 2008;9(9):699–712.
42. Olnes MJ, et al. Eltrombopag and improved hematopoiesis in refractory aplastic anemia. *N Engl J Med.* 2012;367(1):11–19.
43. Fielder PJ, et al. Regulation of thrombopoietin levels by c-mpl-mediated binding to platelets. *Blood.* 1996;87(6):2154–2161.
44. Stoffel R, Wiestner A, Skoda RC. Thrombopoietin in thrombocytopenic mice: evidence against regulation at the mRNA level and for a direct regulatory role of platelets. *Blood.* 1996;87(2):567–573.
45. Hitchcock IS, Chen MM, King JR, Kaushansky K. YRRL motifs in the cytoplasmic domain of the thrombopoietin receptor regulate receptor internalization and degradation. *Blood.* 2008;112(6):2222–2231.
46. Patel SR, Hartwig JH, Italiano JE Jr. The biogenesis of platelets from megakaryocyte proplatelets. *J Clin Invest.* 2005;115(12):3348–3354.
47. Pikman Y, et al. MPLW515L is a novel somatic activating mutation in myelofibrosis with myeloid metaplasia. *PLoS Med.* 2006;3(7):e270.
48. Guglielmelli P, et al. Anaemia characterises patients with myelofibrosis harbouring Mpl mutation. *Br J Haematol.* 2007;137(3):244–247.
49. Tefferi A. JAK and MPL mutations in myeloid malignancies. *Leuk Lymphoma.* 2008;49(3):388–397.
50. Kohn AD, Summers SA, Birnbaum MJ, Roth RA. Expression of a constitutively active Akt Ser/Thr kinase in 3T3-L1 adipocytes stimulates glucose uptake and glucose transporter 4 translocation. *J Biol Chem.* 1996;271(49):31372–31378.
51. Kato Y, et al. Selective activation of STAT5 unveils its role in stem cell self-renewal in normal and leukemic hematopoiesis. *J Exp Med.* 2005;202(1):169–179.
52. Olthof SG, Fatrai S, Drayer AL, Tyl MR, Vellenga E, Schuringa JJ. Downregulation of signal transducer and activator of transcription 5 (STAT5) in CD34+ cells promotes megakaryocytic development, whereas activation of STAT5 drives erythropoiesis. *Stem Cells.* 2008;26(7):1732–1742.
53. Amabile G, et al. In vivo generation of transplantable human hematopoietic cells from induced pluripotent stem cells. *Blood.* 2013;121(8):1255–1264.
54. Suzuki N, et al. Generation of engraftable hematopoietic stem cells from induced pluripotent stem cells by way of teratoma formation molecular therapy: the Journal Of The American Society Of Gene Therapy [published online ahead of print May 14, 2013]. *Mol Ther.* doi:10.1038/mt.2013.71.
55. Kitamura T, et al. Retrovirus-mediated gene transfer and expression cloning: powerful tools in functional genomics. *Exp Hematol.* 2003;31(11):1007–1014.
56. Eto K, et al. The WAVE2/Abi1 complex differentially regulates megakaryocyte development and spreading: implications for platelet biogenesis and spreading machinery. *Blood.* 2007;110(10):3637–3647.

To the editor:

ANKRD26-related thrombocytopenia and myeloid malignancies

Since the discovery that mutations in the 5′ untranslated region (UTR) of *ANKRD26* are responsible for an autosomal-dominant form of thrombocytopenia (*ANKRD26*-RT),[1] 21 affected families were reported.[2] A study analyzing this series of patients suggested that *ANKRD26*-RT is characterized by normal platelet size, moderate thrombocytopenia, and absent or mild bleeding tendency.[2] The study also found that the number of hematologic malignancies in affected families was higher than expected, but the relatively small cohort of patients precluded firm conclusions.

