Pediatric T-cell lymphoblastic leukemia evolves into relapse by clonal selection, acquisition of mutations and promoter hypomethylation

Joachim B. Kunz,[1-3*] Tobias Rausch,[2,4,13*] Obul R. Bandapalli,[1-3] Juliane Eilers,[1,2] Paulina Pechanska,[1,2] Stephanie Schuessele,[1,2] Yassen Assenov,[5] Adrian M. Stütz,[2,4] Renate Kirschner-Schwabe,[6] Jana Hof,[6,7] Cornelia Eckert,[6] Arend von Stackelberg,[6] Martin Schrappe,[8] Martin Stanulla,[9] Rolf Koehler,[10] Smadar Avigad,[11] Sarah Elitzur,[11] Rupert Handgretinger,[12] Vladimir Benes,[13] Joachim Weischenfeldt,[4] Jan O. Korbel,[2,4**] Martina U. Muckenthaler,[1,2**] and Andreas E. Kulozik[1-3**]

[1]Department of Pediatric Oncology, Hematology and Immunology, Children's Hospital, University of Heidelberg, Germany; [2]Molecular Medicine Partnership Unit, EMBL-University of Heidelberg, Germany; [3]German Cancer Consortium (DKTK), Heidelberg, Germany; [4]European Molecular Biology Laboratory (EMBL), Genome Biology Unit, Heidelberg, Germany; [5]Division of Epigenomics and Cancer Risk Factors, The German Cancer Research Center (DKFZ), Heidelberg, Germany; [6]Department of Pediatric Oncology/Hematology, Charité - Universitätsmedizin Berlin, Germany; [7]German Cancer Consortium (DKTK) and German Cancer Research Center (DKFZ), Heidelberg, Germany; [8]Pediatrics, University Hospital Schleswig-Holstein, Campus Kiel, Germany; [9]Department of Pediatric Hematology/Oncology, Medical School Hannover, Germany; [10]Department of Human Genetics, University of Heidelberg, Germany; [11]Molecular Oncology, Felsenstein Medical Research Center and Pediatric Hematology Oncology, Schneider Children's Medical Center of Israel, Petah Tikva, Israel; [12]Children's Hospital, University Hospital Tübingen, Germany; and [13]European Molecular Biology Laboratory (EMBL), Genomics Core Facility, Heidelberg, Germany

*JBK and TR contributed equally to this work.

**JOK, MUM and AEK contrubuted equally to this work.

ABSTRACT

Relapsed precursor T-cell acute lymphoblastic leukemia is characterized by resistance against chemotherapy and is frequently fatal. We aimed at understanding the molecular mechanisms resulting in relapse of T-cell acute lymphoblastic leukemia and analyzed 13 patients at first diagnosis, remission and relapse by whole exome sequencing, targeted ultra-deep sequencing, multiplex ligation dependent probe amplification and DNA methylation array. Compared to primary T-cell acute lymphoblastic leukemia, in relapse the number of single nucleotide variants and small insertions and deletions approximately doubled from 11.5 to 26. Targeted ultra-deep sequencing sensitively detected subclones that were selected for in relapse. The mutational pattern defined two types of relapses. While both are characterized by selection of subclones and acquisition of novel mutations, 'type 1' relapse derives from the primary leukemia whereas 'type 2' relapse originates from a common pre-leukemic ancestor. Relapse-specific changes included activation of the nucleotidase NT5C2 resulting in resistance to chemotherapy and mutations of epigenetic modulators, exemplified by *SUZ12*, *WHSC1* and *SMARCA4*. While mutations present in primary leukemia and in relapse were enriched for known drivers of leukemia, relapse-specific changes revealed an association with general cancer-promoting mechanisms. This study thus identifies mechanisms that drive progression of pediatric T-cell acute lymphoblastic leukemia to relapse and may explain the characteristic treatment resistance of this condition.

© Ferrata Storti Foundation

Introduction

T-cell acute lymphoblastic leukemia (T-ALL) is an aggressive malignancy of thymocytes that accounts for about 15% of pediatric acute lymphoblastic leukemias[1] and can be cured in approximately 80% of affected children[2,3] However, relapsed T-ALL is highly resistant to chemotherapy[4,5] and patients who experience a relapse face a dismal prognosis.[5] In B-cell precursor (BCP)-ALL, integrated genomic analysis identified relapse-specific signatures of gene expression, somatic copy number alterations (CNA) and promoter methylation.[6,7] Similar observations in relapsed T-ALL are scarce: gene expression analysis revealed that relapsed T-ALL is characterized by a gene response pattern that is also found in early, but not in late relapse of BCP-ALL.[8] Genome-wide analysis in relapsed T-ALL, in contrast to relapsed BCP-ALL, found similar numbers of CNA as those in matched samples taken during the primary disease.[9] While DNA methylation in T-ALL has been shown to differ from that in BCP-ALL,[10] systematic comparisons of DNA methylation between primary and relapsed T-ALL are not available. Evolution into relapsed T-ALL has been attributed to clonal selection, because relapse-specific rearrangements of the T-cell receptor were backtracked to primary disease[11] and because leukemic cells from primary disease can develop into a clone resembling relapsed disease after transfer into immunodeficient mice.[12] Rarely, late relapse of T-ALL represents a true secondary leukemia, whereby an underlying genetic predisposition could be documented in at least one patient.[13]

To obtain a more comprehensive picture of the evolution of T-ALL from primary disease to relapse we performed an integrated genomic analysis of 13 patients with T-ALL, who were

©2015 Ferrata Storti Foundation. This is an open-access paper. doi:10.3324/haematol.2015.129692
The online version of this article has a Supplementary Appendix.
Manuscript received on April 24, 2015. Manuscript accepted on August 11, 2015.
Correspondence: Martina.Muckenthaler@med.uni-heidelberg.de, Andreas.Kulozik@med.uni-heidelberg.de or jan.korbel@embl-heidelberg.de

analyzed at the time of primary diagnosis, during remission and at relapse.

Methods

Patients' clinical characteristics

Patients were treated according to ALL-BFM 2000 or related frontline protocols (INS-89, INS-98,[14] IC[15]). One patient was aged 18 at diagnosis, all others were children or adolescents. The 13 patients (Table 1) were recruited between 1993 and 2007 from the ALL-REZ BFM 2002 trials (patients T-ALL-H-A61, -E114, -F110, -KI17, -MD40, -T92, -T128) or from Schneider Children's Medical Center of Israel, Petah Tikva, Israel (patients T-ALL-H-S00169, -S00207, -S00285, -S00438, -S00456, -S00472) and selected on the basis of sufficient material being available from the time points of first diagnosis, remission and relapse. Minimal residual disease (MRD) response was assessed as described previously[2,16] (*Online Supplementary Table S3*).

This study was approved by the institutional review boards of the Charité Universitätsmedizin Berlin and the Medical Faculty Heidelberg. Informed consent was obtained in accordance with the Declaration of Helsinki.

