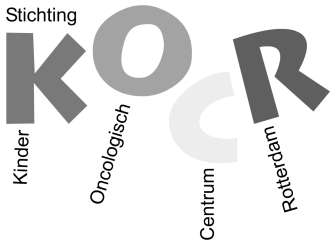
# Gene-expression profiles and oncogenes in pediatric T-cell acute lymphoblastic leukemia

Gratitude is expressed to the Dutch Cancer Society (KWF kankerbestrijding) and the Pediatric Oncology Foundation Rotterdam (KOCR) for financial support of the print and reproduction of this thesis.


The research was financially supported by the Dutch Cancer Society (EMCR 2006-3500).

Printed by Drukkerij Van Denderen, Groningen
ISBN: 978-90-9026438-7

# Gene-expression profiles and oncogenes in pediatric T-cell acute lymphoblastic leukemia

Gen-expressie profielen en oncogenen
in T-cel acute lymfatische leukemie bij kinderen

**Proefschrift**

Ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
Op gezag van de
Rector magnificus
Prof. Dr. H. G. Schmidt
en volgens besluit van het College voor Promoties

De openbare verdediging zal plaatsvinden op
woensdag 30 november 2011 om 15:30

door
Irene Homminga
geboren te Groningen

**Promotiecomissie**

| | |
|---|---|
| Promotor: | Prof. dr. R. Pieters |
| Overige leden: | Prof. dr. J. Cornelissen<br>Dr. M.L. den Boer<br>Dr. A.W. Langerak |
| Copromotor: | Dr. J.P.P. Meijerink |

*Voor Hans*
*Max & Henk*

**CONTENTS**

# General introduction

# 1. GENERAL INTRODUCTION

## 1.1 Blood & leukemia

Blood consists of serum and three other main ingredients: erythrocytes (red blood cells), thrombocytes (platelets) and leukocytes (white blood cells). Leukocytes normally compose less than 1% of the blood volume but have important functions in our defense against foreign and endogenous pathogens. Many different leukocytes can be distinguished, the main types being lymphocytes (~75% T-cells, ~25% B-cells), monocytes and granulocytes[1]. All these different blood cells arise through distinct and tightly controlled developmental stages from hematopoietic precursor cells, which reside in the bone marrow and thymus (Figure 1). However, sequential mutations, chromosomal rearrangements and epigenetic changes can cause a precursor cell to be blocked from further differentiation and to start proliferate in a uncontrollable manner which results in cancer. Uncontrolled proliferation of blood precursor cells is called leukemia. Different types of leukemia are distinguished based on their lineage of origin; myeloid or lymphoblastic leukemia. Lymphoblastic leukemia can be further divided into B-cell precursor (BCP) or T-cell lymphoblastic leukemia. In acute lymphoblastic leukemia malignant cells are arrested at a relative immature stage whereas in chronic lymphocytic leukemia, cells have a more differentiated phenotype.

## 1.2 Acute lymphoblastic leukemia

Acute lymphoblastic leukemia (ALL) is the most common type of cancer in children, comprising approximately 25% of all childhood malignancies.[2] Patients usually present with nonspecific symptoms. Most of these can be explained by the accumulation of leukemic lymphoblasts in the bone marrow, which repress the development of normal, healthy, blood cells. These symptoms can be fever (due to normal leukocyte deficiency), fatigue (due to erythrocyte deficiency) and bleeding (due to thrombocyte deficiency). Infiltration of leukemic cells to sites outside the bone marrow can lead to other symptoms such as bone pain (caused by infiltration of the periosteum), lymphadenopathy (caused by infiltration of the lymph nodes) and headache (caused by infiltration of the central nervous system). Without treatment, leukemia is lethal. Therapy has become more and more effective over the years and the survival of ALL patients has increased to about 80%[3]. Therapy consists mainly of combination chemotherapy.

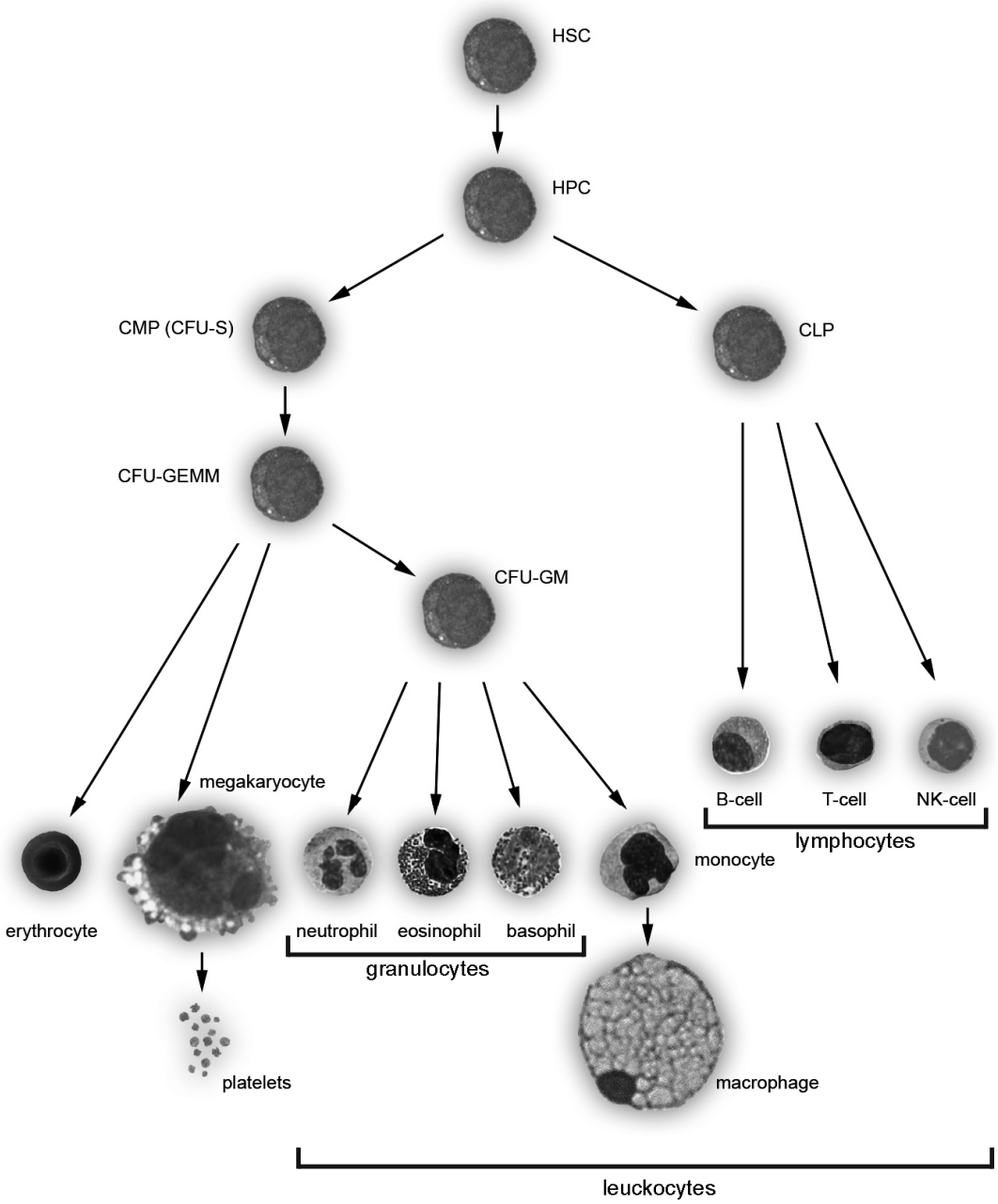



**Figure 1: Schematic overview of hematopoiesis.**
HSC: hematopoietic stem cell, HPC: hematopoietic progenitor cell, CMP: common myeloid progenitor, CFU-GEMM: Colony forming unit generating granulocytes, erythrocytes, monocytes and megakaryocytes, CFU-GM: colony forming unit generating granulocytes and monocytes, CLP: common lymphoid progenitor.
*Adapted picture from Dr. M. William Lensch, George Q. Daley Laboratory/ Harvard Medical School, Boston, US. Printed with permission.*

### 1.3 Pediatric T-cell acute lymphoblastic leukemia

This thesis focuses on T-cell acute lymphoblastic leukemia (T-ALL) in children. T-ALL accounts for approximately 15% of pediatric ALL. Each year, approximately 20 children are diagnosed with T-ALL in the Netherlands. The overall survival is ~70%[4], making T-ALL a relative poor prognostic subgroup. Therefore, improved insight in the biology and pathogenesis of T-ALL is necessary as it may lead to better and more targeted therapies. In addition, the identification of genetic subgroups that have prognostic relevance can also lead to stratification of T-ALL patients and adjusted treatment protocols (higher intensity for poor prognostic subgroups) that may improve overall survival. Especially in the last decade, many different genetic abnormalities have been identified which play a role in T-ALL. These aberrations can be divided in type A and type B.[5,6] Type A aberrations are predominantly mutually exclusive, and often involve chromosomal translocations or large genomic deletions or amplifications and delineate specific T-ALL subgroups. These subgroups are also associated with maturational arrest at particular T-cell developmental stages[7] (see Figure 2). These developmental stages are characterized by the presence or absence of immunophenotypic markers (such as CD1 and mCD3) and different gene-expression profiles. Important genes involved in type A aberrations are *TAL1, LMO2, LMO1, TLX1, TLX3* and *HOXA.* These genes are ectopically expressed due to the genetic aberrations and thus function as oncogenes. In many cases, an enhancer region of a T-cell receptor gene (*TRA@, TRD@, TRB@),* or a specific enhancer region downstream of the *BCL11B* gene, is juxtaposed to the oncogene causing transcriptional activation. This mechanism is common in T-ALL as consequence of the physiologic T-cell receptor rearrangement that occurs in normal T-cell development. The T-cell receptor genes are rearranged to acquire a wide diversity of antigen recognition. Aberrations in this process, which involves breakage and re-ligation of the double stranded DNA, can result in translocations or inversions to ectopic sites. The precise mechanism by which type A oncogenes contribute to leukemogenesis is not clear, but they are thought to function predominantly by blocking differentiation. Type B mutations affect many genes such as *NOTCH1, FBXW7, PTEN, RAS, JAK1* and *CDKN2A-B* and often consist of point mutations, small insertions or deletions. The mutations can be activating as well as inactivating and, in contrast to type A aberrations, they seem not specifically associated with particular T-ALL subgroups (Figure 2). Type B mutations are involved in diverse processes such as cell cycle control, self-renewal, T-cell receptor signaling, differentiation or tyrosine kinase activation[5] .

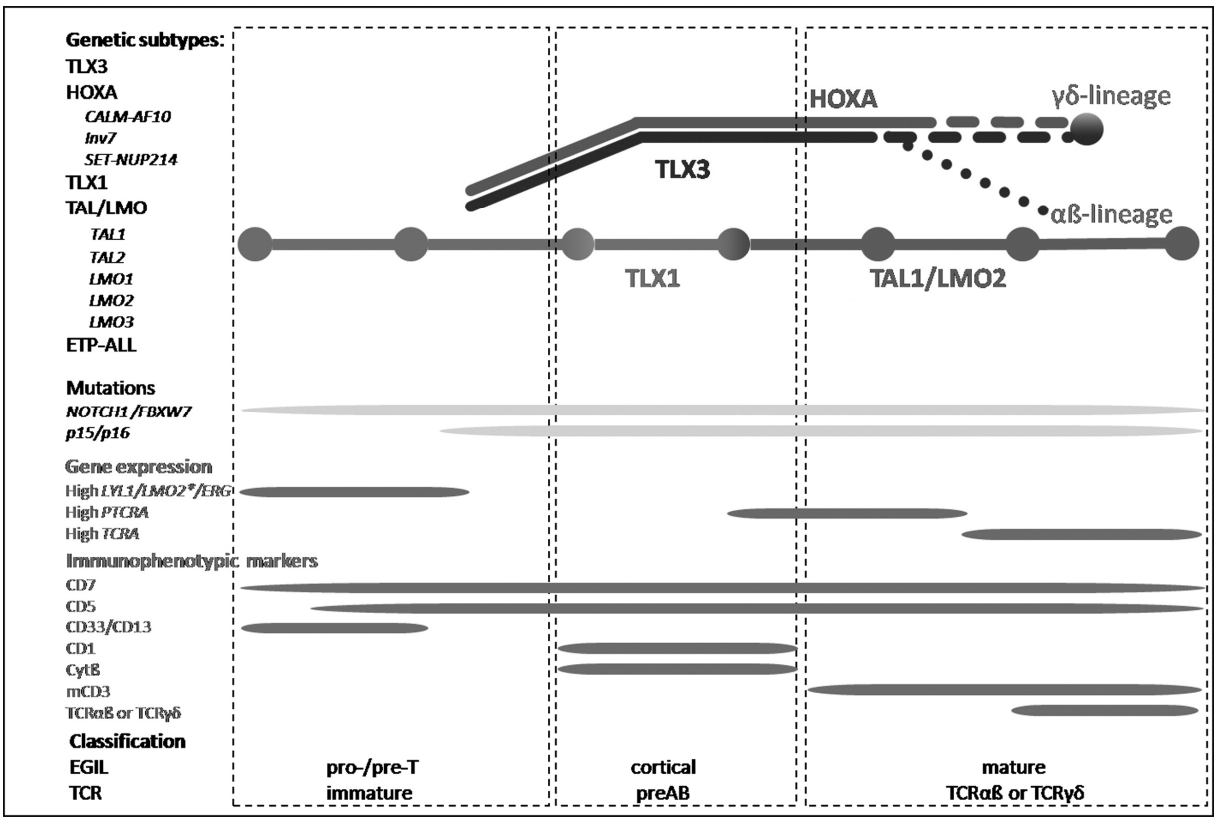



**Figure 2**. Schematic overview of T-ALL genetic subgroups *TLX3*, *TLX1*, *HOXA* and *TAL/LMO* in relation to their T-cell developmental stage based on EGIL or TCR classification systems. The presence of particular mutations such as *p15/p16* deletions, NOTCH1 activating mutations, the high expression of genes such as LYL1, LMO2, ERG, PTCRA and TCRA and the appearance of immunophenoypic markers are indicated. **LMO2* is ectopically expressed in *LMO2*-rearranged cases but is not included in this figure. *Adapted from Meijerink et al, Best Pract Res Clin Haematol. 2010 Sep;23(3):307-18*

## 2. AIM OF THE THESIS

Even though many chromosomal rearrangements and mutations are known in pediatric T-ALL, still about 40% of patients lack a known type A aberration, and additional type B aberrations are currently still being identified. In addition to genetic aberrations, gene-expression profiling can provide important insights in T-ALL and identify subgroups based on shared biology, or shared pathologic pathways, which may overlap with subgroups defined by specific genetic aberrations. Knowledge of these genetic aberrations and gene-signatures is essential for understanding T-ALL, for identifying targets for future targeted therapy and for risk adapted treatment stratification within T-ALL.

The main aim of this thesis is to identify genetic aberration- and gene-signature-based subgroups in pediatric T-ALL patients, to genetically characterize these subgroups and determine their prognostic relevance (Chapter 1-5).

In addition, in Chapter 6, we investigate the *in-vitro* sensitivity of different pediatric leukemia types (T-ALL, BCP-ALL or AML) to two agents that are currently in clinical trials: forodesine and nelarabine, to better predict which patients might benefit from treatment with either one of these compounds in the future.

In the next section we will first provide a short overview of the most important methods that were used to reach our aims. In section 4 of this chapter, a short introduction is given for each chapter of this thesis.

## 3. METHODS OVERVIEW

At the center of our research efforts lies a cohort of 146 pediatric T-ALL patients for which extensive data was present, including karyotypic data, gene-expression data (for almost all human genes) and data on the presence of the most common genetic aberrations that were known at the start of this thesis (*TAL1*, *LMO2*, *HOXA*, *TLX1*, *TLX3* and *NOTCH1* activating aberrations). These data led to the realization that ~40% of our patient cohort lacked a known type A aberration. The data also gave us the first clues where to look for new genetic abnormalities. As these aberrations had thus far remained elusive we needed to apply a combination of different techniques to reveal them. Below we give an overview of the main methods that were used in this thesis to detect *genetic* abnormalities. In table 1 these methods are summarized.

### 3.1 Conventional techniques

#### 3.1.1 Karyotyping

Karyotyping is used to detect gross chromosomal aberrations such as large deletions, large amplifications or translocations. Starting material is an amount of ~5 million viable leukemic cells which are cultured and then arrested in metaphase by addition of colcemide, an inhibitor of spindle assembly that is required for cell division. Nuclei-suspensions are then applied on glass slides so that the chromosomes of nuclei that are in metaphase can be identified as separate chromosomes, and can be stained by Giemsa. This results in a unique and distinct

banding pattern for each chromosome (Figure 3A). Unfortunately, this procedure does not succeed for all patient samples and only large aberrations can be seen on karyotype, up to a resolution of 10-20 Mb. Frequently, aberrations such as translocations or inversions are missed because they do not have visible effects on the banding pattern. These aberrations are called cryptic or hidden rearrangements.

In this thesis, karyotype analysis gave us a clue to look for a translocation in the region 21q22 where the *RUNX1* gene is situated. We performed the 3'-RACE technique (see below) which led to the identification of a novel *RUNX1-AFF3* fusion gene (Chapter 1).

**Table 1: Overview of methods used to detect genetic abnormalities**

| METHOD | COVERAGE | RESO-LUTION | DETECTION* | PRICE /sample | MATERIAL |
|---|---|---|---|---|---|
| **Conventional** | | | | | |
| Karyotype | whole genome | Mb | trans/inv large amps/dels | ~$50 | 5 million viable cells |
| FISH | 1 - 2 loci/genes ~1000Kb | Mb | trans/inv large amps/ dels | ~$50 | ~25.000 viable cells |
| **Array based** | | | | | |
| arrayCGH | whole genome | Kb | large amps/dels | ~$500 | ~4µg DNA |
| 4C | 1 – 10 loci ~5Mb | Kb | trans/inv large amps/dels | ~$500 | 5-20 million viable cells |
| **PCR-based** | | | | | |
| PCR | 1 locus/gene 100-10,000bp | bp | mutations small amps/dels | ~$50 | ~50ng DNA |
| LM-PCR | 1 locus/gene 100-10,000bp | bp | trans/inv, small amps/dels | ~$50 | ~4µg DNA |
| 3'- RACE | 1 (fusion)gene 100-10,000bp | bp | fusiongenes (trans/inv) | ~$50 | ~1µg RNA |

**trans: translocations, inv: inversions, amps: amplifications, dels: deletions*

### 3.1.2 Fluorescent in situ hybridisation (FISH)

This technique can be used to detect large (> 100 Kb) deletions or amplifications but is also suited to detect chromosomal rearrangements such as translocations or inversions. FISH is more restricted than karyotype as only a limited number of selected loci can be visualized at once. Approximately 25,000 viable cells are spun on a glass slide and specific loci on the genome are visualized with fluorescent probes with a size of 100-200 Kb. For the detection of chromosomal

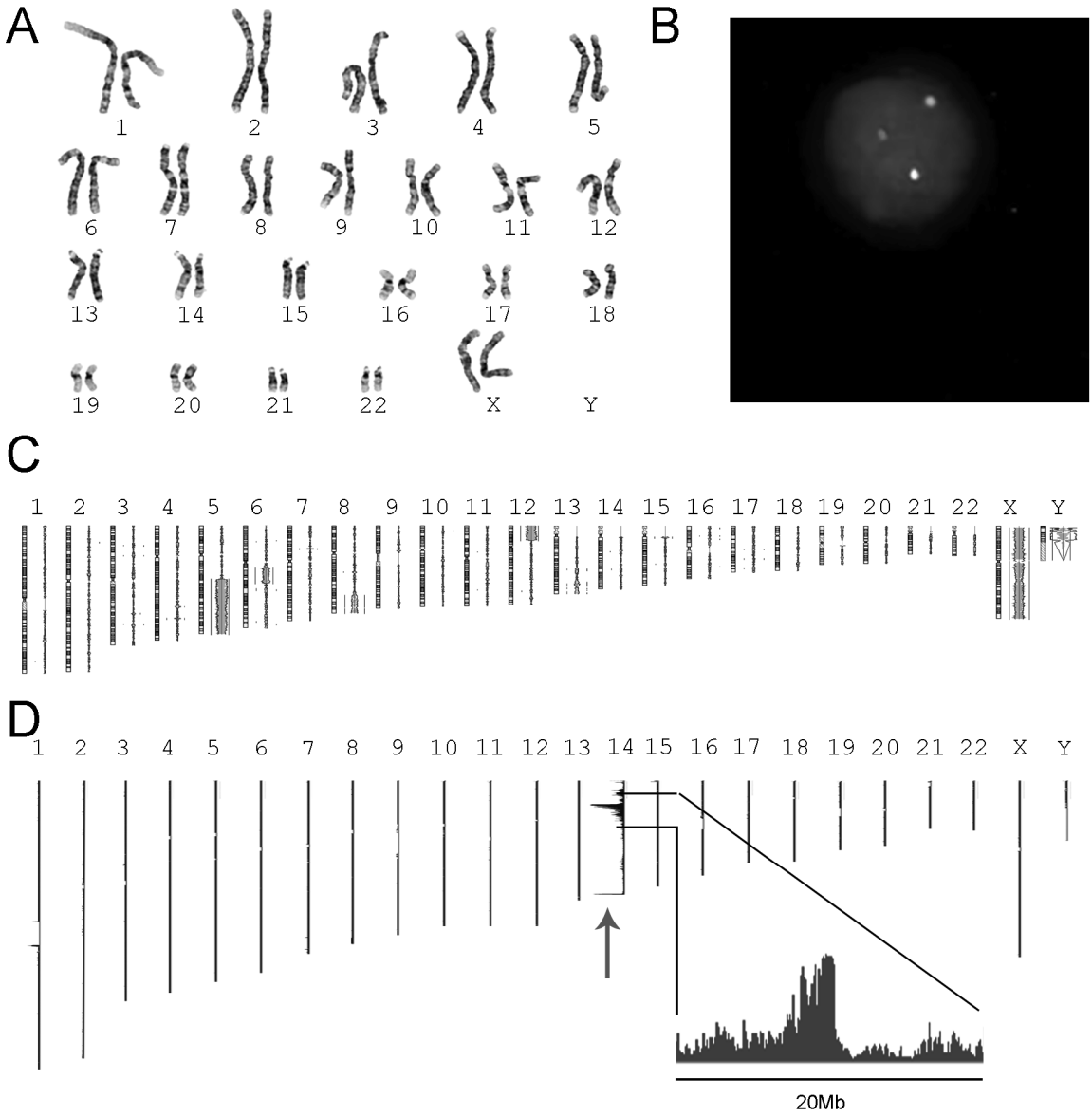



**Figure 3: Results for different types of genetic analyses used in this thesis.**
A, Sorted human chromosomes in a karyotypic analysis. Each chromosome has a unique banding pattern. B, FISH analysis showing a red/green fusion signal (normal chromosome) and separate red and green signals for both derivative chromosomes of a translocation. C, ArrayCGH results for all chromosomes for one patient sample. A red signal to the right indicates a deletion (chromosome 5, 6, 12 and Y), a blue signal to the right an amplification (chromosome 8 and X). D, 4C results for all chromosomes for one viewpoint (*NKX2-1*) on chromosome 14 (upper signal and inset). An additional signal is present on the lower tip (grey arrow) of chromosome 14 indicating an inversion of chromosome 14 (in this case between *NKX2-1* and the *IgH@* locus).

rearrangements, often a split signal FISH is developed. This means that a certain locus for which a chromosomal rearrangement is suspected, is flanked by two probes that are labeled in red and green. When the locus is positioned on a autosomal chromosome, the FISH procedure in a normal nucleus will show two fusion signals (red-green, or yellowish signals). A balanced translocation or inversion will results in the separation of a red and green probe (split signal), and the resulting derivative chromosomes will only have a green or a red signal. FISH analysis then shows one fusion signal (the normal chromosome) and a separate green and a separate red signal (Figure 3B). An advantage of FISH over karyotype analysis, is that one can use metaphase as well as interphase nuclei.

We extensively used FISH in the research described in this thesis. Most importantly, we used it to screen for abnormalities that involve known recurrent translocation loci such as the T-cell receptors and *BCL11B*. Combined with other techniques, this led to the identification of new oncogenic translocation partners such as *NKX2-1* and *NKX2-2* (Chapter 1, 2). Potential oncogenes were also screened for involvement in chromosomal aberrations (Chapter 1) and we used FISH to validate rearrangements that were identified with other techniques (Chapter 1 and 3).

