Features and Outcome of Neonatal Leukemia in Japan: Experience of the Japan Infant Leukemia Study Group

Eiichi Ishii, MD,[1]* Megumi Oda, MD,[2] Naoko Kinugawa, PhD,[3] Takanori Oda, MD,[4]
Tetsuya Takimoto, MD,[5] Nobuhiro Suzuki, MD,[6] Yoshiyuki Kosaka, MD,[7] Akira Ohara, MD,[8]
Atsushi Ogawa, MD,[9] Mutsuo Ishii, MD,[10] Naoki Sakata, MD,[11] Takayuki Okamura, MD,[11]
Kenichi Koike, MD,[12] Seiji Kojima, MD,[13] Keizo Horibe, MD,[14] and Shuki Mizutani, MD[15]

Background. Neonatal leukemia characterized by early stem cell origin and extramedullary infiltration in the first 4 weeks of life is rare. We analyzed the features and outcome of neonatal leukemia in Japan to establish an appropriate treatment strategy for this rare disorder. **Procedure.** Patients with infant leukemia registered and treated in the Japan Infant Leukemia Study between 1996 and 2001 were analyzed. **Results.** Among 162 infant leukemia patients, 11 exhibited neonatal leukemia; frequencies for all infant leukemias were 6.9% (8/116) for acute lymphoblastic leukemia (ALL) and 7.3% (3/41) for acute myeloid leukemia (AML). Positive *MLL* gene rearrangement was observed in all eight patients with ALL; a single patient with AML displayed germline configuration. Acute mono-blastic leukemia was apparent in all three patients with AML (M5a in the FAB classification). Most of the patients demonstrated hepatosplenomegaly and hyperleukocytosis at diagnosis. Cutaneous and central nervous system involvement were detected in half of the patients. Four patients (one with AML, and three with ALL) have survived following stem cell transplantation (SCT); however, growth impairment related to SCT was observed in these patients. **Conclusions.** These results suggest an improvement attributable to treatment of neonatal leukemia. International-based collaborative studies are necessary to investigate the biology of this condition and to establish appropriate therapeutic strategies. Pediatr Blood Cancer 2006;47: 268–272. © 2005 Wiley-Liss, Inc.

Key words: acute leukemia; extramedullary leukemia; *MLL* gene rearrangement; neonate

INTRODUCTION

Leukemia occurring in infants less than 12 months of age is often accompanied by hyperleukocytosis, hepatosplenomegaly, and extramedullary leukemic infiltration at initial diagnosis [1]. Biologically, leukemic cells are characterized by positive 11q23 translocations/*MLL* gene rearrangements, which are associated with poor outcome in instances of acute lymphoblastic leukemia (ALL) in most treatment programs [2]. However, we previously demonstrated the effect of intensive chemotherapy followed by stem cell transplantation (SCT) in infant ALL patients displaying *MLL* gene rearrangements [3,4]. Infant acute myeloid leukemia (AML) also exhibits clinical and biological features distinct from those of childhood AML, which include monoblastic or myelomonoblastic phenotypes with 11q23 translocations/ *MLL* gene rearrangements [2]. Although the outcome and prognostic factors in infants with AML is obscure, we confirmed the effect of intensive chemotherapy without SCT in cases of infant AML, despite the presence of *MLL* gene rearrangements or other possible prognostic factors [5].

Neonatal leukemia, which is also classified as congenital leukemia, is rare and accounts for <1% of all childhood leukemias [6]. The following diagnostic criteria have been proposed for this rare disorder: presentation in the first 4 weeks of life, proliferation of immature myeloid, lymphoid or erythroid cells, and infiltration into extramedullary tissues [7,8]. Greaves et al. clarified the etiology of infant leukemia; analysis of *MLL* gene rearrangements of identical leukemic twins suggested that the *MLL* gene undergoes prenatal rearrangement and that leukemic cells are present in the blood of newborns [9]. *In utero* exposure to some drugs and foods that possess functional similarity to topoisomerase-II inhibitors and to certain environmental factors may affect *MLL* gene rearrangements, which contributes to leukemogenesis [10]. Based on these findings, features and outcome of neonatal leukemia were analyzed in order to clarify pathogenesis and to establish more appropriate treatment modalities.

[1]Department of Pediatrics, Saga University, Saga, Japan; [2]Department of Pediatrics, Okayama University, Okayama, Japan; [3]Department of Medical Information Science, Kyushu University, Fukuoka, Japan; [4]Department of Pediatrics, Hokkaido Children's Hospital and Medical Center, Sapporo, Japan; [5]Division of Pediatrics, Ohtsu Red Cross Hospital, Japan; [6]Department of Pediatrics, Sapporo Medical University, Sapporo, Japan; [7]Department of Pediatrics, Kobe University, Kobe, Japan; [8]Department of Pediatrics, Toho University, Tokyo, Japan; [9]Division of Pediatrics, Niigata Cancer Center Niigata Hospital, Niigata, Japan; [10]Division of Pediatrics, Nagoya Red Cross Second Hospital, Nagoya, Japan; [11]Division of Pediatrics, Osaka Medical Center and Research Institute for Maternal and Child Health, Osaka, Japan; [12]Department of Pediatrics, Shinshu University, Matsumoto, Japan; [13]Department of Pediatrics, Nagoya University, Nagoya, Japan; [14]Department of Pediatrics, Nagoya Medical Center, Nagoya, Japan; [15]Department of Pediatrics, Tokyo Medical and Dental University, Tokyo, Japan

*Correspondence to: Eiichi Ishii, Department of Pediatrics, Faculty of Medicine, Saga University, 5-1-1 Nabeshima, Saga 849-8501, Japan. E-mail: ishiei@med.saga-u.ac.jp

WILEY
InterScience®
DISCOVER SOMETHING GREAT

PATIENTS AND METHODS

All infants less than 12 months of age at diagnosis presenting with acute leukemia were registered in the Japan Infant Leukemia Study Group [3–5]. The period of registration was December 1995 to June 2001 for ALL, and December 1995 to December 1998 for AML. Informed consent was obtained from the parents of all patients at the time of registration. Each patient was evaluated in terms of characteristics of leukemic cells in bone marrow or alternatively, in peripheral blood in case bone marrow was not obtained due to dry tap, which included immuno-phenotypic analysis, cytogenetics and assessment of *MLL* gene rearrangements via Southern blotting. Acute leukemia was diagnosed when leukemic cells exceeded 30% of total nuclear cells in bone marrow. ALL or AML was defined according to morphological, cytochemical, and immunological analyses. Acute mixed lineage leukemia (AMLL) was also diagnosed according to the criteria proposed by the European Group for the immunological Classification of Leukemia (EGIL) [11]. Central nervous system (CNS) invasion was defined as follows: $>5/\mu l$ mononuclear cells with obvious blasts present in the cerebrospinal fluid at onset. Cutaneous involvement was defined as clinically apparent skin infiltration, which disappeared with treatment.

Neonatal leukemia was diagnosed with respect to specific symptoms evident during the first 4 weeks (28 days) of life in accordance with criteria documented in several reports [7,8]. Some of these patients, who were considered congenital leukemia due to the observation of symptoms at birth or within the first few days of life, were classified as having neonatal leukemia. Patients with ALL received either the MLL96 (December 1995 to December 1998) or the MLL98 (January 1999 to June 2001) protocol. The protocol schedule was described in previous publications [3,4]. Those exhibiting AML were treated with the ANLL91 protocol (December 1995–1998), which has also been previously described [5]. Triple intrathecal chemotherapy with methotrexate, cytarabine, and hydrocortisone as prophylaxis for CNS leukemia was also administered during each course of the induction or consolidation therapy; five times in the MLL96 or MLL98, and nine times in the ANLL91 protocols.

RESULTS

One hundred sixty infants were enrolled in this study: 116 displaying ALL, 41 presenting with AML, and 3 demonstrating AMLL. The number of AML patients was relatively low as AML registration was limited to the period from 1996 to 1998. Eleven patients were diagnosed with neonatal leukemia, which included 8 with ALL and 3 with AML, while none with AMLL was observed (Table I). Frequencies of neonatal leukemia for all infant leukemias were 6.9% (8/116) for ALL and 7.3% (3/41) for AML.

TABLE I. Features and Outcome of 11 Patients With Neonatal Leukemia

Age/gender	Type	MLL	Karyotype	Skin inv.	Liver (cm)	Spleen (cm)	WBC (/μl)	PL (/μl)	CNS inv.	Marker	Chemotherapy[a]	CR	HSCT[b]	Outcome
19d/F	M5a	+	t(4;11)	–	0	0	66,500	50,000	–	CD33, CD14, CD13, CD15	P, C	–	–	Dead (1mo)
3d/M	M5a	–	46,XY	–	4	0	91,810	70,000	+	CD7, CD33, CD14, CD13, DR, CD56	ANLL91	+	AlloPBSCT (CR1, BU)	Alive in CR (7y+)
15d/F	M5a	+	46,XX	–	6	2	209,400	116,000	–	CD4, CD13, CD15, CD33, DR, MPO	ANLL91	+	–	Dead (1mo)
12d/F	ALL	+	t(4;11)	+	5	5	198,200	47,000	–	CD19, CD79, CD33, CD41, CD34, TdT, DR	MLL96	+	AlloBMT (noCR, TBI)	Alive in CR (6y+)
8d/F	ALL	+	t(11;19)	+	3	0	72,000	24,000	+	CD19, CD34, CD38, DR/CD33, CD38, CD45	ANLL91	+	AlloPBSCT (CR2, BU)	Alive in CR (3y+)
0d/F	ALL	+	t(11;19)	+	3.5	3	898,000	21,000	+	CD19, DR	MLL96	–	–	Dead (4mo)
24d/F	ALL	+	t(4;11;15)	+	3	2.5	121,600	82,000	+	CD19, CD22, DR, CD15	MLL96	+	AlloBMT (CR1, BU)	Dead (9mo)
28d/M	ALL	+	t(4;11)	–	6	3	20,200	45,000	–	CD19, CD22, DR, CD79a, CD15	P, V, D	–	–	Dead (1mo)
25d/F	ALL	+	t(4;11)	–	7	6	421,500	20,000	–	CD19, DR	MLL96	+	–	Dead (2mo)
22d/F	ALL	+	t(9;11)	+	2	0	97,400	22,000	–	CD19, CD79, CD7, DR	MLL96	+	UCBT (CR1, TBI)	Alive in CR (6y+)
13d/M	ALL	+	t(4;11)	+	4	2	23,900	63,000	+	CD19, CD34, DR	MLL98	+	UCBT (CR2, BU)	Dead (15mo)

[a]MLL96, MLL98, and ANLL91 protocols were described elsewhere. 3–5P, prednisolone; C, cytarabine; V, vincristine; D, daunorubicine.
[b]Timing and preparative regimens of HSCT were shown in parenthesis. CR1, first CR; CR2, second CR; TBI, total body irradiation-containing regimen; BU, Busulfan-containing regimen.

The association between age at onset and the presence of MLL gene rearrangements was analyzed in infant leukemia. In ALL, all 24 patients less than 2 months of age exhibited positive MLL gene rearrangements, whereas only two-thirds in late infancy (9–11 months of age) (16/23) displayed this genetic rearrangement. No relationship was detected between age at onset and incidence of positive MLL gene rearrangements with respect to AML in this investigation. Outcome of infant leukemia was also analyzed. ALL infants less than 5 months of age demonstrated poor outcome; in particular, most of those less than 1 month of age died due to progressive disease or to complications associated with treatment. In AML, an apparent correlation was not detected between age at onset and outcome.

Clinical and laboratory findings of 11 patients with neonatal leukemia are summarized in Table I. Although the onset of disease varied among the patients, three of nine patients diagnosed after the first week of life exhibited skin involvement (petechiae or cutaneous eruption) at birth, suggesting these patients also had leukemia prenatally. All eight patients with ALL exhibited 11q23 translocation associated with MLL gene rearrangements. They revealed negative expression of CD10 antigen, whereas four displayed simultaneous expression of myeloid (CD33) or monocytoid (CD15) antigen. Acute monoblastic leukemia was apparent in all three AML patients (M5a in FAB classification); however, one demonstrated the MLL gene germline configuration. Hepatosplenomegaly and hyperleukocytosis were observed in most of the infants with neonatal leukemia at presentation. Cutaneous and CNS involvement was detected in six (55%) and five (45%) patients, respectively. Of the eight patients achieving complete remission following chemotherapy, six underwent SCT including allogeneic bone marrow transplantation (alloBMT) (2), allogeneic peripheral blood stem cell transplantation (alloPBSCT) (2), and unrelated cord blood transplantation (UCBT) (2). The preparative regimen for SCT consisted of total body irradiation (TBI) in two and busulfan (BU) in four patients, respectively. Two of five patients who relapsed following complete remission (CR) died without SCT. Consequently, four patients (1 with AML and 3 with ALL) have survived after SCT. Timing of transplantation in these patients was first CR in two, second CR in one, and noCR in one.

