診断等を含めたアドバイスが行われるようになった．後天性再生不良性貧血と診断されていた症例からFAなどの先天性再生不良性貧血を見いだすことは，移植を含めた治療法が異なるため極めて重要である．しかしながら，先天性骨髄不全症候群の責任遺伝子は約半数で変異が確認されているにすぎず，新規遺伝子の発見に向けて研究が進められており，その成果に期待したい．

謝　辞

FA遺伝子解析を行っていただきました京都大学放射線生物研究センター 高田穣先生，群馬大学生体調節研究所 山下孝之先生，日本大学医学部小児科 谷ヶ崎博先生，神奈川こども医療センター 浜之上聡先生に，また症例の紹介をいただきました小児科・血液内科の諸先生に深謝いたします．

文　献

1) Fanconi G: Familiare infantile perniziosaartige anamie (pernizioses blutbild und konstitution). Jahrbuch Kinderheik, 117: 257–280, 1927.
2) Schroeder TM, Anchutz F, Knopp A: Spontane chromosomenaberrationen bei familiarer panmyelopathie. Humangenetik, 1: 194–196, 1964.
3) Sasaki MS, Tonomura A: A high susceptibility of Fanconi's anemia to chromosome breakage by DNA cross-linking agents. Cancer Res, 33: 1829–1836, 1973.
4) 小原　明：日本における小児特発性再生不良性貧血など造血障害性疾患の現状．日本小児血液学会再生不良性貧血委員会疫学調査　1988?2005年．日小血会誌，22: 53–62, 2008.
5) 山下孝之，小田　司，関本隆志：Fanconi貧血の分子病態—最近の進歩．臨床血液，50: 538–546, 2009.
6) Shimamura A, Montes de Oca R, Svenson JL, et al: A novel diagnostic screen for defects in the Fanconi anemia pathway. Blood, 100: 4649–4654, 2002.
7) Yabe M, Yabe H, Hamanoue S, et al: In vitro effect of fludarabine, cyclophosphamide, and cytosine arabinoside on chromosome breakagae in Fanconi anemia patients: Relevance to stem cell transplantation. Int J Hematol, 85: 354–361, 2007.
8) Soulier J, Leblanc T, Larghero J, et al: Detection of somatic mosaicism of Fanconi anemia patient by analysis of the FA/BRCA pathway. Blood, 105: 1329–1336, 2005.
9) Pinto FO, Leblanc T, Chamousset D, et al: Diagnosis of Fanconi anemia in patients with bone marrow failure. Haematologica, 94: 487–495, 2009.
10) Mankad A, Taniguchi T, Cox B, et al: Natural gene therapy in monozygotic twins with Fanconi anemia. Blood, 107: 3084–3090, 2006.
11) Gregory JJ, Wagner JE, Verlander PC, et al: Somatic mosaicism in Fanconi anemia: Evidence of genotypic reversion in lymphohematopoietic stem cell. Proc Natl Acad Sci USA, 98: 2532–2537, 2001.
12) Hamanoue S, Yagasaki H, Tsuruta T, et al: Myeloid lineage-selective growth of revertant cells in Fanconi anemia. Br J Haematol, 132: 630–636, 2005.
13) Socie G, Devergie A, Girinski T, et al: Transplantation for Fanconi's anaemia: long-term follow-up of fifty patients transplanted from a sibling donor after low-dose cyclophosphamide and thoraco-abdominal irradiation for conditioning. Br J Haematol, 193: 249–255, 1998.
14) Guardiola P, Pasquini R, Dokal I, et al: Outcome of 69 allogeneic stem cell transplantations for Fanconi anemia using HLA-matched unrelated donors: a study on behalf of the European Group for Blood and Marrow Transplantation. Blood, 95: 422–429, 2000.
15) Locatelli F, Zecca M, Pession A, et al: The outcome of children with Fanconi anemia given hematopoietic stem cell transplantation and the influence of fludarabine in the conditioning regimen: a report from the Italian pediatric group. Haematologica, 92: 1381–1388, 2007.
16) Stepensky P, Shapira MY, Balashov D, et al: Bone marrow transplantation for Fanconi anemia using fludarabine-based conditioning. Biol Blood Marrow Transplant, 17: 1282–1288, 2011.
17) Yabe M, Shimizu T, Morimoto T, et al: Matched sibling donor stem cell transplantation for Fanconi anemia patients with T-cell somatic mosaicism. Pediatric Transplant, 16: 340–345, 2012.
18) Yabe H, Inoue H, Matsumoto M, et al: Allogeneic haematopoietic cell transplantation from alternative donors with a conditioning regimen of low dose irradiation, fludarabine and cyclophosphamide in Fanconi anemia. Br J Haematol, 134: 208–212, 2006.
19) 矢部みはる，高橋義行，稲垣二郎，他：TRUMP登録されたFanconi貧血に対する造血細胞移植の検討．第34回日本造血細胞移植学会総会，2012（抄録）
20) Chaudhury S, Auerbach AD, Kernan NA, et al: Fludarabine-based cytoreductive regimen and T-cell-depleted grafts from alternative donors for the treatment of high-risk patients with Fanconi anemia. Br J Haematol, 140: 644–655, 2008.
21) Waisfisz Q, Morgan NV, Savino M, et al: Spontaneous functional correction of homozygous Fanconi anemia alleles reveals novel mechanistic basis for reverse mosaicism. Nat Genettgenome Res, 22: 379–383, 1999.
22) Gross M, Hanenberg H, Lobitz S, et al: Reverse mosaicism I Fanconi anemia: natural gene therapy via molecular self-correction. Cytogenet Genome Res, 98: 126–135, 2002.
23) Lo Ten Foe JR, Kwee ML, Rooimans MA, et al: Somatic mosaicism in Fanconi anemia: molecular basis and clinical significance. Eur J Hum Genet, 5: 137–148, 1997.
24) Ikeda H, Matsushita M, Waisfisz Q, et al: genetic reversion in an acute myelogenous leukemia cell line from a Fanconi anemia patient with biallelic mutation in BRCA2. Canser Res, 63: 2688–2694, 2003.

INTERNAL MEDICINE

□ CASE REPORT □

A Case of Congenital Dyserythropoietic Anemia Type 1 in a Japanese Adult with a *CDAN1* Gene Mutation and an Inappropriately Low Serum Hepcidin-25 Level

Hiroshi Kawabata[1], Sayoko Doisaki[2], Akio Okamoto[3], Tatsuki Uchiyama[1], Soichiro Sakamoto[1], Asahito Hama[2], Kiminori Hosoda[4], Junji Fujikura[5], Hitoshi Kanno[6], Hisaichi Fujii[6], Naohisa Tomosugi[7], Kazuwa Nakao[5], Seiji Kojima[2] and Akifumi Takaori-Kondo[1]

Abstract

We describe the first case of genetically diagnosed congenital dyserythropoietic anemia (CDA) type 1 in a Japanese man. The patient had hemolytic anemia since he was a child, and he developed diabetes, hypogonadism, and liver dysfunction in his thirties, presumably from systemic iron overload. When he was 48 years old a diagnosis was finally made by genetic analysis that revealed a homozygous mutation of *CDAN1* gene (Pro1129Leu). His serum hepcidin-25 level was inappropriately low. We conclude that physicians should be aware of the possibility of CDA in a patient with anemia and systemic iron overload at any age.

Key words: congenital dyserythropoietic anemia, iron metabolism, hemochromatosis, hepcidin, growth differentiation factor-15

(Intern Med 51: 917-920, 2012)
(DOI: 10.2169/internalmedicine.51.6978)

Introduction

Congenital dyserythropoietic anemia (CDA) is a rare congenital erythropoietic disorder with characteristic morphological abnormalities of the bone marrow cells, ineffective erythropoiesis and systemic iron overload (1). Three types of CDA are known: types 1, 2 and 3. The genes responsible for types 1 and 2 have recently been identified as *CDAN1* and *SEC23B*, respectively (2, 3). Both CDA types 1 and 2 are inherited recessively. The incidence of CDA is very rare, and in a recent pan-European survey, only 124 CDA type 1 cases were recorded (4). To date, several Japanese CDA type 1 cases have also been reported (5-7), but none of them has been genetically proven. Here, we describe in a Japanese adult a case of CDA type 1 with systemic iron overload that was genetically diagnosed in his late forties.

