criteria of ALPS[2] or JMML[6]. It is interesting that Case 2 presented GM-CSF hypersensitivity, which is one of the hallmarks of JMML. Given the strict clinical and laboratory criteria of JMML and ALPS, our two cases should be defined as a new disease entity, like RAS associated ALPS like disease (RALD). Recently defined NRAS mutated ALPS type IV may also be included in a similar disease entity.

There are several cases of JMML reported simultaneously having clinical and laboratory findings compatible with autoimmune disease[8,9]. Autoimmune syndromes are occasionally seen in patients with myelodysplastic syndromes, including chronic myelomonocytic leukemia[10]. These previous findings may suggest a close relationship of autoimmune disease and JMML. Since KRAS G13D has been identified in JMML[11-13], it is tempting to speculate that KRAS G13D mutation is involved in JMML as well as RAS associated ALPS like disease (RALD). It should be noted in JMML, erythroid cells reportedly carry mutant RAS, while B and T cell involvement was variable[13]. In both of our cases, myeloid cells and T cells carried mutant RAS, while B cells were affected variably. These findings would support a hypothesis that the clinical and hematological features are related to the differentiation stages of hematopoietic stem cells where RAS mutation is acquired. JMML-like myelo-monocytic proliferation may predict an involvement of RAS mutation in myeloid stem/precursor cell level whereas ALPS-like phenotype may predict that of stem/precursor cells of lymphoid lineage, especially of T cells. Under the light of subtle differences between the two cases presented, their hematological and clinical features may reflect the characteristics of the stem cell level where *KRAS* mutation is acquired. Involvement of the precursors with higher propensity toward lymphoid lineage may lead to autoimmune phenotypes, while involvement of those with propensity toward the myeloid lineage may lead to GM-CSF hypersensitivity while still

sharing some overlapping autoimmune characteristics.

One may argue from the other points of view with regard to the clinicopathological features of these disorders. First, transformation in a fetal HSC might be obligatory for the development of JMML[14] and that in HSC later in life may not have the same consequences. Second, certain KRAS mutations may be more potent than the others. Codon 13 mutations are generally less deleterious biochemically than codon 12 substitutions, and patients with JMML with codon 13 mutations have been reported to show spontaneous hematologic improvement[12,15]. Thus further studies are needed to reveal in-depth clinicopathological characteristics in this type of lympho-myelo proliferative disorders.

KRAS mutation may initiate the oncogenic pathway as one of the first genetic hits, but is insufficient to cause frank malignancy by itself[16,17]. Considering recent findings that additional mutations of the genes involved in DNA repair, cell cycle arrest, and apoptosis are required for full malignant transformation, one can argue that RAS associated ALPS like disease (RALD) patients will also develop malignancies during the course of the diseases. Occasional association of myeloid blast crisis in JMML and that of lymphoid malignancies in ALPS will support this notion. Thus the two patients are now being followed up carefully. It was recently revealed that half of the patients diagnosed with Evans syndrome, an autoimmune disease presenting with hemolytic anemia and thrombocytopenia, meet the criteria for ALPS diagnosis[18,19]. In this study, FAS-mediated apoptosis analysis was utilized for the screening. Considering the cases we presented, it will be intriguing to re-evaluate Evans syndrome by IL-2 depletion-dependent apoptosis assay focusing on the overlapping autoimmunity with RAS associated ALPS like disease (RALD).

Acknowledgements

This work was supported by a Grant-in-Aid from the Ministry of Education, Science, and Culture, 20390302 (Japan) for SM and by a Grant-in-Aid for Cancer Research from the Ministry of Health, Labor and Welfare, 20-4 and 19-9 (Japan) for SM and MT.

Author ship

MT and SM designed entire experiments and wrote this manuscript. KS, NM, and MT treat those patients, and designed clinical laboratory test. JP performed experiments described in Fig.1b-f. KM, HM, and SD performed colony and mutational analysis. MN, TM, KK, SK, YK and AT supervised clinical and immunological experiments, or coordinated clinical information.

Conflict of interest disclosure

The authors declare no conflict of interest.

References

1. Fisher GH, Rosenberg FJ, Straus SE, et al. Dominant interfering Fas gene mutations impair apoptosis in a human autoimmune lymphoproliferative syndrome. *Cell.* 1995;81(6):935-946.

2. Teachey DT, Seif AE, Grupp SA. Advances in the management and understanding of autoimmune lymphoproliferative syndrome (ALPS). *Br J Haematol.*148(2):205-216.

3. Oliveira JB, Bidere N, Niemela JE, et al. NRAS mutation causes a human autoimmune lymphoproliferative syndrome. *Proc Natl Acad Sci U S A.* 2007;104(21):8953-8958.

4. Oliveira J, Bleesing J, Dianzani U, et al. Revised diagnostic criteria and classification for the autoimmune lymphoproliferative syndrome (ALPS): report from the 2009 NIH International Workshop. *Blood.* 2010;116(14):e35-e40.

5. Miyauchi J, Asada M, Sasaki M, Tsunematsu Y, Kojima S, Mizutani S. Mutations of the N-ras gene in juvenile chronic myelogenous leukemia. *Blood.* 1994;83(8):2248-2254.

6. Emanuel PD. Juvenile myelomonocytic leukemia and chronic myelomonocytic leukemia. *Leukemia.* 2008;22(7):1335-1342.

7. Aoki Y, Niihori T, Narumi Y, Kure S, Matsubara Y. The RAS/MAPK syndromes: novel roles of the RAS pathway in human genetic disorders. *Hum Mutat.* 2008;29(8):992-1006.

8. Kitahara M, Koike K, Kurokawa Y, et al. Lupus nephritis in juvenile myelomonocytic leukemia. *Clin Nephrol.* 1999;51(5):314-318.

9. Oliver JW, Farnsworth B, Tonk VS. Juvenile myelomonocytic leukemia in a child with Crohn disease. *Cancer Genet Cytogenet.* 2006;167(1):70-73.

10. Saif MW, Hopkins JL, Gore SD. Autoimmune phenomena in patients with myelodysplastic syndromes and chronic myelomonocytic leukemia. *Leuk Lymphoma.* 2002;43(11):2083-2092.

11. Flotho C, Kratz CP, Bergstrasser E, et al. Genotype-phenotype

correlation in cases of juvenile myelomonocytic leukemia with clonal RAS mutations. *Blood*. 2008;111(2):966-967.

12. Matsuda K, Shimada A, Yoshida N, et al. Spontaneous improvement of hematologic abnormalities in patients having juvenile myelomonocytic leukemia with specific RAS mutations. *Blood*. 2007;109(12):5477-5480.

13. Flotho C, Valcamonica S, Mach-Pascual S, et al. RAS mutations and clonality analysis in children with juvenile myelomonocytic leukemia (JMML). *Leukemia*. 1999;13(1):32-37.

14. Matsuda K, Sakashita K, Taira C, et al. Quantitative assessment of PTPN11 or RAS mutations at the neonatal period and during the clinical course in patients with juvenile myelomonocytic leukaemia. *Br J Haematol*. 2010;148(4):593-599.

15. Imamura M, Imai C, Takachi T, Nemoto T, Tanaka A, Uchiyama M. Juvenile myelomonocytic leukemia with less aggressive clinical course and KRAS mutation. *Pediatr Blood Cancer*. 2008;51(4):569.

16. Zhang J, Wang J, Liu Y, et al. Oncogenic Kras-induced leukemogeneis: hematopoietic stem cells as the initial target and lineage-specific progenitors as the potential targets for final leukemic transformation. *Blood*. 2009;113(6):1304-1314.

17. Sabnis AJ, Cheung LS, Dail M, et al. Oncogenic Kras initiates leukemia in hematopoietic stem cells. *PLoS Biol*. 2009;7(3):0537-0548.

18. Teachey DT, Manno CS, Axsom KM, et al. Unmasking Evans syndrome: T-cell phenotype and apoptotic response reveal autoimmune lymphoproliferative syndrome (ALPS). *Blood*. 2005;105(6):2443-2448.

19. Seif AE, Manno CS, Sheen C, Grupp SA, Teachey DT. Identifying autoimmune lymphoproliferative syndrome in children with Evans syndrome: a multi-institutional study. *Blood*. 2010;115(11):2142-2145.

Figure Legends

Figure 1

a. Flow cytometric analysis of DNT cells. CD8 and CD4 double staining was performed in TCRαβ-expressing cells.

b. Electropherogram showing KRAS G13D mutation in BM-MNC in case 1 (left panel) and case 2 (right panel).

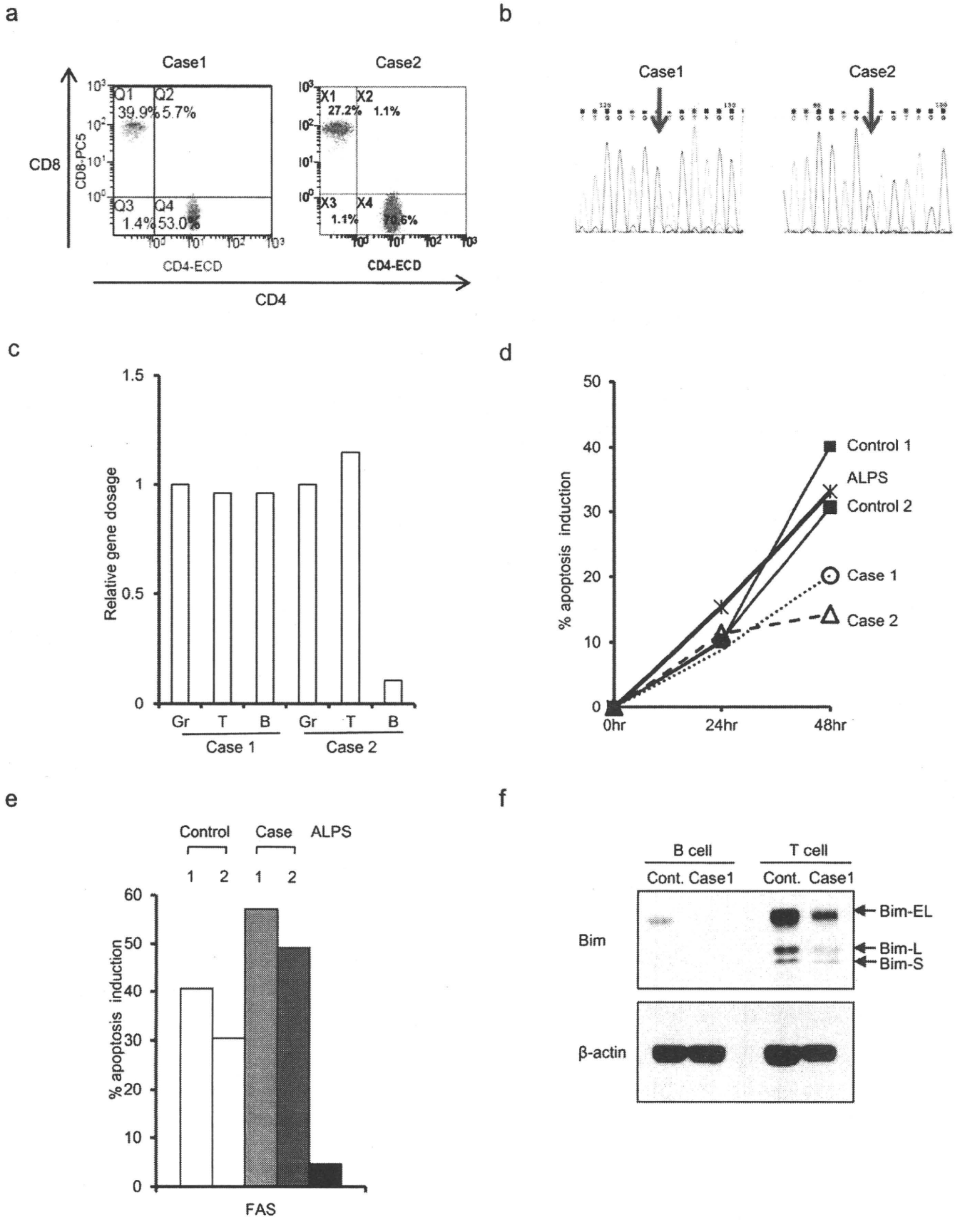
c. Gene dosage of mutated allele in granulocyte (Gr), T cell (T) and B cell (B). Relative gene dosage was estimated by a mutant allele specific PCR method in case 1 and 2 using albumin gene as internal control.

d. Apoptosis assay using activated T cells. Apoptosis percent was measured by flow cytometry with Annexin V staining 24 and 48 hr after IL-2 depletion

e. Apoptosis percent was measured 24hr after addition of anti-FAS CH11 antibody (final 100ng/ml)

f. Western blotting analysis of Bim expression.