Exome capture, target capture and Illumina sequencing

The Agilent SureSelect Target Enrichment Kit (Agilent, Santa Clara, CA, USA; vendor's protocol version 2.0.1) was used to capture all human exons for sequencing. The HaloPlex Target Enrichment Kit (Agilent, Santa Clara, CA, USA; vendor's protocol version D.5, May 2013) was used according to the manufacturer's instructions. The starting material consisted of 225 ng of genomic DNA. The captured fragments were sequenced as 100 bp paired reads using an Illumina HiSeq instrument (Illumina, San Diego, CA, USA).

Analysis of whole exome sequencing and targeted ultra-deep sequencing data

The analysis of sequencing data is detailed in the *Online Supplementary Methods*.

DNA methylation analysis using 450k BeadChip arrays

Genomic DNA (200 ng) was bisulfite-converted using the EZ DNA Methylation Gold Kit (Zymo Research, Irvine, CA, USA). The Infinium methylation assay (Illumina) was carried out as previously described.[17] Data from the 450k Human Methylation Array were normalized by the Beta Mixture Quantile (BMIQ) method[18] using the RnBeads analysis software package.[19]

Multiplex ligation dependent probe amplification

Multiplex ligation dependent probe amplification (MLPA) was done using the MRC Holland (Amsterdam, The Netherlands) SALSA MLPA probe mix P383-A1 T-ALL according to the manufacturer's instructions. Polymerase chain reaction products were separated by capillary electrophoresis on an ABI-3130XL device; the size standard was GeneScan 500-250 (Applied Biosystems). Coffalyser software, available at http://www.mlpa.com, was used for the analyses.

Integrated analysis

In order to evaluate functions that were altered by mutation or DNA methylation at relapse, Ingenuity

© Ferrata Storti Foundation

Table 1. Clinical characteristics of the 13 T-ALL patients analyzed.

Patient ID: T-ALL-H-...	A61	E114	F110	KI17	MD40	T92	T128	S00169	S00207	S00285	S00438	S00456	S00472
Gender	M	M	M	F	M	M	M	M	F	F	M	M	M
Age at diagnosis (years)	11	7	2	5	4	3.5	2	4.4	5.5	9.6	18	12	9.2
Time to relapse (months)	15	29	11	7	29	14	47	6.0	17.1	25.1	16.7	16.6	10.4
White blood cell count at diagnosis (x10^9/L)	256	112	803	2.1	157	83	202	92	83	74	4	465	167
Cell source in relapse	BM	BM	PB	BM	BM	BM	BM	BM	BM	BM	BM	BM	BM
Blast count in relapse (%)	72	98	73	80	59	85	83	unknown	unknown	unknown	unknown	80	83
Immune phenotype*	cortical	mature	pre	mature	cortical	cortical	cortical	pre	mature	mature	cortical	cortical	pre
Remission sample time point (weeks after diagnosis)	12	12	12	12	12	12	12	11	54	83	21	5	24
MRD level in remission	10^{-4}	10^{-4}	neg	10^{-2}	neg	neg	neg	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Risk group in primary disease	MR	HR	HR	HR	MR	MR	HR	HR	HR	MR	MR	HR	HR
Treatment protocol**	ALL-BFM 2000	ALL-BFM 2000	ALL-BFM 2000	ALL-BFM 2000	ALL-BFM 2000	ALL-BFM 2000	ALL-BFM 2000	INS-89	INS-89	INS-98	IC	IC	IC
Outcome	died	died	died	died	CCR after SCT	died in subsequent 4. relapse	CCR after SCT	died	died	died	died	died	died

** Pro/Pre (cyCD3$^+$, CD7$^+$/ cyCD3$^+$, CD2$^+$ and/or CD5$^+$ and/or CD8$^+$), cortical (CD1a$^+$), mature (CD1a, sCD3$^+$). cyCD3$^+$: cytoplasmic CD3$^+$, sCD3$^+$: surface CD3. ** ALL-BFM 2000[2]; INS 89 protocol based on ALL-BFM 86/90, 1989-1998[14,39]; INS 98 protocol based on ALL-BFM 95[40], 1998-2003; and ALL Intercontinental (IC) - BFM 200315, 2003-2005. M. male; F: female; BM: bone marrow; PB: peripheral blood; neg: negative; n.a.: not available; MR: medium risk; HR: high risk; CCR: continuous complete remission; SCT: stem cell transplantation.*

Pathway Analysis (IPA, Version 21249400; Qiagen, Redwood City, CA, USA) was used. For details refer to the *Online Supplementary Methods*.

Results

Relapsed T-cell acute lymphoblastic leukemia acquires single nucleotide variants and small insertions and deletions

We performed whole exome sampling (WES) of DNA samples obtained from 13 patients (see Table 1 for clinical characteristics) at the time of first diagnosis, during remission and at relapse, and identified a total of 340 somatic single nucleotide variants (SNV) and 53 small insertions and deletions (InDels; *Online Supplementary Table S1*). On average, each leukemia specimen carried 11.5 (SD 7.0) somatic SNV and InDels at first diagnosis and 26.0 (SD 26.6) somatic SNV and InDels at the time of relapse – suggesting ongoing evolution of the major clone or of minor subclones. Only 85 of 126 SNV and 14 of 23 InDels present at the time of first diagnosis could also be detected in the corresponding relapse sample, indicating that the (treatment-related) eradication of leukemic clones contributes to the evolution of relapse. We defined a mutation as "relapse-specific" if it was detected by WES or MLPA in a relapse sample, but not in the corresponding primary leukemia sample from the same patient. This could include mutations that were detected in the primary leukemia of other patients.

Comparing all mutations in primary disease to those in relapse, no significant difference in the types of single nucleotide exchanges and in the ratio of transversions and transitions was identified (*Online Supplementary Figure S1A*). However, mutations at the central position of the trinucleotides TCA and GCT were significantly more frequent among relapse-specific SNV than in primary leukemia (*Online Supplementary Figure S1B,C*), consistent with a contribution of cytarabine to the acquisition of mutations during treatment.[20]

Somatic copy number alterations can represent subclonal, "late" events

We used MLPA in order to identify CNA in genes commonly altered in T-ALL. On average, 3.5 CNA were detected in each primary leukemia sample and 3.2 in each relapse sample, whereas 2.7 CNA were detected in both samples from the same patient (*Online Supplementary Table S2*). Most commonly, we found deletions of the *CDKN2A/B* locus[21] on chromosome 9q (all patients), followed by microdeletions within the *LEF1* gene[22] (6/13 patients), amplification of the *MYB* gene[23,24] (4/13 patients), deletions of the *PTEN* gene[25,26] (3/13 patients) and homozygous deletions of the *PTPN2* gene[27] (2/13 patients). Thirty-five of the 45 CNA that were identified in primary disease were preserved in relapse, a proportion comparable to the corresponding numbers of SNV and InDels. In four patients, CNA found to be lost in relapse affected the *CDKN2A*, *LEF1* or *PTEN* genes, indicating that the deletion of these tumor suppressors can be a late event during leukemogenesis. Although several CNA (deletions of *SUZ12*, *LEF1*, *PTEN*, *CASP8AP2*; amplification of *MYB*) were specific for relapse in individual patients, none of these was specific for relapse in more than one patient or significantly enriched in either primary disease or relapse.