Case Report

A Japanese man was referred to the Kyoto University Hospital for hyperglycemia when he was 38 years old. He had had hemolytic anemia since he was a child, but its etiology had not been determined. He had undergone splenectomy when he was 36 years old, which ameliorated his anemia to some extent. At his first visit to our hospital, his white blood cell count was 6,400/μL; red blood cell count, 1.93×10^6/μL; hemoglobin (Hb) level, 7.5 g/dL; hematocrit level, 21.0%; mean corpuscular volume, 108.8 fL; platelet count, 365×10^3/μL; and reticulocyte count, 53×10^3/μL (Ta-

[1]Department of Hematology and Oncology, Graduate School of Medicine, Kyoto University, Japan, [2]Department of Pediatrics, Nagoya University Graduate School of Medicine, Japan, [3]Nantan General Hospital, Japan, [4]Faculty of Human Health Science, Kyoto University Graduate School of Medicine, Japan, [5]Department of Medicine and Clinical Science, Kyoto University Graduate School of Medicine, Japan, [6]Department of Transfusion Medicine and Cell Processing, Tokyo Women's Medical University, Japan and [7]Division of Advanced Medicine, Medical Research Institute, Kanazawa Medical University, Japan
Received for publication November 21, 2011; Accepted for publication January 5, 2012
Correspondence to Dr. Hiroshi Kawabata, hkawabat@kuhp.kyoto-u.ac.jp

Table 1. **Laboratory Data**

Features	Laboratory data at the first visit (38 years old)	Laboratory data at the time of diagnosis (48 years old)
White blood cells (per μL)	6.4×10^3	5.3×10^3
Red blood cells (per μL)	1.93×10^6	2.41×10^6
Hemoglobin (g/dL)	7.5	8.5
Hematocrit (%)	21.0	24.1
Reticulocytes (per μL)	53×10^3	—
Platelet counts (per μL)	365×10^3	312×10^3
Total bilirubin (mg/dL)	1.5	1.6
Direct bilirubin (mg/dL)	0.7	0.1
Haptoglobin (mg/dL)	<7.9	—
AST (IU/L)	49	26
ALT (IU/L)	68	16
LDH (IU/L)	302	277
Ferritin (ng/mL)	4058	186

— indicates that the tests were not performed. Abbreviations; AST, aspartate aminotransferase (reference range, 13-29 IU/L); ALT, alanine aminotransferase (reference range, 8-28 IU/L); LDH, lactate dehydrogenase (reference range, 129-241 IU/L).

ble 1). He had hepatic dysfunction, with a slightly elevated serum alanine aminotransferase level (68 IU/L), hyperglycemia (blood sugar level of 146 mg/dL and HbA1c level of 6.9%) with very low insulin secretion (serum c-peptide level, <0.1 ng/mL), and hypogonadism with a serum testosterone level lower than 0.2 ng/mL, i.e., very low (reference range, 2.7-10.7 ng/mL). His blood test results also suggested iron overload (transferrin saturation of 95.3% and serum ferritin level of 4,058 ng/mL), and the liver biopsy results revealed marked accumulation of iron in the parenchymal cells. Thus, hemochromatosis, along with liver dysfunction, diabetes and hypogonadism, was diagnosed. Insulin therapy was then started. Occasional phlebotomy was also started to remove excess iron and to gradually decrease his serum ferritin and alanine aminotransferase levels to within the reference ranges (Table 1). When he was 46 years old, a series of intensive diagnostic examinations were started. The findings of the biochemical analyses for erythrocyte membrane disorders or unstable hemoglobinopathies were all negative. The bone marrow examination revealed marked erythroid hyperplasia (the myeloid to erythroid ratio of 0.34) and remarkable dysplastic features in the erythroid cells, with megaloblastoid changes and multinuclear cells (Fig. 1A-E). However, no significant dysplasia was observed in the granulocytic or megakaryocytic series (Fig. 1A, B), and no ring sideroblasts were observed in the iron staining. When he was 48 years old, we obtained his written informed consent and approval by the ethics committee of Kyoto University to perform a genetic analysis for indicators of hereditary iron disorders in his peripheral blood cells. The results of the genetic analyses for pyruvate kinase deficiency and thalassemia syndromes were all negative. There were no mutations in the exons and the exon-intron borders of hereditary hemochromatosis genes including *HFE*, *TFR2*, *HJV*, *HAMP*, and *SLC40A1*. However, a homozygous mutation in *CDAN1 ex26 c.3503 C>T* (Pro1129Leu) was detected, consistent with CDA type 1 (Fig. 2). When we reviewed his bone marrow specimen, internuclear bridges that connected two separate erythroblasts were occasionally observed (7 bridges in 500 erythroblasts, Fig. 1F-J). His serum hepcidin-25 level was 0.8 ng/mL [reference range, 2.3-37 ng/mL; analyzed with a quantitative liquid chromatography coupled with tandem mass spectrometry method (8)]. The growth differentiation factor-15 (GDF15) level was 8,469 pg/mL (reference range, 215-835 pg/mL; analyzed with a commercial ELISA kit from R&D, Minneapolis, MN). The patient had two siblings, a brother and a sister; both were in good health. There was no significant family history except that his mother had anemia of undetermined etiology, and his paternal grandfather had diabetes. He declined genetic analysis of his family for the *CDAN1* gene.

Discussion

We encountered an adult patient with hemolytic anemia with various symptoms caused by systemic iron overload, who turned out to have a genetic mutation consistent with CDA type 1. To our knowledge, this is the first documented case of CDA type 1 in a Japanese with *CDAN1* gene mutation. Dgany et al. identified the same *CDAN1* gene mutation as the current case in a French Polynesian family (2). *CDAN1* is located on chromosome 15q15.1-15q15.3, and it codes for a nuclear protein, codanin-1, the human homolog of discs lost (dlt) which is required for cell survival and cell cycle progression in *Drosophila* (9). The diagnosis of CDA type 1 has usually been made from clinical features together with characteristic morphological features of the bone marrow cells such as binucleated erythroblasts, and internuclear bridges between the erythroid cells. As codanin-1 is essential for proper cellular trafficking of the heterochromatin

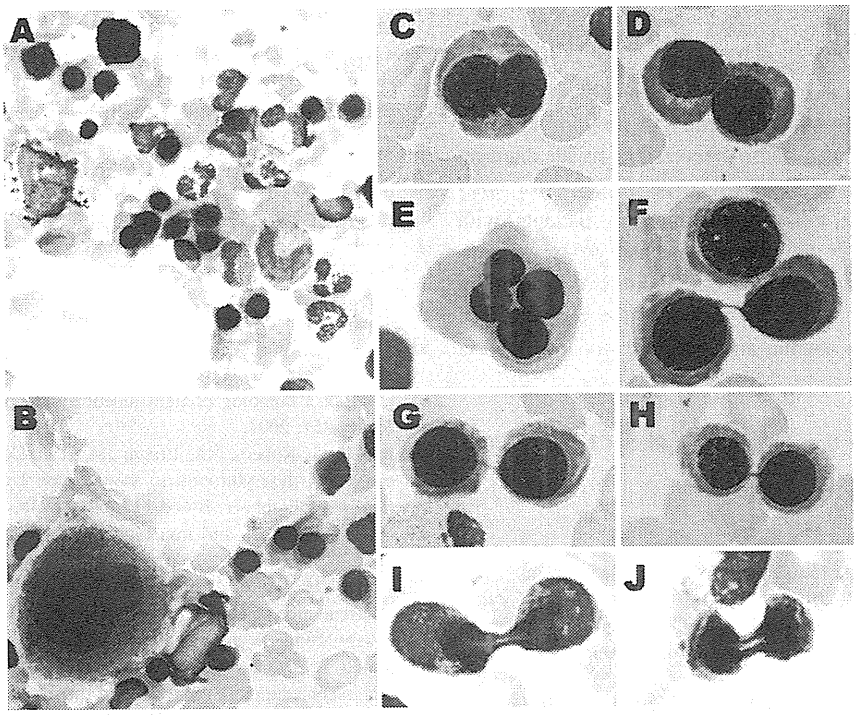

Figure 1. **Bone marrow cell morphology (May-Grünwald-Giemsa staining, original magnification ×1,000; C-J, images were further magnified by photographic enlargement). A and B: Erythroid hyperplasia. No significant dysplasia was observed in the granulocytic or megakaryocytic series. C and D: Binucleated erythroblasts. These cells were found in approximately 12% of the erythroblasts. E: A few tetranucleated erythroblasts were found. F-J: Internuclear bridges between the erythroblasts were found after careful inspection.**