## 3.2 Array based techniques

### 3.2.1. Array comparative genomic hybridization (arrayCGH)

To detect deletions or amplifications with a range that covers the whole human genome, one can use array comparative genomic hybridization. In this technique, DNA of a specific sample of interest is compared to normal genomic control DNA. To illustrate the progress in this technique; one needed 10 µg of DNA (equivalent to $1.5 \cdot 10^6$ leukemic cells) per sample to perform arrayCGH at the start of this thesis, compared to only 100 ng (equivalent to 15,000 cells) with the latest array formats and labeling techniques. DNA samples are digested by restriction enzymes or heat into smaller fragments and labeled with fluorescent dyes. The control DNA is labeled with a different color dye (usually green) than the specific sample (usually labeled in red). Both samples are then mixed in a 1 to 1 ratio, and put on the glass array where labeled fragments can hybridize with DNA probe sets that are spotted on the array and that represent the entire human genome. Each probe set represents a specific site on our genomic DNA. The probe set densities differ per

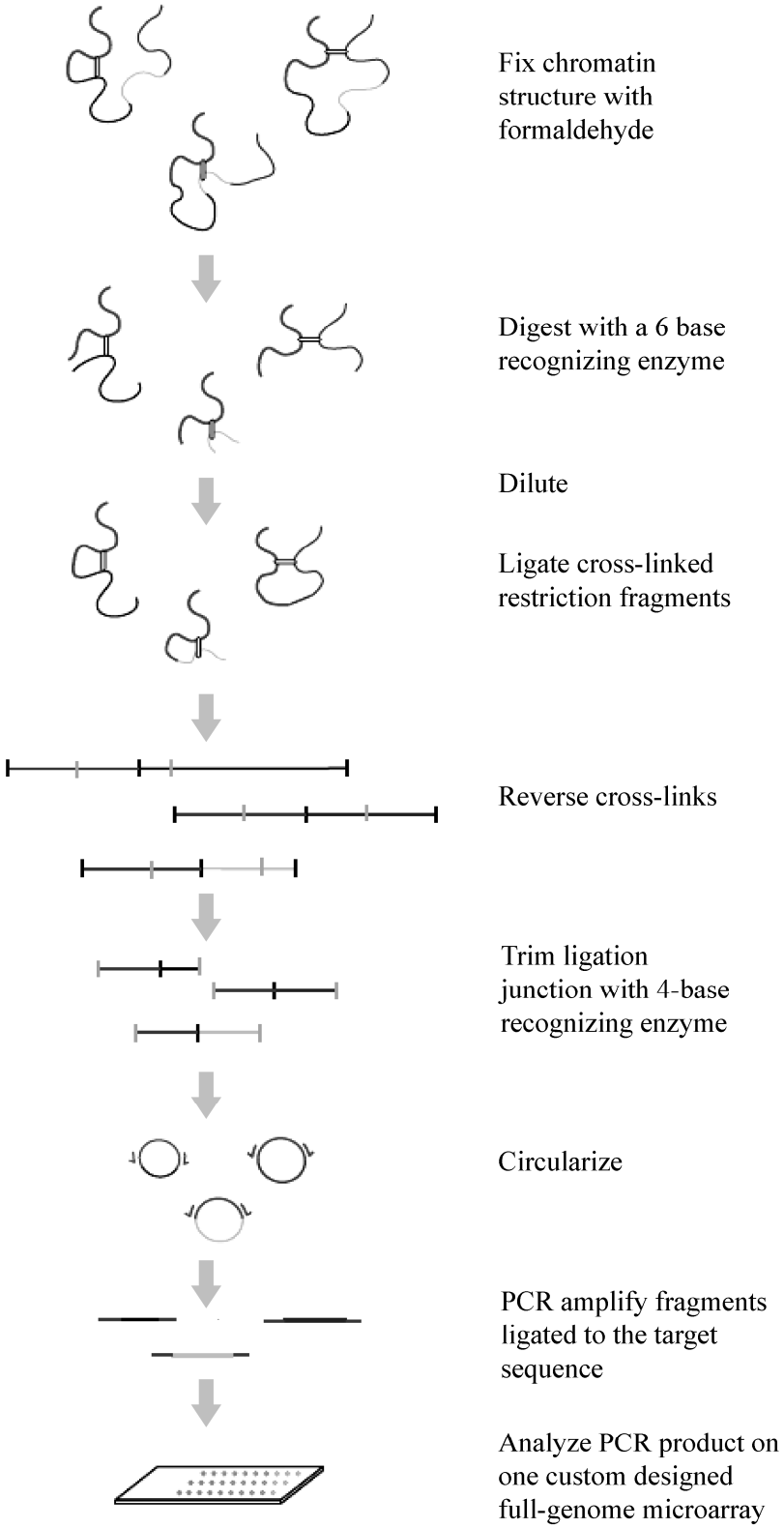
array. ~44,000 probe sets were spotted per array 4 years ago and presently with the latest array formats, this amount has risen to ~400,000 probe sets. When the two different colored DNA fragments bind to the probe sets, the ratios between both colors give information on the copy number status of the genome of the specific DNA versus the normal control DNA. Increased red over green intensities point to local DNA amplification whereas increased green over red intensities point to a local chromosomal deletion in the leukemic cell. Amplifications and deletions in the specific DNA can thus be detected and visualized by a computer program, as depicted in figure 2C. The advantage of this procedure is that the entire genome is visualized at a resolution (~35 Kb) that is much higher than that of karyotype analysis (>10-20·$10^6$ Mb). The main disadvantages of this technique are the relative high costs (Table 1) and that balanced translocations or inversions cannot be detected as no gains or losses of chromosomal DNA occur.

In this thesis we performed arrayCGH on large series of patients. This led to the identification of deletions near a new oncogene *MEF2C* (Chapter 1) and gave clues for an unbalanced translocation involving the T-cell receptor beta locus (*TRB@*) which was further analyzed by 4C (Chapter 1, and see below). Type B aberrations were also detected, such as a deletion of *FBXW7* and *WT1* (Chapter 3). In Chapter 3, we identified a del5q35 that is specific for *TLX3* rearranged patients.

### 3.2.2. Chromosome conformation capture on chip (4C)

Chromosome conformation capture on chip (Figure 4) can be used for multiple purposes. Originally it was designed to study the three-dimensional structure of the genome (chromosome conformation) in living cells, however, it can also be used as a tool to detect translocations and identify unknown translocation partners.[8] For the 4C procedure an amount of 10 million of viable cells is needed. These are fixed with formaldehyde, providing cross links between DNA and proteins. This cross linking preserves the proximity of chromosomal structure when the cells are processed further. The DNA is then digested (HindIII), which will lead to complexes of protein and DNA sequences of several Kb's. The protein-DNA complexes are then diluted and the DNA is randomly ligated again. In this way, DNA fragments that were in close proximity in the viable cell, either by being situated on the same chromosomal region or by chromosomal folding, have a high probability to become fused. Subsequently, the DNA and protein are de-crosslinked and then digested again with a restriction enzyme that can cut at high frequency. The DNA is then diluted and ligated to allow circularization of

**Figure 4: Outline of 4C-technology.**




.

individual restriction fragments. This way circular DNA of ~200-800 bp is obtained that can be easily amplified by an inverse PCR. Most of these circles consist of two pieces of DNA that were originally in proximity to each other in the living cell. Because of the circular configuration, unknown proximal DNA can be amplified by inverse PCR from a specific locus of interest. In general, DNA fragments that flank the region of interest on the same chromosome will have a higher chance to ligate to the fragment of interest than fragments that originally were localized at greater distance from the site of interest. The PCR product is labeled and put on an array containing probesets that represents the entire genomic DNA pool (one probe per HindIII digested DNA fragment). After scanning of the array and computational manipulations, a bell shaped curve of hybridization intensities can be seen around the PCR viewpoint (Figure 3D), with the highest fluorescent intensities directly flanking a region of interest and a gradual decline in intensities for fragments at increasing distance to a distance of ~5 Mb on both sides. When a translocation is present, which has a breakpoint within this high intensity region, part of the signal will be present on a different chromosomal region (the locus of the translocation partner). This way, translocations, inversions and also deletions within a region of ~5 Mb of the viewpoint can be visualized (Figure 3D). Disadvantages of this technique are the relative high cost, and, as with FISH and ligation mediated PCR (LM-PCR, see below), the necessity to start from a known locus or gene. But even though 4C is limited to one or several chromosomal regions of interest, this technique filled a gap in technologies, as thus far balanced chromosomal translocations or inversions to unknown partners could not be resolved at such a high resolution (<10 Kb). Furthermore, the immediate identification of unknown translocation partners offers a great advantage of the 4C technology.

In this thesis, 4C was essential in the identification of a cryptic inversion involving the *TCRA/D@* or *IgH@* locus and the novel oncogene *NKX2-1* on chromosome 14. Other new translocations were also identified with this technique (Chapter 1).

### 3.3 Polymerase chain reaction (PCR) based techniques

The polymerase chain reaction (PCR, reviewed in[9]), developed in the late 80's, enables the amplification of a short (~100-10,000bp) fragment of DNA for further analyses. It is a widely used technique that has become an integral part of molecular-genetic research. Usually, genomic DNA or complementary DNA

(cDNA obtained from RNA) is used and a locus of interest is amplified using two specific primers (16-24 bp long strands of DNA complementary to a specific site). The PCR products can then be separated based on size in an agarose gel, or sequenced for detection of small mutations the size of several base pairs or used for a wide variety of other applications.

### 3.3.1 Ligation mediated PCR (LM-PCR)

Ligation mediated PCR can be used to detect the exact genomic breakpoint sequence of a translocation, deletion or amplification. In case of a translocation to an unknown site, the translocation partner can be identified by this technique. To perform an LM-PCR, 4 µg of genomic DNA of a sample is digested by 4 restriction enzymes that cut roughly every 4000 base pairs. Four different enzymes are used because an enzyme may have a recognition site that is positioned in such a way that it is difficult or impossible to detect the breakpoint, e.i. positioned too close or very far away from the breakpoint. Therefore the separate use of 4 enzymes increases the chance that a breakpoint is detected by this procedure. Upon DNA digestion, DNA linkers are ligated to the ends of the DNA fragments, making it possible to amplify a specific DNA fragment by using one specific primer (on the region of interest) and one reverse primer that is positioned in the linker. This way, a DNA fragment can be amplified that harbors the transition from a known sequence (where the specific primer binds) over the breakpoint region into a DNA sequence that represents the translocation partner. After PCR amplification, products can be visualized on an agarose gel and the sizes can be compared to normal control DNA. Alternate sized bands are excised from the gel, and sequenced to identify the exact breakpoint sequence. It is necessary to start an LM-PCR with a primer located near the breakpoint (~100-1000 bp); this means that knowledge of the relative position of a breakpoint region is required beforehand.

In this thesis we used LM-PCR to identify a novel *TRD@-NKX2-2* translocation, to identify the breakpoint of a novel *TRB@-NKX2-1* translocation (Chapter 1) and a *TRB@-LYL1* translocation (Chapter 2).

### 3.3.2 3'-Rapid amplification of cDNA ends (3'-RACE)

Sometimes, translocations or deletions can result in the formation of fusion genes and subsequently fusion proteins with aberrant function, e.g. the SET-NUP214 fusion as result of the del(9)(q34.11q34.13).[10] Using 3'- RACE, the fusion transcript can be amplified from the 3'- end. Only one partner of this fusion

product has to be known for this technique. One µg of RNA of the patient is converted to cDNA by means of extended primers that bind to the 5' polyA tails of mRNA. The fusion transcript can then be amplified by a specific primer designed for the known fusion partner on the 3'- end and a primer complementary to the 5'-end linker.

In this thesis, 3'-RACE was used to detect a fusion transcript in a patient in which a *RUNX1* translocation was suspected based on karyotype. We identified a novel *RUNX1-AFF3* fusion transcript (Chapter 1).

## 4. SHORT OUTLINE OF THE THESIS

In **Chapter 1** we combined gene expression profiling of 117 pediatric patient samples and detailed molecular cytogenetic analyses to identify novel oncogenic rearrangements for 40% of T-ALL samples for which the driving oncogene had remained elusive. Two T-ALL subtypes were identified, each having a specific expression signature and representing approximately 10% of all pediatric T-ALL cases. In one subtype (the proliferative cluster) we identified novel rearrangements involving oncogenes *NKX2-1* and *NKX2-2*. The second subtype (the immature cluster) associated with immature T-cell development and high expression of the transcription factor *MEF2C*. In this subgroup, we identified multiple novel rearrangements that directly or indirectly target *MEF2C*. In addition, we showed that *MEF2C* blocks T-cell differentiation and provides part of the gene-expression signature of the immature cluster. Our data demonstrated that *NKX2-1*and *MEF2C* represent important novel type A oncogenes in T-ALL.

For years, *LYL1* aberrations including translocations were assumed to be associated with the immature T-ALL subgroup because this subgroup shows high *LYL1* expression. In **Chapter 2** we describe a case with a rare *LYL1* translocation that belonged to the more mature *TAL/LMO* subgroup. This seems logical given the high homology between *TAL1* and *LYL1* basic helix-loop-helix oncogenes. Alike rare *TAL1* rearranged cases that may have synergistic *LMO1* or *LMO2* aberrations as well, our *LYL1* rearranged case possessed a del(11)(p12p13) which activated the *LMO2* oncogene.

Approximately 20% of pediatric T-ALL patients harbor a chromosomal rearrangement (t(5;14)(q35;q32)) that upregulates the oncogenic NK-like homeobox transcription factor *TLX3*. In **Chapter 3** we used arrayCGH to investigate whether these patients harbor additional abnormalities. Five recurrent

genomic deletions were identified of which the cryptic deletion del(5)(q35) was exclusively found in *TLX3* rearranged patients. The deleted region, which lies just telomeric of *TLX3* itself, might harbor a negative regulating element or a novel tumor suppressor gene that may cooperate with *TLX3.*

In addition to the NK-like *TLX3* homeobox transcription factor, other NK-like transcription factors are also involved in T-ALL, such as *TLX1*[5] and *NKX3-1*[11]. In chapter 1, we further identified two additional rearranged NK-like homeobox genes: *NKX2-1* and *NKX2-2*. **Chapter 4** reviews the involvement of NK-like transcription factors in cancer in general and leukemia in particular. In T-ALL an exceptional high number of different NK-like homeoboxgenes is implicated. Normally, most of these genes are not expressed in T-cell development, making it difficult to predict which pathways may be activated or inactivated by these genes. However, one NKL gene, *HHEX*, is highly expressed in early T-cell development and implicated in stem cell self-renewal and leukemogenesis.[12,13] We hypothesize that the ectopically upregulated NK-like homeobox genes in T-ALL might share a common downstream oncogenic pathway by mimicking functions of *HHEX*.

In **Chapter 5** we studied the relevance of *NOTCH1* and *FBXW7* mutations, which are considered as type B aberrations in T-ALL. NOTCH1 is a transmembrane receptor that is cleaved upon ligand binding which gives rise to an active intracellular NOTCH1 (ICN) protein. ICN acts as a transcription factor that activates many genes including *c-MYC* and functions as an oncogenic protein in T-ALL. Mutations in *NOTCH1* increase ICN formation in a ligand-indepent manner. Alternatively, mutations in *FBXW7*, which normally facilitates ICN degradation, also result in ICN accumulation. In our T-ALL cohort, 63% of the patients harbored a NOTCH1 activating mutation. Although these *NOTCH1* and *FBXW7* mutations had been previously associated with a good long term outcome in the German BFM study[14], our study did not support such a superior outcome for NOTCH1 activated T-ALL patients treated with DCOG and CoALL protocols.

In **Chapter 6**, we investigated two drugs that are currently in clinical trials for T-ALL: forodesine and nelarabine. Both drugs impact the same enzyme: purine nucleoside phosphorylase (PNP), an enzyme that can degrade deoxyguanosine (dGuo) into guanosine and deoxyribose-1-phosphate. Forodesine blocks this enzyme, which results in increased dGTP levels that exert a toxic effect especially on T-cells. Nelarabine is the pro-drug of ara-G which is phosphorylated to ara-dGTP, which is incorporated into the DNA and is cytotoxic to cells. To better predict which patients might benefit from these compounds, we investigated the

efficacy of these drugs in pediatric AML, BCP-ALL and T-ALL patient samples *in-vitro*. Especially T-ALL samples appeared sensitive to forodesine and ara-G treatment, but also nearly half of BCP-ALL samples responded to these drugs *in-vitro.* AML samples were markedly resistant. Forodesine sensitive samples accumulated more dGTP than resistant samples, and had higher mRNA levels of the gene *dGK* that may result in higher phosphorylation rates of deoxyguanosine. Ara-G sensitive samples had higher levels of *ENT1* and *ENT2* than resistant samples. ENT1 and ENT2 are transporters that can increase influx of ara-G into the cell and thereby can thereby increase ara-G sensitivity.

In **Chapter 7** a summary, discussion and future perspectives are given for the research presented in this thesis. **Chapter 8** contains a brief summary of the thesis in Dutch for uninitiated.

## REFERENCES


1. Moss P, Pettit J, Hoffbrand V. Essential Haematology (ed 5th): Blackwell Publishers; 2006.
2. Pui CH, Relling MV, Downing JR. Acute lymphoblastic leukemia. N Engl J Med. 2004;350:1535-1548.
3. Pui CH, Carroll WL, Meshinchi S, Arceci RJ. Biology, risk stratification, and therapy of pediatric acute leukemias: an update. J Clin Oncol. 2011;29:551-565.
4. Pieters R, Carroll WL. Biology and treatment of acute lymphoblastic leukemia. Pediatr Clin North Am. 2008;55:1-20, ix.
5. Van Vlierberghe P, Pieters R, Beverloo HB, Meijerink JP. Molecular-genetic insights in paediatric T-cell acute lymphoblastic leukaemia. Br J Haematol. 2008;143:153-168.
6. Meijerink JP. Genetic rearrangements in relation to immunophenotype and outcome in T-cell acute lymphoblastic leukaemia. Best Pract Res Clin Haematol. 2010;23:307-318.
7. van Grotel M, Meijerink JP, Beverloo HB, et al. The outcome of molecular-cytogenetic subgroups in pediatric T-cell acute lymphoblastic leukemia: a retrospective study of patients treated according to DCOG or COALL protocols. Haematologica. 2006;91:1212-1221.
8. Simonis M, Klous P, Homminga I, et al. High-resolution identification of balanced and complex chromosomal rearrangements by 4C technology. Nat Methods. 2009;6:837-842.
9. Meijerink J. Real-time quantitative polymerase chain reaction (RQ-PCR). In: Leeuwen W, Vink C, eds. Molecular Diagnostics - techniques and applications. Rotterdam: IVA group BV; 2009.
10. Van Vlierberghe P, van Grotel M, Tchinda J, et al. The recurrent SET-NUP214 fusion as a new HOXA activation mechanism in pediatric T-cell acute lymphoblastic leukemia. Blood. 2008;111:4668-4680.
11. Kusy S, Gerby B, Goardon N, et al. NKX3.1 is a direct TAL1 target gene that mediates proliferation of TAL1-expressing human T cell acute lymphoblastic leukemia. J Exp Med. 2010;207:2141-2156.
12. McCormack MP, Young LF, Vasudevan S, et al. The Lmo2 oncogene initiates leukemia in mice by inducing thymocyte self-renewal. Science. 2010;327:879-883.
13. George A, Morse HC, 3rd, Justice MJ. The homeobox gene Hex induces T-cell-derived lymphomas when overexpressed in hematopoietic precursor cells. Oncogene. 2003;22:6764-6773.
14. Kox C, Zimmermann M, Stanulla M, et al. The favorable effect of activating NOTCH1 receptor mutations on long-term outcome in T-ALL patients treated on the ALL-BFM 2000 protocol can be separated from FBXW7 loss of function. Leukemia. 2010;24:2005-2013.

# CHAPTER 1

## Integrated transcript and genome analyses reveal *NKX2-1* and *MEF2C* as potential oncogenes in T cell acute lymphoblastic leukemia.

Irene Homminga[1], Rob Pieters[1], Anton W. Langerak[2], Johan J. de Rooi[3,4], Andrew Stubbs[3], Monique Verstegen[1], Maartje Vuerhard[1], Jessica Buijs-Gladdines[1], Clarissa Kooi[1], Petra Klous[5,6], Pieter van Vlierberghe[7], Adolfo A. Ferrando[8], Jean Michel Cayuela[9], Brenda Verhaaf[2], H. Berna Beverloo[10], Martin Horstmann[11-13], Valerie de Haas[14], Anna-Sophia Wiekmeijer[15], Karin Pike-Overzet[15], Frank J.T. Staal[15], Wouter de Laat[5,6], Jean Soulier[9], Francois Sigaux[9,16] and Jules P.P. Meijerink[1,16*]

[1]*Department of Pediatric Oncology/Hematology, Erasmus MC / Sophia Children's Hospital, Rotterdam, The Netherlands.*
[2]*Department of Immunology, Erasmus MC, Rotterdam, The Netherlands.*
[3]*Department of Bioinformatics, Erasmus MC, Rotterdam, The Netherlands.*
[4]*Department of Biostatistics, Erasmus MC, Rotterdam, The Netherlands.*
[5]*Department of Cell Biology, Erasmus MC, Rotterdam, The Netherlands.*
[6]*Hubrecht Institute-KNAW & University Medical Center Utrecht, Utrecht, The Netherlands.*
[7]*Center for Medical Genetics, Ghent University Hospital, Ghent, Belgium.*
[8]*Institute for Cancer Genetics, Columbia University, New York, NY, USA.*
[9]*INSERM U944, Institut Universitaire d'Hématologie Université Paris 7, Hôpital Saint Louis, Paris, France.*
[10]*Department of Clinical Genetics, Erasmus MC, Rotterdam, The Netherlands.*
[11]*Research Institute Children's Cancer Center Hamburg, Germany.*
[12]*Clinic of Pediatric Hematology and Oncology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany.*
[13]*Co-operative study group for childhood acute lymphoblastic leukemia (COALL), Hamburg, Germany.*
[14]*Dutch Childhood Oncology Group (DCOG), The Hague, The Netherlands.*
[15]*Department of Immunohematology and Blood Transfusion, Leiden University Medical Center, Leiden, The Netherlands.*
[16]*These authors contributed equally to this work.*

*Cancer Cell, 2011; 19: 484-497*

**ABSTRACT**

To identify novel oncogenic pathways in T-cell acute lymphoblastic leukemia (T-ALL), we combined expression profiling of 117 pediatric patient samples and detailed molecular cytogenetic analyses including the Chromosome Conformation Capture on Chip (4C) method. Two T-ALL subtypes were identified that lacked rearrangements of known oncogenes. One subtype associated with cortical arrest, expression of cell cycle genes and ectopic *NKX2-1* or *NKX2-2* expression for which rearrangements were identified. The second subtype associated with immature T-cell development and high expression of the *MEF2C* transcription factor as consequence of rearrangements of *MEF2C*, transcription factors that target *MEF2C* or MEF2C-associated cofactors. We propose *NKX2-1*, *NKX2-2* and *MEF2C* as T-ALL oncogenes that are activated by various rearrangements.

**SIGNIFICANCE**

For 40% of pediatric T-ALL cases, underlying oncogenic rearrangements remain unresolved. By combined expression profiling and molecular-cytogenetic techniques, we revealed 2 T-ALL entities lacking known oncogenic rearrangements and representing ~20% of pediatric T-ALL cases. One subtype associated with cortical thymocytic arrest and 10 out of 12 cases ectopically expressed *NKX2-1/NKX2-2* for which 5 rearrangement variants were identified in 7 cases. The second subtype was associated with high *MEF2C* expression (11 out of 12 cases), and rearrangements involving *MEF2C* or transcription factors and transcription cofactors that directly target *MEF2C* were identified in 6 cases. Ectopic expression of NKX2-1 or MEF2C was able to transform cells and interfered with T-cell differentiation. We propose that *NKX2-1*, *NKX2-2* and *MEF2C* are oncogenes in leukemia.

## INTRODUCTION

T-lineage acute lymphoblastic leukemia (T-ALL) is a malignancy of thymocytes. T-ALL represents about 15% of pediatric ALL cases but has an inferior outcome compared to B-ALL as approximately 30% of T-ALL cases relapse during therapy or within the first 2 years following treatment and eventually die (Pieters and Carroll, 2008; Pui and Evans, 2006). T-ALL is mostly characterized by genetic abnormalities that are crucial for T-cell pathogenesis (Van Vlierberghe et al., 2008a). Various genetic rearrangements in T-ALL occur in a mutually exclusive pattern (Van Vlierberghe et al., 2008a) in contrast to frequent *CDKN2A/ARF* deletions (Hebert et al., 1994) or *NOTCH1* activating mutations (Weng et al., 2004). These mutually exclusive rearrangements are considered as driving chromosomal abnormalities that affect the *TAL1*, *LMO2*, *TLX1*, *TLX3*, *MYB* or *HOXA* oncogenes (Van Vlierberghe et al., 2008a). Based on gene expression data (Ferrando et al., 2002; Soulier et al., 2005; Van Vlierberghe et al., 2008b), these oncogenes have been associated with distinct T-ALL subgroups denoted as the *TAL/LMO*, *TLX1*, *TLX3* and the *HOXA* subgroups. Initial profiling data also pointed to the existence of an additional immature T-ALL subgroup (Soulier et al., 2005). This entity probably corresponds to the *LYL1* T-ALL subgroup as previously defined (Ferrando et al., 2002) and to the recently described immature T-ALL subset that is characterized by an early T-cell precursor (ETP) profile and inferior outcome (Coustan-Smith et al., 2009). For approximately 40 percent of all T-ALL patients including the immature T-ALL entity, the driving chromosomal aberrations have thus far remained elusive.

## RESULTS

### Cluster analyses predict new T-ALL genetic subgroups

To identify driving oncogenic mechanisms in T-ALL, we performed unsupervised hierarchical cluster analyses based on microarray expression data of 117 diagnostic pediatric T-ALL samples and 7 normal bone marrow controls. Seventy-seven T-ALL samples were characterized by oncogenic rearrangements, including *TAL1* (n=24), *TAL2* (n=1), *LMO1* (n=1), *LMO1/TAL2* (n=1), *LMO2* (n=9), *TLX3* (n=22), *TLX1* (n=7), *HOXA*-activating rearrangements (including *CALM-AF10*, Inv(7)(p15q34), *SET-NUP214*; n=10) or *MYB* translocations (n=2). No such abnormalities were identified in the remaining 40 T-ALL patient samples. Four robust T-ALL clusters were observed in unsupervised cluster analysis irrespective of the number of genes included or the data normalization methods chosen

(**Figures 1A and S1, Tables S1-S3**). The association with clinical and molecular-cytogenetic data, immunophenotypic markers and expression of *TAL1* and *LYL1* for these 4 subgroups is given in **Table 1 and Figure 1A**.

