Familial platelet disorder (FPD)/acute myelogenous leukaemia (AML) (MIM601399) is an autosomal dominant disorder with inherited thrombocytopenia, abnormal platelet function and a lifelong risk of the development of a variety of haematological malignancies[1], such as AML, myelodysplastic syndromes (MDS) and myeloproliferative neoplasms. Although inherited *RUNX1* mutations are the cause of the congenital thrombocytopenia, it remains unclear whether a mutation in *RUNX1*, which is generally known to have a dominant-negative effect[2–4], is sufficient to induce the development of haematological malignancies in individuals with FPD/AML. It is also not known whether additional gene mutations are required for the transformation, and, if so, which genes are involved. Given that only 40% of FPD/AML patients develop these neoplasms[5] and that a relatively long period is required for subsequent *RUNX1* mutation-mediated development of neoplasms in FPD/AML, the secondary genetic events may function as a driver to promote malignant transformation. We reasoned that identifying gene mutations responsible for the malignant transformation of FPD/AML would provide indispensable information for addressing these questions. However, only about 30 pedigrees with FPD/AML have been reported so far, and the rarity of this disorder has impeded the establishment of clinical diagnostic criteria and the clinical improvement to refine cancer therapy and to identify biomarkers that would allow detection of patients at risk for the onset of malignancies in FPD/AML.

We collected DNA samples and clinical information of 73 individuals, belonging to 57 pedigrees, who have a history of familial thrombocytopenia and/or haematological malignancies, with the aim of identifying pedigrees with FPD/AML and uncovering recurrent mutations that drive the malignant transformation. Next-generation sequencing and single-cell sequencing strategy suggest that somatic mutation in *CDC25C* may be one of the early genetic events for leukaemic initiation in FPD/AML, and further stepwise acquisition of mutations such as *GATA2* leads to FPD/AML-associated leukaemic progression. These observations shed light on a part of leukemogenesis in FPD/AML.

Results

A novel gene target in haematological disorders. Thirteen patients in 7 pedigrees were diagnosed as having FPD/AML after screening for germline *RUNX1* mutations in 73 index patients; 7 of the 13 patients had developed haematological malignancies, while the other 6 only showed thrombocytopenia (Table 1). Most of the detected *RUNX1* mutations were point mutation in Runt homology domain or frame-shift mutation that lost transactivation domain, consistent with the previous reports[2,4]. As haploinsufficiency of *RUNX1* might cause familial thrombocytopenia with propensity to develop AML[1], we also examined whether the pedigrees have *RUNX1* loss of heterozygosity (LOH) or not. A synchronized quantitative-PCR method[6] and single-nucleotide polymorphism (SNP) sequencing detected no case with LOH in *RUNX1* in our cohort (Supplementary Fig. 1 and detailed in Methods). To systematically identify additional genetic alterations, we utilized whole-exome sequencing for two individuals from the same FPD/AML pedigree who shared a common *RUNX1*_p.Phe303fs mutation and who had developed MDS (subject 20) or myelofibrosis (subject 21) at the age of 37 and 17 years, respectively. In both these patients, the disease had progressed to AML[7]. Validation by Sanger sequencing and/or targeted deep sequencing of candidate mutations in paired tumour/normal DNA samples confirmed 10 (subject 20) and 8 (subject 21) somatically acquired nonsynonymous mutations (Table 2; Supplementary Figs 2–4; Supplementary Methods). Surprisingly, both patients carried the identical somatic *CDC25C* mutation (p.Asp234Gly), which had not been reported previously in human cancers (Fig. 1a,b). Prompted by this finding, we investigated *CDC25C* mutations in other FPD/AML cases by deep sequencing. In total, four of seven affected patients with haematological malignancies had *CDC25C* mutations, of which three carried the same p.Asp234Gly mutation. Moreover, *CDC25C* mutations were detected in an additional three FPD/AML patients who had not yet developed haematological malignancies, although the variant allele fractions (VAFs) were much lower in this group of patients than in those who had already developed haematological malignancies (Fig. 1c; Table 1). Thus, 7 of the 13 FPD/AML patients (53%) harboured a *CDC25C* mutation. *CDC25C* was also screened for mutations in 90 sporadic MDS and 53 AML patients, including 13 MDS and 3 AML cases who carried *RUNX1* mutations. No *CDC25C* mutations were identified in the 90 sporadic cases, except for the p.Ala344Val in an MDS patient bearing a *RUNX1* mutation, indicating that *CDC25C* mutations were significantly associated with germline, but not with somatic *RUNX1* mutations ($P = 0.004$; Supplementary Fig. 5; Supplementary Table 1).

Clonal evolution of FPD/AML. Deep sequencing of individual mutations that had been detected by whole-exome sequencing

Table 1 | Mutational status of *CDC25C* in FPD/AML patients.

Pedigree number	Subject number	RUNX1 mutation	Disease status	Age, years*	CDC25C mutation	VAF (%)
18	20	p.Phe303fs	MDS/AML	37/38	p.Asp234Gly	31.7/45.8
	21		MF/AML	17/18	p.Asp234Gly	31.1/39.0
19	22	p.Arg174*	AML	41	p.His437Asn	39.7
54	65	p.Ser140Asn	MDS	25	—	—
	66		AML	56	p.Asp234Gly	24.2
32	38	p.Leu445Pro	HCL	72	—	—
16	18	p.Thr233fs	Thrombocytopenia	—	p.Asp234Gly	5.9
53	62	p.Gly262fs	MDS	12	—	—
	63		Thrombocytopenia	—	—	—
	67		Thrombocytopenia	—	—	—
57	71	p.Gly172Glu	Pancytopenia†	—	p.Asp234Gly	8.3
	72		Thrombocytopenia	—	—	—
	73		Thrombocytopenia	—	p.Lys233Glu	1.8

AML, acute myeloid leukemia; FPD, familial platelet disorder; HCL, hairy cell leukemia; MDS, myelodysplastic syndrome; MF, myelofibrosis; VAF, variant allele fraction.
*Age at the time of diagnosis of each haematological malignancy is shown.
†*Thrombocytopenia, leukopenia and iron-deficiency anemia were diagnosed.*

Table 2 | Validated somatic mutations.

Gene symbol	Ref seq_no.	Amino-acid change	Position (hg19)	Base change	Mutation type	SIFT prediction	VAF at MDS/MF (%)	VAF at AML (%)
Subject 20								
AGAP4	NM_133446	p.Arg484Cys	g.chr10:46321905	C->T	Missense	Damaging	13.2	11.5
CDC25C	NM_001790	p.Asp234Gly	g.chr5:137627720	A->G	Missense	Damaging	31.7	45.8
CHEK2	NM_007194	p.Arg406His	g.chr22:29091740	G->A	Missense	Tolerated	14.6	11.1
COL9A1	NM_001851	p.Gly878Val	g.chr6:70926733	G->T	Missense	Damaging	9.6	26.4
DTX2	NM_001102594	p.Pro74Arg	g.chr7:76110047	C->G	Missense	Damaging	18.3	11.2
FAM22G	NM_001170741	p.Ser508Thr	g.chr9:99700727	T->A	Missense	Tolerated	10.2	27.6
GATA2	NM_001145661	p.Leu321His	g.chr3:128202758	T->A	Missense	Damaging	0.0	28.1
LPP	NM_001167671	p.Val538Met	g.chr3:188590453	G->A	Missense	Damaging	9.7	28.8
RP1L1	NM_178857	p.Ser215fs	g.chr8:10480295	insC	Frameshift	Damaging	14.2	12.7
SIGLEC9	NM_014441	p.Ser437Gly	g.chr19:51633253	A->G	Missense	Tolerated	27.4	42.5
Subject 21								
ANXA8L1	NM_001098845	p.Val281Ala	g.chr10:48268018	T->C	Missense	Damaging	30.8	36.8
CDC25C	NM_001790	p.Asp234Gly	g.chr5:137627720	A->G	Missense	Damaging	31.1	39.1
DENND5A	NM_001243254	p.Arg320Ser	g.chr11:9215218	A->C	Missense	Damaging	29.5	37.3
FER	NM_005246	p.Tyr634Cys	g.chr5:108382876	A->G	Missense	Damaging	1.4	30.4
FNDC1	NM_032532	p.Arg189Cys	g.chr6:159636081	C->T	Missense	Damaging	29.3	35.9
OR8U1	NM_001005204	p.Asn175Ile	g.chr11:56143623	A->T	Missense	Damaging	30.0	34.1
PIDD	NM_145886	p.Arg342Cys	g.chr11:802347	C->T	Missense	Damaging	3.3	28.3
ZNF614	NM_025040	p.Glu202Gly	g.chr19:52520246	A->G	Missense	Damaging	28.7	33.7

AML, acute myeloid leukemia; MDS, myelodysplastic syndrome; MF, myelofibrosis; SIFT, sorting intolerant from tolerant; VAF, variant allele fraction.

allowed accurate determination of their VAFs; on this basis, we could establish an inferred model of clonal evolution in terms of individual mutations in subjects 20 and 21 (Fig. 2a,b; Supplementary Fig. 6a,b). Intratumoral heterogeneity was evident at both MDS and AML phases in subject 20. According to the predicted model, a founding clone with a *CDC25C* mutation acquired additional mutations in *COL9A1*, *FAM22G* and *LPP* (group A), followed by the emergence of a *GATA2* mutation (group B), which was associated with leukaemic transformation, whereas the size of another subclone, defined by mutations in *CHEK2* and three other genes (group C), was unchanged. To validate this hierarchical model, single-cell genomic sequencing was performed using genomic DNA of 63 bone marrow cells from subject 20 when the patient was in the AML phase. Assuming that all cells harbour the *RUNX1* mutation, the false-negative rate of the procedure reached 35%, possibly due to biased allele amplification (Online Methods). However, this technique successfully demonstrated that the group A/B and group C mutations were mutually exclusive (Fig. 2c; Supplementary Table 2). To statistically evaluate this possibility, we assumed two hypotheses (H_0: the mutational status of genes in group A/B and group C is independent; H_1: mutations in group A/B and group C are mutually exclusive) and calculated each probability distribution (P_i: probability that the current results as shown in Fig. 2c were obtained under the hypothesis H_i). Our mutational profile data were achieved with a much higher likelihood under H_1 than H_0 (Supplementary Fig. 7 and detailed in Supplementary Methods). Similarly, the clonal architecture for subject 21 was portrayed in Fig. 2b and Supplementary Fig. 6b. In both scenarios, *CDC25C* mutations seemed to represent a founding mutation with the highest VAF, suggesting that the *CDC25C* mutation contributed to the establishment of a founding tumour population as an early genetic event, whereas progression to AML seemed to be accompanied by the appearance of additional mutations, indicating a multistep process in leukemogenesis.

Along with the somatic mutations found in subjects 20 and 21, a *GATA2* mutation was also identified in subject 22 (Fig. 3a). This patient developed AML with multilineage dysplasia, which led to the diagnosis of AML – MRC (myelodysplasia-related changes). Remission-induction therapies were only partially effective and the blast cell count was reduced from 54 to 5.6%, while dysplastic features persisted (Fig. 3b; Supplementary Fig. 8). Allogeneic stem cell transplantation was successfully performed from a human leukocyte antigen-matched unrelated donor and durable complete remission, with 100% donor chimerism, was achieved. During treatment, the VAF of the *GATA2* mutation decreased virtually in parallel with the blast cell percentage, while the VAF of the *CDC25C* mutation hovered at a high level before transplantation. Thus, we hypothesized that the *GATA2* mutation induced leukaemia progression in this patient, whereas the *CDC25C* mutation was associated with the pre-leukaemic status. Another *GATA2* mutation (p.Leu359Val) was found in subject 18, with a VAF (0.94%), who showed only thrombocytopenia without any signs of leukaemia progression and who had a small subclone with a concurrent *CDC25C* mutation (Fig. 3c). Although *GATA2* mutations are detected in a small number of patients with FPD/AML, the findings described above suggest that mutation of *GATA2* is a key factor promoting disease progression in FPD/AML (Fig. 3d).

Biological consequences of *CDC25C* mutations. We next investigated the possible impact of *CDC25C* mutation on clonal selection and evolution. CDC25C is a phosphatase that prevents premature mitosis in response to DNA damage at the G2/M checkpoint, while it is constitutively phosphorylated at Ser216 throughout interphase by c-TAK1 (refs 8–10). When phosphorylated at Ser216, CDC25C binds to 14-3-3 protein[11], leading to sequestration of CDC25C to the cytoplasm and its inactivation. Ba/F3 cells were transduced with retroviruses encoding the wild-type or mutant CDC25C containing each of the individual mutations (p.Asp234Gly, p.Ala344Val, p.His437Asn and p.Ser216Ala), and assayed for the phosphorylation status, 14-3-3 protein-binding capacity and intracellular localization of each of these proteins. The Ser216Ala mutant form

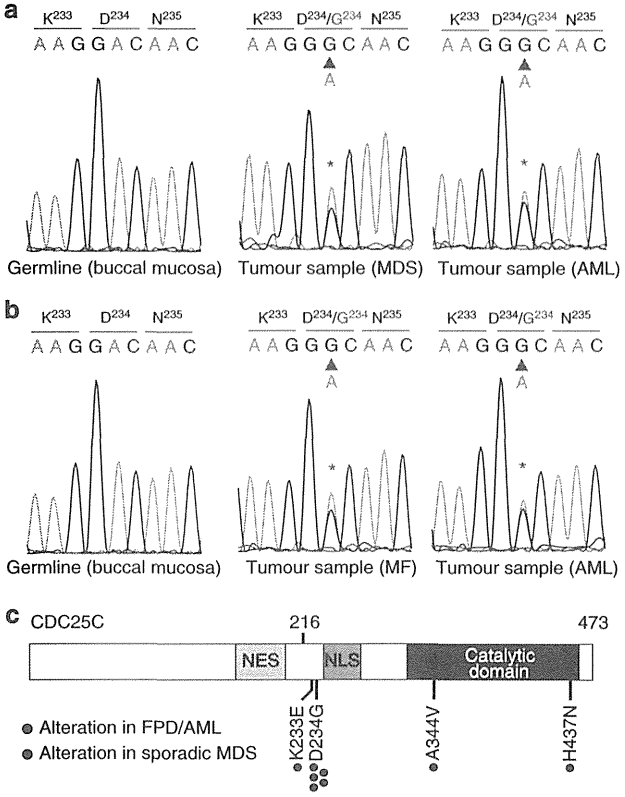

Figure 1 | Mutation in *CDC25C* recurs in cases of FPD/AML. (**a**,**b**) Sanger sequencing of *CDC25C* mutations found in whole-exome sequencing is shown. Both forward and reverse traces were available for each mutation, but only one trace is shown above. The results of buccal mucosa, pre-leukaemic phase and leukaemic phase is demonstrated for subject 20 (**a**) and subject 21 (**b**), respectively. (**c**) The distribution of alterations is shown for the CDC25C protein. NES, a putative nuclear export signal domain between amino acids 177–200; NLS, a putative nuclear localization sequence domain consisting of amino acids 240–244.

of CDC25C, which lacks the phosphorylation site, was used as a negative control. In all of the mutated forms of CDC25C, the capacity for binding to c-TAK1 was reduced (Fig. 4a,b; Supplementary Fig. 9a,b), resulting in decreased phosphorylation of CDC25C at Ser216 (Fig. 4c). Consequently, the mutant proteins failed to bind 14-3-3 protein efficiently (Fig. 4d,e; Supplementary Fig. 8c,d) and remained in the nucleus even during interphase (Fig. 4f; Supplementary Figs 10 and 11). In accordance with these observations, CDC25C mutants enhanced mitotic entry, which was exaggerated by low-dose radiation-induced DNA damage (Fig. 4g,h; Supplementary Fig. 12; Supplementary Methods). These results suggest that mutation of *CDC25C* results in disruption of the DNA checkpoint machinery. Next, we investigated why mutation of *CDC25C* is a frequent genetic event in FPD/AML. It is known that *RUNX1* mutations suppress DNA damage repair and subsequent cell cycle arrest in hematopoietic cells by means of transcriptional suppression of several genes that are involved in DNA repair[12,13]. We confirmed that FPD/AML-associated *RUNX1* mutations have similar effects, as we observed activation of the G2/M checkpoint mechanism in the presence of *RUNX1* mutations (Fig. 4i; Supplementary Fig. 13a,b). We found, however, that introduction of mutations in *CDC25C* resulted in enhanced mitosis entry, despite co-existence of *RUNX1* mutations (Fig. 4i). Therefore, we speculated that compromised DNA damage checkpoint mechanisms caused by mutations in *CDC25C* may contribute to malignant transformation, in concert with increased genomic instability due to *RUNX1* mutations.

Discussion

Whole-exome sequencing, followed by targeted deep sequencing, identified novel aspects of the pathogenesis of malignant transformation in FPD/AML. First, the high frequency of *CDC25C* mutations in FPD/AML underscores their major role in the development of haematological malignancies in FPD/AML patients. To our knowledge, *CDC25C* mutations have not been reported previously and represent a new recurrent mutational target in haematological malignancies, although *CDC25C* mutations have been reported in some solid carcinomas with unknown significance[14,15]. Furthermore, our functional assays support their biological significance, which is characterized by cell cycle progression and premature mitotic entry. Although the 5q31 minimally deleted region, in which CDC25C is located, is frequently detected in MDS, it seems to be associated with other oncogenic mechanisms since our functional assays suggested that *CDC25C* mutations in FPD/AML were gain-of-function type mutations that facilitate the mitotic entry by aberrant accumulation in the nucleus. Impaired DNA repair function mediated by germline *RUNX1* mutation may play a role in the generation of *CDC25C* mutations.