Figure 1


a
Case1
Case2
CD8
CD8-PC5
Q1
39.9%
Q2
5.7%
Q3
1.4%
Q4
53.0%
X1
27.2%
X2
1.1%
X3
1.1%
X4
CD4-ECD
CD4-ECD
CD4
b
Case1
Case2
c
Relative gene dosage
1.5
1
0.5
0
Gr
T
B
Gr
T
B
Case 1
Case 2
d
% apoptosis induction
50
40
30
20
10
0
Control 1
ALPS
Control 2
Case 1
Case 2
0hr
24hr
48hr
e
Control
Case
ALPS
1
2
1
2
% apoptosis induction
60
50
40
30
20
10
0
FAS
f
B cell
T cell
Cont.
Case1
Cont.
Case1
Bim
Bim-EL
Bim-L
Bim-S
β-actin

Oncogene (2010), 1–9

www.nature.com/onc

ONCOGENOMICS

Array-based genomic resequencing of human leukemia

Y Yamashita[1], J Yuan[2], I Suetake[3], H Suzuki[4], Y Ishikawa[5], YL Choi[1,6], T Ueno[1], M Soda[1], T Hamada[1], H Haruta[1], S Takada[1], Y Miyazaki[7], H Kiyoi[8], E Ito[9], T Naoe[5], M Tomonaga[7], M Toyota[10], S Tajima[3], A Iwama[2,11] and H Mano[1,6,11]

[1]Division of Functional Genomics, Jichi Medical University, Tochigi, Japan; [2]Department of Cellular and Molecular Medicine, Graduate School of Medicine, Chiba University, Chiba, Japan; [3]Laboratory of Epigenetics, Institute for Protein Research, Osaka University, Osaka, Japan; [4]First Department of Internal Medicine, Sapporo Medical University, Hokkaido, Japan; [5]Department of Hematology and Oncology, Nagoya University Graduate School of Medicine, Nagoya, Japan; [6]Department of Medical Genomics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan; [7]Department of Molecular Medicine and Hematology, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan; [8]Department of Infectious Diseases, Nagoya University School of Medicine, Nagoya, Japan; [9]Department of Paediatrics, Hirosaki University Graduate School of Medicine, Aomori, Japan; [10]Department of Biochemistry, Sapporo Medical University, Hokkaido, Japan and [11]CREST, Japan Science and Technology Agency, Saitama, Japan

To identify oncogenes in leukemias, we performed large-scale resequencing of the leukemia genome using DNA sequence arrays that determine ~9 Mbp of sequence corresponding to the exons or exon–intron boundaries of 5648 protein-coding genes. Hybridization of genomic DNA from CD34-positive blasts of acute myeloid leukemia ($n = 19$) or myeloproliferative disorder ($n = 1$) with the arrays identified 9148 nonsynonymous nucleotide changes. Subsequent analysis showed that most of these changes were also present in the genomic DNA of the paired controls, with 11 somatic changes identified only in the leukemic blasts. One of these latter changes results in a Met-to-Ile substitution at amino-acid position 511 of Janus kinase 3 (JAK3), and the JAK3(M511I) protein exhibited transforming potential both *in vitro* and *in vivo*. Further screening for *JAK3* mutations showed novel and known transforming changes in a total of 9 out of 286 cases of leukemia. Our experiments also showed a somatic change responsible for an Arg-to-His substitution at amino-acid position 882 of DNA methyltransferase 3A, which resulted in a loss of DNA methylation activity of >50%. Our data have thus shown a unique profile of gene mutations in human leukemia.

Oncogene advance online publication, 19 April 2010; doi:10.1038/onc.2010.117

Keywords: resequencing; AML; JAK3; DNMT3A

Introduction

Leukemias are clonal disorders of hematopoietic stem cells or immature progenitors. Several subtypes of leukemia are associated with disease-specific karyotype anomalies in the malignant blasts. Most cases of acute promyelocytic leukemia a subtype of acute myeloid leukemia (AML), for instance, are associated with a t(15;17) chromosomal rearrangement that results in the production of the PML-RARA fusion-type oncoprotein (Tallman and Altman, 2008). Similarly, another subtype of AML is associated with a t(8;21) rearrangement, resulting in the production of the oncogenic RUNX1-CBFA2T1 protein (Nimer and Moore, 2004).

The karyotype of leukemic blasts is an important determinant of the long-term prognosis of affected individuals. AML with t(15;17), t(8;21) or inv(16) rearrangements thus constitutes a subgroup of leukemias with a 'favorable' karyotype, with a 5-year survival rate of >60%, whereas AML with an 'adverse' karyotype (monosomy 7, monosomy 5 or complex anomalies) has a 5-year survival rate of only <15% (Grimwade *et al.*, 1998). The prognosis of AML with a normal karyotype (constituting ~50% of all AML cases) is substantially worse than that with a favorable karyotype, with a 5-year survival rate of 24% (Byrd *et al.*, 2002), indicating that blasts with a normal karyotype may contain transforming genes generated as a result of (1) sequence alterations, (2) epigenetic abnormalities or (3) small chromosomal rearrangements not detectable by the G-banding technique. Indeed, several genes, including *NPM1* and *KIT*, have been found to be mutated and activated in AML blasts with a normal karyotype (Schlenk *et al.*, 2008).

The identification of transforming genes in AML will require large-scale resequencing of the blast genome. Although a new generation of sequencing technologies is now available, whole-genome resequencing of many samples remains a demanding task (Bentley *et al.*, 2008; Wheeler *et al.*, 2008). Although DNA microarray-based sequencing is suitable for analysis of multiple samples, currently available platforms are limited in the number of nucleotides that each array is able to probe. To overcome such limitations, we have now applied the extra-large arrays ('wafers') manufactured by Perlegen Sciences (Mountain View, CA, USA) (originally developed for typing of single-nucleotide polymorphisms) (Patil *et al.*, 2001) to

Correspondence: Dr H Mano, Division of Functional Genomics, Jichi Medical University, 3311-1 Yakushiji, Shimotsukeshi, Tochigi 329-0498, Japan.
E-mail: hmano@jichi.ac.jp
Received 6 May 2009; revised 1 January 2010; accepted 17 March 2010

resequencing of the human genome. Our two-step analysis of human leukemia specimens ($n=20$) has identified a novel transforming mutation in the gene for Janus kinase 3 (JAK3) and a hypomorphic mutation in that for DNA methyltransferase 3A (DNMT3A).

Results

Sequencing strategy

Oligonucleotide probes on the sequencing wafer for the first phase of our study were designed to detect nonsynonymous nucleotide changes in the coding exons of the genome. Intronic sequences (GT in the splicing donor sequence AG-GT and AG in the splicing acceptor sequence AG-G) adjacent to coding exons were also interrogated with the wafer to capture splicing anomalies. Genes examined by the wafer included those known to be mutated in cancer and reported in the catalog of somatic mutations in cancer (COSMIC, http://www.sanger.ac.uk/genetics/cgp/cosmic) as of September 2006 ($n=338$) and those related to the regulation of DNA repair ($n=419$), chromatin structure ($n=299$), redox regulation ($n=102$), epigenetic regulation ($n=44$), cell signaling ($n=2490$), protein kinases ($n=314$), gene transcription ($n=797$), cell cycle ($n=297$), apoptosis ($n=312$), DNA replication ($n=144$) or other functions ($n=92$) (Figure 1a). A total of 5648 genes were thus analyzed with the wafer.

To efficiently isolate oncogenes generated by point mutation using our sequencing array, we selected leukemic blasts with a karyotype characterized by few chromosome anomalies and by few copy number variations of chromosomes, as determined by comparative genomic hybridization with single-nucleotide polymorphism-typing arrays (Supplementary Figure S1). We isolated 15 cases of *de novo* AML, 4 cases of AML that developed from myelodysplastic syndrome, and 1 case of myeloproliferative disorder negative for the JAK2(V617F) and MPL(W515L) mutations (Kralovics *et al.*, 2005; Pikman *et al.*, 2006) (Supplementary Table S1).

From each of these 20 individuals enrolled in the study, we purified immature blasts positive for the surface expression of CD34 (leukemic fraction) as well as a paired control fraction of mature T cells positive for the surface expression of CD4. Although monocytes-macrophages may also express a low level of CD4 at the cell surface, our magnetic bead-based purification system preferentially enriched mature T cells with a high level of CD4 expression; contamination of the mature T-cell fraction with monocytes-macrophages was judged to be $<9\%$ by flow cytometry (Supplementary Figure S2).

Given the potential presence of substantial numbers of unreported single-nucleotide polymorphisms in the human genome, we adopted a two-step analysis to select somatic changes (Figure 1b). In phase I, genomic DNA was isolated from the CD34$^+$ fraction, subjected to mid-range PCR amplification and hybridized with the wafer to examine ~ 9 Mbp of nucleotide sequence. In phase II, we constructed a smaller wafer to investigate only the

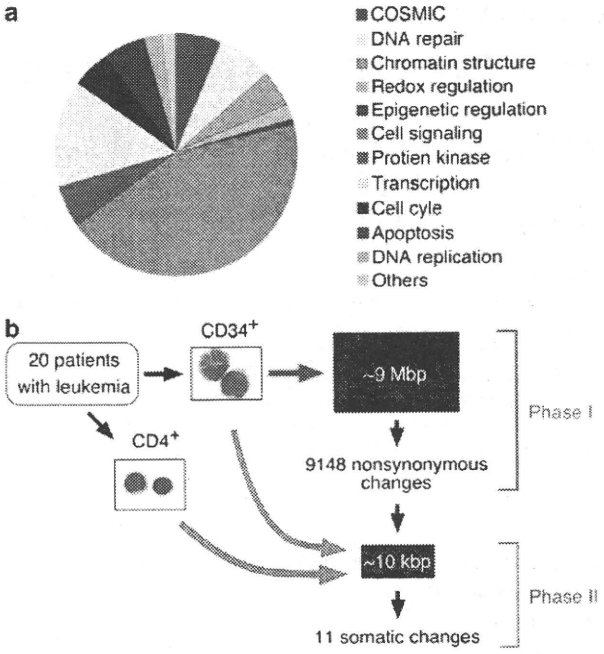


Figure 1 Resequencing of the leukemia genome with wafers. (**a**) Genes interrogated by the phase I wafer ($n=5648$) included those listed in the COSMIC database and those categorized on the basis of function of the encoded protein as indicated. (**b**) CD34$^+$ and CD4$^+$ cell fractions were purified from individuals with leukemia ($n=20$). Genomic DNA of the former fractions was assayed with the phase I wafer including ~ 9 Mbp of sequence, resulting in the isolation of 9148 nonsynonymous nucleotide changes in 3403 independent genes. The phase II wafer was then constructed to analyze these 9148 changes and was hybridized with genomic DNA from both CD34$^+$ and CD4$^+$ fractions separately. Only 11 mutations were found to be present in the former fraction but not in the latter.

nucleotides shown to be changed in phase I relative to the human reference sequence. Genomic DNA isolated from leukemic blasts and paired control fractions was then analyzed individually with the phase II wafer. We assumed that a nucleotide change was a germline polymorphism if it was observed in both leukemic and control fractions of the same individual, and that it was a somatic mutation if it was observed in the former fraction but not in the latter.