**Table 1. Clinical and biological characteristics of unsupervised T-ALL clusters .**

| | **Cohort** | | **TAL/LMO** | | **TLX** | | **Proliferative** | | **Immature** | | **p-value** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Total, *n*** | 117 | | 53 | | 30 | | 19 | | 15 | | |
| **Clinical** | | | | | | | | | | | |
| Gender (n, %) | | | | | | | | | | | |
| Male | 83 | *71%* | 40 | *75%* | 20 | *67%* | 15 | *79%* | 8 | *53%* | |
| female | 34 | *29%* | 13 | *25%* | 10 | *33%* | 4 | *21%* | 7 | *47%* | p=0.306 |
| Age at Dx, years | | | | | | | | | | | |
| Median | 7.8 | | 9.3 | | 7.7 | | 5.5 | | 10.1 | | |
| Range | 1.5-17.8 | | 1.6-16.7 | | 3.2-17.8 | | 1.5-16.7 | | 3.1-16.4 | | p=0.404‡ |
| WBC, $10x10^{E9}$/L | | | | | | | | | | | |
| Median | 115.1 | | 156.9 | | 121.9 | | 64.3 | | 87.6 | | |
| Range | 1.8-900 | | 16.1-900 | | 1.8-417 | | 27.2-192 | | 2.3-435 | | p=0.001‡ |
| **Immunophenotype** | | | | | | | | | | | |
| CD34 (n, %) | 111 | | | | | | | | | | |
| Negative | 77 | *69%* | 37 | *71%* | 19 | *66%* | 14 | *93%* | 7 | *47%* | |
| Positive | 34 | *31%* | 15 | *29%* | 10 | *34%* | 1 | *7%* | 8 | *53%* | p=0.046 |
| CD13/33 (n, %) | 110 | | | | | | | | | | |
| Negative | 92 | *84%* | 47 | *92%* | 19 | *73%* | 17 | *94%* | 9 | *60%* | |
| Positive | 18 | *16%* | 4 | *8%* | 7 | *27%* | 1 | *6%* | 6 | *40%* | p=0.006 |
| CD1 (n, %) | 113 | | | | | | | | | | |
| Negative | 62 | *55%* | 33 | *65%* | 14 | *47%* | 3 | *17%* | 12 | *86%* | |
| Positive | 51 | *45%* | 18 | *35%* | 16 | *53%* | 15 | *83%* | 2 | *14%* | p<0.001 |
| CD4 (n, %) | 115 | | | | | | | | | | |
| Negative | 42 | *37%* | 22 | *42%* | 4 | *13%* | 3 | *17%* | 13 | *87%* | |
| Positive | 73 | *63%* | 30 | *58%* | 26 | *87%* | 15 | *83%* | 2 | *13%* | p<0.001 |
| CD8 (n, %) | 115 | | | | | | | | | | |
| Negative | 45 | *39%* | 14 | *27%* | 16 | *53%* | 2 | *11%* | 13 | *87%* | |
| Positive | 70 | *61%* | 38 | *73%* | 14 | *47%* | 16 | *89%* | 2 | *13%* | p<0.001 |
| CD4/8 (n, %) | 115 | | | | | | | | | | |
| Negative | 61 | *53%* | 26 | *50%* | 17 | *57%* | 5 | *28%* | 13 | *87%* | |
| Positive | 54 | *47%* | 26 | *50%* | 13 | *43%* | 13 | *72%* | 2 | *13%* | p=0.008 |
| CD3 (n, %) | 114 | | | | | | | | | | |
| Negative | 59 | *52%* | 21 | *40%* | 19 | *63%* | 10 | *59%* | 9 | *60%* | |
| Positive | 55 | *48%* | 31 | *60%* | 11 | *37%* | 7 | *41%* | 6 | *40%* | p=0.169 |
| **Oncogenes** | | | | | | | | | | | |
| *TAL1*, % expression of GAPDH x $10^{E-2}$ | | | | | | | | | | | |
| Median | 3.1 | | 13 | | 0.73 | | 1.4 | | 1.14 | | |
| Range | 0.09-1820 | | 0.75-1820 | | 0.088-11 | | 0.17-22 | | 0.10-14 | | p<0.001‡ |
| *LYL1*, % expression of GAPDH x $10^{E-4}$ | | | | | | | | | | | |
| Median | 1.7 | | 1.3 | | 3.1 | | 3.5 | | 8.5 | | |
| Range | 0-126 | | 0-32 | | 0-16.6 | | 0.28-15.3 | | 0.96-126 | | p=0.001‡ |

The p-values are calculated according to the chi-square test or ‡ the Mann-Whitney-U test. See also **Figure S2**.

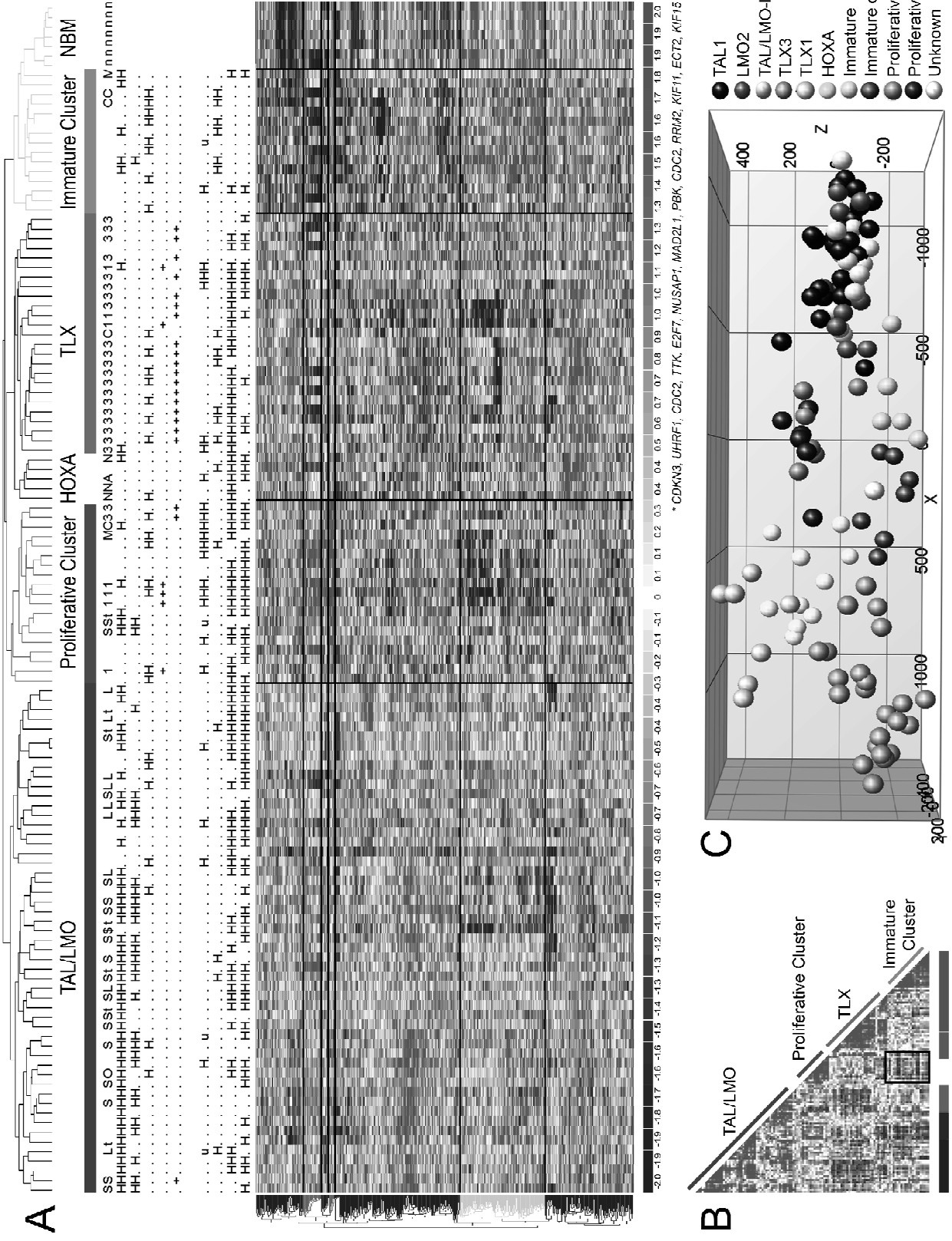

A
TAL/LMO
Proliferative Cluster
HOXA
TLX
Immature Cluster
NBM
Cytogenetics
TAL1_50
TAL1_25
LYL1_25
TLX1
TLX3
CD1
CD33/13
CD4
CD8
Cell Cycle*
CDKs
LMO4
JUP
HHEX
HOP
CCND2
*CDKN3, UHRF1, CDC2, TTK, E2F7, NUSAP1, MAD2L1, PBK, CDC2, RRM2, KIF11, ECT2, KIF15
B
TAL/LMO
Proliferative Cluster
TLX
Immature Cluster
C
TAL1
LMO2
TAL/LMO-like
TLX3
TLX1
HOXA
Immature
Immature core
Proliferative
Proliferative core
Unknown
X
Z

**Figure 1. (page 31) Identification of 2 Entities in Pediatric T-ALL That Lack Known Driving Oncogenic Hits.** (**A**) Unsupervised hierarchical cluster analysis by the average linkage method in dCHIP based on 435 probesets (**Table S3**) for RMA-solo (Soulier et al., 2005) normalized U133 plus 2 Affymetrix data from 117 pediatric T-ALL samples and 7 normal bone marrow controls. Cytogenetic rearrangements indicated are: S, *SIL-TAL1*; T, *TAL1*; t, *TAL2*; O, *LMO1*; L, *LMO2* (includes del(11)(p12p13)); $, *TAL2*/*LMO1*; N, *SET-NUP214*; C, *CALM-AF10*; M, *MYB*; A, Inv(7)(p15q34); 1, *TLX1*; 3, *TLX3*; n, normal bone marrow controls. The 50$^{th}$ and/or the 25$^{th}$ percentiles of samples with the highest *TAL1* or *LYL1* expression, positivity for *TLX1* and *TLX3* expression as measured by RQ-PCR, and expression of the immunophenotypic markers CD13 and/or CD33, CD4 or CD8 are indicate; u, no data available. (**B**) Pearson correlation plot for the patient samples belonging to the 4 unsupervised *TAL/LMO*, *TLX*, proliferative and immature clusters. (**C**) Principal component analysis of pediatric T-ALL patients based upon the top 100 most significant differentially expressed probesets among major T-ALL subgroups (i.e. *TAL1*/*LMO2*, *HOXA*, *TLX1*, and *TLX3* (**Table S3**)). The immature cluster (12 cases) and the proliferative cluster (12 cases) are indicated by green and purple dots, respectively. Samples repeatedly assigned to the proliferative or immature clusters (i.e. the core samples) in multiple unsupervised analyses on RMA-solo (**Figure 1A**), RMA or VSN normalized datasets (not shown) or the supervised cluster analysis (**Figure 1C**) are visualized by dark green or purple dots. See also **Figure S1** and **Tables S1-S4**.

Two clusters represented established T-ALL genetic subgroups (Ferrando et al., 2002; Soulier et al., 2005; Van Vlierberghe et al., 2008b), corresponding to abnormalities of *TAL1/LMO2*, and *TLX3*/*HOXA* transcription factors.

A third cluster included cases that highly expressed *CD1* genes. This corresponded with a CD1a positive immunophenotype for most cases of this cluster ($p<0.001$; **Table 1**), which validated our gene expression data. This cluster also comprised most *TLX1* translocated cases, a genetic entity that was previously associated with CD1-positivity and cortical developmental arrest (Ferrando et al., 2002), and that may share a similar biology with the other samples present in this cluster. In the unsupervised cluster analysis, this cluster is characterized by expression of genes that are involved in cell cycle regulation (*CDKN3*), G1/S transition (*UHRF1*, *CDC2*), cell cycle progression (*TTK*, *E2F7*, *CDC2*), DNA replication and chromosome condensation (*TOP2A*), the spindle-assembly checkpoint (*NUSAP1*, *MAD2L1*, *KIF15*, *KIF11*), the G2/M checkpoint (*PBK*) and genes whose expression are linked to cell cycle (*RRM2*, *ECT2*). Furthermore, differentially expressed genes for this cluster compared to all other T-ALL cases as identified by t-statistics were enriched for genes that are strongly associated with

the cell cycle pathway and spindle assembly (**Table S4**) and this cluster strongly expressed the proliferation marker *MKI67*. This cluster was accordingly denoted as "proliferative cluster". Most of the cases in this cluster lacked currently known driving mutations, which may point towards involvement of new T-ALL oncogenes. This was further supported by the fact that most of these unknown samples clustered as a separate entity (12 cases) distinct from established T-ALL genetic subgroups including the *TLX1*-rearranged cases in a supervised cluster analysis (**Figure 1C**).

The fourth cluster was enriched for immunophenotypic immature CD4/CD8 double negative cases (p=0.008, **Table 1**), and was named the "immature cluster" by reference to previous work (Soulier et al., 2005). Samples in this cluster frequently expressed myeloid markers CD13 and/or CD33 (p=0.006), and were characterized by expression of genes associated with protein binding, protein dimerization and TGFBR1 signal transduction. They expressed low levels of genes associated with cellular proliferation contrary to samples of the proliferative cluster (**Figure 1B**). This cluster comprised 3 *HOXA* activated cases with an immature immunophenotype unlike other *HOXA* activated cases that usually have a more advanced immunophenotype. Other samples in this immature cluster were devoid of known driving mutations. This cluster may comprise a second molecular-cytogenetic T-ALL entity for which driving oncogenes are unknown, and in support of this notion, most of these samples appeared as a separate subgroup (12 cases) in the supervised principal component analysis (PCA) based on differentially expressed genes among the known 4 T-ALL genetic subgroups (**Figure 1C**). This immature cluster largely overlaps with the *LYL1* positive cluster as described earlier (Ferrando et al., 2002) as it expressed the highest *LYL1* levels (**Table 1**). Our immature cluster was highly enriched for early T-cell precursor (ETP) T-ALL cases as previously described (Coustan-Smith et al., 2009), as 13 out of 15 immature cases in contrast to only 3 out of 102 remaining cases were predicted as ETPs in PAM analysis based on the 62 probeset profile that defined the ETP group (p<0.001, not shown). In contrast to that study (Coustan-Smith et al., 2009), the overall survival for immature cases in our cohort was not extremely poor (5yr OS = 73±11%), but seemed equally low to the outcome of *TAL/LMO* and *TLX* subgroups (5yr OS = 65±6%). The proliferative subgroup seemed to have an improved outcome (5yr OS = 88±8%), albeit not significant (p=0.096, **Figure S2**).

We then searched for candidate genes that participate in oncogenic chromosomal abnormalities using several methods including COPA (Tomlins et al., 2005), SAM (Tusher et al., 2001) and PAM statistics (Tibshirani et al., 2002). Both COPA and PAM analyses identified *NKX2-1* and *MEF2C* as characteristic genes for the proliferative and immature clusters, respectively (**Table S5**). The *NKX2-1* homologous *NKX2-2* gene was also identified by COPA as outlier gene for the proliferative cluster. High microarray expression levels of *NKX2-1* and *MEF2C* were validated by RQ-PCR (**Figure 2**) for the proliferative and immature cluster cases, respectively, that lack known oncogenic rearrangements. These cases form separate clusters in the supervised analysis (**Figure 1C**). *NKX2-1* or *MEF2C* were either absent or expressed at relative low levels in most cases belonging to other supervised clusters. However, some *TLX1* positive patient samples that are part of the proliferative cluster in the unsupervised analyses express *NKX2-1*. Also, the *CALM-AF10* positive *HOXA*-activated patient sample #1509 that highly expresses *MEF2C* has an immature phenotype and co-clusters in the immature cluster in unsupervised analyses.

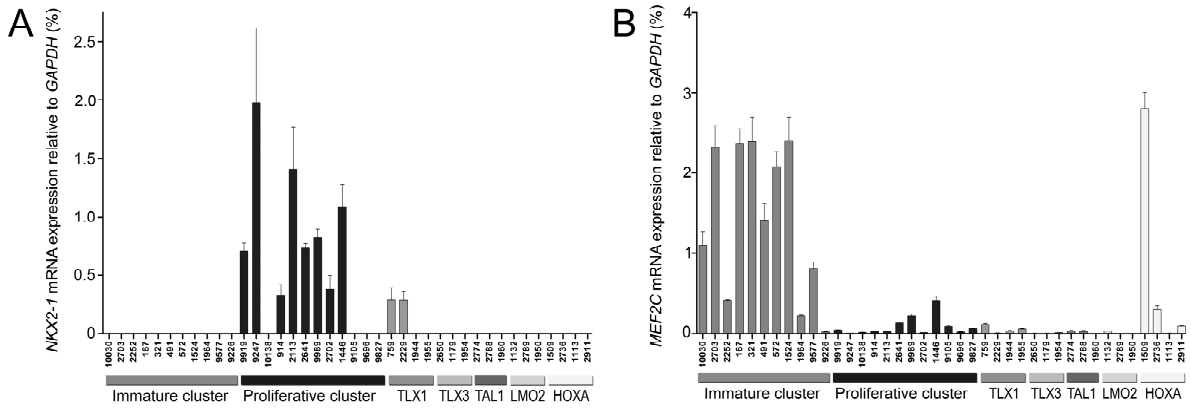



**Figure 2. Validation of Elevated *NKX2-1* and *MEF2C* Levels in Patients from Proliferative and Immature Supervised Clusters.** Relative expression levels of (**A**) *NKX2-1* or (**B**) *MEF2C* is determined by RQ-PCR. *NKX2-1* and *MEF2C* expression levels are indicated for 11 out of 12 immature cluster patient samples (green) and 12 proliferative cluster samples (purple) according to the supervised analysis (**Figure 1C**) compared to cases of other T-ALL molecular-cytogenetic subgroups. The SEM are shown. See also **Table S5**.

## Molecular-cytogenetic identification of *NKX2-1* rearrangements

These data formed the start of detailed molecular-cytogenetic analyses on the 12 immature cluster and the 12 proliferative cluster samples that seemed to form 2 genetic T-ALL entities (**Figure 1C**), and for which driving oncogenic hits were

unknown. We used a variety of molecular cytogenetic techniques including FISH, array-comparative genomic hybridization (array-CGH) and Chromosome Conformation Capture on Chip (4C) (Simonis et al., 2009) to identify potential deletions, amplifications, and T-cell receptor- or *BCL11B*-driven oncogenic events (**Table 2 and Table S6**). The 4C method was originally developed to study the three-dimensional structure of DNA (Simonis et al., 2006), but it was recently shown that it robustly identifies chromosomal rearrangements, in particular inversions and translocations, even when they are balanced (Simonis et al., 2009). In the proliferative cluster, 2 out of 12 samples were characterized by *MYB* translocations, a rearrangement considered as a driving oncogenic hit (Clappier et al., 2007). No further *MYB* translocations were identified in the remaining 10 cases by FISH (**Table S6**). We identified 5 rearrangements of *NKX2-1* or *NKX2-2* genes in 7 out of 12 patient samples that were not observed before in human cancer (**Table 2**, **Figures 3A-E**, and **Figure S3**). The *NKX2-1* gene was inverted to the T-cell receptor gene *TRA@* in 2 cases (#1446, #9247), inverted to the immunoglobulin heavy chain gene *IGH@* in 1 case (#9919) and translocated to the *TRB@* locus (t(7;14)(q34;q13)) in 1 other case (#9989) as identified by 4C analyses (**Figure 3A).** *NKX2-1* rearrangements in these patients were validated by FISH **(Figure 3B**). The der(7) chromosomal breakpoint for this t(7;14)(q34;q13) in patient #9919 was cloned (**Figure 3C**). A fifth patient (#2641) contained a *NKX2-1* rearrangement based on FISH results (**Table 2** and data not shown), whereas a sixth patient (#2702) had an amplification at 14q13 based on array-CGH (**Figure 3D**) presumable due to a *NKX2-1* duplication or an insertion into another chromosome (not shown). These patients highly expressed NKX2-1 protein levels (**Figure 3F**, representative cases are shown). A seventh case (#10138) had a translocation between the homologous *NKX2-2* gene and the *TRD@* locus, for which both reciprocal breakpoint regions were cloned (**Figure 3E**). This patient highly expressed NKX2-2 protein levels (not shown). For the *TLX1* rearranged cases that co-cluster with these *NKX2-1/NKX2-2* rearranged cases in unsupervised cluster analysis that also expressed *NKX2-1* (**Figure 2A**), we did not find evidence for *NKX2-1* rearrangements by FISH (data not shown). This indicates that *TLX1* and *NKX2-1/NKX2-2* oncogenes may exert identical or closely related pathogenic mechanisms.

**Table 2. Identified rearrangements in patient samples of the proliferative and immature clusters.**

| **Proliferative cluster** | | | | | |
|---|---|---|---|---|---|
| **Patient** | ***NKX2-1* expression¥** | **Aberration** | ***Partner gene 1*** | ***Partner gene 2*** | **Methods** |
| **#9919§** | + | inv(14)(q13q32.33) | *IGH@* | *NKX2-1* | FISH, 4C |
| **#9247§** | + | inv(14)(q11.2q13) | *TRA@* | *NKX2-1* | FISH, 4C |
| **#10138§** | +* | t(14;20)(q11;p11) | *TRD@* | *NKX2-2* | FISH, LM-PCR |
| **#914** | + | t(6;7)(q22-23;q34) | *TRB@* | *MYB* | FISH |
| **#2113** | + | - | - | - | - |
| **#2641** | + | rearrangement | *?* | *NKX2-1* | FISH |
| **#9989§** | + | t(7;14)(q34;q13) | *TRB@* | *NKX2-1* | FISH, 4C |
| **#2702§** | + | dup(14)(q13.3q13.3) or ins(?)(?q13.3) | *?* | *NKX2-1* | Array-CGH, FISH |
| **#1446§** | + | inv(14)(q11.2q13) | *TRA@* | *NKX2-1* | FISH, 4C |
| **#9105** | + | t(6;7)(q22-23;q34) | *TRB@* | *MYB* | FISH |
| **#9696** | | - | - | - | - |
| **#9827** | | - | - | - | - |
| **Immature cluster** | | | | | |
| **Patient** | ***MEF2C* expression¥** | **Aberration** | ***Partner gene 1*** | ***Partner gene 2*** | **Methods** |
| **#10030$** | + | - | - | - | - |
| **#2703** | + | - | - | - | - |
| **#2130** | | - | - | - | - |
| **#2252** | + | t(11;14)(p11.2;q32.2) | *BCL11B* | *SPI.1* | FISH, 4C |
| **#167$** | + | - | - | - | - |
| **#321$** | + | - | - | - | - |
| **#491$** | + | del(5)(q14) | - | *MEF2C* | FISH, array-CGH |
| **#572$** | + | t(2;21)(q11.2-12;q22.3) | *RUNX1* | *AFF3* | Karyotype, 3'-RACE |
| **#1524$** | + | t(8;12)(q13;p13) | *ETV6* | *NCOA2* | RT-PCR, FISH |
| **#1964$** | + | der(5)t(4;5)(q26;q14) | 4q26 | *MEF2C* | 4C, array-CGH |
| **#9577** | + | t(5;14)(q34;q32.2) | *BCL11B* | *NKX2-5* | FISH |
| **#9226** | ± | - | - | - | - |
| **Cell lines** | | | | | |
| **LOUCY** | + | t(5;14)(q34;q32.2) | *BCL11B* | *NKX2-5* | (Przybylski et al., 2006) |
| **PEER** | + | del(5)(q14) | - | *MEF2C* | (Nagel et al., 2008) |

¥NKX2-1 or MEF2C expression based on expression array and/or RQ-PCR results. *Sample #10138 expresses the NKX2-1 homologous NKX2-2 gene; $Core immature or §core proliferative cases repeatedly assigned in unsupervised and supervised analyses to the immature or proliferative clusters, respectively. See also Table S6 and Figure S6.

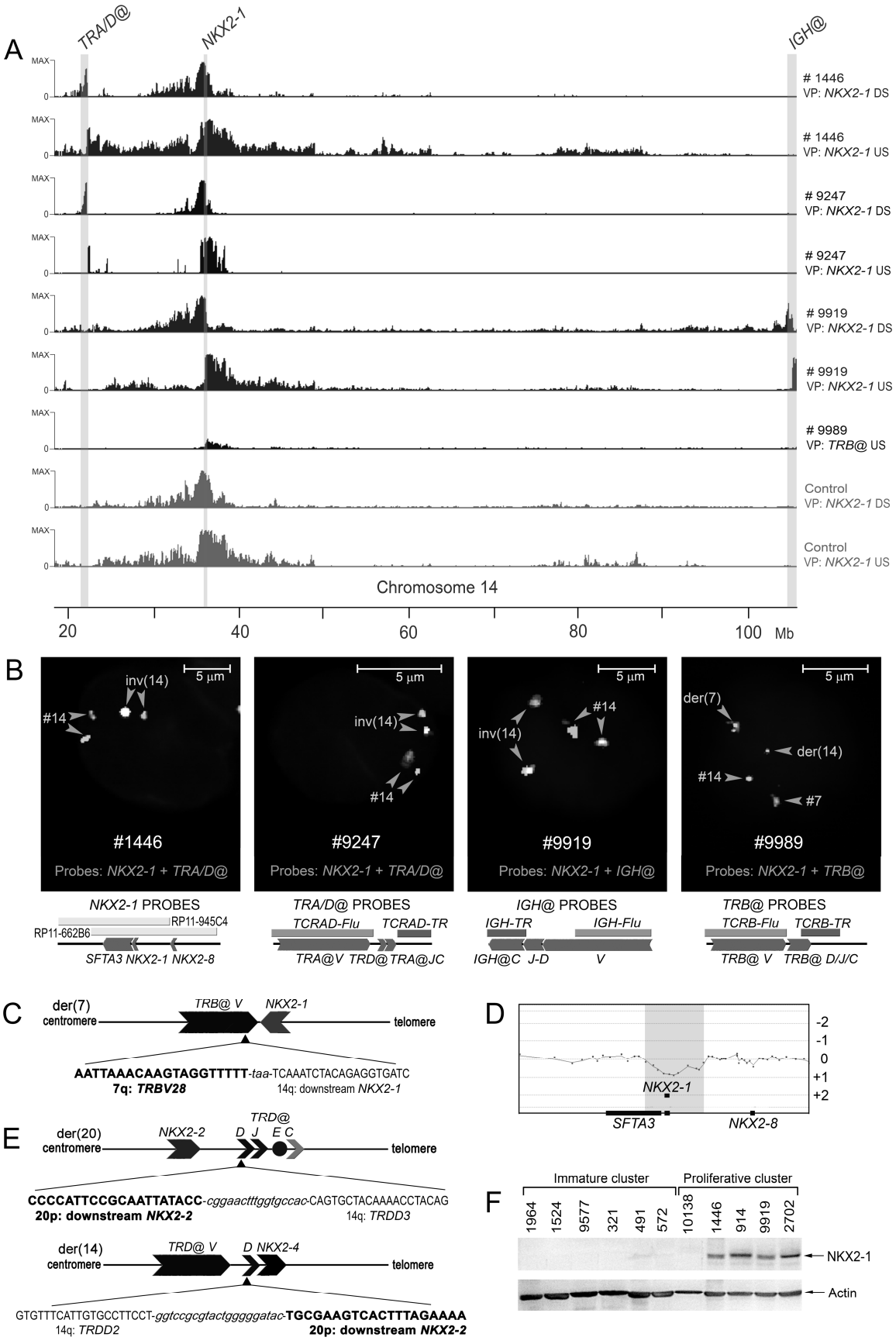

A
TRA/D@
NKX2-1
IGH@
# 1446 VP: NKX2-1 DS
# 1446 VP: NKX2-1 US
# 9247 VP: NKX2-1 DS
# 9247 VP: NKX2-1 US
# 9919 VP: NKX2-1 DS
# 9919 VP: NKX2-1 US
# 9989 VP: TRB@ US
Control VP: NKX2-1 DS
Control VP: NKX2-1 US
Chromosome 14
20
40
60
80
100 Mb
B
inv(14)
#14
#1446
Probes: NKX2-1 + TRA/D@
#9247
Probes: NKX2-1 + TRA/D@
#9919
Probes: NKX2-1 + IGH@
der(7)
der(14)
#7
#9989
Probes: NKX2-1 + TRB@
5 μm
NKX2-1 PROBES
RP11-945C4
RP11-662B6
SFTA3 NKX2-1 NKX2-8
TRA/D@ PROBES
TCRAD-Flu
TCRAD-TR
TRA@V TRD@ TRA@JC
IGH@ PROBES
IGH-TR
IGH-Flu
IGH@C J-D V
TRB@ PROBES
TCRB-Flu
TCRB-TR
TRB@ V TRB@ D/J/C
C
der(7)
centromere
TRB@ V
NKX2-1
telomere
AATTAAACAAGTAGGTTTTT-taa-TCAAATCTACAGAGGTGATC
7q: TRBV28
14q: downstream NKX2-1
D
-2
-1
0
+1
+2
NKX2-1
SFTA3
NKX2-8
E
der(20)
centromere
NKX2-2
TRD@
D J E C
telomere
CCCCATTCCGCAATTATACC-cggaactttggtgccac-CAGTGCTACAAAACCTACAG
20p: downstream NKX2-2
14q: TRDD3
der(14)
centromere
TRD@ V
D NKX2-4
telomere
GTGTTTCATTGTGCCTTCCT-ggtccgcgtactgggggatac-TGCGAAGTCACTTTAGAAAA
14q: TRDD2
20p: downstream NKX2-2
F
Immature cluster
Proliferative cluster
1964
1524
9577
321
491
572
10138
1446
914
9919
2702
NKX2-1
Actin

**Figure 3. (page 37) *NKX2-1* and *NKX2-2* Rearrangements in Proliferative Cluster Patient Samples.** (**A**) 4C-results obtained from *NKX2-1* or *TRB@* viewpoints (VP). Position of *TRA@*, *NKX2-1* and *IGH@* loci are shown by grey vertical bars. 4C-results for a normal control are shown in grey. Higher magnifications of the reciprocal breakpoint regions are given in **Figure S3**. (**B**) Validation of *NKX2-1* rearrangements by FISH. Schematic positions of FISH probes are shown. (**C**) Schematic representation of the der(7) breakpoint region and breakpoint sequence of the unbalanced t(7;14)(q34;q13) for patient #9989. (**D**) Visualization of a single copy *NKX2-1* amplification (green box) in patient #2702 as identified by array-CGH. (**E**) Schematic representation of t(14;20)(q11;20p11) breakpoint regions and cloned breakpoint sequences for the *NKX2-2* rearranged patient #10138. (**F**) NKX2-1 protein expression in representative proliferative cluster and immature cluster patient samples as shown by western blot. Actin was used as loading control.

