to the time of analysis or first event. Failure to achieve remission, relapse or death that occurred during continuous complete remission were evaluated as events. Overall survival (OS) was defined as the time from diagnosis to death. Multivariate survival analysis was performed using the Cox proportional-hazards model. A P value of less than 0·05 (two-sided) was considered statistically significant. All statistical analyses were performed using STATA 8·1 (STATACORP LP, College Station, TX, USA).

Results

FBXW7 and NOTCH1 mutations in T-ALL and T-NHL patients

FBXW7 and/or *NOTCH1* mutations were found in 22 (40·0%) of 55 T-ALL and 7 (50·0%) of 14 T-NHL patients (Tables I–III). *FBXW7* mutations were found in 8 (14·6%) of 55 T-ALL and 3 (21·4%) of 14 T-NHL patients, and *NOTCH1*

Table I. *FBW7* and *NOTCH1* mutations in T-ALL and T-NHL patients.

	FBXW7 mutation		*NOTCH1* mutation	
Patient no.	Nucleotide*	Amino acid	Nucleotide†	Amino acid
T-ALL 4	–	–	4778T > C	L1593P
T-ALL 5	1662C > T	R505C	4817_4818insGCCCCC	1606delinsLPP
T-ALL 8	1450_1451ins AGCTGTT GTCTCTCATCATATG CCTTCTCAC	435AVVSHHMPSHHfsX	–	–
T-ALL 20	1542C > T	R465C	4847T > A	I1616N
T-ALL 22	–	–	4775_4776insGAC	1592delinsLT
T-ALL 23	–	–	7355_7356insCTGGC	2453WRCTLFCPRKAPPCP RRCHPRWSHPfsX
T-ALL 26	–	–	4818_4819insCTTTATCTC	1606_1607insHYL
T-ALL 30	–	–	4732_4734del	1578delV
T-ALL 31	–	–	4732_4734del	1578delV
T-ALL 32	2029T > C	V627A	–	–
T-ALL 33	–	–	4754T > C	L1585P
T-ALL 34	1543G > A 715_718delinsGAC	R465H 189RPQNIQVPLGLYHV QQHQQLLGTSEQPM AKGNNDAELHLSSHL QASRNGfsX	–	–
T-ALL 35	–	–	7412delinsAG	S2471X
T-ALL 37	–	–	4732-4734del	1578delV
T-ALL 38	1585G > A	R479Q	–	–
T-ALL 41	–	–	4754T > C	L1585P
T-ALL 46	–	–	7318C > T	Q2440X
T-ALL 49	1585G > A	R479Q	4732-4734del	1578delV
T-ALL 50	–	–	7330C > T	Q2444X
T-ALL 65	–	–	4814_4815delinsCCCCCCCCGA CCATAAGCC	1606PPDHKPSVTHTASRfsX
T-ALL 67	1543G > A	R465H	–	–
T-ALL 75	–	–	4818_4822delinsAGCACACCA GCCCAAGC	1606delinsLAHQP
T-NHL 18	–	–	4709_4718del	1570_1573delinsVDK
T-NHL 25	–	–	7541_7542del	2515RVPfsX
T-NHL 54	–	–	4793G>C	R1598P
			7541_7542del	2515RVPfsX
T-NHL 55	1543G > A	R465H	–	–
T-NHL 58	1543G > A	R465H	4845_4847delinsCCCCTCGAA	1615_1617delinsIPSNF
T-NHL 59	–	–	7326_7327insCGCGGAGGTGC	2443RGGACSHWAPAAWRC TLFCPRRAPPCP RRCHPR WSHPfsX
T-NHL 61	2107del	653RVNLFETfsX	7403_7404insGGGGG	2469GGHPRWSHPfsX

*Nucleotide number is according to the GenBank accession number NM_033632.
†Nucleotide number is according to the GenBank accession number NM_017617.

Table II. Association of *NOTCH1* and *FBXW7* mutations with clinical characteristics in 55 T-ALL patients.

	NOTCH1			*FBXW7*			*FBXW7* and/or *NOTCH1*		
	Mutation (+)	Mutation (−)		Mutation (+)	Mutation (−)		Mutation (+)	Mutation (−)	
Patient characteristics	*n* (%)	*n* (%)	*P*	*n* (%)	*n* (%)	*P*	*n* (%)	*n* (%)	*P*
Overall	17	38		8	47		22	33	
Gender									
Male	12 (70·6)	25 (65·8)	0·726	7 (87·5)	30 (63·8)	0·250	16 (72·7)	21 (63·6)	0·481
Female	5 (29·4)	13 (34·2)		1 (12·5)	17 (36·2)		6 (27·3)	12 (36·4)	
Age at diagnosis (years)				3 (37·5)	27 (57·4)	0·295	12 (54·5)	18 (54·5)	1·0
<10	10 (58·8)	20 (52·6)	0·670	5 (62·5)	20 (42·6)		10 (45·5)	15 (45·5)	
≥10	7 (41·2)	18 (47·4)							
Presenting at diagnosis WBC ($x10^9$/l)									
<100	12 (70·6)	18 (47·4)	0·110	17 (89·5)	26 (72·2)	0·238	16 (72·7)	14 (42·4)	0·027
≥100	5 (29·4)	20 (52·6)		5 (9·1)	13 (36·1)		6 (27·3)	19 (57·6)	
Mediastinal involvement									
Yes	12 (70·6)	22 (57·9)	0·371	4 (50·0)	30 (63·8)	0·464	14 (63·6)	20 (60·6)	0·821
No	5 (29·4)	16 (42·1)		4 (50·0)	17 (36·2)		8 (36·4)	13 (39·4)	
T cell immunophenotype									
Pro and Pre	3 (17·6)	5 (13·2)	0·665	0 (0)	8 (17·0)	0·287	3 (13·6)	5 (15·2)	0·164
Cortical	8 (47·1)	14 (36·8)		5 (62·5)	17 (36·2)		12 (54·5)	10 (30·3)	
Mature	6 (35·3)	19 (50·0)		3 (37·5)	22 (46·8)		7 (31·8)	18 (54·5)	
Chromosomal abnormalities*									
No	11 (68·8)	21 (55·3)	0·749	8 (100·0)	24 (52·2)	0·031	16 (76·2)	16 (48·5)	0·172
Yes									
Abnormalities involving TCR locus† (+)	1 (6·3)	5 (13·2)		0 (0·0)	6 (13·0)		4 (19·0)	12 (36·4)	
Abnormalities involving TCR locus (−)	4 (25·0)	12 (31·6)		0 (0·0)	16 (34·8)		1 (4·8)	5 (15·2)	
Relapse									
Yes	0 (0)	10 (26·3)	0·022	1 (12·5)	9 (19·1)	1·0	1 (4·5)	9 (27·3)	0·039
No	17 (100)	28 (73·7)		7 (87·5)	38 (80·9)		21 (95·5)	24 (72·7)	

Pro and Pre ($CD7^+$ and $CD1^-$), Corical ($CD1^+$), Mature ($CD1^+$, $sCD3^+$).

P, χ^2 or Fisher's exact test; TCR, T cell receptor.

*Total n = 54.

†Chromosomal abnormalities including 14q11, 7p15, and 7q35.

mutations in 17 (30·9%) of 55 T-ALL and 6 (42·3%) of 14 T-NHL patients. Three (5·4%) T-ALL and two (1·4%) T-NHL patients presented mutations in both *FBXW7* and *NOTCH1* (Table I).

The 12 *FBXW7* mutations detected included nine missense mutations, one 31 bp insertion, one single nucleotide deletion, and one deletion/insertion mutation (Table I). Seven of nine missense mutations were clustered in a 'hot spot' encoding arginines 465 and 479 residues which are highly conserved in the WD40 (tryptophan–aspartic-acid) repeat of FBXW7 (Fig S2A). Of the 12 identified *FBXW7* mutations, one insertion (T-ALL 8), one deletion/insertion (T-ALL 34), and one single nucleotide deletion (T-NHL 61) have not been previously described in T-ALL or other cancers (Fig S2B–D). *FBXW7* missense mutation encoding V627A (T-ALL 32) was also a novel mutation. V627 of FBXW7 is evolutionarily conserved, and V627A was not detected in normal lymphocytes from 20 healthy volunteers. One patient (T-ALL 34) had a *FBXW7* deletion/insertion mutation and a missense mutation that encoded FBXW7 residue R465H (Table I, Fig S2C).

Of the 24 *NOTCH1* mutations detected in 23 cases, 16 (66·7%) were located in sequences encoding the HD domain, 8 (33·3%) in the PEST domain (Table I). In one case (T-NHL 54), mutations were detected in both the HD and PEST domains. Of these 24 mutations, 17 (70·9%) were short in-frame insertion or deletions, 5 (20·8%) were missense mutations, and 2 (8·3%) were nonsense mutations in sequences encoding the HD or PEST domains, respectively. Furthermore, a single nucleotide polymorphism C5097T was observed in the sequence encoding the C-terminal region of the HD domain in 63 (91·3%) of 69 patients, as previously reported for Japanese adult patients with mature T cell malignancies (Shimizu *et al*, 2007).

Table III. Association of *NOTCH1* and *FBXW7* mutations with clinical characteristics in 14 T-NHL patients.

	NOTCH1			*FBXW7*			*FBXW7* and/or *NOTCH1*		
	Mutation (+)	Mutation (−)		Mutation (+)	Mutation (−)		Mutation (+)	Mutation (−)	
Patient characteristics	*n* (%)	*n* (%)	*P*	*n* (%)	*n* (%)	*P*	*n* (%)	*n* (%)	*P*
Overall	6	8		3	11		7	7	
Gender									
Male	5 (83·3)	7 (87·5)	1·0	2 (66·7)	10 (90·9)	0·396	6 (85·7)	6 (85·7)	1·0
Female	1 (16·7)	1 (12·5)		1 (33·3)	1 (9·1)		1 (14·3)	1 (14·3)	
Age at diagnosis (years)									
<10	4 (66·7)	2 (25·0)	0·277	2 (66·7)	4 (36·4)	0·538	4 (57·1)	2 (28·6)	0·592
≥10	2 (33·3)	6 (75·0)		1 (33·3)	7 (63·6)		3 (42·9)	5 (71·4)	
Mediastinal involvement									
Yes	0 (0·0)	1 (12·5)	1·0	0 (0·0)	1 (9·1)	1·0	0 (0·0)	1 (14·3)	1·0
No	6 (100·0)	7 (87·5)		3 (100·0)	10 (90·9)		7 (100·0)	6 (85·7)	
T cell immunophenotype									
Pro and Pre	0 (0·0)	0 (0·0)	1·0	0 (0·0)	0 (0·0)	1·0	0 (0·0)	0 (0·0)	1·0
Cortical	2 (33·3)	2 (28·6)		1 (33·3)	3 (30·0)		2 (28·6)	2 (33·3)	
Mature	4 (66·7)	5 (71·4)		2 (66·7)	7 (70·0)		5 (71·4)	4 (66·7)	
Chromosomal abnormalities*									
No	4 (66·7)	3 (42·9)	0·755	3 (100·0)	4 (40·0)	0·217	5 (71·4)	2 (33·3)	0·470
Yes									
Abnormalities involving TCR locus† (+)	1 (16·7)	1 (14·3)		0 (0·0)	2 (20·0)		1 (14·3)	1 (16·7)	
Abnormalities involving TCR locus (−)	1 (16·7)	3 (42·9)		0 (0·0)	4 (40·0)		1 (14·3)	3 (50·0)	
Relapse									
Yes	0 (0·0)	2 (25·0)	0·473	0 (0·0)	2 (18·2)	1·0	0 (0·0)	1 (14·3)	1·0
No	6 (100·0)	6 (75·0)		3 (100·0)	9 (81·8)		7 (100·0)	6 (85·7)	

Pro and Pre ($CD7^+$ and $CD1^-$), Corical ($CD1^+$), Mature ($CD1^+$, $sCD3^+$).
P, χ^2 or Fisher's exact test; TCR, T cell receptor.
*Total n = 13.
†Chromosomal abnormalities including 14q11, 7p15, or 7q35.

Clinical characteristics of FBXW7 and NOTCH1 mutations

The clinical and biological characteristics of the patients in this study are shown in Tables II and III. *FBXW7* and/or *NOTCH1* mutations were associated only with white blood cell (WBC) counts. *FBXW7* and/or *NOTCH1* mutations, but not *FBXW7* or *NOTCH1* alone, were found more frequently in T-ALL patients with low WBC count, $<10 \times 10^9$/l, than in those with higher WBC count, $>10 \times 10^9$/l ($P = 0{\cdot}027$). *FBXW7* mutations, but not *NOTCH1* mutations, were negatively associated with chromosome abnormalities in both T-ALL and T-NHL. All T-ALL and T-NHL patients having *FBXW7* mutation lacked a chromosome abnormality (100% vs. 52·2%, $P = 0{\cdot}031$ in T-ALL and 100% vs. 40·0%, $P = 0{\cdot}217$ in T-NHL).