Evaluation of the allelic burden of mutated genes demonstrated that *CDC25* mutations are found with high VAFs in FPD/AML-derived leukaemia and with low VAFs in cases of thrombocytopenia. Our hierarchical model and clonal selection highlighted that mutation of *CDC25C* defines an initial event during malignant transformation and predates subclonal mutations in *GATA2* and other genes. On the basis of the observation that four of the seven FPD/AML patients with *CDC25C* mutations have developed leukaemia and that *CDC25C* mutations were actually detected in the leukaemic subclones, we speculated that a FPD/AML patient with a *CDC25C* mutation, but without clinically evident leukaemia, is at high risk for the onset of leukaemic progression. Examination of the allelic burden of *CDC25C* mutation may thus serve to evaluate the risk of leukaemic progression in patients with FPD/AML.

Among the mutations found in FPD/AML, mutations in *GATA2* were identified in 3 of 13 individuals (subjects 18, 20 and 22). *GATA2* mutations were frequently identified in FPD/AML-derived leukaemia (2/7) and in a patient with thrombocytopenia who had a small subclone bearing a *CDC25C* mutation (1/6). Although reports on the clinical relevance of *GATA2* mutations in myeloid malignancy are limited, several lines of evidence in this respect have recently been reported. *GATA2* mutations are frequently found in a subgroup of patients with cytogenetically normal AML with biallelic *CEBPA* gene mutations[16], which account for ~4% of AML. Germline *GATA2* mutations are also observed in disorders linked to an increased propensity for the development of MDS and AML, including Emberger syndrome, MonoMAC syndrome and dendritic cells, monocytes, B and natural killer cells deficiency[17–20]. The alterations in *GATA2* (leading to p.Leu321His and p.Leu359Val), which were found in FPD/AML patients in this study, are located in the part of the gene encoding the N-terminal and C-terminal zinc-finger domains, respectively (Fig. 3d). Mutations affecting the identical amino acids have been reported in AML patients bearing *CEBPA* mutations and chronic myeloid leukaemia patients in blast crisis[16,21]. Thus, *GATA2* mutation may contribute to AML progression in collaboration with *RUNX1* and/or *CDC25C* mutations. Furthermore, although

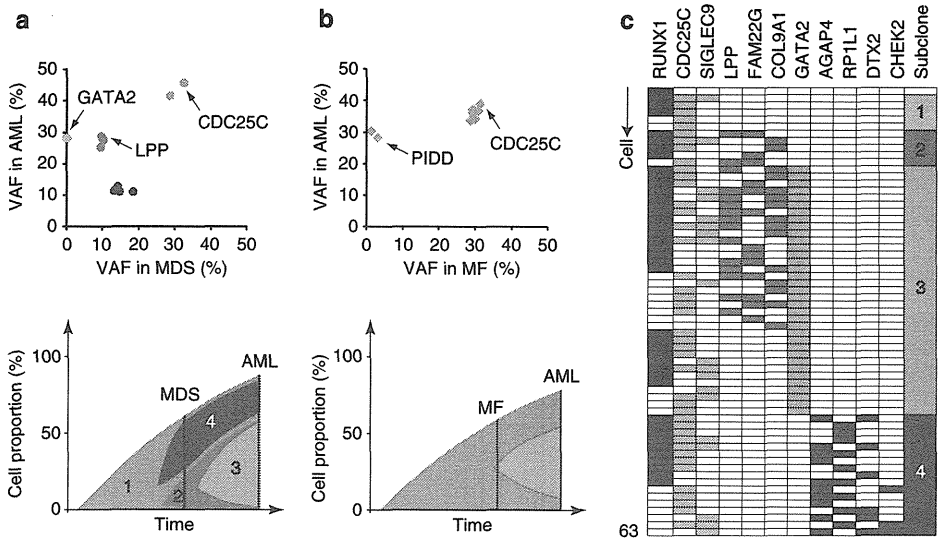

Figure 2 | Clonal evolution of FPD/AML-related myeloid disorders. (**a**,**b**) Observed variant allele fraction (VAF) of validated mutations are listed in Table 2, in both pre-leukaemic and leukaemic phases, are shown in diagonal plots (top) for subject 20 (**a**) and subject 21 (**b**). Predicted chronological behaviours in different leukemia subclones are depicted below each diagonal plot. Distinct mutation clusters are displayed by colour. The vertical axis represents cell proportion of each clone calculated by VAF × 2 (%) (because all the mutations were heterozygous), regarding the whole bone marrow as 100%. (**c**) Mutation status of each bone marrow cell from subject 20 during the acute myeloid leukemia (AML) phase. The vertical axis represents each cell ($n = 63$) and the horizontal axis displays each gene mutation. Coloured columns show that the corresponding cell harbours gene mutation(s) as defined in Online Methods. Subclone numbers shown in the right row correspond to the numbers in the lower figure of **a**.

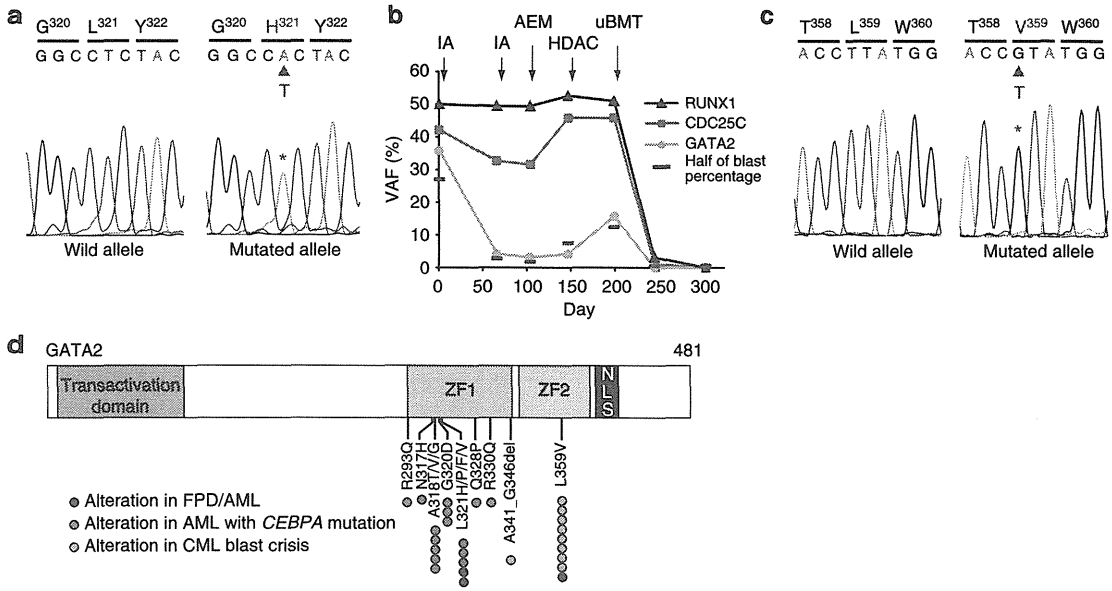

Figure 3 | *GATA2* mutations in FPD/AML. The result of Sanger sequencing for *GATA2* p.Leu321His mutation in subject 22 (**a**) and Leu359Val mutation in subject 18 (**c**) validated with subcloning strategy by methods shown in Supplementary Methods. (**b**) Variant allele fractions (VAFs) of *RUNX1*, *CDC25C* and *GATA2* mutation in subject 22 are demonstrated with the time course of treatment. Half the value of the blast cell percentage, which corresponds to the allele frequency of a heterozygous mutation, is also shown by a red bar. IA, idarubicine + Ara-C; AEM, Ara-C + etoposide + mitoxantrone; HDAC, high-dose Ara-C; uBMT, unrelated bone marrow transplantation. (**d**) Schematic representation of *GATA2* mutations. *GATA2* mutations that were identified in FPD/AML are displayed together with mutations found in AML with *CEBPA* mutation[16] as well as in CML patients in blast crisis[21]. ZF, zinc-finger domain; NLS, a putative nuclear localization sequence domain.

another report identified somatic *CBL* mutation with acquired 11q uniparental disomy as a second hit as being responsible for leukaemic transformation in FPD/AML[22], *CBL* mutations were not detected in our series of FPD/AML samples.

Although the precise pathogenetic roles of *CDC25C* mutations remain unclear, we presume that mutant CDC25C alleles confer a proliferative advantage under certain circumstances in which DNA repair machinery is compromised, such as that mediated by

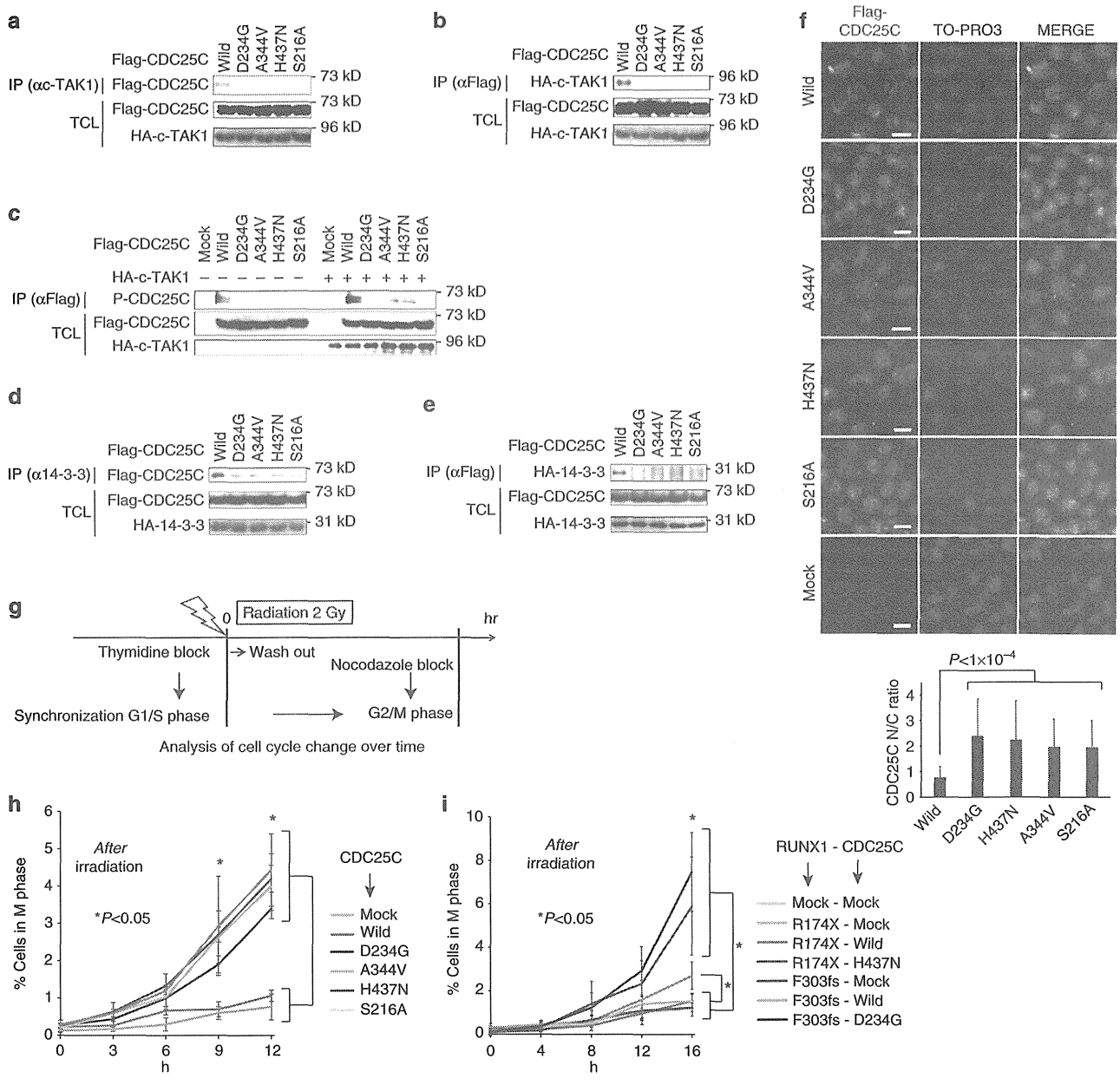

Figure 4 | Mutated CDC25C enhances mitotic entry. (**a**) HEK293T cells were transiently transfected with constructs encoding Flag-tagged CDC25C wild type or mutants, as indicated, and cell lysates were immunoprecipitated with anti-c-TAK1 antibody. Binding capacity of CDC25C was evaluated by western blotting. IP, immunoprecipitation; TCL, total cell lysate. (**b**) Reciprocal immunoprecipitation of **a** using anti-Flag (CDC25C) antibody for immunoprecipitation. (**c**) Left half; cell lysates were immunoprecipitated with anti-Flag antibody. Phosphorylation levels of CDC25C were assessed by phosphorylated-Ser216-specific anti-CDC25C antibody. Right half; the same experiment was performed with cell lysates from HEK293T cells transfected with constructs encoding Flag-tagged CDC25C wild type or mutants and HA-tagged c-TAK1. (**d**) Mutated CDC25C showed reduced capacity for binding to 14-3-3. Cell lysates were immunoprecipitated with anti-14-3-3 antibody and binding capacity of CDC25C was evaluated. (**e**) Reciprocal immunoprecipitation of **d** using anti-Flag (CDC25C) antibody for immunoprecipitation. (**f**) Localization of CDC25C or its mutants was visualized by immunofluorescence. Anti-Flag antibody and Alexa Fluor 555 antibody was used for visualization of CDC25C. N/C ratio of each cell was calculated as detailed in Supplementary Methods and Supplementary Fig. 10. The mean and s.d. of the N/C ratio is presented. Statistical significance of difference was determined by unpaired Student's *t*-test ($n>30$ for each). Scale bar, 10 μm. (**g**) Schematic description of the method used for evaluation of mitotic entry. (**h**) Mitotic entry of CDC25C-mutated cells. Percentage of mutated CDC25C-transduced cells in the M phase was compared with that of wild-type CDC25C-transduced cells. *P* values were calculated using Student's *t*-test and the differences between groups, as indicated, were all statistically significant ($^*P<0.05$) at 10 and 12 h after irradiation ($n=3$). The average and s.d. is presented. (**i**) Mutated RUNX1 and CDC25C were co-expressed in Ba/F3 cells, as indicated, and mitosis entry of these cells was evaluated. The differences between groups, as indicated, were all statistically significant ($^*P<0.05$ at 16 h after washout of thymidine ($n=3$). *P* values were determined using the Student's *t*-test. The average and s.d. is presented.

a germline *RUNX1* mutation. In addition, as Turowski and colleagues reported that CDC25C was involved in S phase entry in addition to mitotic entry[23], release from thymidine-induced G1/S block may be affected by some unknown machinery mediated by mutated CDC25Cs, which might affect the results when we observed G2/M phase fraction of these cells. It is not clear why *CDC25C* mutations are repetitively documented in FPD/AML, but not in sporadic MDS or AML cases. One possibility is that in the presence of a *RUNX1* mutation, as an initial event, an extended period is required before an additional *CDC25C* mutation is acquired. This proposal is supported by the clinical observation that ~40% of patients with FPD/AML develop leukaemia in their 30s[5]; however, the mutational status in *CDC25C* in the reported cohort was unknown.

One of the important problems in the research of FPD/AML is that definitive diagnostic criteria have not been established yet. For this purpose, more extensive studies are required for accumulating clinical characterization, genetic information and functional examination as to whether a *RUNX1* variant in families with thrombocytopenia and/or haematological malignancy is causal[24]. We clarified tentative diagnostic criteria for FPD/AML, which was used in this study (in Methods). Regarding the three missense variants in our study (p.Ser140Asn in pedigree 54, p.Gly172Glu in pedigree 57 and p.Leu445Pro in pedigree 32), Ser140 and Gly172 have been reported to be mutated in sporadic AML and/or MDS cases[25,26]. In addition, induced pluripotent stem cells from a FPD/AML pedigree with p.Gly172Glu recapitulate the phenotype of FPD/AML after hematopoietic differentiation[27]. Ser140 has been also shown to be important for RUNX1 conformation, and a mutation of this site affects hydrogen bonds and results in functional loss[28,29]. Furthermore, all the three missense variants have not been reported in the following SNP database: SNP database (dbSNP) (http://www.ncbi.nlm.nih.gov/projects/SNP), the 1000 Genomes Project (http://www.1000genomes.org), HGVB (http://www.genome.med.kyoto-u.ac.jp/SnpDB/index.html). They were also predicted as 'damaging' by Polyphen-2 (http://genetics.bwh.harvard.edu/pph2/), SIFT (http://sift.jcvi.org/) and PROVEAN (http://provean.jcvi.org/index.php). Therefore, we regarded the pedigrees with these *RUNX1* variants as having FPD/AML in this study. However, regarding pedigree 32 with p.Leu445Pro, we could not completely exclude the possibility of incidental co-occurrence of a possible non-causal *RUNX1* germline variant and hairy cell leukaemia, although co-occurrence of them is supposed to be rare. In addition, we should bear in mind the somatic as well as germline LOH of *RUNX1*, which contributes to thrombocytopenia and/or leukemogenesis in FPD/AML.

In conclusion, our results indicate that FPD/AML-associated leukaemic transformation is due to stepwise acquisition of mutations and clonal selection, which is initiated by a *CDC25C* mutation in the pre-leukaemic phase, and is further driven by mutations in other genes including *GATA2* (Supplementary Fig. 14). The identification of *CDC25C* as the target gene responsible for the leukaemic transformation will facilitate diagnosis and monitoring of individuals with FPD/AML, who are at an increased risk of developing life-threatening haematological malignancy.