## Molecular-cytogenetic identification of *MEF2C* and *MEF2*C-activating rearrangements

We subsequently investigated the 12 immature cluster cases lacking known driving oncogenic hits, and identified chromosomal abnormalities that converge on the activation of the *MEF2C* gene in at least 5 cases. Two cases had chromosomal copy number loss of the 5q14-qter chromosomal arm with breakpoint in a 0.5-2 Mb proximity telomeric of *MEF2C*. A similar deletion was also identified in T-ALL cell line LOUCY (**Figure 4A**). These 5q14-qter deletions were not identified in 90 other T-ALL cases as included in our profiling study for which array-CGH data were available (**Figure S4A**). For patient #1964, this 5q14-qter deletion was part of an unbalanced chromosomal translocation between chromosomal bands 5q14 and 4q27 fusing the telomeric *MEF2C* region to the telomeric region ~0.6 Mb distal of the *PITX2* gene on chromosome 4 (**Figure 4B**). In contrast to other genes in the 5q14 region, *MEF2C* is highly upregulated in both patients indicating that *MEF2C* represents the target of these 5q rearrangements (**Figures S4B-C**).

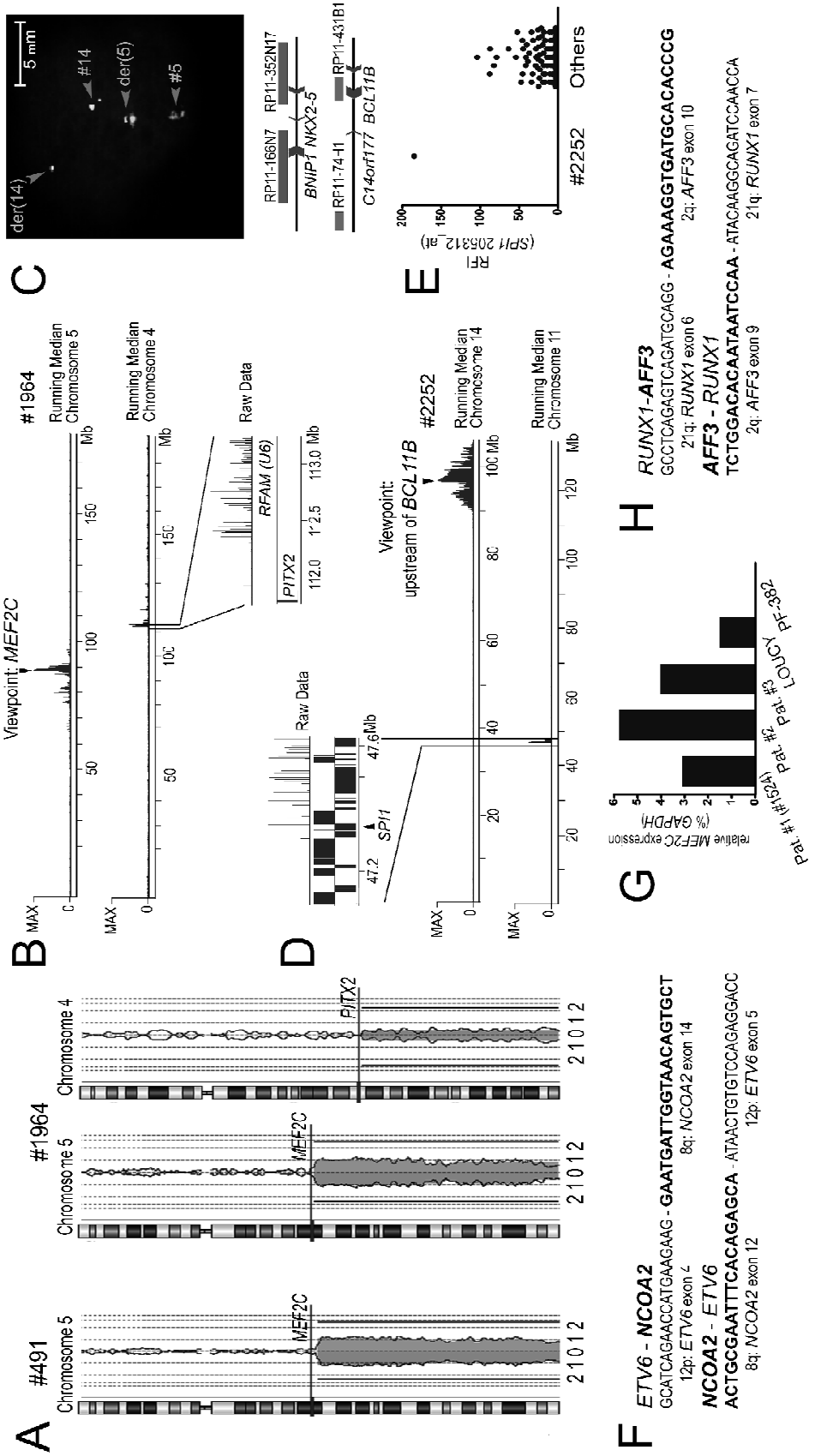
A *NKX2-5/BCL11B* translocation was identified by FISH in a third case (**Figure 4C**), and this case highly expressed *NKX2-5*. This rare translocation has been reported in T-ALL before (Nagel et al., 2003). Knockdown of NKX2-5 levels by siRNA molecules in the *NKX2-5* translocation positive cell line PEER lowered *MEF2C* levels (**Figures 5A-C**), indicating that NKX2-5 controls *MEF2C*. Chromatin immunoprecipitation (ChIP) experiments confirmed that NKX2-5 directly binds in the promoter region of *MEF2C* (**Figure 5D**).

A fourth case harbored a *BCL11B* translocation to *SPI1*, which encodes for PU.1 (**Figure 4D**). This patient uniquely expressed *SPI1* compared to the other T-ALL cases in this study (**Figure 4E**). PU.1 was recently identified as important regulator for *MEF2C* expression in normal lymphoid development (Stehling-Sun et al., 2009), and this T-ALL patient highly expressed *MEF2C* (**Figure 2B, Table 2**).

A fifth case (#1524) harbored a t(8;12)(q13;p13) as identified by FISH (**Figures S4D-E)** resulting in reciprocal *ETV6-NCOA2* fusion products, and both reciprocal breakpoints were cloned for this patient (**Figure 4F**). Similar fusions were recently identified in biphenotypic T-ALL (Strehl et al., 2008). NCOA2 is a known co-regulator of MEF2C (Chen et al., 2000), and *MEF2C* was found consistently upregulated in selected *ETV6–NCOA2* rearranged cases (**Figure 4G**).

A sixth immature case with high *MEF2C* levels had a karyotypic t(2;21) that involved the *RUNX1/AML1* gene (**Figure S4F**). For this patient, we cloned reciprocal in-frame *RUNX1-AFF3* and *AFF3-RUNX1* fusion products as consequence of this translocation (**Figure 4H**). How RUNX1 fusion products could upregulate *MEF2C* expression remains to be determined.

**Figure 4. (page 40) *MEF2C* Activating Rearrangements for Immature Cluster Samples.** (**A**) Array-CGH results for chromosomes 4 and/or 5 for patients #491 and #1964. Blue and red tracings represent dye swopped experiments. Positions of *MEF2C* and *PITX2* have been indicated. (**B**) Visualization of an unbalanced chromosomal translocation t(4;5)(q26;q14) for patient #1964 by 4C-analysis. The *MEF2C* VP is indicated by an arrow. Running median of probeset intensities for chromosome 5 and 4 are indicated in red and blue, respectively. (**C**) Validation of a chromosomal translocation between *NKX2-5* and *BCL11B* in patient #9577 by FISH. Schematic positions of FISH probes are shown. (**D**) Identification of the t(11;14)(p11.2;q32.2) chromosomal translocation between *SPI1* and *BCL11B* in patient #2252 by 4C. The VP is positioned ~0.6 Mb upstream of *BCL11*B as indicated by an arrow. (**E**) Ectopic *SPI1* expression in patient #2252 compared to 116 additional T-ALL patient samples. Raw fluorescent intensities of probeset 205312_at are shown. (**F**) Cloned fusion areas for reciprocal *ETV6-NCOA2* and *NCOA2-ETV6* fusion transcripts in patient #1524. (**G**) Relative *MEF2C* expression by RQ-PCR in 3 selected *ETV6-NCOA2* rearranged T-ALL patients (Pat. #1-3). Cell lines LOUCY and PF382 are positive and negative controls for *MEF2C* expression, respectively. (**H**) Cloned fusion areas for reciprocal *RUNX1-AFF3* and *AFF3-RUNX1* fusion transcripts for patient #572. See also **Figure S4**.


A
#491
Chromosome 5
MEF2C
21012
#1964
Chromosome 5
MEF2C
21012
Chromosome 4
PITX2
21012
B
#1964
Viewpoint: MEF2C
Running Median Chromosome 5
Running Median Chromosome 4
Raw Data
RFAM (U6)
PITX2
112.0
113.0 Mb
MAX
Mb
C
der(14)
5 mm
#14
der(5)
#5
RP11-166N7
RP11-352N17
BNIP1
NKX2-5
RP11-741-1
RP11-431B1
C14orf177
BCL11B
D
Raw Data
47.6Mb
47.2
SPI1
Viewpoint: upstream of BCL11B
#2252
Running Median Chromosome 14
Running Median Chromosome 11
E
RFI
(SPI1 205312_at)
#2252
Others
F
ETV6 - NCOA2
GCATCAGAACCATGAAGAAG - GAATGATTGGTAACAGTGCT
12p: ETV6 exon 4
8q: NCOA2 exon 14
NCOA2 - ETV6
ACTGCGAATTTCACAGAGCA - ATAACTGTGTCCAGAGGACC
8q: NCOA2 exon 12
12p: ETV6 exon 5
G
relative MEF2C expression (% GAPDH)
Pat. #1(#152A)
Pat. #2
Pat. #3
LOUCY
PF-382
H
RUNX1-AFF3
GCCTCAGACTCAGATGCAGG - AGAAAGGTGATGCACACCCG
21q: RUNX1 exon 6
2q: AFF3 exon 10
AFF3 - RUNX1
TCTGGACACAATAATCCAA - ATACAAGGCAGATCCAACCA
2q: AFF3 exon 9
21q: RUNX1 exon 7

To investigate whether MEF2C could indeed regulate the expression of various genes from the immature signature, *MEF2C* stable transfected clones and mock transfected controls were generated for the cell line Jurkat (**Figure 5E**) that does not have an immature signature (not shown). As shown in **Figure 5F**, the *MEF2C* transfected Jurkat clone 2B3 but not the mock transfected control 3G4 highly activates 5 out of 6 selected immature signature genes (*PSCD4*, *HHEX*, *FAM46A*, *LMO2* and *LYL1*), indicating that MEF2C may function as a transcriptional regulator for many genes that are highly expressed in immature T-ALL cases. For the reciprocal setting in cell line PEER, knock-down of NKX2-5 using siRNA molecules that reduced *MEF2C* expression (**Figures 5A-C**) also led to reduced levels of *LMO2*, *LYL1* and *HHEX* (**Figure 5G**). Oncogenic rearrangements of *LMO2*, *LYL1* and the *LYL1* homologous *TAL1* gene are exclusively found in the *TAL/LMO* subgroup but have never been observed in immature T-ALL cases (this work and (Ferrando et al., 2002)). Activation of *LMO2* and *LYL1* through MEF2C may be crucial to prime early committed T-cells for leukemogenesis. By using ChIP, we demonstrated that MEF2C directly binds to the promoter of *HHEX* as well as to the distal and proximal promoters of *LMO2* in the immature cell line LOUCY. This could also be demonstrated for diagnostic leukemic cells of 3 immature cluster patients (#491, #321 and #167, not shown) but not in the control cell line Jurkat (**Figure 5H**). The *MN1* gene, which is targeted by chromosomal alterations in inv(16) M4EO AML subtype (Buijs et al., 2000; Grosveld, 2007) was also identified as a highly activated gene for the immature cluster (**Table S5**). As for *HHEX*, we did not find evidence for chromosomal rearrangements of *MN1* by FISH in immature T-ALL cases (**Table S6**).

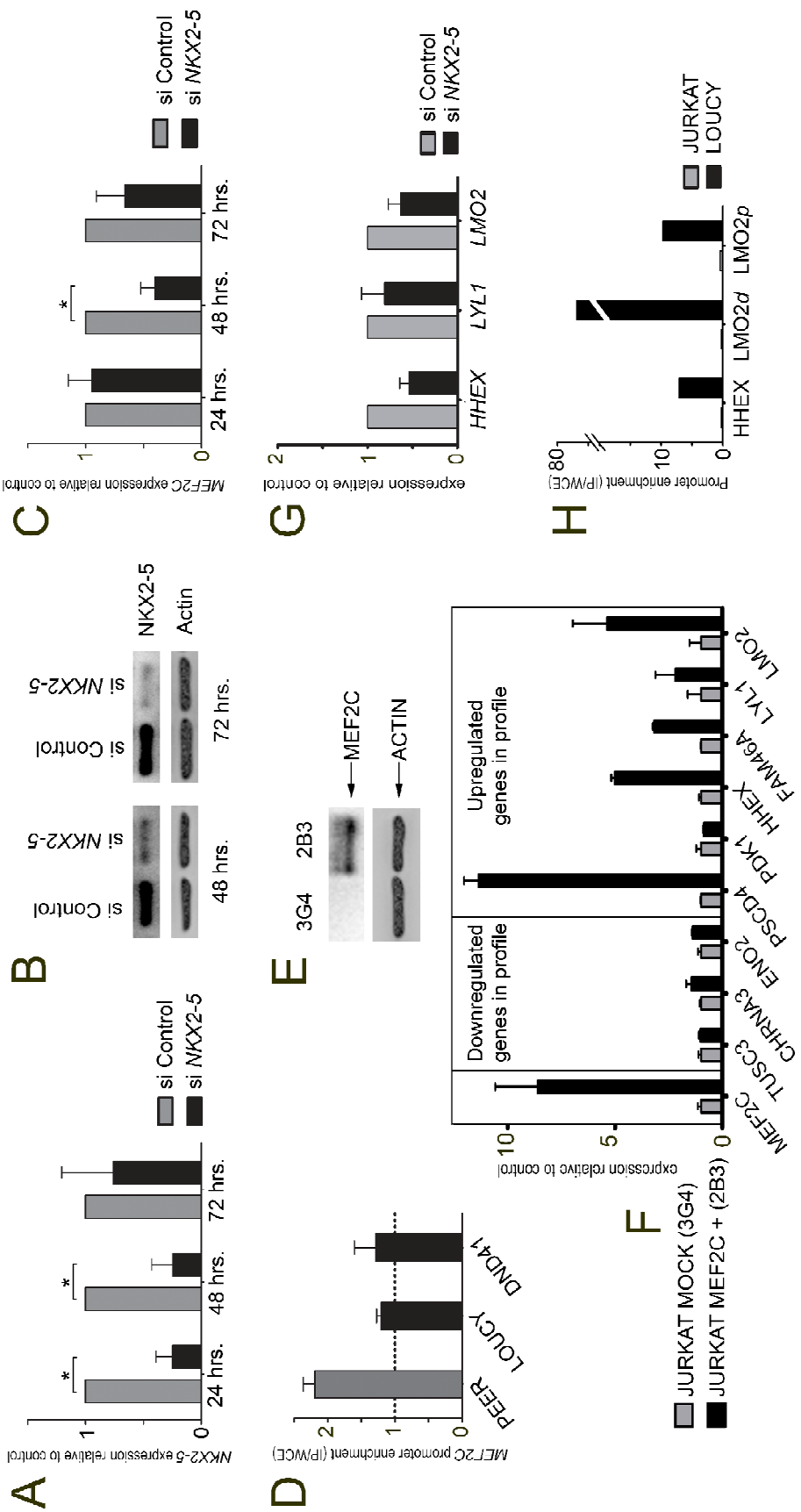

A
si Control
si NKX2-5
NKX2-5 expression relative to control
24 hrs.
48 hrs.
72 hrs.
B
si Control
si NKX2-5
48 hrs.
si Control
si NKX2-5
72 hrs.
NKX2-5
Actin
C
si Control
si NKX2-5
MEF2C expression relative to control
24 hrs.
48 hrs.
72 hrs.
D
MEF2C promoter enrichment (IP/MCE)
PEER
LOUCY
DND41
E
3G4
2B3
MEF2C
ACTIN
F
JURKAT MOCK (3G4)
JURKAT MEF2C + (2B3)
expression relative to control
Downregulated genes in profile
Upregulated genes in profile
MEF2C
TUSC3
CHRNA3
ENO2
PSCD4
PDK1
HHEX
FAM46A
LYL1
LMO2
G
si Control
si NKX2-5
expression relative to control
HHEX
LYL1
LMO2
H
JURKAT
LOUCY
Promoter enrichment (IP/MCE)
HHEX
LMO2d
LMO2p

**Figure 5. (page 42). NKX2-5 Controls *MEF2C* Expression.** (**A**) RQ-PCR results of *NKX2-5* mRNA expression levels or (**B**) NKX2-5 protein levels in cell line PEER at indicated time points following electroporation with siRNAs directed against *NKX2-5* (black bars) relative to control siRNA treated cells (grey bars). For western blot analysis, Actin was used as a loading control. (**C**) RQ-PCR results of *MEF2C* mRNA expression levels at indicated time points following electroporation with anti-*NKX2-5* siRNA molecules (black bars) relative to controls (grey bars). (**D**) Enrichment of MEF2C promoter sequences in NKX2-5 ChIP analysis in the *NKX2-5* translocated cell line PEER, but not in negative control lines LOUCY or DND41. (**E**) Ectopic MEF2C expression in the MEF2C stably transfected Jurkat clone 2B3 as shown by western blot analysis. The mock transfected Jurkat clone 3G4 served as negative control. (**F**) RQ-PCR results for MEF2C positive Jurkat clone 2B3 clone or the mock transfected control (3G4) for *MEF2C* and random selected immature signature genes that are relatively down- (*TUSC3*, *CHRNA3*, *ENO2*) or upregulated (*PSCD4 (CYTH4)*, *PDK1*, *HHEX*, *FAM46A*, *LYL1*, *LMO2*) in immature T-ALL cases compared to other cluster samples. (**G**). Relative expression results for *HHEX*, *LYL1* and *LMO2* in the cell line PEER 72 hours after electroporation with siRNAs directed against *NKX2-5* (black bars) relative to control siRNA treated PEER cells (grey bars). (**H**) Enrichment of *HHEX* promoter and the distal and proximal *LMO2* promoters upon MEF2C ChIP analysis in the immature cell line LOUCY, but not in the negative control line Jurkat. For all panels, the SD are shown.

### Oncogenic activity of *NKX2-1* and *MEF2C*

To substantiate potential oncogenic activity for *NKX2-1* and *MEF2C*, we tested whether both genes had transforming capacity by using cellular transformation assays in NIH3T3 (**Figure 6A**) or BJ-EHT cells (**Figure 6B**). Transfecting *NKX2-1* or *MEF2C* expression constructs into the cells was insufficient to drive cellular transformation. We then tested cellular transformation of *MEF2C* and *NKX2-1* when combined with *RAS* or *MYC*, two oncogenes that are frequently activated in T-ALL through RAS or NOTCH1 activating mutations (Kawamura et al., 1999; Palomero et al., 2006; Weng et al., 2006). *NKX2-1* and *MEF2C* were both able to synergize with *RAS* or *MYC* genes in driving cellular transformation (**Figures 6A-B**).

We then further tested the importance of *MEF2C* for T-cell pathogenesis for which we had a cell line model available. In normal human T-cell development subsets, *MEF2C* is exclusively expressed at the pre-DN1 and DN1 stages, after which it is downregulated (**Figure S5**). We knocked-down MEF2C expression in T-ALL cell line LOUCY using siRNA molecules. MEF2C knock-down induced cellular differentiation as LOUCY cells became positive for membrane CD3 and

TCRγδ expression (**Figures 6C-E**). This indicates that MEF2C can block T-cell differentiation at a very immature stage.

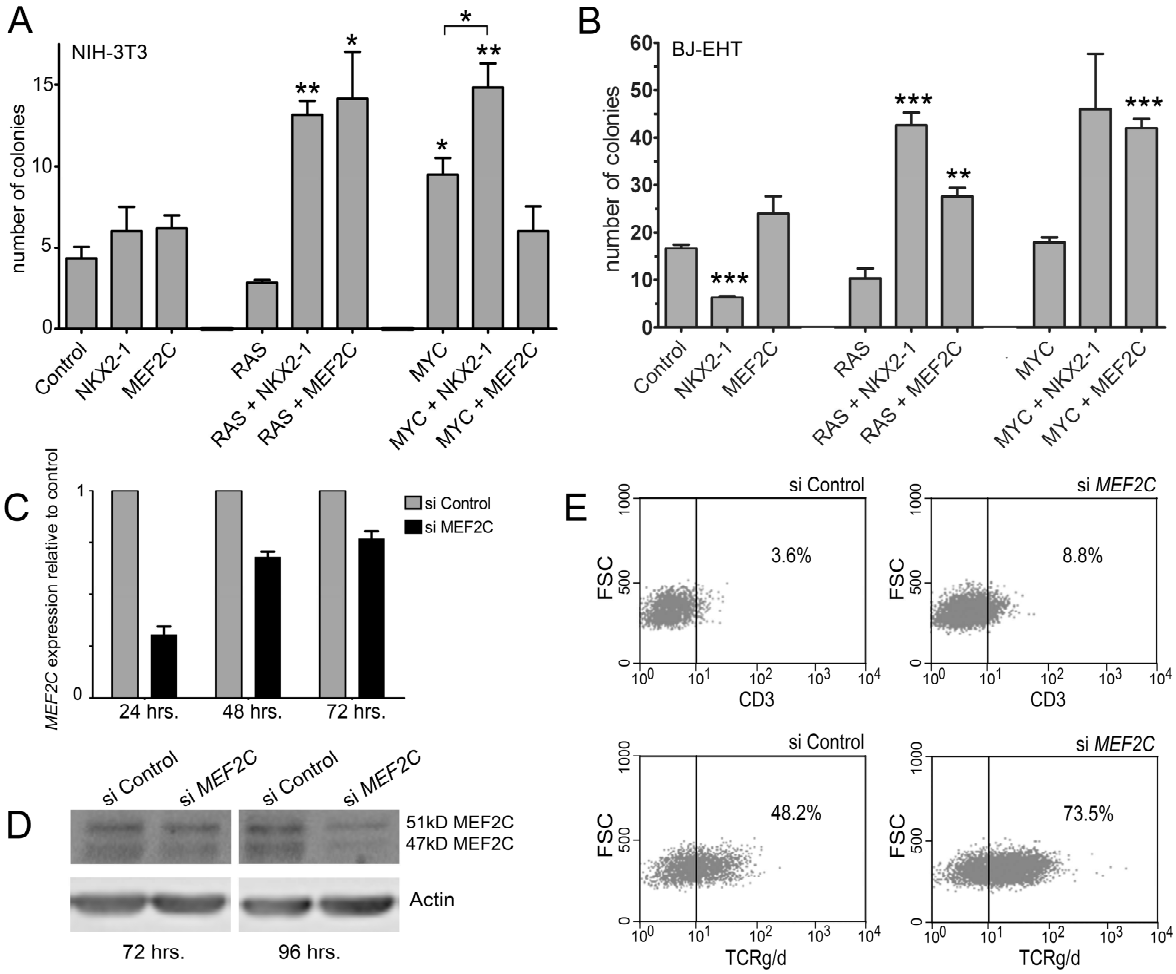



**Figure 6. Cellular Transformation by *MEF2C* and *NKX2-1*.** Cellular transformation of (**A**) NIH3T3 or (**B**) BJ-EHT cells upon transfection of *MEF2C*, *NKX2-1*, *MYC* and/or *RAS* expression vectors as indicated. Significance levels for colony number differences between indicated expression construct combinations relative to the empty vector control are indicated (*p≤0.05, ** p≤0.01, *** p≤0.001). **(C)** *MEF2C* expression knockdown as measured by RQ-PCR for the *MEF2C*-positive cell line LOUCY at indicated time points following electroporation with *MEF2C*-specific siRNA molecules (black bars) relative to control siRNA treated cells (grey bars). (**D**) Downregulation of MEF2C protein following treatment with *MEF2C*-specific siRNA molecules as validated by western blot. Based on the protein size, the predominant α1β (47 KD) and the α1βγ (51.2KD) MEF2C isoform (Zhu and Gulick, 2004) are indicated. (**E**) Increase of mCD3 (p=0.0032) and TCRγδ (p=0.023) expression as demonstrated by FACS analysis in LOUCY cells, 96 hrs following treatment with *MEF2C*-specific siRNA molecules. A representative example from three independent experiments is shown. For all panels, the SEM are shown. See also **Figure S5**.

