European Journal of Haematology 83 (406–419)

ORIGINAL ARTICLE

Two independent gene signatures in pediatric t(4;11) acute lymphoblastic leukemia patients

Luca Trentin[1,*], Marco Giordan[1,*], Theo Dingermann[2], Giuseppe Basso[1], Geertruy te Kronnie[1], Rolf Marschalek[2]

[1]Hemato-Oncology, Dept. of Pediatrics, University of Padova, Padova, Italy; [2]Institute of Pharmaceutical Biology/ZAFES/DCAL/CEF, JWG-University Frankfurt, Frankfurt/Main, Germany

Abstract

Objective: Gene expression profiles become increasingly more important for diagnostic procedures, allowing clinical predictions including treatment response and outcome. However, the establishment of specific and robust gene signatures from microarray data sets requires the analysis of large numbers of patients and the application of complex biostatistical algorithms. Especially in case of rare diseases and due to these constrains, diagnostic centers with limited access to patients or bioinformatic resources are excluded from implementing these new technologies. *Method:* In our study we sought to overcome these limitations and for proof of principle, we analyzed the rare t(4;11) leukemia disease entity. First, gene expression data of each t(4;11) leukemia patient were normalized by pairwise subtraction against normal bone marrow (n = 3) to identify significantly deregulated gene sets for each patient. *Result:* A 'core signature' of 186 commonly deregulated genes present in each investigated t(4;11) leukemia patient was defined. Linking the obtained gene sets to four biological discriminators (*HOXA* gene expression, age at diagnosis, fusion gene transcripts and chromosomal breakpoints) divided patients into two distinct subgroups: the first one comprised infant patients with low *HOXA* genes expression and the *MLL* breakpoints within introns 11/12. The second one comprised non-infant patients with high *HOXA* expression and *MLL* breakpoints within introns 9/10. *Conclusion:* A yet homogeneous leukemia entity was further subdivided, based on distinct genetic properties. This approach provided a simplified way to obtain robust and disease-specific gene signatures even in smaller cohorts.

Key words acute leukemia; gene expression profiles; t(4;11); *MLL* gene; AF4 gene

Correspondence Prof. Dr. Rolf Marschalek, Institute of Pharmaceutical Biology/ZAFES, University of Frankfurt, Marie-Curie Str. 9, 60439 Frankfurt/Main, Germany. Tel: +49 69 798 29647; Fax: +49 69 798 29662; e-mail: Rolf.Marschalek@em.uni-frankfurt.de

*These authors contributed equally to this work.

Accepted for publication 20 June 2009

doi:10.1111/j.1600-0609.2009.01305.x

A milestone in the understanding of the molecular characteristics of leukemias was the use of gene expression profiling (GEP) by microarray and the subsequent identification of specific signatures in different entities of acute leukemia (1–7). The approach of these studies allowed us to identify gene subsets discriminating between different leukemia entities and these gene signatures can be used for diagnostic purposes as well as for predicting clinical parameters such as therapy response (8).

Beside their importance for clinical decision-making established patient gene signatures do not necessarily have any importance for underlying disease processes. This controversial situation is due to the fact that microarray gene expression analyses studies are (i) comparing various subclasses of leukemia to find genes that best distinguish between the considered groups and (ii) aiming to identify specific combinations of genes that guarantee reliability for diagnosis, rather than aiming to identify biologically relevant genes. Another problem of comparing leukemia subclasses by relative statistical analyses is the large number of patient samples needed to identify gene signatures with clinical significance. In case of

© 2009 John Wiley & Sons A/S

MLL-rearranged leukemia patients, relatively uniform signatures have been identified in different studies (1–7, 9), although analyzing patients harboring different *MLL* fusion genes. Thus, the yet applied approaches tend to define a minimum set of genes that can be used to discriminate *MLL*-rearranged leukemia patients from other leukemia entities.

Leukemia patients bearing *MLL* translocations, however, are quite heterogeneous and display acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML) disease phenotypes. Moreover the large number of different *MLL* translocation partner genes (10) ignores important variability and may create difficulties in the interpretation of results. Therefore, we decided to focus only on a single subgroup of pediatric leukemia patients that carry t(4;11) translocations. Based on our assumptions, we expected highly stable signatures that could be further investigated by linking patients to clinical or experimental data. By using this simple approach, a t(4;11)-specific core signature was identified. Surprisingly, the analyzed t(4;11) patients separated into two distinct subgroups when linked to specific discriminators derived from experimental data. Previous studies aiming to identify uniform signatures for MLL rearranged ALL patients have presumably filtered away these novel signatures. Thus, this new approach overcomes the concept of class discovery and introduces the concept of heterogeneity within a single class of cancer patients.

Material and methods

Patient data

Twenty patients diagnosed with B cell precursor (BCP) ALL between January 2005 and September 2008 were included in this study. All samples were t(4;11)(q21;q23) positive as revealed by molecular screening. Main patients' characteristics are summarized in Table 1. All but one patient (P14) included in this study were part of a larger cohort of samples analyzed by GEP during the international 'Microarray Innovation in LEukemia' (MILE) study (11). All patients were randomly selected from a larger cohort of BCP ALL patients. The only further criteria was the availability of nucleic acids for further analyses. Three normal bone marrow (BM1–BM3) samples from pediatric patients were randomly chosen and used as unselected control but included female and male donors to avoid any gender-specific difference. Only residual material from diagnostic procedures was used. Nucleic acids were isolated from bone marrow and/or

Table 1 Patient-specific parameters

		No. significant chip ID'S				Discriminators used in the study				
Patient	Gender	Comp vs. BM1	Comp vs. BM2	Comp vs. BM3	CDTG	HOXA genes	Infant	Non-infant	BRX Introns 11/12	Recipr. FuTx
P1	m	6356	6224	6388	3278	X	X			
P2	m	6505	5978	6267	3259		X			
P3	f	6647	6204	6415	3524		X		X	n.d.
P4	f	5776	4784	6224	2837		X		X	X
P5	f	7109	6347	7411	3681		X		X	X
P6	f	6062	4414	6743	2856	X		X	X	
P7	m	4156	3277	4980	2001	X		X		X
P8	f	3991	3361	4857	1970	X		X		X
P9	m	4357	3410	4926	1943	X		X		
P10	f	4243	3840	4606	2037	X		X	X	X
P11	f	4300	3720	4938	2127	X		X		X
P12	m	6178	5210	6786	3139	X		X		X
P13	m	4923	4710	5285	2584	X		X		X
P14	m	7392	5687	8257	3442		X		X	X
P15	m	4843	3986	5473	2342		X			X
P16	f	6315	5976	6540	3170	X	X			X
P17	f	6498	5830	7212	3390	X		X		X
P18	m	6237	5391	6908	3132	X	X			
P19	m	7723	6435	8408	4124		X			
P20	m	6291	5811	6616	3294		X		X	n.d,
Mean	11:9	5795	5029	6262	2906	60%	55%	45%	35%	66%

Patients' UPN and gender are shown on the left. Next consecutive threes lanes: the number of identified target genes after comparison to three bone marrow samples of unrelated healthy individuals; consecutive lane 4: the number of identified CDTG's for each patient. Consecutive lanes 5–9: discriminators used in this study: presence of *HOXA* gene transcripts, infant vs. non-infant, localization of the chromosomal breakpoint within the *MLL* gene and presence of reciprocal AF4-MLL fusion transcript.

© 2009 John Wiley & Sons A/S

peripheral blood samples. Informed consent was obtained from the parents or the legal guardians of both the patients and control individuals.

RNA preparation, cDNA synthesis and RT-PCR experiments

Total RNA was extracted from bone marrow mononuclear cells using TRIzol RNA isolation (Invitrogen, Karlsruhe, Germany) followed by RNA purification on RNeasy columns (RNeasy Mini Kit, Qiagen, Hilden, Germany). The RNA quality was assessed on the Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany) and RNA concentration was determined using the NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE, USA). One microgram of RNA was reverse transcribed with 200 Units M-MLV reverse transcriptase (Invitrogen) and a mixture of random hexamers (2.5 μM) and oligo dT primers (20 nM) at 37°C for 30 min and 42°C for 15 min. RT-PCR experiments were performed using standard conditions and oligonucleotides specific for transcripts derived from the *MLL·AF4* (MLL8·3 × AF4·5) and *AF4·MLL* fusion genes (AF4·3 × MLL13·5). Oligonucleotide sequences were: MLL8·3 (5′-CCCAAAACCACTCCTAGTGAG-3), MLL13·5 (5′-CAGGGTGATAGCTGTTTCGG-3), AF4·3 (5′-GTTGCAATGCAGCAGAAGCC-3), AF4·5 (5′-ACTGTCACTGTCCTCACTGTCA-3).