Targeted ultra-deep sequencing detects rare subclones

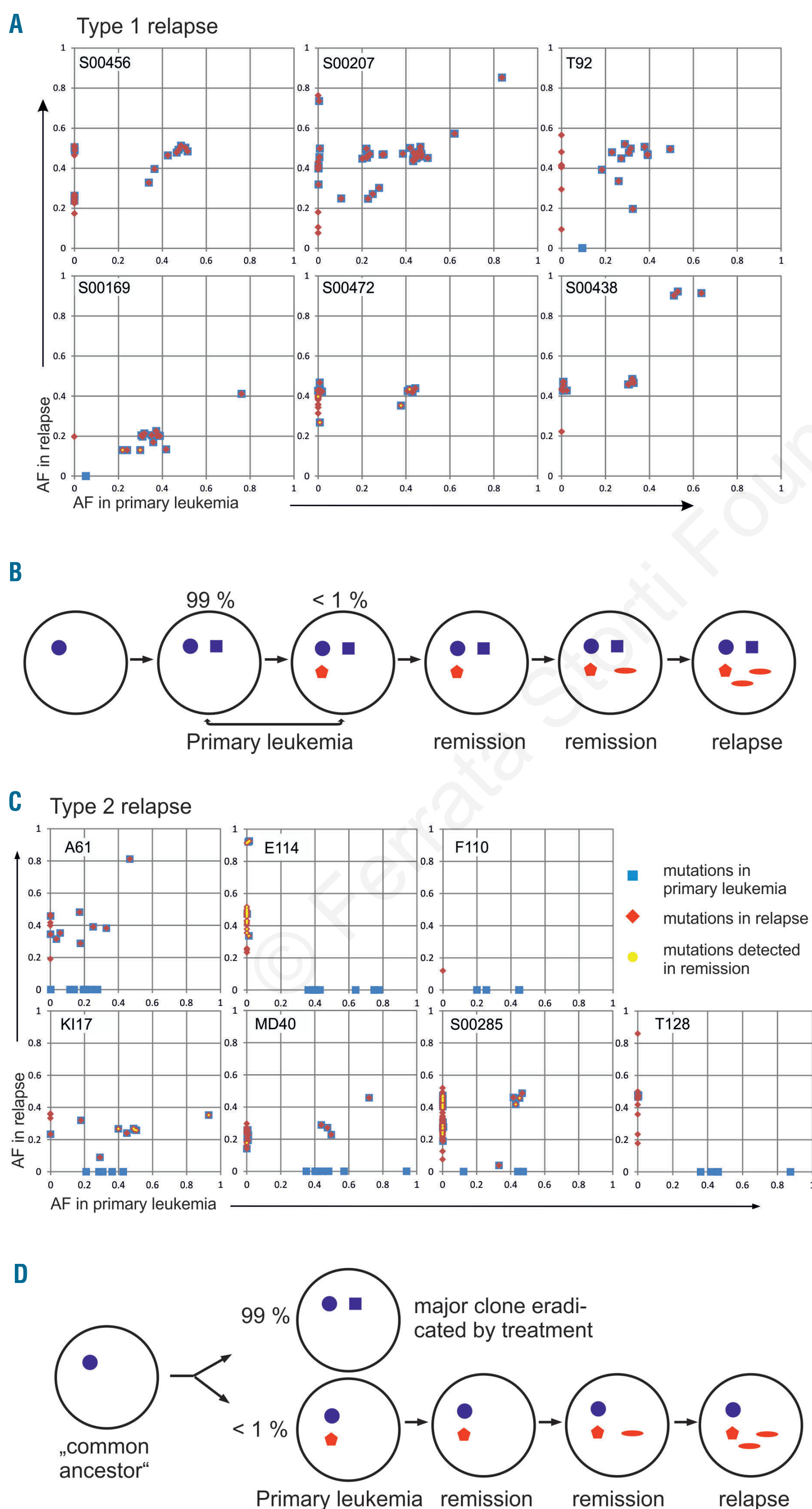
In order to track the clonal evolution of the relapses, we performed targeted ultra-deep sequencing of all mutations that had previously been identified by WES using the HaloPlex target capture (Agilent). Allele frequencies inferred by HaloPlex were highly reproducible (*Online Supplementary Figure S2A*) and corresponded well to allele frequencies inferred by WES (*Online Supplementary Figure S2B*). We analyzed whether HaloPlex reflected the frequencies of rare alleles by measuring the signals in serial dilutions. This analysis indicated that the allele frequency is highly correlated with the dilution (r=0.91). With increasing dilution and decreasing allele frequencies the accuracy of quantification of allele frequencies dropped because mutant reads were overrepresented (*Online Supplementary Figure S2C*). Rare alleles were identified by comparing allele frequencies in the sample of interest with a reference pool consisting of healthy control DNA and unrelated leukemia samples. This technique sensitively detected alleles with a frequency in the range of 0.01 and below: 91% of leukemia-specific SNV were detected after diluting leukemia DNA in control DNA at a ratio of 10^{-2}, 26% were detected in a dilution of 10^{-3} and 3% in a dilution of 10^{-4}. The rate of SNV that were detected as false positives in samples from patients who never carried the respective SNV was 1.8% (*Online Supplementary Figure S2D*), indicating that the detection of a significant proportion of mutant reads by HaloPlex is a reliable qualitative indicator of a small subclone carrying the mutation of interest. The greatest sensitivity of detection was achieved for insertions of more than one nucleotide. For example, in patient's S00169 remission sample two reads carrying a specific 9 bp insertion in the *NOTCH1* gene among 43,567 total reads covering this region indicated persistent MRD in the order of 10^{-4} (*Online Supplementary Table S6*).

In ten of 13 patients we found at least one SNV that had developed from a rare (<0.05) subclonal allele in primary disease to a major clonal allele in relapse (*Online Supplementary Table S5*). Of the 211 SNV and 30 InDels which were found in the major clone exclusively in relapse by WES, 45 and nine, respectively, were detected at a low allele frequency already in primary disease (*Online Supplementary Table S7*). The average allele frequency for such relapse-specific mutations found to be subclonal in primary leukemia was 0.006 (range, 0.0002-0.036). Fifty-five SNV and InDels present in the major clone at primary disease had been lost during the transition to relapse. Fifty-four of these 55 mutations were not detected even in minor clones at the time of relapse, indicating that these clones were eradicated fully or at least to a level below the sensitivity of the HaloPlex analysis.