**Validation of the immature and proliferative clusters in independent T-ALL cohorts**

We then confirmed our T-ALL clustering (**Figures 1A and 1C**) and molecular-cytogenetic findings (**Table 2)** in 2 independent validation cohorts, i.e. a French dataset comprising 107 pediatric and adult T-ALL cases (Clappier et al., 2007; Soulier et al., 2005) and a second Rotterdam cohort comprising 108 pediatric and adult T-ALL cases. Upon testing the comparability of the initial Rotterdam cohort and the French dataset (**Figure S6A**) the proliferative and immature clusters could be reproduced in a combined unsupervised cluster analysis (**Figure S6B**). Based on the unsupervised clustering of our initial Rotterdam cohort, PAM statistics predicted various proliferative cluster cases as well as immature cluster cases in the Rotterdam validation cohort (not shown). Twenty-six proliferative cluster cases were identified of which various samples highly expressed *NKX2-1* (**Figure S6C**). *NKX2-1* translocations/inversions could be demonstrated using FISH in 3 cases (**Figures S6G-I**). Eight out of 10 *TLX1* rearranged cases were part of the proliferative cluster as well (**Figure S6D**, and data not shown), further supporting the notion that *NKX2-1* and *TLX1* oncogenic rearrangements may share common pathogenic mechanisms. Again, some of these *TLX1* rearranged cases also expressed *NKX2-1* at low levels (**Figure S6C**) while none of these samples had *NKX2-1* rearrangements. We also validated high *MEF2C* expression for the 24 cases that were assigned to the immature cluster by PAM analysis (**Figure S6E**), and these samples expressed the highest levels of its downstream target *LYL1* (**Figure S6F**).

**DISCUSSION**

In this study, we have identified *NKX2-1*, its related family member *NKX2-2* and *MEF2C* as potential oncogenes for T-ALL. Supervised cluster analyses based on genes uniquely associated with the known genetic *TAL/LMO*, *TLX3*, *TLX1* and *HOXA* subgroups revealed that samples with high expression of *NKX2-1/NKX2-2* or *MEF2C* characterize 2 T-ALL clusters for which no driving oncogenic hits have been identified so far. Both clusters represent about 20 percent of all T-ALL cases.

Variant rearrangements for *NKX2-1* and *NKX2-2* to T-cell receptor genes *(TRAD@, TRB@)* were identified, and one case had an inversion to the *IGH@* locus. The *IgH*-enhancer seems functional in this T-ALL patient and *IgH*-enhancer (Eμ) driven oncogene expression in a T-cell context has been described before,

both for human T-ALL (Nguyen-Khac et al., 2010) as well as in transgenic mouse models (Katsumata et al., 1992; Strasser et al., 1991). This patient did not express B-cell markers therefore excluding it as a bi-phenotypic leukemia. *NKX2-1* was able to transform NIH3T3 and BJ-EHT cells in synergism with *RAS* or *MYC*, two genes that become activated through RAS or NOTCH1 activating mutations in approximately 15 and 60 percent of T-ALL cases, respectively (Kawamura et al., 1999; Palomero et al., 2006; Weng et al., 2006). Our data therefore strongly support that *NKX2-1/NKX2-2* may represent oncogenes in T-ALL. *NKX2-1* is not expressed during normal T-cell development based on expression data by microarray for flow-sorted thymic subsets (Dik et al., 2005; Soulier et al., 2005).

*NKX2-1* and *NKX2-2* have been associated with other types of cancer before: *NKX2-1* is amplified in human lung cancer (Weir et al., 2007), and *NKX2-2* is a target of the EWS/FLI fusion product in Ewing-Sarcoma (Smith et al., 2006). NKX2-1 and NKX2-2 are 59% identical for the homeodomain region, indicating that both proteins may exert identical oncogenic roles in T-ALL. This is further supported by the fact that rearrangements for both genes were identified in samples that tightly cluster together in unsupervised and supervised analyses. NK-like homeobox transcription factors play important roles in T-ALL as *NKX2-5* was previously identified as part of an oncogenic rearrangement in T-ALL (Nagel et al., 2003). The NK-like homeobox transcription factor *NKX3-1* has been found to be highly activated in *TAL1* rearranged cases (Soulier et al., 2005), as a direct TAL1 target gene (Kusy et al., 2010). The homeodomains of NKX2-5 and NKX3-1 are only distantly related (37% identity) and only 48% and 47% identical to the homeodomain of NKX2-1, respectively. This may explain why *NKX2-5*, *NKX3-1* and *NKX2-1/NKX2-2* are associated with different T-ALL subgroups: ectopic *NKX3-1* expression in the *TAL/LMO* subgroup (Soulier et al., 2005), *NKX2-1/NKX2-2* rearrangements with the proliferative T-ALL cluster (this study) and *NKX2-5* translocations with immature T-cell development (this study) that activates *MEF2C* (this study and (Nagel et al., 2008)).

In unsupervised analyses, *NKX2-1/NKX2-2* rearranged cases cluster together with *TLX1* rearranged cases to form the proliferative cluster. This indicates that *NKX2-1/NKX2-2* and *TLX1* rearranged T-ALLs are biologically related. This is further supported by the fact that *NKX2-1* and *TLX1* rearranged cases share a similar immunophenotypic makeup consistent with cortical arrest as well by the fact that various *TLX1* rearranged cases express *NKX2-1* in the absence of *NKX2-1* rearrangements albeit at low levels. One of the explanations may be that

TLX1 controls *NKX2-1* expression. In addition, several other cases that are part of the proliferative cluster lack *TLX1*, *NKX2-1* or *NKX2-2* rearrangements, indicating that an additional oncogenic rearrangement awaits identification for this cluster.

The second cluster had a very immature immunophenotype, with most cases expressing CD34 and frequently co-expressing the CD13 and/or CD33 myeloid markers. We identified various rearrangements that directly or indirectly activate *MEF2C*. *MEF2C* is a member of the MADS-box transcription factor family that includes the 4 *MEF2A-D* genes that are important regulators of skeletal muscle development (Grounds, 1991). Immature T-ALL subgroups have been identified before (Coustan-Smith et al., 2009; Ferrando et al., 2002; Soulier et al., 2005), and our immature cluster cases could also be predicted based on an ETP expression signature (Coustan-Smith et al., 2009). We now conclude that *MEF2C* is the driving oncogene for immature (ETP) T-ALL cases. Our immature cases also have the highest *LYL1* expression and highly express *LMO2* ((Ferrando et al., 2002), and this study). *LYL1* and *LMO2* are members of the basic helix-loop-helix (bHLH) family and the LIM-domain only family, respectively. Apart from *LYL1* and *LMO2*, the immature cases also highly express the homeobox gene *HHEX*. We have now shown that *HHEX*, *LYL1* and *LMO2* are being regulated by MEF2C, and it was proven that MEF2C directly binds in the promoter regions of at least *HHEX* and *LMO2*. This may support a pathogenic role for established oncogenes as LYL1 and LMO2 in MEF2C deregulated early committed T-cells. To what extend *LMO2* and/or *LYL1* as MEF2C targets will be sufficient to drive a leukemogenic program in these early committed T-cells is presently unclear. Oncogenic rearrangements of *LMO2* and *LYL1* have not been observed in immature T-ALL ((Ferrando et al., 2002) and this work), but are exclusive for the *TAL/LMO* subgroup that also includes rearrangements of the *LYL1*-homologous *TAL1* gene. Therefore, MEF2C may elicit a more comprehensive transcriptional program characteristic for ETP T-ALLs than aberrant expression of LMO2 or LYL1 alone.

MEF2C is a key regulator for lymphoid development that is activated by PU.1 (Stehling-Sun et al., 2009). In B-cell development, MEF2C is activated by calcineurin following BCR-triggering and warrants for cell-viability and proliferation (Wilker et al., 2008). MEF2C has been implicated in human oncogenesis: in myeloid leukemias of *MLL-AF9* transgenic mice, *Mef2c* has been identified as a HoxA9 target gene that regulates selfrenewal of leukemic stem cells (Krivtsov et al., 2006). *MEF2C* is also highly expressed in human *MLL*-rearranged AML that is characterized by upregulation of *HOXA* genes including *HOXA9*

(Schwieger et al., 2009). *Mef2c* is further identified as potential oncogene in insertional mutagenesis studies (Du et al., 2005; Schwieger et al., 2009), and can provoke myeloid leukemias (Schwieger et al., 2009). Also the related family member *MEF2D* is involved in the *MEF2D-DAZAP1* fusion that has been identified in ALL (Prima and Hunger, 2007).

Many oncogenic hits as identified in this study involve early hematopoietic transcription factors including NKX2-5, PU.1 and presumably RUNX1. These factors are important for normal T-cell development (Rothenberg, 2007). All these factors converge on MEF2C in immature T-ALL, and it is tempting to speculate that MEF2C is a central regulator for normal early T-cell development. MEF2C may need to become downregulated to facilitate maturation beyond this immature stage, and we indeed demonstrated that knockdown of MEF2C expression in T-ALL cell line LOUCY provoked differentiation. In support of these notions, MEF2C is expressed in normal human thymocyte pre-DN1 and DN1 subsets, but expression is dramatically decreased beyond the DN2 stage (**Figure S5**). A similar downregulation of *MEF2C* expression could be validated from gene expression data for equivalent flow-sorted thymic subsets as published ((Dik et al., 2005); data not shown). *MEF2C* may represent the central oncogene for immature T-ALL cases that seems to provide a T-cell differentiation block at the immature stage as demonstrated in this article. This was further supported by our transformation assay results in which *MEF2C* transformed NIH3T3 and BJ-EHT cells in combination with *RAS* or *MYC*. We also observed that several genes from the TGFBR1 pathway were upregulated, including *TGFBR1*, *ZEB2*, *SMAD7*, *SMURF2* and *RUNX3*, or downregulated (*SMAD1*). Since both activators (*TGFBR1*) or in some extent inhibitors (like *SMURF2*, *SMAD7*) are overexpressed while *SMAD1* is underexpressed, it is difficult to anticipate the functional consequences of this pathway for the immature T-ALL cases.

In conclusion, we used a strategy integrating molecular genetics with large scale expression profiling and identified two novel oncogenic subgroups and 8 genomic rearrangements that have not been identified before in human cancer. We have shown that these proliferative and immature subtypes reflect different biological entities: the proliferative cluster strongly express proliferation genes and is associated with aberrations and ectopic expression of *NKX2-1* or *NKX2-2*, and expression of CD1. In contrast, the immature cluster was characterized by immature T-cell development, activation of genes involved in protein binding and dimerization, expression of components of the TGFBR1 pathway and high

expression of the MADS transcription factor *MEF2C* due to abnormalities of *MEF2C*, transcription factors that regulate *MEF2C* or MEF2C-associating cofactors. We conclude that *NKX2-1*, *NKX2-2* and *MEF2C* define oncogenic pathways in T-ALL.

## EXPERIMENTAL PROCEDURES

**Patient samples.** Viably frozen diagnostic bone marrow or peripheral blood samples from 117 pediatric T-ALL patients and corresponding clinical and immunophenotypic data were provided by the German Co-operative study group for childhood Acute Lymphoblastic Leukemia (COALL) and the Dutch Childhood Oncology Group (DCOG). The patients' parents or their legal guardians provided informed consent to use leftover material for research purposes according to the declaration of Helsinki, as this study was approved by the Institutional review board of the ErasmusMC Rotterdam. Leukemic cells were isolated and enriched from these samples as previously described (Van Vlierberghe et al., 2006). All resulting samples contained ≥90% leukemic cells, as determined morphologically by May-Grünwald-Giemsa-stained cytospins (Merck, Darmstadt, Germany). Patients were assigned to specific molecular-cytogenetic T-ALL subgroups based on FISH results for *TAL1*, *TAL2*, *LMO1*, *LMO2*, *TLX1*, *TLX3*, *CALM-AF10*, *SET-NUP214*, *MLL*, *MYB*, or Inv(7)(p15;q34)) and positivity by RT-PCR for *SIL-TAL1*, *TLX1*, *TLX3*, *CALM-AF10* or *SET-NUP214* as described before (van Grotel et al., 2006; Van Vlierberghe et al., 2006; Van Vlierberghe et al., 2008b).

**Chromosome Conformation Capture on Chip (4C).** 4C was performed as described before (Simonis et al., 2006). Briefly, DNA and protein in approximately 10 million viable cells was cross linked in a 2% formaldehyde solution to conserve the physical proximity of DNA regions. Cells were lysed and DNA was digested with HindIII. After dilution of DNA, restriction fragments were ligated. This way, DNA fragments that are physically near each other in the viable cell can be ligated. The sample was subsequently de-crosslinked by an overnight incubation at 65°C. DNA was purified and digested with the frequent cutter DpnII. Samples were diluted and ligated to allow circularization of individual restriction fragments. Following linearization with ScaI (located between both inverse PCR primers), DNA sequences ligated to the fragment of interest were amplified by inverse PCR, labeled and hybridized on a microarray (Nimblegen, Madison, USA) containing

probes that roughly represent individual HindIII fragments in the genome. Raw fluorescence intensities are visualized as the running median per 30 neighboring probes, each representing a HindIII restriction fragments. The viewpoint (VP) is the HindIII restriction fragment where 4C PCR primers are located. Data are visualized with SignalMap software (Nimblegen) (NCBI, Build 36). Inverse PCR primer sets developed for *NKX2-1*, *BCL11B* and *MEF2C* are listed in the **Supplemental experimental procedures**.

**Gene expression microarray, data extraction and normalization.** Integrity of patient samples total RNA was checked using the Agilent 2100 Bio-analyzer (Agilent, Santa-Clara, USA). Copy-DNA and ccRNA syntheses from total RNA, hybridization to Humane Genome U133 plus2.0 oligonucleotide microarrays (Affymetrix, Santa-Clara, USA) and washing steps were performed according to the manufacturers' protocol. Probeset intensities were extracted from CEL-files in the statistical data analysis environment *R*, version 2.8.0 (Bioconductor Affy package). All arrays had a 3' to 5' GAPDH ratio lower than 3 fold. Probe intensities were normalized in *R* using RMA-solo, RMA (Irizarry et al., 2003) or VSN (Huber et al., 2002) methods.

**Biostatistical analyses**. Biostatistical analyses have been described in detail in the **Supplemental experimental procedures**. Briefly, unsupervised cluster analyses were performed in Dchip (Li and Wong, 2001). Identification of differentially expressed genes with FDR control was done by various methods including Wilcoxon statistics ("Multtest" in R), SAM statistics (Tusher et al., 2001) (BRB tools, version 3.7, R. Simon & A.P. Lam), and COPA statistics (Tomlins et al., 2005) for outlier analysis using a R routine. Prediction of identified subtypes was done using various algorithms embedded in BRB tools including Diagonal Linear Discriminant Analysis, 1-nearest neighbor, 3-nearest neighbor and nearest centroid as well as tested by prediction analysis for microarrays (PAM) (Tibshirani et al., 2002). Principal component analysis (PCA) based on the top100 most significant differentially expressed genes for the major T-ALL subgroups (i.e. the supervised analysis) was performed using GeneMath XT 1.6.1. software (Applied Maths, Inc, Austin TX, USA). To validate findings from the Rotterdam dataset, this dataset was combined with the French (Paris) Affymetrix U133A dataset (Soulier et al., 2005). Data for overlapping probesets were extracted from both datasets, RMA-solo normalized and corrected for batch effects using the Combat Method (Johnson

et al., 2007). Profiles for similar T-ALL subgroups in both datasets were tested for comparability by using various methods, including the OrderedList method using the Bioconductor package "OrderedList" in R (Scheid, S., Lottaz, C., Yang, X., and Spang, R) as well as the subclass method (Hoshida et al., 2007).

Additional methods and materials are described in the **Supplemental experimental procedures**.

**Accession numbers**

Microarray data are available at http://www.ncbi.nlm.nih.gov/geo/ and the EBI database at http://www.ebi.ac.uk/arrayexpress under accession numbers GSE26713 and E-MEXP-313, respectively.

**Acknowledgements**

I.H. is financed by the Dutch Cancer Society Dutch Cancer Society (KWF-EMCR 2006-3500). C.K. and M.V. are financed by the Stichting Kinderen Kankervrij (KiKa; Grant no. KiKa 2008-029). We also would like to thank the German Jose Carreras Leukemia Foundation (Grant no. SP 04/03) and the French program Carte d'Identité des Tumeurs (CIT) from the Ligue Contre le Cancer for financial support.

## REFERENCES

Buijs, A., van Rompaey, L., Molijn, A. C., Davis, J. N., Vertegaal, A. C., Potter, M. D., Adams, C., van Baal, S., Zwarthoff, E. C., Roussel, M. F., and Grosveld, G. C. (2000). The MN1-TEL fusion protein, encoded by the translocation (12;22)(p13;q11) in myeloid leukemia, is a transcription factor with transforming activity. Mol Cell Biol *20*, 9281-9293.
Chen, S. L., Dowhan, D. H., Hosking, B. M., and Muscat, G. E. (2000). The steroid receptor coactivator, GRIP-1, is necessary for MEF-2C-dependent gene expression and skeletal muscle differentiation. Genes Dev *14*, 1209-1228.
Clappier, E., Cuccuini, W., Kalota, A., Crinquette, A., Cayuela, J. M., Dik, W. A., Langerak, A. W., Montpellier, B., Nadel, B., Walrafen, P., *et al.* (2007). 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 *110*, 1251-1261.
Coustan-Smith, E., Mullighan, C. G., Onciu, M., Behm, F. G., Raimondi, S. C., Pei, D., Cheng, C., Su, X., Rubnitz, J. E., Basso, G., *et al.* (2009). Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia. Lancet Oncol *10*, 147-156.
Dik, W. A., Pike-Overzet, K., Weerkamp, F., de Ridder, D., de Haas, E. F., Baert, M. R., van der Spek, P., Koster, E. E., Reinders, M. J., van Dongen, J. J., *et al.* (2005). New insights on human T cell development by quantitative T cell receptor gene rearrangement studies and gene expression profiling. J Exp Med *201*, 1715-1723.
Du, Y., Spence, S. E., Jenkins, N. A., and Copeland, N. G. (2005). Cooperating cancer-gene identification through oncogenic-retrovirus-induced insertional mutagenesis. Blood *106*, 2498-2505.
Ferrando, A. A., Neuberg, D. S., Staunton, J., Loh, M. L., Huard, C., Raimondi, S. C., Behm, F. G., Pui, C. H., Downing, J. R., Gilliland, D. G., *et al.* (2002). Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. Cancer Cell *1*, 75-87.
Grosveld, G. C. (2007). MN1, a novel player in human AML. Blood Cells Mol Dis *39*, 336-339.
Grounds, M. D. (1991). Towards understanding skeletal muscle regeneration. Pathol Res Pract *187*, 1-22.
Hebert, J., Cayuela, J. M., Berkeley, J., and Sigaux, F. (1994). Candidate tumor-suppressor genes MTS1 (p16INK4A) and MTS2 (p15INK4B) display frequent homozygous deletions in primary cells from T- but not from B-cell lineage acute lymphoblastic leukemias. Blood *84*, 4038-4044.
Hoshida, Y., Brunet, J. P., Tamayo, P., Golub, T. R., and Mesirov, J. P. (2007). Subclass mapping: identifying common subtypes in independent disease data sets. PLoS ONE *2*, e1195.
Huber, W., von Heydebreck, A., Sultmann, H., Poustka, A., and Vingron, M. (2002). Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics *18 Suppl 1*, S96-104.
Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., and Speed, T. P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics *4*, 249-264.

Johnson, W. E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics *8*, 118-127.
Katsumata, M., Siegel, R. M., Louie, D. C., Miyashita, T., Tsujimoto, Y., Nowell, P. C., Greene, M. I., and Reed, J. C. (1992). Differential effects of Bcl-2 on T and B cells in transgenic mice. Proc Natl Acad Sci U S A *89*, 11376-11380.
Kawamura, M., Ohnishi, H., Guo, S. X., Sheng, X. M., Minegishi, M., Hanada, R., Horibe, K., Hongo, T., Kaneko, Y., Bessho, F., *et al.* (1999). Alterations of the p53, p21, p16, p15 and RAS genes in childhood T-cell acute lymphoblastic leukemia. Leuk Res *23*, 115-126.
Krivtsov, A. V., Twomey, D., Feng, Z., Stubbs, M. C., Wang, Y., Faber, J., Levine, J. E., Wang, J., Hahn, W. C., Gilliland, D. G., *et al.* (2006). Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. Nature *442*, 818-822.
Kusy, S., Gerby, B., Goardon, N., Gault, N., Ferri, F., Gerard, D., Armstrong, F., Ballerini, P., Cayuela, J. M., Baruchel, A., *et al.* (2010). NKX3.1 is a direct TAL1 target gene that mediates proliferation of TAL1-expressing human T cell acute lymphoblastic leukemia. J Exp Med *207*, 2141-2156.
Li, C., and Wong, W. H. (2001). Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. Proc Natl Acad Sci U S A *98*, 31-36.
Nagel, S., Kaufmann, M., Drexler, H. G., and MacLeod, R. A. (2003). The cardiac homeobox gene NKX2-5 is deregulated by juxtaposition with BCL11B in pediatric T-ALL cell lines via a novel t(5;14)(q35.1;q32.2). Cancer Res *63*, 5329-5334.
Nagel, S., Meyer, C., Quentmeier, H., Kaufmann, M., Drexler, H. G., and MacLeod, R. A. (2008). MEF2C is activated by multiple mechanisms in a subset of T-acute lymphoblastic leukemia cell lines. Leukemia *22*, 600-607.
Nguyen-Khac, F., Barin, C., Chapiro, E., Macintyre, E. A., Romana, S., and Bernard, O. A. (2010). Cyclin D3 deregulation by juxtaposition with IGH locus in a t(6;14)(p21;q32)-positive T-cell acute lymphoblastic leukemia. Leuk Res *34*, e13-14.
Palomero, T., Lim, W. K., Odom, D. T., Sulis, M. L., Real, P. J., Margolin, A., Barnes, K. C., O'Neil, J., Neuberg, D., Weng, A. P., *et al.* (2006). NOTCH1 directly regulates c-MYC and activates a feed-forward-loop transcriptional network promoting leukemic cell growth. Proc Natl Acad Sci U S A *103*, 18261-18266.
Pieters, R., and Carroll, W. L. (2008). Biology and treatment of acute lymphoblastic leukemia. Pediatr Clin North Am *55*, 1-20, ix.
Prima, V., and Hunger, S. P. (2007). Cooperative transformation by MEF2D/DAZAP1 and DAZAP1/MEF2D fusion proteins generated by the variant t(1;19) in acute lymphoblastic leukemia. Leukemia *21*, 2470-2475.
Przybylski, G. K., Dik, W. A., Grabarczyk, P., Wanzeck, J., Chudobska, P., Jankowski, K., von Bergh, A., van Dongen, J. J., Schmidt, C. A., and Langerak, A. W. (2006). The effect of a novel recombination between the homeobox gene NKX2-5 and the TRD locus in T-cell acute lymphoblastic leukemia on activation of the NKX2-5 gene. Haematologica *91*, 317-321.
Pui, C. H., and Evans, W. E. (2006). Treatment of acute lymphoblastic leukemia. N Engl J Med *354*, 166-178.
Rothenberg, E. V. (2007). Regulatory factors for initial T lymphocyte lineage specification. Curr Opin Hematol *14*, 322-329.

Schwieger, M., Schuler, A., Forster, M., Engelmann, A., Arnold, M. A., Delwel, R., Valk, P. J., Lohler, J., Slany, R. K., Olson, E. N., and Stocking, C. (2009). Homing and invasiveness of MLL/ENL leukemic cells is regulated by MEF2C. Blood.
Simonis, M., Klous, P., Homminga, I., Galjaard, R. J., Rijkers, E. J., Grosveld, F., Meijerink, J. P. P., and De Laat, W. (2009). High-resolution identification of balanced and complex chromosomal rearrangements by 4C technology. Nat Methods *Oct 11 online*.
Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet *38*, 1348-1354.
Smith, R., Owen, L. A., Trem, D. J., Wong, J. S., Whangbo, J. S., Golub, T. R., and Lessnick, S. L. (2006). Expression profiling of EWS/FLI identifies NKX2.2 as a critical target gene in Ewing's sarcoma. Cancer Cell *9*, 405-416.
Soulier, J., Clappier, E., Cayuela, J. M., Regnault, A., Garcia-Peydro, M., Dombret, H., Baruchel, A., Toribio, M. L., and Sigaux, F. (2005). HOXA genes are included in genetic and biologic networks defining human acute T-cell leukemia (T-ALL). Blood *106*, 274-286.
Stehling-Sun, S., Dade, J., Nutt, S. L., DeKoter, R. P., and Camargo, F. D. (2009). Regulation of lymphoid versus myeloid fate 'choice' by the transcription factor Mef2c. Nat Immunol *10*, 289-296.
Strasser, A., Harris, A. W., and Cory, S. (1991). bcl-2 transgene inhibits T cell death and perturbs thymic self-censorship. Cell *67*, 889-899.
Strehl, S., Nebral, K., Konig, M., Harbott, J., Strobl, H., Ratei, R., Struski, S., Bielorai, B., Lessard, M., Zimmermann, M., *et al.* (2008). ETV6-NCOA2: a novel fusion gene in acute leukemia associated with coexpression of T-lymphoid and myeloid markers and frequent NOTCH1 mutations. Clin Cancer Res *14*, 977-983.
Tibshirani, R., Hastie, T., Narasimhan, B., and Chu, G. (2002). Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci U S A *99*, 6567-6572.
Tomlins, S. A., Rhodes, D. R., Perner, S., Dhanasekaran, S. M., Mehra, R., Sun, X. W., Varambally, S., Cao, X., Tchinda, J., Kuefer, R., *et al.* (2005). Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science *310*, 644-648.
Tusher, V. G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A *98*, 5116-5121.
van Grotel, M., Meijerink, J. P., Beverloo, H. B., Langerak, A. W., Buys-Gladdines, J. G., Schneider, P., Poulsen, T. S., den Boer, M. L., Horstmann, M., Kamps, W. A., *et al.* (2006). The outcome of molecular-cytogenetic subgroups in pediatric T-cell acute lymphoblastic leukemia: a retrospective study of patients treated according to DCOG or COALL protocols. Haematologica *91*, 1212-1221.
Van Vlierberghe, P., Pieters, R., Beverloo, H. B., and Meijerink, J. P. (2008a). Molecular-genetic insights in paediatric T-cell acute lymphoblastic leukaemia. Br J Haematol *143*, 153-168.
Van Vlierberghe, P., van Grotel, M., Beverloo, H. B., Lee, C., Helgason, T., Buijs-Gladdines, J., Passier, M., van Wering, E. R., Veerman, A. J., Kamps, W. A., *et al.* (2006). The cryptic chromosomal deletion, del(11)(p12p13), as a new activation mechanism of LMO2 in pediatric T- cell acute lymphoblastic leukemia. Blood.