Identification of the JAK3(M511I) mutation

Screening of the leukemic blasts of the 20 individuals for point mutations in phase I yielded 9148 nonsynonymous changes among 3403 independent genes, a frequency similar to that observed in other large-scale resequencing studies performed with capillary sequencers (Sjoblom *et al.*, 2006; Greenman *et al.*, 2007). However, analysis of CD4$^+$ fractions showed that most of these sequence changes were also present in the paired control genome, leaving only 11 nonsynonymous somatic mutations in 11 genes (Supplementary Table S2). Such small number of somatic mutations is in a good agreement with the eight somatic mutations found in AML through whole-genome resequencing using the

Illumina Genome Analyser (Illumina, San Diego, CA, USA) (Ley *et al.*, 2008). All of our 11 somatic changes were confirmed by analysis of both genomic DNA and cDNA of the corresponding specimens with a capillary sequencer (data not shown). These data thus support the necessity of examining paired noncancerous specimens to pinpoint somatic changes in the cancer genome.

One of the gene mutations found only in the $CD34^+$ fractions results in a Met-to-Ile change at amino-acid position 511 of JAK3. A heterozygous *JAK3* mutation responsible for the amino-acid change was confirmed in both genomic DNA and cDNA from the $CD34^+$ fraction, but not in those from the corresponding $CD4^+$ fraction of patient ID JM07 (Supplementary Figure S3), who had *de novo* AML (M1 subtype) and a normal karyotype (Supplementary Table S1). In contrast to JAK2, activating mutations in which are preferentially associated with myeloproliferative disorder, several gain-of-function mutations (such as I87T, P132T, Q501H, A572V, R657Q and V722I) of JAK3 have recently been associated with acute megakaryoblastic leukemia of children (Walters *et al.*, 2006; Sato *et al.*, 2008). Other JAK3 mutations (such as A573V and A593T) were also identified in the same disorder, and an M576L substitution was detected in an adult with acute megakaryocytic leukemia (AML, M7 subtype) (Kiyoi *et al.*, 2007), although the transforming potential of these changes remains unknown.

Given that the M511I mutant of JAK3 has not previously been described and that the relevance of JAK3 to the pathogenesis of adult AML has not been extensively investigated, we first focused on the function of JAK3(M511I). The M511 residue is located in the linker region between the Src homology 2 (SH2) domain and the pseudokinase domain of JAK3 (Figure 2a). The transforming mutation Q501H that is associated with juvenile acute megakaryoblastic leukemia (Sato *et al.*, 2008) is also located in this region. Given that JAK3 is abundant in and has an essential role in the development of lymphocytes (Russell *et al.*, 1995), we examined the expression level of *JAK3* in AML blasts. The gene was expressed at a high level in most AML specimens ($n = 52$), with its expression level being greater than that of *JAK2* in all but three cases (Supplementary Figure S3).

To examine the transforming potential of JAK3(M511I), we introduced the mutant or wild-type protein into the interleukin-3 (IL-3)-dependent mouse cell line 32D (Greenberger *et al.*, 1983). Although 32D cells forced to express wild-type JAK3 underwent rapid apoptosis after withdrawal of IL-3, those expressing JAK3(M511I) continued to grow even in the absence of IL-3, although at a reduced rate compared with that of cells expressing the artificially generated, highly transforming mutant JAK3(V674A) (Choi *et al.*, 2007) (Supplementary Figure S3). 32D cells differentiate into terminal granulocytes in the presence of granulocyte colony-stimulating factor. However, cells expressing the M511I or V674A mutant of JAK3 maintained an exponential rate of growth, without any sign of differentiation, in the presence of granulocyte colony-stimulating factor (Figure 2b, Supplementary Figure S3), supporting the notion that the M511I mutant has transforming potential.

To directly examine the leukemogenic activity of JAK3(M511I), we generated a recombinant retrovirus encoding this mutant and used it to infect murine hematopoietic stem cells. Reconstitution of the bone marrow of lethally irradiated mice with such infected cells resulted in marked lymphocytosis in peripheral blood and enlargement of the spleen in the recipient animals (Figure 2c). The cells in the peripheral blood, spleen and bone marrow of the recipients responsible for these phenotypes manifested a medium-sized, blastic morphology, and flow cytometric analysis revealed them to be $CD8^+$ T cells (Figure 2c). The clonal nature of these proliferating T cells was further confirmed by Southern blot analysis (Supplementary Figure S4), indicative of the development of T-cell acute lymphoblastic leukemia in the recipient mice.

To assess the prevalence of *JAK3* mutations in adult leukemia, we further examined the nucleotide sequence of the entire coding region of *JAK3* cDNA in an additional 266 specimens of leukemic blasts. The coding region of *JAK3* cDNA was successfully amplified by PCR from 83 specimens. We could further identify 4 distinct *JAK3* sequence changes in 8 of these 83 samples: 1 case with G62S, 4 cases with Q501H, 2 cases with R657Q and 1 case with R918C (Figure 2a). Taking into account the 20 cases evaluated in the phase I analysis, we thus identified a total of 9 cases with a mutant form of JAK3 (3.1%) among 286 cases of leukemia (Supplementary Table S3). Our identification of known transforming JAK3 mutants (Q501H and R657Q) originally associated with acute megakaryoblastic leukemia prompted us to determine the prevalence of these two changes in another cohort of AML ($n = 148$), revealing two cases with JAK3(Q501H) and one case with JAK3(R657Q). In addition, analysis of a hematopoietic cell line (KCL22) (Kubonishi and Miyoshi, 1983) established from a patient with chronic myeloid leukemia in BC revealed yet another mutation (L1017M) of JAK3 (Figure 2a).

To directly compare the transforming potential of these various JAK3 mutants, we introduced each protein into the IL-3-dependent mouse B-cell line BA/F3 and examined the growth properties of the resulting transfectants. Whereas all cells expressing the JAK3 mutants proliferated in a similar manner in the presence of IL-3 (data not shown), culture without IL-3 revealed marked differences in the transforming potential among the mutants. JAK3(M511I) was the most efficient oncokinase, with a transforming activity similar to that of JAK3(V674A). The frequent mutants JAK3(Q501H) and JAK3(R657Q) exhibited weaker but still pronounced transforming potential, whereas the remaining mutants (G62S, R918C and L1017M) showed an even lower potential (Figure 2d).

Somatic mutations of DNMT3A

Another somatic mutation identified in the phase II data set was a heterozygous change in *DNMT3A* that results

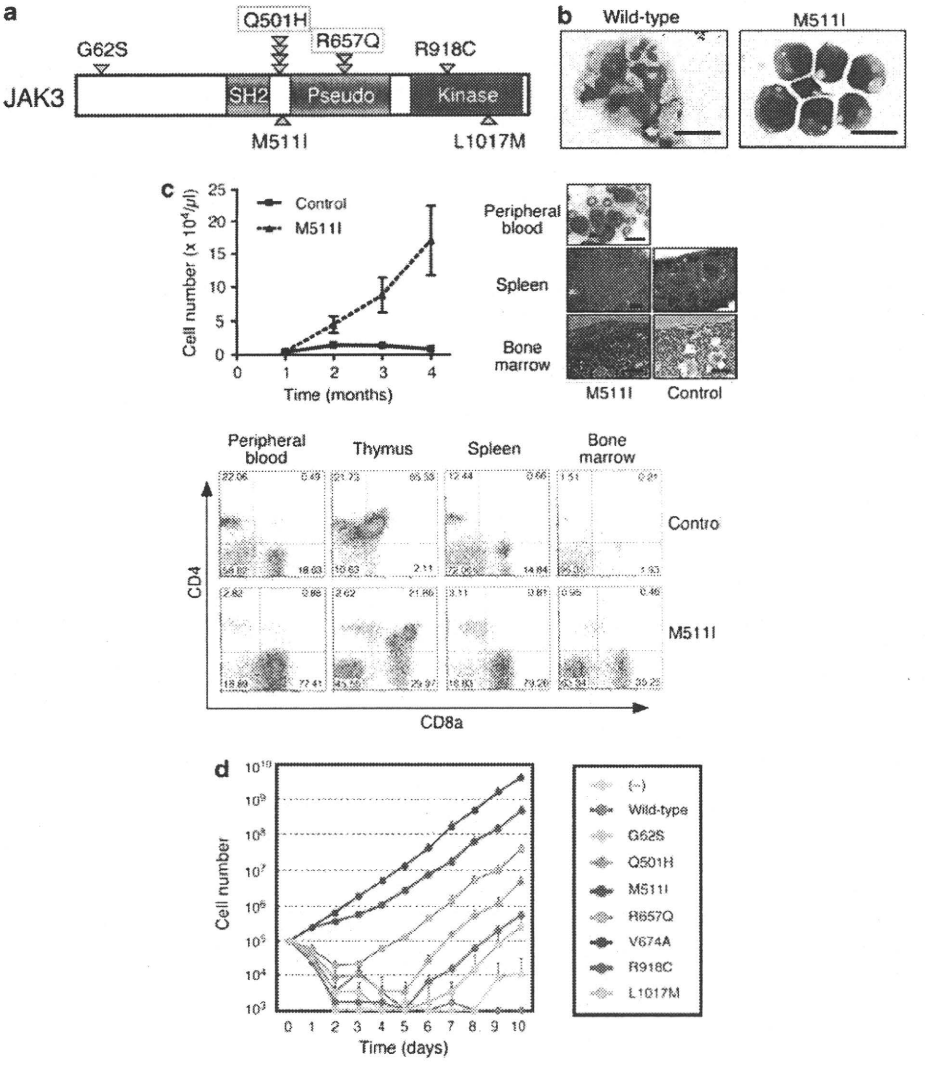


Figure 2 Identification of JAK3 mutants in leukemia. (**a**) Amino-acid substitutions detected in this study are shown relative to the domain organization of JAK3. The mutations M511I (one case) and Q501H (four cases) are located in the linker region between the SH2 and pseudokinase domains of JAK3, whereas G62S (one case), R657Q (two cases) and R918C (one case) are located in the amino-terminal region, the pseudokinase domain and the kinase domain, respectively. The KCL22 cell line also harbors an L1017M mutation within the kinase domain of JAK3. Previously known activating mutations of JAK3 (Q501H and R657Q) are indicated by red rectangles. (**b**) Mouse 32D cells expressing wild-type human JAK3 or the JAK3(M511I) mutant were incubated with G-CSF (0.5 ng/ml) for 14 days, stained with Wright–Giemsa solution and examined by light microscopy. Scale bars, 20 μm. (**c**) C57BL/6 mice were irradiated and then injected intravenously with syngeneic $CD34^-KSL$ hematopoietic stem cells infected with a retrovirus encoding JAK3(M511I) or the corresponding empty virus (control). The number of white blood cells in peripheral blood was counted at the indicated times thereafter; data are means ± s.d. for 10 mice in each group (upper left panel). Peripheral blood, spleen and bone marrow isolated from recipient mice 3 months after cell injection were stained with the Wright–Giemsa solution (peripheral blood) or hematoxylin–eosin (spleen and bone marrow) and were then examined by light microscopy (upper right panel); scale bars represent 10, 200 and 50 μm, respectively. Mononuclear cells isolated from peripheral blood, thymus, spleen and bone marrow of recipient mice 3 months after cell injection were subjected to flow cytometric analysis of surface expression of CD4 and CD8a (lower panel). (**d**) Control BA/F3 cells (−) or those expressing the indicated JAK3 mutants were cultured without IL-3 for the indicated times, after which the cell number was determined. Data are means + s.d. of triplicates from a representative experiment.

in an R882H substitution in the encoded protein (Figure 3a, Supplementary Figure S5). DNMT3A, together with DNMT3B, has an essential role in *de novo* methylation of the human genome (Okano *et al.*, 1999), and an aberrant methylation profile (hypermethylation of CpG islands and hypomethylation of other regions) is a hallmark of cancer cells (El-Osta, 2004). Despite a direct linkage between such methylation changes and silencing of tumor-suppressor genes in cancer, the molecular mechanism responsible for such abnormal methylation remains unknown. Our data thus provide the first evidence of somatic mutation of a DNA methyltransferase gene in cancer cells. Mutations in the catalytic domain of DNMT3B have been shown to be