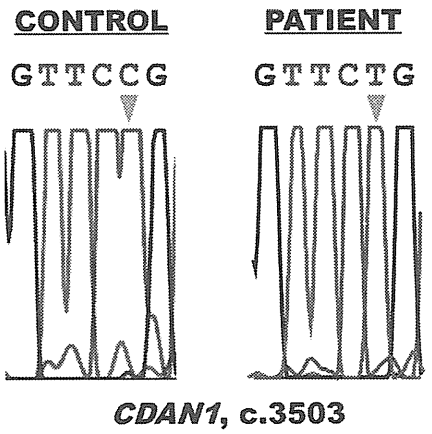

Figure 2. **The homozygous mutation in *CDAN1* (*ex26 c.3503 C>T*, Pro1129Leu) detected in the patient.**

protein HP1-α (10), defects of this protein may result in such morphological abnormalities. In the current case, the bone marrow examination results showed numerous binucleated erythroblasts, but the internuclear bridges, which are much more specific features of this disorder, were observed in less than 3% of the erythroblasts and were overlooked in the first inspection (Fig. 1). Therefore, making a definitive diagnosis of CDA from bone marrow cell morphology alone can sometimes be difficult.

CDA types 1 and 2 are known to be accompanied by iron overload. Similar to hereditary hemochromatosis, inappropriately low production of hepcidin, the central regulator of systemic iron homeostasis, has been proposed as the etiology of iron overload in CDA (11). As the main function of hepcidin is to downregulate the expression of ferroportin, the only known cellular iron exporter of mammals, downregulation of hepcidin results in an increase in ferroportin expression, thereby increasing iron absorption from the intestine and causing systemic iron overload. A previous report demonstrated marked increases of GDF15 in the serum of CDA type 1 patients (12). GDF15, a humoral factor belonging to the transforming growth factor-β superfamily, has been shown to suppress hepatic production of hepcidin (13). Consistent with the previous reports, systemic iron overload was induced in the current case without repeated red cell transfusions, the serum GDF15 level was remarkably elevated, and the serum hepcidin-25 level was inappropriately low. Thus, we postulate that serum hepcidin-25 and GDF15 are useful markers for CDA.

CDA is generally regarded as a pediatric disease because the initial symptoms, such as anemia, jaundice, and splenomegaly, usually appear in the first decade. However, the current case was diagnosed when the patient was in his late forties, and in the pan-European survey, CDA was diag-

nosed in a substantial proportion of patients who were middle-aged or older (4). Early diagnosis of CDA is important because iron chelation therapy (or phlebotomy if anemia is mild) should be started as early as possible to avoid iron overload, which can cause irreversible tissue damage. In addition, interferon-α is known to be effective for ameliorating anemia and iron accumulation in patients with CDA type 1, although the precise mechanism is still unknown (14). The survey data and our findings of the current case suggest that we should be aware of the possibility of CDA in patients with anemia and systemic iron overload at any age.

The authors state that they have no Conflict of Interest (COI).

References

1. Kamiya T, Manabe A. Congenital dyserythropoietic anemia. Int J Hematol **92**: 432-438, 2010.
2. Dgany O, Avidan N, Delaunay J, et al. Congenital dyserythropoietic anemia type I is caused by mutations in codanin-1. Am J Hum Genet **71**: 1467-1474, 2002.
3. Schwarz K, Iolascon A, Verissimo F, et al. Mutations affecting the secretory COPII coat component SEC23B cause congenital dyserythropoietic anemia type II. Nat Genet **41**: 936-940, 2009.
4. Heimpel H, Matuschek A, Ahmed M, et al. Frequency of congenital dyserythropoietic anemias in Europe. Eur J Haematol **85**: 20-25, 2010.
5. Kuribayashi T, Uchida S, Kuroume T, Umegae S, Omine M, Maekawa T. Congenital dyserythropoietic anemia type I: report of a pair of siblings in Japan. Blut **39**: 201-209, 1979.
6. Hiraoka A, Kanayama Y, Yonezawa T, Kitani T, Tarui S, Hashimoto PH. Congenital dyserythropoietic anemia type I: a freeze-fracture and thin section electron microscopic study. Blut **46**: 329-338, 1983.
7. Kato K, Sugitani M, Kawataki M, et al. Congenital dyserythropoietic anemia type 1 with fetal onset of severe anemia. J Pediatr Hematol Oncol **23**: 63-66, 2001.
8. Kanda J, Mizumoto C, Kawabata H, et al. Serum hepcidin level and erythropoietic activity after hematopoietic stem cell transplantation. Haematologica **93**: 1550-1554, 2008.
9. Pielage J, Stork T, Bunse I, Klämbt C. The Drosophila cell survival gene discs lost encodes a cytoplasmic Codanin-1-like protein, not a homolog of tight junction PDZ protein Patj. Dev Cell **5**: 841-851, 2003.
10. Renella R, Roberts NA, Brown JM, et al. Codanin-1 mutations in congenital dyserythropoietic anemia type 1 affect HP1α localization in erythroblasts. Blood **117**: 6928-6938, 2011.
11. Ganz T. Hepcidin and iron regulation, 10 years later. Blood **117**: 4425-4433, 2011.
12. Tamary H, Shalev H, Perez-Avraham G, et al. Elevated growth differentiation factor 15 expression in patients with congenital dyserythropoietic anemia type I. Blood **112**: 5241-5244, 2008.
13. Tanno T, Bhanu NV, Oneal PA, et al. High levels of GDF15 in thalassemia suppress expression of the iron regulatory protein hepcidin. Nat Med **13**: 1096-1101, 2007.
14. Lavabre-Bertrand T, Blanc P, Navarro R, et al. alpha-Interferon therapy for congenital dyserythropoiesis type I. Br J Haematol **89**: 929-932, 1995.

nature genetics

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

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

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

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

RESULTS

Genomic landscape of Down syndrome–related myeloid neoplasms

We performed whole-genome sequencing of 4 trios consisting of samples from TAM, AMKL and complete remission phases (**Supplementary Figs. 1** and **2** and **Supplementary Table 1**). In total,

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

Received 3 May; accepted 19 August; published online 22 September 2013; doi:10.1038/ng.2759

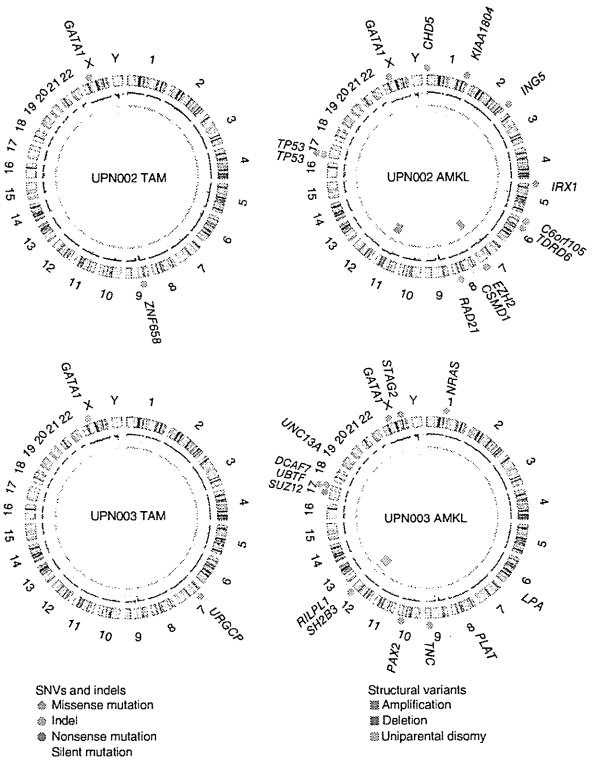

Figure 1 Representative Circos plots of paired TAM and DS-AMKL cases. Locations of somatic mutations, including of missense, frameshift, nonsense and silent mutations (colored circles), are indicated. Total (black) and allele-specific (red and green for alleles showing relatively larger and smaller copy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