Van Vlierberghe, P., van Grotel, M., Tchinda, J., Lee, C., Beverloo, H. B., van der Spek, P. J., Stubbs, A., Cools, J., Nagata, K., Fornerod, M., *et al.* (2008b). The recurrent SET-NUP214 fusion as a new HOXA activation mechanism in pediatric T-cell acute lymphoblastic leukemia. Blood *111*, 4668-4680.
Weir, B. A., Woo, M. S., Getz, G., Perner, S., Ding, L., Beroukhim, R., Lin, W. M., Province, M. A., Kraja, A., Johnson, L. A., *et al.* (2007). Characterizing the cancer genome in lung adenocarcinoma. Nature *450*, 893-898.
Weng, A. P., Ferrando, A. A., Lee, W., Morris, J. P. t., Silverman, L. B., Sanchez-Irizarry, C., Blacklow, S. C., Look, A. T., and Aster, J. C. (2004). Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. Science *306*, 269-271.
Weng, A. P., Millholland, J. M., Yashiro-Ohtani, Y., Arcangeli, M. L., Lau, A., Wai, C., Del Bianco, C., Rodriguez, C. G., Sai, H., Tobias, J., *et al.* (2006). c-Myc is an important direct target of Notch1 in T-cell acute lymphoblastic leukemia/lymphoma. Genes Dev *20*, 2096-2109.
Wilker, P. R., Kohyama, M., Sandau, M. M., Albring, J. C., Nakagawa, O., Schwarz, J. J., and Murphy, K. M. (2008). Transcription factor Mef2c is required for B cell proliferation and survival after antigen receptor stimulation. Nat Immunol *9*, 603-612.
Zhu, B., and Gulick, T. (2004). Phosphorylation and alternative pre-mRNA splicing converge to regulate myocyte enhancer factor 2C activity. Mol Cell Biol *24*, 8264-8275.

# CHAPTER 2

## Characterization of a pediatric T-cell acute lymphoblastic leukemia patient with simultaneous *LYL1* and *LMO2* rearrangements

Irene Homminga[1], Maartje J Vuerhard[1], Anton W Langerak[2], Jessica Buijs-Gladdines[1], Rob Pieters[1], Jules PP Meijerink[1]

*[1]Department of Pediatric Oncology/Hematology, Erasmus MC/Sophia Children's Hospital, Rotterdam, The Netherlands; [2]Department of Immunology, Erasmus MC, Rotterdam, The Netherlands.*

*Submitted*

**ABSTRACT**

*Lymphoid leukemia 1 (LYL1)* translocations are rare in T-cell acute lymphoblastic leukemia (T-ALL), whereas the homologous *TAL1* oncogene is rearranged in approximately 20% of pediatric T-ALL patients. Previous gene-expression studies have identified immature T-ALL patient groups (ETP-ALL) that highly express *LYL1* in the absence of *LYL1* aberrations. Molecular characterization of a t(7;19)(q34;p13) in a pediatric T-ALL patient led to the identification of a T-cell receptor beta enhancer translocation to the *LYL1* locus. Alike incidental T-ALL cases having double *TAL1/2* and *LMO1/2* synergistic translocations, this *LYL1*-translocated case also had a *LMO2* rearrangement pointing to oncogenic cooperation between *LYL1* and *LMO2*. In hierarchical cluster analyses based on gene-expression data, this sample consistently clustered along with *TAL1*- and/or *LMO2*-rearranged cases in the *TAL/LMO* subgroup. We conclude that *LYL1*-rearranged T-ALL cases are not necessarily associated with the immature, ETP-ALL subgroup despite their high *LYL1* expression levels but elicit a *TAL/LMO* expression signature.

## INTRODUCTION

T-cell acute lymphoblastic leukemia (T-ALL) is characterized by chromosomal rearrangements that activate several oncogenes, such as *TAL1, LMO2, HOXA, TLX1* and *TLX3,* which predominantly occur in a mutually exclusive pattern. In our previous study, we used a supervised gene-expression profiling approach to cluster T-ALL patients with these chromosomal aberrations(1). Patients with *HOXA*, *TLX1* and *TLX3* abnormalities formed 3 separate T-ALL clusters. Patients having *TAL1* and/or *LMO2* rearrangements formed a single, fourth *TAL/LMO* cluster, explained by the fact that TAL1 and LMO2 participate in the same transcription complex and effect similar downstream pathways. Co-clustering of 45 additional patients that lack *TAL1, LMO2, HOXA, TLX1* or *TLX3* aberrations, led to the identification of 2 additional T-ALL genetic subgroups that are characterized by *NKX2-1/NKX2-2* or *MEF2C*-activating rearrangements(1). The *MEF2C*-deregulated subgroup overlapped with the early thymic progenitor ALL (ETP-ALL) subgroup as previous described by Dario Campana and co-workers(2). Nineteen of these 45 patient samples strongly co-clustered with *TAL1*- or *LMO2*-rearranged patients in supervised and unsupervised cluster analyses, pointing to a common pathogenic mechanism. These 19 cases were denoted as *TAL/LMO*-likes, and we hypothesized that these patients might harbor rearrangements involving factors homologous to *TAL1* or *LMO2*, or factors that participate in the TAL/LMO transcription complex. This hypothesis was confirmed when we identified translocations that involved *LMO3*(3), *LMO1* or *TAL2* in 3 of these *TAL/LMO*-like patients(4). A fourth patient had double translocations affecting *TAL2* and *LMO1* oncogenes.(4) To identify aberrations in the remaining 15 *TAL/LMO*-like patients, we screened for T-cell receptors driven translocations for which the translocation partner was unknown.

## DESIGN AND METHODS

*Patient material*

Viably frozen diagnostic bone marrow or peripheral blood samples from 117 pediatric T-ALL patients was used(1, 4) Clinical and immunophenotypic data were provided by the German Co-operative study group for childhood Acute Lymphoblastic Leukemia (COALL) and the Dutch Childhood Oncology Group (DCOG). The patients' parents or their legal guardians provided informed consent to use leftover material for research purposes in accordance with the declaration of Helsinki and the study was approved by the ethical committee of the Erasmus

Medical Center. Leukemic cells were isolated and enriched from these samples as previously described(5). All resulting samples contained ≥90% leukemic cells, as determined morphologically by May-Grünwald-Giemsa-stained cytospins (Merck, Darmstadt, Germany). Cytospin slide preparation and DNA and RNA extraction were performed as previously described(5).

*Fluorescent in-situ hybridization (FISH)*
FISH analysis was performed on cytospin slides using the TCRalpha/delta and TCRbeta split signal probes according to the manufacturer's protocol (DAKO, Glostrup, Denmark). Split signal FISH on the *LYL1* locus was performed using the following BAC clones as previously described(6): RP11-352L7, RP11-356L15.

*Ligation mediated PCR (LM-PCR) & Real-time quantitative PCR (RQ-PCR)*
LM-PCR for *TRB@* breakpoint hotspots *(TRB@D1* and *TRB@D2),* and RQ-PCR for *LYL1* were performed as described before(5, 7, 8). For LM-PCR, briefly, genomic DNA was digested with either one of four different restriction enzymes (PvuII, HincII, StuI, DraI) and ligated to adapters. Adaptor primers were then used in combination with *TRB@* loci specific primers to amplify the breakpoint region in two PCR rounds. For the detection of the reciprocal *LYL1-TRB@* breakpoint the following specific primers located near *LYL1* were used: first: 5'-CGG GCT GGA GGA GAG AAG-3', nested: 5'-GTG GCT GAC GAC GTG TAA TTT-3'.

## RESULTS AND DISCUSSION

A FISH strategy was performed to identify novel *TRB@*- or *TRAD@*-driven oncogenic rearrangements in 15 *TAL/LMO*-like patients. These 15 cases strongly clustered in hierarchical cluster analyses with T-ALL cases having *TAL1/2* and/or *LMO1/2/3* rearrangements (Figure 1A-B). One sample (#704), from a 7 year old male patient, showed a *TRB@* split signal pointing to a translocation that had not been revealed by karyotypic analysis (47,XY,+8[6]/46,XY[7]).

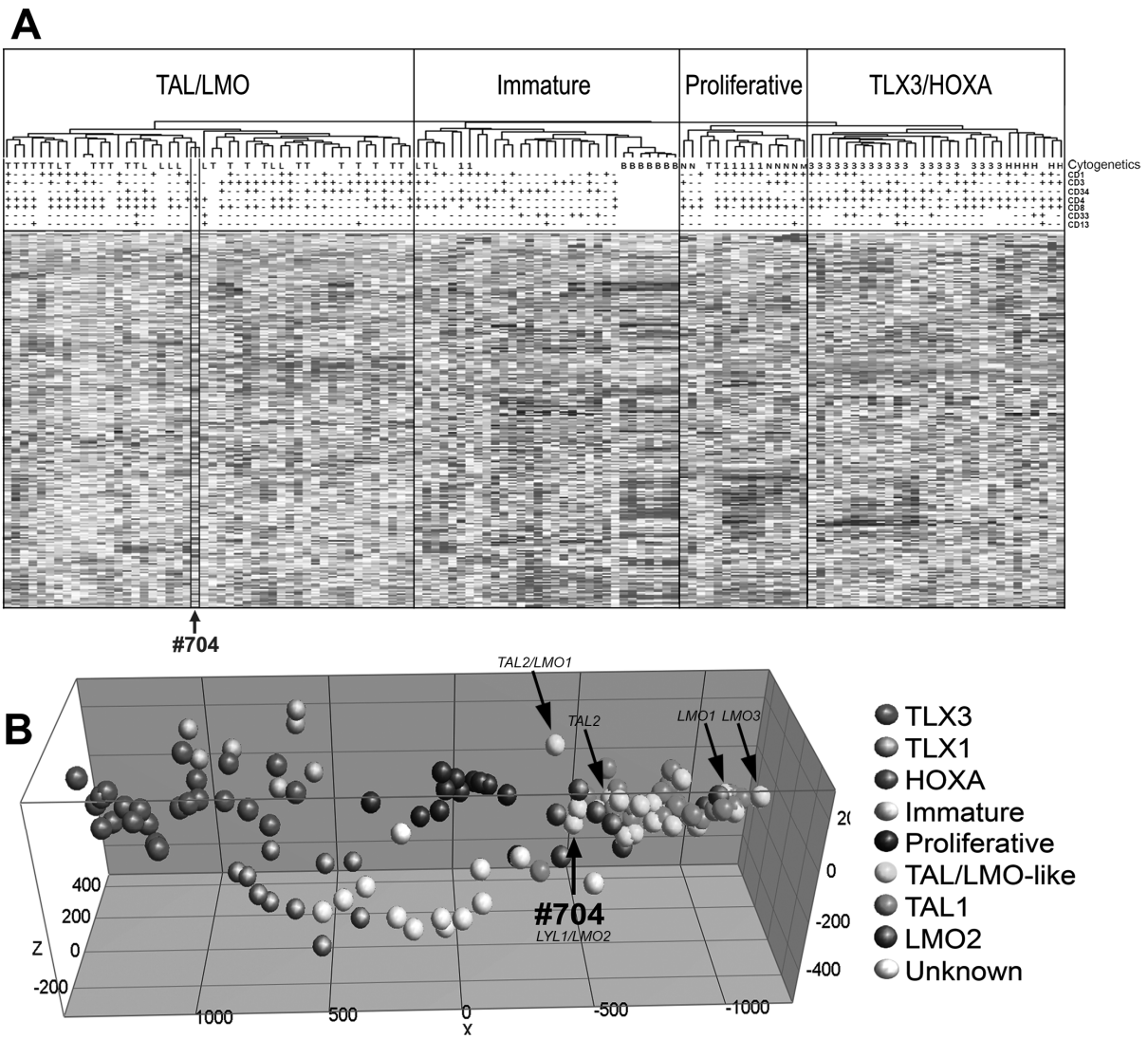



**Figure 1. Unsupervised and supervised hierarchical clustering of 117 pediatric T-ALL samples and 7 normal bone marrow samples and *LYL1* expression of unsupervised subgroups.** (A) Unsupervised hierarchical clustering of 117 pediatric T-ALL samples and 7 normal bone marrow samples (horizontal axis), according to microarray gene-expression (genes on vertical axis, gene names not shown)(1). Red corresponds to high expression, blue to low expression. CD surface markers are shown as present (>25%, "+"), absent (<25% "-") or not performed (blanc). Complete immunophenotype for #704: CD1-, CD2+, CD3-, CD4+, CD5-, CD7+, CD8+, cytoplasmatic CD3+, CD33-, CD14-, CD34-, CD71+, HLA_DR-, TDT+. Cytogentic abnormalities are annotated as follows: T: *SIL-TAL* deletion or *TAL1* translocation, L: *LMO2* translocation/deletion, 1: *TLX1* translocation, 3: *TLX3* translocation, B: normal bone marrow, N: *NKX2-1* translocation/inversion/duplication, M: *MYB* translocation, H: *HOXA* activating aberration (*CALM-AF10, SET-NUP, HOXA* inversion). Patient #704 is highlighted by a blue box. (B) Principal component analysis of supervised analyses of gene-expression data of 117 pediatric T-ALL samples(1). The position of the yellow dots representing *LMO1, TAL2, LMO3, TAL2/LMO1* rearranged cases and sample #704 (LYL1/LMO2) are indicated by arrows.

We then performed ligation-mediated PCR (LM-PCR) from two *TRB@* translocation hotspots (*TRB@D1* and *TRB@D2)* on DNA from this patient, and identified a translocation between *TRB@D1* and the last intron of the *nuclear factor I/X* (*NFIX)* gene for the derivative chromosome 7 (der(7)) (Figure 2A-B). The reciprocal breakpoint of the derivative chromosome 19 couples part of the last intron of the *NFIX* gene to an area between *TRB@J1* and *TRB@J2* (der(19), Figure 2A). The *LYL1* gene is located 240bp centromeric of *NFIX* and is therefore placed under the influence of the *TRB@* enhancer as consequence of this translocation (Figure 2A). *NFIX* is not expressed in any of our patient samples based on microarray data (raw fluorescent intensities<50, 5 probe sets, data not shown) indicating that changes in *NFIX* are not contributing to leukemogenesis. Positioning of the *LYL1* gene under the influence of the *TRB@* enhancer may explain the relative high expression level of *LYL1* in this patient (Figure 2C). FISH analysis of the *LYL1* locus on the remaining 14 *TAL/LMO*-like patients revealed no additional *LYL1* rearrangements.

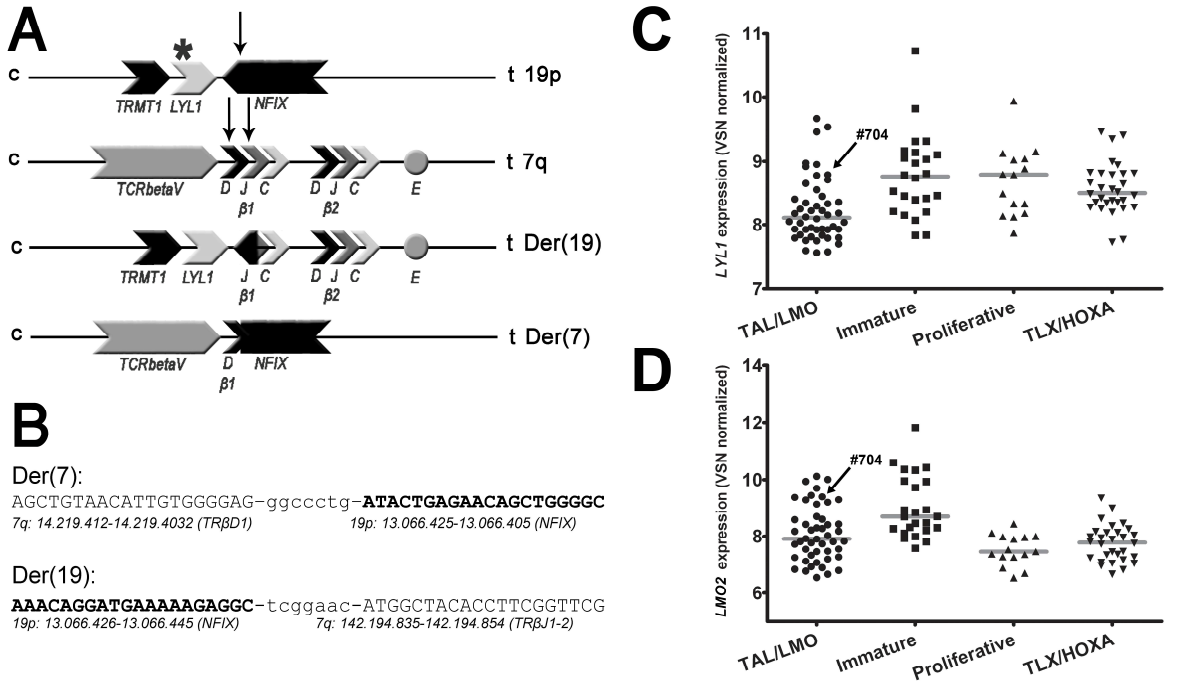



**Figure 2. Schematic overview of *LYL1-TRB@* translocation and breakpoint sequences.** **(**A) Schematic overview of breakpoint loci on germline chromosomes 19 and 7, and the der(19) and der(7). Arrows indicate approximate breakpoint locations, *, approximate breakpoint of previously described LYL1 translocation(23), c, centromeric side and t, telomeric side of chromosal region. (B) Reciprocal breakpoint sequences of the t(7;19)(q34;p13). In caps sequences corresponding to chromosomal regions as described below, in non-caps; randomly inserted nucleotides. *LYL1* (C) and *LMO2* (D) expression according to VSN normalized array data in the four subgroups as depicted in Figure 1A.

Various research groups including ours have reported that *LYL1* and *LMO2* are highly expressed in T-ALL patients with an immature immunophenotype(9-11), despite the fact that *LMO2* rearrangements that are also associated with ectopic *LMO2* expression are exclusively associated with the immunophenotypic more advanced *TAL/LMO* subgroup(1, 12). In another study(2), immature T-ALL cases were described with an early thymic progenitor expression profile that were associated with poor prognosis, and were denoted as ETP-ALL cases. Based on combined expression profiling and molecular-cytogenetic analyses, we recently identified an immature T-ALL subset that was predominantly characterized by rearrangements that activate the *MEF2C* oncogene(1). This subset could also be predicted by the ETP-ALL profile. For these immature, ETP-ALL cases, MEF2C has been shown to directly activate expression of *LYL1*, *LMO2* and *HHEX(1)* that may explain the high *LYL1* expression in immature T-ALL cases. So far, we and others have been unable to reveal *LYL1* rearrangements in these immature, ETP-ALL cases(1, 2, 9). In line with this, the single reported T-ALL case having a *LYL1* translocation had a mature (CD3+, CD1-, CD4+, CD8+ and CD34-) immunophenotype(13).

LYL1 is an basic helix-loop-helix (bHLH) transcription factor that shows 82% amino acid homology in the bHLH domein with TAL1(14). TAL1 and LYL1 also show overlapping expression patterns in hematopoietic development(15) and in some pathways they can exert identical functions(16). The strong co-clustering of this patient sample (#704) along with *TAL1*-rearranged T-ALL cases indicates that *LYL1* rearrangements elicit a similar expression profile as *TAL1* rearrangements during T-cell oncogenesis.

Using array-CGH, we further identified a small del(11)(p12p13) near the *LMO2* locus in this patient #704 ((12), data not shown), accompanied by ectopic *LMO2* expression (Figure 2D). No copy number changes were found at the *TAL1* locus. *LMO2* rearrangements (translocations or del(11)(p12p13)) occur in approximately 9% of pediatric T-ALL(12) and have been exclusively associated with the *TAL/LMO* subgroup(1, 12). The identification of a *LYL1* translocation as well as an *LMO2* rearrangement in this *TAL/LMO*-like patient implies that LYL1 and LMO2 synergize in T-cell oncogenesis. Other incidental cases harbor *TAL1/2* as well as *LMO1/2* aberrations(4), and 2 additional patients out of 55 *TAL/LMO* patients (including the *TAL/LMO*-like patients) as present in our T-ALL cohort (n=117) had combined rearrangements of *TAL* and *LMO* family members: one case had a *SIL-TAL1* deletion and the *LMO2*-activating del(11)(p12p13), and one had a

*TAL2/TRB@* translocation in combination with an *LMO1/TRAD@* translocation(4). This points to strong synergistic effects between these oncogenic family members in line with their participation in similar transcriptional complexes(17-19). *Lmo1/Lmo2* and *Tal1* have also been shown to synergize to T-cell leukemogenesis in mice studies(17, 20-22).

To conclude, we suggest that *LYL1* rearranged cases are not part of the immature, ETP-ALL subgroup, but belong to the *TAL/LMO* subgroup. *LYL1* translocations fulfill a *TAL1*-like role that can synergize with *LMO2* aberrations in T-cell oncogenesis.

**ACKNOWLEDGEMENTS**

This work was supported by a grant from the Dutch Cancer Society (KWF-EMCR 2006-3500).

**AUTHORSHIP AND DISCLOSURES**

IH performed FISH, LM-PCR and wrote manuscript, AL assisted in LM-PCR and participated in discussion, JB-G performed FISH analysis, RP designed study, assisted in discussion and wrote manuscript, JPPM is principal investigator, designed study and wrote manuscript.

Disclosure: authors have nothing to disclose

## REFERENCES:


1. Homminga I, Pieters R, Langerak AW, de Rooi JJ, Stubbs A, Verstegen M, et al. Integrated Transcript and Genome Analyses Reveal NKX2-1 and MEF2C as Potential Oncogenes in T Cell Acute Lymphoblastic Leukemia. Cancer cell. 2011;19(4):484-97.
2. Coustan-Smith E, Mullighan CG, Onciu M, Behm FG, Raimondi SC, Pei D, et al. Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia. The lancet oncology. 2009;10(2):147-56.
3. Simonis M, Klous P, Homminga I, Galjaard RJ, Rijkers EJ, Grosveld F, et al. High-resolution identification of balanced and complex chromosomal rearrangements by 4C technology. Nat Methods. 2009;6(11):837-42.
4. Van Vlierberghe P, van Grotel M, Tchinda J, Lee C, Beverloo HB, van der Spek PJ, et al. The recurrent SET-NUP214 fusion as a new HOXA activation mechanism in pediatric T-cell acute lymphoblastic leukemia. Blood. 2008;111(9):4668-80.
5. van Grotel M, Meijerink JP, Beverloo HB, Langerak AW, Buys-Gladdines JG, Schneider P, et al. The outcome of molecular-cytogenetic subgroups in pediatric T-cell acute lymphoblastic leukemia: a retrospective study of patients treated according to DCOG or COALL protocols. Haematologica. 2006;91(9):1212-21.
6. Van Vlierberghe P, Beverloo HB, Buijs-Gladdines J, van Wering ER, Horstmann M, Pieters R, et al. Monoallelic or biallelic LMO2 expression in relation to the LMO2 rearrangement status in pediatric T-cell acute lymphoblastic leukemia. Leukemia. 2008;22(7):1434-7.
7. Przybylski GK, Dik WA, Wanzeck J, Grabarczyk P, Majunke S, Martin-Subero JI, et al. Disruption of the BCL11B gene through inv(14)(q11.2q32.31) results in the expression of BCL11B-TRDC fusion transcripts and is associated with the absence of wild-type BCL11B transcripts in T-ALL. Leukemia. 2005;19(2):201-8.
8. van Dongen JJ, Langerak AW, Bruggemann M, Evans PA, Hummel M, Lavender FL, et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. Leukemia. 2003;17(12):2257-317.
9. Ferrando AA, Neuberg DS, Staunton J, Loh ML, Huard C, Raimondi SC, et al. Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. Cancer cell. 2002;1(1):75-87.
10. Soulier J, Clappier E, Cayuela JM, Regnault A, Garcia-Peydro M, Dombret H, et al. HOXA genes are included in genetic and biologic networks defining human acute T-cell leukemia (T-ALL). Blood. 2005;106(1):274-86.
11. van Grotel M, Meijerink JP, van Wering ER, Langerak AW, Beverloo HB, Buijs-Gladdines JG, et al. Prognostic significance of molecular-cytogenetic abnormalities in pediatric T-ALL is not explained by immunophenotypic differences. Leukemia. 2008;22(1):124-31.
12. Van Vlierberghe P, van Grotel M, Beverloo HB, Lee C, Helgason T, Buijs-Gladdines J, et al. The cryptic chromosomal deletion del(11)(p12p13) as a new activation mechanism of LMO2 in pediatric T-cell acute lymphoblastic leukemia. Blood. 2006;108(10):3520-9.

13. Smith SD, Morgan R, Gemmell R, Amylon MD, Link MP, Linker C, et al. Clinical and biologic characterization of T-cell neoplasias with rearrangements of chromosome 7 band q34. Blood. 1988;71(2):395-402.
14. Mellentin JD, Smith SD, Cleary ML. lyl-1, a novel gene altered by chromosomal translocation in T cell leukemia, codes for a protein with a helix-loop-helix DNA binding motif. Cell. 1989;58(1):77-83.
15. Giroux S, Kaushik AL, Capron C, Jalil A, Kelaidi C, Sablitzky F, et al. lyl-1 and tal-1/scl, two genes encoding closely related bHLH transcription factors, display highly overlapping expression patterns during cardiovascular and hematopoietic ontogeny. Gene Expr Patterns. 2007;7(3):215-26.
16. Souroullas GP, Salmon JM, Sablitzky F, Curtis DJ, Goodell MA. Adult hematopoietic stem and progenitor cells require either Lyl1 or Scl for survival. Cell stem cell. 2009;4(2):180-6.
17. Wadman I, Li J, Bash RO, Forster A, Osada H, Rabbitts TH, et al. Specific in vivo association between the bHLH and LIM proteins implicated in human T cell leukemia. The EMBO journal. 1994;13(20):4831-9.
18. Wadman IA, Osada H, Grutz GG, Agulnick AD, Westphal H, Forster A, et al. The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. The EMBO journal. 1997;16(11):3145-57.
19. El Omari K, Hoosdally SJ, Tuladhar K, Karia D, Vyas P, Patient R, et al. Structure of the leukemia oncogene LMO2: implications for the assembly of a hematopoietic transcription factor complex. Blood. 2011;117(7):2146-56.
20. Larson RC, Lavenir I, Larson TA, Baer R, Warren AJ, Wadman I, et al. Protein dimerization between Lmo2 (Rbtn2) and Tal1 alters thymocyte development and potentiates T cell tumorigenesis in transgenic mice. The EMBO journal. 1996;15(5):1021-7.
21. Draheim KM, Hermance N, Yang Y, Arous E, Calvo J, Kelliher MA. A DNA-binding mutant of TAL1 cooperates with LMO2 to cause T cell leukemia in mice. Oncogene. 2010.
22. Aplan PD, Jones CA, Chervinsky DS, Zhao X, Ellsworth M, Wu C, et al. An scl gene product lacking the transactivation domain induces bony abnormalities and cooperates with LMO1 to generate T-cell malignancies in transgenic mice. The EMBO journal. 1997;16(9):2408-19.
23. Cleary ML, Mellentin JD, Spies J, Smith SD. Chromosomal translocation involving the beta T cell receptor gene in acute leukemia. The Journal of experimental medicine. 1988;167(2):682-7.