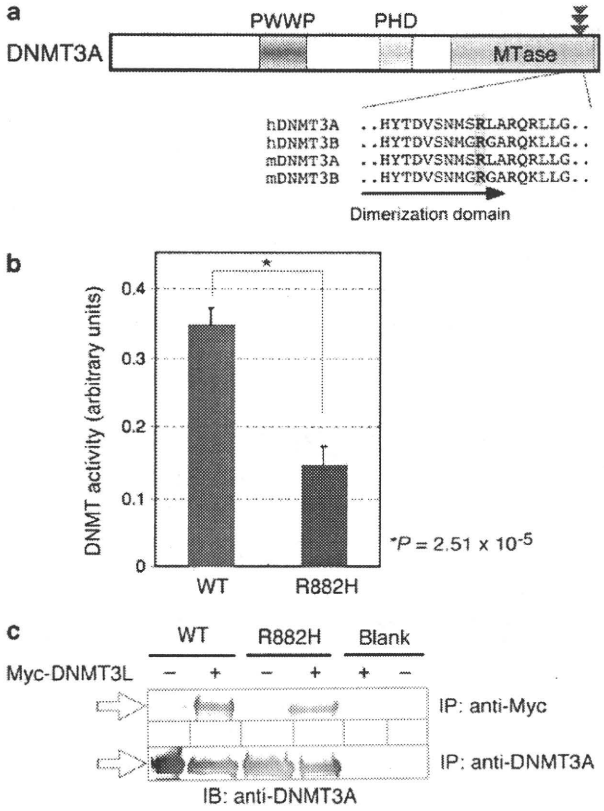


Figure 3 Identification of a DNMT3A mutant in leukemia. (**a**) Domain organization of human DNMT3A showing that the R882 residue found to be mutated in leukemia is conserved among human (h) and mouse (m) members of the DNMT3 family. DNMT3A contains a tetrapeptide PWWP domain, polybromo homology domain (PHD) and methyltransferase (MTase) domain. The R882 residue is located in the homodimerization region present within the MTase domain. (**b**) Wild-type (WT) and R882H forms of DNMT3A were expressed in and purified from insect cells and then subjected to an *in vitro* assay of methyltransferase activity. Data are means + s.d. of triplicates from a representative experiment. The *P*-value was determined by Student's *t*-test. (**c**) Lysates of HEK293 cells expressing Myc epitope-tagged DNMT3L and wild-type or R882H forms of DNMT3A, as indicated, were subjected to immunoprecipitation (IP) with antibodies to Myc or to DNMT3A, and the resulting precipitates were subjected to immunoblot analysis (IB) with antibodies to DNMT3A. The position of DNMT3A (wild-type or mutant) is indicated by an open arrow.

responsible for a hereditary syndrome characterized by ICF (immunodeficiency, instability of the centromeric region of chromosomes and facial anomalies) in humans (Ehrlich, 2003). One of the mutation sites of DNMT3B (R823) associated with the ICF syndrome corresponds to the residue of DNMT3A (R882) shown to be mutated in this study.

The R882 residue of DNMT3A is considered to participate in the homodimerization and activation of the protein (Jia *et al*., 2007) (Figure 3a). To determine whether the R882H mutation affects the catalytic activity of DNMT3A, we expressed mutant and wild-type proteins separately in insect cells, purified them to near homogeneity and subjected them to an *in vitro* assay of methyltransferase activity with a synthetic substrate (Suetake *et al*., 2003). The catalytic activity of DNMT3A(R882H) was $<50\%$ of that of the wild-type protein (Figure 3b). DNMT3L acts as a coactivator for the methyltransferase activity of DNMT3A or DNMT3B through its association with the latter proteins (Jia *et al*., 2007). The R882H mutation did not affect the interaction of DNMT3A with DNMT3L in transfected mammalian cells (Figure 3c) or its sensitivity to DNMT3L as examined by the *in vitro* assay of methyltransferase activity (data not shown). These data thus suggested that the R882H mutation directly inhibits the enzymatic activity of DNMT3A.

Screening of another cohort of leukemia cases ($n = 54$) for mutant forms of *DNMT3A* revealed another two patients with a mutation of the same amino acid (R882H in one patient and R882C in the other) (Supplementary Table S4). Therefore, we identified a total of 3 cases with an R882 mutation (4.1%) among 74 cases of leukemia. Screening for mutations of *DNMT3B* failed to detect any somatic changes in the same individuals (data not shown), suggesting that *DNMT3A* is a preferential target in leukemia.

Multistep transformation in leukemia

Although $>99\%$ of nucleotide changes in the phase I data were also observed in the paired CD4$^+$ cells, it is unlikely that all of these changes are actually germline polymorphisms because they include established oncogenic mutations. They thus include 190 nucleotide changes previously described in cancer cells (Supplementary Table S5), such as those giving rise to NRAS(Q61H) in patient ID JM17 and to FLT3(D835Y) in patient ID JM08 (Figure 4a). Given that both NRAS(Q61H) and FLT3(D835Y) are well-characterized oncoproteins (Yamamoto *et al*., 2001), it is unlikely that these individuals harbored such nucleotide changes in the germ line. There are at least two possible explanations for these findings. First, it is possible that purification of the CD4$^+$ fraction was not efficient, with the result that this fraction was contaminated by CD34$^+$ cells. However, the *CD4* expression ratio for the CD4$^+$ and CD34$^+$ fractions of each individual was $\geqslant 17.1$ (median $= 40.1$) (Figure 4b), and contamination of the CD4$^+$ fraction with CD34$^+$ cells at such a level would not likely produce detectable changes in Sanger sequencing outputs (compare, for instance, the signal intensities of the normal and mutant alleles in Figure 4a).

Furthermore, although CD4 expression has been occasionally observed in AML blasts (Schwonzen *et al*., 2007), quantitation of *CD4* and *CD34* mRNA within our purified CD34$^+$ fractions failed to detect a significant level of the former message in the blasts (Supplementary Figure S6). Therefore, it is unlikely that contamination of CD4$^+$ leukemic blasts within the purified, control CD4$^+$ fraction substantially affected the sequencing results in our phase II experiment.

Rather, it is more likely that leukemia may develop in a stepwise manner with a substantial time interval

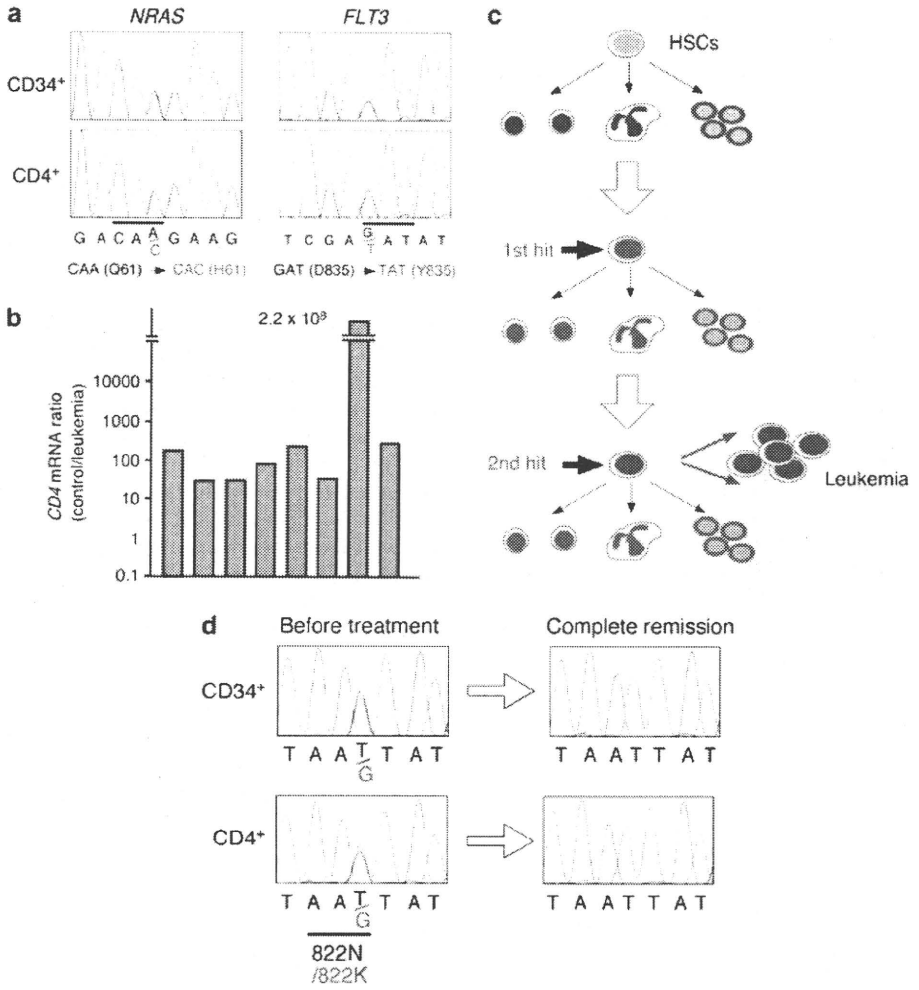


Figure 4 Proposed stepwise nature of leukemogenesis. (**a**) Sequencing electrophoretograms for the regions surrounding codon 61 of *NRAS* or codon 835 of *FLT3* in genomic DNA from the CD34$^+$ and CD4$^+$ fractions of patient IDs JM17 and JM08, respectively. Heterozygous nucleotide changes that give rise to NRAS(Q61H) or FLT3(D835Y) were detected in both fractions of the corresponding patients. (**b**) The amount of *CD4* mRNA in the CD4$^+$ (control) and CD34$^+$ (leukemia) fractions of leukemia patients (with a substantial amount of control *GAPDH* mRNA) was quantitated by reverse transcription and real-time PCR analysis and expressed as the control/leukemia ratio. (**c**) Hematopoietic stem cells (HSCs) give rise to a wide range of mature blood cells. Even after the first hit (mutation) of the genome, HSCs retain their full differentiation capacity, and therefore produce differentiated cells harboring this first hit. After the second hit, the affected cell fraction undergoes full transformation to leukemia. (**d**) Sequencing electrophoretograms for the genome of CD34$^+$ and CD4$^+$ fractions from patient ID JM03 showing a heterozygous mutation for KIT(N822K) before chemotherapy but not after.

between steps (Figure 4c). If a first hit occurs in the genome of hematopoietic stem (or progenitor) cells and if such a somatic change does not result directly in the generation of full-blown leukemia, the preleukemic clones may give rise to terminally differentiated blood cells (including CD4$^+$ cells). After a certain period, a second (or possibly a third) hit occurs in the immature cells and triggers the rapid growth of leukemic clones without differentiation. In such a scenario, terminally differentiated 'normal' cells may still harbor the first hit in their genome.

Support for this latter possibility was provided by patient ID JM03, who had AML (M2 subtype) with a t(8;21) chromosome anomaly. Before chemotherapy, the genomic DNA of both CD34$^+$ and CD4$^+$ fractions from this patient harbored a heterozygous mutation of *KIT* that results in the production of a constitutively activated mutant protein, KIT(N822K) (Shimada *et al.*, 2006) (Figure 4d). The same change was also detected in cDNA prepared from the CD34$^+$ fraction (data not shown). Leukemic blasts in this patient were sensitive to standard chemotherapeutic regimens, and the patient underwent complete remission. Examination of CD34$^+$ and CD4$^+$ fractions obtained during the remission period revealed that the N822K codon change was no longer detectable not only in the CD34$^+$ fraction but also in the CD4$^+$ fraction (Figure 4d). These data thus support the scenario shown in Figure 4c: The N822K

change represents the first hit and was present in differentiated blood cells, and the corresponding pre-leukemic clones were simultaneously eradicated together with the leukemic clones by chemotherapy.

On the other hand, as shown in Supplementary Tables S1 and S2, a heterozygous mutation for NRAS(G12S) was found only in the $CD34^+$ fraction, but not in the $CD4^+$ fraction of the patient ID JM16. Conventional chemotherapy for this patient eradicated the leukemic blasts carrying the mutation (Supplementary Figure S7), also confirming that a successful treatment results in the disappearance of cells with a (possible) 'second hit'.