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

Exome sequencing

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

Spectrum of recurrent mutations in DS-AMKL

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

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

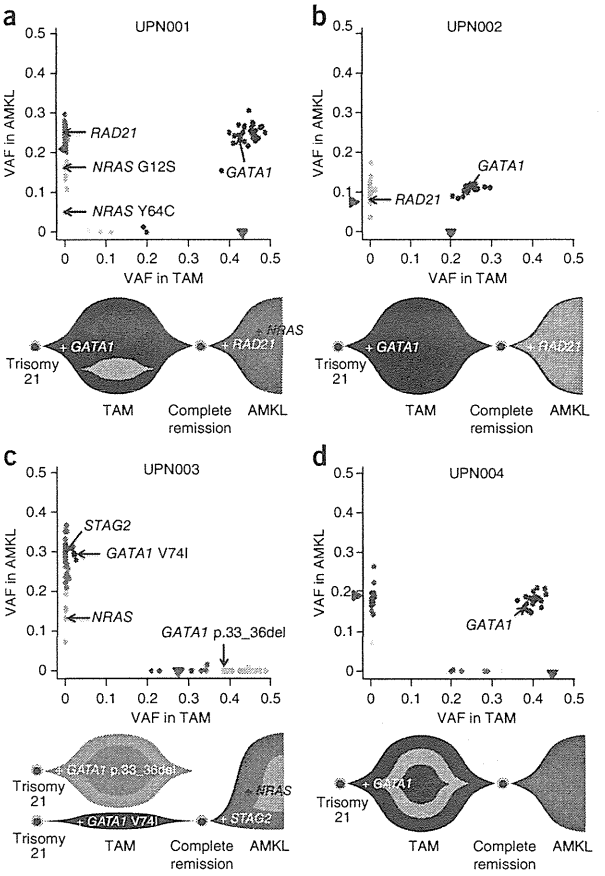

Figure 2 Clonal evolution of Down syndrome–related myeloid disorders. (**a**–**d**) Observed VAFs of validated mutations listed in **Supplementary Table 2** in both TAM and AMKL phases are shown in diagonal plots (top) for UPN001 (**a**), UPN002 (**b**), UPN003 (**c**) and UPN004 (**d**), where VAFs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

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

Mutations of cohesin and associated molecules

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

CTCF mutations

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

Mutations in epigenetic regulators

EZH2, which encodes a catalytic subunit of the Polycomb repressive complex 2 (PRC2) that is responsible for di- and trimethylation of histone H3 lysine 27 (H3K27)[39], is another recurrent mutational target in DS-AMKL (**Table 1**). Inactivating mutations in *EZH2* have

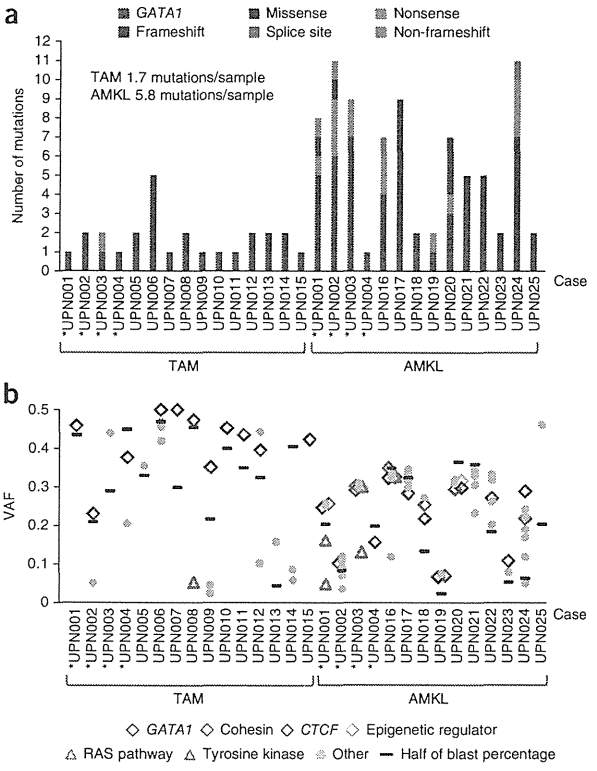

Figure 3 Somatic mutations detected by whole-exome sequencing of Down syndrome–related myeloid disorders. (**a**) Number of validated somatic mutations in 25 individuals with TAM and DS-AMKL identified by whole-exome sequencing. Paired samples are indicated by asterisks. The mutation rates per phase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

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

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

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

Gene	Mutation type	RefSeq	Amino acid change	Nucleotide change	Sample (UPN) number
CTCF	Splice site	NM_006565	p.Gly318_splice	c.953-2A>G	016
CTCF	Frameshift	NM_006565	p.Asn314fs	c.940_941insAC	020
DCAF7	Missense	NM_005828	p.Leu340Phe	c.1018C>T	001
DCAF7	Missense	NM_005828	p.Leu340Phe	c.1018C>T	003
EZH2	Frameshift	NM_004456	p.710_716del	c.2129_2148delATCACAGGATAGGTATTTTT	001
EZH2	Missense	NM_004456	p.Arg25Gln	c.74G>A	002
KANSL1	Frameshift	NM_001193466	p.Arg720fs	c.2159_2160insCG	020
KANSL1	Nonsense	NM_001193466	p.Arg462*	c.1384C>T	024
NRAS	Missense	NM_002524	p.Gly12Ser	c.34G>A	001
NRAS	Missense	NM_002524	p.Tyr64Cys	c.191A>G	001
NRAS	Missense	NM_002524	p.Gly12Ala	c.35G>C	003
RAD21	Nonsense	NM_006265	p.Arg139*	c.415A>T	001
RAD21	Frameshift	NM_006265	p.374_375del	c.1120_1124delTCTTT	002
RAD21	Missense	NM_006265	p.Leu611Arg	c.1832T>G	018
RAD21	Nonsense	NM_006265	p.Arg65*	c.193C>T	024
STAG2	Nonsense	NM_001042750	p.Arg604*	c.1810C>T	003
STAG2	Nonsense	NM_001042750	p.Arg216*	c.646C>T	019
STAG2	Frameshift	NM_001042750	p.Asn863fs	c.2588_2589insT	020
TP53	Nonsense	NM_000546	p.Glu68*	c.202G>T	002
TP53	Non-frameshift	NM_000546	p.157_162del	c.469_486delGTCCGCGCCATGGCCATC	002