Prognostic significance of FBXW7 and NOTCH1 mutations

We next analyzed the correlation between *FBXW7* and/or *NOTCH1* mutations and clinical outcome. T-ALL patients with *NOTCH1* mutation had a better clinical outcome than those without *NOTCH1* mutation [100% vs. 65·8% [95% confidence interval (CI), 48·5–78·5%]; $P = 0{\cdot}008$ for 5-year EFS and 100% vs. 81·6% [95% CI, 65·2–90·8%]; $P = 0{\cdot}065$ for 5-year OS, respectively] (Fig S3), while the prognostic difference between patients with and without *FBXW7* mutation was not significant [87·5% (95% CI, 38·7–98·1%) vs. 74·5% (95% CI, 59·4–84·6%); $P = 0{\cdot}400$ for 5-year EFS and 100% vs. 85·1% (95% CI, 71·3–92·6%); $P = 0{\cdot}259$ for 5-year OS, respectively] (Fig S4). The 5-year EFS and OS for T-ALL patients with *FBXW7* and/or *NOTCH1* mutations were extremely high, suggesting a good prognosis for patients with *FBXW7/NOTCH1* mutation compared to those without [95·5% (95% CI, 71·9–99·4%) vs. 63·6% (95% CI, 45·0–77·5%); $P = 0{\cdot}007$ and 100% vs. 78·8% (95% CI, 60·6–89·3%); $P = 0{\cdot}023$, respectively] (Fig 1). Notably, all three patients with both *FBXW7* and *NOTCH1* mutations were alive without relapse.

Multivariate analysis of prognostic factors adjusted for gender, age at diagnosis, and WBC count presented at diagnosis revealed that *FBXW7* and/or *NOTCH1* mutation status, risk group for treatment, and chromosomal abnormalities retained

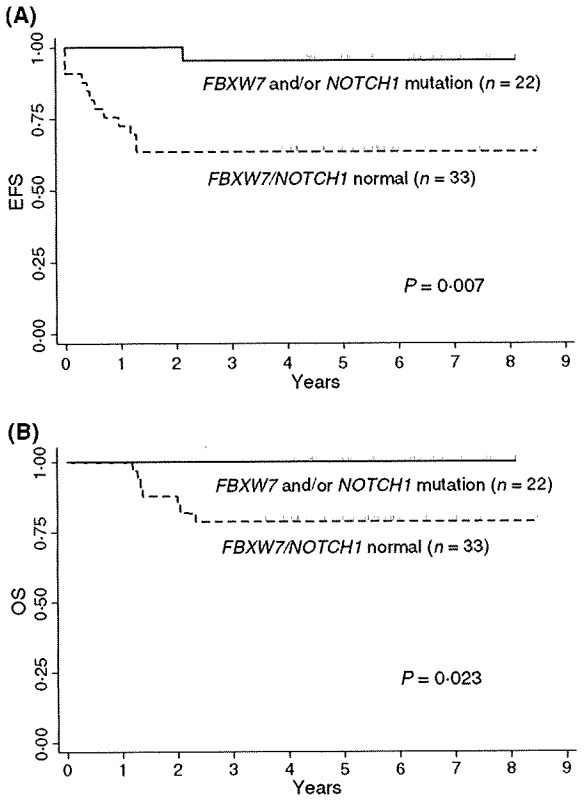

Fig 1. Kaplan–Meier estimate of (A) event-free survival and (B) overall survival of T-ALL patients with or without *FBXW7* and/or *NOTCH1* mutation.

their significant effects on EFS (Table IV). On the other hand, multivariate analysis adjusted for *NOTCH1* and/or *FBXW7* mutation status, risk group for treatment, and chromosomal abnormalities, in addition to gender, age at diagnosis, and WBC count presented at diagnosis, revealed that none of them retained EFS significance (Table IV).

In T-NHL, patients with *NOTCH1* and/or *FBXW7* mutation also had a good prognosis, although the differences in 5-year EFS and OS for patients with and without *NOTCH1* and/or *FBXW7* mutations were not significant [EFS, 85·7% (95% CI, 33·4–97·9%) vs. 57·1% (95% CI, 17·2–83·7%), $P = 0·313$; OS, 85·7% (95% CI, 33·4–97·9%) vs. 53·6% (95% CI, 13·2–82·5%), $P = 0·286$].

Discussion

In this study, we found 14·6% *FBXW7* mutations and 30·9% *NOTCH1* mutations in T-ALL patients, and 21·4% *FBXW7* mutations and 42·3% *NOTCH1* mutations in T-NHL patients. Frequencies of *FBXW7* and *NOTCH1* mutations in T-ALL in this study were similar to those in other recent studies (8·6–30·8% for *FBXW7* mutations, and 30·8–70·8% for *NOTCH1* mutations) (Akhoondi *et al*, 2007; Lee *et al*, 2005; Malyukova *et al*, 2007; Mansour *et al*, 2006; O'Neil *et al*, 2007; Thompson *et al*, 2007; van Grotel *et al*, 2008). This is the first report describing high frequencies of *FBXW7* and *NOTCH1* mutations in T-NHL as well as in T-ALL. The types of mutations identified were similar in T-ALL and T-NHL patients (Table I), although it was previously reported that gene expression profiling revealed intrinsic differences between T-ALL and T-NHL (Raetz *et al*, 2006).

Our results demonstrated that *FBXW7* and/or *NOTCH1* mutations as well as *NOTCH1* mutations alone had a good prognosis in T-ALL patients. The *P* value regarding the significant difference in prognosis for patients with *FBXW7* and/or *NOTCH1* status ($P = 0·007$ for EFS) was less than for those with *NOTCH1* status alone ($P = 0·008$), although the difference in prognosis for *FBXW7* status alone was not significant ($P = 0·397$). All T-ALL and T-NHL patients with

Table IV. Multivariate analysis of effects of *FBXW7* and/or *NOTCH1* mutations on EFS in 55 T-ALL patients.

	Crude HR		Adjusted HR1*		Adjusted HR2†	
	(95% CI)	P‡	(95% CI)	P‡	(95% CI)	P‡
FBXW7 and/or *NOTCH1* mutation						
Negative	1·00§		1·00§		1·00§	
Positive	0·10 (0·01–0·78)	0·028	0·10 (0·01–0·77)	0·027	0·24 (0·05–1·13)	0·071
Chromosomal abnormalities						
No	1·00§		1·00§		1·00§	
Yes						
Abnormalities involving TCR locus (+)	5·99 (1·55–23·22)	0·010	6·04 (1·54–23·70)	0·010	6·41 (1·35–30·58)	0·020
Abnormalities involving TCR locus (−)	7·63 (1·53–38·11)	0·013	10·80 (2·03–57·57)	0·005	3·22 (0·89–11·67)	0·076

HR, hazard ratio; CI, confidence interval.

*Adjusted for sex, age at diagnosis and presenting white blood cell count (categorical: see Table I).

†Adjusted for sex, age at diagnosis, presenting white blood cell count, *FBXW7* and/or *NOTCH1* mutations category, determined risk and chromosomal abnormalities) (categorical: see Table I).

‡*P*, X^2 test.

§Reference category.

FBXW7 mutations, with the exception of one T-ALL patient, have survived without relapse. One patient (T-ALL 38) had an isolated CNS relapse; however, the patient had survived 2 years after the relapse episode.

The paediatric ALL-BFM 2000 study reported the good clinical outcome for T-ALL patients with *NOTCH1* mutations (Breit *et al*, 2006), however, other two reports described results that were not compatible with this (van Grotel *et al*, 2008; Zhu *et al*, 2006). One possible explanation for this discrepancy of prognostic impact is the different treatment protocols used; the survival rates reported in other papers were apparently lower [28·8% 3-year relapse free survival (Zhu *et al*, 2006) and 65% 5-year disease-free survival (van Grotel *et al*, 2008)] for T-ALL patients with *NOTCH1* mutation than those of the ALL-BFM 2000 study (90% relapse-free survival) and our study (100% 5-year EFS). On the other hand, there was no statistically significant impact of *NOTCH1* mutations on prognosis in T-NHL patients, perhaps because the number of T-NHL patients was small in this study. Further study of T-NHL patients is needed to clarify the association of *FBXW7* and *NOTCH1* mutations with T-NHL prognosis.

Four novel mutations were found, and two of the four, V627A in T-ALL 32 and a frame shift mutation at codon 653 in T-NHL 61, were positioned outside of a 'hot spot' region. Codon 627 is localized in the seventh β-propeller blade (β-PB7) of FBXW7 (Orlicky *et al*, 2003), and a R689W mutation in the β-PB8 was also reported in T-ALL cases (Malyukova *et al*, 2007). C-terminal truncation of FBXW7 observed in T-NHL 61 was also reported in an endometrial tumour (nonsense mutation of codon 658) (Akhoondi *et al*, 2007), and these mutations result in the absence of a portion of β-PB7 and all of β-PB8. These findings suggested that a structural change of any β-propeller blades may have similar effects on FBXW7 function. Furthermore, it was also demonstrated that *Fbxw7* deficiency in adult haematopoietic cells leads to T-ALL in mice (Matsuoka *et al*, 2008), suggesting that inactivation of FBXW7 plays a critical role in T-ALL leukaemogenesis.

Chromosomal abnormalities of the *TLX3* (5q35) and *TLX1* (10q24) locus have been reported to be associated with poor and good outcome (van Grotel *et al*, 2008). In this study, chromosomal abnormalities involving the *TLX1* locus were found in one patient and chromosomal abnormalities involving the breakpoint at 5q35·1 (*TLX3*) were not found in any patients. *t*(10;11)(q13;q14) [*PICALM-MLLT10* (previously termed *CALM-AF10*)] was not found. The prognostic significance of these cytogenetic abnormalities was not clear because the number of patients was small.

Notably, *FBXW7* mutations were only observed in T-ALL and T-NHL patients lacking chromosomal abnormalities. *FBXW7* is considered to be a haplo-insufficient tumour suppressor gene (Mao *et al*, 2004). Inactivation of FBXW7 has been reported to cause chromosomal instability in karyotypically stable colorectal cancer cells, resulting in a striking phenotype associated with micronuclei and chromosomal instability (Rajagopalan *et al*, 2004). On the contrary, *FBXW7* mutation has been reported to lack association with chromosomal instability in colorectal cancer (Kemp *et al*, 2005), which was compatible with the present results for T-ALL. Further studies are needed to clarify this issue.

In conclusion, *FBXW7* and *NOTCH1* are functionally related each other, and the mutations of either *FBXW7* or *NOTCH1* genes rather than *FBXW7* or *NOTCH1* alone were associated with good clinical outcome in T-ALL, suggesting that the status of both *FBXW7* and *NOTCH1*, rather than *FBXW7* or *NOTCH1* alone, is a useful prognostic factor in T-ALL.

Acknowledgements

We thank all the participants of the JACLS ALL-97 and NHL-98 study. We also thank Professor Masafumi Taniwaki, Department of Haematology and Oncology, Kyoto Prefectural University of Medicine Graduate School of Medical Science for his help, and Miss Chisato Kubota and Mr Yoshiharu Tokita for technical assistance. This work was supported by a grant for Cancer Research, and a grant for Research on Children and Families from the Ministry of Health, Labour, and Welfare of Japan, a Grant-in-Aid for Scientific Research (C) and Exploratory Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan and by a Research grant for Gunma Prefectural Hospitals.

References

Akhoondi, S., Sun, D., von der Lehr, N., Apostolidou, S., Klotz, K., Maljukova, A., Cepeda, D., Fiegl, H., Dafou, D., Marth, C., Mueller-Holzner, E., Corcoran, M., Dagnell, M., Nejad, S.Z., Nayer, B.N., Zali, M.R., Hansson, J., Egyhazi, S., Petersson, F., Sangfelt, P., Nordgren, H., Grander, D., Reed, S.I., Widschwendter, M., Sangfelt, O. & Spruck, C. (2007) FBXW7/hCDC4 is a general tumour suppressor in human cancer. *Cancer Research*, **67**, 9006–9012.

Breit, S., Stanulla, M., Flohr, T., Schrappe, M., Ludwig, W.D., Tolle, G., Happich, M., Muckenthaler, M.U. & Kulozik, A.E. (2006) Activating NOTCH1 mutations predict favorable early treatment response and long-term outcome in childhood precursor T cell lymphoblastic leukaemia. *Blood*, **108**, 1151–1157.

Cassia, R., Moreno-Bueno, G., Rodriguez-Perales, S., Hardisson, D., Cigudosa, J.C. & Palacios, J. (2003) Cyclin E gene (CCNE) amplification and hCDC4 mutations in endometrial carcinoma. *Journal of Pathology*, **201**, 589–595.

Ellisen, L.W., Bird, J., West, D.C., Soreng, A.L., Reynolds, T.C., Smith, S.D. & Sklar, J. (1991) TAN-1, the human homolog of the Drosophila notch gene, is broken by chromosomal translocations in T lymphoblastic neoplasms. *Cell*, **66**, 649–661.

Ferrando, A.A. & Look, A.T. (2003) Gene expression profiling in T cell acute lymphoblastic leukaemia. *Seminars in Hematology*, **40**, 274–280.

Ferrando, A.A., Neuberg, D.S., Staunton, J., Loh, M.L., Huard, C., Raimondi, S.C., Behm, F.G., Pui, C.H., Downing, J.R., Gilliland, D.G., Lander, E.S., Golub, T.R. & Look, A.T. (2002) Gene expression signatures define novel oncogenic pathways in T-cell acute lymphoblastic leukaemia. *Cancer Cell*, **1**, 75–87.

Ferrando, A.A., Neuberg, D.S., Dodge, R.K., Paietta, E., Larson, R.A., Wiernik, P.H., Rowe, J.M., Caligiuri, M.A., Bloomfield, C.D. &

Look, A.T. (2004) Prognostic importance of TLX1 (HOX11) oncogene expression in adults with T cell acute lymphoblastic leukaemia. *Lancet*, **363,** 535–536.

Fryer, C.J., White, J.B. & Jones, K.A. (2004) Mastermind recruits CycC:CDK8 to phosphorylate the Notch ICD and coordinate activation with turnover. *Molecular Cell*, **16,** 509–520.