Our hypothesis of the stepwise leukemogenesis is also consistent with the previous detection of the *RUNX1-CBFA2T1* oncogene in differentiated blood cells (Kwong *et al.*, 1996; Miyamoto *et al.*, 1996, 2000).

Discussion

Our large-scale genomic resequencing of human leukemia specimens with DNA microarrays has identified recurrent nucleotide changes responsible for the generation of JAK3 and DNMT3A mutants. Whereas JAK3 mutants were unexpectedly found in adult AML, their transforming ability, and possibly their contribution to leukemogenesis, varied substantially. However, our bone marrow transplantation experiments showed that at least one of these JAK3 mutants (M511I) directly participates in the development of leukemia. Identification of the M511I mutation of JAK3 in the leukemic fraction but not in the control fraction of patient ID JM07 suggests that this mutation may be the second hit triggering AML. Given that the blasts of this patient had a normal karyotype, it is likely that the first hit is present in the genome of both fractions. Karyotyping of other patients with *JAK3* mutations showed a total of three cases with a normal karyotype, one case with t(8;21), and one case with a numerical anomaly of several chromosomes (Supplementary Table S3), suggesting that *JAK3* mutations may be preferentially associated with leukemia with a normal karyotype.

Although JAK3(M511I) was identified in AML, our bone marrow transplantation experiments with hematopoietic stem cells expressing this mutant yielded T-cell acute lymphoblastic leukemia. In contrast to human leukemia, in which JAK3 changes may constitute a second hit (probably in progenitor cells), JAK3(M511I) may have been expressed in all hematopoietic cells of the recipient mice. JAK3(M511I) thus likely triggered leukemia within a T-cell fraction the intracellular context of which is optimized for JAK3 signaling.

It has been frequently observed that transgenic mouse or bone marrow transplantation experiments for leukemic oncogenes do not accurately recapitulate the original leukemia subtypes (Wong and Witte, 2001). Transgenic mice expressing $p210^{BCR\text{-}ABL1}$, for instance, usually develop T-cell lymphoma or acute lymphoblastic leukemia, not chronic myeloid leukemia. Furthermore, bone marrow transplantation with hematopoietic progenitor cells expressing $p210^{BCR\text{-}ABL1}$ often leads to development of lymphoma, AML, acute lymphoblastic leukemia or macrophage tumors. Generation of malignancy in such systems may, thus, be elaborately influenced by mouse strains, promoter fragments for artificial expression and/or cell types to be used for gene transduction.

Our detection of recurrent DNMT3A hypomorphic mutations in leukemia clones may indicate the presence of an abnormal methylation profile in the genome of such blasts. However, given the limited amount of the specimens available, we were able to investigate microsatellite stability only at certain loci (Koinuma *et al.*, 2005), revealing no apparent microsatellite instability (data not shown). We also generated BA/F3 cells expressing wild-type or R882H forms of DNMT3A to compare the methylation status of some CpG islands in the genome; again, we detected no discernable differences between the two cell preparations (data not shown). However, given that BA/F3 cells contained two copies of wild-type *Dnmt3a* in addition to multiple copies of mutant *DNMT3A*, whereas the leukemic blasts likely harbor one copy each of the wild-type and mutant *DNMT3A* alleles, the clinical relevance of the R882 mutant requires further examination under the latter condition. Cell proliferation/differentiation is indeed influenced substantially by the copy number of *DNMT3* genes (Okano *et al.*, 1999; Ehrlich, 2003).

Our observations indicate the importance of preparing paired normal fractions in large-scale resequencing projects, but they also reveal a difficulty in the preparation of *bona fide* 'normal' fractions in the case of leukemic disorders. Our data thus indicate that nonleukemic blood cells may harbor early genomic hits, rendering them inappropriate as controls. Furthermore, a substantial proportion of fingernail DNA was recently shown to be derived from donor cells among recipients of allogeneic stem cell transplants (Imanishi *et al.*, 2007), indicating that nonblood cells may contain DNA derived from transplanted cells. Therefore, it is possible that buccal, fingernail or even hair cells may not be suitable as normal cell controls. In contrast to solid tumors, for which blood cells are appropriate as paired normal fractions, leukemic disorders require that caution be taken to discriminate somatic nucleotide changes from germline polymorphisms.

Materials and methods

Wafer sequencing

$CD34^+$ and $CD4^+$ fractions were isolated from leukemic individuals using CD34microbeads and CD4microbeads, respectively, and a MidiMACS separator (Miltenyi Biotec, Gladbach, Germany). All clinical specimens were obtained with written informed consent, and the study was approved by the ethics committees of both the Jichi Medical University and the Nagasaki University. DNA sequencing wafers were designed and processed at Perlegen Sciences. Genes to be interrogated on the wafers were selected from the Entrez Gene database (http://www.ncbi.nlm.nih.gov/sites/entrez?db = gene)

by searching with various keywords characteristic to each subcategory (such as DNA repair, regulation of chromatin structure, etc.), followed by manual inspection. The final gene list for the wafers is shown in Supplementary Table S6. Construction of the wafers, quality control analysis and data processing are described in Supplementary Text.

JAK3 analysis

Complementary DNAs for JAK3 mutants were generated using a QuikChange site-directed mutagenesis kit (Stratagene, La Jolla, CA, USA) and ligated into the pMX retroviral vector (Onishi *et al.*, 1996). Ecotropic recombinant retroviruses encoding each mutant were produced in BOSC23 cells transfected with the corresponding pMX-based plasmid and were used to infect BA/F3 or 32D cells as described previously (Choi *et al.*, 2007). Both types of cell were cultured in RPMI 1640 medium supplemented with 10% fetal bovine serum (both from Life Technologies, Carlsbad, CA, USA) and mouse IL-3 (Sigma, St Loius, MO, USA) at 10 Units/ml; differentiation of 32D cells was induced by culture in the presence of serum and mouse granulocyte colony-stimulating factor (Sigma) at 0.5 ng/ml. A concentrated preparation of a retrovirus with a VSV-G envelope and encoding both JAK3(M511I) and enhanced green fluorescent protein was used to infect $CD34^-$ c-Kit^+ $Sca\text{-}1^+$ $Lineage\text{-}marker^-$ ($CD34^-KSL$) hematopoietic stem cells isolated from the bone marrow of C57BL/6 mice, and the infected cells were transplanted into lethally irradiated mice congenic for the *Ly5* locus (Iwama *et al.*, 2004). *CD4*, *JAK2* and *JAK3* mRNAs were quantitated by reverse transcription and real-time PCR analysis using an ABI7900HT system (Life Technologies) and with the primers 5′-CTGGAATCCAACATCAAGGTTCTG-3′ and 5′-AATTGTAGAGGAGGCGAACAGGAG-3′ for CD4, 5′-CTCCAGAATCACTGACAGAGAGCA-3′ and 5′-CCAC TCGAAGAGCTAGATCCCTAA-3′ for *JAK2* and 5′-GAGC TCTTCACCTACTGCGACAAA-3′ and 5′-AGCTATGAAA AGGACAGGGAGTGG-3′ for JAK3; the cDNA for GAPDH (glyceraldehyde-3-phosphate dehydrogenase) was also amplified with the primers 5′-GTCAGTGGTGGACC TGACCT-3′ and 5′-TGAGCTTGACAAAGTGGTCG-3′. The relative abundance of the cDNAs of interest was calculated from the threshold cycle (C_T) for each cDNA and that for *GAPDH* cDNA.

DNMT3A analysis

Recombinant His_6-tagged DNMT3A or DNMT3A(R882H) was expressed in SF9 cells using the Bac-to-Bac baculovirus expression system (Invitrogen, Carlsbad, CA, USA), and each protein was purified by stepwise column chromatography as described previously (Suetake *et al.*, 2003). The enzymatic activity of each protein was assayed with -*S*-adenosyl-L-methionine (GE Healthcare, Waukesha, WI, USA) and dIdC or dGdC as substrates (Suetake *et al.*, 2003). The association between Myc epitope-tagged human DNMT3L and wild-type or R882H forms of human DNMT3A in transfected HEK293 cells was examined by immunoprecipitation and immunoblot analyses.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

We thank D Cox, KA Frazer, DG Ballinger, J Montgomery, H Tao, C Chen, L Stuve, J Kwon, J Sheehan and Y Zhan for discussion on the wafer experiments, as well as JN Ihle, T Kitamura and SB Baylin for human *JAK3* cDNA, the pMX plasmid and human *DNMT3A* cDNA, respectively. This study was supported in part by a grant for Third-Term Comprehensive Control Research for Cancer from the Ministry of Health, Labor, and Welfare of Japan, and by a grant for Scientific Research on Priority Areas 'Applied Genomics' from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

References

Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG *et al.* (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**: 53–59.

Byrd JC, Mrozek K, Dodge RK, Carroll AJ, Edwards CG, Arthur DC *et al.* (2002). Pretreatment cytogenetic abnormalities are predictive of induction success, cumulative incidence of relapse, and overall survival in adult patients with *de novo* acute myeloid leukemia: results from cancer and leukemia Group B (CALGB 8461). *Blood* **100**: 4325–4336.

Choi YL, Kaneda R, Wada T, Fujiwara S, Soda M, Watanabe H *et al.* (2007). Identification of a constitutively active mutant of JAK3 by retroviral expression screening. *Leuk Res* **31**: 203–209.

Ehrlich M. (2003). The ICF syndrome, a DNA methyltransferase 3B deficiency and immunodeficiency disease. *Clin Immunol* **109**: 17–28.

El-Osta A. (2004). The rise and fall of genomic methylation in cancer. *Leukemia* **18**: 233–237.

Greenberger JS, Sakakeeny MA, Humphries RK, Eaves CJ, Eckner RJ. (1983). Demonstration of permanent factor-dependent multipotential (erythroid/neutrophil/basophil) hematopoietic progenitor cell lines. *Proc Natl Acad Sci USA* **80**: 2931–2935.

Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G *et al.* (2007). Patterns of somatic mutation in human cancer genomes. *Nature* **446**: 153–158.

Grimwade D, Walker H, Oliver F, Wheatley K, Harrison C, Harrison G *et al.* (1998). The importance of diagnostic cytogenetics on outcome in AML: analysis of 1612 patients entered into the MRC AML 10 trial. The Medical Research Council Adult and Children's Leukaemia Working Parties. *Blood* **92**: 2322–2333.

Imanishi D, Miyazaki Y, Yamasaki R, Sawayama Y, Taguchi J, Tsushima H *et al.* (2007). Donor-derived DNA in fingernails among recipients of allogeneic hematopoietic stem-cell transplants. *Blood* **110**: 2231–2234.

Iwama A, Oguro H, Negishi M, Kato Y, Morita Y, Tsukui H *et al.* (2004). Enhanced self-renewal of hematopoietic stem cells mediated by the polycomb gene product Bmi-1. *Immunity* **21**: 843–851.

Jia D, Jurkowska RZ, Zhang X, Jeltsch A, Cheng X. (2007). Structure of Dnmt3a bound to Dnmt3L suggests a model for *de novo* DNA methylation. *Nature* **449**: 248–251.

Kiyoi H, Yamaji S, Kojima S, Naoe T. (2007). JAK3 mutations occur in acute megakaryoblastic leukemia both in Down syndrome children and non-Down syndrome adults. *Leukemia* **21**: 574–576.

Koinuma K, Kaneda R, Toyota M, Yamashita Y, Takada S, Choi YL *et al.* (2005). Screening for genomic fragments that are methylated specifically in colorectal carcinoma with a methylated MLH1 promoter. *Carcinogenesis* **26**: 2078–2085.

Kralovics R, Passamonti F, Buser AS, Teo SS, Tiedt R, Passweg JR *et al.* (2005). A gain-of-function mutation of JAK2 in myeloproliferative disorders. *N Engl J Med* **352**: 1779–1790.

Kubonishi I, Miyoshi I. (1983). Establishment of a Ph1 chromosome-positive cell line from chronic myelogenous leukemia in blast crisis. *Int J Cell Cloning* **1**: 105–117.