Long-term side effects were analyzed in four patients who have survived for an extended period. All patients with neonatal leukemia demonstrated apparent growth impairment with respect to body height and body weight; −1.5 SD and −1.2 SD at 1 year, −1.6 SD and −1.7 SD at two years, and −2.4 SD (body weight was not tested) at 3 years after the completion of therapy, respectively. Other late sequelae were observed in two patients, one with hypothyroidism and the other with chronic GVHD. Lung involvement, cardiac dysfunction, psychosomatic deterioration, or secondary malignancy were not detected in any patient in this study.

DISCUSSION

Neonatal leukemia displays numerous characteristic features that facilitate differentiation from other infant leukemias. Although it has been reported that the frequency of AML is greater than that of ALL in the neonatal period [7,12,13], frequencies of ALL and AML were nearly identical in the current study. Leukemic cells of all patients presenting with AML exhibited monoblastic features consistent with M5a morphology of the FAB classification, which was described in the literature [14,15]. Although no patients with AMLL were observed, there was a relatively high frequency of ALL with myeloid antigen execution noted in this investigation. Co-expression of lymphoid and myeloid antigens or phenotypic switch indicates that leukemic cells in neonatal leukemia develop from an early stem cell that can differentiate in both myeloid and lymphoid directions [13,16,17].

The characteristic clinical feature of neonatal leukemia is a high frequency of cutaneous or CNS involvement at presentation; the former is known as "leukemia cutis" or "blueberry muffin baby." In several investigations, 20–60% of patients with neonatal leukemia exhibited cutaneous infiltration, which appeared more often in AML than in ALL [8,13,18]. In the current study, cutaneous infiltration at presentation, which disappeared with chemotherapy, was observed in half of the patients. Spontaneous waxing and waning of leukemia cutis has been documented; however, the form involving leukemia cutis usually displays an aggressive clinical course [19]. CNS involvement is also common in neonatal leukemia: the incidence was 50% [12]. Half of the patients in this study were characterized by CNS involvement at onset, which was correlated with poor prognosis [3].

The high frequency of 11q23 translocations/MLL gene rearrangements in neonatal leukemia was also evident in this investigation. In a larger study, however, only 30.6% of patients displayed 11q23 translocation; in contrast, MLL gene rearrangement was confirmed in most of the patients [13].The frequency of MLL gene rearrangements is correlated strongly with age in infant ALL [9]. High frequency of neonatal patients exhibiting MLL gene rearrangements strongly suggests prenatal rearrangement of the MLL gene. In fact, some of the patients who were diagnosed after the first week of life also displayed symptoms of disease at birth. MLL gene rearrangements can be detected in blood samples at birth in most infant leukemia patients, which suggests a prenatal origin of leukemic cells in infant leukemia [20,21]. It has been speculated that exposure to specific topoisomerase-II inhibitors could form cleavage complexes leading to facilitation of illegitimate recombination [22].

The prognosis of neonatal leukemia has been generally poor. In the aforementioned larger study, overall survival for 69 AML and 22 ALL patients was 24.4% and 13.6%,

respectively; in particular, disease-free survival of ALL patients was dismal (0%) [13]. Poor prognosis in neonatal ALL has been attributed to unfavorable features at presentation, which include positive *MLL* gene rearrangements, hyperleukocytosis and extramedullary (cutaneous and CNS) involvement [19]. Co-expression of myeloid antigen may also be correlated with poor outcome in these patients [23]. In the current study, three of eight ALL patients survived with intensive chemotherapy followed by SCT, suggesting that improvement in the treatment of these patients is essential in order to achieve high remission rates in infant ALL.

One of three patients presenting with AML has also survived following myeloid-oriented chemotherapy (ANLL91) and SCT in this study. Despite the unfavorable features, infants with AML demonstrate outcomes similar to those of older children with AML upon implementation of intensive multi-agent chemotherapy regimens [5,19]. Pui et al. noted that only two factors predicted a favorable prognosis: M4 or M5 leukemia and the t(9;11) [24]. We recently reported that infant ALL involving t(9;11) possesses a feature with respect to poor prognosis distinct from infant AML with the identical karyotype; furthermore, prognosis of these patients can be improved only via combination of intensive chemotherapy and SCT [25]. Therefore, we hypothesize that a strategy for infant ALL different from that of infant AML should be established.

A major concern related to the utility of SCT in infants is the late toxicity including growth impairment. In particular, TBI was also reported to be a risk factor for short stature, hypopituitarism, and cataract formation; furthermore, these toxicities were inversely correlated with age [26,27]. In this investigation, patients with neonatal leukemia exhibited apparent growth impairment in terms of body height and body weight. In our previous report, outcome of infant ALL did not differ between patients undergoing the TBI- or BU-based regimen [4]. These data indicate that SCT in conjunction with less toxic preparative regimens involving non-myeloablative drugs should be included in the treatment of future neonatal leukemia patients.

To date, the best treatment strategy in terms of improving survival and reducing late toxicity in neonatal leukemia remains elusive. Since limited numbers of neonatal leukemia appear in the literature, international-based collaborative studies are necessary for acquisition of information regarding biology, treatment and natural history of this rare disease.

ACKNOWLEDGMENT

REFERENCES

1. Ishii E, Okamura J, Tsuchida M, et al. Infant leukemia in Japan: Clinical and biological analysis of 48 cases. Med Pediatr Oncol 1991;19:28–32.
2. Pui C-H, Kane JR, Crist WM, et al. Biology and treatment of infant leukemia. Leukemia 1995;9:762–769.
3. Isoyama K, Eguchi M, Hibi S, et al. Risk-directed treatment of infant acute lymphoblastic leukemia based on early assessment of MLL gene status: Results of the Japan Infant Leukemia Study (MLL96). Br J Haematol 2002;118:999–1010.
4. Kosaka K, Koh K, Kinukawa N, et al. Infant acute lymphoblastic leukemia with *MLL* gene rearrangements: Outcome following intensive chemotherapy and hematopoietic stem cell transplantation. Blood 2004;104:3527–3534.
5. Kawasaki H, Isoyama K, Eguchi M, et al. A superior outcome of infant acute myeloid leukemia with intensive chemotherapy. Blood 2001;98:3589–3594.
6. Bader JL, Miller RW. US cancer incidence and mortality in the first year of life. Am J Dis Child 1979;133:157–159.
7. Pierce MI. Leukemia in the newborn infant. J Pediatr 1959;54:691–706.
8. Resnik KS, Brod BB. Leukemia cutis in congenital leukemia: Analysis and review of the world literature with report of an additional case. Arch Dermatol 1993;129:1301–1306.
9. Greaves MF. Infant leukemia biology, aetiology and treatment. Leukemia 1996;10:372–377.
10. Ishii E, Eguchi M, Eguchi-Ishimae M, et al. In vitro cleavage of the MLL gene by topoisomerase II inhibitor (etoposide) in normal cord and peripheral blood mononuclear cells. Int J Hematol 2002;76:74–79.
11. Matutes E, Morilla R, Farahat N, et al. Definition of acute biphenotypic leukemia. Haematologica 1997;82:64–66.
12. van Wering ER, Kamps WA. Acute leukemia in infants: A unique pattern of acute nonlymphocytic leukemia. Am J Pediatr Hematol Oncol 1986;8:220–224.
13. Bresters D, Reus ACW, Veeman AJP, et al. Congenital leukaemia: The Dutch experience and review of the literature. Br J Haematol 2002;117:513–524.
14. Taki T, Ida K, Bessho F, et al. Frequency and clinical significance of the MLL gene rearrangements in infant acute leukemia. Leukemia 1996;10:1303–1307.
15. Swansbury GJ, Slater R, Bain BJ, et al. Hematological malignancies with t(9;11)(p21-22;q23)—a laboratory and clinical study of 125 cases. Leukemia 1998;12:792–800.
16. Ridge SA, Cabrera ME, Ford AM, et al. Rapid intraclonal switch of lineage dominance in congenital leukemia with a MLL gene rearrangement. Leukemia 1995;9:2023–2026.
17. Matamoros N, Matutes E, Hernandez M, et al. Special case report: Neonatal mixed lineage leukemia. Leukemia 1994;8:1236–1242.
18. Frangoul HA, Patterson K. Complications of acute leukemia. Case one: Congenital acute myelogenous leukemia with cutaneous involvement. J Clin Oncol 1998;16:3199–3202.
19. Sande JE, Arceci RJ, Lampkin BC. Congenital and neonatal leukemia. Sem Perinatol 1999;23:274–285.
20. Yagi T, Hibi S, Tabata Y, et al. Detection of clonotypic IGH and TCR rearrangements in the neonatal blood spots of infants and children with B-cell precursor acute lymphoblastic leumemia. Blood 2000;96:264–268.
21. Gale KB, Ford AM, Repp R, et al. Backtracking leukemia to birth: Identification of clonotypic gene fusion sequences in neonatal blood spots. Proc Natl Acad Sci USA 1997;94:13950–13954.

22. Greaves M. Molecular genetics, natural history and the demise of childhood leukaemia. Eur J Cancer 1999;35:173–185.

23. Basso G, Rondelli R, Covezzoli A, Putti MC. The role of immunophenotype in acute lymphoblastic leukemia of infant age. Leuk Lymphoma 1994;15:51–60.

24. Pui C-H, Raimondi SC, Srivastava DK, et al. Prognostic factors in infants with acute myeloid leukemia. Leukemia 2000;14:684–687.

25. Kitazawa J, Tono C, Terui K, et al. Successful outcome of mismatched hematopoietic stem cell transplantation from a related donor in an infant with acute lymphoblastic leukemia and a 9;11 translocation: A case report and review of the literature. In J Hematol 2005;81:428–432.

26. Cohen A, Rovelli A, Bakker B, et al. Final height of patients who underwent bone marrow transplantation for hematological disorders during childhood: A study by the working party for late effects- EBMT. Blood 1999;93:4109–4115.

27. Leung W, Hudson M, Zhu Y, et al. Late effects in survivors of infant leukemia. Leukemia 2000;14:1185–1190.

Brief report

KIT mutations, and not FLT3 internal tandem duplication, are strongly associated with a poor prognosis in pediatric acute myeloid leukemia with t(8;21): a study of the Japanese Childhood AML Cooperative Study Group

Akira Shimada, Tomohiko Taki, Ken Tabuchi, Akio Tawa, Keizo Horibe, Masahiro Tsuchida, Ryoji Hanada, Ichiro Tsukimoto, and Yasuhide Hayashi

Patients with t(8;21) acute myeloid leukemia (AML) are considered to have a good prognosis; however, approximately 50% of them relapse. The genetic alterations associated with a poor outcome in t(8;21) AML remain unknown. Recently, aberrations of receptor tyrosine kinases (RTKs) were frequently found in patients with AML. However, the prevalence and prognostic impact of RTK aberrations in pediatric t(8;21) AML remains undetermined. Here, we found the kinase domain mutations of the KIT gene in 8 (17.4%) of 46 patients with t(8;21) AML among newly diagnosed pediatric patients with AML treated on the AML99 protocol in Japan. Significant differences between patients with or without KIT mutations were observed in the 4-year overall survival (50.0% versus 97.4%, P = .001), disease-free survival (37.5% versus 94.7%, P < .001) and relapse rate (47.0% versus 2.7%, P < .001). Furthermore, FLT3 internal tandem duplication was found in only 2 (4.3%) patients. These results suggested that KIT mutations are strongly associated with a poor prognosis in pediatric t(8;21) AML. (Blood. 2006; 107:1806-1809)

© 2006 by The American Society of Hematology

Introduction

Patients with t(8;21) acute myeloid leukemia (AML) have been reported to have a good prognosis; however, approximately 50% of them relapse.[1,2] A high presenting leukocyte count, CD56 expression, or extramedullary disease has been reported to be associated with a poor prognosis in t(8;21) AML.[1,3,4] However, the genetic alterations associated with a poor outcome in patients with t(8;21) AML remain unknown. Recent studies revealed that internal tandem duplication (ITD) of FLT3 is considered to be one factor predicting poor prognosis in adult and pediatric patients with AML.[5-9] More recently, KIT mutations were found in 12.7% to 48.1% of adult patients with AML with t(8;21)[10-12] and were reported to be associated with a poor prognosis.[13,14] The prevalence and prognostic impact of KIT mutations in pediatric t(8;21) AML remain unknown. We performed the mutational analysis of KIT and FLT3 in pediatric patients with t(8;21) AML who were treated on the Japanese Childhood AML Cooperative Study Group Protocol, AML99.

We report here that KIT mutations are strongly associated with a poor prognosis in pediatric patients with t(8;21) AML.