# CHAPTER 3

## Cooperative genetic defects in *TLX3* rearranged pediatric T-ALL

P Van Vlierberghe[1], I Homminga[1], L Zuurbier[1], J Gladdines-Buijs[1], E R van Wering[2], M Horstmann[3], H B Beverloo[4], R Pieters[1] and J P P Meijerink[1]

1. Department of Pediatric Oncology/Hematology, Erasmus MC/Sophia Children's Hospital, Rotterdam, The Netherlands
2. Dutch Childhood Oncology Group (DCOG), The Hague, The Netherlands
3. German Co-operative Study Group for childhood acute lymphoblastic leukemia (COALL), Hamburg, Germany
4. Department of Clinical Genetics, Erasmus MC, Rotterdam, The Netherlands

*Leukemia, 2008: 22; 762–770*

**ABSTRACT**

T-cell acute lymphoblastic leukemia (T-ALL) is an aggressive neoplastic disorder, in which multiple genetic abnormalities cooperate in the malignant transformation of thymocytes. About 20% of pediatric T-ALL cases are characterized by *TLX3* expression due to a cryptic translocation t(5;14)(q35;q32). Although a number of collaborating genetic events have been identified in *TLX3* rearranged T-ALL patients (*NOTCH1* mutations, *p15*/*p16* deletions, *NUP214-ABL1* amplifications), further elucidation of additional genetic lesions could provide a better understanding of the pathogenesis of this specific T-ALL subtype. In this study, we used array-CGH to screen *TLX3* rearranged T-ALL patients for new chromosomal imbalances. Array-CGH analysis revealed five recurrent genomic deletions in *TLX3* rearranged T-ALL, including del(1)(p36.31), del(5)(q35), del(13)(q14.3), del(16)(q22.1) and del(19)(p13.2). From these, the cryptic deletion, del(5)(q35), was exclusively identified in about 25% of *TLX3* rearranged T-ALL cases. In addition, 19 other genetic lesions were detected once in *TLX3* rearranged T-ALL cases, including a cryptic *WT1* deletion and a deletion covering the *FBXW7* gene, an U3-ubiquitin ligase that mediates the degradation of NOTCH1, MYC, JUN and CyclinE. This study provides a genome-wide overview of copy number changes in *TLX3* rearranged T-ALL and offers great new challenges for the identification of new target genes that may play a role in the pathogenesis of T-ALL.

## INTRODUCTION

T-cell acute lymphoblastic leukemia (T-ALL) is an aggressive disorder of T-cells, and represents about 15% of pediatric ALL cases. T-ALL is characterized by a rapid progression of disease and shows a 30% relapse rate.[1] Over the last decade, a large number of new genomic aberrations were identified in T-ALL, including chromosomal translocations (involving the genes *TAL1*, *LYL1*, *LMO1*, *LMO2*, *TLX1/HOX11*, *TLX3/HOX11L2*, *MYB*, *Cyclin D2*), deletions (*SIL-TAL1*, del(6q), del(9)(p21), del(11)(p12p13)), amplifications (*NUP214-ABL1*), duplications (*MYB*) and mutations (*RAS*, *NOTCH1*).[2, 3, 4, 5, 6, 7, 8, 9, 10, 11] Several of these abnormalities represent unique and mutually exclusive aberrations possibly delineating distinct T-ALL subgroups. Others occur in combination with various of these subgroups, for example, the del(9)(p21) that includes the *CDKN2A/p15* and *CDKN2B/p16* loci both deregulating the cell cycle.[3, 12] Also *NOTCH1* activation mutations are present in more than half of all T-ALL cases of all subgroups.[9] The genetic defects as identified in T-ALL so far target different cellular processes, including cell cycle regulation, T-cell differentiation, proliferation and survival. It is hypothesized that these genetic events cooperate in the leukemic transformation of thymocytes.[13]

*TLX3* is a homeobox gene that is not expressed in normal T-cell development. In T-ALL patients, it becomes aberrantly activated due to the cryptic translocation, t(5;14)(q35;q32), mostly juxtaposing *TLX3* to the *BCL11B* gene. *BCL11B* is normally expressed during T-cell maturation.[14] Some alternative *TLX3* translocations have been described including the t(5;14)(q32;q11) juxtaposing *TLX3* to the *TCRα/δ* locus,[15] and the t(5;7)(q35;q21) coupling *TLX3* to the *CDK6* gene.[16] Although there is a clear relationship between the presence of *TLX3* translocations and *TLX3* expression levels,[17] incidental *TLX3* expression has been described in the absence of chromosomal abnormalities,[18, 19] suggesting that alternative mechanisms for *TLX3* activation exist in T-ALL. Conflicting data have been published about the relation between *TLX3* expression and treatment outcome. In some studies, *TLX3* rearranged T-ALL patients showed a poor prognosis, whereas in other studies *TLX3* translocations had no effect on outcome or was even associated with an improved outcome.[17, 20, 21] These discrepancies have not been clarified thus far, but may be therapy-dependent.

In this study, we used microarray-based comparative genome hybridization (array-CGH) to screen *TLX3* rearranged pediatric T-ALL patients for new chromosomal imbalances that could provide further insight in the development of *TLX3*-mediated T-cell leukemia.

## DESIGN AND METHODS

### Patients

Viably frozen diagnostic bone marrow or peripheral blood samples from 146 pediatric T-ALL patients were provided by the Dutch Childhood Oncology Group and the German Co-operative Study Group for childhood acute lymphoblastic leukemia.[17] The patients and patients' parents or their legal guardians provided informed consent to use leftover material for research purposes. Leukemic cells were isolated and enriched from these samples as previously described,[8] and genomic DNA and total cellular RNA were isolated as described before.[8]

### Quantitative real-time RT-PCR (RQ-PCR).

cDNA synthesis and RQ-PCR in an ABI 7700 sequence detection system (Applied Biosystems, Foster City, CA, USA) was used to quantify the expression levels of *TLX3* transcripts relative to the endogenous housekeeping gene glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*), as described previously.[17] *NUP214–ABL1* fusions were determined as previously described.[4]

### Oligo array-CGH

Oligo array-CGH analysis was performed on the human genome CGH Microarray 44A (Agilent Technologies, Palo Alto, CA, USA) according to the manufacturer's protocol, as previously described.[8] Microarray images were analyzed using feature extraction software (version 8.1, Agilent) and the data were subsequently imported into array-CGH analytics software v3.1.28 (Agilent). For the detection of copy number abnormalities, we have used a *Z*-score cutoff value of 3. All copy number aberrations were compared to the database of genomic variants (http://projects.tcag.ca/variation) and all genomic regions previously linked to copy number variations[22] were not included in Table 1.

### Fluorescence *in situ* hybridization

Fluorescence *in situ* hybridization (FISH) was performed using a standard procedure, as described previously.[17] *TLX3* translocations were determined using the *TLX3-U/TLX3-D* translocation probes (DakoCytomation, Glostrup, Denmark). BAC probes RP11-299P16 and RP11-98C11 (BACPAC resources, Oakland, CA, USA) were used to confirm the presence of Wilms' tumor 1 (*WT1*) deletions, whereas RP11-300I24 and RP11-650G8 were used to confirm the *FBXW7* deletion. RP11-1072I20 (*RANBP17*/*TLX3*), RP11-10N18 (*RANBP17*) and RP11-117L6 (downstream of *TLX3*) as well as CTD-2243O22 (5qter) (Invitrogen, Breda, the Netherlands) were used to further characterize the deletion, del(5)(q35).

**Real-time quantification of DNA copy number**

Deletion analysis was performed using real-time quantitative PCR of the *NSD1* gene relative to the internal control gene, albumin, as previously described.[23]

**Mutation analysis**

For the detection of *WT1* mutations, the purified DNA was subjected to 40 cycles of PCR of 15 s at 95 °C and 1 min at 60 °C, using forward primer 5'-AAG CCTCCCTTCCTCTTACTCT-3' and reverse primer 5'-TGGGTCCTTAG CAGTGTGAGA-3' for *WT1* exon 7. *FBXW7* mutation detection was performed using forward primer 5'-TTTTCCAGTGTCTGAGAACAT-3' and reverse primer 5'-CCCAAATTCACCAATAATAGA-3' for exon 9, forward primer 5'-TAAA CGTGGGTTTTTTTGTT-3' and reverse primer 5'-TCAGCAATTTGA CAGTGATT-3' for exon 10 and forward primer 5'-GGACATGGGTTTCT AAATATGTA-3' and reverse primer 5'-CTGCACCACTGAGAACAAG-3' for exon 12, using similar PCR conditions as described above. *NOTCH1* mutation screening in T-ALL was performed as previously described.[24] PCR products were purified by standard methods and directly sequenced from both strands. The sequence data were analyzed using Seqscape V2.5 (Applied Biosystems).

**RESULTS**

**Collaborating genetic events in *TLX3* rearranged T-ALL**

In our previous study, we screened a large pediatric T-ALL cohort (*n*=146) for *TLX3* rearrangements using FISH and identified 29 of 146 (19%) rearranged cases,[17] in line with previous studies.[18, 19] All *TLX3* rearranged cases uniquely expressed *TLX3* whereas other T-ALL cases were negative.[17] To identify additional genetic abnormalities that may cooperate with *TLX3* expression during T-cell leukemogenesis, we performed array-CGH analysis to detect genomic amplifications or deletions on those *TLX3* rearranged T-ALL cases for which material was available (*n*=21). All genomic deletions and/or amplifications as identified by array-CGH are summarized in Table 1, except for known polymorphic copy number variations.[22] Genomic deletions are more abundant as compared to amplifications, as only two regions of genomic amplification in contrast to 22 regions of genomic deletion were identified in our *TLX3* rearranged patient cohort. To confirm if these additional aberrations are truly *TLX3* specific, we analyzed whether these additional abnormalities were also identified in a large-scale T-ALL array-CGH study (*n*=85, unpublished data) of non-*TLX3* rearranged T-ALL patients (Table 1). Other known T-ALL-specific genetic aberrations were

determined using an RT–PCR or PCR and sequencing strategy or using FISH, and included *NOTCH1* mutations,[24] *NUP214-ABL1* amplifications[4] and *p15/p16* deletions (Supplementary Table 1).

**Table 1. Novel genetic lesions in *TLX3* positive pediatric T-ALL**

| Chromosome Band | Gain Loss | Start (Mb) | End (Mb) | Size (Mb) | Patient ID(s) (% of cases) | Gene(s) in region |
|---|---|---|---|---|---|---|
| 1p36.31 | Loss | 5.95 | 7.21 | 1.26 | 2738, 2112, 2757, 585 (19) | 18 genes including ***HES2, HES3 and CAMTA1*** |
| 1p36.12-1p36.13 | Loss | 16.55 | 24.91 | 8.36 | 2757 (5) | > 50 genes |
| 1p34.2-1p34.3 | Loss | 35.99 | 39.66 | 3.67 | 2757 (5) | > 30 genes |
| 2p23.3-2p24.1 | Loss | 20.50 | 24.50 | 4.00 | 9012 (5) | 12 genes including ***TP53I3*** and ***NCOA1*** |
| 2q37.1 | Loss | 233.30 | 234.45 | 1.15 | 2112 (5) | 12 genes; non leukemia associated |
| 3q13.32-3q21.2 | Loss | 119.90 | 126.40 | 6.50 | 585 (5) | > 50 genes, ***miR-198*** |
| 3q26.2-3q26.31 | Loss | 172.10 | 176.40 | 4.30 | 585 (5) | 10 genes, ***miR-569*** |
| 4q31.3-4q32.1 | Loss | 153.08 | 155.51 | 2.43 | 2786 (5) | 12 genes including ***FBXW7*** |
| 5q35.2-5qter | Loss | 172.60 | 180.90 | 8.30 | 2112, 2640, 2650, 222, 9012 (24) | 25 genes |
| 6q25.1-6qter | Loss | 149.94 | 170.80 | 20.86 | 222 (5) | > 100 genes |
| 7q31.33-7q36.2 | Loss | 125.30 | 153.99 | 28.69 | 9012 (5) | > 100 genes, 11 microRNAs |
| 8 | Gain | 0 | 146.25 | 146.25 | 2105 (5) | complete chromosome 8 |
| 9p24.1-9p24.2 | Loss | 2.71 | 5.12 | 2.41 | 2100 (5) | 9 genes and *mIR-101-2*; breakpoint in ***JAK2*** |
| 9q21.13-9q31.1 | Loss | 67.12 | 119.72 | 52.60 | 222 (5) | > 100 genes and 10 microRNAs |
| 10p15.2-10p15.3 | Loss | 1.08 | 3.20 | 2.12 | 9858 (5) | *WDR37*, *ADARB2*, *PFKP* and *PITRM1* |
| 10p11.23-10p12.1 | Loss | 27.99 | 30.05 | 2.06 | 9012 (5) | 7 genes and *mIR-604* |
| 11p13 | Loss | 32.00 | 33.50 | 1.50 | 2723 (5) | 10 genes including ***WT1*** |
| 11q21-11q22.3 | Loss | 94.42 | 107.76 | 13.34 | 1179 (5) | > 20 genes, breakpoint in the ***ATM*** gene |
| 12p13.1-12p13.2 | Loss | 12.50 | 13.10 | 0.60 | 9858 (5) | 12 genes including ***CDKN1B***, *miR-613/614* |
| 13q14.3 | Loss | 49.45 | 50.35 | 0.90 | 2112, 2723 (9) | *KCNRG*, ***miR-15/16a***, *DLEU7* |
| 16q22.1 | Loss | 66.20 | 66.60 | 0.40 | 2100, 2112, 9012 (14) | 12 genes including *CTCF* |
| 17q11.2 | Loss | 26.08 | 27.47 | 1.39 | 2780 (5) | 11 genes including ***NF1,*** *mIR-193a* / *mIR-365-2* |
| 19p13.2 | loss | 10.75 | 11.90 | 1.15 | 222, 378 (9) | 33 genes, *miR-199a* |
| 20p12.3-20pter | gain | 0 | 5.94 | 5.94 | 9012 (5) | > 50 genes and *mIR-103-2* |

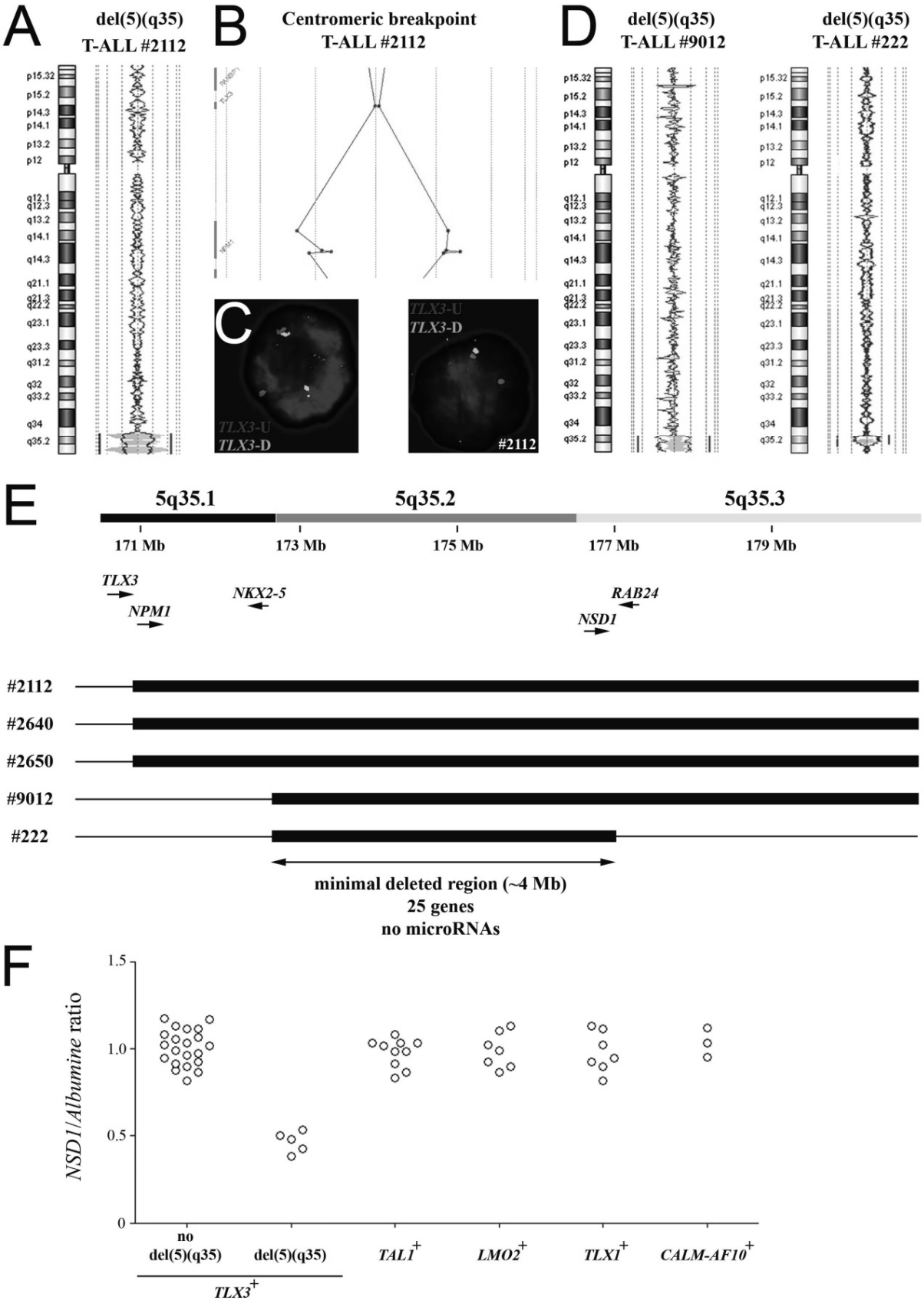
**The cryptic deletion, del(5)(q35), is associated with *TLX3* expression in T-ALL**

The most frequent recurrent genetic abnormality identified in *TLX3* rearranged cases was a heterozygous deletion at band 5q35, which was present in 5 of 21 (24%) *TLX3* rearranged T-ALL cases. The deletional area differed in size among cases. In three cases (nos. 2112, 2640 and 2650), the deletion started just downstream of the *TLX3* gene, as shown by the normal hybridization pattern of the

*TLX3* probe, and loss of all four array-CGH probes covering the nucleophosmin (*NPM1*) gene located 80 kb telomeric of *TLX3* (Figures 1a and b). For these cases, FISH analysis using the *TLX3-U/TLX3-D* translocation probes (Supplementary Figure 1) confirmed the presence of this cryptic deletion (Figure 1c). In the other two cases (nos. 9012 and 222), the deletion started upstream of *NKX2-5* (Figures 1d and e). However, gene expression array data revealed no *NKX2-5* expression in any of these five cases (data not shown). For cases nos. 2112, 2640, 2650 and 9012, the deletion seemed to include the complete telomeric region (Figure 1e). For case no. 222, the terminal breakpoint was situated downstream of the *NSD1* gene. Therefore, the minimal deleted region at 5q35 for these five cases is about 4 Mb in size and contains 30 known genes including the *NSD1* gene. Gene expression array data revealed no difference in *NSD1* expression levels between del(5)(q35) positive and negative T-ALL patients (data not shown).

Quantitative PCR analysis of the *NSD1* gene, which is present in the minimal deleted region, on 26 TLX3 rearranged T-ALL cases and 27 TLX3 negative cases (including *TAL1* rearranged, *LMO2* rearranged, *TLX1* rearranged and *CALM-AF10* positive cases), confirmed a one-copy *NSD1* loss in all *TLX3* rearranged T-ALL cases having the cryptic del(5)(q35) deletion (Figure 1f). None of the non-*TLX3* rearranged cases showed loss of NSD1, indicating that none of these had a similar del(5)(q35).

**Figure 1. (page 74) The recurrent cryptic deletion, del(5)(q35), in *TLX3* rearranged pediatric T-cell acute lymphoblastic leukemia (T-ALL).** (**a**) Chromosome 5 ideogram and corresponding oligo microarray-based comparative genome hybridization (array-CGH) plot of case DNA:control DNA ratios (blue tracing) versus the dye-swap experiment (red tracing) for T-ALL cases 2112. Hybridization signals around the -2X or +2X lines represent loss of the corresponding region in the case DNA. (**b**) Detailed analysis of the centromeric breakpoint of the deletion in case 2112. (**c**) Dual-color fluorescence *in situ* hybridization (FISH) analysis on interphase cells of case 9858 (left panel) and case 2640 (right panel) using the *TLX3-U* (Red) and *TLX3-D* (green) translocation probe set. Case 9858 showed a split signal, indicative for a *TLX3* translocation, whereas case 2640 showed loss of the *TLX3-D* (green) signal. (**d**) Similar chromosome 5 ideograms as in (**a**) for T-ALL cases 9012 and 222. (**e**) Schematic overview of the minimal deleted region on chromosomal band 5q35 for the 5 TLX3 rearranged T-ALL cases showing a del(5)(q35). Depicted genome positions and gene locations are based on the UCSC Genome Browser at http://genome.ucsc.edu/. (**f**) Quantitative PCR analysis of *NSD1*, present in the minimal deleted region, on 26 TLX3 rearranged T-ALL cases and 27 TLX3 negative cases.


A
del(5)(q35)
T-ALL #2112
B
Centromeric breakpoint
T-ALL #2112
C
TLX3-U
TLX3-D
TLX3-U
TLX3-D
#2112
D
del(5)(q35)
T-ALL #9012
del(5)(q35)
T-ALL #222
E
5q35.1
5q35.2
5q35.3
171 Mb
173 Mb
175 Mb
177 Mb
179 Mb
TLX3
NPM1
NKX2-5
RAB24
NSD1
#2112
#2640
#2650
#9012
#222
minimal deleted region (~4 Mb)
25 genes
no microRNAs
F
NSD1/Albumine ratio
1.5
1.0
0.5
0
no del(5)(q35)
del(5)(q35)
TAL1+
LMO2+
TLX1+
CALM-AF10+
TLX3+

Next, we studied whether the deletions that started just downstream of the *TLX3* gene (nos. 2112, 2640 and 2650), truly represented cryptic 5q35 deletions or rather corresponded to unbalanced *TLX3* translocations. Therefore, we performed FISH analysis using BAC clones covering the RANBP17/TLX3 breakpoint region and the telomeric end of chromosome 5 (Supplementary Figure 2A). For case 2650, FISH analysis, using RP11-1072I20 and RP11-10N18, revealed two fusion signals indicating that both *TLX3* gene copies were normally present (Supplementary Figure 2A,B). In addition, FISH analysis using RP11-1072I20, RP11-117L6 and CTD-2243O22 confirmed the presence of a del(5)(q35.1q35.3) (Supplementary Figure 2A,B). In contrast, FISH analysis on case 2640 (Supplementary Figure 2C) revealed an additional RP11-1072I20 (RANBP17/TLX3) hybridization signal, indicative for a cryptic unbalanced chromosomal rearrangement involving the RANBP17/TLX3 loci. FISH analysis for the known translocation partners *TCRα/δ*, *TCRβ* and BCL11B showed that these loci were not involved in this chromosomal rearrangement (Supplementary Figure 2C), indicating that patient 2640 has a novel variant of *TLX3* rearrangement with subsequent loss of 5q35.1. For case 2112, no material was left to perform additional *TLX3*-specific FISH analyses.

**Other recurrent genomic deletions in *TLX3* rearranged T-ALL**

Besides, the cryptic deletion, del(5)(q35), four other recurrent genetic abnormalities were identified in various *TLX3* rearranged T-ALL cases (Table 1 and Supplementary Table 1). At chromosome 1, an identical cryptic deletion of ~1 Mb was detected at chromosomal band 1p36 in three cases (nos. 2738, 2112 and 585) (Supplementary Figure 3A). This deletion area was also comprised in a larger deletion in case no. 2757 that in addition demonstrated multiple deletions on chromosome 1p (Table 1, Supplementary Figure 3A). The minimal deleted area on 1p36 for these four cases comprised 18 genes, including *HES2*, *HES3* and chromodomain helicase DNA binding domain 5 (*CHD5*) (Supplementary Figure 3B). The centromeric breakpoints of the del(1)(p36.31) in cases nos. 2738, 2112 and 585 all clustered in the *CAMTA1* gene (Supplementary Figure 3A).

Cryptic deletions of chromosome 13q were identified in 2 *TLX3* rearranged T-ALL cases (nos. 2112 and 2723, Supplementary Figure 3C). These deletions differed in size and the minimal deleted region contained the microRNA cluster, *miR-15/miR-16a* (Supplementary Figure 3D).

Three T-ALL cases showed cryptic deletions at chromosomal band 16q22.1 (nos. 2100, 2112 and 9012) (Supplementary Figure 4A,B; Table 1). This del(16)(q22.1) seemed identical in two cases (nos. 2100 and 9012), but was smaller (~400 kb) in a third case (no. 2112). The minimal deleted area comprised 12 genes, and included the *CTCF* gene.

Finally, two other cases contained similar deletions at 19p13.2 (nos. 222 and 378) covering a region of approximately 1.4 Mb (Supplementary Figure 4C,D; Table 1) that covers 33 genes and the microRNA gene *miR-199a*.

**Genetic abnormalities identified in single TLX3 rearranged T-ALL cases**

Apart from the recurrent abnormalities, other genetic abnormalities were observed in single cases (Table 1 and Supplementary Table 1) and occasionally contained known tumor suppressor genes (*FBWX7*, *WT1*, *ATM*, *p27KIP1*, *NF1*). None of these abnormalities have been reported as normal copy number variation in the healthy population.[22]

One case (no. 2786) showed a cryptic deletion on the long arm of chromosome 4 of about 2.5 Mb in size, del(4)(q31.3q32.1), which contained among others the *FBXW7* gene (Figure 2). The loss of one *FBXW7* gene copy in this case was confirmed using FISH (Figure 2b). *FBXW7* mutations have recently been identified in 8–30% of primary T-ALL samples.[25, 26, 27, 28] Therefore, we screened case no. 2786 for the currently known *FBXW7* mutations. This analysis revealed no additional *FBXW7* mutation on the remaining allele of this case.

Other rearrangements identified in single *TLX3* rearranged cases included a cryptic deletion, del(12)(p13.1p13.2), including the *CDKN1B/p27/KIP1* gene (no. 9858) and a cryptic deletion, del(17)(q12), including the *NF1* gene (no. 2780).[29] In one case (no. 2100), the breakpoint of a cryptic deletion on chromosome 9, del(9)(p24.1p24.2), was situated in *JAK2*. In another case (no. 1179), the breakpoint of a cryptic deletion on chromosome 11, del(11)(q21q22.3), was located in the *ATM* gene.