In ten of 13 of the first disease/relapse pairs a DNA sample obtained at the time of remission (for 7 patients, MRD level from the same sample was available and $\leq 10^{-2}$, Table 1) was also available for ultra-deep sequencing. In these samples we searched for the presence of SNV that had been detected in the same patient at the time of either primary disease or relapse. In six patients (T-ALL-H-E114, -MD40, -KI17, -S00169, -S00285, -S00472) we detected a total of 52 (range, 1 – 21) SNV or InDels, which were significantly more abundant in the remission samples than in the reference consisting of the average of all samples from healthy controls and from patients who never carried the respective SNV of interest (*Online Supplementary Table S6*). Allele frequencies in remission samples ranged between

©Ferrata Storti Foundation

0.00005 and 0.0095. By comparison with a dilution series of leukemia samples in healthy DNA we estimate that these allele frequencies correspond to a frequency of leukemic cells of 0.0001 to 0.01 (1:10,000 – 1:100), consistent with the MRD levels that were available (Table 1). All of the mutations detected in remission samples were also found in the corresponding relapse sample with an allele frequency in the range of 0.13 to 0.91. Thirteen mutations that were detectable in remission were already found in a major clone (allele frequency 0.22-0.93) at primary disease, and an additional ten SNV were detected in primary disease samples at low allele frequencies (range, 0.0003 –



© Ferrata Storti Foundation

Figure 1. Ultra-deep sequence analysis of primary and relapsed T-ALL distinguishes two types of relapse. (A, C) Variant allele frequencies in relapse were plotted over the variant allele frequency in the corresponding primary leukemia sample. Allele frequencies were determined by HaloPlex sequencing, only in the rare case that a certain allele was not covered by HaloPlex, allele frequencies from WES were used. (B, D) The simplified models show each mutation as a single symbol (circle, square, pentagon, ellipse). Blue symbols denote mutations that were detected in the major clone of primary leukemia, red symbols denote mutations that were specific for relapse. (A, B) Type 1 relapse: all mutations present in the major clone from primary disease were also present in relapse. The clone giving rise to relapse carried all mutations that were detected in the major clone of primary leukemia. (C, D) Type 2 relapse: the major clone from primary leukemia was lost in relapse, as indicated by the mutations that were present in the major clone in primary leukemia but absent in relapse. The clone giving rise to relapse shared some but not all mutations with the major clone of primary leukemia and was derived from a common ancestor, but evolved independently already before the initial diagnosis.

0.0120). Twenty-six SNV that were detected in remission were not detected in the corresponding primary disease sample with a sensitivity of 0.01 or higher. We propose that these 26 SNV may have originated from a mutational event during treatment, although clonal selection starting from a very small subclone present at initial diagnosis cannot be ruled out.

In all 13 patients a minimal set of common genetic changes (at least one concordant MRD marker, *Online Supplementary Table S3*, or at least six concordant SNV and InDels, Figure 1) were found in both the primary leukemia and relapse samples, which indicated that all relapses were *bona fide* recurrences of the leukemia and not a second, unrelated neoplasm (as has previously been described in a small proportion of relapsed patients based on the identification of discordant MRD markers[13]). By analyzing allele frequency plots, we can distinguish two types of relapse: type 1 and type 2. Type 1 relapse, observed in six of 13 patients (Figure 1A,B), contained all mutations that were already detectable at the time of primary leukemia. This type of relapse developed either from a major subclone or from a smaller subclone that had acquired additional mutations late in the process of leukemogenesis. In type 2 relapse, observed in the remaining seven patients, mutations that had been present in the major clone in primary leukemia were lost at relapse (Figure 1C,D). Here, relapse developed from an ancestral pre-leukemic clone that had already diverged into distinguishable subclones at an early time point prior to the initial diagnosis. In both types of relapse, clonal selection and acquisition of novel mutations contributed to the mutational load. Type 1 showed a trend to be more frequent in early relapses (time to relapse <24 months; P=0.07, see *Online Supplementary Figure S3* for a logistic regression model) and in Israeli and Palestinian patients (P=0.029, Table 2).

The example of mutations in the nucleotidase NT5C2 illustrates the genetic plasticity of T-ALL. *NT5C2* mutations were identified in five of 13 relapse samples (R367Q in patients A61, S00207, S00285, T92; D407Y in patient T92; P414S in S00456; *Online Supplementary Tables S1* and *S5*). R367Q has been shown to activate the nucleotidase activity of NT5C2 and to confer resistance against nucleoside analogs.[28,29] In two patients, an *NT5C2* mutation was already detected in primary disease samples at low allele frequency (A61: D407Y, allele frequency 0.3%; S00456: P414S, allele frequency 0.1%). While patient S00456 carried the same mutation in the corresponding relapse sample, patient A61 lost the D407Y mutation and acquired the R367Q mutation. *NT5C2* mutations were clonal in three relapses, but subclonal in two other relapse samples (S00207: allele frequency 0.1; T92: allele frequencies 0.41 for R367Q and 0.09 for D407Y). This is compatible with the notion that acquisition of resistance to chemotherapy by *NT5C2* activation can be a late, not-initiating event on the way to relapse.

Patient E114 demonstrated that the evolution of the relapse-specific clone from a pre-leukemic ancestor may be facilitated by intensive induction treatment of the primary leukemia. In this patient, two preserved MRD markers confirmed the relationship between primary leukemia and relapse. In addition, targeted ultra-deep sequencing identified five mutations that had been present at a subclonal level in primary disease, persisted in remission and became predominant in relapse (Figure 1C, *Online Supplementary Tables S6* and *S7*). Already in the remission sample, which was taken immediately after induction treatment had been completed, 16 newly acquired mutations were detected and later predominated at relapse (*Online Supplementary Table S6*). During the 2 years of remission before relapse occurred, only five more mutations accumulated. Interestingly, although MLPA identified deletion of *CDKN2A/B* and amplification of *MYB* in the primary disease sample of this patient (*Online Supplementary Table S2*), none of them was conserved between primary disease and relapse, indicating that in this specific leukemia these copy number variants were probably not the initiating event in leukemogenesis.

Relapse-specific alterations in T-acute lymphoblastic leukemia do not show association with leukemogenesis, but with cancerogenesis in general

In nine of 13 patients we identified relapse-specific mutations that, based on current knowledge, are likely to contribute to the evolution of relapse (Table 3). Besides known mechanisms of leukemogenesis in T-ALL such as NOTCH-activation or *MYB*-amplification, the most common alterations were mutations activating the nucleotidase *NT5C2* and mutations in epigenetic modifiers, such as *WHSC1*, *SUZ12*, *SMARCA4*, *ARID4B* and *USP7*.

In order to obtain an unsupervised view of the contribution of mutations to relapse we investigated whether the genetic alterations that are specific for relapse can be linked to certain biological functions. To this end, we grouped genetic alterations, which had been detected either by WES or by MLPA, according to the time points at which they were found. Using Ingenuity Pathway Analysis (IPA) software, genes that were mutated or deleted in the major clones at both times, primary leukemia

Table 2. Type of relapse is associated with treatment and with time to relapse.