Study design

Patients and samples

The diagnosis of AML was based on the French-American-British (FAB) classification, and cytogenetic analysis was performed using a routine G-banding method. From January 2000 to December 2002, 318 patients were newly diagnosed as having de novo AML. Of 240 patients, 77 (32.1%), except for 29 AML-M3 and 49 Down syndrome, had t(8;21)(q22; q22) according to cytogenetics or AML1-MTG8 fusion transcript with the reverse-transcriptase–polymerase chain reaction (RT-PCR) (Figure S1; see the Supplemental Materials link at the top of the online article, at the Blood website). Samples were available from 135 (56.3%) of 240 patients with AML, including 46 (59.7%) of 77 patients with t(8;21) AML. Of 46 patients with t(8;21) AML, 3 patients were classified into M1, 39 into M2, and 4 into

Figure 1. Kaplan-Meier analysis. This analysis shows 4-year overall survival (A) and disease-free survival (B) of the patients with or without *KIT* mutation. The difference is statistically significant (A: P = .001; B: P < .001).
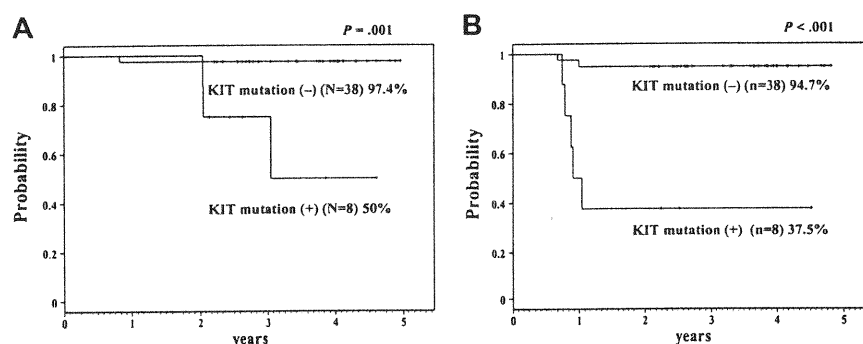
M4. There were no statistical differences between 46 analyzed patients with t(8;21) AML and the 31 nonanalyzed patients in age (median 7.5 years [range: 2-15 years] versus 9 years [range: 1-15 years]), initial white blood cell (WBC) count (median: 14.4 × 10⁹/L; range: 1.65 × 10⁹/L-107.7 × 10⁹/L; versus 9.1 × 10⁹/L; range: 1.4 × 10⁹/L-136 × 10⁹/L), induction rate (100% versus 93.5%), relapse rate (15.2% versus 19.4%), and 4-year overall survival rate (4y-OS; 87% versus 91%). In the AML99 protocol, patients with t(8;21) with initial WBC count lower than 50 × 10⁹/L were categorized into a low-risk group. Thus, after patients with t(8;21) AML obtained complete remission (CR) with induction chemotherapy (cytarabine, etoposide, and mitoxantrone), they were treated with 5 additional courses of intensive chemotherapy (high-dose cytarabine [HDCA], etoposide, idarubicine, and mitoxantron; Figure S2 and Tsukimoto et al[15]). If the initial WBC count was greater than 50 × 10⁹/L, patients were categorized into an intermediate-risk group and received allogeneic stem cell transplantation (allo-SCT) in the case of the presence of a donor. Informed consent was obtained from the patients or patients' parents, according to guidelines based on the tenets of the revised Helsinki protocol. The institutional review board of Gunma Children's Medical Center approved this project.

KIT mutation analysis

Mutational analysis of the extracellular (EC) domain (exons 8, 9), transmembrane (TM) domain (exon 10), juxtamembrane (JM) domain (exon 11), and the second intracellular kinase (TK) 2 domain (exons 17 and 18) of *KIT* gene was performed with RT-PCR followed by direct sequencing. Primers used are shown in Table S1.

FLT3 mutation analysis

Mutational analysis of ITD within the JM domain and D835 mutation (D835Mt) within the TK2 domain of the *FLT3* gene was performed as previously reported.[16-18]

Statistical analysis

Estimation of survival distributions was performed using the Kaplan-Meier method and the differences were compared using the log-rank test. Disease-free survival (DFS), event-free survival (EFS), and overall survival

(OS) were defined as the times from diagnosis to relapse, from diagnosis to event (relapse or death of any cause), and from diagnosis to death of any cause or the last follow-up. Statistical difference analysis was performed using the χ^2 test.

Results and discussion

KIT and *FLT3* expressions were found in all of the 46 t(8;21) AML samples. Although *KIT* mutations have been reported in a small number of pediatric patients with t(8;21) AML,[8,19] TK2 domain mutations of the *KIT* gene were found in 8 (17.4%) of 46 patients in this study (Table 1). However, we could not find any mutation other than the TK2 domain. The N822K mutation, which has been frequently reported so far,[12] was found in 3 of 8 patients in this study.

The statistical differences between patients with or without *KIT* mutations were not significant in age (median 8 years [range: 1-15 years] versus 7 years [range: 2-15 years]), and the initial WBC count (median: 20.65 × 10⁹/L; range: 4.6 × 10⁹/L-66.2 × 10⁹/L; versus 14.3 × 10⁹/L; range: 1.65 × 10⁹/L-107.7 × 10⁹/L). Interestingly, *KIT* mutations were observed only in M2 patients according to FAB classification. Another report also suggested that *KIT* mutations were frequently found in M2 patients with t(8;21).[19] Significant differences between patients with or without *KIT* mutations were observed in 4-year OS (50.0% versus 97.4%, P = .001, Figure 1), DFS (37.5% versus 94.7%, P < .001), and relapse rate (47.0% versus 2.7%, P < .001). Short CR duration and high relapse rate were more significant than those of the previous report in adults.[14] *KIT* mutations have recently been reported not to influence the clinical outcome in pediatric core-binding factor (CBF) leukemia patients.[20] Although they found *KIT* mutations in 5 of 16 cases of t(8;21) AML, they did not describe the clinical outcome of patients with t(8;21) AML with or without *KIT* mutations. Furthermore, the clinical outcome of the patients

Table 1. Clinical characteristics of patients with t(8;21) AML with *KIT* mutations

Patient no.	Age, y	Sex	WBC count, × 10⁹ cells/L	Additional chromosome abnormalities	Time of relapse, mo	Status of allo-SCT	Survival, mo	*KIT* mutation
1	8	F	14.10	None	12	Second CR	37	A814S
2	8	M	27.60	−Y	14	Second CR	47*	N822K
3	8	F	10.77	−X	10	Second CR	25	D816H
4	6	M	34.50	−Y, +4	12	Second CR	26*	N822K
5	3	F	20.50	None	11	—	25	N822K
6	1	F	4.60	−X, t(7;9)	—	—	32*	N822T
7	15	M	20.80	−Y	—	First CR	56*	D816V
8	13	M	66.20	None	—	First CR	30*	V825A

— indicates not applicable.
*Patient still alive.

without *KIT* mutations in their study was poorer than the outcome of those in our study (EFS 63% versus 92.1%). Our result may depend on our good clinical outcome of patients with t(8;21) AML without *KIT* mutations.

Except for 2 patients who received allo-SCT in first CR (patients no. 7 and no. 8 in Table 1), 5 of 6 (83.3%) patients with the mutation relapsed within 14 months after diagnosis. Allo-SCT was performed in 6 of 8 patients with t(8;21) AML with *KIT* mutations (2 in first CR, 4 in second CR) and 4 patients are still alive. In contrast, allo-SCT was also performed in only 1 of 38 patients with t(8;21) AML without *KIT* mutation in second CR, and this patient is still alive.

A high presenting leukocyte count and extramedullary disease were not associated with the poor prognosis in this study. Notably, *KIT* was mapped to chromosome 4 at band q11 and trisomy 4 was reported to be associated with *KIT* mutation.[21] One patient with trisomy 4 in addition to t(8;21) had N822K mutation (patient no. 4). As for additional chromosome abnormality, loss of sex chromosome was observed in 5 (62.5%) of 8 patients with *KIT* mutation and 14 (37%) of 38 patients without mutations, although the difference between them was not statistically significant. Recently, it has been reported that AML blasts with N822K mutation are sensitive to the tyrosine kinase inhibitor Gleevec/STI571/imatinib mesylate.[12] The effectiveness of imatinib mesylate for the patient with AML with *KIT* mutation was also reported.[22] Thus, tyrosine kinase inhibitors may be applicable for these patients in the future.

Two samples examined at relapse showed the same mutations as those at diagnosis (patients no. 3 and no. 5), and these *KIT* mutations disappeared in samples in remission, suggesting that *KIT* mutation was not a constitutional abnormality.

Recently, clonal leukemic cells with *AML1-MTG8* fusion transcript have been reported to arise in utero.[23] Moreover, it was reported that this fusion transcript was not sufficient for full leukemogenesis, and that additional genetic events were required.[24,25] *KIT* mutations may be one of the secondary genetic events of the stepwise leukemogenesis of t(8;21) AML.

FLT3-ITD was found in only 2 (4.6%) of 46 patients with t(8;21). One patient died during chemotherapy, and the other patient was disease free for 42 months from diagnosis. *FLT3*-ITD is considered to be strongly associated with a poor prognosis in AML.[6,7] However, *FLT3*-ITD was rarely reported in patients with t(8;21) AML.[8,9,13,14,20] Our data also confirmed the low incidence of *FLT3*-ITD in patients with t(8;21) AML. As for D835Mt of the *FLT3* gene, we found the mutation in 1 of 46 patients, who was alive for 31 months after diagnosis.

In total, 11 (23.9%) of 46 patients with t(8;21) AML in this study had *KIT* or *FLT3* mutations, suggesting that the pediatric patients with t(8;21) AML had genetic heterogeneity. In conclusion, *KIT* mutations are considered to be strongly associated with poor prognosis in pediatric t(8;21) AML.

Acknowledgment

Appendix

Members of the Japanese Childhood AML Cooperative Study Group who contributed data to the study include Akira Morimoto, Department of Pediatrics, Kyoto Prefectural University of Medicine; Ryoji Kobayashi, Department of Pediatrics, Hokkaido University School of Medicine; Hiromasa Yabe, Department of Pediatrics, Tokai University School of Medicine; Kazuko Hamamoto, Department of Pediatrics, Hiroshima Red Cross Hospital; Shigeru Tsuchiya, Department of Pediatric Oncology, Institute of Development, Aging, and Cancer, Tohoku University; Yuichi Akiyama, Department of Pediatrics, National Hospital Organization Kyoto Medical Center; Hisato Kigasawa, Department of Hematology, Kanagawa Children's Medical Center; Akira Ohara, Department of First Pediatrics, Toho University School of Medicine; Hideki Nakayama, Department of Pediatrics, Hamanomachi Hospital; Kazuko Kudo, Department of Pediatrics, Nagoya University Graduate School of Medicine; and Masue Imaizumi, Department of Hematology/Oncology, Miyagi Prefectural Children's Hospital.

References

1. Rubnitz JE, Raimondi SC, Halbert AR, et al. Characteristics and outcome of t(8;21)-positive childhood acute myeloid leukemia: a single institution's experience. Leukemia. 2002;16:2072-2077.

2. Schlenk RF, Benner A, Krauter J, et al. Individual patient data-based meta-analysis of patients aged 16 to 60 years with core binding factor acute myeloid leukemia: a survey of the German Acute Myeloid Leukemia Intergroup. J Clin Oncol. 2004;22:3741-3750.

3. Nguyen S, Leblanc T, Fenaux P, et al. A white blood cell index as the main prognostic factor in t(8;21) acute myeloid leukemia (AML): a survey of 161 cases from the French AML Intergroup. Blood. 2002;99:3517-3523.

4. Baer MR, Stewart CC, Lawrence D, et al. Expression of the neural cell adhesion molecule CD56 is associated with short remission duration and survival in acute myeloid leukemia with t(8;21)(q22; q22). Blood. 1997;90:1643-1648.

5. Yokota S, Kiyoi H, Nakao M, et al. Internal tandem duplication of the FLT3 gene is preferentially seen in acute myeloid leukemia and myelodysplastic syndrome among various hematological malignancies: a study on a large series of patients and cell lines. Leukemia. 1997;11:1605-1609.

6. Kottaridis PD, Gale RE, Frew ME, et al. The presence of a FLT3 internal tandem duplication in patients with acute myeloid leukemia (AML) adds important prognostic information to cytogenetic risk group and response to the first cycle of chemotherapy: analysis of 854 patients from the United Kingdom Medical Research Council AML 10 and 12 trials. Blood. 2001;98:1752-1759.

7. Thiede C, Steudel C, Mohr B, et al. Analysis of FLT3-activating mutations in 979 patients with acute myelogenous leukemia: association with FAB subtypes and identification of subgroups with poor prognosis. Blood. 2002;99:4326-4335.

8. Meshinchi S, Stirewalt DL, Alonzo TA, et al. Activating mutations of RTK/ras signal transduction pathway in pediatric acute myeloid leukemia. Blood. 2003;102:1474-1479.

9. Zwaan CM, Meshinchi S, Radich JP, et al. FLT3 internal tandem duplication in 234 children with acute myeloid leukemia: prognostic significance and relation to cellular drug resistance. Blood. 2003;102:2387-2394.

10. Gari M, Goodeve A, Wilson G, et al. c-kit proto-oncogene exon 8 in-frame deletion plus insertion mutations in acute myeloid leukaemia. Br J Haematol. 1999;105:894-900.