Kwong YL, Wong KF, Chan V, Chan CH. (1996). Persistence of AML1 rearrangement in peripheral blood cells in t(8;21). *Cancer Genet Cytogenet* **88**: 151–154.

Ley TJ, Mardis ER, Ding L, Fulton B, McLellan MD, Chen K *et al.* (2008). DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature* **456**: 66–72.

Miyamoto T, Nagafuji K, Akashi K, Harada M, Kyo T, Akashi T *et al.* (1996). Persistence of multipotent progenitors expressing AML1/ETO transcripts in long-term remission patients with t(8;21) acute myelogenous leukemia. *Blood* **87**: 4789–4796.

Miyamoto T, Weissman IL, Akashi K. (2000). AML1/ETO-expressing nonleukemic stem cells in acute myelogenous leukemia with 8;21 chromosomal translocation. *Proc Natl Acad Sci USA* **97**: 7521–7526.

Nimer SD, Moore MA. (2004). Effects of the leukemia-associated AML1-ETO protein on hematopoietic stem and progenitor cells. *Oncogene* **23**: 4249–4254.

Okano M, Bell DW, Haber DA, Li E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell* **99**: 247–257.

Onishi M, Kinoshita S, Morikawa Y, Shibuya A, Phillips J, Lanier LL *et al.* (1996). Applications of retrovirus-mediated expression cloning. *Exp Hematol* **24**: 324–329.

Patil N, Berno AJ, Hinds DA, Barrett WA, Doshi JM, Hacker CR *et al.* (2001). Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. *Science* **294**: 1719–1723.

Pikman Y, Lee BH, Mercher T, McDowell E, Ebert BL, Gozo M *et al.* (2006). MPLW515L is a novel somatic activating mutation in myelofibrosis with myeloid metaplasia. *PLoS Med* **3**: e270.

Russell SM, Tayebi N, Nakajima H, Riedy MC, Roberts JL, Aman MJ *et al.* (1995). Mutation of Jak3 in a patient with SCID: essential role of Jak3 in lymphoid development. *Science* **270**: 797–800.

Sato T, Toki T, Kanezaki R, Xu G, Terui K, Kanegane H *et al.* (2008). Functional analysis of JAK3 mutations in transient myeloproliferative disorder and acute megakaryoblastic leukaemia accompanying Down syndrome. *Br J Haematol* **141**: 681–688.

Schlenk RF, Dohner K, Krauter J, Frohling S, Corbacioglu A, Bullinger L *et al.* (2008). Mutations and treatment outcome in cytogenetically normal acute myeloid leukemia. *N Engl J Med* **358**: 1909–1918.

Schwonzen M, Diehl V, Dellanna M, Staib P. (2007). Immunophenotyping of surface antigens in acute myeloid leukemia by flow cytometry after red blood cell lysis. *Leuk Res* **31**: 113–116.

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

Sjoblom T, Jones S, Wood LD, Parsons DW, Lin J, Barber TD *et al.* (2006). The consensus coding sequences of human breast and colorectal cancers. *Science* **314**: 268–274.

Suetake I, Miyazaki J, Murakami C, Takeshima H, Tajima S. (2003). Distinct enzymatic properties of recombinant mouse DNA methyltransferases Dnmt3a and Dnmt3b. *J Biochem* **133**: 737–744.

Tallman MS, Altman JK. (2008). Curative strategies in acute promyelocytic leukemia. *Hematol Am Soc Hematol Educ Program* **2008**: 391–399.

Walters DK, Mercher T, Gu TL, O'Hare T, Tyner JW, Loriaux M *et al.* (2006). Activating alleles of JAK3 in acute megakaryoblastic leukemia. *Cancer Cell* **10**: 65–75.

Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A *et al.* (2008). The complete genome of an individual by massively parallel DNA sequencing. *Nature* **452**: 872–876.

Wong S, Witte ON. (2001). Modeling Philadelphia chromosome positive leukemias. *Oncogene* **20**: 5644–5659.

Yamamoto Y, Kiyoi H, Nakano Y, Suzuki R, Kodera Y, Miyawaki S *et al.* (2001). Activating mutation of D835 within the activation loop of FLT3 in human hematologic malignancies. *Blood* **97**: 2434–2439.

Supplementary Information accompanies the paper on the Oncogene website (http://www.nature.com/onc)

nature medicine

PU.1-mediated upregulation of *CSF1R* is crucial for leukemia stem cell potential induced by MOZ-TIF2

Yukiko Aikawa[1], Takuo Katsumoto[1], Pu Zhang[2], Haruko Shima[1], Mika Shino[1], Kiminori Terui[3], Etsuro Ito[3], Hiroaki Ohno[4], E Richard Stanley[5], Harinder Singh[6], Daniel G Tenen[2,7] & Issay Kitabayashi[1]

Leukemias and other cancers possess self-renewing stem cells that help to maintain the cancer[1,2]. Cancer stem cell eradication is thought to be crucial for successful anticancer therapy. Using an acute myeloid leukemia (AML) model induced by the leukemia-associated monocytic leukemia zinc finger (MOZ)-TIF2 fusion protein, we show here that AML can be cured by the ablation of leukemia stem cells. The MOZ fusion proteins MOZ-TIF2 and MOZ-CBP interacted with the transcription factor PU.1 to stimulate the expression of macrophage colony–stimulating factor receptor (CSF1R, also known as M-CSFR, c-FMS or CD115). Studies using PU.1-deficient mice showed that PU.1 is essential for the ability of MOZ-TIF2 to establish and maintain AML stem cells. Cells expressing high amounts of CSF1R (CSF1R^{high} cells), but not those expressing low amounts of CSF1R (CSF1R^{low} cells), showed potent leukemia-initiating activity. Using transgenic mice expressing a drug-inducible suicide gene controlled by the *CSF1R* promoter, we cured AML by ablation of CSF1R^{high} cells. Moreover, induction of AML was suppressed in CSF1R-deficient mice and CSF1R inhibitors slowed the progression of MOZ-TIF2–induced leukemia. Thus, in this subtype of AML, leukemia stem cells are contained within the CSF1R^{high} cell population, and we suggest that targeting of PU.1-mediated upregulation of CSF1R expression might be a useful therapeutic approach.

Chromosomal translocations that involve the *MOZ* gene[3] (official gene symbol *Myst3*) are typically associated with acute myelomonocytic leukemia and predict a poor prognosis[4]. Whereas MOZ is essential for the self-renewal of hematopoietic stem cells[5,6], MOZ fusion proteins enable the transformation of non–self-renewing myeloid progenitors into leukemia stem cells[7]. We previously generated a mouse model for AML by introducing c-Kit$^+$ mouse myeloid stem/progenitor cells infected with a retrovirus encoding MOZ-TIF2 and EGFP into lethally irradiated mice[8].

To identify leukemia-initiating cells (LICs), we investigated the bone marrow cells of these mice for various cell surface markers by FACS analysis. CSF1R^{high} and CSF1R^{low} cells were present in the bone marrow (**Fig. 1a**) and expressed equivalent amounts of MOZ-TIF2 protein (**Fig. 1b**). To determine the LIC activity of these cell populations, we isolated CSF1R^{high} and CSF1R^{low} cells by cell sorting and transplanted limited numbers (10 to 1 × 10^4 cells) into irradiated mice. One hundred CSF1R^{high} cells were sufficient to induce AML in all transplanted mice (**Fig. 1c**). Conversely, no mice developed AML after 1 × 10^3 CSF1R^{low} cells were transplanted per mouse, and only half of the mice developed AML with delayed onset when 1 × 10^4 CSF1R^{low} cells were transplanted (**Fig. 1d**). Thus, the CSF1R^{high} cells showed a >100-fold stronger LIC activity than CSF1R^{low} cells.

FACS analysis indicated that the CSF1R^{high} cell population had the phenotype of both granulocyte-macrophage progenitors (GMPs, Kit$^+$Sca-1$^-$CD16/CD32$^+$) and differentiated monocytes (Mac-1^{low}Gr-1$^+$) (**Supplementary Fig. 1a**). Comparison of the CSF1R^{high} and CSF1R^{low} cell populations indicated that Mac-1 expression was lower in CSF1R^{high} than in CSF1R^{low} cells (**Fig. 1e**). However, we did not observe significant differences between the CSF1R^{high} and CSF1R^{low} cell populations with respect to their cell morphology (**Fig. 1f**), colony-forming ability in methylcellulose medium (**Fig. 1g**), cell cycle distribution (**Supplementary Fig. 1b**) or homeobox A9 (HoxA9) expression (**Supplementary Fig. 1c**). To investigate whether downstream pathways of CSF1R signaling were activated, we measured phosphorylation levels of signal transducer and activator of traranscription-5 (STAT5) and extracellular signal–regulated kinase (ERK) in CSF1R^{high} and CSF1R^{low} cells. STAT5 was highly phosphorylated in the CSF1R^{high} cell population but not in the CSF1R^{low} population, whereas ERK was equivalently phosphorylated in the two cell populations (**Fig. 1h**).

Side population cells, which are present in some types of normal and malignant stem cell populations, were present in the bone marrow of MOZ-TIF2–induced AML mice (**Supplementary Fig. 2a**). Whereas most side population cells were CSF1R^{high}, the non–side population fraction contained both CSFR1^{high} and CSF1R^{low} cells (**Supplementary Fig. 2b**). LICs were approximately tenfold more enriched in the side population fraction than in the non–side population fraction (**Supplementary Fig. 2c,d**). Because the side population fraction was very small (~0.12% of total bone marrow cells), the fraction of LICs in the side population fraction was also small (~1% of all LICs), and most LICs were present in the non–side population fraction (~99%).

[1]Molecular Oncology Division, National Cancer Center Research Institute, Tsukiji, Chuo-ku, Tokyo, Japan. [2]Harvard Stem Cell Institute, Boston, Massachusetts, USA. [3]Department of Pediatrics, Hirosaki University School of Medicine, Hirosaki, Japan. [4]Pharmacological Research Laboratories, Research Division, Kyowa Hakko Kirin, Gunma, Japan. [5]Albert Einstein College of Medicine, Bronx, New York, USA. [6]Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, Illinois, USA. [7]Cancer Science Institute, National University of Singapore, Singapore. Correspondence should be addressed to I.K. (ikitabay@ncc.go.jp).

Received 11 January; accepted 18 February; published online 25 April 2010; doi:10.1038/nm.2122

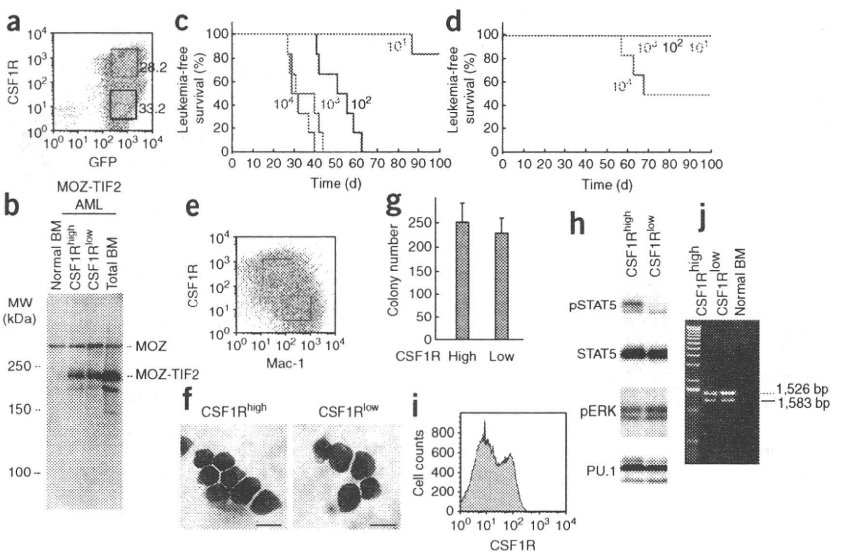


Figure 1 CSF1R^{high} cells show potent leukemia-initiating activity. (**a**) FACS analysis of bone marrow cells from mice with MOZ-TIF2–induced AML for expression of GFP and CSF1R. The red and black boxes signify CSF1R^{high} and CSF1R^{low} cell fractions, respectively. (**b**) Immunoblot analysis of MOZ-TIF2 expression in CSF1R^{high} and CSF1R^{low} cell populations (sorted by flow cytometry) with a MOZ-specific antibody. MW, molecular weight; BM, bone marrow. (**c**,**d**) Leukemia-free survival after the indicated numbers of flow-sorted CSF1R^{high} (**c**) and CSF1R^{low} (**d**) cells were transplanted into sublethally irradiated mice. $n = 6$, $P = 0.0001$ (1×10^4, 1×10^3 and 1×10^2) and 0.3173 (1×10^1) (CSF1R^{high} versus CSF1R^{low} cells). (**e**) FACS analysis of Mac-1 and CSF1R expression in bone marrow cells from mice with MOZ-TIF2–induced AML. The red and blue boxes signify CSF1R^{high} and CSF1R^{low} cell fractions, respectively. (**f**–**h**) CSF1R^{high} and CSF1R^{low} cells were sorted and analyzed for morphology by staining with May-Giemsa (**f**), colony-forming activity in methylcellulose medium (**g**) and levels of total and phosphorylated STAT5, phosphorylated ERK and PU.1 (**h**). Scale bars represent 10 μm in **f**. The error bars represent s.d. in **g**. (**i**) FACS analysis of CSF1R expression in bone marrow cells from an individual with AML with a t(8;16) translocation; the cells were cultured for 3 d in 10 ng ml^{-1} human M-CSF. (**j**) RT-PCR analysis of MOZ-CBP transcripts in CSF1R^{high} and CSF1R^{low} cells of the individual with t(8;16) AML. The results are representative of 25 (**a**,**e**), four (**b**), three (**c**,**d**,**f**–**h**) and two (**i**,**j**) independent experiments.