RNA gene expression array analyses

Gene expression analysis was performed using the Affymetrix HG-U133 Plus 2.0 oligonucleotide microarrays. From each sample 2.0 μg of purified RNA were converted by reverse transcription into double-stranded cDNA (Roche Applied Science, Mannheim, Germany) and then purified using the GeneChip Sample Cleanup module (Affymetrix, Santa Clara, CA, USA). Then, labeled cRNA was generated using the Microarray RNA target synthesis kit (Roche Applied Science) and an *in vitro* transcription labeling nucleotide mixture (Affymetrix). The cRNA was then purified using the GeneChip Sample Cleanup module (Affymetrix) and quantified using the NanoDrop spectrophotometer. For the following fragmentation we used 11 μg of labeled cRNA. Hybridization, washing, staining and scanning protocols were performed on Affymetrix GeneChip instruments (Hybridization Oven 640, Fluidics Station 450Dx, Scanner GCS3000Dx respectively), following the manufacturer's instructions.

Data analysis

After scanning, absolute and comparison analyses were carried out using the Affymetrix GeneChip Analysis Suite 5.0 software. To determine differentially expressed genes, comparison files were further filtered using the Affymetrix Data Mining Tool 3.0 software. Filter criteria for robustly up- and down-regulated genes included change P-value < 0.0001 and change P-value > 0.9999. Resulting CEL-files (P1-P20) were then compared to CEL-files obtained from normal bone marrow (BM1–BM3) of pediatric volunteers. The 3 resulting output files (Table 1, lanes 1–3) were then compared to each other and only genes found to be present in all three output files were exported into a single file (Table 1, lane 4: commonly deregulated target genes; CDTG1-20). All 20 CDTG files were imported into a relational database program (FileMaker Pro 9) for further analyses. All database entries – deriving from the imported CDTGs – generated a data space of 10 692 gene entries. Each patient was then assigned to the discriminators listed in Table 1. Gene entries present in every patients were automatically assigned by '1' in a sorting field named 'A', whereas gene entries not present in every patient were automatically classified by '0' in the same sorting field; a sorting routine asking for gene entries assigned with number '1' in sorting field 'A' was then used to identify specific target genes. Similar procedures were applied to the four discriminators, and gene entries present/absent in a selected patient subgroup were assigned with '1'/'0' in the respective sorting fields B, C, D, etc. By combining these sorting routines (e.g. are there gene entries assigned with '1' in both sorting fields B and C, B and D, C and D or B and C and D, etc.), specific associations were identified. The database program, including all t(4;11) patient CDTGs, can be made available upon request to interested researchers. It contains 39 different files, including the main program, 20 patient data files and a gene ontology. In order to use these program files the FileMaker Pro 9 software (either Macintosh or PC) will be required.

Heatmap generating subroutines

In order to find discriminating genes that allow to generate a heatmap, gene entries were selected that are predominantly present in either subgroup ('HOXA high' vs. 'HOXA low') but less present or absent in the other group. 'HOXA high' gene entries were selected by a hit frequency of at least 75% in the data files of the 12 'HOXA high' patients, and present at a maximum of 25% in the 'HOXA low' subgroup. *Vice versa*, 'HOXA low' gene entries were selected against 'HOXA low' patients. If genes were present in both subgroups, a difference in their mean value of gene expression of at least 8-fold (log2 change $\geq$3) was requested. A final routine was searching for genes that were either up-regulated in one subgroup, while down-regulated in the other subgroup or *vice versa*. By

© 2009 John Wiley & Sons A/S

applying these three search algorithms, probe sets were selected and exported into the R-program (http://www.R-project.org). The R-program produced a heatmap based on hierachical cluster analysis using the original CEL-files and the defined probe sets.

Data interpretation

Gene names were used to screen the STRING-database (string.embl.de). Potential functions were retrieved from publications available in PUBMED (http://www.ncbi.nlm.nih.gov).

Results

Molecular analysis of fusion gene transcripts

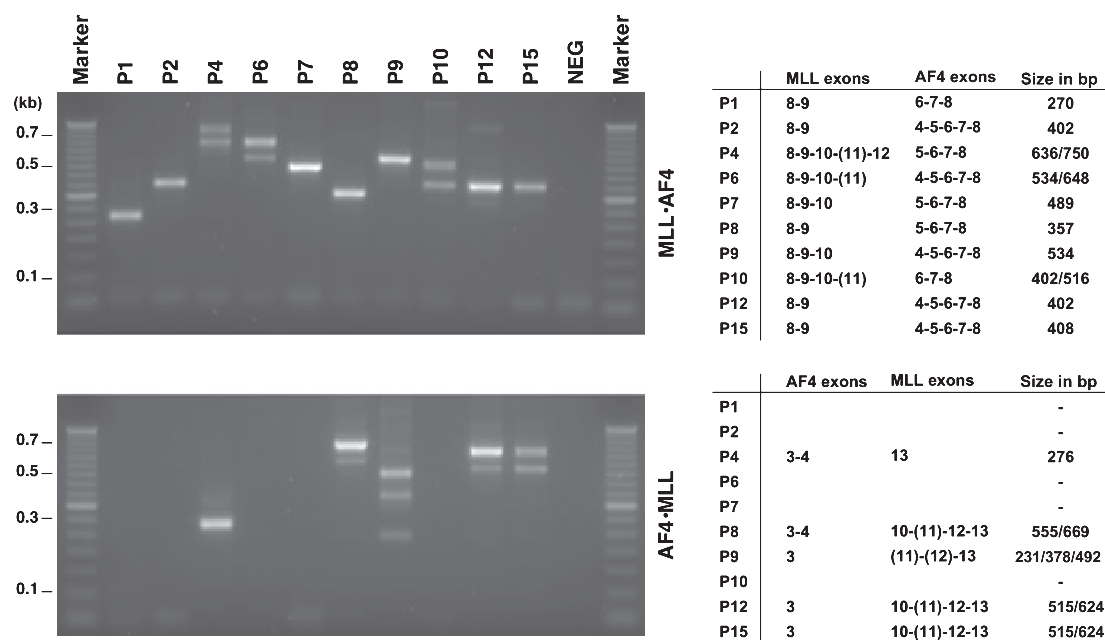
All t(4;11) patients were investigated for their property to transcribe the reciprocal fusion genes *MLL·AF4* and *AF4·MLL*. As exemplarily shown in Fig. 1, the investigated t(4;11)-patients transcribed the *MLL·AF4* fusion gene, whereas the reciprocal *AF4·MLL* fusion allele was expressed only in 11 out of 18 analyzed samples. For two patients, no analysis could be performed due to insufficient amount of material. All PCR amplimers were cut out from the gels and subsequently analyzed by DNA sequencing in order to find the location of the *MLL* breakpoint of each leukemia patient; this information was implemented in our relational database and used as one of the four independent discriminators in the following analyses.

Comparison of t(4;11)-GEPs against 3 bone marrow samples of healthy individuals

Pediatric leukemia patients bearing a t(4;11) translocation and clinically classified as infant (n = 11) and non-infant leukemias (n = 9) were used for investigations. CEL-files of these t(4;11)-patients (P1–P20) were obtained by hybridization experiments using HG-U133 Plus 2.0 microarrays (Affymetrix). Each patient CEL-file was then compared to CEL-files of three unrelated normal bone marrow samples (BM1–BM3) by using the GeneChip Analysis Suite 5.0 program for single comparison analysis. This led to the identification of differentially expressed genes (P-value < 0.0001 or > 0.9999). The resulting three output files of each patient were then analyzed for commonly deregulated target genes (CDTG) – representing only those gene entries that were present in all three output files. (see Table 1).