Gaynon, P.S., Trigg, M.E., Heerema, N.A., Sensel, M.G., Sather, H.N., Hammond, G.D. & Bleyer, W.A. (2000) Children's cancer group trials in childhood acute lymphoblastic leukaemia: 1983–1995. *Leukaemia*, **14,** 2223–2233.

Hubbard, E.J., Wu, G., Kitajewski, J. & Greenwald, I. (1997) sel-10, a negative regulator of lin-12 activity in Caenorhabditis elegans, encodes a member of the CDC4 family of proteins. *Genes and Development*, **11,** 3182–3193.

Kemp, Z., Rowan, A., Chambers, W., Wortham, N., Halford, S., Sieber, O., Mortensen, N., von Herbay, A., Gunther, T., Ilyas, M. & Tomlinson, I. (2005) CDC4 mutations occur in a subset of colorectal cancers but are not predicted to cause loss of function and are not associated with chromosomal instability. *Cancer Research*, **65,** 11361–11366.

Lee, S.Y., Kumano, K., Masuda, S., Hangaishi, A., Takita, J., Nakazaki, K., Kurokawa, M., Hayashi, Y., Ogawa, S. & Chiba, S. (2005) Mutations of the Notch1 gene in T cell acute lymphoblastic leukaemia: analysis in adults and children. *Leukaemia*, **19,** 1841–1843.

Maloney, K.W., Shuster, J.J., Murphy, S., Pullen, J. & Camitta, B.A. (2000) Long-term results of treatment studies for childhood acute lymphoblastic leukaemia: paediatric oncology group studies from 1986–1994. *Leukaemia*, **14,** 2276–2285.

Malyukova, A., Dohda, T., von der Lehr, N., Akhoondi, S., Corcoran, M., Heyman, M., Spruck, C., Grander, D., Lendahl, U. & Sangfelt, O. (2007) The tumour suppressor gene hCDC4 is frequently mutated in human T cell acute lymphoblastic leukaemia with functional consequences for Notch signaling. *Cancer Research*, **67,** 5611–5616.

Mansour, M.R., Linch, D.C., Foroni, L., Goldstone, A.H. & Gale, R.E. (2006) High incidence of Notch-1 mutations in adult patients with T cell acute lymphoblastic leukaemia. *Leukaemia*, **20,** 537–539.

Mao, J.H., Perez-Losada, J., Wu, D., Delrosario, R., Tsunematsu, R., Nakayama, K.I., Brown, K., Bryson, S. & Balmain, A. (2004) Fbxw7/Cdc4 is a p53-dependent, haploinsufficient tumour suppressor gene. *Nature*, **432,** 775–779.

Matsuoka, S., Oike, Y., Onoyama, I., Iwama, A., Arai, F., Takubo, K., Mashimo, Y., Oguro, H., Nitta, E., Ito, K., Miyamoto, K., Yoshiwara, H., Hosokawa, K., Nakamura, Y., Gomei, Y., Iwasaki, H., Hayashi, Y., Matsuzaki, Y., Nakayama, K., Ikeda, Y., Hata, A., Chiba, S., Nakayama, K.I. & Suda, T. (2008) Fbxw7 acts as a critical fail-safe against premature loss of hematopoietic stem cells and development of T-ALL. *Genes and Development*, **22,** 986–991.

Moberg, K.H., Bell, D.W., Wahrer, D.C., Haber, D.A. & Hariharan, I.K. (2001) Archipelago regulates Cyclin E levels in Drosophila and is mutated in human cancer cell lines. *Nature*, **413,** 311–316.

Moghrabi, A., Levy, D.E., Asselin, B., Barr, R., Clavell, L., Hurwitz, C., Samson, Y., Schorin, M., Dalton, V.K., Lipshultz, S.E., Neuberg, D.S., Gelber, R.D., Cohen, H.J., Sallan, S.E. & Silverman, L.B. (2007) Results of the Dana–Farber Cancer Institute ALL Consortium Protocol 95-01 for children with acute lymphoblastic leukaemia. *Blood*, **109,** 896–904.

Oda, M., Yumura-Yagi, K., Hara, J., Nishimura, S., Tawa, A., Komada, Y., Kudo, T., Yoshida, M., Kobayashi, R., Takimoto, T., Nakagawa, A. & Horibe, K. & The Japan Association of Childhood Leukaemia, S. (2006) Mediastinal mass with bone marrow involvement is a favorable factor for childhood T-lymphoid malignancy in Japan. *ASH Annual Meeting Abstracts*, **108,** 4525.

O'Neil, J., Grim, J., Strack, P., Rao, S., Tibbitts, D., Winter, C., Hardwick, J., Welcker, M., Meijerink, J.P., Pieters, R., Draetta, G., Sears, R., Clurman, B.E. & Look, A.T. (2007) FBW7 mutations in leukemic cells mediate NOTCH pathway activation and resistance to gamma-secretase inhibitors. *Journal of Experimental Medicine*, **204,** 1813–1824.

Orlicky, S., Tang, X., Willems, A., Tyers, M. & Sicheri, F. (2003) Structural basis for phosphodependent substrate selection and orientation by the SCFCdc4 ubiquitin ligase. *Cell*, **112,** 243–256.

Pui, C.H., Sandlund, J.T., Pei, D., Campana, D., Rivera, G.K., Ribeiro, R.C., Rubnitz, J.E., Razzouk, B.I., Howard, S.C., Hudson, M.M., Cheng, C., Kun, L.E., Raimondi, S.C., Behm, F.G., Downing, J.R., Relling, M.V. & Evans, W.E. (2004) Improved outcome for children with acute lymphoblastic leukaemia: results of total therapy study XIIIB at St Jude Children's Research Hospital. *Blood*, **104,** 2690–2696.

Raetz, E.A., Perkins, S.L., Bhojwani, D., Smock, K., Philip, M., Carroll, W.L. & Min, D.J. (2006) Gene expression profiling reveals intrinsic differences between T cell acute lymphoblastic leukaemia and T cell lymphoblastic lymphoma. *Paediatric Blood & Cancer*, **47,** 130–140.

Rajagopalan, H., Jallepalli, P.V., Rago, C., Velculescu, V.E., Kinzler, K.W., Vogelstein, B. & Lengauer, C. (2004) Inactivation of hCDC4 can cause chromosomal instability. *Nature*, **428,** 77–81.

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

Shimizu, D., Taki, T., Utsunomiya, A., Nakagawa, H., Nomura, K., Matsumoto, Y., Nishida, K., Horiike, S. & Taniwaki, M. (2007) Detection of NOTCH1 mutations in adult T cell leukaemia/lymphoma and peripheral T cell lymphoma. *International Journal of Hematology*, **85,** 212–218.

Taketani, T., Taki, T., Sugita, K., Furuichi, Y., Ishii, E., Hanada, R., Tsuchida, M., Sugita, K., Ida, K. & Hayashi, Y. (2004) FLT3 mutations in the activation loop of tyrosine kinase domain are frequently found in infant ALL with MLL rearrangements and paediatric ALL with hyperdiploidy. *Blood*, **103,** 1085–1088.

Tetzlaff, M.T., Yu, W., Li, M., Zhang, P., Finegold, M., Mahon, K., Harper, J.W., Schwartz, R.J. & Elledge, S.J. (2004) Defective cardiovascular development and elevated cyclin E and Notch proteins in mice lacking the Fbw7 F-box protein. *Proceedings of the National Academy of Sciences of the United States of America*, **101,** 3338–3345.

Thompson, B.J., Buonamici, S., Sulis, M.L., Palomero, T., Vilimas, T., Basso, G., Ferrando, A. & Aifantis, I. (2007) The SCFFBW7 ubiquitin ligase complex as a tumour suppressor in T cell leukaemia. *Journal of Experimental Medicine*, **204,** 1825–1835.

van Grotel, M., Meijerink, J.P., van Wering, E.R., Langerak, A.W., Beverloo, H.B., Buijs-Gladdines, J.G., Burger, N.B., Passier, M., van Lieshout, E.M., Kamps, W.A., Veerman, A.J., van Noesel, M.M. & Pieters, R. (2008) Prognostic significance of molecular-cytogenetic abnormalities in paediatric T-ALL is not explained by immunophenotypic differences. *Leukaemia*, **22,** 124–131.

Weng, A.P., Ferrando, A.A., Lee, W., Morris, J.P.t., Silverman, L.B., Sanchez-Irizarry, C., Blacklow, S.C., Look, A.T. & Aster, J.C. (2004)

Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukaemia. *Science*, **306**, 269–271.

Winter, S.S., Jiang, Z., Khawaja, H.M., Griffin, T., Devidas, M., Asselin, B.L. & Larson, R.S. (2007) Identification of genomic classifiers that distinguish induction failure in T-lineage acute lymphoblastic leukaemia: a report from the Children's Oncology Group. *Blood*, **110**, 1429–1438.

Zhu, Y.M., Zhao, W.L., Fu, J.F., Shi, J.Y., Pan, Q., Hu, J., Gao, X.D., Chen, B., Li, J.M., Xiong, S.M., Gu, L.J., Tang, J.Y., Liang, H., Jiang, H., Xue, Y.Q., Shen, Z.X., Chen, Z. & Chen, S.J. (2006) NOTCH1 mutations in T cell acute lymphoblastic leukaemia: prognostic significance and implication in multifactorial leukemogenesis. *Clinical Cancer Research*, **12**, 3043–3049.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Fig S1. Treatment plan for JACLS ALL T-97 protocol.

Fig S2. Newly identified *FBW7* mutations in T-ALL patients.

Fig S3. Kaplan–Meier estimate of (A) event-free survival and (B) overall survival of T-ALL patients with or without *NOTCH1* mutation.

Fig S4. Kaplan–Meier estimate of (A) event-free survival and (B) overall survival of T-ALL patients with or without *FBW7* mutation.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

bjh research paper

Age-associated difference in gene expression of paediatric acute myelomonocytic lineage leukaemia (FAB M4 and M5 subtypes) and its correlation with prognosis

Aoi Jo,[1,2] Ichiro Tsukimoto,[3] Eiichi Ishii,[4] Norio Asou,[5] Sachiyo Mitani,[1] Akira Shimada,[6] Takashi Igarashi,[2] Yasuhide Hayashi[6] and Hitoshi Ichikawa[1]

[1]Genetics Division, National Cancer Centre Research Institute, Tokyo, [2]Department of Paediatrics, University of Tokyo, Tokyo, [3]Department of Paediatrics, Toho University School of Medicine, Tokyo, [4]Department of Paediatrics, Ehime University, Toon, [5]Department of Haematology, Kumamoto University, Kumamoto, and [6]Department of Haematology/Oncology, Gunma Children's Medical Centre, Shibukawa, Japan

Received 11 September 2008; accepted for publication 23 October 2008

Correspondence: Dr Hitoshi Ichikawa, Section Head, Genetics Division, National Cancer Centre Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan.
E-mail: hichikaw@ncc.go.jp

Summary

Acute myeloid leukaemia, French-American-British M4 and M5 subtypes (AML-M4/M5) is frequently associated with *MLL* gene rearrangement and its incidence is relatively high among infants. Clinically, paediatric AML-M4/M5 has been considered as an intermediate or undefined prognostic group. In this study, we analysed gene expression of 40 paediatric AML-M4/M5 patients excluding inv(16) and t(8;21) patients, and found striking differences among the patients in an age-associated manner. In particular, most of the infants displayed very distinct gene expression. On the basis of this difference, we divided paediatric patients into three subgroups (A, B and C) with the average age of 0·3, 3·1 and 6·6 years old respectively. All subgroups included patients with *MLL* gene rearrangement as well as normal and other karyotypes. Surprisingly, gene expression signatures of *MLL* gene rearrangement differed substantially among these subgroups. In addition, subgroup C presented extremely poor outcome (3-year event-free survival 28%) whilst eight patients with *MLL* gene rearrangement in subgroup C had all relapsed within 18 months. These results suggest that age is an important factor contributing to the biology of AML-M4/M5 and the sub-grouping procedures developed in this study could be a powerful tool to identify unfavourable risk patients within paediatric AML-M4/M5.

Keywords: acute myeloid leukaemia, gene expression profiling, microarray, MLL, prognostic factors.

Acute myeloid leukaemia (AML) is a heterogeneous disease with different morphological features usually classified according to the French-American-British (FAB) classification system into subtypes M0–M7. In addition, a variety of genetic alterations have been ascertained including specific chromosomal translocations, such as t(8;21), t(15;17), inv(16) and 11q23 rearrangement. These major chromosomal translocations are usually associated with specific FAB subtypes. With the introduction of high throughput gene expression analysis using DNA microarrays, many studies have reported unique gene expression signatures for major chromosomal translocations, suggesting that the underlying molecular biology of these leukaemias can be defined by the genetic alterations (Golub *et al*, 1999; Armstrong *et al*, 2002; Schoch *et al*, 2002; Yagi *et al*, 2003; Ross *et al*, 2004). AML classified as FAB M4 or M5 subtypes (AML-M4/M5) have myelomonoblastic or monoblastic morphology. In this subtype, 11q23 rearrangements resulting in the fusion of the *MLL* with other partner genes (*MLL* gene rearrangement) are often observed. Unlike other chromosomal translocations, *MLL* gene rearrangement is also observed in acute lymphoblastic leukaemia (ALL), especially in infant ALL. Different studies have identified common gene expression signatures for acute leukaemia that contain *MLL* gene rearrangement irrespective of their lineage (Armstrong *et al*, 2002; Ross *et al*, 2004; Kohlmann *et al*, 2005).

