LYMPHOID NEOPLASIA

Clinical and molecular genetic characterization of wild-type *MLL* infant acute lymphoblastic leukemia identifies few recurrent abnormalities.

SUPPLEMENTAL DATA

Marieke H. van der Linden[1], Judith M. Boer[1], Pauline Schneider[1], Merel Willekes[1], Lidija Seslija[1], Paola De Lorenzo[2,3], Maria Grazia Valsecchi[2], Giovanni Cazzaniga[3], Andrea Biondi[3], Monique L. den Boer[1], Rob Pieters[4], Ronald W. Stam[1]

[1] Department of Pediatric Oncology/Hematology, Erasmus MC-Sophia Children's Hospital, Rotterdam, The Netherlands; [2] Interfant Trial Data Center, Department of Health Sciences, University of Milano-Bicocca, Monza, Italy; [3] Centro Ricerca Tettamanti, Clinica Pediatrica Università Milano-Bicocca, Ospedale San Gerardo/Fondazione MBBM, Monza, Italy; [4] Princess Maxima Center for Pediatric Oncology, Utrecht, The Netherlands

Correspondence: Ronald W. Stam, Erasmus MC - Sophia Children's Hospital, Department of Pediatric Oncology/Hematology, Room Na-1611, Wytemaweg 80, P.O. Box 2060, 3000 CB, Rotterdam, The Netherlands; phone: +31-10-7044654; fax: +31-10-7044708; e-mail: r.stam@erasmusmc.nl.

Materials and Methods

Patient samples and sample preparation

Bone marrow and peripheral blood samples from wild-type *MLL* infant ALL patients (<1 year of age) with newly diagnosed ALL were included from the international collaborative Interfant-99[1] and Interfant-06 studies up till February 2010 (n=78). Out of these samples 36 samples were used for gene expression profiling, of which for thirty patients survival data was available. Array CGH was done on 31 out of 83 samples. Patient samples were representative for the groups in terms of clinical characteristics. The *MLL*-rearranged patient data used in this study was drawn from the Interfant-99 study (n=70). Pediatric (non-infant) patient data was obtained from the Dutch Childhood Oncology Group (DCOG) treated according to the ALL-10 protocol (n=484), which were all used for their clinical data. Array CGH was done on a selected group of 115 patients, from which T-ALL patients were excluded. B-others are those patients that cannot be classified into any known genetic subgroup. Further details of this group are described in the paper published by Den Boer *et al*.[2] Leukemic samples preparation was essentially carried out as described before.[1] Samples from pediatric ALL patients older than 1 year (i.e., non-infants) were selected from our cell bank present at the Erasmus MC – Sophia Children's Hospital, Rotterdam, The Netherlands. All samples were freshly processed within 24 hours after sampling as previously described. [3] Briefly, mononuclear cells were isolated by density gradient centrifugation using Lymphoprep (NycomedPharma), and non-leukemic cells were removed using immunomagnetic beads. [4] All leukemia samples used in this study contained more than 90% leukemic cells, as determined morphologically on May-Grünwald-Giemsa (Merck)-stained cytospins.

RNA and DNA extraction

Total RNA and gDNA were extracted from a minimum of $5x10^6$ leukemic cells using TRIzol reagent (Invitrogen Life Technologies) according to the manufacturer's instructions with minor modifications. Quantification of DNA was performed using a spectrophotometer. Quantification and assessment of integrity of the extracted RNA was assessed using the Agilent 2100 Bio-analyzer (Agilent).

Gene expression profiling

The gene expression profiling (Affymetrix platform) data used in the present study has previously been published, [1] and has been deposited in the National Center for Biotechnology Information Gene Expression Omnibus[5] and is accessible via GEO Series accession number GSE19475. Additional unpublished gene expression data can be found under GEO accession number GSE58565. Raw array data were collectively normalized using variance-stabilizing normalization [6] and additionally corrected for batch effects using ComBat[7]. Removal of batch effects was visually verified by unsupervised principle component analysis. Differential gene expression was statistically evaluated using linear models for microarray analyses. [8, 9]

Prednisone response

In vivo prednisone responses were determined after seven days of prednisone monotherapy (including a single intrathecal dose of methotrexate), prior to the initiation of combination chemotherapy. Patients were defined as good responders when <1000 leukemic blasts/µL were detectable in the peripheral blood. Patients still burdened with ≥1000 leukemic blasts/µL after prednisone monotherapy were defined as poor responders.