Analysis of CDTG1 to CDTG20 according to clinical and molecular discriminators

CDTG1 to CDTG20 were imported into the relational database program and assigned to the following discrimi-



	MLL exons	AF4 exons	Size in bp
P1	8-9	6-7-8	270
P2	8-9	4-5-6-7-8	402
P4	8-9-10-(11)-12	5-6-7-8	636/750
P6	8-9-10-(11)	4-5-6-7-8	534/648
P7	8-9-10	5-6-7-8	489
P8	8-9	5-6-7-8	357
P9	8-9-10	4-5-6-7-8	534
P10	8-9-10-(11)	6-7-8	402/516
P12	8-9	4-5-6-7-8	402
P15	8-9	4-5-6-7-8	408

	AF4 exons	MLL exons	Size in bp
P1			-
P2			-
P4	3-4	13	276
P6			-
P7			-
P8	3-4	10-(11)-12-13	555/669
P9	3	(11)-(12)-13	231/378/492
P10			-
P12	3	10-(11)-12-13	515/624
P15	3	10-(11)-12-13	515/624

Figure 1 Transcriptional analysis of investigated t(4;11) patients. RT-PCR analysis of a subset of analyzed patients. Upper left panel: transcripts derived from the *MLL·AF4* allele; lower left panel: transcripts derived from the *AF4·MLL* allele; all PCR amplimers were sequenced to analyze precisely their exon compositions. The exact exon composition of all PCR amplimers was summarized in the displayed tables on the right. Some faint PCR bands (e.g. in lanes P9 and P12) turned out to be PCR artifacts.

© 2009 John Wiley & Sons A/S

nators: (1) *HOXA* gene overexpression (12–16); (2) the clinical onset of disease ('infant' vs. 'non-infant' t(4;11) leukemia patients); (3) the localization of the chromosomal breakpoint within the *MLL* breakpoint cluster region (17); and transcription of *MLL·AF4* alone or of both reciprocal fusion genes (18–20).

Sorting routines were then used to identify gene signatures that are present in 100% of all patients and regulated in the same direction (either up- or down-regulated). This supervised approach revealed 186 gene entries, termed the 'core signature' of t(4;11) patients which are summarized in Table S1. Next we used the four discriminators to identify further subsets of genes related to the 20 t(4;11) patients.

First discriminator: HOXA gene overexpression

When analyzing the CDTG data sets for *HOXA* gene expression, to our surprise, only 12 t(4;11) patients exhibit ectopically up-regulated *HOXA5*, *HOXA9* and *HOXA10* genes (P1, P6–P13, P16–P18), whereas 8 t(4;11) patients (P2–P5, P14, P15, P19, P20) displayed a significant down-regulation of these genes when compared to normal bone marrow signatures. The same finding was also attained (Fig. 2A) using the original CEL-files (P1–P20) normalized with affy package in R (http://www.bioconductor.org). Since *HOXC8* is known to be an MLL target gene (21) we used the normalized CEL-files to analyze for *HOXC8* transcription expression. As shown in Fig. 2A, no significant differences in *HOXC8* gene expression were found between all investigated t(4;11) patients and control bone marrow samples. Two genes known to be transcriptionally activated in *MLL*-mediated leukemia, *MEIS1* and *MEF2C*, were also investigated by using the original and normalized CEL-files (Fig. 2B). In all but one patient (P6), both genes were transcriptionally activated to a similar extent when compared to the normal controls. Thus, the observed differences in *HOXA* gene expression are separating the investigated t(4;11) patients into two distinct subgroups, named 'HOXA high' and 'HOXA low', that are characterized by the differential expression of 102 and 321 genes respectively. A heatmap (Fig. 3) was generated using the 57 most differentially regulated genes between 'HOXA low' and 'HOXA high' patients, thereby confirming the presence of two different patients' subgroups within our initial cohort of samples.

Second discriminator: age at diagnosis

Eleven patients were grouped into the 'infant group' (below 1 year of age) and nine patients were classified into the 'non-infant group'. The two groups of patients are characterized by additional 182 and 196 genes respectively.

Third discriminator: MLL *breakpoint site*

We used the sequence information deriving from *MLL·AF4* fusion transcripts to determine the *MLL* breakpoints in the *MLL* breakpoint cluster region. Patients were either grouped into breakpoints within *MLL* 'introns 9–10' or 'introns 11–12'. Breakpoint distribution also defined additional 98 and 254 gene entries for chromosomal breakpoints localized in *MLL* introns 9–10 and 11–12 respectively.

Fourth discriminator: transcription of the t(4;11) fusion alleles

Based on the results of the RT-PCR experiments, seven patients expressed the *MLL·AF4* fusion allele alone, whereas 11 patients expressed both. Two patients were not investigated for *AF4·MLL* transcripts due to limited material available for RT-PCR analysis. The sorting routine distinguished 260 and 87 additional gene entries that were related to the expression of *MLL-AF4* alone or the presence of both fusion transcripts.

All the target genes related to the four biological discriminating factors are summarized in Tables S2–S9.

Identification of association within the discriminator-specific signatures

A relational database program allows to find associations between discriminator-specific signatures; since all the identified subsets were chosen by a 100% criteria (present in every patient of a given subgroup), the question was raised whether there is any cross-correlation between these signatures. By testing any possible combinations of the sorting routines, two strong correlations were identified.

The first association comprised infant leukemia, displaying low *HOXA* expression and breakpoints localized in *MLL* introns 11 or 12 (Fig. 4A); the second one was identified in non-infant leukemia, displaying high *HOXA* expression and breakpoints within *MLL* introns 9 or 10 (Fig. 4B); all gene entries of the two associations are summarized in Table S10 and S11.

Then, associations between the gene signatures identified in relation to the four discriminators have been tested as shown in Table 2. Each of the six comparisons was analyzed by Fisher's exact test and the resulting *P*-values have been corrected to control family wise error rate (FWER) using Bonferroni's method. Significant adjusted *P*-values were found for the correlation between the 'Infant/Non-infant'/'HOXA low/high' signature (P-value = 3.043^{-60}), the correlation between 'Infant/Non-infant'/'breakpoint within *MLL* introns 11–12/introns 9–10' (P-value = 6.823^{-13}) and the 'HOXA low/high'/'breakpoint within introns 11–12/introns 9–10' signatures (P-value = 1.337^{-33}) respectively. Thus, the