The biological and clinical features of acute leukaemia differ between children and adults (Downing & Shannon, 2002) and considerable variations exist among different age groups in children. It is well known that infant ALL possess molecular abnormalities of *MLL* gene rearrangement that result in

 First published online 16 December 2008
British Journal of Haematology, **144**, 917–929 doi:10.1111/j.1365-2141.2008.07531.x

extremely poor outcome (Taki *et al*, 1996; Pui *et al*, 2002). For these differences, infant ALL is clinically and statistically handled separately from all the other paediatric ALL. In AML, distribution of FAB subtype differs among the age group. Infant AML generally show myelomonoblastic or monoblasic features (FAB M4 or M5) and often involves *MLL* gene rearrangement (Pui *et al*, 1995; Horibe *et al*, 2001). However, the presence of *MLL* gene rearrangement failed to correlate with treatment response in infant AML and the clinical outcomes were very similar to those of childhood AML (Satake *et al*, 1999; Pui *et al*, 2000; Kawasaki *et al*, 2001). For these reasons, not many studies distinguish infant from older paediatric AML clinically and statistically, and the biological differences between infant AML and childhood AML remain obscure.

In this study, we examined gene expression of paediatric AML-M4/M5, excluding inv(16) and t(8;21) patients, using microarray technology and found that age-associated differences exist among paediatric AML-M4/M5 patients. With the use of expression patterns, we have subdivided the patients into three subgroups and identified differences in the clinical outcome among these subgroups. Also, we have reviewed the gene expression profiles specific to patients with *MLL* gene rearrangement and showed that different signatures exist within different subgroups.

Materials and methods

Patients and samples

Forty paediatric AML patients (≤15 years old) diagnosed as FAB M4 or M5 subtype, between 1995 and 2001 in Japan, were enrolled in this study. Patients with inv(16) or t(8;21) were not included in this study. Characteristics of the 40 patients are shown in Table SI. These patients were treated with different but similar protocols, ANLL91 (Nagao, 1995; Satake *et al*, 1999) or AML99 (Tsukimoto *et al*, 2005; Shimada *et al*, 2006), which combine cytarabine, etoposide and anthracyclines. All leukaemic samples used in this study were obtained at the time of diagnosis. Morphological diagnosis according to FAB classification, immunophenotype and cytogenetic analysis were performed locally. *MLL* gene rearrangement was determined by cytogenetic analysis and multiplex reverse transcription polymerase chain reaction (RT-PCR) analysis (Pallisgaard *et al*, 1998; Salto-Tellez *et al*, 2003) using the HemaVision kit (DNA Technology, Aarhus, Denmark). Southern blot or fluorescence *in situ* hybridization analysis was also performed in some patients. In addition to the paediatric patients, 14 adult patients (20–79 years old), diagnosed as AML FAB M4 or M5 subtype [excluding inv(16) and t(8;21) patients], were also enrolled to compare paediatric AML to adult AML (Table SII).

This study was approved by the ethics committee of the National Cancer Centre. Patient identities were masked, and study numbers were assigned to all collected samples. To maintain anonymity, only age, sex, FAB subtype, chromosomal abnormality, treatment protocol and clinical outcome of the patients were linked to RNA samples. Informed consent was provided and obtained from the patients and/or the parents as appropriate.

Microarray analysis

For microarray gene expression analysis, mononuclear cells were isolated from the bone marrow cells at the time of diagnosis (median percentage of leukaemic blast cells: 82·3%, see Table SI), and then total RNA was prepared using an RNeasy Mini kit (Qiagen, Hilden, Germany). The integrity of the purified RNA was confirmed using a 2100 Bioanalyzer and an RNA 6000 Nano LabChip kit (Agilent Technologies, Santa Clara, CA, USA). The DNA microarray used was a Human Genome U133 plus 2.0 array (Affymetrix, Santa Clara, CA, USA). Target cRNA was prepared from 20 ng of the purified RNA with a two-cycle cDNA synthesis kit and 3′-amplification reagents for IVT labeling (Affymetrix). Hybridization to the microarrays, washing and staining with the antibody amplification procedure and scanning, were performed according to the manufacturer's instructions. Using the GeneChip Operating Software version 1.4 (Affymetrix), the scanned image data were processed and the expression value (Signal) and detection call (Present, Marginal or Absent) of each probe set were calculated. The Signal values were normalized so that the mean in each experiment was set at 100 to adjust for minor differences between the experiments.

Statistical analysis

For statistical analysis, the Signal values were log-transformed after the addition of 10 to reduce adverse effect caused by the noises at low expression levels. Most of the statistical analysis including principal component analysis (PCA), hierarchical clustering analysis, and Student's *t*-test were performed on the GeneSpring GX software version 7.31 (Silicon Genetics, Redwood City, CA, USA) after the log-transformed expression values were normalized to the median of all samples enrolled in each of the analyses. For the hierarchical clustering analysis, Pearson's correlation was used. To select karyotype discriminating genes, Significance Analysis of Microarray (SAM) was also used (Tusher *et al*, 2001). The Kaplan–Meier method and log-rank test were used for comparison of event-free survival (EFS), which was defined as the time from diagnosis to event (relapse or death of any cause).

Quantitative RT-PCR

Quantitative RT-PCR was carried out using the 7500 Fast Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) with FastStart TaqMan Probe Master (Rox) and Universal ProbeLibrary (Roche Applied Science, Mannheim, Germany). cDNA was prepared from 50 ng of total RNA using SuperScript III (Invitrogen), and 1/100 of the cDNA was used as a

template for each PCR reaction. Probes and primers were designed using the Probe Finder software (Roche Applied Science). Sequences of the primers are listed in Table SIII. Transcript levels were normalized to that of the *GAPDH* transcript. Quantitative RT-PCR was performed on 15 of the 40 paediatric patients.

Results

Expression profile of paediatric AML-M4/M5

Gene expression of 40 paediatric AML patients diagnosed as FAB M4 or M5 subtype was analysed with an oligonucleotide microarray composed of 54675 probe sets. Patients with inv(16) or t(8;21) were excluded from this study because they are known to display distinct expression profiles and clinical behavior (Yagi *et al*, 2003; Ross *et al*, 2004; Valk *et al*, 2004). Among the 40 paediatric patients, 14 patients (35%) were infants (<1 year old) and 10 out of 14 infants (71%) possessed *MLL* gene rearrangement (Table I and SI).

To obtain the overview of the variations in gene expression among the 40 patients, PCA was performed. Patients were clearly separated into two clusters by the first principal component, a tight cluster and a relatively loose cluster (Fig 1A). The tight cluster was composed of 10 patients that were all infants. Although the remaining four infants were included in the loose cluster, the gathering of infants into one cluster was statistically significant ($P = 6{\cdot}5^{-07}$ in chi-square test). On the other hand, the distribution of *MLL* gene rearrangement (presence or absence) and FAB subtype (M4 or M5) had no significant correlations between the two clusters ($P = 0{\cdot}097$ and $P = 0{\cdot}58$ respectively). Clinical outcomes were also not significantly different between the two clusters ($P = 0{\cdot}23$ in log-rank test), although the patients in the loose cluster seemed to have poorer outcome.

Table I. AML-M4/M5 patients enrolled in this study.

Variables	Patients (*n*)	Ratio (%)
Total	40	
Gender		
Male	22	55·0
Female	18	45·0
French-American-British subtype		
M4	17	42·5
M5	23	57·5
Age at diagnosis (years)		
<1	14	35·0
1–5	16	40·0
≥6	10	25·0
Karyotype/gene rearrangement		
MLL gene rearrangement	23	57·5
Normal	13	32·5
Others	4	10·0

To assess the differences in gene expression among the 40 patients, we then performed unsupervised two-dimensional hierarchical clustering analysis. In this analysis, the same 10 infants who assembled a strong cluster in PCA again formed a distinct cluster (Fig 1B). This analysis also showed that this infant cluster was characterized by high expression of a very large set of genes (Fig 1B). These results suggest that most of infant AML-M4/M5 possess biological features different from older children with AML-M4/M5. This typical infant cluster comprised both *MLL* gene rearrangement-positive [*MLL*gr(+)] and *MLL* gene rearrangement-negative [*MLL*gr(−)] patients, indicating that a typical infant profile was not restricted by *MLL* gene rearrangement status. In addition, we could not detect any prominent cluster formation of *MLL*gr(+) patients or gene expression signature specific to *MLL* gene rearrangement in this analysis.

Comparison with adult AML-M4/M5

To compare the gene expression of paediatric AML-M4/M5 with that of adult AML-M4/M5, gene expression of 14 adult patients were examined and added to the PCA and hierarchical clustering analysis. PCA with 40 paediatric and 14 adult patients revealed three clusters correlating with patient age. Ten infants and the remaining paediatric patients once again formed separate clusters while 11 out of 14 adult patients also formed a distinct cluster (Fig 1C). These results suggest that older paediatric AML-M4/M5 as well as infant AML-M4/M5 have distinct characteristics that were different from adult AML-M4/M5. In the unsupervised two-dimensional hierarchical clustering analysis, 12 out of 14 adult patients formed a distinct cluster (Fig S1). A large set of genes was commonly overexpressed between adults and infants, and a relatively small set of genes was commonly overexpressed between adults and older children (Fig S1). These results imply that expression patterns are not defined by the linear increment of patients' age. The genes shared by the infant and adult clusters included those encoding transcription factors and regulatory proteins involved in haematopoietic cell differentiation and proliferation (*AFF1*, *MLLT4*, *MLLT10*, *ARHGAP26*, *ELF2*, *ETV6*, *FLI1*, *FOXP1*, *FYN*, *LRRFIP1/2*, *LYN*, *MED13*, *MED13L*, *RUNX1*).

Sub-grouping of paediatric AML-M4/M5

When we looked into the expression patterns of paediatric patients, the solid clustering of 10 infants was inevitable but we also realized that the remaining patients seemed to be heterogeneous. For those genes that were specifically overexpressed in 10 infants, some of the other patients showed quite a different expression while the others showed somewhat closer expression patterns to 10 infants (Fig 1B). Thus, we proceeded with our analysis to separate the 40 paediatric patients into three subgroups based solely on gene expression without any previous knowledge of the patients' clinical

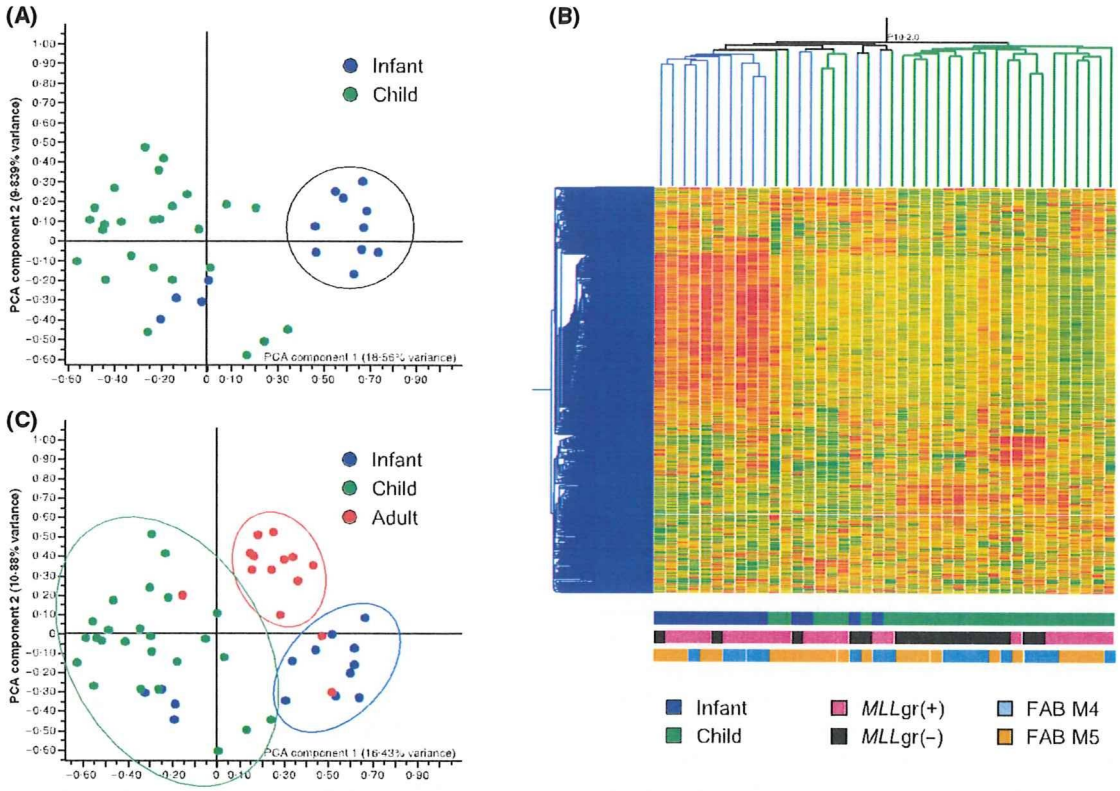

Fig 1. Overview of AML-M4/M5 gene expression. (A) PCA analysis of 40 paediatric AML-M4/M5 patients. 25754 probe sets with Present calls in at least 10 of the 40 patients were used in this analysis. Infant (<1 year old) and child (≥1 year old) patients are indicated by the colours. (B) Two-dimensional hierarchical clustering analysis of 40 paediatric AML-M4/M5 patients. 1865 probe sets with high variation in expression (≥2-fold changes to the median value in at least 10 of the 40 patients) were selected from the 25754 probe sets and used in the analysis. Columns and rows in the matrix represent patients and probe sets, respectively. Red and green indicates high and low expression, respectively. Indicated below the matrix are the age group (infant or child), karyotype [*MLL*gr(+) or *MLL*gr(−)] and FAB subtype (M4 or M5) for each patient by colour codes. (C) PCA analysis of 40 paediatric and 14 adult AML-M4/M5 patients. 25605 probe sets with Present calls in at least 15 of the 54 patients were used in the analysis. Infant, child and adult patients are indicated.

information. First, we isolated 2421 probe sets overexpressed in 10 distinct infants by selecting probe sets that exhibited *P*-values of <0·01 in Students *t*-test and twofold higher expression when the 10 infants were compared with the other remaining patients (Fig 2A and genes listed in Table SIV). With the average normalized expression value of these genes, we then divided the 40 patients into subgroups, designated as A, B and C. Subgroup A had an average normalized value greater than 2·0, whereas subgroup C had an average normalized value less than 1·0 (Fig 2A and B, see also Table SI). All of the 10 distinct infants were selected as subgroup A, and 12 and 18 of the remaining 30 patients were divided into subgroup B and C, respectively.