To gain further information on this matter, we screened for mutations in the 5′ UTR of *ANKRD26* 215 subjects with an inherited thrombocytopenia of unknown origin and found 11 mutations (3 not previously described) in 23 cases (Table 1[3]). Analysis of family members identified another 52 affected subjects. Moreover, we found 43 additional subjects, first- or second-degree relatives of *ANKRD26*-RT patients, who were known to be thrombocytopenic from an early age but were not available for genetic investigation (they were dead or not willing to perform the analysis). Analysis of the new series of 75 patients with *ANKRD26* mutations confirmed that *ANKRD26*-RT is characterized by moderate thrombocytopenia with normal platelet size and a mild bleeding phenotype.

In the extended series of 118 subjects certainly or very likely affected, we identified 10 patients who had developed myeloid malignancies: 4 acute myeloid leukemias (age at onset, 40-60 years), 4 myelodysplastic syndromes (MDS) (age at onset, 35-70 years), and 2 chronic myeloid leukemias (CML) (age at onset, 30-65 years). The total observation time was 4741 years, and the incidence of acute myeloid leukemia, MDS, and CML was 84 (confidence interval [CI], 23-216), 84 (CI, 23-216), and 42 (CI, 5-152) per 100 000, respectively. Putting together the 118 cases examined in this study and the 104 cases already reported, 4.9% of patients had acute leukemias, 2.2% MDS, and 1.3% CML. The total observation time was 8915 years, with an incidence of acute leukemias, MDS, and CML of 123 (CI, 62-221), 56 (CI, 18-131), and 34 (CI, 7-98) per 100 000, respectively, thus higher than expected in the general population (5.2, 4.5, and 1.6 per 100 000, respectively, according to the National Cancer Institute[4]). Unlike myeloid malignancies, the incidence of lymphoproliferative disorders and nonhematologic cancers in *ANKRD26*-RT pedigrees was not higher than expected (data not reported). Available data are therefore compatible with the hypothesis that *ANKRD26*-RT, as the inherited thrombocytopenia deriving from *RUNX1* mutations,[5] predisposes to myeloid malignancies, in particular acute leukemias. The observation that this unfavorable outcome occurred in a limited proportion of subjects and has been observed in only 12 of 44 families indicates that the penetrance for malignancies was incomplete and other genetic and/or environmental factors contributed to the development of these disorders. Of note, the the case series described in this paper confirmed the already reported finding of unexplained high leukocyte counts in a large proportion of subjects. Further observation is required to ascertain whether the higher leukocyte counts are related to eventual development of leukemia.

In conclusion, *ANKRD26*-RT is an insidious form of inherited thrombocytopenias that exposes patients to a low risk of bleeding but predisposes them to hematologic myeloid malignancies. Recognizing this disorder and its attendant risks is important for proper management of affected subjects.[6]

Table 1. Main clinical and laboratory features of patients with *ANKRD26*-RT grouped by family