 © 2009 John Wiley & Sons A/S

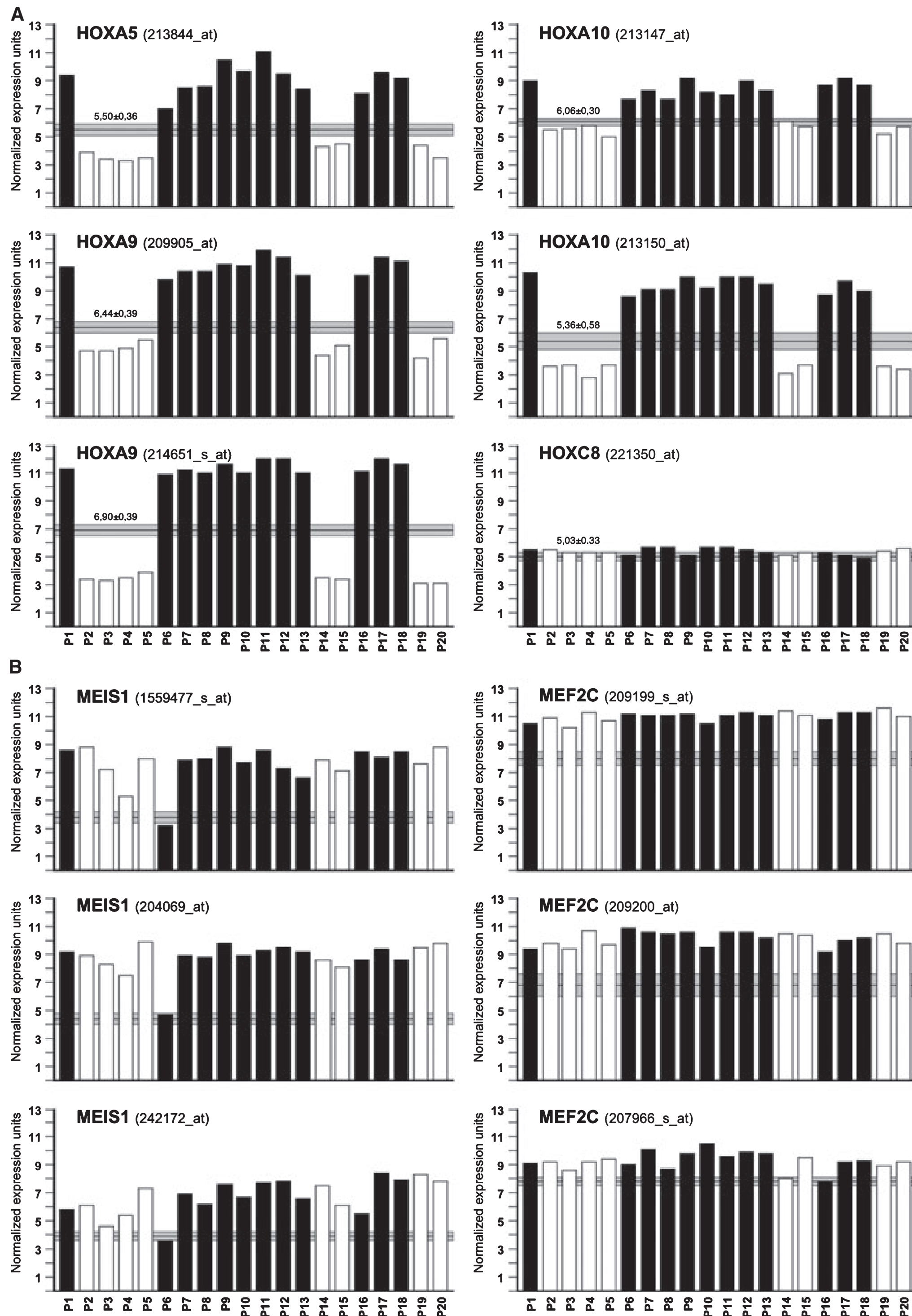


Figure 2 Normalized gene expression data of selected MLL target genes. Original CEL-files of all patients were loaded into the R-program, which automatically normalizes all GEP data. Black bars represent patients with high *HOXA* gene expression; white bars represent patients with low *HOXA* gene expression. (A) Normalized gene expression data for *HOXA5* (213844_at), *HOXA9* (209905_at; 214651_s_at), HOXA10 (213147_at; 213150_at) and *HOXC8* (221350_at) are shown for all investigated patients, expressed by their normalized log2-change. Normalized bone marrow expression data are shown as black horizontal lines with standard deviations (gray bars). (B) Normalized gene expression data for *MEIS1* (1559477_at; 204069_at; 242172_at) and *MEF2C* (209199_s_at; 209200_at; 207966_s_at) are shown for all investigated patients, expressed by their normalized log2-change. Normalized bone marrow expression data are shown as black horizontal lines with standard deviations (gray bars).

© 2009 John Wiley & Sons A/S

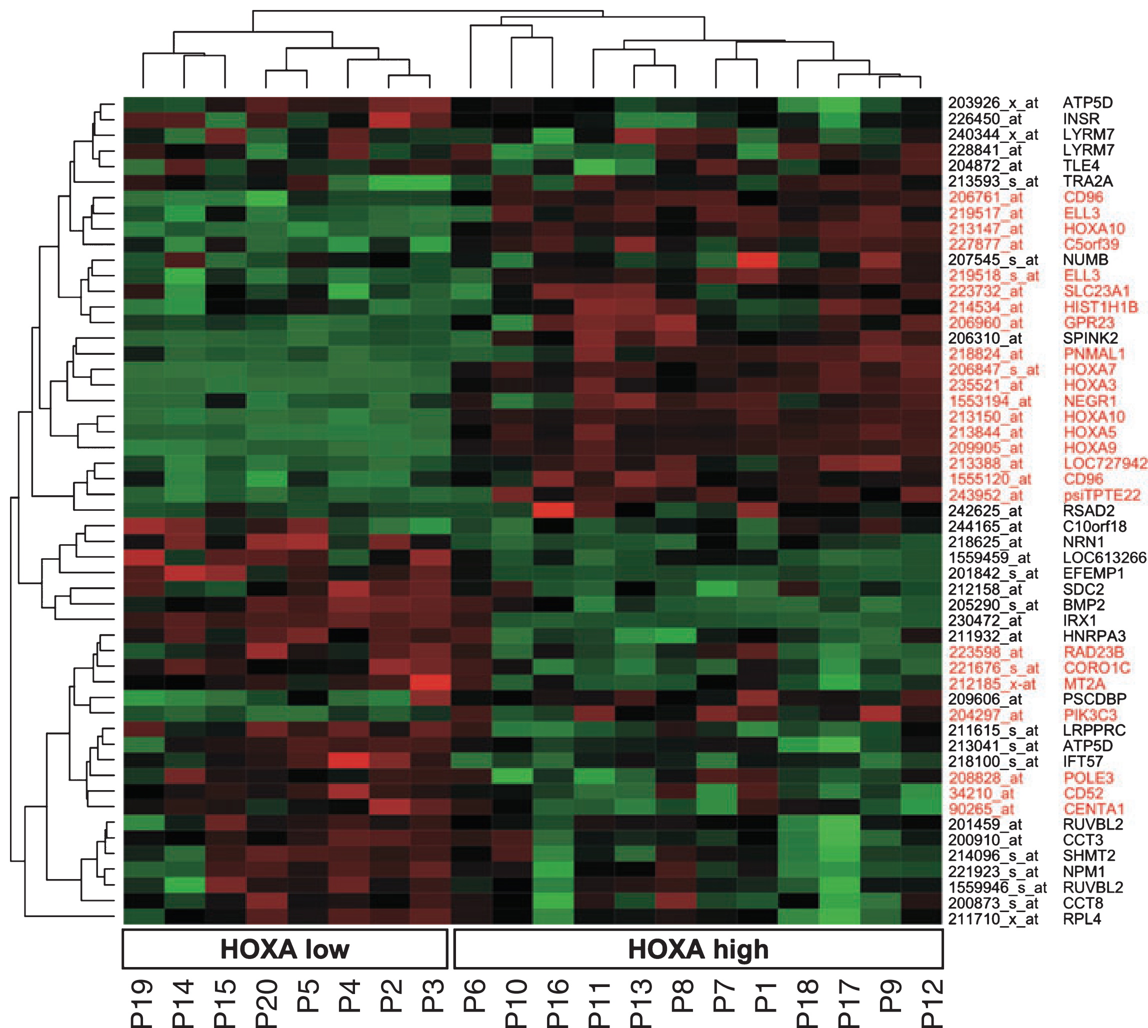


Figure 3 Heatmap of gene entries that distinguishes between both identified t(4;11) leukemia patient subgroups. A heatmap was created by using the open source R-program. Patients cluster according to the chosen gene entries into two independent groups, top: cluster analysis; right: probe sets and corresponding gene names. Below: patients used in this study. Gene names marked in red are idiosyncratic for the 'HOXA high'-signature, while gene names marked in black belong to the 'HOXA low'-signature. Up- and down-regulated genes of both subgroups were used.

defined associations are highly statistical significant and, verify the presence of two independent subgroups within the investigated t(4;11) patients. Finally, the expression of the *MLL* fusion alleles, either *MLL·AF4* alone or both fusion alleles, was significantly correlated to *HOXA* gene expression (*P*-value = 0.0182).

Identified gene signatures and their biological function: t(4;11) core signature

The relational database highlighted the presence of two distinct subgroups with specific associations to clinical and molecular parameters (associations 1 and 2). Nevertheless, these two subgroups shared a common core of 36 up-regulated and 150 down-regulated genes (Supplemental Table S1). Biological processes such as cell proliferation, cellular growth, apoptosis and regulation of transcription could be associated to the up-regulated genes of the core signature. Genes involved in the regulation of cell proliferation and growth are for example: *CTGF*, *CD72*, *BLK*, *IGF2BP3, MAP4* and *SOCS2.*

The *connective tissue growth factor (CTGF)* is a growth factor for connective tissue and interacts with VEGF and TGFß. It is a mediator of local angiogenesis, it has been implicated in osteolytic metastasis by breast cancer cells (22) and CTGF over-expression predicts poor outcome in adult ALL patients (23). CD72 is presumably a prognostic marker of progenitor B-cell leukemias (24) and associates with protein tyrosin phosphatase SHP1 and stimulate the phosphorylation of *B Lymphoid Kinase (BLK)* that is a Src-related kinase and causes proliferation of B progenitor cells and