	Type of relapse		
	type 1	type 2	*P*
Gender			
male	5	5	
female	1	2	1
Age (years)			
<9	3	5	
>9	3	2	0.59
Time to relapse (months)			
<24	6	3	
>24	0	4	0.07
WBC at diagnosis (x10^9/L)			
<100	4	2	
>100	2	5	0.29
Risk group in primary disease			
medium risk	2	3	
high risk	4	4	1.0
Protocol			
ALL BFM 2000	1	6	
INS89, INS98, IC	5	1	0.029
Outcome			
dead	6	5	
alive	0	2	0.46

Patients having suffered from either type 1 or type 2 relapse were categorized according to clinical characteristics. P: Fisher exact test.

and relapse, were compared to those genes that were found to be mutated or deleted in the major clone at relapse but not in the corresponding primary leukemia (*Online Supplementary Table S8*).

Using genes altered both in primary leukemia and in relapse, IPA constructed a dense network involving the nodes *NOTCH*, *IL7R*, *MTOR*, *GATA3* and *AKT*, reflecting the frequent occurrence of somatic DNA alterations in established leukemia drivers in this gene set (*Online Supplementary Figure S4A*; Table 4). In contrast, the genes that were mutated in a relapse-specific way showed similarities to gene lists typically carrying somatic mutations in gastrointestinal and other cancers, but not leukemia (*Online Supplementary Figure S4B*; Table 4). These findings suggest that relapsed T-ALL share a profile of mutations with non-hematologic cancers. Among genes that were mutated in a relapse-specific manner, we did not observe a strong enrichment of canonical pathways (minimal P=0.003 for actin cytoskeleton signaling).

Hypermutation caused by somatic DNA repair deficiency can contribute to genetic instability in relapsed T-cell acute lymphoblastic leukemia

Patient S00285 carried a moderate number of eight mutations in primary disease, but an extraordinarily high number of 106 mutations in relapse, many of which were subclonal (*Online Supplementary Tables S1 and S5*). One of the 102 relapse-specific mutations affected the *BLM* gene. While this mutation was not present at the time of primary disease, it was the most abundant newly acquired mutation in remission and present in the main clone at relapse. *BLM* codes for a RecQ DNA helicase and is required for DNA replication and DNA repair. Inactivating germline mutations in *BLM* cause Bloom syndrome, a recessively inherited tumor predisposition syndrome (OMIM #210900). The K39N mutation observed in this patient is predicted to be damaging by PolyPhen[30] and likely caused a somatic DNA-repair defect, resulting in the accumulation of a large number of mostly subclonal mutations during treatment.

Hypomethylated promoters in relapsed T-cell acute lymphoblastic leukemia do not show association with leukemogenesis, but with cancerogenesis in general

We used Illumina 450k arrays to compare DNA methylation in relapse samples to that in the corresponding primary T-ALL samples. In contrast to the situation in relapsed BCP-ALL,[6] the average degree of methylation in relapse did not differ from average methylation in primary leukemia (Figure 2A, r=0.996). When primary disease and relapse samples from individual patients were compared, both hypo- and hypermethylation of promoters in relapse could be observed (Figure 2B).

In order to identify promoters that may undergo differential methylation in relapse, we filtered for promoters that: (i) were represented on the 450k array by at least three different probes; (ii) had a gene symbol assigned; and (iii) had a decrease or an increase of the β-value of at least 0.2 in absolute numbers in at least three different patients. According to these criteria, a total of 239 promoters were recurrently hypermethylated and 579 promoters recurrently hypomethylated in relapse. The lists of hyper- and hypomethylated promoters (*Online Supplementary Table S8*) were subjected to IPA. The association of hypermethylated promoters with diseases and functions reached marginal significance for one functional category only ($P=6.62^{-5}$ for "development of neurons"). In contrast, hypomethylated promoters, similar to the mutations specifically found in relapse, showed highly significant associations with different types of cancer, namely

©Ferrata Storti Foundation

Table 3. Mutational mechanisms likely contributing to relapse.

Patient	Gene	Alteration	Evidence for subclonality in primary leukemia	Mechanism
A61	*NT5C2*	NM_012229:exon15:c.1100G>A:p.R367Q	no	Chemotherapy resistance[28,29]
E114	*FBXW7*	NM_018315:exon8:c.1147A>G:p.T383A	no	NOTCH1 activation[41,42]
	NOTCH1	NM_017617:exon26:c.4775T>C:p.F1592S	no	NOTCH1 activation[43]
	WT1	NM_001198552:exon6:c.423_424insT:p.R141fs	no	WT1 inactivation[44]
	SUZ12	Heterozygous deletion chr 17q11.2	no	PRC2 inactivation[45]
T92	*NT5C2*	NM_012229:exon17:c.1219G>T:p.D407Y	no	Chemotherapy resistance[28,29]
S00169	*WHSC1*	NM_133330:exon20:c.3295G>A:p.E1099K	no	Histone methylation[35,36]
S00207	*SMARCA4*	NM_001128847:exon19:c.2896C>T:p.R966W	yes	Chromatin remodeling[46]
S00285	*BLM*	NM_000057:exon3:c.117G>T:p.K39N	no	DNA repair defect[47]
	USP7	NM_003470:exon15:c.1663C>T:p.R555W	no	DNA methylation[48]
	ARID4B	Frameshift by deletion NM_001206794:Exon7:c.420_423delAATA:p.N141Efs*32	no	Chromatin remodeling[49]
	NT5C2	NM_012229:exon15:c.1100G>A:p.R367Q	no	Chemotherapy resistance[28,29]
S00438	*JAK3*	NM_000215:exon11:c.1533G>A:p.M511I	yes	*JAK3* activation[50]
	NOTCH1	NM_017617:exon26:c.4811_4820delinsTCCCTCACGCTTGAGG: p.V1605Pfs*4	no	NOTCH1 activation[43]
S00456	*NT5C2*	NM_012229:exon16:c.1240G>A:p.P414S	yes	Chemotherapy resistance[28,29]
S00472	*NOTCH1*	NM_017617:exon34:c.7524_7534delCTTCCTCACCC:p.F2509Vfs*	no	NOTCH1 activation[434343]
	MYB	Amplification chr 6q23.3	no	Oncogenic transcription factor

For each patient, relapse-specific mutations or CNA are listed that, according to published evidence, are likely to contribute to leukemogenesis or to emergence of relapse. No such mutations were observed in patients KI17, MD40 and T128.

Table 4. Ingenuity Pathway Analysis.