As shown in Fig 2C and D, age distribution and karyotypes were different among the three subgroups. The average age of subgroup A was 0·3 years old with eight out of 10 patients' possessing *MLL* gene rearrangement. The average age of subgroup B was 3·1 years old, and seven of the 12 patients possessed *MLL* gene rearrangement with t(9;11) as being the dominant karyotype. Subgroup C included many of the older children with the average age of 6·6 years old, and eight of the 18 patients possessed *MLL* gene rearrangement. Interestingly, all the t(6;11) patients (average age 10·0 years old) were placed in subgroup C. FAB subtypes were equally distributed among each subgroup (Fig 2E).

To compare the clinical outcomes among the subgroups, we performed Kaplan–Meier analysis and log-rank test. Subgroup C had the worst outcome (3-year EFS 28%), which was significantly lower than those of subgroup A (3-year EFS 70%) and subgroup B (3-year EFS 75%) ($P = 0·023$) (Fig 2F). The outcome of subgroup C was definitely lower than that of the intermediate-risk patients in AML 99 study (3-year EFS 60%) (Tsukimoto *et al*, 2005).

To investigate the underlying biology in paediatric AML-M4/M5, we selected the genes that differentially expressed in subgroups A and C (Table II and SV). Three thousand and thirty-three probe sets were identified by selecting genes with *P*-values of <0·0001 and fold change values of >2·0 or <0·5. Among these genes, 2487 probe sets (82%) were higher in subgroup A and 546 probe sets (18%) were higher in subgroup C. Almost all of

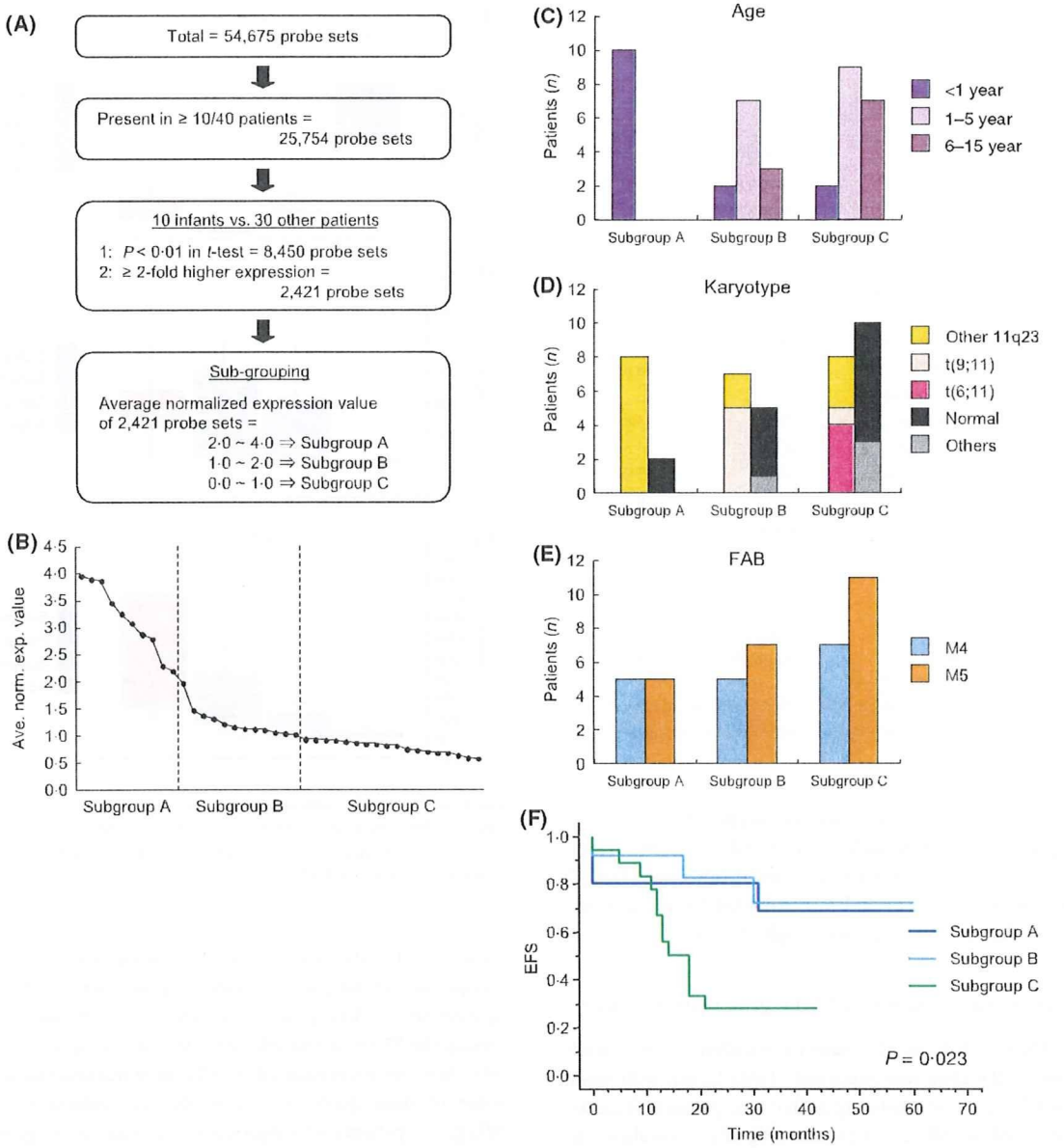

Fig 2. Sub-grouping of paediatric AML-M4/M5. (A) Sub-grouping scheme. First, subgroup-discriminating probe sets were selected as follows: (i) probe sets without Present calls in at least 10 of 40 paediatric patients were removed (25754 probe sets); (ii) probe sets with *P*-value of <0·01 in Student's *t*-test comparing 10 distinct infants and the remaining 30 patients were selected (8450 probe sets) and (iii) probe sets with ≥2-fold higher expression in the 10 distinct infants were chosen (2421 probe sets). Then, the average of normalized expression values of the 2421 subgroup-discriminating probe sets was calculated for each patient and used for sub-grouping. 10 patients with the average values between 2·0 and 4·0 were assigned as subgroup A, 12 patients with the average values between 1·0 and 2·0 as subgroup B, 18 patients with the average values lower than 1·0 were assigned as subgroup C. (B) Distribution of the average normalized expression values of 40 patients. (C) Age distribution of each subgroup. The average age was 0·3, 3·1 and 6·6 years old for subgroup A, B and C, respectively. (D, E) Karyotype and FAB subtype distribution of each subgroup. (F) Kaplan–Meier analysis of event-free survival (EFS). *P*-value was obtained by log-rank test.

the selected genes (2995 probe sets, 99%) showed medium expression level in subgroup B whether their expression were higher in subgroup A or in subgroup C, confirming that subgroups A, B and C have a linear correlation in gene expression. Among the genes overexpressed in subgroup A, the top-ranked gene (25·3-fold) was a DNA-binding zinc finger homeobox protein gene *ZEB2* (Table II). The expression of this gene was higher in adult patients than in subgroup C patients (Fig 3A). Interestingly, *MLL* translocation partner genes such as *ARHGAP26* and *FOXO3* were also overexpressed

Table II. Genes expressed differentially between subgroups A and C.

Fold rank	Probe set ID	Gene symbol	Fold change
Highly expressed in subgroup A			
1	239296_at	*ZEB2*	25·3
2	242008_at	*AGTPBP1*	21·3
3	240652_at	Unknown	21·0
4	1556590_s_at	Unknown	18·1
5	1562194_at	*ZEB2*	17·3
6	244548_at	*ARHGAP26*	17·0
7	1557797_a_at	*ZEB2*	16·6
8	239555_at	*LYN*	16·5
9	1569477_at	*FOXO3*	16·4
10	232307_at	*MED13L*	16·4
Highly expressed in subgroup C			
1	206310_at	*SPINK2*	23·8
2	203373_at	*SOCS2*	13·5
3	201427_s_at	*SEPP1*	12·8
4	205051_s_at	*KIT*	9·6
5	209160_at	*AKR1C3*	8·9
6	206067_s_at	*WT1*	7·1
7	206772_at	*PTH2R*	6·8
8	210140_at	*CST7*	6·7
9	200923_at	*LGALS3BP*	6·5
10	236738_at	*LOC401097*	6·2

Probe sets were selected with Student's *t*-test and fold changes. Top 10 probe sets were listed in the order of their fold change values.

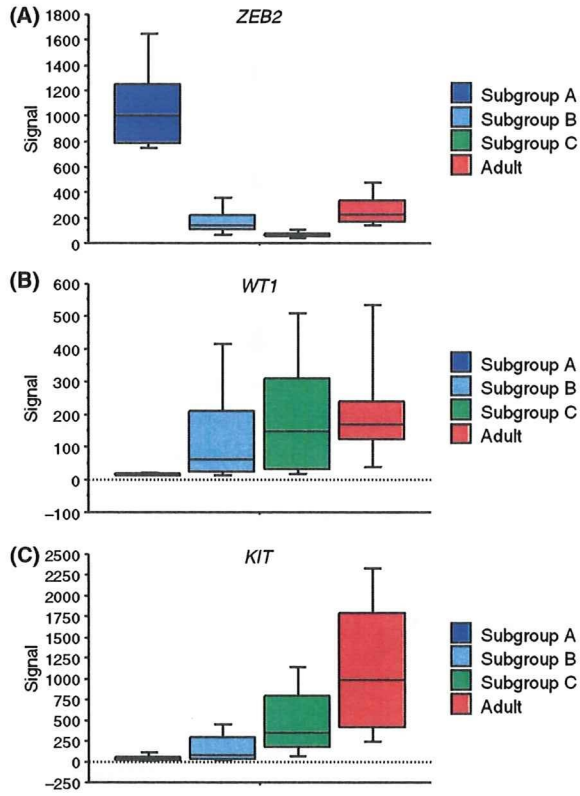

Fig 3. Differentially expressed genes between subgroup A and subgroup C. Box and whiskers plots are shown for *ZEB2* (A), *WT1* (B) and *KIT* (C). Signal values were compared among subgroup A, B and C paediatric patients and adult patients.

in subgroup A regardless of the karyotype. In subgroup C, poor prognostic factors such as *WT1* and *KIT* were overexpressed, coinciding with its poor clinical outcome (Table II and Fig 3B and C). *WT1* and *KIT* exhibited linear increment with the adult patients expressing highest of all.

Gene expression signatures of MLL gene rearrangements

More than a half of the patients enrolled in this study possessed *MLL* gene rearrangement (Table I), but with unsupervised hierarchical clustering analysis, no prominent cluster was observed for *MLL*gr(+) patients (Fig 1B). To evaluate the differences between *MLL*gr(+) AML-M4/M5 and *MLL*gr(−) AML-M4/M5, we used SAM and selected differentially expressed probe sets, at 5% false discovery rate (FDR), between *MLL*gr(+) and *MLL*gr(−) patients among all patients and also within each subgroup (Table III and SVI). When all the patients were compared, 112 probe sets were selected. Comparison within subgroup A did not exhibit any significant difference between *MLL*gr(+) and *MLL*gr(−) patients and this was probably due to the scarcity of *MLL*gr(−) patients. Only 11 and 55 probe sets were selected at 5% FDR when compared within subgroup B and subgroup C, respectively (Table III). Surprisingly, commonly selected probe sets were very small (Fig 4A), indicating that the gene expression signatures of *MLL* gene rearrangement differed considerably among the different subgroups identified. Expression patterns of selected genes are shown in Fig 4B. Among the 112 probe sets selected in comparison of all patients, only 17 probe sets (15%) were upregulated by *MLL* gene rearrangement. On the other hand, among the 55 probe sets selected within subgroup C, 53 probe sets (96%) were upregulated by *MLL* gene rearrangement, and some of these probe sets were also upregulated in some *MLL*gr(+) patients of subgroup B but not in subgroup A patients (Fig 4B), indicating a clear difference of *MLL* gene rearrangement signatures among the subgroups.

Table III. The numbers of probe sets showing differential expression between *MLL*gr(+) and *MLL*gr(−).

	Patient no. *MLL*gr(+) vs. *MLL*gr(−)	5% FDR
All patients	(23 vs. 17)	112
Subgroup A	(8 vs. 2)	0
Subgroup B	(7 vs. 5)	11
Subgroup C	(8 vs. 10)	55

Probe sets were selected using SAM with 5% false discovery rate (FDR).