 © 2009 John Wiley & Sons A/S

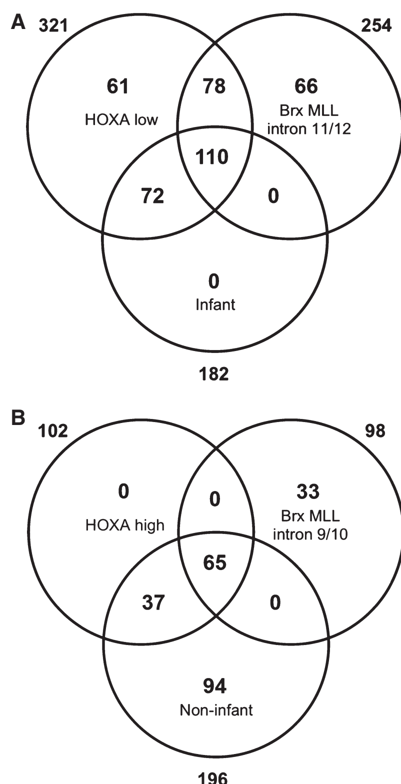


Figure 4 Cross-correlations between distinct subsets identified by the discriminators. (A) Correlations between the subsets 'HOXA low', 'Infant leukemia' and '*MLL* breakpoint within introns 11 and 12'. Numbers represent the amount of gene entries overlapping between different subsets. The discriminators 'Infant' and 'breakpoint distribution' resulted in gene entries that are either completely or partially part of the 'HOXA low' gene signature. (B) Correlations between the subsets 'HOXA high', 'Non-infant leukemia' and '*MLL* breakpoint within introns 9 and 10'. Numbers represent the amount of gene entries overlapping between different subsets. The discriminators 'HOXA high' and 'breakpoint distribution' resulted in gene entries that are either completely or partially part of the 'Non-Infant' gene signature.

enhances responsiveness to Interleukin-7 (25). The *insulin-like growth factor 2 mRNA binding protein 3 (IGF2BP3, IMP-3)* is an oncofetal protein expressed during embryonic development and found to be ectopically expressed in some tumors. When IGF2BP3 is knocked-down, a significant decrease of cell proliferation was observed (26).

Microtubule-associated protein 4 (MAP4) is necessary for increased microtubule polymerization and decreases vinca alkaloid sensitivity (27). MAP4 is a downstream target of p53. MAP4 binds to the heterotrimer SEPT2, SEPT6 and SEPT7. The direct interaction of CDC2/CyclinB with microtubules depends on MAP4, which becomes phosphorylated at Ser696 and Ser787. *Suppressor of cytokine signaling 2 (SOCS2)* are inhibitors of JAK and TYK kinases. Overexpressed SOCS2 is highly effective in blocking signaling from a large variety of different surface receptors, including the prolactin receptor, growth hormone receptor, insulin-like growth factor 1 receptor and the insulin receptor. SOCS2 is also able to block STAT1-, STAT3-, STAT5a- and STAT5b-mediated signaling by interfering with JAK1, JAK3 and TYK2. A variety of other receptors, like EPOR, EGFR, CSFR, IL3R, IL6R, gp130, IL9R, IL10R and LEPR may also be impaired in function by increased SOCS2 expression. Possibly, a block in signaling activity may induce the activation of FOXO proteins able to establish quiescence in t(4;11) positive ALL cells (28). Quiescent cells are resistant to therapy, and therefore, may explain the treatment difficulties and poor prognosis of these patients.

Genes associated with the apoptosis pathway are for example *FAIM* and *SOX4;* the *Fas apoptosis inhibitory molecule (FAIM)* has been identified in Fas-resistant B lymphocytes (29); further *four SRY-box 4 (SOX4)* directly activates TLE3 and PUMA; SOX4 overexpression is associated with resistance against apoptosis leading to growth-transformation (30).

Finally, genes involved in transcription regulation processes are for example *JMJD1C*, *TFEB* and *MEF2C*. *Jumonji domain containing 1C (JMJD1C)* is a histone H3K9 demethylase involved in the removement of repressive histone signatures. *Transcription factor EB (TFEB)* has been identified to be transcriptionally activated in t(6;11)(p21;q13) chromosomal translocations and it has been related to renal cell carcinoma (31). TFEB has DNA-binding and oligomerization properties of a unique helix-loop-helix/leucine-zipper family and binds directly to TFE3 (32). Both TFEB and TFE3 are activated by LIF and regulate the E-Cadherin gene. *Myocyte enhancer factor 2C (MEF2C)* is a B-cell restricted transcription factor in lymphocytes (33). Transactivating activity depends on p38-mediated MAPK phosphorylation. MEF2C binds to other MEF2 proteins, to a variety of different HDACs (including HDAC4, HDAC5, HDAC9), to GATA4, ID3, FOXH1, NKX2-5, HAND1, HAND2, NFATC1, MYOD1, MYF5, MYF6, MYOG and to different nuclear complexes like CARM1 or p300. Moreover, murine Mef2c significantly accelerates myeloid leukemia induced by Sox4 (34).

Among the core down-regulated genes, we found a large number of myeloid-specific genes (i.e. *CD14, CD31, CD59, CD163, CD302, FGR, MAFB, MNDA* and *MPO*), *MXD1, MXL1, MKRN1, hTERT* and *FAS*. MXD1 and MXL1 are negative regulators of the c-MYC protein and this finding may suggest that active c-MYC cannot be counter-regulated appropriately. MKRN1 encodes a protein that negatively regulates hTERT suggesting that hTERT is active in t(4;11) leukemia cells. Finally, down-regulated FAS may prevent extrinsic induction of apoptosis.

© 2009 John Wiley & Sons A/S

Table 2 Comparison of discriminator-specific gene sets by statistical analyses

Discriminator	No. identified genes	HOX high	HOX low	Infant	Non-infant	Brx introns 9/10	Brx introns 11/12	der11 alone	der4/11
HOX high	102	102	0	0	102	65	11	64	9
HOX low	321	0	321	182	20	20	188	109	47
P-values		–		**0.000**		**0.000**		**0.018**	
Infant	182	0	182	182	0	20	110	79	34
Non-Infant	196	102	0	0	196	65	40	94	36
P-values		**0.000**		–		**0.000**		**1.000**	
Brx introns 9/10	98	65	20	20	65	98	0	57	24
Brx introns 11/12	254	11	188	110	40	0	254	119	32
P-values		**0.000**		**0.000**		–		**1.000**	
der11 alone	260	64	109	79	86	57	119	260	0
der4/11	87	9	47	34	28	24	32	0	87
P-values		**0.018**		**1.000**		**1.000**		–	

For each of the four discriminator (HOXA high vs. HOXA low, infant vs. non-infant, breakpoint localization, and presence (absence of reciprocal AF4-MLL fusion transcript), a distinct number of genes were identified. All identified gene sets were then compared to the other identified gene sets by statistical analysis (Fisher's exact test: resulting *P*-values were subsequently corrected to control family wise error rate (FWER) using Bonferroni's method). *P*-values for each of the six comparisons are shown in the center of each subset.
Number of overlapping genes between the discriminator-specific gene sets.

Association 1: infant ALL with low HOXA expression and breakpoints within MLL introns 11/12

This association was identified as common data set of three different signatures and summarized in Table S10. This association is comprised by 110 probe sets representing 21 up-regulated and 72 down-regulated genes. Highest activation of transcription was found for *PPP1R14A* (+13-fold), *CAMK2D* (+12-fold), *KLRK1* (+10-fold), *LCN8* (+10-fold) and *LOC144481* (+9-fold). Strongest suppression of transcription was observed for *PPBP* (−340-fold), *RWDD3* (−138-fold), *SYNE1* (−112-fold), *MME* (−107-fold) and *CD36* (−96-fold).

Up-regulated PPP1R14A and CAMK2D are both involved in IP3 and Ca^{2+} signaling pathways. SMC6 is involved in DNA repair and checkpoint response. Over-expressed BAALC has been identified as adverse risk factor in AML with normal cytogenetics and distinguishes AML patients into a specific subgroup (35). It has been proved to be expressed only in early hematopoietic progenitor-cells able to differentiate into myeloid, lymphoid, and erythroid pathways (36).

Among down-regulated genes we can distinguish again several myeloid markers (i.e. *BPI*, *CD36* and *CEBPE)*, *CEACAM1*, *FAM129A*, *RHOU*, *ANXA1*, *MME* and *PRAM1*. The down-regulation of *CEACAM1*, *FAM129A*, *RHOU* have already been described for solid tumors. ANXA1 is a surface protein and its down-regulation allows transmigration. MME is a negative regulator of focal adhesion kinase signaling and blocks cell migration. Thus, its down-regulation may allow migration of leukemic cells. The PRAM1 protein is involved in B- and T-cell signaling and was found to be down-regulated by the PML·RARa fusion protein.