Methods

Subjects. Studies involving human subjects were done in accordance with the ethical guidelines for biomedical research involving human subjects, which was developed by the Ministry of Health, Labour and Welfare, Japan; the Ministry of Education, Culture, Sports, Science, and Technology, Japan; and the Ministry of Economy, Trade, and Industry, Japan, and enforced on 29 March 2001. This study was approved by ethical committee of the University of Tokyo and each participating institution. Written informed consent was obtained from all patients whose samples were collected after the guideline was enforced. All animal experiments were approved by the University of Tokyo Ethics Committee for Animal Experiments. The clinical data, peripheral blood sample and buccal mucosa of the patients whose pedigree contained two or more individuals with thrombocytopenia and/or any haematological malignancies were collected from participating institutions. Platelet threshold depended on each institution's judge and any haematological malignancies were allowed. The diagnoses were self-reported. When all the following four criteria were fulfilled, the patient was considered as having FPD/AML in this study: (1) the pedigree has two or more individuals with thrombocytopenia and/or any haematological malignancies; (2) a germline *RUNX1* variant, including missense, nonsense, frameshift, insertion and deletion, is confirmed by Sanger sequencing and a synchronized quantitative-PCR method in at least one family member; (3) the *RUNX1* variant has not been reported in public dbSNP; (4) no germline mutations were detected in the following 16 genes: *GATA2, GATA1, CEBPA, MPL, MYH9, MYL9, GP1BA, GP9, MASTL, HOXA11, CBL, DIDO1, TERT, ANKRD26, GFI1B* and *SRP72*. Regarding the last criterion, 16 genes were selected because they have been reported to be responsible for familial thrombocytopenia and/or haematological malignancies.

Whole-exome sequencing. Genomic DNA was extracted from samples using the QIAamp DNA Mini kit (Qiagen). Exome capture was performed. Enriched exome fragments were subjected to sequencing using HiSeq2000 (Illumina). We removed any potential somatic mutations that were observed in dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP) or in the 1000 Genomes Project (http://www.1000genomes.org) data. All candidate single-nucleotide variations and indels, which were predicted to be deleterious by the Polyphen-2 algorithm, were validated by deep sequencing and Sanger sequencing. Genomic DNA samples from the buccal mucosa of the two patients (subject 20 and subject 21) were used as references. All candidate somatic mutations were validated by Sanger sequence and deep sequencing using primers listed in Supplementary Tables 3 and 4.

Deep sequencing. Using genomic DNA of the patients as template, each targeted region was PCR amplified with specific primers (Supplementary Table 4). The amplification products from an individual sample were combined and purified with the AMpure XP Kit (Beckman Coulter) and library preparation was carried out using the Ion Xpress Fragment Library Kit (Life Technologies) according to the manufacturer's instructions. The Agilent 2100 Bioanalyzer (Agilent Technologies) and the associated High Sensitivity DNA kit (Agilent Technologies) were used to determine quality and concentration of the libraries. The amount of the library required for template preparation was calculated using the template dilution factor calculation described in the protocol. Emulsion PCR and enrichment steps were carried out using the Ion OneTouch 200 Template Kit v2 DL (Life Technologies). Sequencing was undertaken using Ion Torrent PGM and Ion 318 chips Kit v2 (Life Technologies). The Ion PGM 200 Sequencing Kit (Life Technologies) was used for sequencing reactions, following the recommended protocol. The presence of *CDC25C* and *GATA2* mutations was also validated by a subclone strategy for DNA sequence analysis.

Single-cell sequencing and genome amplification. Single cells were separated from the bone marrow of subject 20 at AML phase using FACSAria II (BD biosciences) (Supplementary Fig. 15a). Each cell was deposited into individual wells of a 96-well plate. Single cells were lysed and whole genome from single cell was amplified using GenomePlex Single Cell Whole-Genome Amplification Kit (Sigma-Aldrich). Mutation status of each gene was analysed by direct sequencing with specific primers (Supplementary Table 5). To improve the sensitivity of this procedure, we used multiple primer sets for detecting a single-nucleotide variation. We estimated the false-negative rate of this procedure based on the ratio of RUNX1 mutation, which is supposed to be observed in all of the cells. The false-negative rate was estimated to be 35% (22 cells out of 63 cells, Supplementary Table 2), which is consistent with the manufacturer's bulletin reporting the allelic dropout of 30%. In light of these results, we regard those cells with at least one gene mutation in a mutational group (coloured in red, orange, green, blue or purple) as being positive for gene mutations of the corresponding group. To assess whether mutations in *LPP, FAM22G, COL9A1* and *GATA2* and mutations in *AGAP4, RP1L1, DTX2* and *CHEK2* were mutually exclusive, we performed a statistical analysis as follows. First of all, we determine a matrix **A** that virtually represents the mutational status of eight genes (1: *LPP*, 2: *FAM22G*, 3: *COL9A1*, 4: *GATA2*, 5: *AGAP4*, 6: *RP1L1*, 7: *DTX2* and 8: *CHEK2*) of 57 cells. Concretely, **A** is defined as follows:

$$\mathbf{A}=\begin{pmatrix} a_{1,1} & \cdots & a_{8,1} \\ \vdots & \ddots & \vdots \\ a_{1,57} & \cdots & a_{8,57} \end{pmatrix} \quad a_{i,j}=\begin{cases} 0 : \text{if gene i of cell j is wildtype} \\ 1 : \text{if gene i of cell j is mutated} \end{cases} \qquad (1)$$

On the other hand, a matrix **R** indicates data from the actual experimental results of mutational analysis as shown in Fig. 2c. Elements of **R** is provided in

Supplementary Table 2.

$$\mathbf{R}=\begin{pmatrix} r_{1,1} & \cdots & r_{8,1} \\ \vdots & \ddots & \vdots \\ r_{1,57} & \cdots & r_{8,57} \end{pmatrix} \quad r_{i,j}=\begin{cases} 0: \text{if gene i of cell j is wild type} \\ 1: \text{if gene i of cell j is mutated} \\ 2: \text{if mutational status of gene i of cell j is undetermined} \end{cases} \tag{2}$$

Then we assumed two hypotheses: H_0 and H_1.

H_0: the mutational status of genes $1\sim4$ and genes $5\sim8$ is independent. Each matrix elements of A are randomly assigned 0 or 1 (at ratio of 1:1) independently of each other.

H_1: mutations in genes $1\sim4$ and genes $5\sim8$ are mutually exclusive, and cells $1\sim40$ harbour mutations of genes $1\sim4$, while cells $41\sim57$ harbour mutations of genes $5\sim8$. In mathematical representation,

$$a_{i,j}=\begin{cases} 0: (5\le i\le 8 \text{ and } 1\le j\le 40) \text{ and } (1\le i\le 4 \text{ and } 41\le j\le 57) \\ 0 \text{ or } 1 \text{ randomely}: (1\le i\le 4 \text{ and } 1\le j\le 40) \text{ and } (5\le i\le 8 \text{ and } 41\le j\le 57) \end{cases} \tag{3}$$

We assumed matrices A_0 and A_1 that represent virtually generated mutational status under the hypotheses H_0 and H_1, and calculate the probability of substantializing R for given A_0 and A_1.

$P_0(\mathbf{R}/A_0)$ and $P_1(\mathbf{R}/A_1)$ can be calculated for given matrices A_0 and A_1 under the condition as follows:

Probability that we cannot determine whether a cell has mutation in gene X when the cell does not actually have a mutation; 28% (based on our data shown in Supplementary Table 2).

Probability that we judge that a cell has a mutation in gene X when the cell does not actually have a mutation; 5% (because it is very unlikely to happen).

Probability that we can judge correctly that a cell does not have a mutation in gene X when the cell does not actually have a mutation; 67% ($100-28-5=67\%$).

Probability that we cannot determine whether a cell has mutation in gene X when the cell actually has a mutation; 28% (based on our data shown in Supplementary Table 2).

Probability that we judge that a cell has a mutation in gene X when the cell actually has a mutation; 35% (the estimated false-negative rate based on the ratio of *RUNX1* mutation).

Probability that we can judge correctly that a cell has a mutation in gene X when the cell actually has a mutation; 37% ($100-28-35=37\%$).

Put it simply, $\mathbf{P}_0$ represents the probability that one can get the mutational profile R when a cell harbours mutations independently of each other, while $\mathbf{P}_1$ indicates the probability that R is realized under the condition where mutations in gene groups $1\sim4$ and $5\sim8$ are exclusive. Because A_0 and A_1 that meet the hypotheses H_0 and H_1 can be generated innumerably, we conducted a computational simulation to acquire the distribution of P_0 and P_1 by generating A_0 and A_1 100,000 times. For visibility, horizontal axis is converted to $-\ln(P)$.

Synchronized quantitative-PCR. These experiments were performed mostly as described previously[6]. Briefly, genomic DNA was denatured 95 °C for 5 min and iced immediately. Using the LightCycler 480 Instrument II (Roche), thermal cycling was performed with denatured genomic DNA, forward and reverse primers (Supplementary Table 6), THUNDERBIRD SYBR qPCR mix (TOYOBO). Threshold cycle scores were determined as the average of triplicate samples. We designed 27 primers for *RUNX1* and 3 reproducible primers (that is, primer RUNX-9, RUNX-19 and RUNX-20) were chosen by preparatory experiments. RPL5-2 and PRS7-1 primers, which were authorized previously[6], were also utilized as controls. In addition, genomic DNA extracted from the bone marrow sample of a MDS patient with a chromosome 21 deletion was also examined with the same primers as a control of *RUNX1* locus copy-number loss. Crossing points (Cps) of designed primers were examined by quantitative PCR. *RUNX1* locus copy-number relative to RPL5-2 was calculated using Cps of RUNX-9 and RPL5-2, with RPL5-2 values set at 2. Similar results were obtained when Cps of RUNX-19, RUNX-20 or RPS7-1 values were used.

LOH detection with SNP sequencing. To examine the existence of uniparental disomy, we designed four specific primers to detect nine SNPs in *RUNX1*, which are frequently seen (>40%) (Supplementary Table 7). Direct sequencing was performed with the primers, and heterogeneity of SNPs was examined.

Chemicals and immunological reagents. Thymidine and nocodazole were purchased from Sigma-Aldrich. Anti-CDC25C, anti-phospho-CDC25C (Ser216) and anti-beta-actin antibodies were purchased from Cell Signaling Technology. Anti-HA monoclonal antibody was purchased from MBL. Rabbit anti-Flag monoclonal antibody was purchased from Sigma-Aldrich. Anti-HA was purchased from Roche. Mouse anti-phospho-histone H2AX (Ser139) antibody and Alexa Fluor 488 mouse anti-phospho-H3 (Ser10) antibody were purchased from Merck Millipore. Alexa Fluor 488 rabbit anti-mouse immunoglobulin (Ig)G, Alexa Fluor 488 goat anti-rabbit IgG and Alexa Fluor 555 goat anti-rabbit IgG were purchased from Invitrogen. TO-PRO3 was purchased from Molecular Probes. Rabbit anti-14-3-3 Sigma antibody was purchased from Bethyl laboratories. Sheep anti-c-TAK1 antibody was purchased from Exalpha Biologicals. Anti-sheep IgG-HRP was purchased from RSD. Nonviable cell exclusion was performed by 7-AAD Viability Staining Solution (BioLegend).

Subclone strategy and direct sequencing. Using genomic DNA of the patients as template, each targeted region was amplified by PCR with specific primers (Supplementary Table 4). PCR products were purified with illustra ExoStar (GE Healthcare) and subcloned into *Eco*RV site of pBluescript II KS(−) (Stratagene). Ligated plasmids were transformed into *E. coli* strain XL1-Blue by 45 s heat shock at 42 °C. Positive transformants were incubated on LB plates containing 100 μg ml^{-1} ampicillin supplemented with X-gal (Sigma-Aldrich) and isopropyl β-D-1-thiogalactopyranoside (Sigma-Aldrich). For colony PCR, a portion of a white colony was directly added to a PCR mixture as the DNA template. Insert region was amplified by PCR procedure with T3 and T7 universal primers, purified with illustra ExoStar (GE Healthcare Life Sciences), and sequenced by the Sanger method with T3 and T7 primers using BigDye Terminator v3.1 Cycle Sequencing kit (Applied Biosystems) and ABI Prism 310 Genetic Analyzer (Life Technologies).

Immunoprecipitation and western blotting. These experiments were performed as described previously[30]. Briefly, HEK293T cells were transiently transfected with mammalian expression plasmids encoding Flag-tagged CDC25C and its mutants, HA-tagged 14-3-3 or c-TAK1. All plasmids were sequence verified. After 48 h, cell lysates were collected and incubated with an antibody (anti-HA antibody (1:200, 3 h), anti-Flag antibody (1:200, 3 h), anti-c-TAK1 antibody (1:150, 3 h) and anti 14-3-3 antibody (1:150, 3 h)). After incubation, the cell lysates were incubated with protein G-Sepharose (GE Healthcare) for 1 h. The precipitates were stringently washed with high salt-containing wash buffer and analysed by western blotting. Anti-Flag (HRP-conjugated, Sigma-Aldrich), anti-HA (MBL), anti-HA (HRP-conjugated, Roche), anti-CDC25C (Cell Signaling Technology), anti-phospho-CDC25C (Ser216) (Cell Signaling Technology), anti-c-TAK1 antibody (Exalpha Biologicals) or anti-14-3-3 antibody (Bethyl laboratories) antibodies and Immunostar LD (Wako) was used for detection. Original gel images of western blot analysis are shown in Supplementary Fig. 16.

Cell cycle synchronization and analysis for mitosis entry. After transduction of wild-type CDC25C or its mutated forms to murine lymphoid cell line Ba/F3 cells (RIKEN BioResource Center), double-thymidine block was performed to obtain cell cycle synchronization at G1/S phase. In brief, 2 mM of thymidine was added to the medium. After 16 h, cells were washed and released from the first thymidine for 8 h. A second block was initiated by adding 2 mM of thymidine, and cells were maintained for 16 h. Then thymidine was washed out and the cells were incubated with 1 mM nocodazole with or without 2 Gy of irradiation (Supplementary Fig. 10a). Ba/F3 cells were fixed over time with 75% ethanol in phosphate-buffered saline (PBS) at 4 °C overnight and permeabilized with 2% Triton-X at 4 °C for 15 min. The cells were stained with anti-phospho-H3 (Ser10) Alexa Fluor 488 conjugated antibody (dilution, 1:200) in PBS with 2% fetal calf serum at 4 °C for 30 min and then treated with 5% propidium iodide and 1% RNase in PBS at room temperature (RT) for 30 min. Cell cycle was analysed using a BD LSR II Flow cytometer (BD biosciences) (Supplementary Fig. 15b). To assess the cooperation of *CDC25C* and *RUNX1* mutation, wild-type or mutant (D234G, H437N) pMXs-neo-Flag-CDC25C and mutant (F303fsX566, R174X) pGCDNsam-IRES-KusabiraOrange-Flag-RUNX1 were retrovirally transduced into Ba/F3 cells.

Immunofluorescent microscopic analysis. These experiments were performed as described previously[30]. Briefly, Ba/F3 cells were fixed, permeabilized and blocked. Staining for phosphorylated histone H2AX was performed with anti-phospho-histone H2AX (Ser139) antibody (dilution, 1:500; Merck Millipore) at RT for 3 h. After washing with PBS three times and with 1% bovine serum albumin in PBS, the cells were treated with Alexa Fluor 488 rabbit anti-mouse IgG (dilution, 1:500; Invitrogen) and TO-PRO3 (dilution, 1:1,000; Molecular Probes) for 1 h. The proteins were visualized using FV10i (Olympus) or BZ-9000 (Keyence). The percentage of γH2AX foci-positive cells was determined by examining 100 cells per sample. Three independent experiments were performed. To evaluate the localization of CDC25C, Ba/F3 cells were treated with 2 mM thymidine for 12 h and stained. Staining was underwent with anti-Flag antibody or anti-CDC25C antibody at RT for 3 h. After washing, the cells were treated with Alexa Fluor 488 or 555 antibody and TO-PRO3 for 1 h. The mean intensity of CDC25C in the nucleus and cytoplasm of each cell was measured within a region of interest placed within the nucleus and cytoplasm (Supplementary Fig. 10). Similarly, the background intensity was quantified within the region of interest placed outside the cells. All the measurements were performed using the Fluoview FV10i software or ImageJ. The background-subtracted intensity ratio of the nucleus to cytoplasm was calculated in >30 cells in each specimen.

Retrovirus production. The procedures were performed as described previously[30]. Briefly, Plat-E packaging cells were transiently transfected with each retroviral construct using the calcium phosphate precipitation method, and supernatant

containing retrovirus was collected 48 h after transfection and used for infection after it was centrifuged overnight at 10,000 r.p.m.

Statistical analysis. To compare data between groups, unpaired Student's *t*-test was used when equal variance were met by the *F*-test. When unequal variances were detected, the Welch *t*-test was used. Differences were considered statistically significant at a *P* value of <0.05.