Genes mutated in major clone of primary leukemia and of relapse (n=145)		Genes mutated in relapse only (n=182)		Recurrently hypomethylated promoters (n=579)	
Disease or function	P	Disease or function	P	Disease or function	P
T-cell leukemia	2.26E-18	adenocarcinoma	4.25E-14	melanoma	6.89E-15
Precursor T-cell lymphoblastic leukemia-lymphoma	5.56E-18	colorectal carcinoma	1.16E-11	endometrium tumor	1.92E-09
Acute lymphocytic leukemia	9.77E-18	gastrointestinal adenocarcinoma	1.29E-11	morphology of nervous system	2.29E-09
Melanoma	4.10E-15	colon carcinoma	1.39E-11	endometrial cancer	3.04E-09
Acute leukemia	1.23E-13	gastrointestinal carcinoma	2.02E-11	endometrioid carcinoma	4.20E-09
Lymphoproliferative disorder	2.30E-13	colon cancer	3.31E-11	development of central nervous system	4.54E-09
Leukemia	2.55E-13	malignant solid tumor	4.25E-11	adenocarcinoma in endometrium	5.18E-09
Hematologic cancer	5.06E-13	melanoma	1.22E-10	development of body axis	1.05E-08
Bone marrow cancer	7.87E-13	gastrointestinal neoplasia	1.27E-10	uterine cancer	1.42E-08
Lymphocytic leukemia	8.54E-13	gastrointestinal tract cancer	1.39E-10	synaptic transmission	1.67E-08

Genes that were found to be mutated or deleted by WES or MLPA both in primary disease and in relapse (n=145), genes that were found to be mutated or deleted by WES or MLPA only in relapse (n=182) and promoters that were found to be hypomethylated in at least three patients (n=579) were subjected to Ingenuity Pathway Analysis for associated functions and diseases (see Online Supplementary Table S8*). The ten most significant associations are reported.*

©Ferrata Storti Foundation

melanoma (P=6.89x10^{-15}) and endometrial carcinoma (P=1.92x10^{-9}; Table 4). We did not observe a strong enrichment of canonical pathways (minimal P=0.003 for glutamate receptor signaling) among promoters that were hypomethyled in a relapse-specific manner. Although the lack of RNA samples precludes an analysis of a correlation between methylation and gene expression data, our results show that DNA methylation changes may cooperate with mutational events in driving T-ALL relapse.

Discussion

By applying a deep coverage target enrichment technique to sensitively and quantitatively detect rare mutations in primary disease, remission and relapse of pediatric T-ALL, we distinguish between relapses arising from the major clone of the primary leukemia (type 1) and relapses arising from a pre-leukemic ancestral clone (type 2). In both types, selection of subclones and acquisition of novel mutations contributed to clonal evolution. Similar observations have been made before for BCP-leukemia[6,7,9] and for acute myeloid leukemia.[31] The size of the subclones that later gave rise to relapse can be estimated to range between 1:100 and 1:1,000 of the primary leukemia in our series. The association of type 1 relapse with a short duration of remission may intuitively be explained by a relative resistance of the primary leukemia to treatment. The more frequent occurrence of type 1 relapses in the Israeli/Palestinian group of patients is unlikely to be related to differences of treatment, because the protocols used in Israel and in Germany are very similar and give comparable overall results.[2,14]

More than three-quarters (187 of 241) of relapse-specific mutations could not be detected in primary disease samples despite a sensitivity of detection that exceeded 1:100 for most mutations. The mutational load during relapse is more than doubled compared to that during primary leukemia, suggesting that mutations are truly acquired *de novo* during treatment and/or during remission. Relapse-specific patterns of mutations attributed to chemotherapy have been described by others.[20,31] Comparing relapse-specific mutations to mutations in primary leukemia, we found a significant increase in single nucleotide exchanges at the central position of the trinucleotide sequence TCA, a mutational signature that has been attributed to the mutagenic effects of cytarabine.[20] However, the overall mutational pattern changed only slightly in relapse, indicating that the contribution of chemotherapy to the acquisition of mutations may be minor in this cohort of patients.

The leukemia that acquired a *BLM* mutation during treatment was of particular interest. This mutation first appeared at the time of remission and was clonal at the time of relapse. The presence of this mutation was associated with an unusually high number of mostly subclonal relapse-specific mutations, indicating that acquired mutations of DNA repair genes may result in somatic hypermutation during the clonal evolution of T-ALL. While haploinsufficiency is not evident in heterozygous carriers of germline *BLM* mutations,[32,33] the effect of somatic *BLM* mutations in the context of leukemia may be amplified by mutagenic stress induced by chemotherapy.

While the genetic alterations shared by primary leukemia and relapse are, as expected, enriched in genes that are known to be implicated in leukemogenesis,[3,26,27,34] the genes that we specifically found to be mutated in T-ALL relapse are not enriched for leukemia-specific genes. Relapse-specific genes are more closely related to other cancer types, implying that relapsed T-ALL, in addition to the genetic repertoire required to induce leukemia in T cells, have acquired properties that are linked to cancer in a less specific sense. Potentially, these genes contain a previously unrecognized category of genes that are specific for T-ALL relapses. These observations suggest that the type of malignancy is determined by the early genetic changes of the normal hematopoietic precursor and that the progression may be driven by more general abnormalities of carcinogenesis. These progression-related changes may also explain the clinical observation that relapsed T-ALL is frequently refractory to leukemia-directed chemotherapy.

An analysis of the 450k methylome data revealed that, in contrast to BCP-ALL,[6] overall DNA methylation levels in T-ALL hardly differ between primary leukemia and

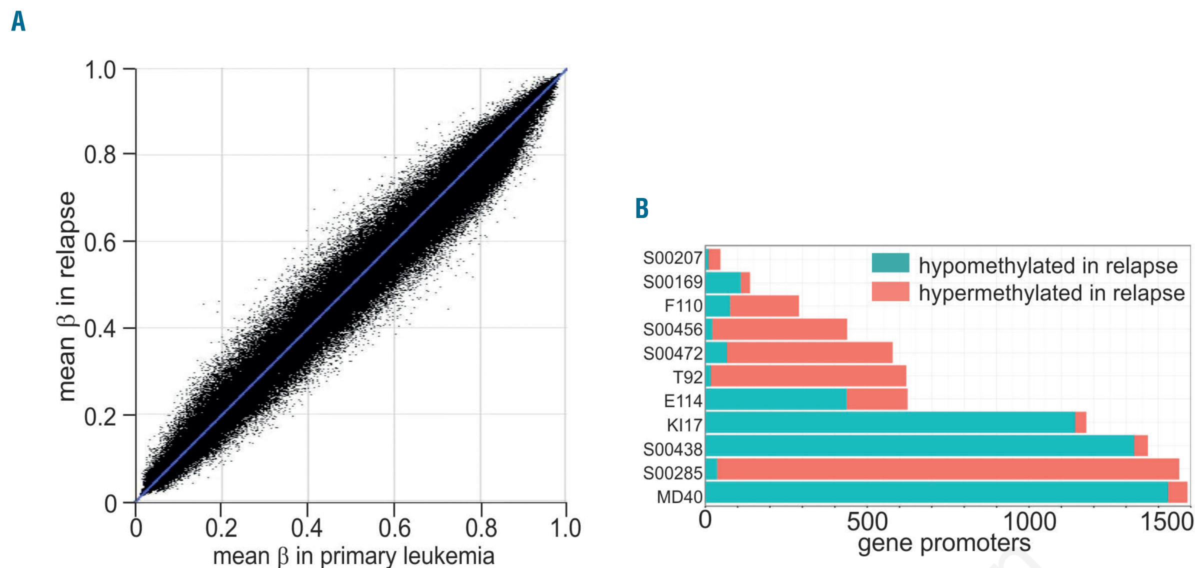


Figure 2. **Methylation analysis identifies promoter specific hypomethylation in relapsed T-ALL. β values in relapse were plotted over β values in primary leukemia in scatter plots. (A) Average β values (single sites) of all patients analyzed. (B) The β value for each promoter in relapse was compared to that in primary leukemia. Promoters with a β value in relapse that was at least 0.2 higher than in primary leukemia were considered to be hypermethylated. Promoters with a β value in relapse that was at least 0.2 lower than in primary leukemia were considered to be hypomethylated. Only promoters of those genes for which at least one gene symbol is assigned and which are represented by at least three probes on the 450k array are represented here.**