Family/No. of patients/Country	*ANKRD26* 5′ UTR mutation	Mean age, y (range)	WHO bleeding score[3] (no. of patients)	Mean platelet count, $\times 10^9$/L (range)	Mean MPV, fL (range)	Mean hemoglobin, g/dL (range)	Mean WBC, $\times 10^9$/L (range)
1/3/Canada	c.-116C>G*	18 (6-37)	1 (2), 2 (1)	70.6 (45-107)	9.55 (9.4-9.7)†	12.85 (12.1-13.6)†	11.7 (10.6-12.8)†
2/2/France	c.-118C>A	34 (17-51)	2 (2)	63.5 (45-82)	10.2 (9.9-10.5)	15.5 (14.8-16.3)	7.25 (6.8-7.7)
3/1/US	c.-118C>T	57	0 (1)	38	7.7	14.3	10.5
4/3/France	c.-119C>A	32.6 (2-65)	2 (2), 3	55.6 (36-81)	10.1 (9-11.3)	13.7 (12.8-14.2)	7.5 (4.7-9.6)
5/1/Italy	c.-126T>G	28	1 (1)	14	8.1	16.6	7.3
6/9/Argentina	c.-127A>G	37.7 (6-74)	0 (2), 1 (4), 2 (3)	96.2 (68-147)	8.6 (7.7-9.5)‡	15.3 (12.9-17.7)	8.47 (6.1-11.1)
7/2/Italy	c.-127A>G	32 (17-47)	0 (1), 2 (1)	79.5 (68-91)	8.8 (8.6-9)	14.15 (12.4-15.9)	8.59 (7.92-9.26)
8/3/France	c.-127A>T	19 (6-38)	0 (1), 1 (2)	65.6 (47-85)	7.5 (7.1-7.8)	14.8 (12.8-17.2)	7 (5.3-8.7)
9/8/France	c.-127A>T	39.25 (1-68)	0 (5), 1 (3)	39.37 (24-84)	10.8 (9.3-11.7)	14.15 (11-16.8)	9.55 (6.7-12.1)
10/6/France	c.-127delAT*	34.6 (1-66)	0 (6)	54.6 (26-96)	10.2 (8.7-11)	14.2 (10.4-16.6)	7.86 (5.8-13.4)
11/3/France	c.-128G>A	63 (21-97)	2 (2)	19 (12-30)	8.1 (7.7-8.5)†	13.6 (13.6-13.6)†	7.75 (7.54-7.96)†
12/2/France	c.-128G>A	36.5 (26-47)	1, 2	71.5 (68-75)	8.4§	14.8§	10§
13/3/Italy	c.-128G>A	20 (12-26)	0 (3)	34 (14-70)	7.6 (6.3-8.5)	15.3 (14-16.5)	14.6 (9.21-21)
14/1/Japan	c.-128G>A	25	0 (1)	54	8.9	15.8	12.5
15/2/Italy	c.-128G>A	52 (50-54)	1 (2)	20 (19-21)	8.4§	14.2§	5.1§
16/10/US	c.-128G>A	23.2 (1-62)	0 (4), 1 (6)	35.5 (19-65)	8.7 (7.5-10.2)	14.43 (12.7-16.1)	9.93 (7-12.4)
17/8/Italy	c.-128G>A	40.2 (6-93)	0 (1), 1 (1), 2 (3), 3 (19), 4 (1)	18.7 (5-34)	8.5 (7-10.3)‖	13.47 (11.8-16.3)	7.64 (5.73-11.31)
18/1/Italy	c.-128G>A	38	2 (1)	48	9.04	15.9	8.51
19/2/France	c.-128G>C*	20 (1-39)	0 (2)	52.5 (24-81)	12 (10-14)	13.15 (12.3-4)	9.35 (6.6-12.1)
20/1/Japan	c.-134G>A	13	0 (1)	81	9.2	14.2	7.64
21/2/US	c.-134G>A	58.5 (50-67)	2 (2)	10 (7-13)	8.2 (5.7-10.7)	11.7 (10.7-12.7)	7.85 (5-10-75)
22/1/The Netherlands	c.-134G>A	34	2 (1)	24	11.1	15.5	12.5
23/1/US	c.-134G>A	64	1 (1)	20	8.8	13.4	8.2

MPV, mean platelet volume; US, United States; WBC, white blood cell; WHO, World Health Organization.
*New mutations.
†Data from 2 of 3 patients.
‡Data from 7 of 9 patients.
§Data from 1 of 2 patients.
‖Data from 7 of 8 patients.

Patrizia Noris
Department of Internal Medicine, University of Pavia–Istituto Di Ricovero e Cura a Carattere Scientifico Policlinico San Matteo Foundation, Pavia, Italy

Remi Favier
French Reference Center for Inherited Platelet Disorders, Armand Trousseau Children's Hospital, Assistance Publique-Hôpitaux de Paris, Paris, France
Institut National de la Santé et de la Recherche Médicale, Unités Mixtes de Recherche 1009, Villejuif, France

Marie-Christine Alessi
Aix-Marseille University, Faculty of Medicine, Institut National de la Santé et de la Recherche Médicale, Unités Mixtes de Recherche 1062, Marseille, France

Amy E. Geddis
Division of Pediatric Hematology Oncology, University of California at San Diego, Rady Children's Hospital San Diego, San Diego, CA