References

1. Song, W. J. *et al.* Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. *Nat. Genet.* **23,** 166–175 (1999).
2. Ichikawa, M. *et al.* A role for RUNX1 in hematopoiesis and myeloid leukemia. *Int. J. Hematol.* **97,** 726–734 (2013).
3. Cameron, E. R. & Neil, J. C. The Runx genes: lineage-specific oncogenes and tumor suppressors. *Oncogene* **23,** 4308–4314 (2004).
4. Nickels, E. M., Soodalter, J., Churpek, J. E. & Godley, L. A. Recognizing familial myeloid leukemia in adults. *Ther. Adv. Hematol.* **4,** 254–269 (2013).
5. Liew, E. & Owen, C. Familial myelodysplastic syndromes: a review of the literature. *Haematologica* **96,** 1536–1542 (2011).
6. Kuramitsu, M. *et al.* Extensive gene deletions in Japanese patients with diamond-blackfan anemia. *Blood* **119,** 2376–2384 (2012).
7. Kirito, K. *et al.* A novel RUNX1 mutation in familial platelet disorder with propensity to develop myeloid malignancies. *Haematologica* **93,** 155–156 (2008).
8. Boutros, R., Lobjois, V. & Ducommun, B. CDC25 phosphatases in cancer cells: key players? Good targets? *Nat. Rev. Cancer* **7,** 495–507 (2007).
9. Kastan, M. B. & Bartek, J. Cell-cycle checkpoints and cancer. *Nature* **432,** 316–323 (2004).
10. Peng, C. Y. *et al.* C-TAK1 protein kinase phosphorylates human Cdc25C on serine 216 and promotes 14-3-3 protein binding. *Cell Growth Differ.* **9,** 197–208 (1998).
11. Lopez-girona, A., Furnari, B., Mondesert, O. & Early, P. R. Nuclear localization of Cdc25 is regulated by DNA damage and a 14-3-3 protein. *Nature* **397,** 172–175 (1999).
12. Satoh, Y., Matsumura, I., Tanaka, H. & Harada, H. C-terminal mutation of RUNX1 attenuates the DNA-damage repair response in hematopoietic stem cells. *Leukemia* **26,** 303–311 (2011).
13. Krejci, O. *et al.* p53 signaling in response to increased DNA damage sensitizes AML1-ETO cells to stress-induced death. *Blood* **111,** 2190–2199 (2008).
14. Park, J. *et al.* Mutation profiling of mismatch repair-deficient colorectal cancers using an in silico genome scan to identify coding microsatellites advances in brief mutation profiling of mismatch repair-deficient colorectal cancers using an in silico genome scan to Ide. *Cancer Res.* **62,** 1284–1288 (2002).
15. Vassileva, V., Millar, A., Briollais, L., Chapman, W. & Bapat, B. Genes involved in DNA repair are mutational targets in endometrial cancers with microsatellite instability. *Cancer Res.* **62,** 4095–4099 (2002).
16. Greif, P. A. *et al.* GATA2 zinc finger 1 mutations associated with biallelic CEBPA mutations define a unique genetic entity of acute myeloid leukemia. *Blood* **120,** 395–403 (2012).
17. Ostergaard, P. *et al.* Mutations in GATA2 cause primary lymphedema associated with a predisposition to acute myeloid leukemia (Emberger syndrome). *Nat. Genet.* **43,** 929–931 (2011).
18. Hahn, C. N. *et al.* Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. *Nat Genet.* **43,** 1012–1017 (2011).
19. Hsu, A. P. *et al.* Mutations in GATA2 are associated with the autosomal dominant and sporadic monocytopenia and mycobacterial infection (MonoMAC) syndrome. *Blood* **118,** 2653–2655 (2011).
20. Dickinson, R. E. *et al.* Exome sequencing identifies GATA-2 mutation as the cause of dendritic cell, monocyte, B and NK lymphoid deficiency. *Blood* **118,** 2656–2658 (2011).
21. Zhang, S.-J. *et al.* Gain-of-function mutation of GATA-2 in acute myeloid transformation of chronic myeloid leukemia. *Proc. Natl Acad. Sci. USA* **105,** 2076–2081 (2008).
22. Hasegawa, D. *et al.* CBL mutation in chronic myelomonocytic leukemia secondary to familial platelet disorder with propensity to develop acute myeloid leukemia (FPD/AML). *Blood* **119,** 2612–2614 (2012).
23. Turowski, P. *et al.* Functional cdc25C dual-specificity phosphatase is required for S-phase entry in human cells. *Mol. Biol. Cell* **14,** 2984–2998 (2003).
24. Michaud, J. *et al.* In vitro analyses of known and novel RUNX1/AML1 mutations in dominant familial platelet disorder with predisposition to acute myelogenous leukemia: Implications for mechanisms of pathogenesis. *Blood* **99,** 1364–1372 (2002).
25. Kohlmann, A. *et al.* Monitoring of residual disease by next-generation deep-sequencing of RUNX1 mutations can identify acute myeloid leukemia patients with resistant disease. *Leukemia* **28,** 129–137 (2014).
26. Chen, C. Y. *et al.* RUNX1 gene mutation in primary myelodysplastic syndrome - The mutation can be detected early at diagnosis or acquired during disease progression and is associated with poor outcome. *Br. J. Haematol.* **139,** 405–414 (2007).
27. Sakurai, M. *et al.* Impaired hematopoietic differentiation of RUNX1-mutated induced pluripotent stem cells derived from FPD/AML patients. *Leukemia.* (epub ahead of print 15 April 2014; doi:10.1038/leu.2014.136).
28. Bravo, J., Li, Z., Speck, N. A. & Warren, A. J. The leukemia-associated AML1 (Runx1)-CBF beta complex functions as a DNA-induced molecular clamp. *Nat. Struct. Biol.* **8,** 371–378 (2001).
29. Akamatsu, Y., Tsukumo, S. I., Kagoshima, H., Tsurushita, N. & Shigesada, K. A simple screening for mutant DNA binding proteins: application to murine transcription factor PEBP2?? subunit, a founding member of the Runt domain protein family. *Gene* **185,** 111–117 (1997).
30. Yoshimi, A. *et al.* Evi1 represses PTEN expression and activates PI3K/AKT/mTOR via interactions with polycomb proteins. *Blood* **117,** 3617–3628 (2011).

Acknowledgements

This work was supported in part by grants-in-aid from the Ministry of Health, Labor and Welfare of Japan (H23-Nanchi-Ippan-104; M. Kurokawa) and KAKENHI (24659457; M. Kurokawa). We thank R. Lewis (University of Nebraska Medical Center) and T. Kitamura (Institute of Medical Science, The University of Tokyo) for providing essential materials; T. Koike (Nagaoka Red Cross Hospital), K. Nara (Ootemachi Hospital), K. Suzuki (Japanese Red Cross Medical Center), H. Harada (Fujigaoka Hospital), Y. Morita (Kinki University), M. Matsuda (PL Hospital), H. Kashiwagi (Osaka University), T. Kiguchi (Chugoku Central Hospital), T. Masunari (Chugoku Central Hospital), K. Yamamoto (Yokohama City Minato Red Cross Hospital), T. Takahashi (Mitsui Memorial Hospital) and T. Takaku (Juntendo University) for providing patient samples; M. Kuramitsu (National Institutte of Infectious Diseases) for providing kind support of synchronized quantitative PCR; and K. Tanaka and Y. Shimamura for their technical assistance.

Author contributions

A.Y., T.T., M.I. and M. Kurokawa analysed genetic materials and performed functional studies. A.T., H.I., M.N., Y.N. and S.A. were involved in sequencing and/or functional studies. M. Kawazu, T.U. and H.M. took part in whole-exome sequencing, deep sequencing and bioinformatics analyses of the data. A.Y., T.T., M.I., H.H., K.U., Y.H., E.I., K.K. and H.N. collected specimens. A.Y. and T.T. generated figures and tables. M. Kurokawa designed and led the entire project. A.Y., T.T. and M. Kurokawa wrote the manuscript. All authors participated in the discussion and interpretation of the data.

Additional information

Accession codes: Sequence data for FPD/AML patients has been deposited in GenBank/EMBL/DDBJ sequence read archive (SRA) under the accession code SRP043031

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions

How to cite this article: Yoshimi, A. *et al.* Recurrent *CDC25C* mutations drive malignant transformation in FPD/AML. *Nat. Commun.* 5:4770 doi: 10.1038/ncomms5770 (2014).

Leukemia (2014) **28**, 2344–2354

www.nature.com/leu

ORIGINAL ARTICLE

Impaired hematopoietic differentiation of *RUNX1*-mutated induced pluripotent stem cells derived from FPD/AML patients

M Sakurai[1], H Kunimoto[1], N Watanabe[2], Y Fukuchi[1], S Yuasa[3], S Yamazaki[4], T Nishimura[5], K Sadahira[1], K Fukuda[3], H Okano[6], H Nakauchi[4,5], Y Morita[7], I Matsumura[7], K Kudo[8], E Ito[8], Y Ebihara[9], K Tsuji[9,10], Y Harada[11,12], H Harada[11,12], S Okamoto[1] and H Nakajima[1]

Somatic mutation of *RUNX1* is implicated in various hematological malignancies, including myelodysplastic syndrome and acute myeloid leukemia (AML), and previous studies using mouse models disclosed its critical roles in hematopoiesis. However, the role of *RUNX1* in human hematopoiesis has never been tested in experimental settings. Familial platelet disorder (FPD)/AML is an autosomal dominant disorder caused by germline mutation of *RUNX1*, marked by thrombocytopenia and propensity to acute leukemia. To investigate the physiological function of *RUNX1* in human hematopoiesis and pathophysiology of FPD/AML, we derived induced pluripotent stem cells (iPSCs) from three distinct FPD/AML pedigrees (FPD-iPSCs) and examined their defects in hematopoietic differentiation. By *in vitro* differentiation assays, FPD-iPSCs were clearly defective in the emergence of hematopoietic progenitors and differentiation of megakaryocytes, and overexpression of wild-type (WT)-*RUNX1* reversed most of these phenotypes. We further demonstrated that overexpression of mutant-*RUNX1* in WT-iPSCs did not recapitulate the phenotype of FPD-iPSCs, showing that the mutations were of loss-of-function type. Taken together, this study demonstrated that haploinsufficient *RUNX1* allele imposed cell-intrinsic defects on hematopoietic differentiation in human experimental settings and revealed differential impacts of *RUNX1* dosage on human and murine megakaryopoiesis. FPD-iPSCs will be a useful tool to investigate mutant *RUNX1*-mediated molecular processes in hematopoiesis and leukemogenesis.

Leukemia (2014) **28,** 2344–2354; doi:10.1038/leu.2014.136

INTRODUCTION

RUNX1 is a founding member of Runt-family transcription factors, which was cloned from a break point of t(8;21) chromosomal translocation observed in acute myeloid leukemia (AML). Studies over a decade have revealed critical roles of *RUNX1* in hematopoiesis. During embryonic development, *Runx1* is absolutely essential in the emergence of hematopoietic stem and progenitor cells through hemogenic endothelium. In contrast, conditional disruption of *Runx1* in adult hematopoietic system revealed that it was critical in the differentiation of megakaryocytes (MgKs) and lymphocytes as well as in the homeostasis of hematopoietic stem cells.[1] However, these results were mostly derived from gene-disruption studies in mice, and the role of *RUNX1* in human hematopoiesis has never been tested in experimental settings.

Somatic mutation of *RUNX1* has been implicated in a variety of hematological malignancies, including myelodysplastic syndrome (MDS) and AML. It was found in 15–35% cases of AML M0 subtype,[2] 10–20% of MDS,[3–6] 37% of chronic myelomonocytic leukemia[7] and 14% of MDS/myeloproliferative neoplasm,[8] which makes *RUNX1* as one of the most frequently mutated genes in hematological malignancies.

The mutations are distributed throughout *RUNX1* protein, being roughly classified into two categories, missense mutation in N-terminal Runt-homology domain and frame-shift or non-sense mutations leading to a C-terminal truncation. Runt-homology domain mutation impairs DNA-binding and nuclear localization, while the C-terminal truncation disrupts transcriptional activation or repression activity.[3,9] Biochemical studies have shown that most of the *RUNX1* mutations observed in MDS or AML are loss-of-function mutation or dominant negative to the residual wild-type (WT) allele.[9–12] However, their capacities to suppress transcriptional activity of WT-*RUNX1* vary among mutants by *in vitro* assays.[3] In addition, it is possible that some mutations confer the protein with non-physiological functions acting as gain-of-function mutants. Therefore, precise function of mutant *RUNX1* in MDS or AML remains obscure, and it must be tested in physiological and, ideally, in human settings.

Familial platelet disorder/AML(FPD/AML) is a rare autosomal dominant disorder caused by germline mutation of *RUNX1*, marked by thrombocytopenia and propensity to acute leukemia.[13] Approximately 30 FPD/AML pedigrees have been reported to date, and the affected patients retained *RUNX1* mutation at similar positions as reported in MDS and AML. Most patients present no

[1]Division of Hematology, Department of Internal Medicine, Keio University School of Medicine, Tokyo, Japan; [2]Department of Transfusion Medicine and Cell Therapy, Keio University School of Medicine, Tokyo, Japan; [3]Division of Cardiology, Department of Internal Medicine, Keio University School of Medicine, Tokyo, Japan; [4]Japan Science and Technology Agency, ERATO, Tokyo, Japan; [5]Division of Stem Cell Therapy, Center for Stem Cell Biology and Regenerative Medicine, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan; [6]Department of Physiology, Keio University School of Medicine, Tokyo, Japan; [7]Division of Hematology and Rheumatology, Department of Internal Medicine, Kinki University Faculty of Medicine, Osaka, Japan; [8]Department of Pediatrics, Hirosaki University Graduate School of Medicine, Aomori, Japan; [9]Department of Pediatric Hematology/Oncology, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan; [10]Division of Stem Cell Processing, Center for Stem Cell Biology and Regenerative Medicine, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan and [11]Department of Hematology and Oncology, Division of Clinical Research, Research Institute for Radiation Biology and Medicine, Hiroshima University, Hiroshima, Japan. Correspondence: Professor H Nakajima, Division of Hematology, Department of Internal Medicine, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan.
E-mail: hnakajim@z2.keio.jp
[12]Present address: Department of Hematology, Juntendo University School of Medicine, Tokyo, Japan.
Received 30 October 2013; revised 30 March 2014; accepted 9 April 2014; accepted article preview online 15 April 2014; advance online publication, 13 May 2014

evident clinical symptoms or developmental abnormalities except mild thrombocytopenia from their childhood. However, approximately half of FPD/AML patients develop MDS or acute leukemia after a long latency, generally after the third decade of their lives. These facts together with insights from mouse studies[14] clearly indicate that *RUNX1* mutation *per se* is not sufficient for leukemia development, but it establishes preleukemic state that predisposes cells to full-blown leukemia by acquiring additional genetic events.[15,16] It is therefore expected that studying the pathogenesis of FPD/AML would provide a valuable insight into the molecular mechanism of leukemia or MDS with *RUNX1* mutation. However, rarity of FPD/AML pedigrees and limited opportunity to obtain their patient samples has tremendously hampered the study.

Induced pluripotent stem cells (iPSCs) provide us with novel opportunities for disease modeling and drug discovery[17,18] Hematopoietic differentiation of iPSCs can be induced by co-culture on stromal cells, and iPSC-derived hematopoietic progenitors can be used further for recapitulating disease phenotypes.[19–27] As iPSCs can be an indefinite source for differentiated cells, they are particularly useful when disease samples cannot be easily obtained from patients.

To investigate the physiological function of *RUNX1* in human hematopoiesis and the pathophysiology of FPD/AML, we derived iPSCs from three distinct FPD/AML pedigrees (FPD-iPSCs) and examined their defects in the emergence of hematopoietic progenitor cells (HPCs) and hematopoietic differentiation. These pedigrees have distinct heterozygous mutations in *RUNX1* gene, two in the N-terminal RUNT domain affecting its DNA-binding activity and one in the C-terminal region affecting its transactivation capacity. Three FPD-iPSC lines uniformly presented a variety of defects in the emergence in HPCs and MgK differentiation, which were rescued by overexpression of WT-*RUNX1*. We further demonstrated that overexpression of mutant-*RUNX1* in WT-iPSCs did not recapitulate the phenotype of FPD-iPSCs, showing that the mutations were of loss-of-function type. Taken together, this study, for the first time, demonstrated that haploinsufficient *RUNX1* allele imposed cell-intrinsic defects on the emergence of HPCs and MgK differentiation in human experimental settings and revealed differential impacts of *RUNX1* dosage on human and murine megakaryopoiesis.

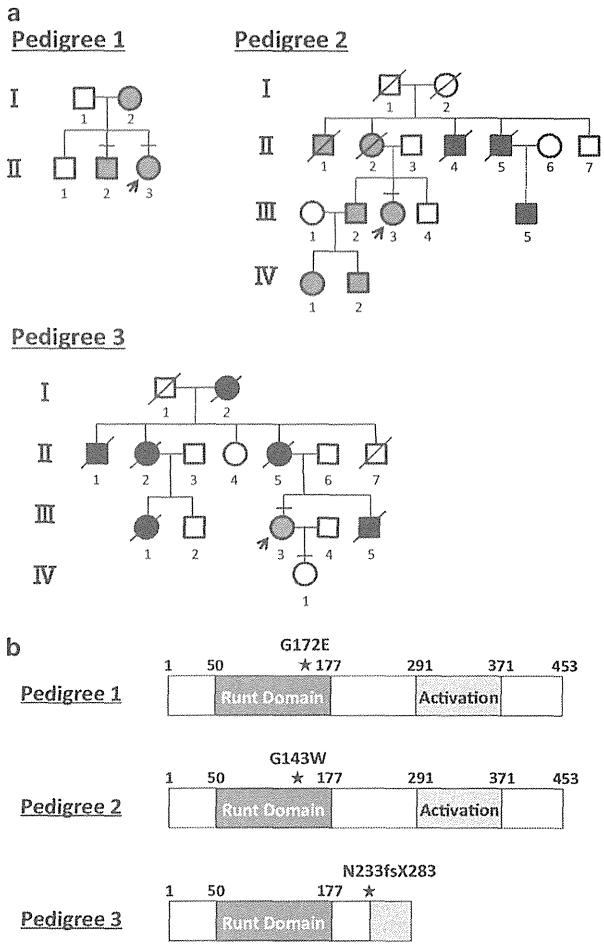

Figure 1. Pedigrees of FPD/AML and sequence analysis of *RUNX1* gene. (**a**) Family trees of three FPD/AML pedigrees. Open symbols, unaffected individuals; gray symbols, patients with thrombocytopenia; black symbols, patients who developed acute leukemia or MDS. Slash lines represent deceased individuals. Arrows denote patients enrolled in the study. (**b**) Summary of *RUNX1* mutations identified in FPD/AML pedigrees. Amino acid numbers are shown on the top of each panel. Positions of mutated amino acids are shown in filled stars. An area with oblique lines denotes irrelevant amino-acid sequence added due to a frame-shift mutation. Activation; transactivation domain.

MATERIALS AND METHODS

Patients

Family trees of three FPD/AML pedigrees are depicted in Figure 1a. Peripheral blood samples from affected patients were collected after obtaining written informed consent. The study was conducted with approval from the internal review board of Keio University School of Medicine, Tokyo, Japan and conformed to the principles outlined in the Declaration of Helsinki for use of human tissue or subjects.

DNA sequence

Genomic DNA was purified by phenol–chloroform method or by QIAamp DNA Micro Kit (Qiagen, Tokyo, Japan) according to the manufacturer's protocol. RUNX1 mutations of iPSCs were verified by direct sequencing of PCR product of *RUNX1* gene amplified from genomic DNA of iPSCs. PCR was performed by PicoMaxx high-fidelity PCR system (Agilent Technologies, Santa Clara, CA, USA) with primers for *RUNX1* that were previously described.[3] PCR products were purified and subjected to direct sequencing by using BigDye Terminator v1.1 Cycle sequencing kit (Life Technologies Japan, Tokyo, Japan) and ABI Prism 310 Genetic Analyzer (Life Technologies Japan).