relapse. However, a set of promoters was found to be recurrently hypomethylated in relapse compared to primary disease. This set contains many genes related to cancer, albeit not specifically involved in leukemogenesis. The correlation between DNA methylation and RNA abundance could not be analyzed due to lack of RNA samples. However, a strong inverse correlation between promoter methylation and RNA expression has been described before in ALL.[6,10] Hence, both relapse-specific mutations and changes of the DNA methylome are consistent with the activation of additional oncogenic mechanisms in relapsed T-ALL.

The importance of epigenetic changes is specifically highlighted by the findings in patient S00169 (Figure 1A, *Online Supplementary Table S1*), who suffered from a very early relapse and in whom the only detectable mutational difference between primary leukemia and relapse was the acquisition of the activating mutation E1099K in the H3K36 histone methyl transferase WHSC1. This hotspot mutation has been recurrently reported in mantle cell lymphoma[35] and in pediatric BCP-ALL[36,37] and results in cellular transformation.[36] The emergence of a clone carrying the WHSC1 E1099K mutation may have been sufficient to induce early relapse in this patient.

The only gene that was recurrently mutated in a relapse-specific manner was *NT5C2*,[28,29] partially explaining the relative success in the use of nelarabine, which is not a substrate for NT5C2, in T-ALL relapse.[38] In our small series of patients, no other mutation was recurrent and many mutations present in the primary leukemia were lost in relapse. The application of novel, targeted therapies to patients with relapsed T-ALL will, therefore, likely require a thorough genetic characterization of relapse-specific targets.

In conclusion, the data presented here identify two molecularly defined types of relapse in pediatric T-ALL and implicate the selection of subclones, the acquisition of novel somatic mutations and the hypomethylation of promoters as mechanisms driving the progression of T-ALL from primary disease to relapse. It is noteworthy that the relapse-specific alterations tend to activate general mechanisms of carcinogenesis rather than known leukemia-specific drivers.

Acknowledgments

The authors would like to thank the following institutions for grants: German Consortium for Translational Cancer Research (DKTK), German Ministry of Education and Research (BMBF, NGFN Plus), "Tour der Hoffnung", Manfred Lautenschläger Stiftung, European Commission (FP7, ERA-NET on Translational Cancer Research, TRANSCALL to MUM; Health-F2-2010-260791 and a European Research Council grant to JOK). JE and JBK received fellowships from the Heidelberg Research Center for Molecular Medicine. ORB received grants from the Excellence Initiative, Innovation Fund FRONTIER, University of Heidelberg. We thank the EMBL GeneCore sequencing team, the EMBL IT unit, the DKFZ genomics and proteomics core facility, Elena Orlova and Margit Happich for excellent technical assistance.

Authorship and Disclosures

Information on authorship, contributions, and financial & other disclosures was provided by the authors and is available with the online version of this article at www.haematologica.org.