Many genes (i.e. *LBR*, *MARCKS*, *MBOAT2*, *MCTP2*, *S100A12* and *TYROBP*) coding for proteins involved in different signaling pathways were also down-regulated, and indicated that leukemic cells of this t(4;11) subpopulation are presumably less actively signaling.

Association 2: non-infants with high HOXA expression and breakpoints within MLL introns 9/10

This signature was identified as association between three different signatures and is summarized in Table S11. This association is comprised by 65 probe sets representing 13 up-regulated genes and 39 down-regulated genes. Highest activation of transcription was found for *LUZP1* (+191-fold), *PROM1* (+15-fold), *PRO1073* (+9-fold), *LOC441108* (+8-fold). Strongest suppression of transcription was observed for *LPCAT2* (−38-fold), *RRAGD* (−35-fold), *KCNE3* (−30-fold), *SLC22A4* (−22-fold) and *APP* (−22-fold). LUZP1 is a leucine zipper protein that seems to be strongly over-expressed in this group. ATRX is a chromatin remodeling factor. Dysfunctions of ATRX are associated with myelodysplasia associated with alpha-thalassemia (ATMDS) and somatic mutations of the gene encoding the chromatin remodeling factor ATRX cause an unexpectedly severe hematological phenotype (37). MEF2A associates again with HDAC4 and shows a similar profile as MEF2C, mentioned above. They regulate muscle and adipose tissue during states of insulin deficiency by the regulation of the GLUT4 receptor. MZF1 delays ATRA-mediated apoptosis in myeloid cells (38). Moreover, MZF1 regulates the CD34 promoter and interacts with FHL3 to suppress transcription. MZF1 also induces N-Cadherin expression. Within the

 © 2009 John Wiley & Sons A/S

down-regulated genes, FNDC3B normally exerts anti-neoplastic activity, whereas the kinase SLK promotes apoptosis via the activation of MAPK signaling. Thus, down-regulation of these two genes may support malignant cell growth.

Discussion

Acute Lymphoblastic Leukemia in infants (< 1 year of age) are characterized by a high incidence of translocations involving the *MLL* gene. *MLL* rearranged ALL represents an aggressive and difficult to treat subtype of ALL. By far, the most common *MLL* translocation in ALL patients is the chromosomal translocation t(4;11)(q21;q23), fusing the *MLL* gene with the *AF4* gene (10). As current therapies fail in a large portion of these very young children (> 60%), novel genetic targets – against which innovative therapeutic strategies may presumably be developed – are urgently needed. Therefore, a firm understanding about the genes able to induce and maintain the leukemic phenotype is required. An attractive tool for this task are GEP experiments using available microarray systems. Over the past years, several gene signatures associated with *MLL*-rearranged leukemias have been established (1–7, 9), which have proven to be useful for the clinical classification of this and other leukemias (8). However, the principle strategies used in these studies does not necessarily allow to draw conclusions on the pathology of *MLL*-rearranged leukemia cells, because these studies aimed to establish a uniform gene signature able to discriminate the complex *MLL*-leukemia entity from other leukemia subtypes.

Here, we report on a supervised approach that can be performed to analyze GEPs from leukemia patients with fewer patients and limited access to bioinformatic resources. We used only very stable gene sets (CDTGs) that were obtained by subtracting patient CEL-files against normal bone marrow of unrelated healthy individuals. This step eliminated a pronounced number of genes (~50%) and focuses predominantly on genes that are not abundantly transcribed in normal marrow cells. Moreover, normal bone marrow samples used in this study derived from healthy individuals and were not sorted for any specific hematopoietic compartment. Thus, these pediatric control CEL-files represented all hematopoietic compartments including stem cells, multipotent progenitors, but dominantly premature and mature cells. Thus, the subtraction against normal bone marrow cells presumably enriches for gene sets that represent more immature compartments.

Analyses of the resulting data sets by a relational database program allowed to establish different gene signatures. First, a 'core signature' was identified that was present in every investigated t(4;11) ALL patient. This core signature comprises 186 target genes and overlapped with published data established for *MLL*-rearranged ALL (*HOXA9, DNTT, BLK* and *GUCY1A3*) (2) and data obtained from *in vitro* experiments when both t(4;11) fusion proteins were transfected into murine fibroblasts (*CXCL1, CD302, PLAG1, ITGAM, QPCT, MGST1, CD1* and *THBS1*) (19). Additional overlaps were identified with Polycomb repressor complex II ChIP-on-chip data (*FLT3, MAFB, TACSTD2* and *CD14*) (39), MLL target genes (*RABEF2, FHL1, SOCS2, MEF2C, HOXA9, EBF1, RBMS1, ITGB2, RIOK3, PLEK, PHACTR2, FAS, TNFAIP6, LGALS3, PLAG1, RAB27A, VAMP3, LCN2, SLC40A1, EIF4E3* and *TOB1*) (19), and AF4 target genes (*CSRP2, HOXA9, EBF1, CCL20, IFIT20, CXCL1, SNX10, SAMSN1, CTSS, PLEK, BLVRA, UCGC, TNFAIP6, PLAG1, C3AR1, TNFAIP8* and *PBEF1*; Bursen unpublished data).

Additional signatures were obtained when applying the 4 different discriminators used in this study (see Table 1). The most intriguing finding was the identification of two patients' subgroups characterized by the different *HOXA5*, *HOXA9* and *HOXA10* gene transcription levels (see Fig. 2A). Differences in transcription were in the range of 60- to 100-fold. By contrast, the *HOXC8* gene was similarly expressed in all investigated patients (see Fig. 2A). Moreover, transcriptional levels of genes known to be up-regulated in *MLL*-mediated acute leukemias, *MEIS1* (except patient P6) and *MEF2C*, were transcriptionally overexpressed in all patients (see Fig. 2B). Thus, the applied method seems to be valid and led to results which are in line with earlier findings.

The MEIS1 homeoprotein heterodimerizes with PBX homeoproteins, and in conjunction with HOXA9, is capable of inducing myeloid leukemias in mice (40). In humans, co-expression of MEIS1 and HOXA9 has been shown to be sufficient to immortalize hematopoietic progenitor cells (41). MEIS1 and HOXA9 presumably exert their transforming influence by activating a specific set of target genes. Interestingly, Wang *et al.* demonstrated that a dominant transactivated form of MEIS1 was able to induce leukemic transformation even in the absence of the HOXA9 protein. Moreover, the same study demonstrated that MEIS1 is able to induce a particular set of target genes, most of which were further activated by enforced expression of HOXA9 (42). Among the MEIS1 target genes are FLT3 and SOX4 which were also identified as transcriptional activated target genes within the 'core signature' (see Table S1), suggesting that expression of FLT3 and SOX4 may be a consequence of overexpressed MEIS1. Both FLT3 (a receptor tyrosine kinase) and SOX4 (a HMG-box containing transcription factor) are involved in leukemogenesis. High-level FLT3 expression in *MLL* rearranged leukemias is associated with the constitutive activation of growth-promoting signaling cascades and enhanced survival (43, 44). Mice

© 2009 John Wiley & Sons A/S

receiving Sox4 virus-infected bone marrow cells were shown to develop myeloid leukemia (34).

The separation of t(4;11) leukemia patients into a 'HOXA high' and 'HOXA low' argue for the notion that the transcriptional activation of *MEIS1* is an independent genetic event that could be separated from the ectopic activation of *HOXA* genes. With exception of patient P6, *MEIS1* indeed was highly transcribed in all investigated patients. Based on the presented data, leukemogenic transformation could then further enhanced by two independent genetic programs: transcriptional activation of *HOXA* genes ('HOXA high' patients) or by an alternative pathway that does not require the transcriptional activation of *HOXA* genes ('HOXA low' patients).

This is reflected by the two additional signatures apart from the common 'core signature'. The first signature was found in infant ALL patients that display low *HOXA* transcription signals and breakpoints localized within *MLL* introns 11 and 12 (see Fig. 4A). This patient-specific subset is characterized by 93 gene entries (see Table S10). The second signature was identified in non-infant ALL patients that display high *HOXA* transcription signals and breakpoints localized in *MLL* introns 9 and 10. This patient-specific subset is characterized by 52 gene entries (see Table S11). These signatures have presumably been overseen in the past due to the applied algorithms. Since the first signature has been identified in infants, whereas the other is characteristic for non-infants, it is also plausible that they reflect on a different origin of the malignant cells (e.g. fetal liver vs. a bone marrow). These findings have to be addressed in further studies aiming to dissect the observed results.