Generation of iPSCs and cell culture

iPSCs were established from peripheral T cells obtained from patients. Detailed protocol for generating iPSCs from human peripheral blood mononuclear cells was previously described.[28] Control iPSCs were generated from peripheral T cells of healthy male donors after informed consent. Established iPSCs were maintained on inactivated mouse embryonic fibroblasts in iPSC medium, and cells were passaged by treating cells with 1 mg/ml collagenase IV (Life Technologies Japan) every 5–6 days.

Reverse transcription-PCR (RT-PCR)

Total RNA was isolated using TRIZOL reagent (Life Technologies Japan) according to the manufacturer's instructions. cDNA was reverse-transcribed using SuperScript II reverse transcriptase (Life Technologies Japan). PCR was performed using PicoMaxx high-fidelity PCR system (Agilent Technologies) as previously described.[29] Quantitative RT-PCR (qRT-PCR) was performed as previously described.[29] Primer sequences are listed in Supplementary Table S1.

Teratoma formation assay

iPSCs (1×10^7) were injected into the testis of NOD-SCID mice (CLEA Japan, Tokyo, Japan) under anesthesia with pentobarbital sodium (Kyoritsu Seiyaku Corporation, Tokyo, Japan). Eleven weeks after injection, tumors were dissected and fixed in 4% paraformaldehyde in phosphate-buffered-

saline. Fixed tissues were then embedded in paraffin, sectioned and stained with hematoxylin and eosin for analysis.

Immunofluorescence staining

Immunofluorescence staining was performed using the following primary antibodies: anti-NANOG (Abcam, Cambridge, MA, USA), anti-OCT3/4 (Santa Cruz Biotechnology, Dallas, TX, USA), anti-SSEA 3 (Millipore, Billerica, MA, USA), anti-SSEA 4 (Millipore), anti-Tra1-60 (Millipore), and anti-Tra1-81 (Millipore). The secondary antibodies used were: anti-mouse immunoglobulin G (IgG), anti-mouse IgM, anti-rabbit IgG, and anti-rat IgM monoclonal antibodies conjugated with Alexa Fluor 488 or 594 (Life Technologies Japan). Fluorescent images were captured using fluorescence microscope (IX70, Olympus, Tokyo, Japan) with CCD camera (DP70, Olympus).

Hematopoietic differentiation of iPSCs

We used AGM-S3 co-culture [3031] or embryonic stem (ES) sac protocols[32] to assess hematopoietic differentiation of iPSCs as previously described. For AGM-S3 co-culture, iPSCs were plated onto inactivated AGM-S3 cells and cultured for 2–3 days with iPSC medium. On day 2 or 3, medium was replaced with Iscove's modified Dulbecco's medium (Sigma, St Louis, MO, USA) supplemented with 10% fetal bovine serum (Sigma), 20 ng/ml human vascular endothelial growth factor (PeproTech, Rocky Hill, NJ, USA), 1% nonessential amino acid solution (Life Technologies Japan), 100 μM 2-ME (Wako, Osaka, Japan) and 1 mM L-glutamine (Wako). Hematopoietic cells recognized as 'cobble-stone' area surrounding iPS colonies emerge on day 10–14 of co-culture, which are then harvested using 0.05% trypsin/EDTA (Wako) for further experiments.

For ES sac formation, small iPS colonies were transferred onto irradiated C3H10T1/2 cells and co-cultured in Iscove's modified Dulbecco's mediu supplemented with 15% fetal bovine serum, 10 μg/ml human insulin, 5.5 μg/ml human transferrin, 5 ng/ml sodium selenite (Sigma), 2 mM L-glutamine (Life Technologies Japan), 0.45 mM α-monothioglycerol (Sigma), 50 μg/ml ascorbic acid (Sigma) and 20 ng/ml human vascular endothelial growth factor. On days 14 – 15 of culture, sac-like structure containing hematopoietic cells (iPS-sac) formed on the feeders were manually collected into 50-ml tubes, gently crushed with pipetting and passed through a 40-μm cell strainer to obtain hematopoietic progenitors.

Colony-forming assay

$CD34^+$ cells derived from iPSCs by AGM-S3 co-culture were sorted by flow cytometry and subjected to colony-forming assays using Methocult H4435 (Stem Cell Technologies, Vancouver, BC, Canada). Numbers and types of colonies were assessed on day 14.

Flow cytometry

Cells were stained in phosphate-buffered-saline/5% fetal bovine serum with the following monoclonal antibodies: anti-CD34-fluorescein isothiocyanate (FITC), anti-CD45-phycoerythrin (PE), anti-CD31-PE, anti-CD41a-PE, anti-CD42b-FITC (BD Pharmingen, San Jose, CA, USA), anti-glycophorin A (GPA)-FITC and anti-CD43- allophycocyanin (BioLegend, San Diego, CA, USA). Stained cells were analyzed by fluorescence-activated cell sorting (FACS) Calibur with the CellQuest software (BD Biosciences, San Jose, CA, USA) or sorted by MoFlo (Beckman Coulter, Brea, CA, USA). The data were analyzed by the FlowJo software (Tomy Digital Biology, Tokyo, Japan).

Differentiation of MgKs from $CD34^+$ cells

$CD34^+$ cells generated from iPSCs were sorted by flow cytometry and cultured in minimum essential medium alpha (Life Technologies Japan) supplemented with 10% bovine serum albumin, 100 μM 2-ME, 100 ng/ml stem cell factor (PeproTech) and 10 ng/ml thrombopoietin (PeproTech) at 37 °C under hypoxia condition (5% O_2). After 3 days of culture, cells were counted, harvested and analyzed by flow cytometry.

Stable transfection of iPSCs

pcDNA3.1/Flag-WT-*RUNX1* expressing human WT-*RUNX1b* isoform, pcDNA3.1/Flag-$RUNX1^{G172E}$ or pcDNA3.1/$RUNX1^{N233fsX283}$ was transfected into iPSCs using FuGENE HD transfection reagent (Promega, Madison, WI, USA) according to the manufacturer's protocol. After 2 days of transfection, stable transformants were selected in human ES medium supplemented with 100 μg/ml of G418 (Roche, Basel, Switzerland). Surviving colonies were picked up around day 14 of selection and subjected to further analyses of mRNA and protein expression.

Western blotting

Preparation of protein extracts and western blotting were performed as previously described.[33] Briefly, iPSCs were lysed in the lysis buffer (1% Nonidet P-40; 20 mM Tris-HCl, PH7.5; 150 mM NaCl; 1 mM phenylmethylsulfonyl floride; 1 μg/ml leupeptin). Proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and transferred to PROTRAN BA85 membrane (Schleicher and Schuell, Dassel, Germany). Membranes were blocked with 5% non-fat milk in TBS-T (0.1% Tween-20) and hybridized with anti-FLAG M2 antibody (Sigma), anti-RUNX1 rabbit polyclonal antibody (gift from H Harada) or anti-α-tubulin monoclonal antibody (Sigma) followed by a horseradish peroxidase-conjugated anti-mouse or anti-rabbit immunoglobulin G secondary antibody (GE Healthcare, Pittsburgh, PA, USA). Bound antibodies were detected by enhanced chemiluminescence (GE Healthcare).

Statistical analysis

All statistical analyses were performed using unpaired Student's *t*-test. *P*-values < 0.05 were considered statistically significant.

RESULTS

Derivation of iPSCs from patients with FPD/AML

In order to investigate the physiological function of *RUNX1* in human hematopoiesis and the pathophysiology of FPD/AML, we derived iPSCs from three distinct FPD/AML pedigrees to examine their defects in the emergence of blood cells and hematopoietic differentiation. Three FPD/AML pedigrees that we utilized in the study are depicted in Figure 1a. These pedigrees carried distinct heterozygous mutations in *RUNX1* gene, two in the N-terminal RUNT domain and one in the C-terminal region (Figure 1b). N-terminal mutations in RUNT domain in pedigrees 1 and 2 (G172E and G143W) were considered to affect DNA-binding activity of RUNX1, and C-terminal mutation in pedigree 3 (N233fsX283) was reported to abrogate the transactivation/ repression capacity.

After obtaining informed consent from the affected patients, we established iPSCs from their peripheral T cells by infecting Sendai viruses expressing four reprogramming factors (*OCT3/4*, *SOX2*, *KLF4* and *c-MYC*) (Figure 2a).[34] FPD-iPSCs could be established in comparable frequency as the one from normal individuals (WT-iPSCs), and their behavior in the culture and the morphology of the colonies were indistinguishable from those of WT-iPSCs (Figure 2b). We confirmed that each iPSCs harbored the same *RUNX1* mutation identified in somatic cells of the original patients (Figure 2c). Initial characterization of FPD-iPSCs revealed that the established clones retained typical characteristics of pluripotent stem cells such as the expression of immature ES cell markers (for example, Nanog, Oct3/4, SSEA-3, SSEA-4, Tra-1-60 or Tra-1-81) as examined by immunostaining (Figure 2d), RT-PCR (Figure 2e), qRT-PCR (Supplementary Figure S1a) or flow cytometry (Supplementary Figure S1b) and the ability to form teratomas with differentiation to three germ layers in immunodeficient mice (Figure 2f). Sendai virus-induced reprogramming does not accompany viral integration into the host genome, and we confirmed that the transduced genes were not expressed on mRNA level in the established FPD-iPSC clones (Figure 2e).

Defective emergence of hematopoietic progenitors from FPD-iPSCs

To investigate the impact of *RUNX1* mutation on the emergence of hematopoietic progenitors (HPCs), we induced hematopoietic differentiation of FPD-iPSCs by co-culture on AGM-S3 cells, a stromal cell line established from aorta–gonad–mesonephros (AGM) region[21,35] (Figure 3a). Briefly, WT-iPSCs and FPD-iPSCs were dispersed and plated on inactivated AGM-S3 cells and were co-cultured in the presence of vascular endothelial growth factor.

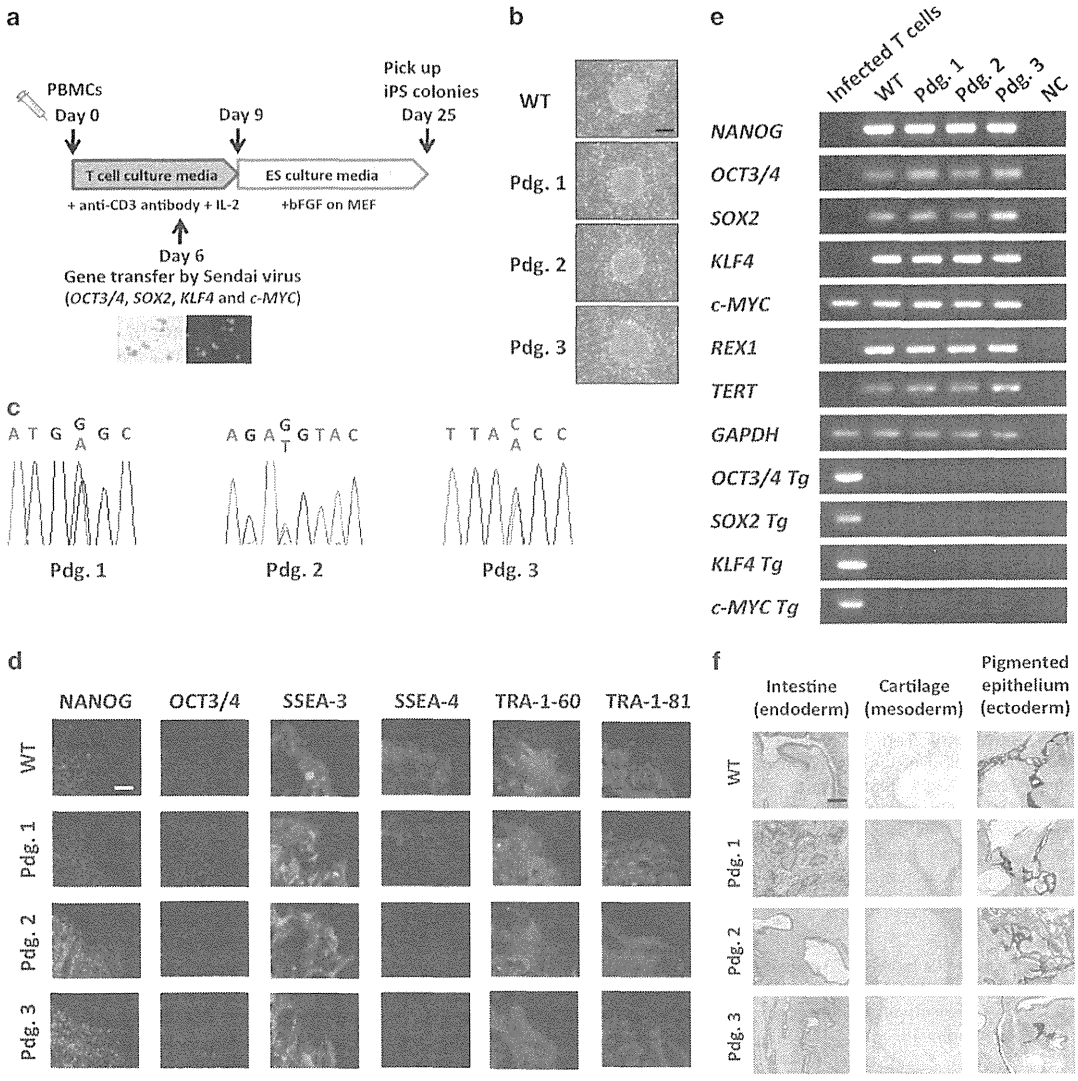

Figure 2. Generation and characterization of iPSCs from FPD/AML patients. (**a**) Schematic diagrams of iPSC derivation from FPD/AML patients using terminally differentiated peripheral T cells. Pictures are T cells infected with control retrovirus expressing green fluorescent protein (GFP) (left; light microscope, right; GFP). (**b**) Morphology of the colonies of WT-iPSCs or FPD-iPSCs. Scale bar = 500 μm. Pdg.; pedigree. (**c**) *RUNX1* mutations in FPD-iPSCs. Each FPD-iPSC retained the same mutation as somatic cells of the original patient. (**d**) Immunofluorescence staining for human embryonic stem cell (hESC) markers. Scale bar = 200 μm. (**e**) RT-PCR analysis for the endogenous hESC maker genes (*NANOG*, *OCT3/4*, *SOX2*, *KLF4*, *c-MYC*, *REX1* and *TERT*), and SeV-transgenes (*OCT3/4 Tg*, *SOX2 Tg*, *KLF4 Tg* and *c-MYC Tg*). NC; negative control. Infected T cells represent T cells 3 days after SeV infection. (**f**) Teratoma formation assay. iPSCs were injected into the testes of NOD-SCID mice. Teratomas were resected, fixed, sectioned and stained with hematoxylin–eosin. WT-iPSC and all FPD-iPSCs showed differentiation to three germ layers, including pigmented epithelium (ectoderm), cartilage (mesoderm) and intestinal glandular structure (endoderm). Scale bar = 50 μm.

On day 15 or day 16 of culture, cells were collected and analyzed for the emergence of HPCs by flow cytometry. Interestingly, the frequencies of $CD34^+$ and $CD45^+$ cells emerged from FPD-iPSCs were decreased to about 40–60% and 20–40% of WT-iPSCs, respectively, suggesting that HPC emergence was profoundly impaired by *RUNX1* mutation (Figures 3b and c). Notably, these defects were observed to a similar extent in all three FPD-iPSC lines, showing that either N-terminal or C-terminal mutations of *RUNX1* do not make any differential effects on the emergence of $CD34^+$ and $CD45^+$ cells. In contrast, expression of CD235a/GPA, an early erthroid-MgK specification marker during human ES cell differentiation,[36] was not impaired in all three FPD-iPSCs as compared with WT (Figure 3c), which suggests that the emergence of early erthroid-MgK progenitors from iPSCs was not affected by *RUNX1* mutation.

In order to quantitatively evaluate the frequency of lineage-committed HPCs derived from each FPD-iPSC, $CD34^+$ cells generated by AGM-S3 co-culture were sorted by flow cytometry and subjected to colony-forming assays. As shown in Figure 3d, the frequencies of granulocyte-monocyte (GM), erythroid (E) or mix colony-forming-cells (CFCs) in $CD34^+$ fraction were significantly lower in FPD-iPSCs as compared with those of WT. It is of note that differences in size and morphology of the colonies were not discernible between WT- and FPD-iPSCs (Supplementary Figure S2).