Other mutations in DS-AMKL

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

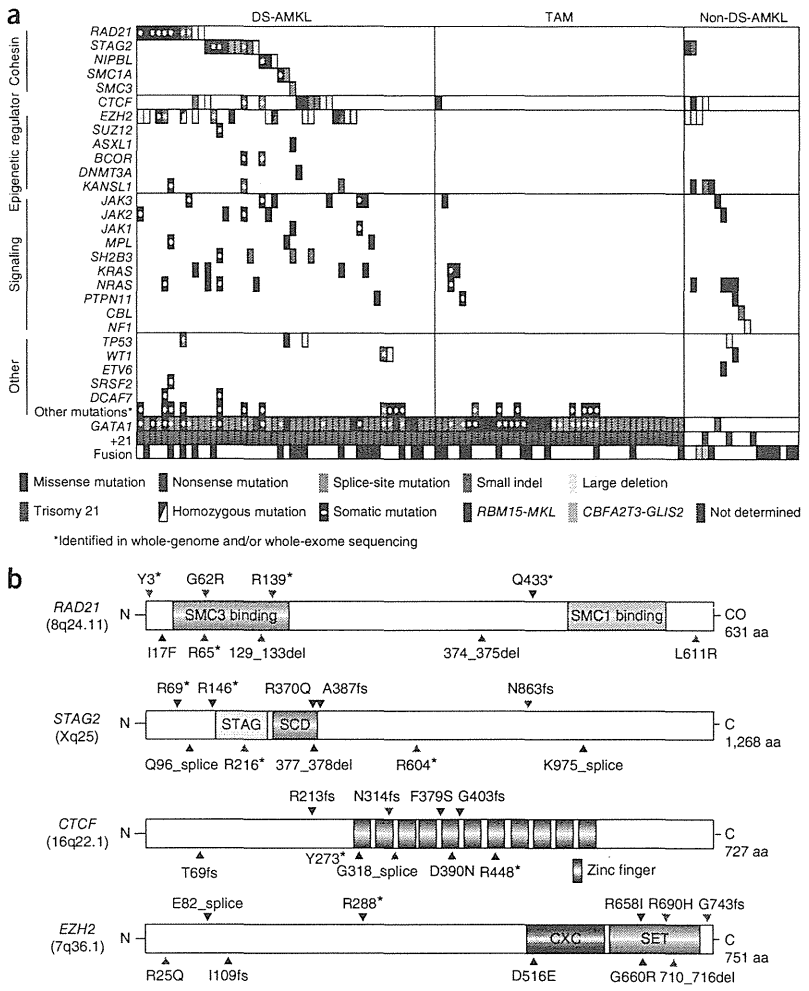

Figure 4 Driver mutations in Down syndrome–related myeloid disorders and non-DS-AMKL. (**a**) Driver mutations in 109 samples of 49 DS-AMKL, 41 TAM and 19 non-DS-AMKL cases. Types of mutations are distinguished by color. Each sample is also described in **Supplementary Table 12**. (**b**) Distribution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

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

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

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

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

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

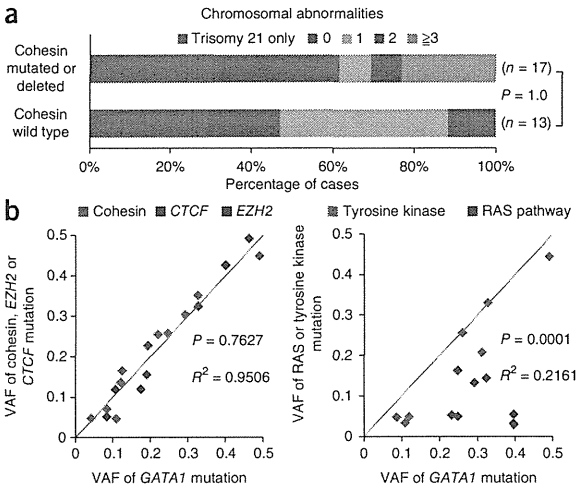

Figure 5 Relationship of cohesin mutations with karyotypes and comparison of mutation loads between major gene targets in DS-AMKL and *GATA1*. (**a**) The number of chromosomal abnormalities is compared between cases with and without cohesin mutations or deletions for DS-AMKL cases. Zero signifies chromosomal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

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

DISCUSSION

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

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

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

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

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

METHODS

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

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

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

ACKNOWLEDGMENTS

We thank Y. Mori, M. Nakamura, O. Hagiwara and N. Mizota for their technical assistance. This work was supported by the Research on Measures for Intractable Diseases Project and Health and Labor Sciences Research grants (Research on Intractable Diseases) from the Ministry of Health, Labour and Welfare, by Grants-in-Aid from the Ministry of Health, Labor and Welfare of Japan and

KAKENHI (22134006, 23249052, 23118501, 23390266 and 25461579) and by the Japan Society for the Promotion of Science (JSPS) through the Funding Program for World-Leading Innovative Research and Development on Science and Technology (FIRST Program), initiated by the Council for Science and Technology Policy (CSTP) and research grants from the Japan Science and Technology Agency CREST.

AUTHOR CONTRIBUTIONS

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

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

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

ONLINE METHODS

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

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

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

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

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

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

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

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

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

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

where k_s is determined to satisfy

$$k_0 = \sum_j D_j^s$$

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

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

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

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

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

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

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

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

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

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

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

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

54. Sato, Y. *et al.* Integrated molecular analysis of clear-cell renal cell carcinoma. *Nat. Genet.* **45**, 860–867 (2013).
55. Yoshida, K. *et al.* Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **478**, 64–69 (2011).
56. Kent, W.J. BLAT—the BLAST-like alignment tool. *Genome Res.* **12**, 656–664 (2002).
57. Shiraishi, Y. *et al.* An empirical Bayesian framework for somatic mutation detection from cancer genome sequencing data. *Nucleic Acids Res.* **41**, e89 (2013).
58. Sakaguchi, H. *et al.* Exome sequencing identifies secondary mutations of *SETBP1* and *JAK3* in juvenile myelomonocytic leukemia. *Nat. Genet.* **45**, 937–941 (2013).
59. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
60. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
61. Forbes, S.A. *et al.* COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* **39**, D945–D950 (2011).
62. Robinson, J.T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
63. Torres, L. *et al.* Acute megakaryoblastic leukemia with a four-way variant translocation originating the *RBM15-MKL1* fusion gene. *Pediatr. Blood Cancer* **56**, 846–849 (2011).
64. Nannya, Y. *et al.* A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. *Cancer Res.* **65**, 6071–6079 (2005).
65. Yamamoto, G. *et al.* Highly sensitive method for genomewide detection of allelic composition in nonpaired, primary tumor specimens by use of Affymetrix single-nucleotide-polymorphism genotyping microarrays. *Am. J. Hum. Genet.* **81**, 114–126 (2007).

Ann Hematol
DOI 10.1007/s00277-012-1564-5

ORIGINAL ARTICLE

Clinical and genetic characteristics of congenital sideroblastic anemia: comparison with myelodysplastic syndrome with ring sideroblast (MDS-RS)

Rie Ohba · Kazumichi Furuyama · Kenichi Yoshida · Tohru Fujiwara · Noriko Fukuhara · Yasushi Onishi · Atsushi Manabe · Etsuro Ito · Keiya Ozawa · Seiji Kojima · Seishi Ogawa · Hideo Harigae

Received: 30 May 2012 / Accepted: 21 August 2012

Abstract Sideroblastic anemia is characterized by anemia with the emergence of ring sideroblasts in the bone marrow. There are two forms of sideroblastic anemia, i.e., congenital sideroblastic anemia (CSA) and acquired sideroblastic anemia. In order to clarify the pathophysiology of sideroblastic anemia, a nationwide survey consisting of clinical and molecular genetic analysis was performed in Japan. As of January 31, 2012, data of 137 cases of sideroblastic anemia, including 72 cases of myelodysplastic syndrome (MDS)–refractory cytopenia with multilineage dysplasia (RCMD), 47 cases of MDS–refractory anemia with ring sideroblasts (RARS), and 18 cases of CSA, have been collected. Hemoglobin and MCV level in CSA are significantly lower than those of MDS, whereas serum iron level in CSA is significantly higher than those of MDS. Of 14 CSA for which DNA was available for genetic analysis, 10 cases were diagnosed as X-linked sideroblastic anemia due to *ALAS2* gene mutation. The mutation of *SF3B1* gene, which was frequently mutated in MDS-RS, was not detected in CSA patients. Together with the difference of clinical data, it is

Electronic supplementary material The online version of this article (doi:10.1007/s00277-012-1564-5) contains supplementary material, which is available to authorized users.

R. Ohba · T. Fujiwara · N. Fukuhara · Y. Onishi · H. Harigae (✉)
Department of Hematology and Rheumatology,
Tohoku University Graduate School of Medicine,
1-1 Seiryo-machi, Aoba-ku,
Sendai 980-8574, Japan
e-mail: harigae@med.tohoku.ac.jp

K. Furuyama
Department of Molecular Biology and Applied Physiology,
Tohoku University Graduate School of Medicine,
Sendai, Japan

K. Yoshida · S. Ogawa
Cancer Genomics Project, Graduate School of Medicine,
The University of Tokyo,
Tokyo, Japan

T. Fujiwara · H. Harigae
Department of Molecular Hematology/Oncology,
Tohoku University Graduate School of Medicine,
Sendai, Japan

A. Manabe
Department of Pediatrics, St Luke's International Hospital,
Hirosaki University Graduate School of Medicine,
Hirosaki, Japan

E. Ito
Department of Pediatrics,
Hirosaki University Graduate School of Medicine,
Hirosaki, Japan

K. Ozawa
Division of Hematology, Jichi Medical University,
Shimotsuke, Japan

S. Kojima
Department of Pediatrics,
Nagoya University Graduate School of Medicine,
Nagoya, Japan

Published online: 16 September 2012

suggested that genetic background, which is responsible for the development of CSA, is different from that of MDS-RS.