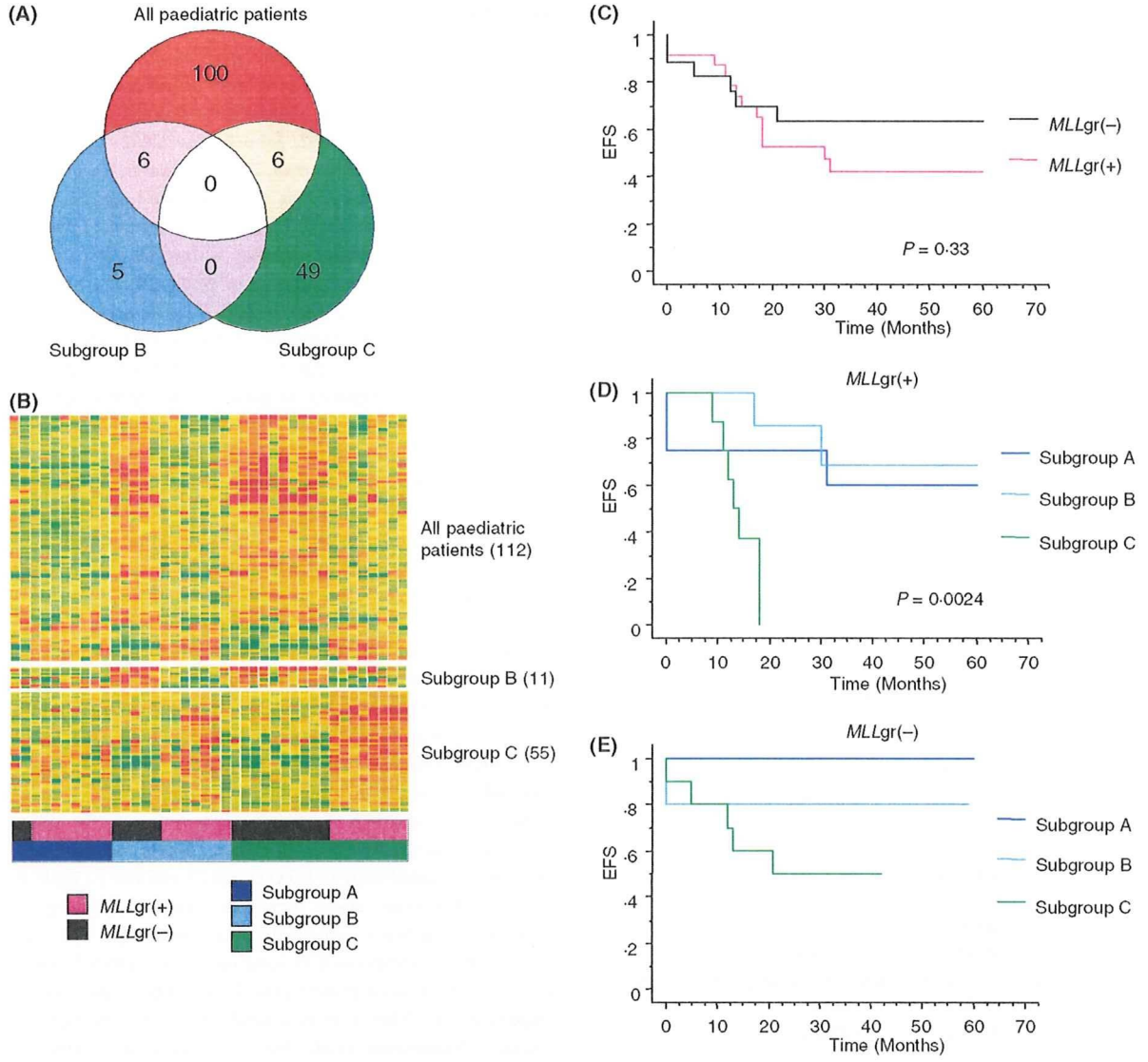

Fig 4. The significance of *MLL* gene rearrangements in paediatric AML-M4/M5. (A) Venn diagram comparing the differentially expressed genes between *MLL*gr(+) and *MLL*gr(−) patients. The probe sets selected by SAM with 5% FDR were compared among subgroup B, subgroup C and all paediatric patients. (B) Expression pattern of the SAM-selected probe sets. Columns and rows in the matrix represent patients and probe sets, respectively. The normalized expression level for each probe set is represented by colour, with red representing deviation above the median and green representing deviation below the median. The *MLL* gene rearrangement [*MLL*gr(+) or *MLL*gr(−)] and paediatric AML-M4/M5 subgroup (A, B or C) of each patient are indicated by colour codes below the matrix. (C–E) Kaplan–Meier analysis of event-free survival (EFS). *MLL*gr(+) and *MLL*gr(−) patients within paediatric patients (C), subgroups A, B and C within *MLL*gr(+) patients (D) and subgroups A, B and C within *MLL*gr(−) patients (E) were compared. *P*-value was obtained by log-rank test.

Genes specifically overexpressed throughout all of the *MLL*gr(+) patients included a zinc finger protein gene *ZNF521* (Table IV and Fig 5A), which has been reported previously (Kohlmann *et al*, 2005). However, frequently reported genes such as *HOXA* family members and *MEIS1* (Armstrong *et al*, 2002; Yeoh *et al*, 2002; Tsutsumi *et al*, 2003; Bullinger *et al*, 2004; Ross *et al*, 2004; Valk *et al*, 2004; Andersson *et al*, 2005) were not selected with SAM and *HOXA9* had no correlation with *MLL* gene rearrangement (Fig 5D). Rather, the *HOXB2* and *HOXB3* genes were selected as underexpressed genes among *MLL*gr(+) patients (Table IV and Fig 5B). A transcription factor gene *EVI1*, whose overexpression is known to associate with poor prognosis, was upregulated in subgroup C *MLL*gr(+) patients (Table IV and Fig 5C) (Doorn-Khosrovani *et al*, 2003; Valk *et al*, 2004). These data were also confirmed by quantitative RT-PCR (Fig 6).

Table IV. Genes expressed differentially between *MLL*gr(+) and *MLL*gr(−) patients.

Fold rank	Probe set ID	Gene symbol	Fold change
Highly expressed in *MLL*gr(+) patients among all patients			
1	226676_at	*ZNF521*	19·0
2	226677_at	*ZNF521*	15·8
3	1552665_at	*MGC14425*	7·9
4	227929_at	Unknown	4·9
5	217520_x_at	*LOC646278*	4·9
6	206440_at	*LIN7A*	3·1
7	205472_s_at	*DACH1*	3·0
8	1557450_s_at	*WHDC1L1*	2·8
9	225785_at	*REEP3*	2·7
10	205471_s_at	*DACH1*	2·5
Highly expressed in *MLL*gr(−) patients among all patients			
1	228904_at	*HOXB3*	13·1
2	205453_at	*HOXB2*	10·8
3	205033_s_at	*DEFA1*	10·2
4	236892_s_at	Unknown	8·5
5	223828_s_at	*LGALS12*	5·1
6	202833_s_at	*SERPINA1*	5·0
7	200923_at	*LGALS3BP*	4·8
8	211429_s_at	*SERPINA1*	4·5
9	206761_at	*CD96*	4·4
10	239791_at	*LOC404266*	4·3
Highly expressed in *MLL*gr(+) patients among subgroup C patients			
1	235700_at	*CT45-4*	28·2
2	1567912_s_at	*CT45-4*	25·4
3	226676_at	*ZNF521*	22·3
4	236858_s_at	Unknown	17·9
5	226677_at	*ZNF521*	16·9
6	226420_at	*EVI1*	12·1
7	1552665_at	*MGC14425*	11·2
8	1559266_s_at	*LOC730417*	10·5
9	236859_at	*RUNX2*	10·0
10	1557261_at	*WHDC1L1*	10·0
Highly expressed in *MLL*gr(−) patients among subgroup C patients			
1	209298_s_at	*ITSN1*	2·7
2	209357_at	*CITED2*	2·6

Probe sets were selected with SAM at 90th percentile with 5% false discovery rate (FDR). Top 10 or all selected probe sets were listed in the order of their fold change values.

When compared between all patients, EFS did not differ statistically ($P = 0{\cdot}33$) between *MLL*gr(+) and *MLL*gr(−) patients whilst *MLL*gr(+) patients seemed to have a poorer prognosis (Fig 4C). The EFS of each subgroup was compared within each *MLL* gene arrangement status and subgroup C presented the poorest outcome for both *MLL*gr(+) and *MLL*gr(−) karyotypes (Fig 4D and E). It was astounding to see that all subgroup C *MLL*gr(+) patients (eight patients) had relapsed within 18 months (3-year EFS 0%). Subgroup C represented an extremely poor prognostic subgroup with early relapse especially with *MLL* gene rearrangement.

Discussion

Microarray-based gene expression profiling has been applied to the diagnosis of many different types of leukaemia and used to explore their underlying molecular pathology (Golub *et al*, 1999; Armstrong *et al*, 2002; Schoch *et al*, 2002; Yeoh *et al*, 2002; Yagi *et al*, 2003; Ross *et al*, 2004). However, no previous study has been reported that focussed solely on the gene expression profile of AML-M4/M5 or on the difference between infants and older children with AML. The present study analysed the gene expression of 40 paediatric AML-M4/M5 patients and found that typical infants displayed very unique and distinct expression patterns that are clearly different from the older patients. These distinctly expressed genes were used to separate 40 patients into three subgroups without any previous knowledge of the patients' clinical information, and showed that these subgroups differed in age distribution and clinical outcome. In addition, the gene expression signature of *MLL* gene rearrangement and outcomes of *MLL*gr(+) patients were different among the different subgroups. These results indicate that paediatric AML-M4/M5 is heterogeneous in an age-associated manner and this heterogeneity is correlated with their prognosis.

The three subgroups that were separated by gene expression were different in age distribution and in clinical outcome, but not in FAB subtype (Fig 2). The difference between subgroup A and subgroup C was extremely prominent. A large set of differentially expressed genes was isolated and these genes were medianly expressed in subgroup B. Among the genes overexpressed in subgroup A, many transcription factor genes involved in haematopoietic cell differentiation and proliferation were included such as *FOXO3*, *MLL3* and *ELF2* (Table II and S5). These observations suggest that underlying biology is transitionally different among the three subgroups. The top-ranked gene overexpressed in subgroup A is a DNA-binding zing finger homeobox protein gene *ZEB2*. *ZEB2* mutations are responsible for Mowat–Wilson syndrome and Hirschsprung disease (Wakamatsu *et al*, 2001; Mowat *et al*, 2003). It functions in the transforming growth factor β signaling pathway and may promote tumor invasion in hepatic, renal and breast cancer (Miyoshi *et al*, 2004; Elloul *et al*, 2005; Krishnamachary *et al*, 2006), but the association between *ZEB2* and leukaemia is still unknown. Genes that were overexpressed in subgroup C included *WT1* and *KIT*; overexpression of these genes is known to be associated with poor prognosis (Ashman *et al*, 1991; Inoue *et al*, 1994; Weisser *et al*, 2005; Paschka *et al*, 2006; Shimada *et al*, 2006) and this overexpression may be responsible for the dismal outcome of subgroup C patients.

When adult AML-M4/M5 patients were included in the analysis, it was very interesting to find that adult patients not only shared common expression pattern with older children (subgroup C) but many of the genes upregulated in infants (subgroup A) were similarly highly expressed among the adult patients (Fig S1). It is well known that certain types of leukaemia are common only in the paediatric population and

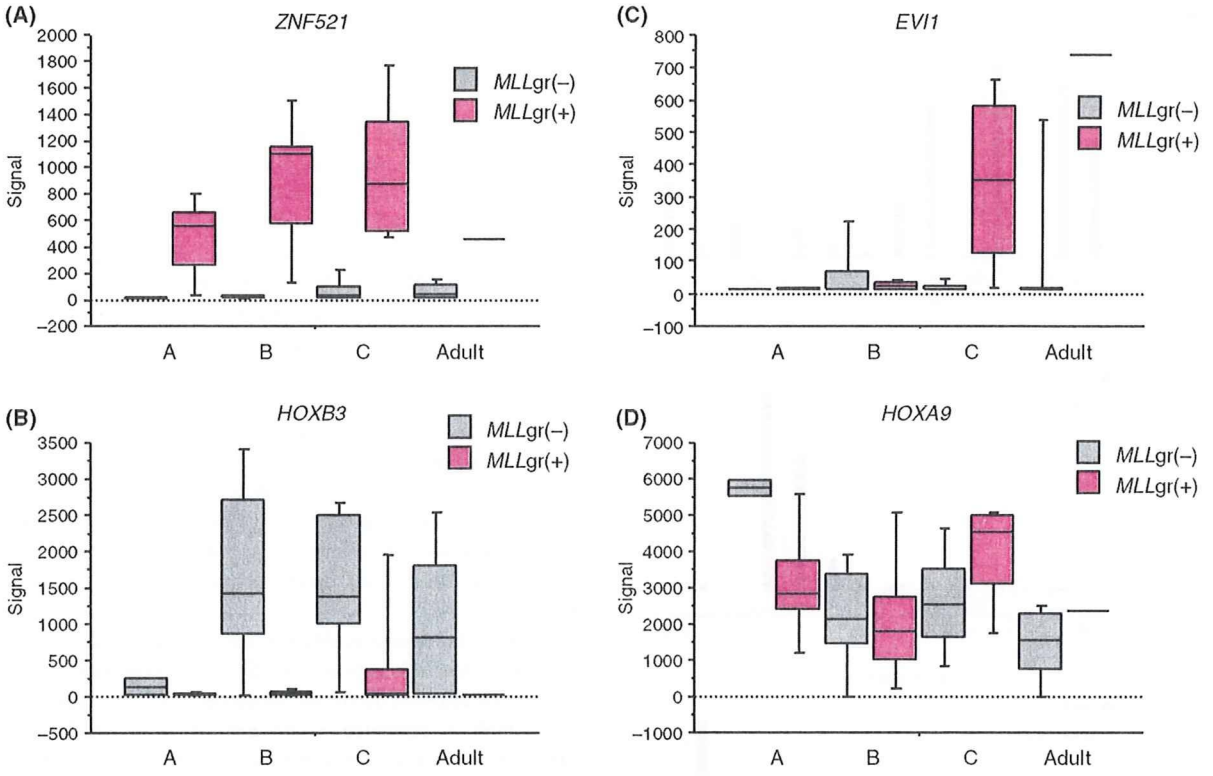

Fig 5. Differentially expressed genes between *MLL*gr(+) and *MLL*gr(−) patients. Box and whiskers plot are shown for *ZNF521* (A), *HOXB3* (B), *EVI1* (C) and *HOXA9* (D). Signal values were compared between *MLL*gr(+) and *MLL*gr(−) patients within subgroup A, B and C paediatric patients and adult patients.