In support of this notion, a recently established t(4;11) cell line did not display activated *HOXA* genes (45). In another *in vitro* study, *HOXA* gene expression dropped to control levels when both t(4;11) fusion genes were co-expressed in stably transfected cells in a doxycyclin-dependent manner, although the *MEIS1* gene was strongly activated (19).

A critical view on core- and discriminator-specific signatures (Tables S1–S11) revealed that many identified genes were already classified either as tumor markers in different malignancies (including solid tumors) or were correlated with worse outcome. Beside these tumor markers, genes coding for signaling proteins, surface markers, cytoskeleton proteins and regulatory proteins were identified. These candidate genes can now be tested for their prognostic value or used in further experiments to investigate their particular role in t(4;11) leukemia.

Moreover, to the best of our knowledge, this is the first time that a breakpoint-dependent gene signature (~350 genes) has been identified. Breakpoints in *MLL* intron 11 are frequently associated with infant ALL and therapy-related leukemia (17). In these patients, the coding sequences of *MLL* exons 10 and 11 remain in the *MLL·AF4* fusion transcript. By contrast, juvenile and adult t(4;11) leukemia patients bear their chromosomal breakpoints mostly in *MLL* introns 9 and 10, and thus, amino acid sequences encoded by *MLL* exon 11 or exons 12 make part of the reciprocal *AF4·MLL* fusion transcript. Importantly, *MLL* exons 11–16 encode three PHD fingers. Exons 11 and 12 encode portions of the first and second PHD finger (see Fig. 5). This is in line with recent findings claiming that PHD fingers may have suppressive functions for clonogenic growth by influencing *HOXA* expression (46, 47). Chromosomal breakpoints within *MLL* introns 11 or 12 will disrupt the structure of the first and second PHD finger, and thus, may cause misfolding of the resulting protein sequence due to the high content of cysteine/histidine residues within this protein region. Thus, misfolded PHD domains in the reciprocal AF4·MLL fusion protein may not be able to suppress clonogenic growth. Only breakpoints localized within *MLL* introns 9 and 10 will not disturb the PHD finger domains. Therefore, it is quite plausible that different breakpoints may influence the biological consequences of the resulting t(4;11) fusion

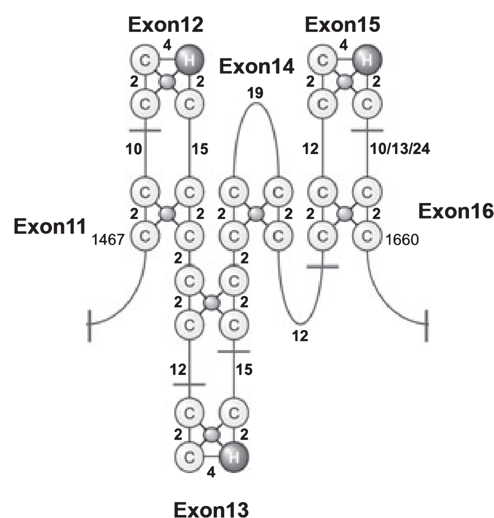


Figure 5 PHD domains of the MLL protein. A potential structure of the PHD finger domain 1–3 is shown. Exon borders are indicated. Bold numbers reflect the distances between the cysteine and histidine residues. The first cysteine of PHD finger 1 is amino acid 1467; the last amino acid of PHD finger 3 is 1660 according to the full-length MLL protein sequence (4.005 amino acids) encoded by 37 *MLL* exons. Diverging numbers between exon 15 and 16 are due to alternative splicing, resulting in 3 or 14 amino acids omitted from the protein sequence. Missing exon 11 or 12 will presumably result in misfolded PHD finger domains.

 © 2009 John Wiley & Sons A/S

proteins, which is presumably reflected by the identified 'breakpoint signatures'.

In conclusion, this study has revealed several new findings that allows to make novel predictions for t(4;11)-mediated ALL. Our approach enabled us to identify a core signature and two independent signatures that define two subgroups within t(4;11) patients. The main discriminator was the up- and down-regulation of transcripts deriving from the *HOXA5, HOXA9, HOXA10* genes. Further analysis, e.g. on the clinical behavior of both patient subgroups will presumably provide novel insights into t(4;11) pathology, and if so, define new molecular targets for further investigations. Finally, linking gene expression data to experimental or clinical informations seems to be a valid method to investigate potential disease pathways and the use of a relational database program could be a valuable tool to analyze gene expression data, when only few patients are available.

Supporting Information

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

Table S1. Identified gene signature of all investigated t(4;11) patients.
Table S2. Identified 'Infant' gene signature.
Table S3. Identified 'Non-infant' gene signature.
Table S4. Identified 'HOXA high' gene signature.
Table S5. Identified 'HOXA low' gene signature.
Table S6. Identified 'Breakpoints within MLL introns 9/10' gene signature.
Table S7. Identified 'Breakpoints within MLL introns 11/12' gene signature.
Table S8. Identified 'der11 fusion transcript only' gene signature.
Table S9. Identified 'both der4/11 fusion transcripts' gene signature.
Table S10. Identified 'association 1' gene signature.
Table S11. Identified ' association 2' gene signature.

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.

Acknowledgements

This study was supported by Grants MA 1876/7-1 and MA 1876/9-1 from the DFG, 01GS0875 from the BMBF, and Grant 102362 from the German Cancer Foundation (Deutsche Krebshilfe) to RM. Additional grant support from Città della Speranza, Padova, Italy to LT and MG and from Ministero dell'Università e della Ricerca, PRIN (Programmi di ricerca di Rilevante Interesse Nazionale), Rome, Italy to GteK. RM is PI within the CEF on Macromolecular Complexes funded by DFG Grant EXC 115.

References

1. Armstrong SA, Staunton JE, Silverman LB, *et al.* MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat Genet* 2002;**30**:41–7.
2. Yeoh EJ, Ross ME, Shurtleff SA, *et al.* Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell* 2002;**1**:133–43.
3. Ferrando AA, Armstrong SA, Neuberg DS, *et al.* Gene expression signatures in MLL-rearranged T-lineage and B-precursor acute leukemias: dominance of HOX dysregulation. *Blood* 2003;**102**:262–8.
4. Ross ME, Zhou X, Song G, *et al.* Classification of pediatric acute lymphoblastic leukemia by gene expression profiling. *Blood* 2003;**102**:2951–9.
5. Rozovskaia T, Ravid-Amir O, Tillib S, *et al.* Expression profiles of acute lymphoblastic and myeloblastic leukemias with ALL-1 rearrangements. *Proc Natl Acad Sci USA* 2003;**100**:7853–8.
6. Haferlach T, Kohlmann A, Schnittger S, *et al.* Global approach to the diagnosis of leukemia using gene expression profiling. *Blood* 2005;**106**:1189–98.
7. Kohlmann A, Schoch C, Dugas M, *et al.* New insights into MLL gene rearranged acute leukemias using gene expression profiling: shared pathways, lineage commitment, and partner genes. *Leukemia* 2005;**19**:953–64.
8. Holleman A, Cheok MH, den Boer ML, *et al.* Gene-expression patterns in drug-resistant acute lymphoblastic leukemia cells and response to treatment. *N Engl J Med* 2004;**351**:533–42.
9. Ayton PM, Cleary ML. Transformation of myeloid progenitors by MLL oncoproteins is dependent on Hoxa7 and Hoxa9. *Genes Dev* 2003;**17**:2298–307.
10. Meyer C, Kowarz E, Hofmann J, *et al.* New insights into the *MLL* recombinome of acute leukemias. *Leukemia* 2009; March 5. [Epub ahead of print].
11. Haferlach T, Kohlmann A, Basso G, *et al.* The clinical utility of microarray-based gene expression profiling in the diagnosis and sub-classification of leukemia: Final Report on 3252 Cases from the International MILE Study Group. *Blood* (ASH Annual Meeting Abstracts). 2008; **112**: Abstract 753.
12. Hess JL. Mechanisms of transformation by MLL. *Crit Rev Eukaryot Gene Expr* 2004;**14**:235–54.
13. Horton SJ, Grier DG, McGonigle GJ, *et al.* Continuous MLL-ENL expression is necessary to establish a "Hox