11. Beghini A, Peterlongo P, Ripamonti CB, et al. C-

kit mutations in core binding factor leukemias. Blood. 2000;95:726-727.

12. Wang YY, Zhou GB, Yin T, et al. AML1-ETO and C-KIT mutation/overexpression in t(8;21) leukemia: implication in stepwise leukemogenesis and response to Gleevec. Proc Natl Acad Sci U S A. 2005;102:1104-1109.

13. Care RS, Valk PJ, Goodeve AC, et al. Incidence and prognosis of c-KIT and FLT3 mutations in core binding factor (CBF) acute myeloid leukaemias. Br J Haematol. 2003;121:775-777.

14. Nanri T, Matsuno N, Kawakita T, et al. Mutations in the receptor tyrosine kinase pathway are associated with clinical outcome in patients with acute myeloblastic leukemia harboring t(8;21)(q22; q22). Leukemia. 2005;19:1361-1366.

15. Tsukimoto I, Tawa A, Hanada R, et al. Excellent outcome of risk stratified treatment for childhood acute myeloid leukemia-AML99 trial. For the Japanese Childhood AML Cooperative Study Group [abstract]. Blood. 2005;106:261a. Abstract 889.

16. Xu F, Taki T, Yang HW, et al. Tandem duplication of the FLT3 gene is found in acute lymphoblastic leukaemia as well as acute myeloid leukaemia but not in myelodysplastic syndrome or juvenile chronic myelogenous leukaemia in children. Br J Haematol. 1999;105:155-162.

17. Taketani T, Taki T, Sugita K, et al. FLT3 mutations in the activation loop of tyrosine kinase domain are frequently found in infant ALL with MLL rearrangements and pediatric ALL with hyperdiploidy. Blood. 2004;103:1085-1088.

18. Yamamoto Y, Kiyoi H, Nakano Y, et al. Activating mutation of D835 within the activation loop of FLT3 in human hematologic malignancies. Blood. 2001;97:2434-2439.

19. Beghini A, Ripamonti CB, Cairoli R, et al. KIT activating mutations: incidence in adult and pediatric acute myeloid leukemia, and identification of an internal tandem duplication. Haematologica. 2004;89:920-925.

20. Goemans BF, Zwaan CM, Miller M, et al. Mutations in KIT and RAS are frequent events in pediatric core-binding factor acute myeloid leukemia. Leukemia. 2005;19:1536-1542.

21. Langabeer SE, Beghini A, Larizza L. AML with t(8;21) and trisomy 4: possible involvement of c-kit? Leukemia. 2003;17:1915; author reply 1915-1916.

22. Nanri T, Matsuno N, Kawakita T, Mitsuya H, Asou N. Imatinib mesylate for refractory acute myeloblastic leukemia harboring inv(16) and a C-KIT exon 8 mutation. Leukemia. 2005;19:1673-1675.

23. Wiemels JL, Xiao Z, Buffler PA, et al. In utero origin of t(8;21) AML1-ETO translocations in childhood acute myeloid leukemia. Blood. 2002;99: 3801-3805.

24. Yuan Y, Zhou L, Miyamoto T, et al. AML1-ETO expression is directly involved in the development of acute myeloid leukemia in the presence of additional mutations. Proc Natl Acad Sci U S A. 2001;98:10398-10403.

25. Higuchi M, O'Brien D, Kumaravelu P, Lenny N, Yeoh EJ, Downing JR. Expression of a conditional AML1-ETO oncogene bypasses embryonic lethality and establishes a murine model of human t(8;21) acute myeloid leukemia. Cancer Cell. 2002;1:63-74.

Common gene expression signatures in t(8;21)- and inv(16)-acute myeloid leukaemia

Hitoshi Ichikawa,[1] Kenji Tanabe,[2] Hiroshi Mizushima,[3] Yasuhide Hayashi,[4] Shuki Mizutani,[5] Eiichi Ishii,[6] Teruaki Hongo,[7] Akira Kikuchi[8] and Masanobu Satake[2]

[1]Cancer Transcriptome Project, National Cancer Centre Research Institute, Chuo-ku, Tokyo, [2]Institute of Development, Aging and Cancer, Tohoku University, Aoba-ku, Sendai, [3]Centre for Medical Genomics, National Cancer Centre Research Institute, Chuo-ku, Tokyo, [4]Gunma Children's Medical Centre, Hokkitsu, Shibukawa, Gunma, [5]Department of Paediatrics and Developmental Biology, Tokyo Medical and Dental University, Bunkyo-ku, Tokyo, [6]Department of Paediatrics, Faculty of Medicine, Saga University, Saga, [7]Department of Paediatrics, Iwata City Hospital, Iwata City, Shizuoka, and [8]Division of Haematology and Oncology, Saitama Children's Medical Centre, Iwatsuki-ku, Saitama, Japan

Summary

Human acute myeloid leukaemia (AML) involving a core-binding factor (CBF) transcription factor is called CBF leukaemia. In these leukaemias, AML1 (RUNX1, PEBP2αB, CBFα2)-MTG8 (ETO) and CBFβ (PEBP2β)-MYH11 chimaeric proteins are generated by t(8;21) and inv(16) respectively. We analysed gene expression profiles of leukaemic cells by microarray, and selected genes whose expression appeared to be modulated in association with t(8;21) and inv(16). In a pair-wise comparison, 15% of t(8;21)-associated transcripts exhibited high or low expression in inv(16)-AML, and 26% of inv(16)-associated transcripts did so equivalently in t(8;21)-AML. These common elements in gene expression profiles between t(8;21)- and inv(16)-AML probably reflect the situation that AML1-MTG8 and CBFβ-MYH11 chimaeric proteins affect a common set of target genes in CBF leukaemic cells. On the other hand, 38% of t(8;21)-associated and 24% of inv(16)-associated transcripts were regulated in t(8;21)- and inv(16)-specific manners. These distinct features of t(8;21)- and inv(16)-associated genes correlate with the bimodular structures of the chimaeric proteins (CBF-related AML1 and CBFβ portions, and CBF-unrelated MTG8 and MYH11 portions).

Keywords: transcription factor, chimaera protein, gene expression, microarray.

Human acute myeloid leukaemia (AML) is a heterogeneous group of diseases. Diagnosis of AML-subtypes is of clinical importance, as they vary in their responsiveness to therapy and prognosis. These subtypes are recognised by the morphology of leukaemic cells and classified using the French–American–British (FAB) system. Each FAB subtype corresponds to the differentiation blockage of leukaemic cells at a specific stage in a certain lineage. In addition, various types of chromosomal rearrangements are seen in AML, and a particular type of chromosomal translocation is sometimes associated with a FAB subtype. Among AML-subtypes, genes most frequently encountered in the clinic are those encoding a core-binding factor (CBF; PEBP2) transcription factor (Look, 1997; Speck & Gilliland, 2002). AML involving CBF is known as CBF leukaemia. The AML1 (RUNX1, PEBP2αB, CBFα2) and CBFβ (PEBP2β) proteins constitute both DNA-binding and non-binding subunits of a heterodimeric CBF. Due to chromosomal rearrangements, the AML1-MTG8 (ETO) and CBFβ-MYH11 chimaeric genes are generated in t(8;21)-AML with an M2 subtype and in inv(16)-AML with an M4Eo subtype respectively (Miyoshi et al, 1991, 1993; Erickson et al, 1992; Liu et al, 1993).

Functional studies using reporter assays indicate that the AML1-MTG8 and CBFβ-MYH11 chimaeric polypeptides exhibit dominant negative activity against CBF-dependent transcription (Peterson & Zhang, 2004; Shigesada et al, 2004). Furthermore, when cDNAs of *AML1-MTG8* and *CBFβ-MYH11* are knocked into the respective *Aml1* and *Cbfβ* loci, knocked-in heterozygotes exhibit essentially similar phenotypes as those seen in *Aml1* (−/−) and *Cbfβ*(−/−) mice (Castilla et al, 1996; Yergeau et al, 1997). These phenotypes do not include leukaemia but rather involve the failure to develop definitive-type haematopoiesis, in strong support of the hypothesis that chimaeric proteins function as dominant negatives. This observation suggests that the two chimaeric proteins modulate expression of a common set of target genes in CBF leukaemic cells.

Alternatively, AML1-MTG8 and CBFβ-MYH11 proteins could also modulate gene expression in a specific fashion, as the amino acid sequence of C-terminal portion of each chimaeric protein is not homologous to CBF or to the other. MTG8 belongs to the Nervy family proteins (Kitabayashi et al, 1998) and reportedly interacts with various factors including transcriptional co-repressors (Peterson & Zhang, 2004). On the other hand, the MYH11 portion of the chimaera represents a rod domain seen in smooth muscle myosin heavy chain protein.

Several attempts to correlate karyotypic classification of AML with the gene expression profiles have been reported. Such studies have focused on t(8;21), t(15;17), inv(16), 11q23-alteration and normal karyotypes (Schoch et al, 2002; Debernardi et al, 2003; Kohlmann et al, 2003; Bullinger et al, 2004; Ross et al, 2004; Gutierrez et al, 2005). A recent report classified all manifestations of AML into 16 distinct subgroups based on gene expression profiles (Valk et al, 2004). Given that the AML1-MTG8 and CBFβ-MYH11 proteins are structurally bimodular, this study aimed to examine whether these chimaeric proteins regulate common as well as unique sets of targets in CBF leukaemia. To do so, we analysed gene expression profiles of AML clinical samples by microarray and found that several genes were commonly regulated in t(8;21)- and inv(16)-AML, and that others were regulated in t(8;21)-specific and inv(16)-specific manners. Below we describe our approach and discuss implications of our results.

Materials and methods

Patient samples

A total of 50 paediatric AML patients were enroled in this study (Table SI). Among these, 45 were derived from 54 patients reported previously (nine of the 54 patients were excluded because their FAB subtype or karyotype information could not be obtained) (Yagi et al, 2003). The other five patients (U06, U09, U11, U20 and U21) were included in this study to increase the number of samples grouped as G2, G3 and G4 (see below). Four normal bone marrow samples

enroled in this study were also described previously (Yagi et al, 2003). This study was approved by the ethics committee of the National Cancer Centre and conducted according to tenets of the Declaration of Helsinki. Informed consent was obtained from each patient.

Microarray and statistical analysis

The microarray used in this study was Human Genome U95Av2 (Affymetrix, Santa Clara, CA, USA) containing 12 566 probe sets. For the 45 previously reported samples, the scanned image data obtained previously (Yagi et al, 2003) were re-used. Microarray analysis of the newly included samples was performed as described (Yagi et al, 2003). The analysis included preparation of mononuclear cells from bone marrow or peripheral blood, total RNA isolation, monitoring RNA integrity, preparation of biotin-labelled cRNA from total RNA, hybridisation to the microarray, and washing, staining and scanning of samples. Scanned image data were processed using Affymetrix Microarray Suite software version 5.0, and an expression value (signal) of each probe set was calculated and normalised, such that the mean of signal values in each experiment was 100, to adjust for minor differences between experiments. Statistical analyses and fold change calculations were performed using expression values that were log-transformed after the addition of 10. Hierarchical clustering analysis and matrix presentation were performed using CLUSTER and TREE VIEW software.

To validate selected genes, microarray data of adult AML reported by Valk et al (2004) were down-loaded from the NCBI Gene Expression Omnibus database (GSE 1159 in http://www.ncbi.nlm.nih.gov/projects/geo/), and the data of 222 AML and eight normal samples were used after excluding data from 57 samples whose FAB subtype or karyotype information was not given. Promoter analysis, including prediction of a transcription initiation site, extraction of genomic sequence around the initiation site, and assignment of transcription factor-binding sites, was performed using the GENOMATIX-SUITE software (Genomatix, Munich, Germany). To assign AML1-, POU4F1-, and HOXB2-binding sites, V$AML1.01, V$BRN3.01, and V$HOXA9.01 matrixes in the SOFTWARE were used respectively.

RT-PCR analysis

For semi-quantitative RT-PCR analysis, cDNA was prepared from 0·5 μg of total RNA, and one fiftieth of the cDNA was used as template for each PCR reaction. Forward and reverse primers were designed using GENETYX MAC 9.0/SEARCH PRIMER software. Sequences of forward and reverse primers were: *CD34*, 5′-ATTTCCTGATGAATCGCCGC-3′ and 5′-GCCTTTCCCTGA-GCCTCAGG-3′; *CAV1*, 5′-ACCTCAACGATGACGTGGTC-3′ and 5′-CAAGTTGATGCGGACATTGC-3′; *CLIPR-59*, 5′-GTC-TTCGCACCAGCATCCCG-3′ and 5′-AGGTTTCTGATCCA-GGGTTG-3′; and *HOXA9*, 5′-GCACCGCTTTTTCCGAGTG-3′

and 5'-GCGGTGTACCACCACCATC-3'. PCR using Amplitaq Gold (Applied Biosystems, Foster City, CA, USA) was performed using conditions appropriate for each transcript. PCR products were run on agarose gels.