vice versa. For example, *MLL* gene rearrangements are much more frequent among younger populations while *NPM1* mutations and *FLT3*-internal tandem duplication (ITD) are typically prevalent in the adult population (Iwai *et al*, 1999; Meshinchi *et al*, 2001; Cazzaniga *et al*, 2005; Falini *et al*, 2005; Brown *et al*, 2007). With these differences in genetic alterations, we had expected that gene expression would be transitional in an age-increasing manner. In previous studies, the gene expression signatures of genetic alterations were compared between adult and paediatric leukaemia. Kohlmann *et al* (2004) have reported that gene expression signatures identified in paediatric ALL were applicable to adult ALL patients. Ross *et al* (2004) have also shown that the expression signatures generated from paediatric AML were able to accurately classify adult AML with the same genetic lesions. The similarities or differences between infant AML and adult AML has never been thoroughly discussed before, and thus we have identified for the first time that infant AML-M4/M5 and adult AML-M4/M5 share expression patterns. Among the commonly shared genes between infant AML-M4/M5 (subgroup A) and adult AML-M4/M5, transcription factor genes involved in haematopoietic cell differentiation and proliferation such as *ETV6* and *RUNX1* were included, and this may suggest that these two groups might share common underlying biology.

Many studies have already reported that AML with *MLL* gene rearrangement present distinct gene expression profiles (Armstrong *et al*, 2002; Yeoh *et al*, 2002; Tsutsumi *et al*, 2003; Bullinger *et al*, 2004; Ross *et al*, 2004; Valk *et al*, 2004; Andersson *et al*, 2005; Kohlmann *et al*, 2005), but with our unsupervised analysis, no prominent cluster of *MLL*gr(+) patients was observed. This inconsistency may be attributable to the fact that we focused solely on AML-M4/M5, while previous studies investigated the whole AML population to identify the signature of *MLL* gene rearrangements. We speculated that *MLL*gr(+) AML-M4/M5 and *MLL*gr(−) AML-M4/M5 might share similar gene expression especially in paediatric patients. When we examined the differences between *MLL*gr(+) and *MLL*gr(−) patients in AML-M4/M5, relatively small differences (112 probe sets at 5% FDR) were found and almost all of these genes were under-expressed in *MLL*gr(+) patients. In addition, *HOXA* family genes and *MEIS1*, which have been repeatedly identified as overexpressed genes in *MLL*gr(+) leukaemia (Armstrong *et al*, 2002; Yeoh *et al*, 2002; Tsutsumi *et al*, 2003; Bullinger *et al*, 2004; Ross *et al*, 2004; Valk *et al*, 2004; Andersson *et al*, 2005; Kohlmann *et al*, 2005), were not included in the selected gene lists. Rather, under-expression of *HOXB2* and *HOXB3* in *MLL*gr(−) patients was observed (Table IV). Among the

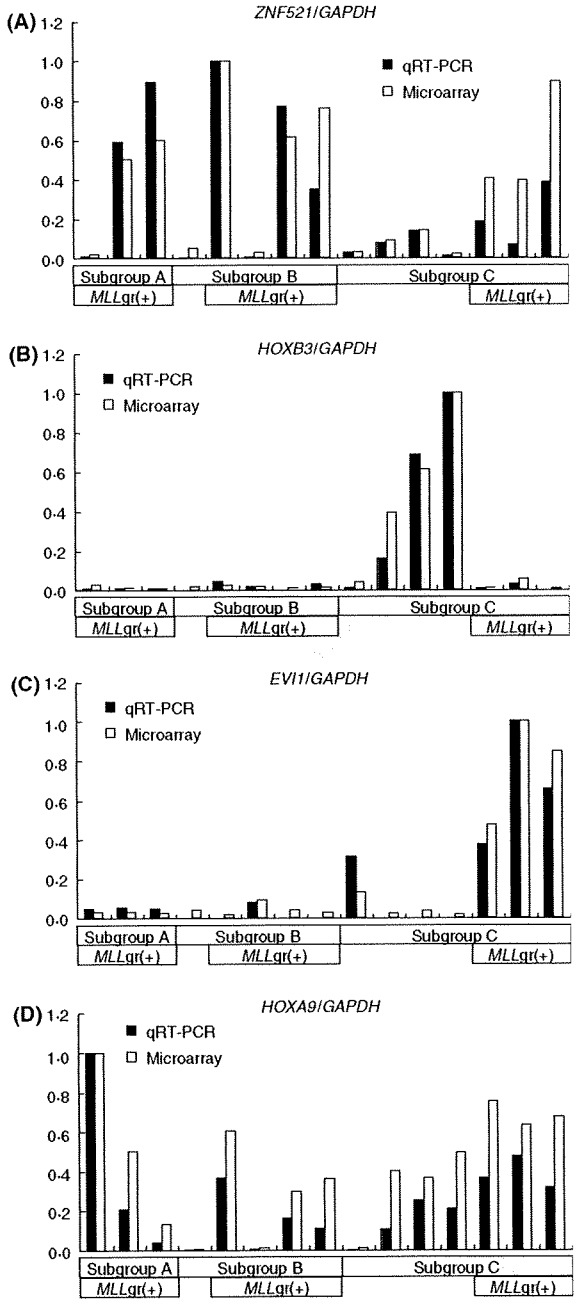

Fig 6. Quantitative RT-PCR confirmation of differential expression between *MLL*gr(+) and *MLL*gr(−) patients. Quantitative RT-PCR was performed for *ZNF521* (A), *HOXB3* (B), *EVI1* (C) and *HOXA9* (D) on RNA samples from 15 patients. Included were 3 subgroup A [3 *MLL*gr(+)], 5 subgroup B [4 *MLL*gr(+) and 1 *MLL*gr(−)] and 7 subgroup C [3 *MLL*gr(+) and 4 *MLL*gr(−)] patients. Transcript levels of those 4 genes were normalized to that of *GAPDH*, and presented as ratios to the highest patient. Their expression levels determined by microarray analysis were also shown as the same ratios after normalization to the *GAPDH* expression level.

previously reported *MLL* gene rearrangement-associated genes, *ZNF521* was specifically overexpressed among all of the *MLL*gr(+) patients in our study. *ZNF521* is an early haematopoietic zinc finger protein gene that is highly expressed in primitive human haematopoietic cells (Bond *et al*, 2004), and *PAX5-ZNF521* gene fusion has been recently reported among B-progenitor ALL (Mullighan *et al*, 2007). This gene, which is likely to be involved in normal haematopoiesis and ALL development, may also be active in the leukaemogenesis of *MLL*gr(+) AML-M4/M5. On the other hand, we also found that some *MLL* gene rearrangement-associated genes were subgroup-specific. When compared within subgroup C, 55 probe sets were selected as differentiating genes for *MLL*gr(+) and *MLL*gr(−) patients at 5% FDR, but the expression patterns of these probe sets were not shared by the other subgroups, especially subgroup A. This is noteworthy because EFS of *MLL*gr(+) patients were very different among different subgroups and subgroup C *MLL*gr(+) patients had the poorest outcome (Fig 4D). These findings suggest that the target genes of *MLL* fusion proteins might be different among each subgroup.

Our gene expression profiling analysis revealed that paediatric AML-M4/M5 consists of different subgroups with distinctively different outcomes. Subgroup C, which comprised older children, presented 3-year EFS of 28%, which was extremely lower than the previously reported EFS of intermediate risk patients (60%) (Tsukimoto *et al*, 2005). Our sub-grouping procedure enabled us to extract these poor prognostic patients from the other intermediate risk patients. It is known that some mutations, such as *FLT3*-ITD, *KIT* mutation and *NPM1* mutation, influence the prognosis of AML and their incidences increase in an age-dependent manner. In our analysis, however, these mutations were rare not only in subgroup B but also in subgroup C (Table SI), indicating that these factors are not the main reason for the dismal outcome of subgroup C. Subgroup C contained both *MLL*gr(+) and *MLL*gr(−) patients but subgroup C *MLL*gr(+) patients especially had a disappointing outcome, as all eight patients had relapsed within 18 months. Subgroup C included different types of *MLL* gene rearrangement, but t(6;11) was present only in subgroup C. Previous reports have also suggested that the prognosis of t(6;11) is dismal (Welbron *et al*, 1993; Martineau *et al*, 1998). In our study, the clinical outcome of t(6;11) was also very disappointing as all four patients had relapsed within 18 months (Fig S2A). In addition, we were able to isolate the genes discriminating t(6;11) and the other *MLL* gene rearrangements with SAM (Table SVII and SVIII). These results provide another possibility that patients with t(6;11) exhibit different expression profiles and dismal outcome. However, some of the upregulated genes in t(6;11) patients were also upregulated in *MLL*gr(−) patients of subgroup C (Fig S2B). So some of the selected genes for t(6;11) represent not only t(6;11) patients but rather represent subgroup C patients. Moreover, the outcome of other subgroup C *MLL*gr(+) patients is also dismal, although they

did not show high expression of those genes. These data suggest that t(6;11) karyotype is not an independent prognostic factor but rather a surrogate marker for subgroup C *MLL*gr(+) patients.

In summary, paediatric AML-M4/M5 was identified as a group of heterogeneous diseases comprised different outcomes. Adverse prognostic subgroup could be distinguished with the use of gene expression profiling. Our results have shown that age is an important factor contributing to the biology of AML-M4/M5. These findings also suggest that gene expression profiling may be very useful for risk estimation and our sub-grouping procedures could be a powerful risk stratification tool to identify unfavourable risk patients within the heterogeneous paediatric AML-M4/M5. Further studies are needed to explain the biology underlying different gene expression between infants and older children and to verify the usefulness of our sub-grouping procedures for risk-stratified therapeutics.

Acknowledgements

The authors thank the members of Japan Infant Leukemia Study Group (MLL96/98) and Japanese Childhood AML Cooperative Study Group (AML99). This work was supported by a Grant-in-Aid for Third Term Comprehensive Control Research for Cancer from the Ministry of Health, Labour and Welfare and by the programme for promotion of Fundamental Studies in Health Sciences of National Institute of Biomedical Innovation (NiBio).

References

Andersson, A., Olofsson, T., Lindgren, D., Nilsson, B., Ritz, C., Eden, P., Lassen, C., Rade, J., Fontes, M., Morse, H., Heldrup, J., Behrendtz, M., Mitelman, F., Hoglund, M., Johansson, B. & Fioretos, T. (2005) Molecular signatures in childhood acute leukemia and their correlations to expression patterns in normal hematopoietic subpopulations. *Proceedings of the National Academy of Sciences of the United States of America*, **102,** 19069–19074.

Armstrong, S.A., Staunton, J.E., Silverman, L.B., Pieters, R., den Boer, M.L., Minden, M.D., Sallan, S.E., Lander, E.S., Golub, T.R. & Korsmeyer, S.J. (2002) MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nature Genetics*, **30,** 41–47.

Ashman, L.K., Cambareri, A.C., To, L.B., Levinsky, R.J. & Juttner, C.A. (1991) Expression of the YB5.B8 antigen (c-kit proto-oncogene product) in normal human bone marrow. *Blood*, **78,** 30–37.

Bond, H.M., Mesuraca, M., Carbone, E., Bonelli, P., Agosti, V., Amodio, N., De Rosa, G., Di Nicola, M., Gianni, A.M., Moore, M.S., Hata, A., Grieco, M., Morrone, G. & Venuta, S. (2004) Early hematopoietic zinc finger protein (EHZF), the human homolog to mouse Evi3, is highly expressed in primitive human hematopoietic cells. *Blood*, **103,** 2062–2070.

Brown, P., McIntyre, E., Rau, R., Meshinchi, S., Lacayo, N., Dahl, G., Alonzo, T.A., Chang, M., Arececi, R.J. & Small, D. (2007) The incidence and clinical significance of nucleophosmin mutations in childhood AML. *Blood*, **110,** 979–985.

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.

Cazzaniga, G., Dell'Oro, M.G., Mecucci, C., Giarin, E., Masetti, R., Rossi, V., Locatelli, F., Martelli, M.F., Basso, G., Pession, A., Biondi, A. & Falini, B. (2005) Nucleophosmin mutations in childhood acute myelogenous leukemia with normal karyotype. *Blood*, **106,** 1419–1422.