Keywords Congenital sideroblastic anemia · Myelodysplastic syndrome · ALAS2

Introduction

Sideroblastic anemia is characterized by anemia with the emergence of ring sideroblasts in the bone marrow. Ring sideroblasts are formed by the irregular accumulation of iron in mitochondria. There are two forms of sideroblastic anemia i.e., congenital sideroblastic anemia (CSA) and acquired sideroblastic anemia. Most of acquired sideroblastic anemia cases were included in myelodysplastic syndrome (MDS). To date, mutations of genes involved in heme biosynthesis, Fe–S cluster biogenesis, or the biology of mitochondria have been reported in CSA [1–5]. Impaired function of these genes is speculated to result in disutilization of iron, leading to accumulation of iron in mitochondria. Acquired sideroblastic anemia in MDS is categorized either as refractory cytopenia with multilineage dysplasia (RCMD) or refractory anemia with ring sideroblasts (RARS) depending on the level of dysplasia. In contrast CSA, mechanism of forming ring sideroblasts in MDS is not clarified, although it was recently suggested that the mutations of splicing pathway are involved in the pathogenesis of MDS [6]. It is possible that there is a common mechanism between CSA and MDS; however, mutations in genes, which are responsible for development of the CSA, have not been identified in MDS.

The most common form of CSA is X-linked sideroblastic anemia (XLSA), which is caused by mutation of erythroid-specific 5-aminolevulinate synthase (*ALAS2*), the first enzyme of heme synthesis in erythroid cells [7–10]. More than half of the patients with XLSA respond to the administration of pyridoxine [vitamin B6 (Vit.B6)], or pyridoxal 5-phosphate (PLP), which is the coenzyme of *ALAS2* [11]. In XLSA, adult onset cases have been reported [12, 13]; therefore, it is possible that some cases of CSA may be misdiagnosed as MDS, especially RARS. However, the clinical and pathological features of congenital and acquired sideroblastic anemia have not been fully clarified because there have been no comprehensive studies, including clinical and genetic analyses, focusing on sideroblastic anemia.

Here, we performed a nationwide survey of sideroblastic anemia in Japan to investigate the epidemiology and pathogenesis of this disease. The difference of clinical data and results of genetic analysis suggest that genetic background, which is responsible for the development of CSA, is distinct from that of MDS-RS.

Materials and methods

Data acquisition

This study consisted of three investigations. First, patients with sideroblastic anemia were searched by questionnaire sent to hospitals with hematology department (493 hospitals) and pediatric hematology department (593 hospitals) asking for information about patients diagnosed as sideroblastic anemia (first investigation) over the past 10 years. Next, detailed clinical data of sideroblastic anemia patients were collected from the hospital based on responses to the first investigation (second investigation). Survey items were age of onset, gender, family history, hematological and biochemical findings, treatment, and cause of death. Then, genetic analysis of patients, who were diagnosed as CSA and MDS without chromosomal anomaly, was performed in cases for which genome sample was available (third investigation).

This study was approved by the ethics committee of Tohoku University Graduate School of Medicine, the center responsible for clinical and genetic analysis. Informed consent for the genetic analysis was obtained in all cases.

Diagnostic procedure

Ring sideroblasts were defined following the 2001 World Health Organization (WHO) classification. Sideroblastic anemia patients were diagnosed in the respective institutions. In all cases, bone marrow smears were investigated, and at least 15 % ring sideroblasts were confirmed by iron staining. Furthermore, diagnosis for RARS was made when dysplasia restricted to erythroid lineage in bone marrow was recognized. Diagnosis for RCMD was made when there is multilineage dysplasia. Thereafter, in the present study, RCMD correspond to refractory cytopenia with multilineage dysplasia and ringed sideroblasts (RCMD-RS) of the 2001 WHO classification. Diagnosis for CSA was made when the patient had a family history or the disease onset during infancy, or fulfilled the characteristic features of XLSA, such as onset at a young age, microcytic anemia, and responsiveness to Vit.B6.

Genetic analysis of patients with sideroblastic anemia

In the genetic analysis, mutations in *ALAS2*, *SLC25A38*, *GLRX5*, *ABCB7*, *PUS1*, and *SLC19A2*, which are known to be responsible for CSA, were examined in 14 cases of CSA and 10 cases of MDS. In addition, *SF3B1*, which was very recently reported to be mutated in sideroblastic anemia in MDS at a high incidence, were analyzed as well. Mutation analysis for the *ALAS2* gene was performed first in all candidates, and then the analysis proceeded to the other

genes if no mutations in *ALAS2* were detected. For mutation analysis of *ALAS2*, genomic DNA was extracted from the proband's peripheral blood using QIAamp DNA blood midi kit (QIAGEN, Valencia, CA, USA). The proximal promoter region [14], erythroid enhancer in intron 8 [15], and all exons and exon–intron boundaries of the *ALAS2* gene were amplified using ExTaq DNA polymerase (Takara Bio, Shiga, Japan) [16]. Amplified products were purified using a QIAquick gel extraction kit (QIAGEN) after agarose gel electrophoresis. They were then subjected to direct sequencing analysis using BigDye Terminator Cycle sequencing kit v1.1 with an ABI3100 genetic analyzer (Life Technologies Corp., Carlsbad, CA, USA). Mutation of the gene was confirmed by repeated polymerase chain reaction (PCR) followed by direct sequencing analysis. Genes other than *ALAS2* were sequenced by Hiseq2000® [6]. Briefly, genomic DNA was amplified using REPLI-g mini kit® (QIAGEN Science). After adjusting the concentration of amplified DNA, DNA from consecutive 12 samples was combined into one DNA pool, and the entire coding sequences were amplified by primers to which *Not*I linker was attached. The products were digested with *Not*I, and ligated with T4 ligase. Then, DNA was sonicated into ~200-bp fragments, and sequencing libraries were generated. Libraries were subjected to deep sequencing on Hiseq2000®. Sequencing data was analyzes as described previously. Detected mutations were validated by direct sequence.

Analysis of enzymatic activity of recombinant ALAS2 protein

For preparing recombinant ALAS2 proteins, complementary DNA (cDNA) encoding mature human ALAS2 protein was amplified using a following primer set (5′-GGTGGTCATATGATCCACCTTAAGGCAACAAAGG-3′and 5′-GGCATAGGTGGTGACATACTG-3′). The amplified cDNA was then treated with *Nde*I restriction enzyme and was cloned between *Nde*I and blunt-ended *Sap*I site of pTXB1 plasmid (New England Biolabs, Ipswich, MA, USA), resulting in pTXB-AEm. From this plasmid, mature ALAS2 protein was expressed as an inducible fusion protein with Intein and chitin-binding domain in *E. coli*. To obtain the mutant protein, the identified mutation was introduced into pTXB-AEm using PrimeStar Max site-directed mutagenesis kit (Takara Bio, Shiga, Japan). For expression and purification of wild-type and mutant ALAS2 proteins, *E. coli* BL21 (DE3) was transformed with each plasmid. The induction and purification of the recombinant proteins were performed using Impact system (New England Biolabs) according to manufacturer's instruction. Briefly, each recombinant protein was induced in *E. coli* with 0.1 mM IPTG at 25 °C for overnight. Then, cells were resuspended with lysis buffer (20 mM Tris–HCl pH8.5, 500 mM NaCl, 1 mM EDTA, 0.1 % Triton X-100, 1 mM PMSF, 1 μg/ml of antipain, pepstatin, and leupeptin). After the sonication and centrifugation, cleared cell lysates were incubated with chitin beads for 1 h at 4 °C, then washed with wash buffer (20 mM Tris–HCl pH8.5, 500 mM NaCl, 1 mM EDTA, and 0.1 % Triton X-100). Tag-free recombinant mature ALAS2 protein was obtained by on-column cleavage with 50 mM DTT in wash buffer at room temperature for 16 h. After the elution from the column, protein concentration was determined using Bio-Rad Protein assay reagent (Bio-Rad Laboratories, Inc., Hercules, CA, USA). The ALAS activity of each recombinant protein was measured in vitro, as described previously [8].