Results

Selection of t(8;21)- and inv(16)-associated genes

A goal of this study was to extract genes whose expression was modulated in the presence of t(8;21) and inv(16) and evaluate how many genes were similarly up or downregulated in t(8;21)- and inv(16)-AML. For this purpose, we used microarray gene expression data of 50 paediatric AML patients including eight cases of t(8;21) and seven cases of inv(16).

Initially, unsupervised hierarchical clustering analysis was performed as shown in Fig 1, in which samples are identified together with their FAB-subtype. All eight t(8;21) samples and six of seven inv(16) samples (except sample S12) were clustered as distinct groups, indicating that t(8;21)- and inv(16)-AML constitute an independent AML subgroup. In addition, neighbouring of t(8;21) and inv(16) clusters suggests that there may exist a common element in their gene expression.

To extract t(8;21)- and inv(16)-associated genes, we first categorised AML samples into five groups, G1–G5, according to FAB subtype and karyotype [G1, M2 with t(8;21); G2, other M2; G3, M4 with inv(16); G4, other M4; and G5, other FAB subtypes]. Sample numbers in G1, G2, G3, G4 and G5 were 8, 5, 7, 7 and 23 respectively. The method of gene extraction
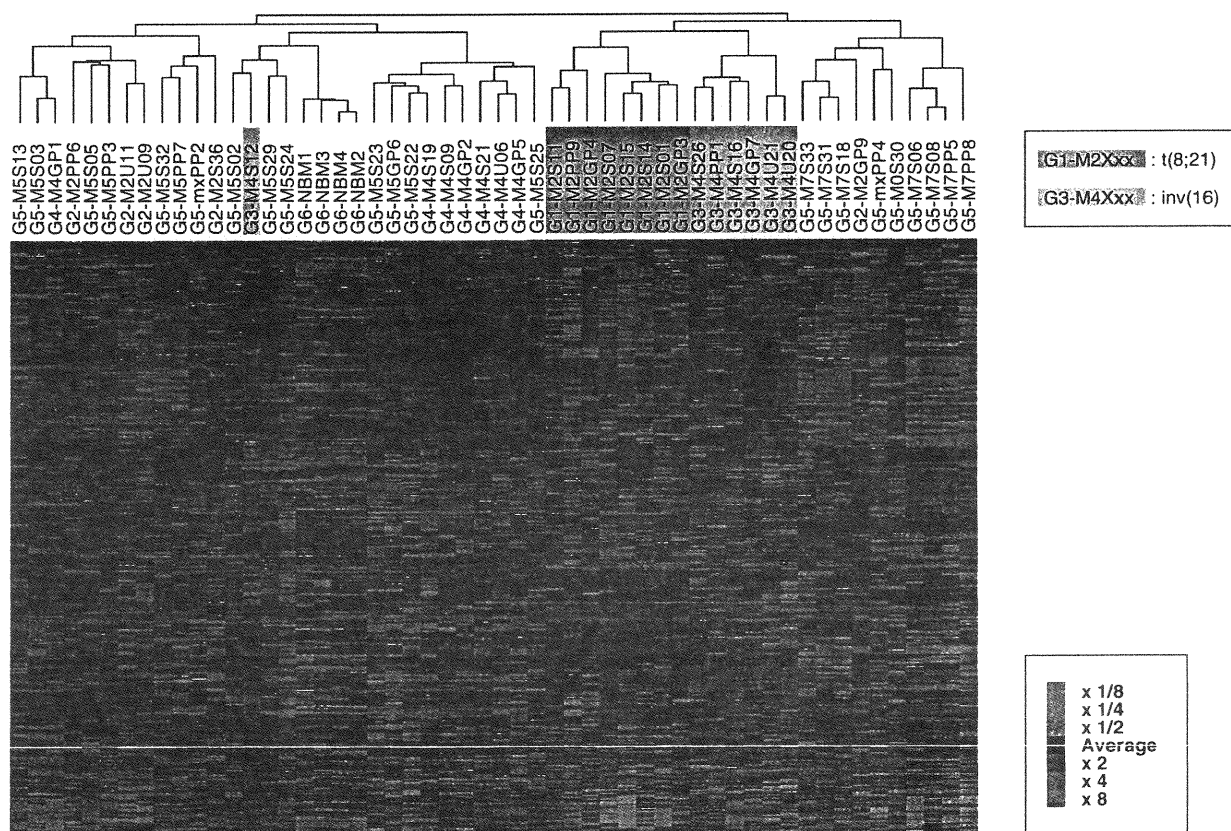

Fig 1. Unsupervised two-dimensional hierarchical clustering analysis of 50 acute myeloid leukaemia (AML) and four normal bone marrow samples. Expression data of 804 probe sets that were determined by the following two criteria were used. Namely, their median expression in AML samples was twofold higher compared with that in normal bone marrow samples, and their expression in the 5th highest AML sample showed more than a fivefold difference compared with those in the lowest 5th AML sample. Thus, these 804 sets represent those whose expression was highly varied across the samples. Each row represents a respective probe set, and each column a respective sample. Relative expression levels normalised to the average for each probe set were indicated by colour, where red and green represent increased and decreased expression respectively. At the top, the group number (G), FAB-classification (M) and the identification of the AML sample are indicated. Relationships between samples are shown by dendrograms. It must be noted that clustering analyses were performed using different sets of probes that were selected by varying the criteria. In most cases of those analyses, t(8;21) and inv(16) samples were clustered as distinct groups, and they often neighboured to each other. Peculiarly, the inv(16) sample, S12, always behaved as an outlier in the shown as well as not-shown clustering analyses. The reason for this is not obvious, and we did not notice any particular clinical nor karyotypic features of S12 other than AML-M4 with inv(16).

employed is shown schematically in Fig 2A. First, G1 was compared with G2 and also with a combined group of G2 + G4 + G5 and extracted transcripts whose average expression in G1 was more than twofold higher with a $P < 0.01$ according to the Student's t-test. As both G1 and G2 belong to the same M2-subtype, transcripts extracted by comparing G1 and G2 are considered to represent those associated with the t(8;21) abnormality but not with the M2-subtype. Transcripts extracted by these two rounds of comparison were defined as t(8;21)-associated highly expressed transcripts. Similarly, inv(16)-associated highly expressed transcripts were extracted by two rounds of comparison between G3 and G4, and between G3 and the G2 + G4 + G5 group. t(8;21)- and inv(16)-associated low expression transcripts, whose average expression was more than twofold lower with a $P < 0.01$, were also extracted. G3 samples were not included to extract t(8;21)-associated transcripts, and G1

samples were not used to extract inv(16)-associated transcripts. Thus, t(8;21)- and inv(16)-associated genes were selected independently of each other. As summarised in Fig 2B, 59 t(8;21)-associated highly expressed transcripts, 58 inv(16)-associated highly expressed transcripts, 15 t(8;21)-associated low expression transcripts, and 18 inv(16)-associated low expression transcripts were selected.

Selection of commonly and specifically regulated genes

The gene extraction method shown in Fig 2A identified six highly expressed and two low expression transcripts associated with both t(8;21) and inv(16) (Fig 2B). A review of the extracted genes, however, suggested that the numbers of commonly regulated transcripts was probably underestimated. For example, 19 of the 58 inv(16)-associated highly expressed transcripts showed more than twofold greater expression not
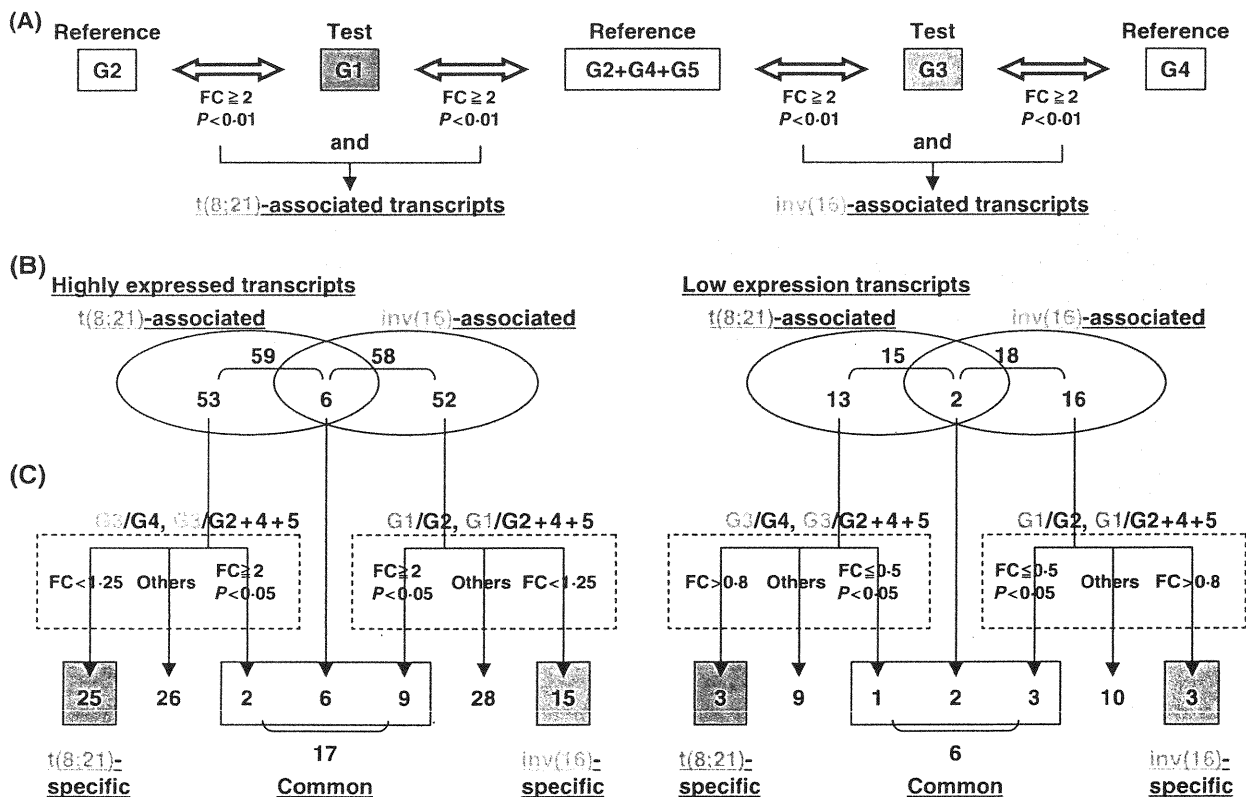
Fig 2. (A) Schematic illustration of the gene extraction procedure. Test (G1 or G3) and reference (G2, G4 or G2 + G4 + G5) samples were subjected to pair-wise comparisons as indicated by the double-headed arrows. To extract t(8;21)-associated transcripts, comparison was made between G1 and G2 and between G1 and G2 + G4 + G5. Genes extracted by two rounds of comparison were defined as t(8;21)-associated. Inv(16)-associated transcripts were similarly extracted. Criteria of gene extraction were that average expression of test samples was significantly ($P < 0.01$) greater than twofold increased or decreased compared with reference samples. (B) Venn diagram comparison of transcripts extracted by the method described in A. Indicated are numbers of transcripts belonging to t(8;21)-associated highly expressed, inv(16)-associated highly expressed, t(8;21)-associated low expression and inv(16)-associated low expression transcripts, consisting of 59 (53 + 6), 58 (52 + 6), 15 (13 + 2) and 18 (16 + 2) respectively. (C) Modification of the extraction procedure. The starting material was genes shown in panel B, and each criterion defined in broken rectangles was applied to them. As a result, 17 common, 25 t(8;21)-specific and 15 inv(16)-specific highly expressed transcripts, and 6 common, 3 t(8;21)-specific and 3 inv(16)-specific low expression transcripts were reclassified. The numbers of transcripts shown in panels B and C represent numbers of probe sets.

only in G3 but also in G1 (data not shown). This is not surprising, as the extraction procedure described above adopted the strict standard of $P < 0.01$. Therefore, we modified the gene extraction procedure as shown in Fig 2C with the goal of selecting transcripts both commonly and specifically up or downregulated in the presence of t(8;21)/inv(16). Therefore, in the common group, transcripts showing more than a twofold increase or decrease in expression at $P < 0.05$ rather than $P < 0.01$ were added to the previously described six and two common transcripts. Thus, 17 and 6 transcripts were defined as commonly increased and decreased in expression respectively. In addition, transcripts whose average expression in G3 was more than 1.25-fold higher compared with G4 or G2 + G4 + G5 were excluded from the t(8;21)-associated highly expressed transcripts, and the remaining 25 transcripts were defined as t(8;21)-specific highly expressed transcripts. Using similar modifications, inv(16)-specific highly expressed transcripts ($n = 15$), t(8;21)-specific low expression transcripts ($n = 3$), and inv(16)-specific low expression transcripts ($n = 3$) were defined.