Oligo array-CGH

Oligo array-CGH (array competitive genomic hybridization) analysis was performed for the infant cohort using the human genome CGH Microarray 105k-A (Agilent Technologies, Palo Alto, CA) according to the manufacturer's protocol using a dye-swap experimental design to minimize false positive results, as previously described.[10] For the ALL-10 cohort Agilent SurePrint G3 180k arrays (Agilent Technologies, Palo Alto, CA) were used and no dye swap was done. Raw microarray image files were processed with Feature Extraction (Agilent Technologies, Santa Clara, CA). Results were analyzed using Agilent Genomic Workbench version 6.5 (Agilent Technologies, Santa Clara, CA). Numerical aberrations were defined as loss or gain of a full chromosome. Structural aberrations were defined as any gain or loss of three or more subsequent probes with a minimum average absolute log ratio of 0.3 excluding known copy-number variations regions using track Hs_hg18_cnv_20090312.

Multiplex ligation-dependent probe amplification (MLPA) analysis and JAK2 mutations

MLPA was performed using the SALSA MLPA P335 ALL-IKZF1 probemix kit (MRC-Holland, Amsterdam, The Netherlands) according to the manufacturer's protocol. Mutations in JAK2 exon 16 were analyzed by polymerase chain reaction amplification and subsequent sequencing as described previously.[11]

TCF3-PBX1 PCR

TCF3-PBX1 was determined positive upon detection of a band after PCR using the following primers: forward: 5'-CACCAGCCTCATGCACAA-3', reverse: 5'-TCGCAGGAGATTCATCACG-3'.

Statistical analyses

Disease-free survival (DFS) was calculated from date of first remission to the date of event which included relapse, death in complete remission, or second malignancy, whichever occurred first. Overall survival (OS) was calculated from the date of diagnosis to the date of death from any cause. Observations of patients were censored at the date of last contact when no events were observed. Follow-up was on December 31th 2009 for the Interfant-99 cohort, and on December 31th 2013 for the Interfant-06 cohort, with a median (interquartile range) follow-up of 7·0 years (5·1 – 8·0) and 4·6 years (3·8 – 5·5), respectively. The Kaplan-Meier method was used to estimate the probabilities of DFS and OS, with standard errors (SE) calculated according to Greenwood. Curves were compared using the log-rank test. Cumulative incidence of relapse (CIR) were estimated adjusting for competing risks of death and second malignancy were statistically analyzed by the Gray test. We used the Fisher exact test to assess the association between patients' characteristics and cohorts. All tests were two-sided. Survival analyses were performed using SAS 9.2 (SAS institute, Cary, NC, USA) and R 3.1.0 (R Core Team, R Foundation for Statistical Computing, Vienna, Austria) at the Interfant Trial Center. The significance analysis of microarray (SAM) [12] was used to identify probe sets significantly associated with EFS, score was calculated with 200 permutations.

References

1. Stam RW, Schneider P, Hagelstein JA, et al. Gene expression profiling-based dissection of MLL translocated and MLL germline acute lymphoblastic leukemia in infants. Blood. 2010 Apr 8;115(14):2835-2844.

2. Den Boer ML, van Slegtenhorst M, De Menezes RX, et al. A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide classification study. The lancet oncology. 2009 Feb;10(2):125-134.

3. Stam RW, den Boer ML, Schneider P, et al. Targeting FLT3 in primary MLL-gene-rearranged infant acute lymphoblastic leukemia. Blood. 2005 Oct 1;106(7):2484-2490.

4. Kaspers GJ, Veerman AJ, Pieters R, et al. Mononuclear cells contaminating acute lymphoblastic leukaemic samples tested for cellular drug resistance using the methyl-thiazol-tetrazolium assay. Br J Cancer. 1994 Dec;70(6):1047-1052.

5. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002 Jan 1;30(1):207-10. PubMed PMID: 11752295. Pubmed Central PMCID: 99122. Epub 2001/12/26. eng.

6. Huber W, von Heydebreck A, Sultmann H, et al. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics. 2002;18 Suppl 1:S96-104.

7. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007 Jan;8(1):118-127.

8. Wettenhall JM, Smyth GK. limmaGUI: a graphical user interface for linear modeling of microarray data. Bioinformatics. 2004 Dec 12;20(18):3705-3706.

9. Smyth GK. Limma: linear models for microarray data. Bioinformatics and Computational Biology Solutions Using R and Bioconductor. ed. Gentleman R, Dudoit S, Irizarry R, Huber W, editors. New York, NY: Springer; 2005. 24 p.

10. Van Vlierberghe P, van Grotel M, Beverloo HB, et al. The cryptic chromosomal deletion del(11)(p12p13) as a new activation mechanism of LMO2 in pediatric T-cell acute lymphoblastic leukemia. Blood. 2006 Nov 15;108(10):3520-3529.

11. Blink M, Buitenkamp TD, van den Heuvel-Eibrink MM, et al. Frequency and prognostic implications of JAK 1-3 aberrations in Down syndrome acute lymphoblastic and myeloid leukemia. Leukemia. 2011 Aug;25(8):1365-1368.

12. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A. 2001 Apr 24;98(9):5116-5121.

Supplemental table 1. Distribution of prognostic factors in low versus high *MEIS1* expressing wild-type infant ALL patients

	MEIS1 < median n=18	MEIS1 > median n=18	P*
Age at diagnosis			0.711
< 6 months	4 (22%)	6 (33%)	
> 6 months	14 (78%)	12 (67%)	
Unknown	0	0	
WBC count (cells/L)			0.168
< $100x10^9$	9 (60%)	9 (56%)	
$100\text{-}300x10^9$	2 (13%)	6 (38%)	
> $300x10^9$	4 (27%)	1 (6%)	
Unknown	3	2	
Prednisone response			0.596
good response	12 (92%)	11 (79%)	
poor response	1 (8%)	3 (21%)	
Unknown	5	4	
Immunophenotype			0.009
pro-B cell	0	3 (20%)	
common B cell	2 (15%)	8 (53%)	
pre-B cell	9 (69%)	1 (7%)	
T-lineage	2 (15%)	3 (20%)	
Unkonwn	5	3	

All data are number (%). *P-value comparing the distribution of poor prognostic factors between low-level expression of *MEIS1* (n=18) and high-level expression of *MEIS1* (n=18) in wild-type *MLL* infant ALL using the Fisher's Exact test on patients with known data.

Supplemental Table 2. Summary of aberrations found by array-CGH of 31 wild-type *MLL* infant ALL patients

Case (n)	Sex	Age (months)	Immuno-phenotype	Chromosome	Cytoband	Start	Stop	+/-
1	female	11	unknown	NO ABERRATIONS				
2	male	5	pre-B	7	p22·3 - p15·3	149068	20814521	-
				9	p24·3 - p13·2	193993	36847230	-
				9	p21·3	21827673	21998367	--
3	male	5	pre-B	9	p21·3 - p21·2	21482343	27254039	-
				16	p13·3	70150	4404795	+
				16	p13·3 - p13·2	5622326	8521959	-
4	female	6	T	NO ABERRATIONS				
5	male	8	pro-B	9	q31·1 - q34·3	105109369	140193874	+
				17	p13·3 - p13·1	28969	8074153	-
6	female	8	pre-B	14	complete chromosome			+
				22	complete chromosome			+
7	female	7	pre-B	NO ABERRATIONS				
8	male	6	unknown	1	complete chromosome			-
				2	complete chromosome			-
				4	complete chromosome			-
				9	complete chromosome			-
				13	complete chromosome			-
				15	complete chromosome			-
				16	complete chromosome			-
				18	complete chromosome			+
				19	complete chromosome			-
				20	complete chromosome			-
				X	complete chromosome			+
				Y	complete chromosome			-
9	female	7	pre-B	9	p24·3 - p21·2	152931	26455189	--
				9	p24·3 - p24·1	345879	6206019	-
				9	p24·1	6234464	6781084	--
				9	p24·1 - p23	6967740	9962217	-
				9	p23 - p22·3	10013642	14387386	--
				9	p22·3	14496495	16425025	-
				9	p22·3 - p22·2	16437991	18063498	--
				9	p22·2 - p21·3	18490144	21537536	-
				9	p21·3	21573783	22755422	--
				9	p21·3 - p21·2	22889384	26264718	-
				9	p21·2 - p11·1	27498421	47002387	-
				9	q12 - q21·11	67701166	72430420	-
				9	q21·13 - q21·2	78718532	80054660	-
				9	q22·1 - q22·32	90543552	98124268	-
				9	q33·1	119125575	120550685	-