Statistical analysis

Results are presented as means±SD with the exception of the age of onset, which is expressed as the median. Statistical analysis was performed using Student's *t* test, and $P<0.05$ was taken to indicate statistical significance.

Results

Epidemiology of sideroblastic anemia

As of 31 January 2012, detailed data for 148 sideroblastic anemia, including MDS and secondary sideroblastic anemia, patients have been collected. Excluding 10 cases of refractory anemia with excess blasts (RAEB) and one case of sideroblastic anemia due to alcohol, the remaining 137 cases were classified as 18 cases of CSA, 47 cases of RARS, and 72 cases of RCMD. Of 18 CSA cases, 7 were already confirmed to be XLSA due to mutation of *ALAS2* before registration in this study, and the others were diagnosed as CSA based on family history or clinical findings, including responsiveness to Vit.B6 treatment. Clinical findings and family history, which suggest the porphyria, were not observed in any CSA patients.

Analysis of the pathology of congenital sideroblastic anemia

Laboratory data of CSA, RARS, and RCMD are shown in Tables 1 and 2. Median age at onset of CSA was younger than those of RARS and RCMD (19, 72.5, and 71 years old, respectively). Hemoglobin and mean corpuscular volume (MCV) values of CSA were significantly lower than those of RARS and RCMD cases (7.1 g/dl and 69.0 fl, 8.7 g/dl and 106.8 fl; and 8.3 g/dl and 106.5 fl, respectively). Serum iron level in CSA was significantly higher than that in RARS or RCMD (210.7, 162.8, and 171.1 μg/dl, respectively). These data have possibilities of reflecting the states of the iron over-loaded of CSA; however, as serum iron concentration is very instable and depends from different factors, this finding should be carefully evaluated.

Table 1 Clinical data of CSA, RARS, and RCMD (1)

	CSA (n=18)	RARS (n=47)	RCMD (n=72)	p-value (between CSA and RARS)	p-value (between CSA and RCMD)
Gender					
Male	17	33	44		
Female	1	14	28		
Median age at onset (year)	19.0 (±20.2)	72.5 (±10.4)	71.0 (±13.0)	<0.01	<0.01
White blood cells (/μl)	5547 (±2022)	4741 (±2561)	4105 (±1847)	0.24	<0.01
Red blood cells ($\times 10^4$/μl)	383.4 (±100.0)	245.6 (±45.6)	239.4 (±56.4)	<0.01	<0.01
Hemoglobin (g/dl)	7.1 (±2.21)	8.7 (±1.7)	8.3 (±1.8)	<0.01	0.02
Mean corpuscular volume (fl)	69.0 (±11.6)	106.8 (±9.0)	106.5 (±9.2)	<0.01	<0.01
Platelet ($\times 10^4$/μl)	28.5 (±12.62)	25.9 (±15.5)	23.9 (±24.1)	0.53	0.44
Reticulocyte (‰)	12.1 (±10.9)	17.7 (±10.8)	21.5 (±20.1)	0.07	0.05

When iron-related laboratory data were examined in transfusion independent cases (CSA, 13; RARS, 26; RCMD, 34), Serum iron level in CSA was tended to be higher than that in RARS or RCMD (210.6, 180.3, and 166.6 μg/dl, respectively), although the difference was not significant (p=0.07, data not shown). Serum ferritin level in CSA, RARS and RCMD were elevated in these transfusion independent cases (1,087.9, 898.1, and 732.2 ng/ml, respectively), suggesting that most of sideroblastic anemia patients were iron-overloaded before transfusion. There were no significant differences in other biochemical data among the three groups.

Chromosomal abnormalities of MDS

Data regarding cytogenetic abnormalities were available for all RARS patients and for 68 of 72 RCMD patients. Figure 1 shows the cytogenetic findings of RARS and RCMD. In RARS cases, chromosomal abnormalities were found in 17 patients (36.2 %). Abnormalities consisted of abnormality including +8 (3 cases), complex abnormality with deletion 5 (2 cases), and complex abnormality with 20q− (3 cases). Chromosomal abnormalities in RCMD were found in bone marrow samples from 27 RCMD patients (39.7 %). Abnormality including +8 was detected in nine cases (33.3 %) and abnormality of idic (X) (q13), associated with the *ABCB7* gene [17], was found in one case. In addition, −7, which was not identified in RARS, was identified in four RCMD patients (14.8 %).

Treatment and outcome

Analysis of the available data regarding treatment indicated that 17 of 47 RARS cases and 26 of 72 RCMD cases were administered Vit.B6 (data not shown). The effectiveness was judged according to the criteria of IWG [18], and one RARS patient obtained a major response, and three RARS patients and one RCMD patient obtained minor responses. Thus, 4 of 17 RARS patients and 1 of 26 RCMD patients responded to Vit.B6 treatment. However, improvement of Hb was not sustained in two RARS patients; Hb level gradually returned to or dropped below the pretreatment level. Therefore, Vit.B6 treatment may not be effective for MDS, or the effect if any may be very limited. The clinical outcomes of patients are shown in Supplemental Table 1. The median follow-up from the time of diagnosis in CSA patients was 30.5 months, and two patients died due to sepsis (one case) and cardiac failure (one case). One patient

Table 2 Clinical data of CSA, RARS, and RCMD (2)

	CSA (n=18)	RARS (n=47)	RCMD (n=72)	p-value (between CSA and RARS)	p-value (between CSA and RCMD)
Total bilirubin (mg/dl)	1.1 (±0.8)	1.3 (±0.9)	1.1 (±0.7)	0.47	0.78
AST (GOT) (IU/l)	33.0 (±24.3)	24.9 (±11.7)	27.9 (±20.8)	0.08	0.38
LDH (IU/l)	218.3 (±98.9)	263.5 (±119.2)	246.1 (±97.7)	0.16	0.28
CRP (mg/dl)	0.13 (±0.15)	0.40 (±1.16)	1.17 (±3.81)	0.37	0.30
Serum iron (mg/dl)	210.7 (±77.6)	162.8 (±73.6)	171.1 (±66.2)	0.03	0.04
UIBC (mg/dl)	80.4 (±113.6)	102.4 (±82.7)	79.9 (±60.7)	0.48	0.93
Ferritin (ng/ml)	1239.8 (±1306.8)	743.4 (±815.3)	804.3 (±990.2)	0.08	0.13

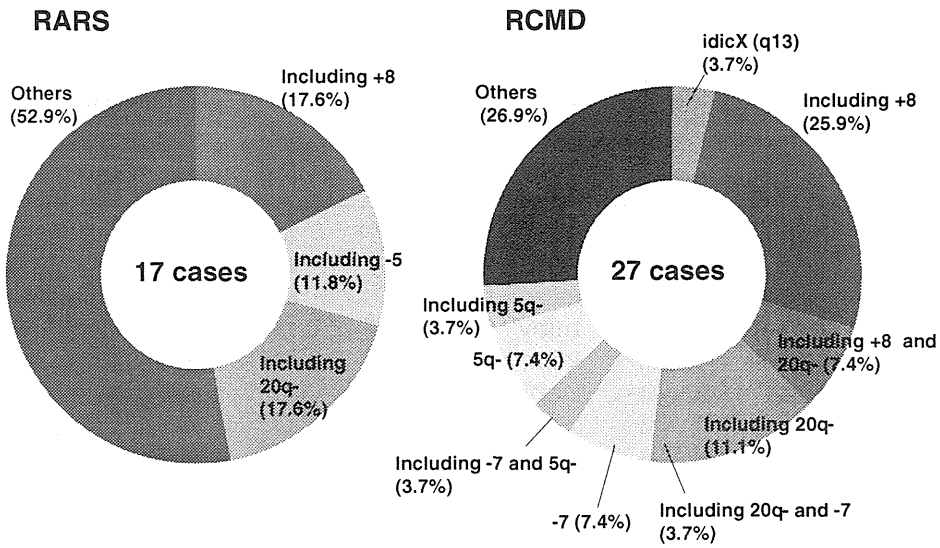

Fig. 1 Chromosomal abnormalities in RARS and RCMD. Data of chromosomal analysis in RARS and RCMD are shown. +8 was most common both in RARS and RCMD. -7 was only seen in RCMD

who died due to cardiac failure was heavily iron overloaded as defined by serum ferritin level, suggesting that cardiac complications may be caused by hemochromatosis. The median follow-up from the time of diagnosis in RARS patients was 23 months, and 6 patients (12.8 %) died due to pneumonia (two cases), evolution to leukemia (one case), and others (three cases). The median follow-up from the time of diagnosis in RCMD patients was 19.5 months, and 20 patients (27.8 %) died due to pneumonia (7 cases), cardiac failure (3 cases), evolution to leukemia (2 cases), sepsis (1 case), and others (7 cases). These results suggest that the prognosis of RCMD is worse than that of RARS.