According to our criteria, 8 of 59 t(8;21)-associated highly expressed transcripts were also highly expressed in inv(16)-AML, and 3 of the 15 t(8;21)-associated low expression transcripts were also low expression transcripts in inv(16)-AML. Similarly, 15 of 58 inv(16)-associated highly expressed transcripts were also highly expressed in t(8;21)-AML, and five of 18 inv(16)-associated low expression transcripts were also of low expression in t(8;21)-AML. Overall, 15% (11/74) of t(8;21)-associated transcripts exhibited similar expression in inv(16)-AML, and 26% (20/76) of inv(16)-associated transcripts did so in t(8;21)-AML. On the other hand, 38% (25 + 3/74) of t(8;21)-associated and 24% (15 + 3/76) of inv(16)-associated transcripts appeared to be regulated in t(8;21)- and inv(16)-specific manners respectively. In summary, these results indicate that there exists a significant number of commonly regulated transcripts in addition to those regulated specifically by either t(8;21)- or inv(16)-AML. Although the genes whose expression was unique to either t(8;21)- or inv(16)-AML have been reported repeatedly, common gene expression signatures to both AML-subtypes are the first demonstration.

It must be noted that in some cases multiple probe sets were used for a specific gene. Taking redundancy into account, the numbers of individually selected genes were as follows: 15 common, 21 t(8;21)-specific and 13 inv(16)-specific genes as highly expressed; and 6 common, 3 t(8;21)-specific and 3 inv(16)-specific genes as low expressed. Table I details a list of these genes and includes the ratios of average expression values.

Evaluation of selected genes

Expression data of selected genes were processed for matrix presentation, as shown in Fig 3. For this analysis samples and probe sets were aligned by alphabetical and numerical order respectively, within the each group. Each member of G1 and G3 groups was clearly separated from other groups in terms of the expression of selected genes. Also, t(8;21)-specific and inv(16)-specific as well as commonly regulated genes, showed patterns unique to the relevant groups, indicating that the method of gene extraction functioned effectively.

We next investigated whether the expression evaluated by the microarray analysis reflected RNA levels. To do so, four representative transcripts were chosen: *CD34* from the common highly expressed group, *CAV1* from t(8;21)-specific highly expressed genes, *CLIPR-59* from inv(16)-specific highly expressed genes, and *HOXA9* from the common low expression group. cDNA was synthesised from RNA from 17 samples and processed for semi-quantitative RT-PCR (Fig 4). Overall the relative level of each transcript in a respective sample paralleled the relative expression values obtained by microarray analysis (see also the legend to Fig 4).

A control analysis was performed to evaluate the significance of common gene expression signatures that were found in t(8;21)- and inv (16)-AML. This was done by examining how many genes might be selected as commonly regulated between G2 and G4 subgroups. Each of G2 and G4 were used as test samples, and G1, G3 and G1 + G3 + G5 were used as reference samples. The gene extraction procedure used was (Fig S1) similar to that shown in Fig 2, and the selected genes are listed in Table SII. Only three transcripts (two genes) were selected as commonly high expressed, and no transcript was extracted as commonly low expressed (5 G2-specific and 26 G4-specific highly expressed transcripts and 10 G2-specific and 11 G4-specific low expression transcripts were selected at the same time). The gene number common to G2 and G4 was much smaller compared with that common to t(8;21) and inv(16) [23 transcripts (21 genes)]. This indicates the following: firstly, very few common elements exist in gene expression of G2- and G4-subgroups. Secondly, the above described 23 transcripts were not selected by chance but probably reflect common features of t(8;21)- and inv(16)-AML.

Validation of selected genes using a different set of microarray data

We next evaluated whether the selected genes were valid indicators of t(8;21)- and inv(16)-AML activities. To do so we employed another set of microarray data from AML patients reported by Valk *et al* (2004). Their data set consisted of 285 patients and contained information on FAB-classification and karyotype for each AML sample. We used the data of 222 of those patients and excluded the other 63 due to lack of FAB subtype and karyotype information. According to our classification, the numbers of AML samples were 20 in G1, 35 in G2, 14 in G3, 31 in G4 and 122 in G5.

Figure 5 shows a matrix presentation of the expression data. Each of the aforementioned six categories of genes again appeared to behave as a distinct cluster, suggesting that most genes selected were indeed modulated in the presence of

Table I. Transcripts whose expression was modulated commonly to t(8;21)- and inv(16)-AML and specifically to each.

Group	Probe set	Public ID	Gene symbol	Gene name	G1/G2	G1/G2 + 4 + 5	G3/G4	G3/G2 + 4 + 5
High	Common	36650_at	D13639	CCND2	Cyclin D2			
High	Common	38747_at	M81945	CD34	CD34 antigen			
High	Common	538_at	S53911	CD34	CD34 antigen			
High	Common	479_at	U53446	DAB2	Disabled homologue 2, mitogen-responsive phosphoprotein			
High	Common	37762_at	Y07909	EMP1	Epithelial membrane protein 1			
High	Common	38052_at	M14539	F13A1	Coagulation factor XIII, A1 polypeptide			
High	Common	39070_at	U03057	FSCN1	Fascin homologue 1			
High	Common	38833_at	X00457	HLA-DPA1	Major histocompatibility complex, class II, DP alpha 1			
High	Common	38095_i_at	M83664	HLA-DPB1	Major histocompatibility complex, class II, DP beta 1			
High	Common	38096_f_at	M83664	HLA-DPB1	Major histocompatibility complex, class II, DP beta 1			
High	Common	32773_at	AA868382	HLA-DQA2	Major histocompatibility complex, class II, DQ alpha 2			
High	Common	41723_s_at	M32578	HLA-DRB1	Major histocompatibility complex, class II, DR beta 1			
High	Common	37749_at	D78611	MEST	Mesoderm specific transcript homologue (mouse)			
High	Common	37283_at	X82209	MN1	Meningioma (disrupted in balanced translocation) 1			
High	Common	35523_at	AF150241	PGDS	Prostaglandin D2 synthase, haematopoietic			
High	Common	32905_s_at	M30038	TPSAB1	Tryptase alpha/beta 1			
High	Common	32323_at	M63582	TRH	Thyrotropin-releasing hormone			
High	t(8;21)-specific	34512_at	J03853	ADRA2C	Adrenergic, alpha-2C, receptor			1·00	0·98
High	t(8;21)-specific	37543_at	D25304	ARHGEF6	Rac/Cdc42 guanine nucleotide exchange factor (GEF) 6			0·87	1·00
High	t(8;21)-specific	41690_at	AL049471	ARID5B	AT rich interactive domain 5B (MRF1-like)			1·16	1·03
High	t(8;21)-specific	36119_at	AF070648	CAV1	Caveolin 1, caveolae protein, 22 kDa			1·09	0·99
High	t(8;21)-specific	39038_at	AF093118	FBLN5	Fibulin 5			1·07	0·99
High	t(8;21)-specific	37251_s_at	AF016004	GPM6B	Glycoprotein M6B			0·65	0·82
High	t(8;21)-specific	32845_at	M85289	HSPG2	Heparan sulfate proteoglycan 2 (perlecan)			1·10	1·05
High	t(8;21)-specific	406_at	X53587	ITGB4	Integrin, beta 4			1·17	1·12
High	t(8;21)-specific	33226_at	AB020683	JMJD2B	Jumonji domain containing 2B			0·58	0·74
High	t(8;21)-specific	39237_at	U43784	MAPKAPK3	Mitogen-activated protein kinase-activated protein kinase 3			1·16	1·10
High	t(8;21)-specific	38803_at	AF052142	NCALD	Neurocalcin delta			1·20	1·14
High	t(8;21)-specific	40081_at	L26232	PLTP	Phospholipid transfer protein			1·03	1·07
High	t(8;21)-specific	36980_at	U03105	PNRC1	Proline-rich nuclear receptor coactivator 1			1·03	1·22
High	t(8;21)-specific	35939_s_at	L20433	POU4F1	POU domain, class 4, transcription factor 1			1·09	1·05
High	t(8;21)-specific	35940_at	X64624	POU4F1	POU domain, class 4, transcription factor 1			1·05	0·93
High	t(8;21)-specific	32176_at	AB011110	RASA4	RAS p21 protein activator 4			1·03	0·83
High	t(8;21)-specific	35638_at	D43638	RUNX1T1/MTG8	Runt-related transcription factor 1/MTG8			0·95	0·89
High	t(8;21)-specific	36192_at	D83777	SCRN1	Secernin 1			0·48	0·50
High	t(8;21)-specific	41246_at	AL040655	SERPINE2	Serine (or cysteine) proteinase inhibitor, member 2			1·25	0·98
High	t(8;21)-specific	38118_at	U73377	SHC1	SHC transforming protein 1			0·87	0·93
High	t(8;21)-specific	38997_at	X96924	SLC25A1	Solute carrier family 25, member 1			0·80	0·91

Table I. Continued.

Group	Probe set	Public ID	Gene symbol	Gene name	Ratio			
					G1/G2	G1/G2 + 4 + 5	G3/G4	G3/G2 + 4 + 5
High	t(8;21)-specific	38998_g_at	X96924	*SLC25A1*	Solute carrier family 25, member 1			0·83	0·91
High	t(8;21)-specific	39825_at	U25147	*SLC25A1*	Solute carrier family 25, member 1			0·87	0·96
High	t(8;21)-specific	1953_at	AF024710	*VEGF*	Vascular endothelial growth factor		2·46	0·48	0·66
High	t(8;21)-specific	36100_at	AF022375	*VEGF*	Vascular endothelial growth factor		2·04	0·64	0·80
High	inv(16)-specific	34246_at	AA418437	*C6orf145*	Chromosome 6 open reading frame 145	0·91	0·95		
High	inv(16)-specific	33569_at	D50532	*CLEC10A*	C-type lectin domain family 10, member A	0·67	0·85		
High	inv(16)-specific	36095_at	N99340	*CLIPR-59*	CLIP-170-related protein	1·14	1·10		
High	inv(16)-specific	40936_at	AI651806	*CRIM1*	Cysteine-rich motor neuron 1	0·46	0·84		
High	inv(16)-specific	37187_at	M36820	*CXCL2*	Chemokine (C-X-C motif) ligand 2	0·59	1·03		
High	inv(16)-specific	39610_at	X16665	*HOXB2*	Homeo box B2	0·61	0·72		
High	inv(16)-specific	2058_s_at	M35011	*ITGB5*	Integrin, beta 5	1·13	0·99		
High	inv(16)-specific	32582_at	X69292	*MYH11*	Myosin, heavy polypeptide 11, smooth muscle	0·92	0·92		
High	inv(16)-specific	37407_s_at	AF013570	*MYH11*	Myosin, heavy polypeptide 11, smooth muscle	1·05	1·08		
High	inv(16)-specific	31886_at	X55740	*NT5E*	5'-nucleotidase, ecto (CD73)	1·06	1·04		
High	inv(16)-specific	38862_at	Y11215	*SCAP1*	src family associated phosphoprotein 1	0·74	1·04		
High	inv(16)-specific	2092_s_at	J04765	*SPP1*	Secreted phosphoprotein 1 (osteopontin)	0·83	0·84		
High	inv(16)-specific	34342_s_at	AF052124	*SPP1*	Secreted phosphoprotein 1 (osteopontin)	0·56	0·54		
High	inv(16)-specific	41531_at	AI445461	*TM4SF1*	Transmembrane 4 superfamily member 1	0·73	1·17		
High	inv(16)-specific	892_at	M90657	*TM4SF1*	Transmembrane 4 superfamily member 1	0·68	0·97		
Low	Common	40585_at	D25538	*ADCY7*	Adenylate cyclase 7	0·39	0·30	0·39	0·48
Low	Common	37105_at	M16117	*CTSG*	Cathepsin G	0·07	0·26	0·03	0·14
Low	Common	41448_at	AC004080	*HOXA10*	Homeo box A10	0·24	0·34	0·33	0·47
Low	Common	873_at	M26679	*HOXA5*	Homeo box A5	0·40	0·35	0·18	0·32
Low	Common	37809_at	U41813	*HOXA9*	Homeo box A9	0·09	0·11	0·05	0·15
Low	Common	32193_at	AF030339	*PLXNC1*	Plexin C1	0·32	0·30	0·24	0·31
Low	t(8;21)-specific	36802_at	M23197	*CD33*	CD33 antigen (gp67)	0·46	0·50	0·89	0·92
Low	t(8;21)-specific	39664_at	U28413	*ERCC8*	Excision repair deficiency, complementation group 8	0·37	0·50	1·32	0·95
Low	t(8;21)-specific	1463_at	M93425	*PTPN12*	Protein tyrosine phosphatase, non-receptor type 12	0·43	0·47	0·87	1·08
Low	inv(16)-specific	41175_at	L20298	*CBFB*	Core-binding factor, beta subunit	1·26	1·15	0·25	0·30
Low	inv(16)-specific	35282_r_at	M33680	*CD81*	CD81 antigen		1·07	0·26	0·29
Low	inv(16)-specific	41396_at	AB006629	*CYLN2*	Cytoplasmic linker 2	1·05	1·01	0·50	0·35