				20	q11·21 - q13·33	31170057	62363774	-
10	female	6	T	1	p33	47475493	47540696	-
				10	q23·31	89615444	89666862	-
11	female	11	pre-B	9	p24·3 - p13·2	193993	37317972	-
				13	q14·11 - q34	40064678	114124062	-
12	female	9	unknown	NO ABERRATIONS				
13	male	7	common	4	complete chromosome			+
				6	complete chromosome			+
				10	complete chromosome			+
				14	complete chromosome			+
				17	complete chromosome			+
				18	complete chromosome			+
				21	complete chromosome			+
				X	complete chromosome			+
14	male	10	pre-B	NO ABERRATIONS				
15	male	0	unknown	1	p36·11	24235782	25193094	-
16	female	10	unknown	4	complete chromosome			+
				6	complete chromosome			+
				14	complete chromosome			+
				17	complete chromosome			+
				18	complete chromosome			+
				21	complete chromosome			+
17	male	2	pre-B	14	complete chromosome			+
				22	complete chromosome			+
18	female	5	unknown	22	q11·22	20730547	21443935	-
19	female	7	common	3	q25·2 - q26·33	156242699	180759981	-
20	male	10	Other	1	p35·3 - p35·2	29103090	31160825	-
				16	p11·2	30849299	33517567	-
				17	p13·1	9934363	10343321	+
				18	q21·32	55088482	55882821	+
21	female	10	pro-B	NO ABERRATIONS				
22	male	10	common	6	complete chromosome			+
				14	complete chromosome			+
				17	complete chromosome			+
				18	complete chromosome			+
				21	complete chromosome			+
23	female	9	common	9	p24·3 - p13·2	229226	36930463	-
24	male	10	T	11	q14·1 - q22·1	84284829	99548539	-

25	female	11	common	NO ABERRATIONS				
26	male	5	unknown	8	complete chromosome			+
				19	complete chromosome			+
27	female	9	common	NO ABERRATIONS				
28	female	5	common	1	complete q-arm			+
29	male	9	pre-B	NO ABERRATIONS				
30	female	11	common	9	p24·1 - p21·1	6631559	32777373	-
				9	p21·3	20375131	22638651	--
				9	p21·3	21399600	21490892	-
31	male	4	pro-B	NO ABERRATIONS				

Alterations in copy-number variation regions were omitted. + = amplification of one allele, - = loss of one allele heterozygosity, -- = loss of both alleles. Start and stop positions according to human genome build 18.

Supplemental Table 3. Multiplex ligation-dependent probe amplification (MLPA) analysis of genes associated with B-cell differentiation.

	wild-type *MLL* infant ALL n=32	pediatric non-infant ALL n=232[#]	Fisher P-value
Deletions			
CDKN2A	6 (19%)	81 (35%)	0.07
CDKN2B	6 (19%)	78 (34%)	0.11
PAX5	6 (19%)	55 (24%)	0.7
ETV6	1 (3%)	61 (26%)	0.002
BTG1	1 (3%)	25 (11%)	0.3

[#]Pediatric non-infant B-ALL cohort consisting of 11 (5%) *TCF3-PBX1* translocated patient, 57 (25%) patients with a hyperdiploid karyotype, 5 (2%) *MLL*-rearranged patients, 61 (26%) 3 (1%) *BCR-ABL1* translocated patients, *ETV6-RUNX1* translocated patients, 53 (23%) B-other patients and 42 (18%) T-ALL patients.