Shinji Kunishima
Department of Advanced Diagnosis, Clinical Research Center, National Hospital Organization Nagoya Medical Center, Nagoya, Japan

Paula G. Heller
Instituto de Investigaciones Médicas Alfredo Lanari, Consejo Nacional de Investigaciones Científicas y Técnicas, University of Buenos Aires, Buenos Aires, Argentina

Paola Giordano
Dipartimento di Scienze Biomediche e Oncologia Umana, Sezione di Pediatria, Università di Bari, Bari, Italy

Karen Y. Niederhoffer
Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada

James B. Bussel
Division of Pediatric Hematology-Oncology, Weill Medical College of Cornell University, New York, NY

Gian Marco Podda
UO Medicina III, Ospedale San Paolo, Milan, Italy

Nicola Vianelli
Department of Haematology and Clinical Oncology "L. and A. Seràgnoli," S. Orsola-Malpighi Hospital, University of Bologna, Bologna, Italy

Rogier Kersseboom
Department of Clinical Genetics, Erasmus Medical Centre, Rotterdam, The Netherlands

Alessandro Pecci
Department of Internal Medicine, University of Pavia–Istituto Di Ricovero e Cura a Carattere Scientifico Policlinico San Matteo Foundation, Pavia, Italy

Chiara Gnan
Institute for Maternal and Child Health, Istituto Di Ricovero e Cura a Carattere Scientifico "Burlo Garofolo," Trieste, Italy

Caterina Marconi
Medical Genetics Unit, Policlinico S. Orsola-Malpighi, University of Bologna, Bologna, Italy

Anne Auvrignon
French Reference Center for Inherited Platelet Disorders, Armand Trousseau Children's Hospital, Assistance Publique-Hôpitaux de Paris, Paris, France

William Cohen
Aix-Marseille University, Faculty of Medicine, Institut National de la Santé et de la Recherche Médicale, Unités Mixtes de Recherche 1062, Marseille, France

Jennifer C. Yu
Division of Pediatric Hematology Oncology, University of California at San Diego, Rady Children's Hospital San Diego, San Diego, CA

Akihiro Iguchi
Department of Pediatrics, Hokkaido University Graduate School of Medicine, Sapporo, Japan

Allison Miller Imahiyerobo
Division of Pediatric Hematology-Oncology, Weill Medical College of Cornell University, New York, NY

Francoise Boehlen
Division of Angiology and Hemostasis, University Hospital of Geneva and Faculty of Medicine, Geneva, Switzerland

Dorsaf Ghaloussi
Aix-Marseille University, Faculty of Medicine, Institut National de la Santé et de la Recherche Médicale, Unités Mixtes de Recherche 1062, Marseille, France

Daniela De Rocco
Institute for Maternal and Child Health, Istituto Di Ricovero e Cura a Carattere Scientifico "Burlo Garofolo," Trieste, Italy

Pamela Magini
Medical Genetics Unit, Policlinico S. Orsola-Malpighi, University of Bologna, Bologna, Italy

Elisa Civaschi
Department of Internal Medicine, University of Pavia–Istituto Di Ricovero e Cura a Carattere Scientifico Policlinico San Matteo Foundation, Pavia, Italy

Ginevra Biino
Institute of Molecular Genetics, and Pavia and Institution of Population Genetics, Consiglio Nazionale delle Ricerche, Sassari, Italy

Marco Seri
Medical Genetics Unit, Policlinico S. Orsola-Malpighi, University of Bologna, Bologna, Italy

Anna Savoia
Institute for Maternal and Child Health, Istituto Di Ricovero e Cura a Carattere Scientifico "Burlo Garofolo" and Department of Medical Sciences, University of Trieste, Trieste, Italy

Carlo L. Balduini
Department of Internal Medicine, University of Pavia–Istituto Di Ricovero e Cura a Carattere Scientifico Policlinico San Matteo Foundation, Pavia, Italy

P.N. and R.F. contributed equally to this manuscript.