To examine the HPC emergence in more detail, we undertook ES-sac differentiation protocol to induce hematopoietic differentiation.[32]

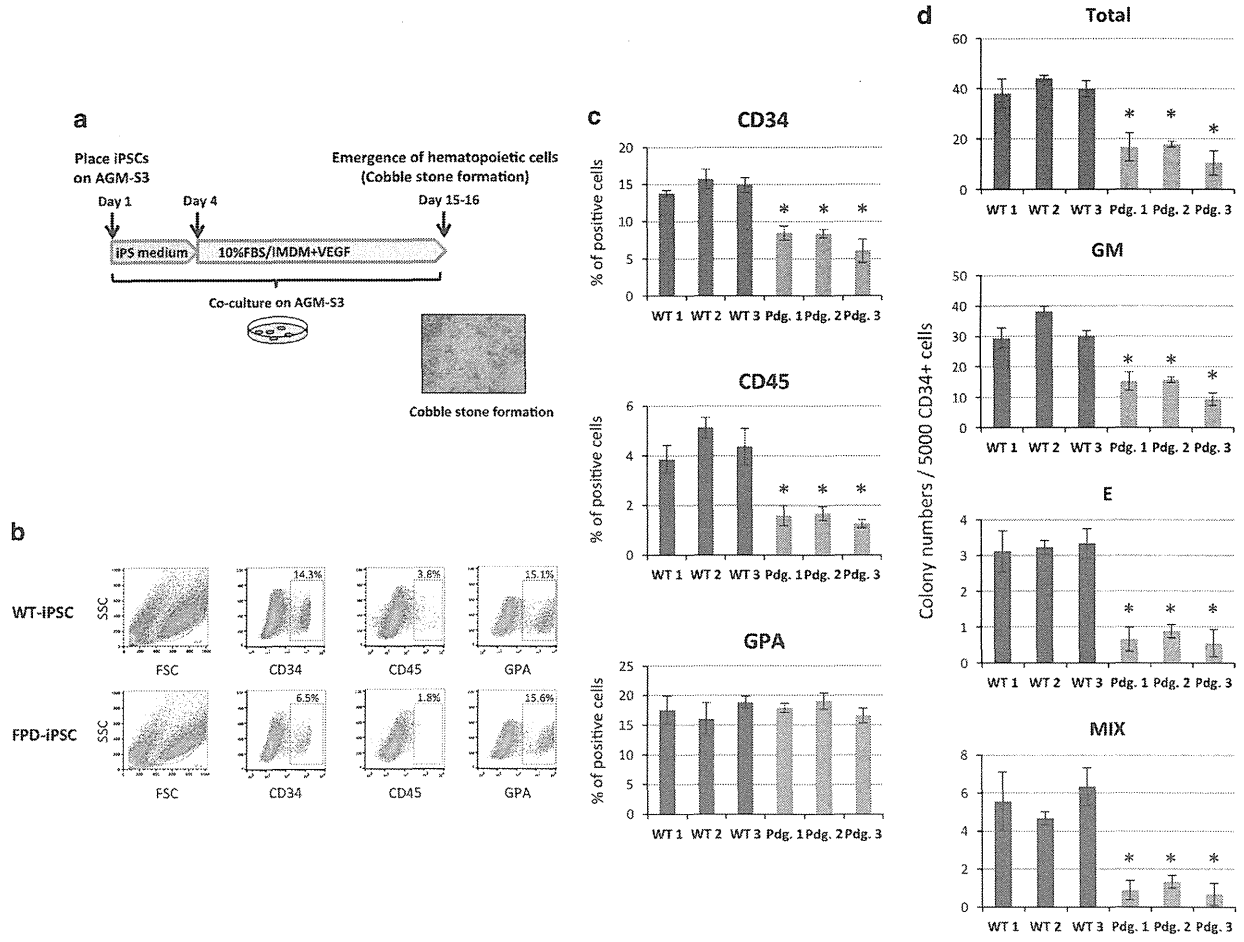

Figure 3. Hematopoietic differentiation of iPSCs by AGM-S3 co-culture. (**a**) Schematic diagram of hematopoietic differentiation of iPSCs by co-culture with AGM-S3. Photograph on the right shows a cobblestone area that appeared around iPSC colonies on day 13. (**b**) Representative flow cytometric profile of cells harvested from AGM-S3 co-cultures. Analyses of WT-iPSC (WT1) and FPD-iPSCs (Pdg. 1) are shown. Mononuclear cell (MNC) fractions were gated and analyzed for CD34, CD45 and GPA. (**c**) Percentages of $CD34^+$, $CD45^+$ or GPA^+ cells against MNCs are shown ($n=3$, mean ± s.d.). All FPD-iPSCs uniformly showed impaired differentiation to CD34 or CD45 cells, while differentiation of GPA^+ cells remained intact. WT1, WT2 and WT3 represent WT-iPSCs, established from three distinct individuals. $^*P<0.01$. (**d**) Sorted $CD34^+$ cells (5000 cells/ plate) from AGM-S3 co-culture were subjected to colony-forming assay as described in Methods. GM, CFU-GM; E, BFU-E; Mix, CFU-mix. Data are shown as mean ± s.d. ($n=3$). $^*P<0.01$.

The efficiency of ES-sac induction was comparable between WT- and FPD-iPSCs (Supplementary Figure S3a). As shown in Figures 4a–c and Supplementary Figure 3b, the percentage of $CD34^+CD43^+CD45^-$ cells, earliest HPCs detected during ES cell/iPSC differentiation (HP1), was decreased to 10–40% of WT in FPD-iPSCs.[36,37] Furthermore, $CD34^+CD43^+CD45^+$ or $CD34^-CD43^+CD45^+$ cells, representing late-committed HPCs (HP2) or myeloid-restricted HPCs (Lin P), respectively, were drastically decreased to 3–5% of WT in FPD-iPSCs (Figure 4c). Notably, these frequencies were not statistically different between FPD-iPSCs with N-terminal *RUNX1* mutation (FPD-N-iPSCs) (pedigree 1 and pedigree 2) and those with C-terminal *RUNX1* mutation (FPD-C-iPSCs) (pedigree 3) in this assay. We have also checked the frequency of various lineage-committed progenitors in the fixed number of $CD34^+CD43^+CD45^-$ cells by colony-forming assays. This revealed that the frequencies of GM-, E- and mix-CFCs in $CD34^+CD43^+CD45^-$ cells were comparable between WT- and FPD-iPSCs (Supplementary Figure S3c), and no apparent difference was noted in the morphology of the colonies (Supplementary Figure S3d).

Taken together, these results clearly indicate that the net emergence of HPCs from human iPSCs is profoundly impaired by *RUNX1* mutation.

Defective differentiation and maturation of MgKs from FPD-iPSCs

It has been shown that *RUNX1* was critical for MgK differentiation and maturation by gene-disruption studies in mice.[1] We asked whether this finding could be applied to human settings by using FPD-iPSC-differentiation model. To do this, $CD34^+$ cells induced in AGM-S3 co-culture system were assessed for their ability to differentiate into MgKs in liquid culture with thrombopoietin and SCF (Figure 5a). Interestingly, $CD34^+$ cells from FPD-iPSCs generated $CD41a^+$ MgKs in significantly lower frequencies (30–50%) as compared with WT in this assay (Figures 5b and d). Actual number of MgKs generated from $CD34^+$ cells was also decreased in the FPD-iPSC-group as compared with WT (Supplementary Figure S4). Of note, MgKs differentiated from FPD-iPSCs were less mature and smaller in size as evidenced by CD42b and mean-forward scatter (FSC), respectively, by flow cytometry (Figure 5d). However, differences in size or morphology were not readily apparent by cytospin preparation (Figure 5c).

These results indicate that differentiation of MgKs is impaired both quantitatively and qualitatively in FPD-iPSCs. Again, all three FPD-iPSC lines shared the same phenotype in these assays, suggesting that N-terminal and C-terminal *RUNX1* mutations

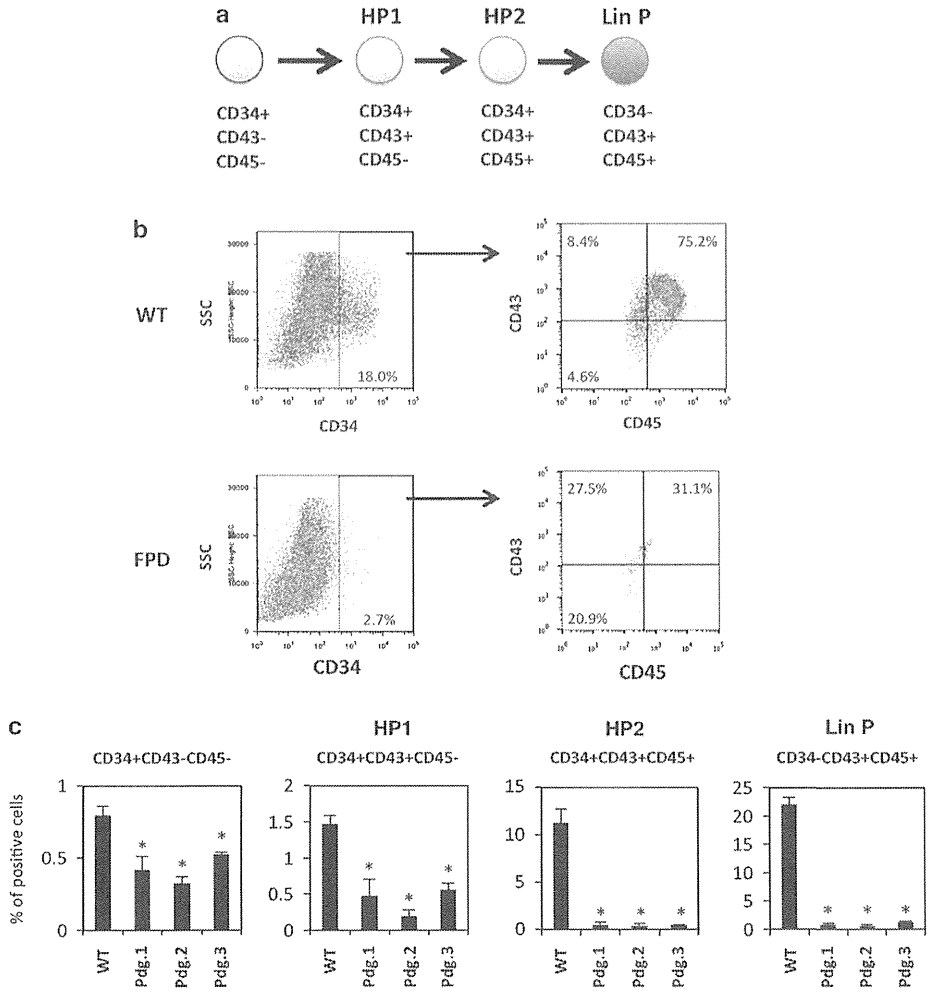

Figure 4. Emergence of hematopoietic progenitors from iPSCs. (**a**) A model of hematopoietic differentiation from human ES cells.[36] HP1, early HPC; HP2, late-committed HPC; Lin P, myeloid-restricted HPC. (**b**) Representative flow cytometric profiles of HPCs generated from WT- or FPD-iPSCs by ES-sac protocol. (**c**) Percentages of HPCs generated from WT- or FPD-iPSCs analyzed by flow cytometry. Data are mean ± s.d. ($n = 3$). *$P < 0.01$. SSC, side scatter.

impose similar defects in MgK differentiation and maturation in FPD-iPSCs.

Phenotypic rescue of FPD-iPSCs by overexpression of WT *RUNX1*

As mutant *RUNX1* has been reported to act in a loss-of-function or dominant-negative manner for WT-*RUNX1*, we tried to rescue the phenotypes of FPD-iPSCs by overexpressing WT-*RUNX1*. FPD-N-iPSC and FPD-C-iPSC clones (pedigrees 1 and 3) overexpressing WT-*RUNX1* were established by transfecting the vector expressing Flag-tagged WT-*RUNX1*. Three clones for each FPD-iPSC were established, and they presented highly similar phenotype. The expression of Flag-RUNX1 was confirmed by western blotting and RT-PCR (Figures 6a and b). FPD-iPSCs overexpressing Flag-*RUNX1* was morphologically indistinguishable from parental FPD-iPSCs, and their immature phenotype was confirmed by the expression of pluripotent genes such as *NANOG* or *OCT3/4* by RT-PCR (Figure 6b). We then investigated whether these established clones recovered the capacity to differentiate into hematopoietic lineage. As expected, overexpressing WT-*RUNX1* in FPD-iPSCs rescued the emergence of $CD34^+$ and $CD45^+$ cells by AGM-S3 co-culture both in FPD-N-iPSC (pdg. 1) and FPD-C-iPSC (pdg. 3), whereas mock control did not (Figure 6c). Moreover, CFC numbers (CFU-GM, CFU-E, CFU-mix) in $CD34^+$ cells and differentiation of MgKs as examined by CD41a expression and cell numbers were also rescued by WT-*RUNX1* overexpression (Figures 6d and e, Supplementary Figure S5). Morphology of the colonies was not different between parental FPD-iPSC and mock- or Flag-*RUNX1*-transfected FPD-iPSCs (Supplementary Figure S6). Interestingly, however, mean-FSC by flow cytometry was not recovered, and CD42 expression was rescued only in pedigree 3 (Figure 6e). This suggests that overexpression of WT-*RUNX1* only partially rescues Mgk maturation. Of note, RT-PCR analysis showed that the transgene for WT-*RUNX1* was not silenced either before or after the induction of HPCs (Supplementary Figure S7).

Taken together, these results support the notion that mutant *RUNX1* in FPD/AML acts in a loss-of-function or dominant-negative manner to the WT allele in hematopoietic differentiation, although some aspects of impaired Mgk maturation in FPD-iPSCs may not be the consequence of impaired *RUNX1* function.

Expression of RUNX1 target genes in FPD-iPSC-derived HPCs

To obtain clues whether mutant *RUNX1* acts in a loss-of-function or dominant-negative manner in hematopoietic differentiation,

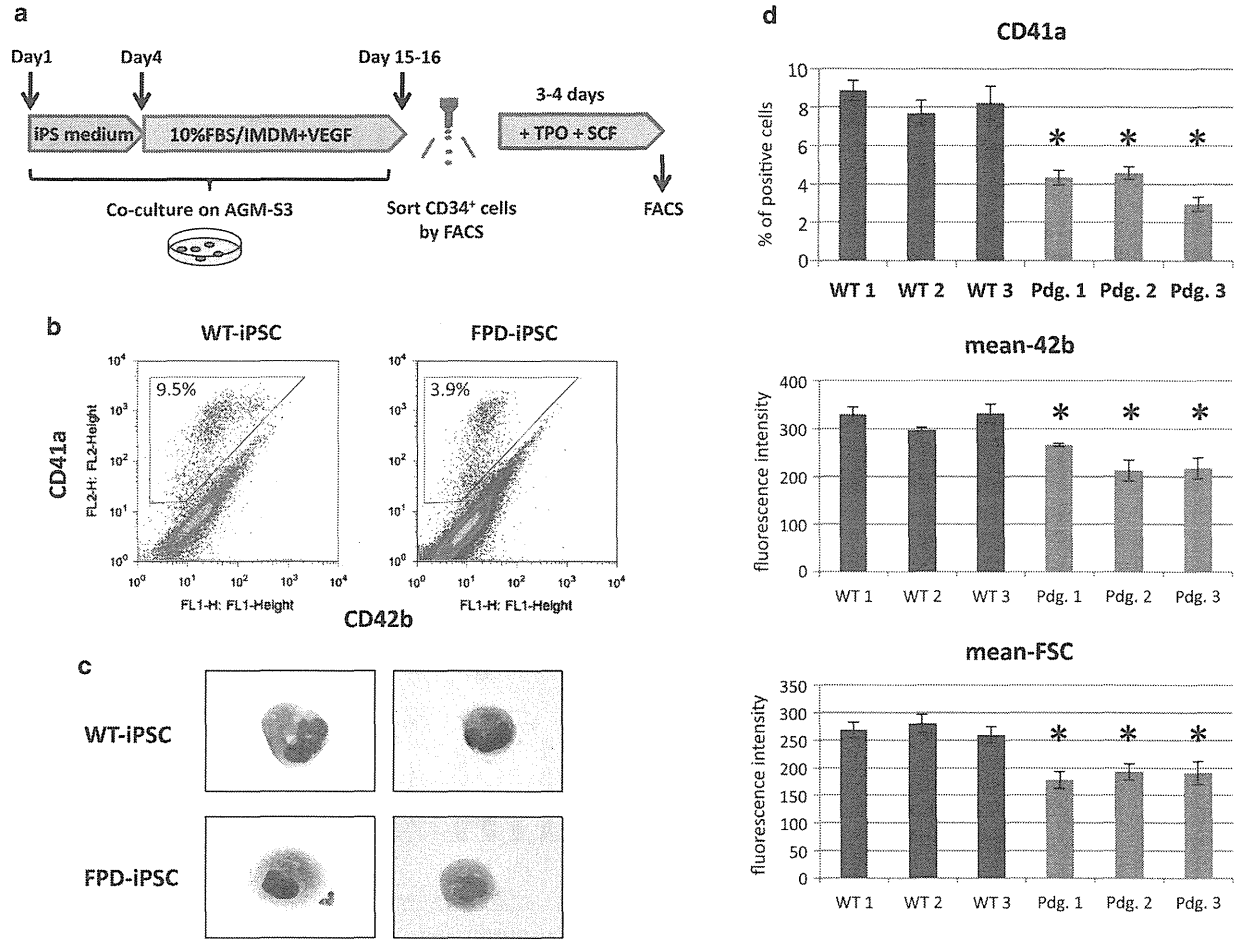

Figure 5. Differentiation of MgKs from iPSC-derived CD34$^+$ cells. (**a**) Schematic diagram of MgK differentiation from iPSC-derived CD34$^+$ cells. CD34$^+$ cells from AGM-S3 co-culture were sorted and cultured in MgK differentiation medium containing thrombopoietin and stem cell factor. Cells were harvested and analyzed on day 3 or 4. (**b**) Representative FACS profile of MgKs generated from iPSC-derived CD34$^+$ cells. Analyses of WT-iPSC (WT1) and FPD-iPSCs (Pdg. 1) are shown. (**c**) Morphology of MgKs derived from WT-iPSC (WT1) and FPD-iPSCs (Pdg. 1) (Giemsa staining. Original magnification; × 1000). (**d**) Percentage of CD41a$^+$ cells, mean fluorescence intensity of CD42b and mean-FSC of WT-iPSC- or FPD-iPSC-derived Mgks are shown. Data are mean ± s.d. ($n = 3$). $^*P < 0.01$.

we examined the expression of *RUNX1* target genes by qRT-PCR in hematopoietic cells derived from FPD-iPSCs. Interestingly, expressions of well-known *RUNX1* target genes such as *PU.1*, GM-colony stimulating factor (*GM-CSF*) or myeloperoxidase (*MPO*)[38] in FPD-iPSC-derived hematopoietic cells were decreased to approximately half of those of WT (Figure 7). These data strongly suggest that *RUNX1* alleles of FPD-iPSCs are haploinsufficient.