Values of ratio highlighted by pink and blue represent more than twofold high and low expression respectively.
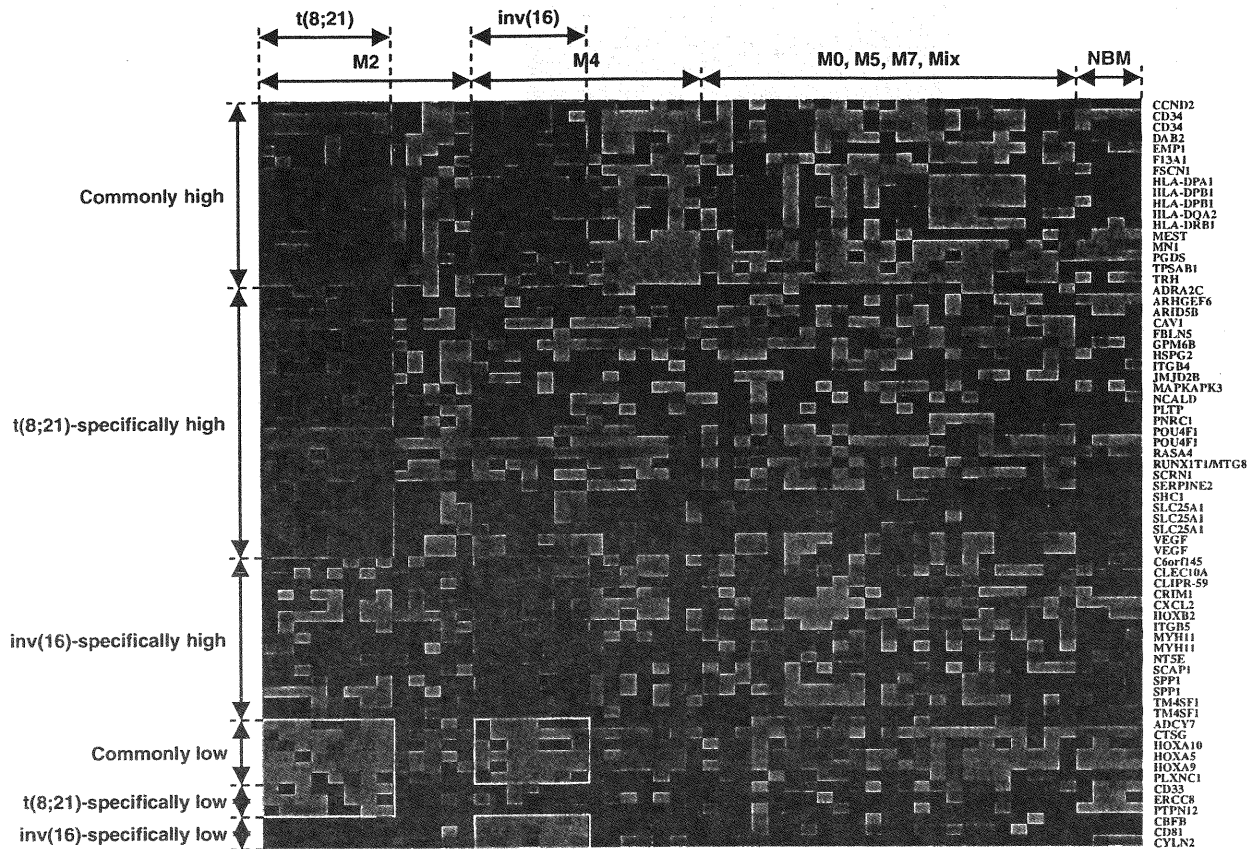
Fig 3. Matrix presentation of expression of 69 selected transcripts representing 57 high- and 12 low-expression examples in 50 AML and four normal bone marrow samples. Each row represents a respective gene, and each column is a respective patient. Relative levels of expression are indicated by colour, where red and green represent increased and decreased expression respectively. FAB classification is indicated at the top, and grouping of gene expression is on the left. Gene symbols are shown on the right.

t(8;21) and inv(16) activities. Genes selected as inv(16)-specific highly expressed, however, did not yield such reproducible results. Half appeared to be high in the G3 samples, whereas the other half did not. Although the reason for this discrepancy is not clear, the samples used in our study were from paediatric patients, whereas those in Valk et al (2004) were from adults. In the case of inv(16)-AML, factors regulating gene expression may differ between childhood and adulthood AML (see Fig 5 legend regarding the comparison of our analysis and that of Valk et al (2004).

Analysis of AML1-binding sites in a promoter region

Finally, we investigated whether the promoter region of a selected gene harbors AML1-binding sites. Only highly expressed genes were examined, as the number of low expression genes was relatively small. Putative AML1-binding sites were evaluated in a genomic region covering −2000 to +500 bp with respect to the predicted transcription initiation site, using GENOMATIXSUITE software. We determined the number of AML1-binding sites for a specific promoter and calculated the average ± standard deviation for highly

expressed genes of the common, t(8;21)-specific and inv(16)-specific groups (Table II). As references, 44 and 46 probe sets were used whose expression values showed approximately the average and median respectively, of all probe sets on the microarray. As seen in Table II, a significant difference was not detected in the average number of AML1-binding sites between promoters of selected genes and those of the reference genes. POU4F1 and HOXB2, which encode transcription factors, were extracted as a t(8;21)-specific highly expressed gene and an inv(16)-specific highly expressed gene respectively. The numbers of POU4F1- and HOXB2-binding sites in t(8;21)-specific and inv(16)-specific highly expressed genes were the same as those seen in the reference genes. This result suggests that the number of putative AML1-, POU4F1- and HOXB2-binding sites in the promoter region cannot account for the expression levels of the extracted genes.

Discussion

To compare gene expression between t(8;21)- and inv(16)-AML, AML samples were classified as groups G1–G5 and a pair-wise comparison was performed between these groups.

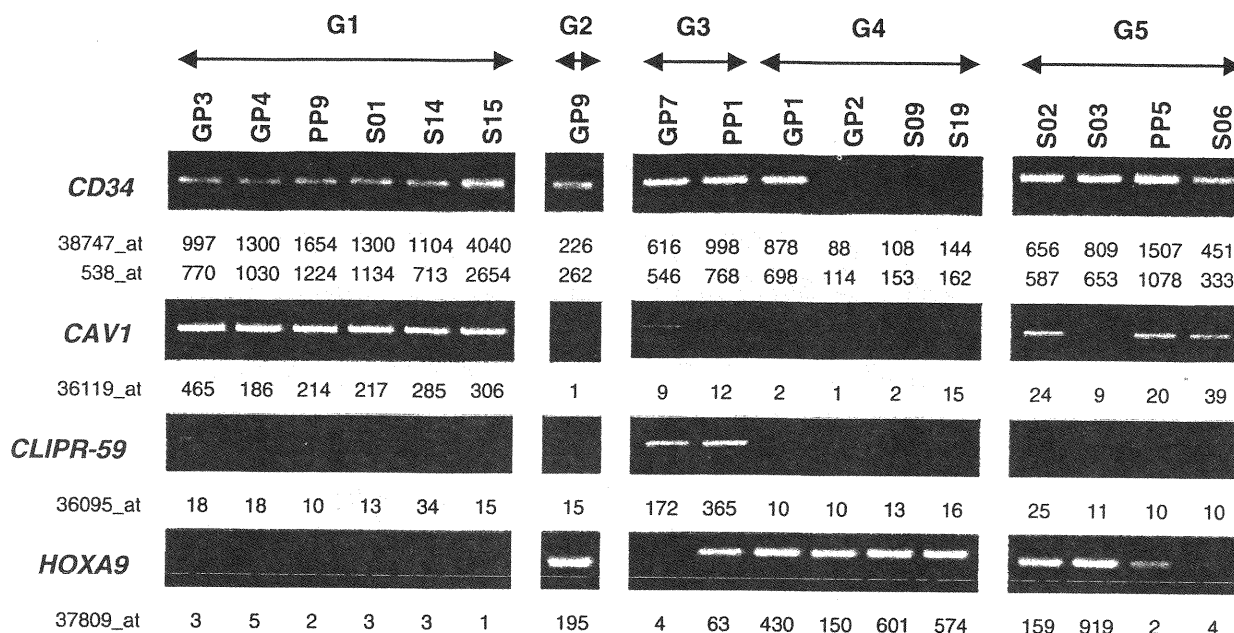|	G1						G2	G3		G4				G5			
	GP3	GP4	PP9	S01	S14	S15	GP9	GP7	PP1	GP1	GP2	S09	S19	S02	S03	PP5	S06
CD34																	
38747_at	997	1300	1654	1300	1104	4040	226	616	998	878	88	108	144	656	809	1507	451
538_at	770	1030	1224	1134	713	2654	262	546	768	698	114	153	162	587	653	1078	333
CAV1																	
36119_at	465	186	214	217	285	306	1	9	12	2	1	2	15	24	9	20	39
CLIPR-59																	
36095_at	18	18	10	13	34	15	15	172	365	10	10	13	16	25	11	10	10
HOXA9																	
37809_at	3	5	2	3	3	1	195	4	63	430	150	601	574	159	919	2	4

Fig 4. Semi-quantitative RT-PCR. Relative amounts of *CD34*, *CAV1*, *CLIPR-59* and *HOXA9* transcripts were compared with G1, G2, G3, G4 and G5 groups. At the top, the group number (G) and the identification of the AML samples are indicated. Microarray expression values (signal values) are also shown beneath the gels pictures. The G5 samples shown here appear to be rather exceptional cases in terms of their expression of *CD34* and *CAV1*. Due to a lack of sufficient amount of RNA, we could not process other than the samples shown for RT-PCR.

Because of the combinations of test and reference samples used here, the genes identified are probably associated with t(8;21) and inv(16), but not with the lineage/stage specificity of leukaemic cells as exemplified by the FAB-classification. The existence of a gene expression signature characteristic of CBF leukaemia was reported previously (see Fig 4 in Ross *et al*, 2004, although a common element in t(8;21) and inv(16) is not immediately clear). However, our approach is unique in that the t(8;21)- and inv(16)-associated genes were selected independently of inv(16)- and t(8;21)-AML respectively, and commonly modulated genes were selected by comparing t(8;21)- and inv(16)-associated genes. Thus, we demonstrate that t(8;21)- and inv(16)-AML exhibit significant overlap in gene expression signatures. This result indicates that AML1-MTG8 and CBFβ-MYH11 chimaeric proteins affect a common set of targets in leukaemic cells. In addition, our method identifies new genes not previously reported.

Detection of commonly regulated genes agrees with the notion that both AML1-MTG8 and CBFβ-MYH11 exert a similar dominant negative effect on wild type CBF. Furthermore, our identification of specifically regulated genes is in line with the prediction that each chimaeric protein may also have a unique activity. Promoter analysis of selected genes, however, suggests that regulation is complex. The number of putative AML1-binding sites in a predicted promoter region did not differ significantly between selected and reference genes, regardless of the type of genes selected as common or specific highly expressed genes. This observation was also true for the POU4F1- and HOXB2-binding sites for t(8;21)- and inv(16)-

specific highly expressed gene promoters respectively. Several mechanisms could explain the relationship between transcriptional activity of a chimaeric protein and regulation of gene expression. One is that these proteins regulate promoters through sites other than AML1 binding sites by interacting with other transcription factors and/or co-factors. For example, AML1-MTG8 interacts with p300/CBP and interferes with transcription mediated by an E-box transcription factor whose binding site is distinct from that of AML1 (Zhang *et al*, 2004). Alternatively, although both chimaeric proteins could target the promoters of the same set of genes, environmental factors, such as cell lineage or developmental stage, that influence the magnitude of gene expression may vary for each chimaera. A chimaeric protein may also target an unidentified molecule, which in turn could modulate expression of extracted genes.