Doorn-Khosrovani, S.B., Erpelinck, C., van Putten, W.L., Valk, P.J., Luytgaarde, S., Hack, R., Slater, R., Smit, E.M., Beverloo, B., Verhoef, G., Verdonck, L.F., Ossenkoppele, G.J., Sonneveld, P., de Greef, G.E., Lowenberg, B. & Delwel, R. (2003) High EVI1 expression predicts poor survival in acute myeloid leukemia: a study of 319 de novo AML patients. *Blood*, **101,** 837–845.

Downing, J.R. & Shannon, K.M. (2002) Acute leukemia: a pediatric perspective. *Cancer Cell*, **2,** 437–445.

Elloul, S., Elstrand, M.K., Nesland, J.M., Trope, C.G., Kvalheim, G., Goldberg, I., Reich, R. & Davidson, B. (2005) Snail, Slug, and Smad-interacting protein 1 as novel parameters of disease aggressiveness in metastatic ovarian and breast carcinoma. *Cancer*, **103,** 1631–1643.

Falini, B., Mecucci, C., Tiacci, E., Alcalay, M., Rosati, R., Pasqualucci, L., La Starza, R., Diverio, D., Colombo, E., Santucci, A., Bigerna, B., Pacini, R., Pucciarini, A., Liso, A., Vignetti, M., Fazi, P., Meani, N., Pettirossi, V., Saglio, G., Mandelli, F., Lo-Coco, F., Pelicci, P.G. & Martelli, M.F. (2005) Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. *The New England Journal of Medicine*, **352,** 254–266.

Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A., Bloomfield, C.D. & Lander, E.S. (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*, **286,** 531–537.

Horibe, K., Tsukimoto, I. & Ohno, R. (2001) Clinicopathologic characteristics of leukemia in Japanese children and young adults. *Leukemia*, **15,** 1256–1261.

Inoue, K., Sugiyama, H., Ogawa, H., Nakagawa, M., Yamagami, T., Miwa, H., Kita, K., Hiraoka, A., Masaoka, T., Nasu, K., Kyo, T., Dohy, H., Nakauchi, H., Ishidate, T., Akiyama, T. & Kishimoto, T. (1994) WT1 as a new prognostic factor and a new marker for the detection of minimal residual disease in acute leukemia. *Blood*, **84,** 3071–3079.

Iwai, T., Yokota, S., Nakao, M., Okamoto, T., Taniwaki, M., Onodera, N., Watanabe, A., Kikuta, A., Tanaka, A., Asami, K., Sekine, I., Mugishima, H., Nishimura, Y., Koizumi, S., Horikoshi, Y., Mimaya, J., Ohta, S., Nishikawa, K., Iwai, A., Shimokawa, T., Nakayama, M., Kawakami, K., Gushiken, T., Hyakuna, N., Katano, N., Tsurusawa, M. & Fujimoto, T. (1999) Internal tandem duplication of the FLT3 gene and clinical evaluation in childhood acute myeloid leukemia. *Leukemia*, **13,** 38–43.

Kawasaki, H., Isoyama, K., Eguchi, M., Hibi, S., Kinukawa, N., Kosaka, Y., Oda, T., Oda, M., Nishimura, S., Imaizumi, M., Okamura, T., Hongo, T., Okawa, H., Mizutani, S., Hayashi, Y., Tsukimoto, I., Kamada, N. & Ishii, E. (2001) Superior outcome of infant acute myeloid leukemia with intensive chemotherapy: results of the Japan Infant Leukemia Study Group. *Blood*, **98,** 3589–3594.

Kohlmann, A., Schoch, C., Schnittger, S., Dugas, M., Hiddemann, W., Kern, W. & Haferlach, T. (2004) Pediatric acute lymphoblastic

leukemia (ALL) gene expression signatures classify an independent cohort of adult ALL patients. *Leukemia*, **18,** 63–71.

Kohlmann, A., Schoch, C., Dugas, M., Schnittger, S., Hiddemann, W., Kern, W. & Haferlach, T. (2005) New insights into MLL gene rearranged acute leukemias using gene expression profiling: shared pathways, lineage commitment, and partner genes. *Leukemia*, **19,** 953–964.

Krishnamachary, B., Zagzag, D., Nagasawa, H., Rainey, K., Okuyama, H., Baek, J.H. & Semenza, G.L. (2006) Hypoxia-inducible factor-1-dependent repression of E-cadherin in von Hippel-Lindau tumor suppressor-null renal cell carcinoma mediated by TCF3, ZFHX1A, and ZFHX1B. *Cancer Research*, **66,** 2725–2731.

Martineau, M., Berger, R., Lillington, D.M., Moorman, A.V. & Secker-Walker, L.M. (1998) The t(6;11)(q27;q23) translocation in acute leukemia: a laboratory and clinical study of 30 cases. *Leukemia*, **12,** 788–791.

Meshinchi, S., Woods, W.G., Stirewalt, W.D., Sweetser, D.A., Buckley, J.D., Tjoa, T.K., Bernstein, I.D. & Radich, J.P. (2001) Prevalence and prognostic significance of FLT3 internal tandem duplication in pediatric acute myeloid leukemia. *Blood*, **97,** 89–94.

Miyoshi, A., Kitajima, Y., Sumi, K., Sato, K., Hagiwara, A., Koga, Y. & Miyazaki, K. (2004) Snail and SIP1 increase cancer invasion by upregulating MMP family in hepatocellular carcinoma cells. *British Journal of Cancer*, **90,** 1265–1273.

Mowat, D.R., Wilson, M.J. & Goossens, M. (2003) Mowat-Wilson syndrome. *Journal of Medical Genetics*, **40,** 305–310.

Mullighan, C.G., Goorha, S., Radtke, I., Miller, C.B., Coustan-Smith, E., Dalton, J.D., Girtman, K., Mathew, S., Ma, J., Pounds, S.B., Su, X., Pui, C.H., Relling, M.V., Evans, W.E., Shurtleff, S.A. & Downing, J.R. (2007) Genome-wide analysis of genetic alterations in acute lymphoblastic leukemia. *Nature*, **446,** 758–764.

Nagao, T. (1995) Footprints of the Children's Leukemia Study Group of Japan. *The Japanese Journal of Pediatric Hematology*, **9,** 332–340.

Pallisgaard, N., Hokland, P., Riishoj, D.C., Pedersen, B. & Jorgensen, P. (1998) Multiplex reverse transcription-polymerase chain reaction for simultaneous screening of 29 translocations and chromosomal aberrations in acute leukemia. *Blood*, **92,** 574–588.

Paschka, P., Marucci, G., Ruppert, A.S., Mrozek, K., Chen, H., Kittles, R.A., Vukosavljevic, T., Perrotti, D., Vardiman, J.W., Carroll, A.J., Kolitz, J.E., Larson, R.A. & Bloomfield, C.D. (2006) Adverse prognostic significance of KIT mutations in adult acute myeloid leukemia with inv(16) and t(8;21): a Cancer and Leukemia Group B Study. *Journal of Clinical Oncology*, **24,** 3904–3911.

Pui, C.H., Kane, J.R. & Crist, W.M. (1995) Biology and treatment of infant leukemias. *Leukemia*, **9,** 762–769.

Pui, C.H., Raimondi, S.C., Srivastava, D.K., Tong, X., Behm, F.G., Razzouk, B., Rubnitz, J.E., Sandlund, J.T., Evans, W.E. & Ribeiro, R. (2000) Prognostic factors in infants with acute myeloid leukemia. *Leukemia*, **14,** 684–687.

Pui, C.H., Gaynon, P.S., Boyett, J.M., Chessells, J.M., Baruchel, A., Kamps, W., Silverman, L.B., Biondi, A., Harms, D.O., Vilmer, E., Schrappe, M. & Camitta, B. (2002) Outcome of treatment in childhood acute lymphoblastic leukaemia with rearrangements of the 11q23 chromosomal region. *Lancet*, **359,** 1909–1915.

Ross, M.E., Mahfouz, R., Onciu, M., Liu, H., Zhou, X., Song, G., Shurtleff, S.A., Pounds, S., Cheng, C., Ma, J., Ribeiro, R.C., Rubnitz, J.E., Girtman, K., Williams, W.K., Raimondi, S.C., Liang, D.C., Shih, L.Y., Pui, C.H. & Downing, J.R. (2004) Gene expression profiling of pediatric acute myelogeneous leukemia. *Blood*, **104,** 3679–3687.

Salto-Tellez, M., Shelat, S.G., Benoit, B., Rennert, H., Carroll, M., Leonard, D.G., Nowell, P. & Bagg, A. (2003) Multiplex RT-PCR for the detection of leukemia-associated translocations. *Journal of Molecular Diagnostics*, **5,** 231–236.

Satake, N., Maseki, N., Nishiyama, M., Kobayashi, H., Sakurai, M., Inaba, H., Katano, N., Horikoshi, Y., Eguchi, H., Miyake, M., Seto, M. & Kaneko, Y. (1999) Chromosome abnormalities and MLL rearrangements in acute myeloid leukemia of infants. *Leukemia*, **13,** 1013–1017.

Schoch, C., Kohlmann, A., Schnittger, S., Brors, B., Dugas, M., Mergenthaler, S., Kern, W., Hiddemann, W., Eils, R. & Haferlach, T. (2002) Acute myeloid leukemias with reciprocal rearrangements can be distinguished by specific gene expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*, **99,** 10008–10013.

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

Taki, T., Ida, K., Bessho, F., Hanada, R., Kikuchi, A., Yamamoto, K., Sako, M., Tsuchida, M., Seto, M., Ueda, R. & Hayashi, Y. (1996) Frequency and clinical significance of the MLL gene rearrangements in infant acute leukemia. *Leukemia*, **10,** 1303–1307.

Tsukimoto, I., Tawa, A., Hanada, R., Tabuchi, K., Kigasawa, H., Tsuchiya, S., Tsuchida, M., Yabe, H., Nakayama, H., Kudo, K., Kobayashi, R., Hamamoto, K., Imaizumi, M., Morimoto, A. & Horibe, K. (2005) Excellent outcome of risk stratified treatment for childhood acute myeloid leukemia-AML99 trial. For the Japanese Childhood AML Cooperative Study Group. *Blood*, **106.** (Abstract 889).

Tsutsumi, S., Taketani, T., Nishimura, K., Ge, X., Taki, T., Sugita, K., Ishii, E., Hanada, R., Ohki, M., Aburatani, H. & Hayashi, Y. (2003) Two distinct gene expression signatures in pediatric acute lymphoblastic leukemia with MLL rearrangements. *Cancer Research*, **63,** 4882–4887.

Tusher, V.G., Tibshirani, R. & Chu, G. (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences of the United States of America*, **98,** 5116–5121.

Valk, P.J., Verhaak, R.G., Beijen, M.A., Erpelinck, C.A., Doorn-Khosrovani, S.B., Boer, J.M., Beverloo, H.B., Moorhouse, M.J., van der Spek, P.J., Lowenberg, B. & Delwel, R. (2004) Prognostically useful gene-expression profiles in acute myeloid leukemia. *The New England Journal of Medicine*, **350,** 1617–1628.

Wakamatsu, N., Yamada, Y., Yamada, K., Ono, T., Nomura, N., Taniguchi, H., Kitoh, H., Mutoh, N., Yamanaka, T., Mushiake, K., Kato, K., Sonta, S. & Nagaya, M. (2001) Mutations in SIP1, encoding Smad interacting protein-1, cause a form of Hirschsprung disease. *Nature Genetics*, **27,** 369–370.

Weisser, M., Kern, W., Rauhut, S., Schoch, C., Hiddemann, W., Haferlach, T. & Schnittger, S. (2005) Prognostic impact of RT-PCR-based quantification of WT1 gene expression during MRD monitoring of acute myeloid leukemia. *Leukemia*, **19,** 1416–1423.

Welbron, J.L., Jenks, H.M. & Hagemeijer, A. (1993) Unique clinical features and prognostic significance of the translocation (6;11) in acute leukemia. *Cancer Genetics and Cytogenetics*, **65,** 125–129.

Yagi, T., Morimoto, A., Eguchi, M., Hibi, S., Sako, M., Ishii, E., Mizutani, S., Imashuku, S., Ohki, M. & Ichikawa, H. (2003)

Identification of a gene expression signature associated with pediatric AML prognosis. *Blood*, **102,** 1849–1856.

Yeoh, E.J., Ross, M.E., Shurtleff, S.A., Williams, W.K., Patel, D., Mahfouz, R., Behm, F.G., Raimondi, S.C., Relling, M.V., Patel, A., Cheng, C., Campana, D., Wilkins, D., Zhou, X., Li, J., Liu, H., Pui, C.H., Evans, W.E., Naeve, C., Wong, L. & Downing, J.R. (2002) Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell*, **1,** 133–143.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Expression profiles of AML-M4/M5.

Figure S2. The significance of *MLL* gene rearrangements partners.

Table SI. Clinical information of paediatric patients enrolled in this study.

Table SII. Clinical information of adult patients enrolled in this study.

Table SIII. Primer sequences used for quantitative RT-PCR.

Table SIV. Top 50 probe sets in overexpressed genes in 10 typical infants.

Table SV. Top 50 probe sets in differentially expressed genes between subgroup A and subgroup C patients.

Table SVI. Differentially expressed genes between *MLL*gr(+) and *MLL*gr(−) patients.

Table SVII. Differentially expressed genes among *MLL*gr(+) patients.

Table SVIII. The numbers of probe sets showing differential expression among *MLL*gr(+) patients.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