Overexpression of mutant *RUNX1* in WT-iPSCs does not recapitulate the phenotype of FPD-iPSCs

In order to gain further insight into the role of mutant *RUNX1* in hematopoietic differentiation of iPSCs, we derived WT-iPSCs overexpressing mutant *RUNX1* and examined their differentiation to hematopoietic lineage (Figure 8a). In this experiment, we tested both N-terminal G172E *RUNX1* mutant (*RUNX1*-N^m) from pedigree 1 and C-terminal N233fsX283 *RUNX1* mutant (*RUNX1*-C^m) from pedigree 3.

Three stable iPSC clones for each *RUNX1* mutant were established, and the expression of mutant RUNX1 protein in each clone was confirmed by western blotting (Figure 8a). These clones were then subjected to hematopoietic differentiation assays by AGM-S3 co-culture system. Surprisingly, both *RUNX1* mutants, either *RUNX1*-N^m or *RUNX1*-C^m, scarcely affected the differentiation of WT-iPSCs to HPCs (Figures 8b and c) or MgKs (Figure 8d) as examined by flow cytometry or colony assays, whereas FPD-iPSCs were defective in the same settings (Figures 8b–d). Of note, the expression of mutant *RUNX1* was not silenced in CD34$^+$ cells derived from iPSC clones analyzed (Supplementary Figure S8). These results clearly demonstrated that overexpression of mutant *RUNX1* in WT-iPSCs did not recapitulate the phenotype of FPD-iPSCs. Taken together with the effects of mutant *RUNX1* on the expression of target genes (Figure 7), these results strongly suggest that *RUNX1* mutants act in a loss-of-function, not dominant-negative, manner in hematopoietic differentiation of iPSCs.

DISCUSSION

Knowledge on *RUNX1* function has been mostly derived from genetically modified animals, such as knockout mice or mutant zebrafish. It was previously shown that *Runx1* has a critical role in the establishment of definitive hematopoietic stem cells during embryonic development,[39,40] and the differentiation of MgKs and lymphocytes in adult hematopoiesis.[1] Further studies have shown that *Runx1* was required for the emergence of definitive hematopoietic stem cells from the so-called 'hemogenic

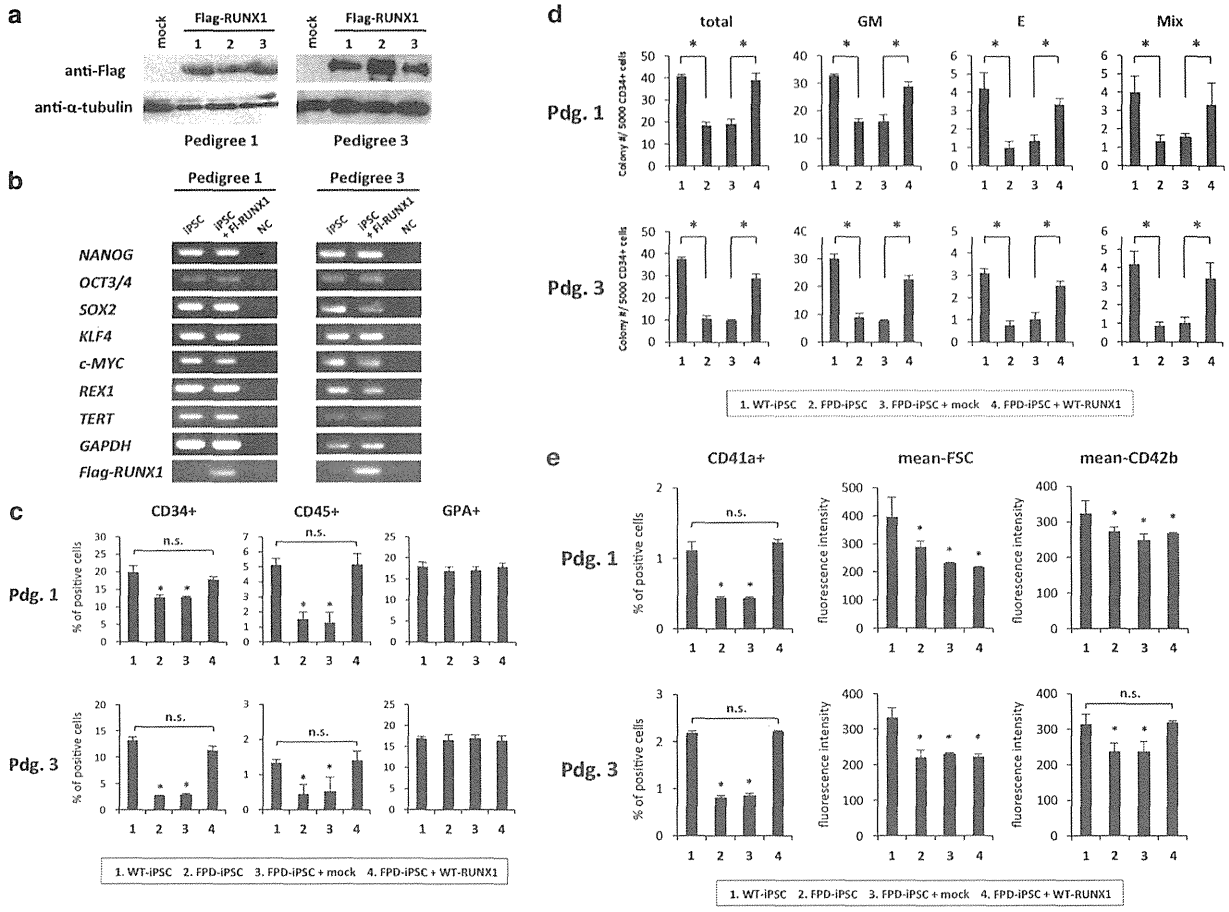

Figure 6. Phenotypic rescue of FPD-iPSCs by WT *RUNX1*. (**a**) Expression of Flag-RUNX1 protein in FPD-iPSCs transfected with pcDNA3/Flag-*RUNX1*. Three different clones were isolated for each FPD-iPSC (pedigrees 1 and 3). The clone numbers are shown on each gel. (**b**) RT-PCR analysis for the endogenous hESC maker genes and Flag-*RUNX1* before and after transfection of Flag-*RUNX1*. Fl-RUNX1, Flag-*RUNX1*; NC, negative control. (**c**) Hematopoietic differentiation of rescued FPD-iPSCs by AGM-S3 co-culture. Percentages of $CD34^+$ cells, $CD45^+$ cells and GPA^+ cells derived from each iPSC are shown. Data are mean ± s.d. ($n = 3$). $^*P < 0.05$. NS, not significant. (**d**) Colony-forming assay of sorted $CD34^+$ cells. Data are mean ± s.d. ($n = 3$). $^*P < 0.05$. GM, CFU-GM; E, BFU-E; Mix, CFU-mix. (**e**) Differentiation of MgKs from rescued FPD-iPSC-derived $CD34^+$ cells. Percentage of $CD41a^+$ cells, mean fluorescence intensity of CD42b and mean-FSC are shown. Data are mean ± s.d. ($n = 3$). $^*P < 0.05$. NS, not significant.

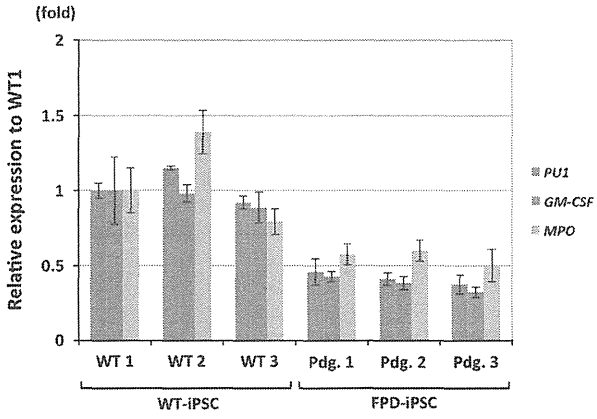

Figure 7. Expression of *RUNX1* target genes in HPCs derived from FPD-iPSCs. Cells recovered from hematopoietic colonies generated from WT- or FPD-iPSCs were subjected to RNA extraction and quantitative RT-PCR analyses for *PU.1*, *GM-CSF* or *MPO*. Data are presented as the relative expression to those of WT1 (mean ± s.d., $n = 3$).

endothelium'.[41–44] However, physiological function of *RUNX1* has been rarely investigated in human experimental settings. In order to investigate the impact of *RUNX1* mutation on human hematopoiesis and to delineate the pathophysiology of FPD/AML, we derived iPSCs from FPD/AML patients and examined their defects in hematopoietic differentiation. In this study, we, for the first time, demonstrated that human iPSCs with *RUNX1* mutation are defective in the emergence of HPCs and MgK differentiation.

We also demonstrated that mutant *RUNX1* acts in a loss-of-function manner in hematopoietic differentiation of human iPSCs, strongly suggesting that the phenotypes of FPD-iPSCs are the consequence of haploinsufficiency of *RUNX1*. This is compatible with the previous observation that significant number of FPD/AML pedigrees harbor germline heterozygous deletion of entire or a part of *RUNX1* allele. Furthermore, it was reported that genetically modified mice with heterozygous mutant *RUNX1*-knock-in (KI) alleles, which resembled human hematologic diseases, displayed 60–70% decrease of hematopoietic CFC numbers in AGM regions or FLs during murine embryogenesis, suggesting that they act as haploinsufficient alleles *in vivo*. In contrast, previous *in vitro* biochemical studies have suggested that some of the *RUNX1* mutations observed in MDS or AML act as weak dominant-negative allele.[3,9] They showed that most of the *RUNX1* mutants

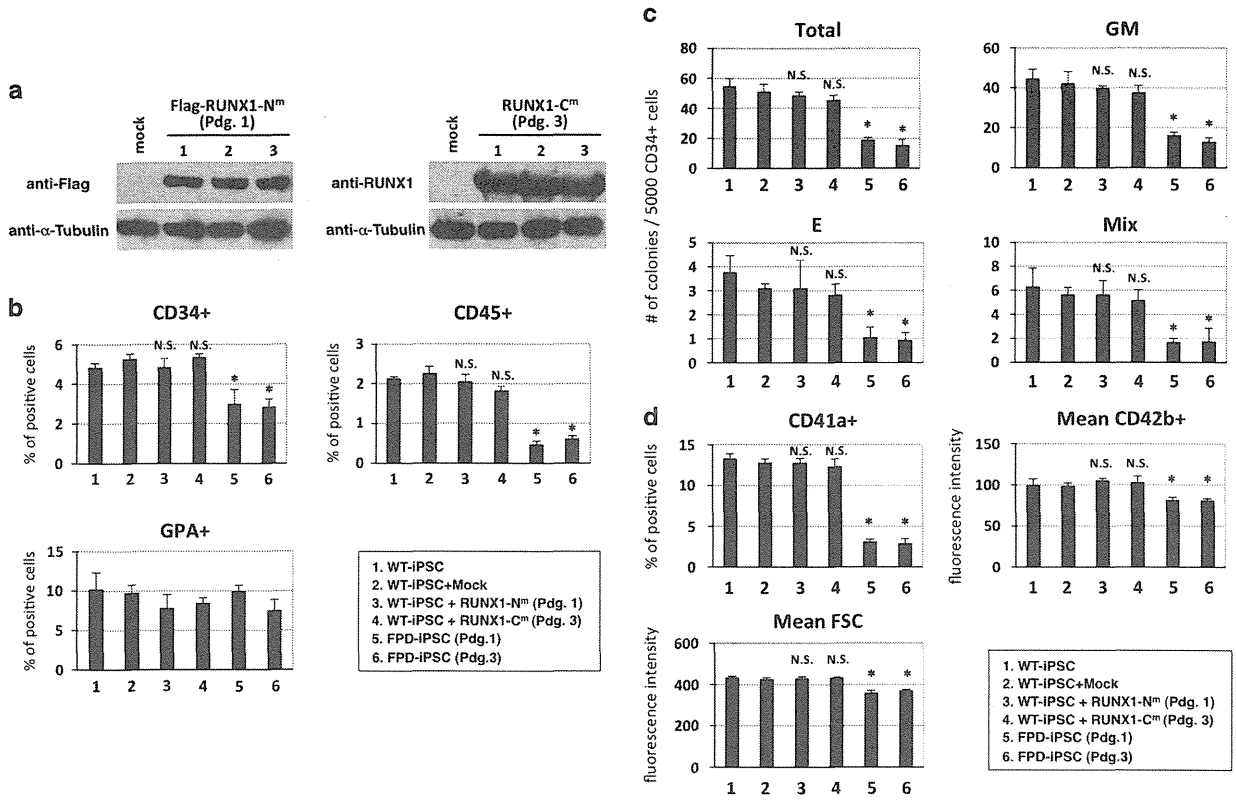

Figure 8. Effects of mutant *RUNX1* on hematopoietic differentiation of WT-iPSCs. (**a**) Expression of mutant RUNX1 protein in WT-iPSCs transfected with pcDNA3/Flag-*RUNX1*-N^m or pcDNA3/*RUNX1*-C^m. Three different clones were isolated for each *RUNX1*-mutant, and the expression of transfected genes was examined by western blotting. Flag-RUNX1-N^m and RUNX1-C^m proteins were detected by anti-Flag and anti-RUNX1 antibodies, respectively. α-Tubulin was used as a loading control. The clone numbers are shown on each lane. WT-iPSCs transfected with mock vector (mock) were used as a negative control. (**b**) Hematopoietic differentiation of WT-iPSCs expressing *RUNX1*-N^m and *RUNX1*-C^m by AGM-S3 co-culture. Cells were co-cultured on AGM-S3 cells for 10–14 days and subjected to FACS analyses. Frequencies (%) of CD34$^+$, CD45$^+$ or GPA$^+$ cells derived from WT-iPSCs expressing *RUNX1*-N^m and *RUNX1*-C^m or FPD-iPSCs (Pdg. 1 and 3) against MNCs are shown ($n=3$, mean ± s.d.). Parental WT-iPSCs and WT-iPSCs transfected with mock vector were used as controls. *$P<0.05$. NS, not significant against 'WT-iPSCs' or 'WT-iPSCs + mock'. (**c**) Colony-forming assay of HPCs derived from WT-iPSCs expressing *RUNX1*-N^m and *RUNX1*-C^m. Sorted CD34$^+$ cells (5000 cells/plate) from AGM-S3 co-culture were subjected to the assay as described in Methods. GM, CFU-GM; E, BFU-E; Mix, CFU-mix. Data are mean ± s.d. ($n=3$). *$P<0.05$. NS, not significant against 'WT-iPSCs' or 'WT-iPSCs + mock'. Description of the columns is the same as that in panel (**b**) or (**d**). (**d**) Differentiation of MgKs from CD34$^+$ cells derived from WT-iPSCs expressing *RUNX1*-N^m and *RUNX1*-C^m. Percentage of CD41a$^+$ cells, mean fluorescence intensity of CD42b and mean-FSC are shown. Data are mean ± s.d. ($n=3$). *$P<0.05$. NS, not significant against 'WT-iPSCs' or 'WT-iPSCs + mock'.

are defective in DNA binding, dimerization with CBFβ or transactivation capacities,[3,9] and thereby suppress WT-*RUNX1* by varying degrees when they are overexpressed *in vitro*.[3] Taken together, it is plausible that the effects of mutant *RUNX1* on hematopoiesis significantly vary depending on the expression levels and the cellular context. This notion is supported by the previous observation that retroviral overexpression of mutant *Runx1* in BM cells leads to the development of MDS/AML in mice, while heterozygous mutant *Runx1*-KI mice do not develop leukemia during their lifetime. Further investigation is required to reveal precise molecular mechanism for differential effect of *RUNX1* under various conditions.

Ran *et al.*[45] have recently reported that enforced expression of *RUNX1a*, a naturally occurring isoform lacking C-terminal activation/repression domain, enhanced the production of CD34$^+$CD45$^+$ HPCs from human ESCs/iPSCs, which are transplantable to immune-deficient mice. However, in our hands, overexpression of *RUNX1*^{N233fsX283} mutant (*RUNX1*-C^m), which closely resembles *RUNX1a*, in WT-iPSCs did not affect the differentiation to CD34$^+$ cells. This discrepancy could be due to a slight difference between the sequences of *RUNX1a* and *RUNX1*-C^m, the different strategies taken to deliver *RUNX1* mutants into iPSCs (lentiviral transduction vs plasmid overexpression), integration sites of the transduced gene or different iPSC clones utilized. This is certainly another issue of future investigation.

Another critical finding of this study is a cell-autonomous effect of *RUNX1* mutation on hematopoiesis. Although previous studies using mutant mice or patient samples suggested cell-autonomous effects of mutant *RUNX1* on HPC emergence and MgK differentiation, it has still been possible that extrinsic factors such as altered microenvironmental cues affected the differentiation process *in vivo*. This study demonstrated that FPD-iPSCs were defective in hematopoietic differentiation in *in vitro* assays, showing that disease-specific *RUNX1* mutation impaired the emergence of HPCs and MgK differentiation indeed in a cell-autonomous manner.

It is noteworthy that functional roles of mutant *RUNX1* on MgK differentiation and platelet production are critically different between humans and mice. In mice, heterozygous DNA-binding *Runx1* mutation never led to thrombocytopenia, and only homozygous animals presented mild MgK/platelet defects.[9] In contrast, heterozygous *RUNX1* mutation is sufficient to cause MgK defects in human settings as demonstrated by the current study. These differences could be due to a differential sensitivity of MgK

differentiation and platelet production to *RUNX1* dosage in human or mouse hematopoiesis. Molecular mechanism underlying this discrepancy definitely requires further investigation.

In MgK differentiation assays, mean-FSC by flow cytometry was significantly lower in MgKs derived from FPD-iPSCs as compared with those from WT-iPSCs, suggesting that FPD-iPSC-derived MgKs are smaller in size. Complementation of *RUNX1* activity in FPD-iPSCs by WT-*RUNX1* rescued MgK differentiation as examined by CD41a expression. Interestingly, however, mean-FSC of FPD-iPSC-derived MgKs did not return to the level comparable to that of WT in the same experiment. These results indicated that differentiation and cell size of MgKs were differentially regulated by *RUNX1* and raise a possibility that reduced size of FPD-iPSC-derived MgKs might be the consequence of novel function acquired by mutant *RUNX1*.