Of selected genes, *MTG8* and *MYH11* are of interest. The probe sets for these genes correspond to the *MTG8* portion of the *AML1-MTG8* chimaeric transcript and to the *MYH11* portion of the *CBFβ-MYH11* chimaeric transcript respectively. The fact that the probe sets for *MTG8* and *MYH11* were selected as t(8;21)- and inv(16)-specific highly expressed genes respectively, indicates that our gene extraction procedure worked efficiently. Expression of *MTG8* and *MYH11* have been assigned repeatedly as the most characteristic signatures of t(8;21) and inv(16) respectively, in other microarray analyses (Schoch *et al*, 2002; Debernardi *et al*, 2003; Kohlmann *et al*, 2003; Bullinger *et al*, 2004; Ross *et al*, 2004; Valk *et al*, 2004; Gutierrez *et al*, 2005). Immunophenotyping studies of t(8;21)- and inv(16)-leukaemic cells document CD34 and HLA-DR as
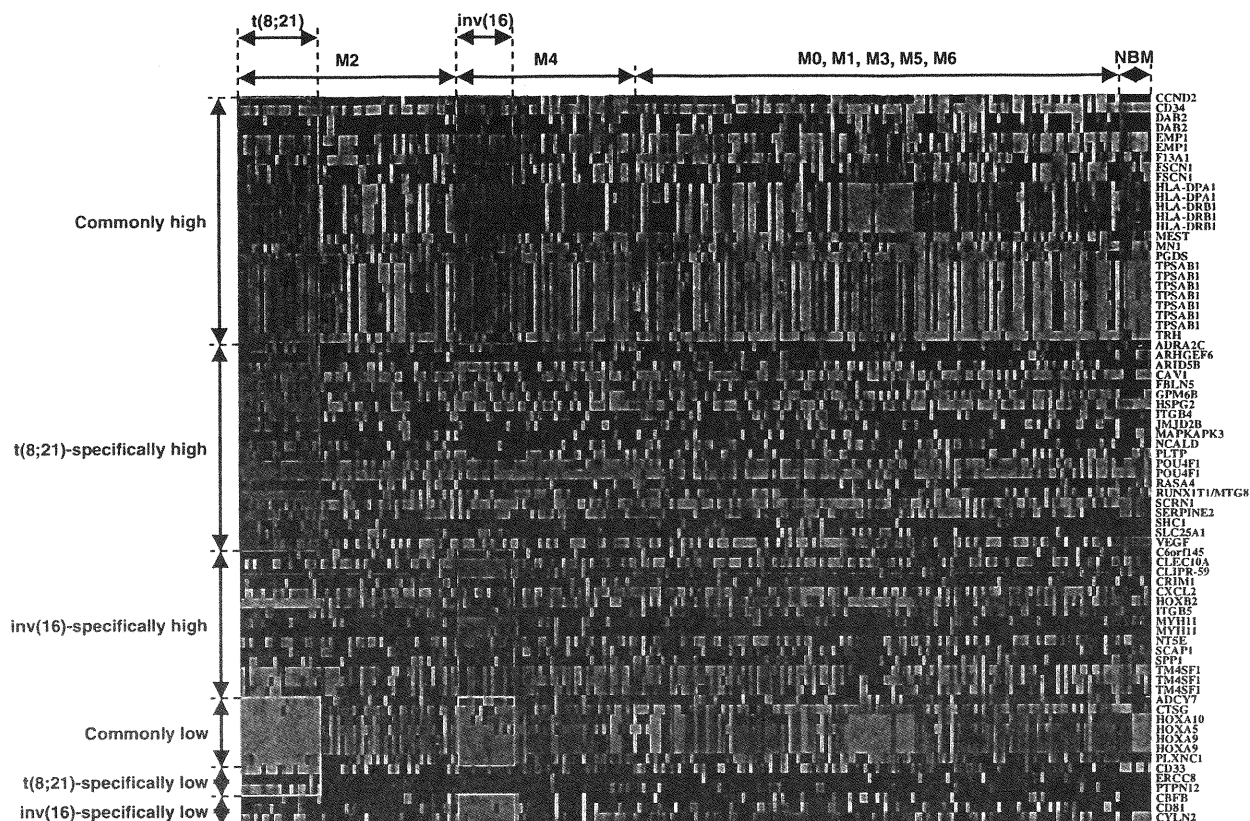
Fig 5. Validation of selected genes using microarray data of adult acute myeloid leukaemia (AML) patients obtained using Affymetrix Human Genome U133A microarray provided by Valk *et al* (2004). The 222 AML samples consisted of 20 G1, 35 G2, 14 G3, 31 G4, and 122 G5 samples, and there were eight normal bone marrow samples including four fractionated CD34$^+$ cell samples. Relative expression levels are indicated by colour, where red and green represent increased and decreased expression respectively. FAB-classification is at the top, and gene expression grouping is indicated on the left. Gene symbols are shown on the right. Valk *et al* (2004) reported the identification of t(8;21)- and inv(16)-unique transcripts in their Supplementary Tables M1 and I1. We compared their Tables with our Table I and evaluated the degree of overlap of selected genes. Among the 26 t(8;21)- and 28 inv(16)-unique genes reported by Valk *et al* (2004) (their original 40 and 40 probes shown in their Tables were reclassified to 26 and 28 genes respectively, after taking into consideration the redundancy of probes for a specific gene and a difference of their U133A and our U95Av2 microarray), 10 and 6 were also included in our list of t(8;21)- and inv(16)-specific genes respectively. This suggests that a significant number of identical genes were selected by two different analyses. In contrast, only one from 26 and one from 28 genes were contained in our gene list common to both t(8;21) and inv(16). This indicates that our method used for extracting common signatures worked efficiently.

Table II. Number of transcription factor-binding sites per gene.

Group	No. of genes	AML1	POU4F1	HOXB2
Commonly high expressed (A)	15	1·60 ± 1·66	3·07 ± 2·57	1·13 ± 0·88
t(8;21)-specifically high expressed (B)	17	1·24 ± 1·11	2·33 ± 2·47	1·56 ± 1·83
inv(16)-specifically high expressed (C)	10	1·10 ± 1·30	2·10 ± 2·02	2·10 ± 1·70
A + B + C	42	1·33 ± 1·39	2·57 ± 2·46	1·55 ± 1·59
Around average	44	1·64 ± 3·03	3·50 ± 4·60	2·68 ± 4·40
Around median	46	1·43 ± 1·33	4·37 ± 5·25	1·59 ± 2·00

Transcription factor-binding sites were predicted using the GENOMATIXSUITE software 1. The average number of sites per gene was calculated and presented together with SD for each group of genes.

specific markers that may reflect the immaturity of cells (Hurwitz *et al*, 1992; Osato *et al*, 1997). Identification of corresponding transcripts as common highly expressed genes is probably the basis of this immunophenotype of cells. High expression of *CCND2* also may confer a growth advantage on leukaemic cells, although ectopic expression of CBFβ-MYH11 in 32Dcl3 and Ba/F3 cell lines has been reported to retard the G1 to S transition and stimulate expression of *p21WAF1* (Cao *et al*, 1997; Lou *et al*, 2000). It also is noteworthy that many genes related to endothelial cells and signal transduction were extracted as t(8;21)-specific highly expressed genes. The endothelial group includes *VEGF*, *FBLN5*, and *ITGB4*, whereas the genes that encoded signalling factors were *VEGF*, *ADRA2C*, *SHC1*, *CAV1*, *RASA4*, *ARHGEF6* and *MAPKAPK3*. Also potentially critical are genes that were extremely highly expressed, including *TRH*, *TPSAB1*, *MN1*, *POU4F1* and *CAV1*.

Among the low expression genes, *HOXA5*, *HOXA9* and *HOXA10* were categorised as common genes. Enhanced expression of *HOXA9* is well established in many AML-subtypes including those with normal karyotypes (Lawrence *et al*, 1999; Debernardi *et al*, 2003; Bullinger *et al*, 2004; Gutierrez *et al*, 2005), and a *NUP98-HOXA9* chimaera is generated in t(7;11)-AML (Nakamura *et al*, 1996). Nevertheless, the average level of *HOXA9* transcripts in G1 and G3 was as low as 10% of that seen in other groups (evidenced by quantification of the results in Fig 4). Thus, an interesting possibility is that a pathway mediated by HOXA9 is not necessary or may even be antagonistic to the molecular mechanisms involved in t(8;21)- and inv(16)-AML. Low expression of *CBFβ* in G3 but not G1 may mean that the CBFβ-MYH11 protein negatively autoregulates expression of the wild type *CBFβ* allele.

In conclusion, we extracted and categorised t(8;21)/inv(16)-associated genes as t(8;21)-specific, inv(16)-specific, and genes common to both. This categorisation was enabled by a unique approach to gene extraction. While the presence of specific groups of selected genes correlated with the bimodular structures of the chimaeric proteins, it was notable t(8;21)- and inv(16)-AML display a significant degree of overlap in their gene expression signatures.

Acknowledgements

References

Bullinger, L., Dohner, K., Bair, E., Frohling, S., Schlenk, R.F., Tibshirani, R., Dohner, H. & Pollack, J.R. (2004) Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *The New England Journal of Medicine*, 350, 1605–1616.

Cao, W., Britos-Bray, M., Claxton, D.F., Kelley, C.A., Speck, N.A., Liu, P.P. & Friedman, A.D. (1997) CBFβ-SMMHC, expressed in M4Eo AML, reduced CBF DNA-binding and inhibited the G1 to S cell cycle transition at the restriction point in myeloid and lymphoid cells. *Oncogene*, 15, 1315–1327.

Castilla, L.H., Wijmenga, C., Wang, Q., Stacy, T., Speck, N.A., Eckhaus, M., Marin-Padilla, M., Collins, F.S., Wynshaw-Boris, A. & Liu, P.P. (1996) Failure of embryonic hematopoiesis and lethal hemorrhages in mouse embryos heterozygous for a knocked-in leukemia gene *CBFβ-MYH11*. *Cell*, 87, 687–696.

Debernardi, S., Lillington, D.M., Chaplin, T., Tomlinson, S., Amess, J., Rohatiner, A., Lister, T.A. & Young, B.D. (2003) Genome-wide analysis of acute myeloid leukemia with normal karyotype reveals a unique pattern of homeobox gene expression distinct from those with translocation-mediated fusion events. *Genes, Chromosomes & Cancer*, 37, 149–158.

Erickson, P., Gao, J., Chang, K.S., Look, T., Whisenant, E., Raimondi, S., Lasher, R., Trujillo, J., Rowley, J. & Drabkin, H. (1992) Identification of breakpoints in t(8;21) acute myelogenous leukemia and isolation of a fusion transcript, *AML1/ETO*, with similarity to *Drosophila* segmentation gene, runt. *Blood*, 80, 1825–1831.

Gutierrez, N.C., Lopez-Perez, R., Hernandez, J.M., Isidro, I., Gonzalez, B., Delgado, M., Ferminan, E., Garcia, J.L., Vazquez, L., Gonzalez, M. & San Miguel, J.F. (2005) Gene expression profile reveals deregulation of genes with relevant functions in the different subclasses of acute myeloid leukemia. *Leukemia*, 19, 402–409.

Hurwitz, C.A., Raimondi, S.C., Head, D., Krance, R., Mirro, Jr J., Kalwinsky, D.K., Ayers, G.D. & Behm, F.G. (1992) Distinctive immunophenotypic features of t(8;21)(q22;q22) acute myeloblastic leukemia in children. *Blood*, 80, 3182–3188.

Kitabayashi, I., Ida, K., Morohoshi, F., Yokoyama, A., Mitsuhashi, N., Shimizu, K., Nomura, N., Hayashi, Y. & Ohki, M. (1998) The AML1-MTG8 leukemic fusion protein forms a complex with a novel member of the MTG8(ETO/CDR) family, MTGR1. *Molecular and Cellular Biology*, 18, 846–858.

Kohlmann, A., Schoch, C., Schnittger, S., Dugas, M., Hiddemann, W., Kern, W. & Haferlach, T. (2003) Molecular characterization of acute leukemias by use of microarray technology. *Genes, Chromosomes & Cancer*, 37, 396–405.

Lawrence, H.J., Rozenfeld, S., Cruz, C., Matsukuma, K., Kwong, A., Komuves, L., Buchberg, A.M. & Largman, C. (1999) Frequent co-expression of the *HOXA9* and *MEIS1* homeobox genes in human myeloid leukemias. *Leukemia*, 13, 1993–1999.

Liu, P., Tarle, S.A., Hajra, A., Claxton, D.F., Marlton, P., Freedman, M., Siciliano, M.J. & Collins, F.S. (1993) Fusion between transcription factor *CBFβ/PEBP2β* and a myosin heavy chain in acute myeloid leukemia. *Science*, 261, 1041–1044.

Look, A.T. (1997) Oncogenic transcription factors in the human acute leukemias. *Science*, 278, 1059–1064.

Lou, J., Cao, W., Bernardin, F., Ayyanathan, K., Rauscher, III F.J. & Friedman, A.D. (2000) Exogenous cdk4 overcomes reduced cdk4 RNA and inhibition of G1 progression in hematopoietic cells expressing a dominant-negative CBF – a model for overcoming inhibition of proliferation by CBF oncoproteins. *Oncogene*, 19, 2695–2703.

Miyoshi, H., Shimizu, K., Kozu, T., Maseki, N., Kaneko, Y. & Ohki, M. (1991) t(8;21) breakpoints on chromosome 21 in acute myeloid leukemia are clustered within a limited region of a single gene, *AML1*. *Proceedings of the National Academy of Sciences of the United States of America*, 88, 10431–10434.

Miyoshi, H., Kozu, T., Shimizu, K., Enomoto, K., Maseki, N., Kaneko, Y., Kamada, N. & Ohki, M. (1993) The t(8;21) translocation in acute myeloid leukemia results in production of an *AML1-MTG8* fusion transcript. *The EMBO Journal*, 12, 2715–2721.

Nakamura, T., Largaespada, D.A., Lee, M.P., Johnson, L.A., Ohyashiki, K., Toyama, K., Chen, S.J., Willman, C.L., Chen, I.M., Feinberg, A.P., Jenkins, N.A., Copeland, N.G. & Shaughnessy, Jr, J.D. (1996) Fusion of the nucleoporin gene *NUP98* to *HOXA9* by the chromosome translocation t(7;11)(p15;p15) in human myeloid leukaemia. *Nature Genetics*, 12, 154–158.