© 2009 John Wiley & Sons A/S

Code" and maintain immortalization of hematopoietic progenitor cells. *Cancer Res* 2005;**65**:9245–52.
14. Milne TA, Hughes CM, Lloyd R, *et al.* Menin and MLL cooperatively regulate expression of cyclin-dependent kinase inhibitors. *Proc Natl Acad Sci USA* 2005;**102**:749–54.
15. Milne TA, Dou Y, Martin ME, Brock HW, Roeder RG, Hess JL. MLL associates specifically with a subset of transcriptionally active target genes. *Proc Natl Acad Sci USA* 2005;**102**:14765–70.
16. Milne TA, Martin ME, Brock HW, Slany RK, Hess JL. Leukemogenic MLL fusion proteins bind across a broad region of the Hox a9 locus, promoting transcription and multiple histone modifications. *Cancer Res* 2005;**65**:11367–74.
17. Reichel M, Gillert E, Angermüller S, *et al.* Biased distribution of chromosomal breakpoints involving the MLL gene in infants versus children and adults with t(4;11) ALL. *Oncogene* 2001;**20**:2900–7.
18. Bursen A, Moritz S, Gaussmann A, Moritz S, Dingermann T, Marschalek R. Interaction of AF4 wild-type and AF4·MLL fusion protein with SIAH proteins: indication for t(4;11) pathobiology? *Oncogene* 2004;**23**:6237–49.
19. Gaussmann A, Wenger T, Eberle I, *et al.* The combined effects of the two reciprocal t(4;11) fusion proteins, MLL·AF4 and AF4·MLL, confer resistance to apoptosis, cell cycling capacity and growth transformation. *Oncogene* 2006;**26**:3352–6333.
20. Kowarz E, Burmeister T, Lo Nigro L, *et al.* Complex MLL rearrangements in t(4;11) leukemia patients with absent AF4.MLL fusion allele. *Leukemia* 2007;**21**:1232–8.
21. Hanson RD, Hess JL, Yu BD, *et al.* Mammalian Trithorax and polycomb-group homologues are antagonistic regulators of homeotic development. *Proc Natl Acad Sci USA* 1999;**96**:14372–7.
22. Shimo T, Kubota S, Yoshioka N, *et al.* Pathogenic role of connective tissue growth factor (CTGF/CCN2) in osteolytic metastasis of breast cancer. *J Bone Miner Res* 2006;**21**:1045–59.
23. Sala-Torra O, Gundacker HM, Stirewalt DL, *et al.* Connective tissue growth factor (CTGF) expression and outcome in adult patients with acute lymphoblastic leukemia. *Blood* 2007;**109**:3080–3.
24. Schwarting R, Castello R, Moldenhauer G, *et al.* Human Lyb-2 homolog CD72 is a marker for progenitor B-cell leukemias. *Am J Hematol* 1992;**41**:151–8.
25. Tretter T, Ross AE, Dordai DI, Desiderio S. Mimicry of pre-B cell receptor signaling by activation of the tyrosine kinase Blk. *J Exp Med* 2003;**198**:1863–73.
26. Liao B, Hu Y, Herrick DJ, Brewer G. The RNA-binding protein IMP-3 is a translational activator of insulin-like growth factor II leader-3 mRNA during proliferation of human K562 leukemia cells. *J Biol Chem* 2005;**280**:18517–24.
27. Bash-Babula J, Toppmeyer D, Labassi M, *et al.* A Phase I/pilot study of sequential doxorubicin/vinorelbine: effects on p53 and microtubule-associated protein 4. *Clin Cancer Res* 2002;**8**:1057–64.
28. Tothova Z, Kollipara R, Huntly BJ, *et al.* FoxOs are critical mediators of hematopoietic stem cell resistance to physiologic oxidative stress. *Cell* 2007;**128**:325–39.
29. Schneider TJ, Fischer GM, Donohoe TJ, Colarusso TP, Rothstein TL. A novel gene coding for a Fas apoptosis inhibitory molecule (FAIM) isolated from inducibly Fas-resistant B lymphocytes. *J Exp Med* 1999;**189**:949–56.
30. Liu P, Ramachandran S, Ali Seyed M, *et al.* Sex-determining region Y box 4 is a transforming oncogene in human prostate cancer cells. *Cancer Res* 2006;**66**:4011–9.
31. Kuiper RP, Schepens M, Thijssen J, *et al.* Upregulation of the transcription factor TFEB in t(6;11)(p21;q13)-positive renal cell carcinomas due to promoter substitution. *Hum Mol Genet* 2003;**12**:1661–9.
32. Fisher DE, Carr CS, Parent LA, Sharp PA. TFEB has DNA-binding and oligomerization properties of a unique helix-loop-helix/leucine-zipper family. *Genes Dev* 1991;**5**:2342–52.
33. Swanson BJ, Jäck HM, Lyons GE. Characterization of myocyte enhancer factor 2 (MEF2) expression in B and T cells: MEF2C is a B cell-restricted transcription factor in lymphocytes. *Mol Immunol* 1998;**35**:445–58.
34. Du Y, Spence SE, Jenkins NA, Copeland NG. Cooperating cancer-gene identification through oncogenic-retrovirus-induced insertional mutagenesis. *Blood* 2005;**106**:2498–505.
35. Baldus CD, Tanner SM, Ruppert AS, *et al.* BAALC expression predicts clinical outcome of de novo acute myeloid leukemia patients with normal cytogenetics: a cancer and leukemia group B study. *Blood* 2003;**102**:1613–8.
36. Baldus CD, Tanner SM, Kusewitt DF, *et al.* BAALC, a novel marker of human hematopoietic progenitor cells. *Exp Hematol* 2003;**31**:1051–6.
37. Gibbons RJ, Pellagatti A, Garrick D, *et al.* Identification of acquired somatic mutations in the gene encoding chromatin-remodeling factor ATRX in the alpha-thalassemia myelodysplasia syndrome (ATMDS). *Nat Genet* 2003;**34**:446–9.
38. Robertson KA, Hill DP, Kelley MR, *et al.* The myeloid zinc finger gene (MZF-1) delays retinoic acid-induced apoptosis and differentiation in myeloid leukemia cells. *Leukemia* 1998;**12**:690–8.
39. Lee TI, Jenner RG, Boyer LA, *et al.* Control of developmental regulators by polycomb in human embryonic stem cells. *Cell* 2006;**125**:301–13.
40. Kroon E, Krosl J, Thorsteinsdottir U, Baban S, Buchberg AM, Sauvageau G. Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b. *EMBO J* 1998;**17**:3714–25.
41. Zeisig BB, Milne T, Garcia-Cuellar MP, *et al.* Hoxa9 and Meis1 are key targets for MLL-ENL-mediated cellular immortalization. *Mol Cell Biol* 2004;**24**:617–28.

 © 2009 John Wiley & Sons A/S

42. Wang GG, Pasillas MP, Kamps MP. Persistent transactivation by meis1 replaces hox function in myeloid leukemogenesis models: evidence for co-occupancy of meis1-pbx and hox-pbx complexes on promoters of leukemia-associated genes. *Mol Cell Biol* 2006;**26**:3902–16.
43. Armstrong SA, Kung AL, Mabon ME, *et al*. Inhibition of FLT3 in MLL. Validation of a therapeutic target identified by gene expression based classification. *Cancer Cell* 2003;**3**:173–83.
44. Stam RW, den Boer ML, Schneider P, *et al*. Targeting FLT3 in primary MLL-gene-rearranged infant acute lymphoblastic leukemia. *Blood* 2005;**106**:2484–90.
45. Bertrand FE, Spengeman JD, Shah N, LeBien TW. B-cell development in the presence of the MLL/AF4 oncoprotein proceeds in the absence of HOX A7 and HOX A9 expression. *Leukemia* 2003;**17**:2454–9.
46. Chen J, Santillan DA, Koonce M, *et al*. Loss of MLL PHD finger 3 is necessary for MLL-ENL-induced hematopoietic stem cell immortalization. *Cancer Res* 2008;**68**:6199–207.
47. Muntean AG, Giannola D, Udager AM, Hess JL. The PHD fingers of MLL block MLL fusion protein-mediated transformation. *Blood* 2008;**112**:4690–3.

© 2009 John Wiley & Sons A/S