Gene analysis of congenital sideroblastic anemia

Eighteen CSA patients were candidates for gene analysis; however, mutation analysis for genes responsible for CSA was not performed in four patients. One patient was diagnosed as having PMPS based on clinical findings, and DNA samples were not available for the remaining three patients. Therefore, gene analysis was performed in 14 of 18 CSA patients. Ten of these 14 patients were diagnosed as XLSA due to *ALAS2* mutation. Table 3 summarizes the results of gene analysis in XLSA. Case 2 (R411C), case 4 (D190V), case 6 (M567I), and case 7 (V562A) were reported previously [19–21]. Since amino acid substitution at Arg170, 411, and 452 were observed in plural patients, there are hot spots of mutation of *ALAS2* gene.

Patient with D190V (case 4), R170L (Case 10) and two patients with R452C (cases 3 and 5) did not respond to Vit.B6 treatment, whereas six patients responded to Vit.B6 treatment, although the increment of hemoglobin varied from 1.7 to 8.1 g. Interestingly, case 8 responded to Vit.B6 treatment, whereas case 10 did not, although both of them harbor the same mutation, R170L. Therefore, the activity of R170L mutant proteins was examined to determine the property, especially the Vit.B6 responsiveness. The enzymatic activities of wild type and R170L mutant protein were 7,193±138 nmol ALA/mg protein/h and 2,240±145 nmol ALA/mg protein/h, respectively, in the absence of PLP (Fig. 2). With an excess amount of PLP (100 μM) in the assay mixture, higher enzymatic activities were obtained with wild-type and mutant proteins (12,662±311 nmol ALA/mg protein/h and 7,700± 49 nmol ALA/mg protein/h, respectively) (Fig. 2). In addition, the enzymatic activity of R170C, which is another substitution at Arg170 found in this study, was also examined. As shown in Fig. 2, The enzymatic activity of mutant protein was significantly lower than wild-type protein without PLP (4,612± 87 nmol ALA/mg protein/h vs 7,193±138 nmol ALA/mg protein/h), and the activity was restored by addition of excess amount of PLP (100 μM) in the assay mixture. These in vitro data suggest that amino acid substitution at Arg 170, at least R170L and R170C, results in the decrease in enzymatic activity, but the decrease can be recovered by excess amount of PLP. The enzymatic activity of mutant proteins, which were identified in this study, is summarized in Table 3. The enzymatic activities of R411C, D190V, M567I, and V562A were referred from previous reports [19–21]. The levels of activity and PLP responsiveness in vitro were not correlated with clinical responsiveness to PLP in some cases. It is possible that the variety of mechanisms, such as the decrease in enzymatic activity of mutant ALAS2 protein, the changes of amount of *ALAS2* transcript, and physiological and environmental status of the patients, are responsible for the development of the disease.

Data for CSA patients other than XLSA are summarized in Table 4. Case 15 was diagnosed as PMPS. Gene analysis was not performed for cases 16 and 17; however, XLSA was strongly suspected because these patients were male and had microcytic anemia that was responsive to Vit.B6 treatment.

Table 3 Congenital sideroblastic anemia (XLSA)

Case number	Age at diagnosis (y.o.)	Gender	Position of *ALAS2* mutation	*SF3B1* mutation	Hb at onset (g/dl)	MCV at onset (fl)	Increment of Hb by Vit.B6 treatment (g/dl)	In vitro enzymatic activity of mutant protein[a]	
								Without PLP	With PLP
1	0	M	R170C	N/D	4.8	52.5	1.7	64.1 %	72.5 %[b]
2	20	M	R411C	N/D	4.8	52.5	5.2	11.9 %	25.1 % [19]
3	68	M	R452C	–	6.0	67.3	No effect	99.9 %	94.0 % [21]
4	17	M	D190V	N/D	8.9	66.9	No effect	98.6 %	98.5 % [20]
5	36	M	R452C	–	7.4	70.0	No effect	99.9 %	94.0 % [21]
6	36	M	M567I	N/D	6.5	64.4	3.4	38.1 %	25.2 % [21]
7	14	M	V562A	–	8.1	61.2	4.7	150.6 %	116.9 % [21]
8	31	M	R170L	–	4.1	50.8	8.1	31.1 %	60.8 %[b]
9	3	M	R411C	–	5.4	54.4	2.9	11.9 %	25.1 % [19]
10	62	M	R170L	N/D	8.0	73.9	No effect	31.1 %	60.8 %[b]

[a] % of WT

[b] Present study

ALAS2 mutations were not identified in cases 11, 12, 13, and 14. Therefore, mutations of *SLC25A38*, *GLRX5*, *ABCB7*, *PUS1*, *SLC19A2*, and *SF3B1* were examined in these cases; however, no mutations were identified in these cases. In contrast to other cases, case 18 was female and showed normocytic anemia. She was diagnosed with CSA due to family history; however, gene mutation analysis was not performed because DNA samples were not available. *SF3B1* gene mutation was examined in nine cases including five XLSA, however, no mutation was identified (Tables 3 and 4). On the other hand, *SF3B1* gene mutation was frequently detected in MDS-RS (Table 5).

Discussion

Because of its rarity, there have been few clinical and pathological investigations focusing on sideroblastic anemia. This study was performed to investigate the epidemiological and pathological characteristics of sideroblastic anemia. Based on the data of 137 patients, it was revealed that hemoglobin level in CSA was significantly lower than those seen in MDS, and serum iron level was higher in CSA compared to MDS. These results revealed that anemia in CSA is more severe than that in MDS at onset, although significant cases improved by Vit.B6 treatment. Reflecting the high incidence of XLSA in CSA, MCV level was significantly lower in CSA than MDS. These findings suggest that CSA should be strongly suspected rather than MDS, at least in Japan, in male patients exhibiting microcytic anemia and an elevated serum iron level.

MDS-RCMD is the most common form of acquired sideroblastic anemia. Chromosomal abnormalities were observed in 39.7 % of RCMD cases and 36.2 % of RARS cases. The types of chromosomal abnormality frequently observed in RCMD and RARS did not differ from those reported previously, such as +8, −7, 20q− and −5. Among them, +8 was observed in nine cases of RCMD (33.3 %). As the frequency of +8 in MDS was reported to be 6.5–16.7 %,

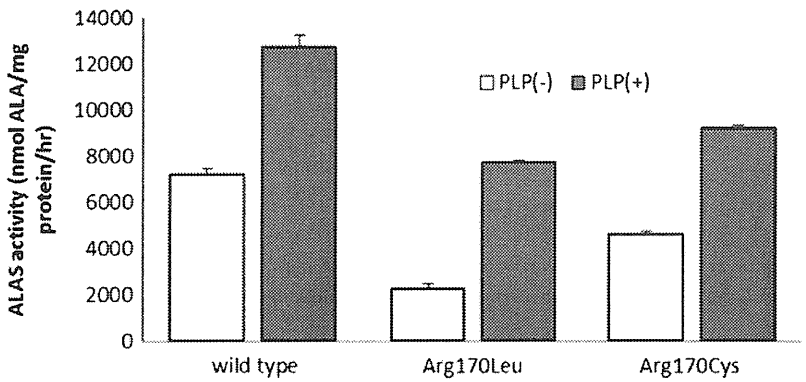

Fig. 2 Enzymatic activity of mutant ALAS2 proteins. Enzymatic activity of wild-type and mutant ALAS2 proteins was measured as described in Materials and Methods. Both of R170L and R170C ALAS2 mutant proteins showed decreased enzymatic activity; however, the activity was partially restored by the addition of PLP