To determine whether a high level of CSF1R expression also occurs in human AML cells with *MOZ* translocations, we investigated CSF1R expression in bone marrow cells from a subject with AML harboring a t(8;16) translocation, yielding a MOZ-CREB–binding protein (CBP, encoded by the *Crebbp* gene) fusion[9]. FACS analysis indicated that both CSF1R^{high} and CSF1R^{low} cells were present among the bone marrow cells with this translocation (**Fig. 1i**). We detected *MOZ-CBP* fusion transcripts in both the CSF1R^{high} and CSF1R^{low} cell populations (**Fig. 1j**).

These results suggest that leukemia stem cells in this subtype of AML express a high amount of CSF1R, indicating that leukemia might be cured by inducing apoptosis of CSF1R^{high} cells. To test this idea, we used transgenic mice expressing a drug-inducible FK506-binding protein (FKBP)-Fas suicide gene and EGFP under the control of the *CSF1R* promoter[10] (**Fig. 2a**). The suicide gene products are inactive monomers under normal conditions but can be activated by injection of the AP20187 dimerizer, inducing apoptosis of cells expressing high amounts of CSF1R[10]. We infected c-Kit$^+$ bone marrow cells of transgenic mice with the MOZ-TIF2 retrovirus and transplanted them into lethally irradiated wild-type mice. These mice developed AML ~2 months after transplantation. In the bone marrow of these mice, we observed morphologically indistinguishable CSF1R^{high} and CSF1R^{low} cells. As expected, endogenous CSF1R expression was proportional to EGFP and FKBP-Fas expression (**Fig. 2b** and **Supplementary Fig. 3a**).

Next, we transplanted the bone marrow cells of these AML mice (1×10^5 cells per mouse) into secondary sublethally irradiated recipient mice. Seven days after transplantation, we injected the mice with

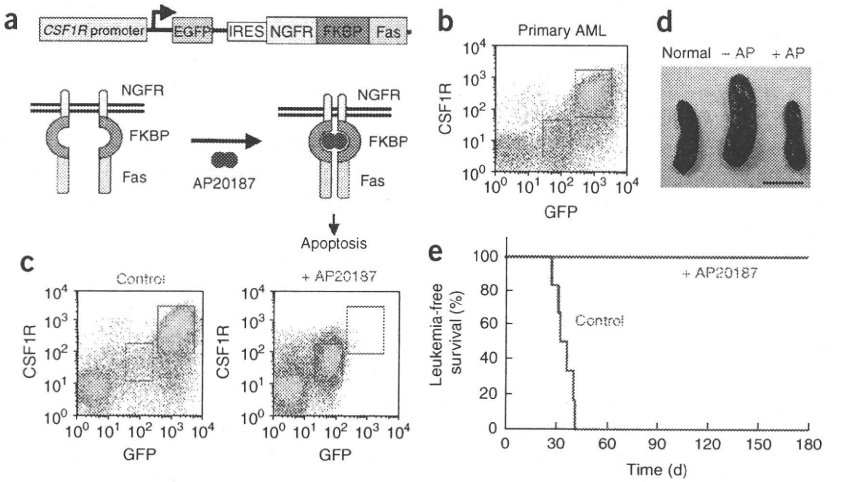


Figure 2 Cure of AML by ablation of CSF1R^{high} cells. (**a**) Top, structure of the *CSF1R* promoter-EGFP-NGFR-FKBP-Fas suicide construct. Bottom, schematic showing the activation of the NGFR-FKBP-Fas fusion protein: in transgenic mice carrying this suicide construct, ablation of cells expressing high levels of CSF1R can be induced by exposure to the AP20187 dimerizer. (**b**) FACS analysis of GFP and CSF1R expression in bone marrow cells of mice with AML 2 months after the transplantation of MSCV-MOZ-TIF2-IRES-GFP–transfected bone marrow cells derived from transgenic mice into lethally irradiated C57BL/6 mice. The red boxes signify CSF1R^{high} and CSF1R^{low} cell fractions. (**c**–**e**) Bone marrow cells (1×10^5) of primary transplanted mice with AML, generated as in **b**, were transplanted into sublethally irradiated C57BL/6 mice. Administration of AP20187 or solvent (control) to the secondary transplanted mice was started by intravenous injection 3 weeks after transplantation. Expression of GFP and CSF1R in bone marrow cells (**c**) and spleen sizes (**d**) were analyzed 4 weeks after transplantation. Scale bars, 1 cm. (**e**) Leukemia-free survival of the untreated ($n = 6$) and AP20187-treated ($n = 6$) secondary transplanted mice. $P < 0.0001$. The results are representative of five (**b**), four (**c**) and three (**d**,**e**) independent experiments.

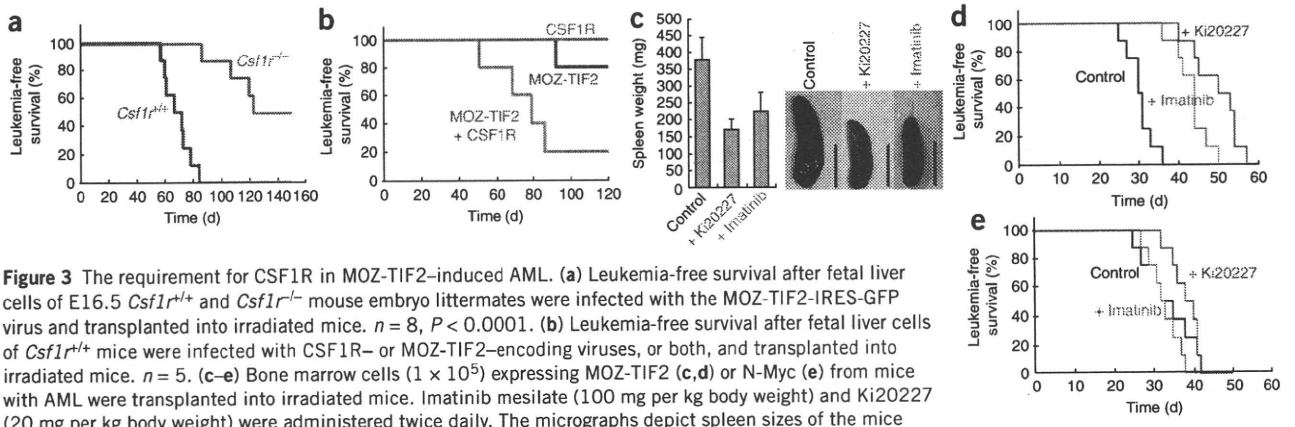


Figure 3 The requirement for CSF1R in MOZ-TIF2–induced AML. (**a**) Leukemia-free survival after fetal liver cells of E16.5 *Csf1r*$^{+/+}$ and *Csf1r*$^{-/-}$ mouse embryo littermates were infected with the MOZ-TIF2-IRES-GFP virus and transplanted into irradiated mice. $n = 8$, $P < 0.0001$. (**b**) Leukemia-free survival after fetal liver cells of *Csf1r*$^{+/+}$ mice were infected with CSF1R– or MOZ-TIF2–encoding viruses, or both, and transplanted into irradiated mice. $n = 5$. (**c–e**) Bone marrow cells (1×10^5) expressing MOZ-TIF2 (**c,d**) or N-Myc (**e**) from mice with AML were transplanted into irradiated mice. Imatinib mesilate (100 mg per kg body weight) and Ki20227 (20 mg per kg body weight) were administered twice daily. The micrographs depict spleen sizes of the mice transplanted with MOZ-TIF2–expressing cells, analyzed three weeks after transplantation (**c**). Scale bars, 1 cm. (**d,e**) Leukemia-free survival of the control and drug-treated mice was analyzed. In **d**, $n = 8$, $P < 0.0001$ (control versus + Ki20227 and control versus + imatinib). In **e**, $n = 8$, $P = 0.3825$ (control v.s. + Ki20227) and 0.4051 (control versus + imatinib).

AP20187 or a control solvent, as previously described[10]. We observed an increase in the number of CSF1R^{high} cells (**Fig. 2c**) and splenomegaly (**Fig. 2d**) in the control-treated mice 3 weeks after transplantation. However, we detected neither CSF1R^{high} cells nor splenomegaly in the AP20187-treated mice after a 1-week course of treatment (**Fig. 2c,d**). Although we observed CSF1R^{low} cells in the bone marrow and peripheral blood after the 1-week treatment course, we did not detect these cells after three months of treatment (**Fig. 2c** and **Supplementary Fig. 3b**). All control-treated mice developed AML 4–6 weeks after transplantation, but none of the AP20187-treated mice died of AML within 6 months of transplantation (**Fig. 2e**). These results indicate that ablation of the CSF1R^{high} cells was sufficient to cure MOZ-TIF2–induced AML, and that a high level of CSF1R expression is a key contributor to leukemia stem cell potential.

As it has been reported that N-Myc overexpression rapidly causes AML in mice[11], we next tested the specificity of the requirement for CSF1R^{high} cells in AML progression. We transfected the bone marrow cells of suicide gene–expressing transgenic mice with a retrovirus encoding N-Myc and EGFP, and transplanted the cells into lethally irradiated recipient mice, which developed AML. In these mice, GFP$^+$ leukemia cells were Mac1$^+$Gr1$^+$CSF1R$^-$ blast cells (**Supplementary Fig. 4a,b**), and treatment with AP20187 did not affect AML induction (**Supplementary Fig. 4c**). These results indicate a specific role of CSFR expression in MOZ-TIF2–induced AML.

To investigate the role of CSF1R in the development of MOZ-TIF2–induced AML, we infected wild-type and *Csf1r*$^{-/-}$ (ref. 12) mouse fetal liver cells of embryonic day 16.5 (E16.5) littermate embryos with the MOZ-TIF2 virus and transplanted them into lethally irradiated mice. All mice transplanted with wild-type cells developed AML within 3 months (**Fig. 3a**). In contrast, AML induction was initially suppressed in mice transplanted with *Csf1r*$^{-/-}$ cells, but half of the mice developed AML after a longer latency period (**Fig. 3a**). The suppression of AML was rescued by co-infection with the retrovirus encoding CSF1R (**Fig. 3b**). STAT5, which was highly phosphorylated in CSF1R^{high} cells but not in CSF1R^{low} cells (**Fig. 1h**), was phosphorylated in the bone marrow of recipient mice transplanted with *Csf1r*$^{+/+}$ cells but not with *Csf1r*$^{-/-}$ cells (**Supplementary Fig. 5**). To test the specificity of the requirement of CSF1R for AML induction by MOZ-TIF2, we transfected *Csf1r*$^{+/+}$ and *Csf1r*$^{-/-}$ fetal liver cells with the retrovirus encoding N-Myc and transplanted them into irradiated recipient mice. All of the mice transplanted with either *Csf1r*$^{+/+}$ or *Csf1r*$^{-/-}$ cells expressing N-Myc developed AML (**Supplementary Fig. 4d**). These results indicate that CSF1R has a key role in AML induction by MOZ-TIF2, but not by N-Myc.