We were able to analyze three FPD-iPSC lines derived from three distinct FPD/AML pedigrees. Two pedigrees carried mutation in RUNT domain that disrupts DNA binding, and the other carried frame-shift mutation resulting in premature termination before C-terminal activation/repression domain. Through our *in vitro* hematopoietic differentiation analyses, we could find no major difference between the three FPD-iPSC lines in terms of HPC emergence and MgK differentiation. These results suggest that disease-specific *RUNX1* mutations impose highly similar impact on the hematopoietic differentiation of iPSCs regardless of the sites of mutation. However, it still leaves a possibility that various *RUNX1* mutations differentially affect other aspects of hematopoiesis. Particularly, as mutant *RUNX1* is involved in the malignant transformation of hematopoietic cells, it would be intriguing to examine differential impacts of various *RUNX1* mutations on the development of AML or MDS using our FPD-iPSC models.

In summary, we have successfully established iPSCs from three distinct FPD/AML pedigrees and have shown that these FPD-iPSCs are uniformly defective in HPC emergence and MgK differentiation. This report is the first to demonstrate critical roles of *RUNX1* in hematopoiesis in human experimental settings and revealed differential impact of heterozygous mutant *RUNX1* on human and mouse MgK differentiation. We also demonstrated that the phenotypes of FPD-iPSCs are the consequence of haploinsufficiency of *RUNX1*. We expect that these FPD-iPSC lines are extremely useful as an unlimited source for human HPCs with various *RUNX1* mutations, and they will serve as a novel platform for investigating multistep leukemogenesis based on *RUNX1* mutation.

CONFLICT OF INTEREST

HO is the scientific consultant of San Bio, Inc., Eisai Co Ltd. and Daiichi Sankyo Co Ltd. HN is a founder, a member of scientific advisory board and shareholder of ReproCELL and is a scientific advisor of Megakaryon, iCELL and Shionogi & Co. The remaining authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank excellent technical assistance by J Kawakita. We also thank S Suzuki (FACS Core Laboratory, Keio University School of Medicine) for FACS sorting, W Akamatsu and M Sato (Department of Physiology, Keio University School of Medicine) and H Nakata (Division of Cardiology, Keio University School of Medicine) for cell culture or establishment of iPSCs. This work was supported in part by a grant from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

REFERENCES

1 Ichikawa M, Asai T, Saito T, Seo S, Yamazaki I, Yamagata T *et al.* AML-1 is required for megakaryocytic maturation and lymphocytic differentiation, but not for maintenance of hematopoietic stem cells in adult hematopoiesis. *Nat Med* 2004; **10**: 299–304.
2 Osato M. Point mutations in the RUNX1/AML1 gene: another actor in RUNX leukemia. *Oncogene* 2004; **23**: 4284–4296.
3 Harada H, Harada Y, Niimi H, Kyo T, Kimura A, Inaba T. High incidence of somatic mutations in the AML1/RUNX1 gene in myelodysplastic syndrome and low blast percentage myeloid leukemia with myelodysplasia. *Blood* 2004; **103**: 2316–2324.
4 Christiansen DH, Andersen MK, Pedersen-Bjergaard J. Mutations of AML1 are common in therapy-related myelodysplasia following therapy with alkylating agents and are significantly associated with deletion or loss of chromosome arm 7q and with subsequent leukemic transformation. *Blood* 2004; **104**: 1474–1481.
5 Harada H, Harada Y, Tanaka H, Kimura A, Inaba T. Implications of somatic mutations in the AML1 gene in radiation-associated and therapy-related myelodysplastic syndrome/acute myeloid leukemia. *Blood* 2003; **101**: 673–680.
6 Zharlyganova D, Harada H, Harada Y, Shinkarev S, Zhumadilov Z, Zhunusova A *et al.* High frequency of AML1/RUNX1 point mutations in radiation-associated myelodysplastic syndrome around Semipalatinsk nuclear test site. *J Radiat Res* 2008; **49**: 549–555.
7 Kuo MC, Liang DC, Huang CF, Shih YS, Wu JH, Lin TL *et al.* RUNX1 mutations are frequent in chronic myelomonocytic leukemia and mutations at the C-terminal region might predict acute myeloid leukemia transformation. *Leukemia* 2009; **23**: 1426–1431.
8 Ernst T, Chase A, Zoi K, Waghorn K, Hidalgo-Curtis C, Score J *et al.* Transcription factor mutations in myelodysplastic/myeloproliferative neoplasms. *Haematologica* 2010; **95**: 1473–1480.
9 Matheny CJ, Speck ME, Cushing PR, Zhou Y, Corpora T, Regan M *et al.* Disease mutations in RUNX1 and RUNX2 create nonfunctional, dominant-negative, or hypomorphic alleles. *EMBO J* 2007; **26**: 1163–1175.
10 Harada Y, Harada H. Molecular pathways mediating MDS/AML with focus on AML1/RUNX1 point mutations. *J Cell Physiol* 2009; **220**: 16–20.
11 Michaud J, Wu F, Osato M, Cottles GM, Yanagida M, Asou N *et al.* In vitro analyses of known and novel RUNX1/AML1 mutations in dominant familial platelet disorder with predisposition to acute myelogenous leukemia: implications for mechanisms of pathogenesis. *Blood* 2002; **99**: 1364–1372.
12 Imai Y, Kurokawa M, Izutsu K, Hangaishi A, Takeuchi K, Maki K *et al.* Mutations of the AML1 gene in myelodysplastic syndrome and their functional implications in leukemogenesis. *Blood* 2000; **96**: 3154–3160.
13 Song WJ, Sullivan MG, Legare RD, Hutchings S, Tan X, Kufrin D *et al.* Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. *Nat Genet* 1999; **23**: 166–175.
14 Higuchi M, O'Brien D, Kumaravelu P, Lenny N, Yeoh EJ, Downing JR. Expression of a conditional AML1-ETO oncogene bypasses embryonic lethality and establishes a murine model of human t(8;21) acute myeloid leukemia. *Cancer Cell* 2002; **1**: 63–74.
15 Motoda L, Osato M, Yamashita N, Jacob B, Chen LQ, Yanagida M *et al.* Runx1 protects hematopoietic stem/progenitor cells from oncogenic insult. *Stem Cells* 2007; **25**: 2976–2986.
16 Jacob B, Osato M, Yamashita N, Wang CQ, Taniuchi I, Littman DR *et al.* Stem cell exhaustion due to Runx1 deficiency is prevented by Evi5 activation in leukemogenesis. *Blood* 2010; **115**: 1610–1620.
17 Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 2007; **131**: 861–872.
18 Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* 2007; **318**: 1917–1920.
19 Chou ST, Byrska-Bishop M, Tober JM, Yao Y, Vandorn D, Opalinska JB *et al.* Trisomy 21-associated defects in human primitive hematopoiesis revealed through induced pluripotent stem cells. *Proc Natl Acad Sci USA* 2012; **109**: 17573–17578.
20 Ye Z, Zhan H, Mali P, Dowey S, Williams DM, Jang YY *et al.* Human-induced pluripotent stem cells from blood cells of healthy donors and patients with acquired blood disorders. *Blood* 2009; **114**: 5473–5480.
21 Park IH, Arora N, Huo H, Maherali N, Ahfeldt T, Shimamura A *et al.* Disease-specific induced pluripotent stem cells. *Cell* 2008; **134**: 877–886.
22 Hiramoto T, Ebihara Y, Mizoguchi Y, Nakamura K, Yamaguchi K, Ueno K *et al.* Wnt3a stimulates maturation of impaired neutrophils developed from severe congenital neutropenia patient-derived pluripotent stem cells. *Proc Natl Acad Sci USA* 2013; **110**: 3023–3028.
23 Maclean GA, Menne TF, Guo G, Sanchez DJ, Park IH, Daley GQ *et al.* Altered hematopoiesis in trisomy 21 as revealed through in vitro differentiation of isogenic human pluripotent cells. *Proc Natl Acad Sci USA* 2012; **109**: 17567–17572.
24 Chang CJ, Bouhassira EE. Zinc-finger nuclease-mediated correction of alpha-thalassemia in iPS cells. *Blood* 2012; **120**: 3906–3914.
25 Carette JE, Pruszak J, Varadarajan M, Blomen VA, Gokhale S, Camargo FD *et al.* Generation of iPSCs from cultured human malignant cells. *Blood* 2010; **115**: 4039–4042.

26 Kumano K, Arai S, Hosoi M, Taoka K, Takayama N, Otsu M *et al.* Generation of induced pluripotent stem cells from primary chronic myelogenous leukemia patient samples. *Blood* 2012; **119**: 6234–6242.
27 Gandre-Babbe S, Paluru P, Aribeana C, Chou ST, Bresolin S, Lu L *et al.* Patient-derived induced pluripotent stem cells recapitulate hematopoietic abnormalities of juvenile myelomonocytic leukemia. *Blood* 2013; **121**: 4925–4929.
28 Seki T, Yuasa S, Fukuda K. Generation of induced pluripotent stem cells from a small amount of human peripheral blood using a combination of activated T cells and Sendai virus. *Nat Protoc* 2012; **7**: 718–728.
29 Fukuchi Y, Shibata F, Ito M, Goto-Koshino Y, Sotomaru Y, Ito M *et al.* Comprehensive analysis of myeloid lineage conversion using mice expressing an inducible form of C/EBP alpha. *EMBO J* 2006; **25**: 3398–3410.
30 Ma F, Wang D, Hanada S, Ebihara Y, Kawasaki H, Zaike Y *et al.* Novel method for efficient production of multipotential hematopoietic progenitors from human embryonic stem cells. *Int J Hematol* 2007; **85**: 371–379.
31 Ma F, Ebihara Y, Umeda K, Sakai H, Hanada S, Zhang H *et al.* Generation of functional erythrocytes from human embryonic stem cell-derived definitive hematopoiesis. *Proc Natl Acad Sci USA* 2008; **105**: 13087–13092.
32 Takayama N, Nishikii H, Usui J, Tsukui H, Sawaguchi A, Hiroyama T *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**: 5298–5306.
33 Nakajima H, Ito M, Smookler DS, Shibata F, Fukuchi Y, Morikawa Y *et al.* TIMP-3 recruits quiescent hematopoietic stem cells into active cell cycle and expands multipotent progenitor pool. *Blood* 2010; **116**: 4474–4482.
34 Seki T, Yuasa S, Oda M, Egashira T, Yae K, Kusumoto D *et al.* Generation of induced pluripotent stem cells from human terminally differentiated circulating T cells. *Cell Stem Cell* 2010; **7**: 11–14.
35 Xu MJ, Tsuji K, Ueda T, Mukouyama YS, Hara T, Yang FC *et al.* Stimulation of mouse and human primitive hematopoiesis by murine embryonic aorta-gonad-mesonephros-derived stromal cell lines. *Blood* 1998; **92**: 2032–2040.
36 Vodyanik MA, Thomson JA, Slukvin II. Leukosialin (CD43) defines hematopoietic progenitors in human embryonic stem cell differentiation cultures. *Blood* 2006; **108**: 2095–2105.
37 Choi KD, Yu J, Smuga-Otto K, Salvagiotto G, Rehrauer W, Vodyanik M *et al.* Hematopoietic and endothelial differentiation of human induced pluripotent stem cells. *Stem Cells* 2009; **27**: 559–567.
38 Dowdy CR, Xie R, Frederick D, Hussain S, Zaidi SK, Vradii D *et al.* Definitive hematopoiesis requires Runx1 C-terminal-mediated subnuclear targeting and transactivation. *Hum Mol Genet* 2010; **19**: 1048–1057.
39 Okuda T, van Deursen J, Hiebert SW, Grosveld G, Downing JR. AML1 the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. *Cell* 1996; **84**: 321–330.
40 Wang Q, Stacy T, Binder M, Marin-Padilla M, Sharpe AH, Speck NA. Disruption of the Cbfa2 gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis. *Proc Natl Acad Sci USA* 1996; **93**: 3444–3449.
41 North T, Gu TL, Stacy T, Wang Q, Howard L, Binder M *et al.* Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. *Development* 1999; **126**: 2563–2575.
42 North TE, de Bruijn MF, Stacy T, Talebian L, Lind E, Robin C *et al.* Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. *Immunity* 2002; **16**: 661–672.
43 Lancrin C, Sroczynska P, Stephenson C, Allen T, Kouskoff V, Lacaud G. The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. *Nature* 2009; **457**: 892–895.
44 Chen MJ, Yokomizo T, Zeigler BM, Dzierzak E, Speck NA. Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. *Nature* 2009; **457**: 887–891.
45 Ran D, Shia WJ, Lo MC, Fan JB, Knorr DA, Ferrell PI *et al.* RUNX1a enhances hematopoietic lineage commitment from human embryonic stem cells and inducible pluripotent stem cells. *Blood* 2013; **121**: 2882–2890.

Supplementary Information accompanies this paper on the Leukemia website (http://www.nature.com/leu)

Please cite this article in press as: Huang et al., Modularized Functions of the Fanconi Anemia Core Complex, Cell Reports (2014), http://dx.doi.org/10.1016/j.celrep.2014.04.029

Cell Reports
Report

Modularized Functions of the Fanconi Anemia Core Complex

Yaling Huang,[1] Justin W.C. Leung,[1,6] Megan Lowery,[2] Nobuko Matsushita,[4] Yucai Wang,[1] Xi Shen,[1] Do Huong,[1] Minoru Takata,[5] Junjie Chen,[1] and Lei Li[1,3,*]
[1]Department of Experimental Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
[2]Department of Molecular Carcinogenesis, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
[3]Department of Genetics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
[4]Laboratory of Molecular Biochemistry, School of Life Science, Tokyo University of Pharmacy and Life Science, Tokyo 192-0392, Japan
[5]Laboratory of DNA Damage Signaling, Department of Late Effect Studies, Radiation Biology Center, Kyoto University, Kyoto 606-8501, Japan
[6]Present address: The University of Texas at Austin, Section of Molecular Genetics and Microbiology, 2506 Speedway Stop A5000, Austin, TX 78712, USA
*Correspondence: leili@mdanderson.org
http://dx.doi.org/10.1016/j.celrep.2014.04.029

SUMMARY

The Fanconi anemia (FA) core complex provides the essential E3 ligase function for spatially defined FANCD2 ubiquitination and FA pathway activation. Of the seven FA gene products forming the core complex, FANCL possesses a RING domain with demonstrated E3 ligase activity. The other six components do not have clearly defined roles. Through epistasis analyses, we identify three functional modules in the FA core complex: a catalytic module consisting of FANCL, FANCB, and FAAP100 is absolutely required for the E3 ligase function, and the FANCA-FANCG-FAAP20 and the FANCC-FANCE-FANCF modules provide nonredundant and ancillary functions that help the catalytic module bind chromatin or sites of DNA damage. Disruption of the catalytic module causes complete loss of the core complex function, whereas loss of any ancillary module component does not. Our work reveals the roles of several FA gene products with previously undefined functions and a modularized assembly of the FA core complex.

INTRODUCTION

Fanconi anemia (FA) is a complex genetic disorder encompassing 16 tumor suppressor genes that act together to protect cells against genotoxic stress, particularly complexed DNA lesions such as DNA interstrand crosslinks (Bogliolo et al., 2013; D'Andrea, 2010) and potentially DNA-protein crosslinks created by endogenous metabolites (Langevin et al., 2011; Rosado et al., 2011). Classical manifestations of FA include pancytopenia, chromosomal abnormalities, congenital abnormalities, and a high predisposition to a broad spectrum of cancers. Despite the identification of genetic defects in patients with FA, the molecular mechanism underpinning FA pathway functions remains unclear.

A group of classical FA genes is connected by a DNA damage-induced monoubiquitination reaction in the nucleus (Garcia-Higuera et al., 2001; Smogorzewska et al., 2007; Taniguchi et al., 2002). Monoubiquitination of the FANCD2/I complex has the presumed functions of recruiting DNA lesion-processing endonucleolytic activities (Knipscheer et al., 2009; Kratz et al., 2010; Liu et al., 2010; MacKay et al., 2010; Smogorzewska et al., 2010) and transcriptional activation of tumor suppressor genes (Park et al., 2013). The E3 ligase activity of this reaction resides in the FA core complex consisting of seven FA proteins (FANCA, FANCB, FANCC, FANCE, FANCF, FANCG, and FANCL) and two FA-associated proteins (FAAP20 and FAAP100), with the RING domain protein FANCL bearing the E3 ligase activity (Alpi et al., 2008; Meetei et al., 2003). Aside from FANCL and FAAP20, most other components of the core complex have neither recognizable motifs nor clearly defined functions as to how they contribute to the DNA damage-mediated FANCD2/I monoubiquitination.

Studies of protein-protein interactions within the FA core complex have suggested the existence of three subcomplexes (Figure 1A). FANCA, FANCG, and FAAP20, form a subcomplex (A-G-20) (Ali et al., 2012; Garcia-Higuera et al., 1999; Kruyt et al., 1999; Reuter et al., 2000; Waisfisz et al., 1999). The UBZ domain of FAAP20 is suggested to bind to ubiquitinated histone (Leung et al., 2012; Yan et al., 2012). FANCG contains seven TPR repeats and is considered a possible scaffold for the subcomplex (Blom et al., 2004; Léveillé et al., 2004). The FANCB-FANCL-FAAP100 subcomplex (B-L-100) contains the E3 ligase FANCL (Ling et al., 2007; Medhurst et al., 2006). Given that FANCL alone acts sufficiently in reconstituted ubiquitination reactions (Alpi et al., 2008; Longerich et al., 2009; Sato et al., 2012), whether FANCB and FAAP100 contribute to the E3 activity is unclear. A third subcomplex is formed by FANCC, FANCE, and FANCF (C-E-F). FANCF has been shown to interact with FANCM (Deans and West, 2009) and was also suggested to act as an adaptor protein (Léveillé et al., 2004). Despite these observations, the