References

1. Dores GM, Devesa SS, Curtis RE, Linet MS, Morton LM. Acute leukemia incidence and patient survival among children and adults in the United States, 2001-2007. Blood. 2012;119(1):34-43.
2. Schrappe M, Valsecchi MG, Bartram CR, et al. Late MRD response determines relapse risk overall and in subsets of childhood T-cell ALL: results of the AIEOP-BFM-ALL 2000 study. Blood. 2011;118(8):2077-2084.
3. Van Vlierberghe P, Pieters R, Beverloo HB, Meijerink JP. Molecular-genetic insights in paediatric T-cell acute lymphoblastic leukaemia. Br J Haematol. 2008;143(2):153-168.
4. Nguyen K, Devidas M, Cheng SC, et al. Factors influencing survival after relapse from acute lymphoblastic leukemia: a Children's Oncology Group study. Leukemia. 2008;22(12):2142-2150.
5. Einsiedel HG, von Stackelberg A, Hartmann R, et al. Long-term outcome in children with relapsed ALL by risk-stratified salvage therapy: results of trial acute lymphoblastic leukemia-relapse study of the Berlin-Frankfurt-Munster Group 87. J Clin Oncol. 2005;23(31):7942-7950.
6. Hogan LE, Meyer JA, Yang J, et al. Integrated genomic analysis of relapsed childhood acute lymphoblastic leukemia reveals therapeutic strategies. Blood. 2011;118(19):5218-5226.
7. Ma X, Edmonson M, Yergeau D, et al. Rise and fall of subclones from diagnosis to relapse in pediatric B-acute lymphoblastic leukaemia. Nat Commun. 2015;6:6604
8. Bhojwani D, Kang H, Moskowitz NP, et al. Biologic pathways associated with relapse in childhood acute lymphoblastic leukemia: a Children's Oncology Group study. Blood. 2006;108(2):711-717.
9. Mullighan CG, Phillips LA, Su X, et al. Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. Science. 2008;322(5906):1377-1380.
10. Nordlund J, Backlin CL, Wahlberg P, et al. Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia. Genome Biol. 2013;14(9):r105.
11. Eckert C, Flohr T, Koehler R, et al. Very early/early relapses of acute lymphoblastic leukemia show unexpected changes of clonal markers and high heterogeneity in response to initial and relapse treatment. Leukemia. 2011;25(8):1305-1313.
12. Clappier E, Gerby B, Sigaux F, et al. Clonal selection in xenografted human T cell acute lymphoblastic leukemia recapitulates gain of malignancy at relapse. J Exp Med. 2011;208(4):653-661.
13. Szczepanski T, van der Velden VH, Waanders E, et al. Late recurrence of childhood T-cell acute lymphoblastic leukemia frequently represents a second leukemia rather than a relapse: first evidence for genetic predisposition. J Clin Oncol. 2011;29(12):1643-1649.
14. Stark B, Nirel R, Avrahami G, et al. Long-term results of the Israeli National Studies in childhood acute lymphoblastic leukemia: INS 84, 89 and 98. Leukemia. 2010;24(2):419-424.
15. Fronkova E, Mejstrikova E, Avigad S, et al. Minimal residual disease (MRD) analysis in the non-MRD-based ALL IC-BFM 2002 protocol for childhood ALL: is it possible to avoid MRD testing? Leukemia. 2008;22(5):989-997.
16. Breit S, Stanulla M, Flohr T, et al. Activating NOTCH1 mutations predict favorable early treatment response and long-term outcome in childhood precursor T-cell lymphoblastic leukemia. Blood. 2006;108(4):1151-1157.
17. Bibikova M, Le J, Barnes B, et al. Genome-wide DNA methylation profiling using Infinium(R) assay. Epigenomics. 2009;1(1):177-200.
18. Teschendorff AE, Marabita F, Lechner M, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics. 2013;29(2):189-196.
19. Assenov Y, Muller F, Lutsik P, Walter J, Lengauer T, Bock C. Comprehensive analysis of DNA methylation data with RnBeads. Nat Methods. 2014;11(11):1138-1140.
20. Fordham SE, Cole M, Irving JA, Allan JM. Cytarabine preferentially induces mutation at specific sequences in the genome which are identifiable in relapsed acute myeloid leukaemia. Leukemia. 2015;29(2): 491-494.
21. Fizzotti M, Cimino G, Pisegna S, et al. Detection of homozygous deletions of the cyclin-dependent kinase 4 inhibitor (p16) gene in acute lymphoblastic leukemia and association with adverse prognostic features. Blood. 1995;85(10):2685-2690.
22. Gutierrez A, Sanda T, Ma W, et al. Inactivation of LEF1 in T-cell acute lymphoblastic leukemia. Blood. 2010;115(14):2845-2851.
23. Lahortiga I, De Keersmaecker K, Van Vlierberghe P, et al. Duplication of the MYB oncogene in T cell acute lymphoblastic leukemia. Nat Genet. 2007;39(5):593-595.
24. Clappier E, Cuccuini W, Kalota A, et al. The C-MYB locus is involved in chromosomal translocation and genomic duplications in human T-cell acute leukemia (T-ALL), the translocation defining a new T-ALL subtype in very young children. Blood. 2007;110(4):1251-1261.
25. Palomero T, Sulis ML, Cortina M, et al. Mutational loss of PTEN induces resistance to NOTCH1 inhibition in T-cell leukemia. Nat Med. 2007;13(10):1203-1210.
26. Zuurbier L, Petricoin EF 3rd, Vuerhard MJ, et al. The significance of PTEN and AKT aberrations in pediatric T-cell acute lymphoblastic leukemia. Haematologica. 2012;97(9):1405-1413.
27. Kleppe M, Lahortiga I, El Chaar T, et al. Deletion of the protein tyrosine phosphatase gene PTPN2 in T-cell acute lymphoblastic leukemia. Nat Genet. 2010;42(6):530-535.
28. Meyer JA, Wang J, Hogan LE, et al. Relapse-specific mutations in NT5C2 in childhood acute lymphoblastic leukemia. Nat Genet. 2013;45(3):290-294.
29. Tzoneva G, Perez-Garcia A, Carpenter Z, et al. Activating mutations in the NT5C2 nucleotidase gene drive chemotherapy resistance in relapsed ALL. Nat Med. 2013;19(3):368-371.
30. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7(4):248-249.
31. Ding L, Ley TJ, Larson DE, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature. 2012;481(7382):506-510.
32. Gruber SB, Ellis NA, Scott KK, et al. BLM heterozygosity and the risk of colorectal cancer. Science. 2002;297(5589):2013.
33. Zauber NP, Sabbath-Solitare M, Marotta S, et al. Clinical and genetic findings in an Ashkenazi Jewish population with colorectal neoplasms. Cancer. 2005;104(4):719-729.
34. Van Vlierberghe P, Palomero T, Khiabanian H, et al. PHF6 mutations in T-cell acute lymphoblastic leukemia. Nat Genet. 2010;42(4):338-342.
35. Bea S, Valdes-Mas R, Navarro A, et al. Landscape of somatic mutations and clonal evolution in mantle cell lymphoma. Proc Natl Acad Sci USA. 2013;110(45):18250-18255.
36. Jaffe JD, Wang Y, Chan HM, et al. Global chromatin profiling reveals NSD2 mutations in pediatric acute lymphoblastic leukemia. Nat Genet. 2013;45(11):1386-1391.
37. Oyer JA, Huang X, Zheng Y, et al. Point mutation E1099K in MMSET/NSD2 enhances its methyltranferase activity and leads to altered global chromatin methylation in lymphoid malignancies. Leukemia. 2014;28(1):198-201.
38. Gokbuget N, Basara N, Baurmann H, et al. High single-drug activity of nelarabine in relapsed T-lymphoblastic leukemia/lymphoma offers curative option with subsequent stem cell transplantation. Blood. 2011;118(13):3504-3511.
39. Schrappe M, Reiter A, Ludwig WD, et al. Improved outcome in childhood acute lymphoblastic leukemia despite reduced use of anthracyclines and cranial radiotherapy: results of trial ALL-BFM 90. German-Austrian-Swiss ALL-BFM Study Group. Blood. 2000;95(11):3310-3322.
40. Moricke A, Reiter A, Zimmermann M, et al. Risk-adjusted therapy of acute lymphoblastic leukemia can decrease treatment burden and improve survival: treatment results of 2169 unselected pediatric and adolescent patients enrolled in the trial ALL-BFM 95. Blood. 2008;111(9):4477-4489.
41. O'Neil J, Grim J, Strack P, et al. FBW7 mutations in leukemic cells mediate NOTCH pathway activation and resistance to gamma-secretase inhibitors. J Exp Med. 2007;204(8):1813-1824.
42. Thompson BJ, Buonamici S, Sulis ML, et al. The SCFFBW7 ubiquitin ligase complex as a tumor suppressor in T cell leukemia. J Exp Med. 2007;204(8):1825-1835.
43. Weng AP, Ferrando AA, Lee W, et al. Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. Science. 2004;306(5694):269-271.
44. Tosello V, Mansour MR, Barnes K, et al. WT1 mutations in T-ALL. Blood. 2009;114(5):1038-1045.
45. Ntziachristos P, Tsirigos A, Van Vlierberghe P, et al. Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia. Nat Med. 2012;18(2):298-301.
46. Huether R, Dong L, Chen X, et al. The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. Nat Commun. 2014;5:3630.
47. Goss KH, Risinger MA, Kordich JJ, et al. Enhanced tumor formation in mice heterozygous for Blm mutation. Science. 2002;297(5589):2051-2053.
48. Holmfeldt L, Wei L, Diaz-Flores E, et al. The genomic landscape of hypodiploid acute lymphoblastic leukemia. Nat Genet. 2013;45(3):242-252.
49. Wu MY, Eldin KW, Beaudet AL. Identification of chromatin remodeling genes Arid4a and Arid4b as leukemia suppressor genes. J Natl Cancer Inst. 2008;100(17):1247-1259.
50. Bains T, Heinrich MC, Loriaux MM, et al. Newly described activating JAK3 mutations in T-cell acute lymphoblastic leukemia. Leukemia. 2012;26(9):2144-2146.