The above results suggest that signaling through CSF1R might be a therapeutic target for kinase inhibitors in leukemogenesis induced by MOZ fusions. To test this, we used the CSF1R-specific inhibitor Ki20227 (ref. 13) and the tyrosine kinase inhibitor imatinib mesylate (STI571), which inhibits CSF1R[14–16]. Oral administration of Ki20227 or imatinib inhibited MOZ-TIF2–induced splenomegaly (**Fig. 3c**) and slowed MOZ-TIF2–induced AML onset (**Fig. 3d**). However, the drugs did not affect the progress of N-Myc–induced AML (**Fig. 3e**).

Next, we investigated the molecular mechanism of CSF1R expression in the leukemia cells. Monocyte-specific expression of CSF1R is reportedly regulated by transcription factors such as AML1, PU.1 and CCAAT/enhancer-binding proteins (C/EBPs)[17]. We previously found that MOZ interacts with AML1 and PU.1, but not with C/EBPα or C/EBPε, to stimulate transcription of their target genes[5,18]. Deletion analysis indicated that PU.1 interacted with the N-terminal and central regions of MOZ (**Fig. 4a** and **Supplementary Fig. 6**), and that the acidic amino acid–rich region (DE region) of PU.1 was required for its high-affinity interaction with MOZ (**Fig. 4a** and **Supplementary Fig. 7a–d**). Although binding of PU.1 to N-terminal MOZ (amino acids 1–513) was inhibited by several deletions in the PU.1 protein (**Supplementary Fig. 7c**), binding to full-length MOZ was not completely inhibited by these deletions (**Supplementary Fig. 7b**), suggesting that there may be other PU.1-binding sites in MOZ, its associated proteins or both. A pull-down assay with *Escherichia coli*–produced GST-PU.1 or GST-AML1 and *in vitro*–produced N-terminal MOZ indicated a direct interaction between both PU.1 and MOZ and between AML1 and MOZ (**Supplementary Fig. 8**). However, we cannot rule out a possibility that other factors may facilitate interactions between PU.1 or AML1 and MOZ *in vivo*.

To investigate transcriptional regulation of CSF1R, we performed reporter analysis with a *CSF1R* promoter–luciferase construct and found that MOZ, MOZ-TIF2 and MOZ-CBP could all activate the *CSF1R* promoter in the presence of PU.1 but not in the presence of AML1 (**Fig. 4b**). Moreover, MOZ, MOZ-TIF2 and MOZ-CBP did not activate a *CSF1R* promoter mutant lacking PU.1-binding sites (**Fig. 4c**). These results suggest that MOZ and MOZ fusion

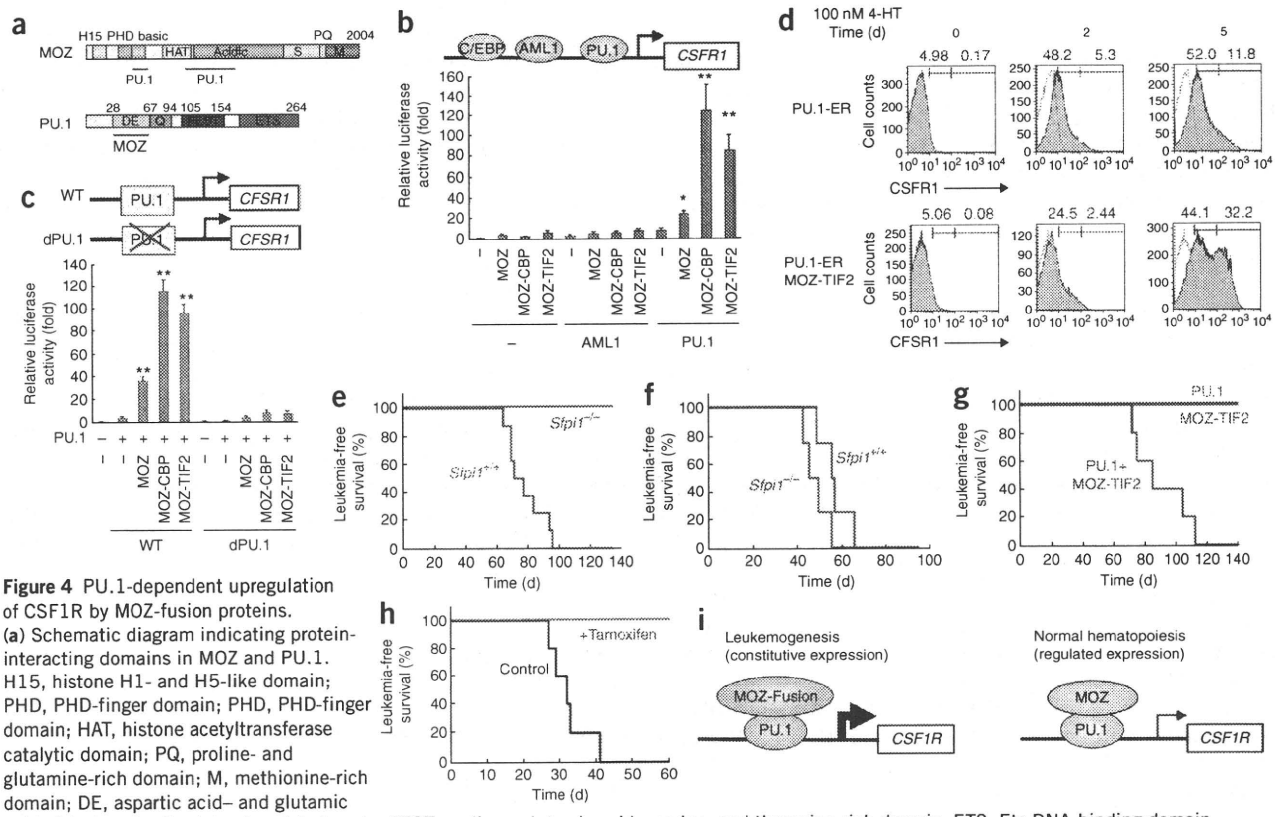


Figure 4 PU.1-dependent upregulation of CSF1R by MOZ-fusion proteins. (**a**) Schematic diagram indicating protein-interacting domains in MOZ and PU.1. H15, histone H1- and H5-like domain; PHD, PHD-finger domain; PHD, PHD-finger domain; HAT, histone acetyltransferase catalytic domain; PQ, proline- and glutamine-rich domain; M, methionine-rich domain; DE, aspartic acid– and glutamic acid–rich domain; Q, glutamine-rich domain; PEST, proline-, glutamic acid–, serine- and threonine-rich domain; ETS, Ets DNA-binding domain. (**b**) Effects of MOZ, MOZ-CBP and MOZ-TIF2 on AML1- and PU.1-mediated transcription of the *CSF1R* promoter. Osteocarcinoma SaOS2 cells were transfected with the CSF1R-luciferase construct and the indicated effector constructs encoding AML1 or PU.1 together with MOZ, MOZ-CBP or MOZ-TIF2. Luciferase activity was analyzed 24 h after transfection. Error bars represent s.d. $*P < 0.01$ and $**P < 0.005$ (comparison to PU.1 only). The results are representative of six independent experiments in which three samples were tested for each group in each experiment. (**c**) PU.1-dependent activation of CSF1R promoter. SaOS2 cells were transfected with the wild-type (WT) *CSF1R*-luciferase construct or its mutant lacking the PU.1-binding site (dPU.1), together with the indicated effectors. Error bars represent s.d. $*P < 0.01$ and $**P < 0.005$ (comparison to PU.1 only). The results are representative of three independent experiments in which three samples were tested for each group in each experiment. (**d**) FACS analysis of CSF1R expression in PUER cells infected with MSCV-GFP (top) or MSCV-MOZ-TIF2-IRES-GFP (bottom) retroviruses and exposed to 100 nM 4-HT for 0, 2 or 5 d. Population (%) of $CSF1R^{high}$ and $CSF1R^{low}$ cells were indicated. The results are representative of three independent experiments. The horizontal lines and the numbers above the graphs indicate $CSF1R^{high}$ (right) and $CSF1R^{low}$ (left) cell fractions and their populations (%), respectively. (**e**,**f**) Leukemia-free survival after fetal liver cells of E12.5 $Sfpi1^{+/+}$ and $Sfpi1^{-/-}$ mouse embryo littermates were infected with either MOZ-TIF2– (**e**) or N-Myc– (**f**) encoding viruses and transplanted into irradiated mice. (**g**) Leukemia-free survival after fetal liver cells of $Sfpi1^{-/-}$ mice were infected with PU.1– or MOZ-TIF2–encoding viruses, or both, and transplanted into irradiated mice. In **e**, $n = 8$, $P < 0.0001$; in **f**, $n = 4$, $P = 0.0943$; in **g**, $n = 5$, $P = 0.0001$ (PU.1 + MOZ-TIF2 versus either PU.1 or MOZ-TIF2). (**h**) Fetal liver cells of E14.5 $Sfpi1^{flox/flox}$ ER-Cre mice were infected with the MOZ-TIF2–encoding virus and transplanted into irradiated mice, which developed AML. The bone marrow cells of these mice were then transplanted into sublethally irradiated wild-type mice. Tamoxifen or solvent (control) was administered to the secondary transplanted mice every 2 d by intravenous injection starting 17 d after transplantation, when GFP^+ cells were detected in peripheral blood. Leukemia-free survival of the secondary transplanted mice is shown. $n = 5$, $P = 0.0018$. (**i**) Model for transcriptional regulation by normal and fusion MOZ proteins. MOZ fusion proteins stimulate constitutive CSF1R expression to induce leukemia (left). Normal MOZ protein controls CSF1R expression by binding to PU.1 to regulate normal hematopoiesis (right).

proteins activate *CSF1R* transcription in a PU.1-dependent manner. It was recently reported that although chromatin reorganization of *Csf1r* requires prior PU.1 expression together with AML1 binding, stable transcription factor complexes and active chromatin can be maintained at the *Csf1r* locus without AML1 once the full hematopoietic program has been established[19]. This might explain why we found that AML1 was not required for MOZ-TIF2–mediated activation of *Csf1r*. Deletion analysis indicated that the DE-rich, Q-rich and ETS DNA-binding domains of PU.1, as well as the histone H1 and H5-like (H15) and the central PU.1-binding domains of MOZ and MOZ fusion proteins, are required for the activation of *CSF1R* transcription (**Supplementary Figs. 7e** and **9**). A truncated version of MOZ (1–1518) lacking the C-terminal region failed to activate transcription, indicating that the transcriptional activity of MOZ-TIF2 and MOZ-CBP, which do not contain that C-terminal region, requires the TIF2 or CBP portion of the fusion protein.

To test the requirement of PU.1 for the expression of endogenous *CSF1R*, we used PU.1-deficient ($Sfpi1^{-/-}$) myeloid progenitors expressing the PU.1–estrogen receptor fusion protein (PUER). Upon restoration of PU.1 activity by exposure to 4-hydroxytamoxifen (4-HT), PUER cells can differentiate into macrophages[20]. We infected PUER cells with the MOZ-TIF2 retrovirus or control retrovirus, sorted them for GFP expression and cultured the GFP^+ cells in the presence of 4-HT. The results of FACS (**Fig. 4d**) and quantitative RT-PCR (**Supplementary Fig. 10**) analyses indicated that CSF1R expression was induced after exposure to 4-HT, and that MOZ-TIF2 enhanced