**WT1 inactivation in pediatric T-ALL**

Another abnormality that was identified in a single *TLX3* rearranged case (no. 2723) was a cryptic deletion of about 1.5 Mb in size, del(11)(p13p13), and included the *WT1* gene (Figure 3a). Because conflicting data have been reported on the role of WT1 as a tumor suppressor and/or oncogene in human leukemias,[30] we wondered whether *WT1* was indeed the target gene of this genomic deletion. The telomeric breakpoint of this deletion was situated downstream of the *WT1* gene,

whereas the centromeric breakpoint was located downstream of the *CD59* gene (Figure 3a). FISH analysis confirmed the one-copy loss of *WT1* in this case (Figure 3b). To investigate *WT1* inactivation in this case, the remaining *WT1* allele was analyzed for the presence of inactivation mutations. A small frameshift mutation (delCinsTAG) was identified in exon 7, disrupting the *WT1* coding region (Figure 3c).

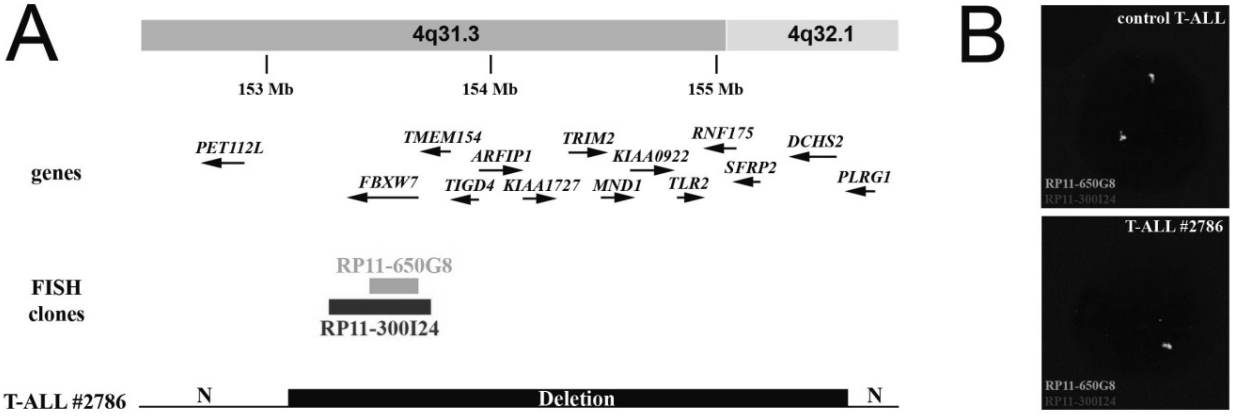



**Figure 2. *FBXW7* deletion in pediatric T-cell acute lymphoblastic leukemia (T-ALL).** (**a**) Schematic overview of the chromosomal deletion, del(4)(q31.3q32.1), as detected in case 2786. Genomic positions of genes situated in this chromosomal region and bacterial artificial chromosome (BAC) clones used for fluorescence *in situ* hybridization (FISH) analysis are depicted. (**b**) FISH analysis using RP11-650G8 (green) and RP11-300I24 (red) confirms the presence of the del(4)(q31.3q32.1) in case 2786.

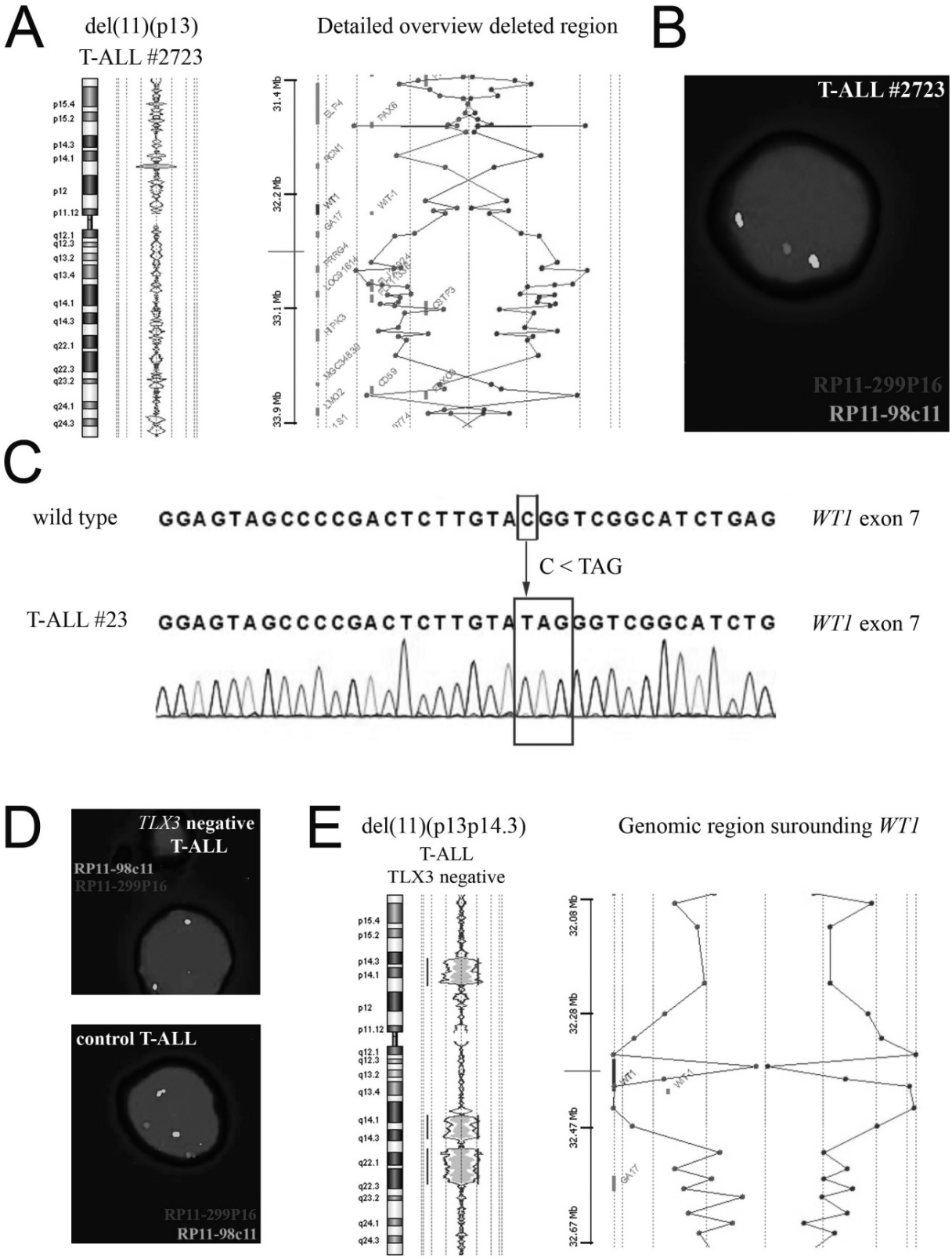
To investigate whether *WT1* inactivation is restricted to *TLX3* rearranged T-ALL cases, we performed additional *WT1*-specific FISH analysis on 25 *TLX3* negative pediatric T-ALL cases. This revealed one additional case in which FISH analysis revealed a loss of both *WT1* gene copies (Figure 3d). Subsequent array-CGH analysis confirmed the presence of a large monoallelic deletion on the short arm of chromosome 11, del(11)(p13p14.3), in combination with a loss of the genomic region surrounding the *WT1* gene on the other allele (Figure 3e). Gene expression array data showed that *WT1* expression was virtually absent in this T-ALL case showing a homozygous *WT1* deletion (Supplementary Figure 5). For case no. 2723, *WT1* was equally expressed compared to *WT1* wild-type cases of all T-ALL subgroups (*TAL1*, *LMO2*, *TLX3*, *TLX1* and unknown) (Supplementary Figure 5).

**Figure 3. (page 79) Wilms' tumor 1 (*WT1*) inactivation in pediatric T-cell acute lymphoblastic leukemia (T-ALL).** (**a**) Chromosome 11 ideogram and oligo microarray-based comparative genome hybridization (array-CGH) plot for the deletion, del(11)(p13), as detected in case 2723 (left panel). The right panel shows a detailed overview of the deleted region for this 11p13 deletion. (**b**) Fluorescence *in situ* hybridization (FISH) analysis using RP11-98C11 (green) and RP11-299P16 (red, covering *WT1*) confirms the presence of the del(11)(p13) in case 2723. (**c**) Sequence analysis shows a truncating *WT1* exon 7 mutation on the remaining allele of case 2723. (**d**) Similar FISH analysis as in (**b**) on TLX3 wild-type T-ALL cases identified one additional case showing a biallelic *WT1* deletion. (**e**) Array-CGH analysis confirmed the presence of a large monoallelic deletion, del(11)(p13p14.3), in combination with an additional loss of the genomic region surrounding the *WT1* gene on the other allele.

## DISCUSSION

To get more insight in new genetic defects that may cooperate with *TLX3* gene expression in the leukemic transformation of thymocytes, we performed array-CGH analysis on a *TLX3* rearranged T-ALL patient cohort.

About 25% of *TLX3* rearranged T-ALL cases showed a deletion at the terminal end of the long arm of chromosome 5. Interestingly, for a number of cases, the genomic breakpoint of this deletion was situated just downstream of the *TLX3* oncogene. The deletions in these cases differ from the previously described *TLX3* deletions that involved a genomic region upstream of the *TLX3* gene near the translocation breakpoint.[16] Although most T-ALL cases that show *TLX3* expression harbor a cryptic translocation at this genomic locus, a number of studies have reported *TLX3* activation in the presence of a seemingly normal *TLX3* locus.[18, 19] For case no. 2650, combined array-CGH and FISH analysis strongly suggest that the *TLX3* expression is associated with a interstitial del(5)(q35.1q35.3) in the absence of a *TLX3* translocation. It is therefore tempting to speculate that *TLX3* is normally under transcriptional control of a negative regulatory domain downstream of *TLX3*. Deletion of this negative regulatory element may lead to ectopic *TLX3* expression. In addition, a potential tumor suppressor gene could be present in the minimal deleted area at 5q35 that specifically cooperates with *TLX3* expression in the leukemogenesis of T-ALL. This hypothesis is strengthened by the fact that two


A
del(11)(p13)
T-ALL #2723
Detailed overview deleted region
B
T-ALL #2723
RP11-299P16
RP11-98c11
C
wild type
GGAGTAGCCCCGACTCTTGTACGGTCGGCATCTGAG
WT1 exon 7
C < TAG
T-ALL #23
GGAGTAGCCCCGACTCTTGTATAGGGTCGGCATCTG
WT1 exon 7
D
TLX3 negative
T-ALL
RP11-98c11
RP11-299P16
control T-ALL
RP11-299P16
RP11-98c11
E
del(11)(p13p14.3)
T-ALL
TLX3 negative
Genomic region surounding WT1

cases have smaller 5q35 deletions with breakpoints near *NKX2-5*. A potential candidate gene in this 5q35 genomic region is *NSD1*. Mutations or deletions of the *NSD1* gene are the major cause of Sotos syndrome, a constitutional overgrowth disorder,[31] and patients with this syndrome have a higher risk for the development of leukemia.[32, 33, 34] In addition, *NSD1* is involved in a cryptic translocation, t(5;11)(q35;p15.5), generating a *NUP98–NSD1* fusion gene in acute myeloid leukemia (AML).[35] Although gene expression array data revealed no difference in *NSD1* gene expression between patients with and without the del(5)(q35), a future mutation screening of *NSD1* in *TLX3* rearranged T-ALL is mandatory to evaluate a potential role for *NSD1* inactivation in T-ALL.

Cryptic deletions on chromosome 1 were identified in four T-ALL cases with a commonly deleted region surrounding chromosomal band 1p36. Similar 1p36 deletions were previously identified in about 30% of human neuroblastomas,[36] 25% of colorectal cancer cases,[37] and a variety of hematological malignancies including AML,[38] chronic myelogenous leukemia[39] and non-Hodgkin's lymphoma.[40] In neuroblastoma and colorectal cancer, reduced expression levels of the CAMTA1 gene correlated with adverse outcome, suggesting that *CAMTA1* could act as the 1p36-specifc tumor suppressor gene in these malignancies.[36, 37] Another interesting target gene in this genomic region is the *CHD5* gene, which has been shown to be a tumor suppressor that controls proliferation and apoptosis via the p19Arf/p53 pathway.[41] Other potential target genes within this genomic region are *HES2* and *HES3*, both of which are highly similar to *HES1* that is a basic helix-loop-helix transcriptional repressor and known *NOTCH1* target gene.[42] Also the *TNFRSF25* gene may represent a target gene of this deletion. *TNFRSF25* is a member of the tumor necrosis factor-receptor family which controls lymphocyte proliferation and regulates cell apoptosis.[43]

The minimal deleted region of the cryptic deletions on chromosome 13 contained the microRNA cluster, *miR-15/miR-16a*. In chronic lymphocytic leukemia, deletion of this *miR-15/miR-16a* cluster leads to the activation of antiapoptotic BCL2.[44] For example, in case no. 2723 (Supplementary Table 1), activation of BCL2 could cooperate with a homozygous deletion of the *p15/p16* locus, a *NOTCH1* mutation and activated *TLX3* expression in the development of T-ALL.

Three *TLX3* rearranged T-ALL cases showed cryptic deletions on chromosomal band 16q22.1. In AML, this genomic region is recurrently targeted by cytogenetic abnormalities including an inversion, inv(16)(p13q22); a

translocation, t(16;16)(p13q22) and a deletion, del(16)(q22).[40] The inv(16) and t(16;16) both result in a *CBFB–MYH11* fusion gene, which is associated with a more favorable prognosis.[45] In contrast, the deletion del(16)(q22) does not provide a favorable outcome and it remains to be elucidated whether *CBFB* is targeted in the 16q deletions in AML.[46] In the *TLX3* rearranged T-ALL cases, the minimal deleted region on 16q22.1 contained 21 genes but lacked the *CBFB* gene. One interesting candidate genes in this genomic region is *CTCF*, which is a conserved transcriptional repressor of the *MYC* oncogene.[47] MYC has been described in T-ALL to become aberrantly activated due to a TCR-mediated translocation[48] and has been shown to represent an important downstream target of activated *NOTCH1*.[49, 50] Therefore, inactivation of *CTCF* could represent an alternative mechanism for MYC activation in T-ALL.

The majority of novel regions of genomic amplification or deletion were only detected in single *TLX3* rearranged T-ALL cases (*n*=19). Given their low frequency, one could argue that their oncogenic role in T-ALL is negligible. However, *NOTCH1*, which was originally identified due to its involvement in a rare chromosomal translocation (<1%), was later identified as the most predominant mutational target in T-ALL (>50% of cases).[9] Similarly, the cryptic deletion, del(4)(q31.3q32.1), that was detected in a single case, includes the *FBXW7* gene. FBXW7 is an F-box protein that binds specific substrates including CyclinE, NOTCH1, cMYC and cJUN target these for ubiquitin-mediated proteolysis. Heterozygous missense mutations of the *FBXW7* gene are present in 8–30% of T-ALL cases,[25, 26, 27, 28] demonstrating the importance of this gene in T-ALL albeit inactivation through chromosomal deletions is rare. Mutant *FBXW7* has lost the potential to bind the PEST domain of NOTCH1-IC and target NOTCH1 for proteolytic degradation. This results in stabilized NOTCH1-IC in the nucleus, providing an alternative mechanism of NOTCH1 activation in T-ALL that is insensitive for γ-secretase inhibition. The present study describes the first case of a heterozygous *FBXW7* deletion in human T-ALL. Haploinsufficiency of *FBXW7* may be sufficient for NOTCH1 stabilization as no *FBXW7* mutation could be identified in the remaining allele. *FBXW7* mutations can occur in combination with NOTCH1 heterodimerization (HD) mutations but not with PEST truncating mutations, and may complement the relatively weak transcriptional activity of HD mutant NOTCH1 molecules.[28] In our del(4)(q31.3q32.1)-positive T-ALL case, an activating NOTCH1 mutation was identified in the HD domain.

Deletions on the short arm of chromosome 12 are frequently detected in a wide range of hematological malignancies. A recent genome-wide copy number analysis showed 12p deletions in about 25% of B-ALL cases and suggested the *TEL* gene as the main target of this genomic abnormality.[51] However, the 12p deletion that we identified was about 600 kb in size and included the *CDKN1B/p27/KIP1* gene and the microRNA genes, *miR-613* and *miR-614*, whereas it did not include the *TEL* gene. This indicates that the target gene(s) for 12p deletions probably differs between T- and B-cells ALL. The *CDKN1B/p27/KIP1* gene encodes a cell cycle regulator that, similar to p15/p16, inhibits cyclin-dependent kinases (CDK). Loss of these CDK inhibitors may result in uncontrolled cell cycle. The T-ALL case with this *CDKN1B* deletion (no. 9858) also contained a homozygous *p15/p16* deletion, indicating that different T-cell cycle defects can collaborate with a *NOTCH1* mutation and *TLX3* overexpression in the development of T-ALL (Supplementary Table 1).

WT1, a transcription factor involved in normal cellular development and cell survival, was initially discovered as a tumor suppressor in Wilms' tumor, a pediatric kidney malignancy.[52] In acute leukemias, there is evidence that this gene can both act as an oncogene as well as a tumor suppressor gene.[30] *WT1* mutations have been described in AML, leading to a truncated WT1 protein.[53, 54] In addition, specific AML subtypes show low levels of *WT1* expression.[30] Both observations are consistent with a tumor suppressor role of WT1 in AML. In contrast, a variety of leukemias are characterized by activated *WT1* expression compared to normal bone marrow or normal progenitor cells[55, 56] that has been associated with poor outcome.[57, 58] Our single T-ALL case with a deletion of *WT1* combined with an inactivational mutation in the remaining *WT1* allele points toward a potential tumor suppressor role of *WT1* in T-ALL. *WT1* inactivation is not restricted to *TLX3* rearranged T-ALL cases, as biallelic WT1 deletions were also observed in a *TLX3* negative T-ALL case indicating that WT1 inactivation may be a more general collaborating genetic event in T-cell leukemia. Nevertheless, the *TLX3* rearranged T-ALL case showing *WT1* inactivation also harbored a *miR-15/miR-16a* deletion, further extending the range of different genetic defects that collaborate in T-ALL development.

In conclusion, we performed a genome-wide copy number screening on *TLX3* rearranged T-ALL cases and identified the cryptic deletion, del(5)(q35), as a new and recurrent genetic aberration that is exclusively associated with *TLX3* expression in T-ALL. In addition, we identified a number of genetic events,

including *FBXW7* and *WT1* inactivation that could collaborate with *TLX3* expression, *NOTCH1* activation and *p15/p16* deletion in the development of T-cell leukemia. As shown for *FBXW7*, the identification of new genomic deletions/amplifications, even at low frequency, can still highlight important target genes with a broader role in T-ALL. Therefore, it is likely that the current overview of genetic defects will be further helpful for a better understanding of the molecular pathways leading to T-cell leukemia.

**ACKNOWLEDGEMENTS**

PVV is financed by the Sophia Foundation for Medical Research (SSWO-440). This study was further supported by the Ter Meulen Fund, Royal Netherlands Academy of Arts and Sciences and the Foundation 'De Drie Lichten'.

## REFERENCES


1. Pui CH, Relling MV, Downing JR. Acute lymphoblastic leukemia. N Engl J Med 2004; 350: 1535–1548.
2. Clappier E, Cuccuini W, Cayuela JM, Vecchione D, Baruchel A, Dombret H *et al*. Cyclin D2 dysregulation by chromosomal translocations to TCR loci in T-cell acute lymphoblastic leukemias. Leukemia 2006; 20: 82–86.
3. De Keersmaecker K, Marynen P, Cools J. Genetic insights in the pathogenesis of T-cell acute lymphoblastic leukemia. Haematologica 2005; 90: 1116–1127.
4. Graux C, Cools J, Melotte C, Quentmeier H, Ferrando A, Levine R *et al*. Fusion of NUP214 to ABL1 on amplified episomes in T-cell acute lymphoblastic leukemia. Nat Genet 2004; 36: 1084–1089.
5. Lahortiga I, De Keersmaecker K, Van Vlierberghe P, Graux C, Cauwelier B, Lambert F *et al*. Duplication of the MYB oncogene in T cell acute lymphoblastic leukemia. Nat Genet 2007; 39: 593–595.
6. Soulier J, Clappier E, Cayuela JM, Regnault A, Garcia-Peydro M, Dombret H *et al*. HOXA genes are included in genetic and biologic networks defining human acute T-cell leukemia (T-ALL). Blood 2005; 106: 274–286.
7. Speleman F, Cauwelier B, Dastugue N, Cools J, Verhasselt B, Poppe B *et al*. A new recurrent inversion, inv(7)(p15q34), leads to transcriptional activation of HOXA10 and HOXA11 in a subset of T-cell acute lymphoblastic leukemias. Leukemia 2005; 19: 358–366.
8. Van Vlierberghe P, van Grotel M, Beverloo HB, Lee C, Helgason T, Buijs-Gladdines J *et al*. The cryptic chromosomal deletion del(11)(p12p13) as a new activation mechanism of LMO2 in pediatric T-cell acute lymphoblastic leukemia. Blood 2006; 108: 3520–3529.
9. Weng AP, Ferrando AA, Lee W, Morris JPt, Silverman LB, Sanchez-Irizarry C *et al*. Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. Science 2004; 306: 269–271.
10. Clappier E, Cuccuini W, Kalota A, Crinquette A, Cayuela JM, Dik WA *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: 1251–1261.
11. Graux C, Cools J, Michaux L, Vandenberghe P, Hagemeijer A. Cytogenetics and molecular genetics of T-cell acute lymphoblastic leukemia: from thymocyte to lymphoblast. Leukemia 2006; 20: 1496–1510.
12. Armstrong SA, Look AT. Molecular genetics of acute lymphoblastic leukemia. J Clin Oncol 2005; 23: 6306–6315.
13. Grabher C, von Boehmer H, Look AT. Notch 1 activation in the molecular pathogenesis of T-cell acute lymphoblastic leukaemia. Nat Rev Cancer 2006; 6: 347–359.
14. Bernard OA, Busson-LeConiat M, Ballerini P, Mauchauffe M, Della Valle V, Monni R *et al*. A new recurrent and specific cryptic translocation, t(5;14)(q35;q32), is associated with expression of the Hox11L2 gene in T acute lymphoblastic leukemia. Leukemia 2001; 15: 1495–1504.
15. Hansen-Hagge TE, Schafer M, Kiyoi H, Morris SW, Whitlock JA, Koch P *et al*. Disruption of the RanBP17/Hox11L2 region by recombination with the TCRdelta

locus in acute lymphoblastic leukemias with t(5;14)(q34;q11). Leukemia 2002; 16: 2205–2212.
16. Su XY, Busson M, Della Valle V, Ballerini P, Dastugue N, Talmant P *et al*. Various types of rearrangements target TLX3 locus in T-cell acute lymphoblastic leukemia. Genes Chromosomes Cancer 2004; 41: 243–249.
17. van Grotel M, Meijerink JP, Beverloo HB, Langerak AW, Buys-Gladdines JG, Schneider P *et al*. The outcome of molecular-cytogenetic subgroups in pediatric T-cell acute lymphoblastic leukemia: a retrospective study of patients treated according to DCOG or COALL protocols. Haematologica 2006; 91: 1212–1221.
18. Berger R, Dastugue N, Busson M, Van Den Akker J, Perot C, Ballerini P *et al*. t(5;14)/HOX11L2-positive T-cell acute lymphoblastic leukemia. A collaborative study of the Groupe Francais de Cytogenetique Hematologique (GFCH). Leukemia 2003; 17: 1851–1857.
19. Mauvieux L, Leymarie V, Helias C, Perrusson N, Falkenrodt A, Lioure B *et al*. High incidence of Hox11L2 expression in children with T-ALL. Leukemia 2002; 16: 2417–2422.
20. Ballerini P, Blaise A, Busson-Le Coniat M, Su XY, Zucman-Rossi J, Adam M *et al*. HOX11L2 expression defines a clinical subtype of pediatric T-ALL associated with poor prognosis. Blood 2002; 100: 991–997.
21. Cave H, Suciu S, Preudhomme C, Poppe B, Robert A, Uyttebroeck A *et al*. Clinical significance of HOX11L2 expression linked to t(5;14)(q35;q32), of HOX11 expression, and of SIL-TAL fusion in childhood T-cell malignancies: results of EORTC studies 58881 and 58951. Blood 2004; 103: 442–450.
22. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD *et al*. Global variation in copy number in the human genome. Nature 2006; 444: 444–454.
23. Waggoner DJ, Raca G, Welch K, Dempsey M, Anderes E, Ostrovnaya I *et al*. NSD1 analysis for Sotos syndrome: insights and perspectives from the clinical laboratory. Genet Med 2005; 7: 524–533.
24. van Grotel M, Meijerink JP, van Wering ER, Langerak AW, Beverloo HB, Buijs-Gladdines JG *et al*. Prognostic significance of molecular-cytogenetic abnormalities in pediatric T-ALL is not explained by immunophenotypic differences. Leukemia 2007 Oct 11 (advance online publication).
25. Malyukova A, Dohda T, von der Lehr N, Akhondi S, Corcoran M, Heyman M *et al*. The tumor suppressor gene hCDC4 is frequently mutated in human T-cell acute lymphoblastic leukemia with functional consequences for Notch signaling. Cancer Res 2007; 67: 5611–5616.
26. Maser RS, Choudhury B, Campbell PJ, Feng B, Wong KK, Protopopov A *et al*. Chromosomally unstable mouse tumours have genomic alterations similar to diverse human cancers. Nature 2007; 447: 966–971.
27. O'Neil J, Grim J, Strack P, Rao S, Tibbitts D, Winter C *et al*. FBW7 mutations in leukemic cells mediate NOTCH pathway activation and resistance to gamma-secretase inhibitors. J Exp Med 2007; 204: 1813–1824.
28. Thompson BJ, Buonamici S, Sulis ML, Palomero T, Vilimas T, Basso G *et al*. The SCFFBW7 ubiquitin ligase complex as a tumor suppressor in T cell leukemia. J Exp Med 2007; 204: 1825–1835.

29. Balgobind B, Van Vlierberghe P, van den Ouweland A, Beverloo HB, Terlouw-Kromosoeto J, van Wering ER *et al*. Leukemia associated NF1 inactivation in pediatric T-ALL and AML patients lacking evidence for neurofibromatosis. Blood 2007 (in press).
30. Yang L, Han Y, Suarez Saiz F, Minden MD. A tumor suppressor and oncogene: the WT1 story. Leukemia 2007; 21: 868–876.
31. Kurotaki N, Imaizumi K, Harada N, Masuno M, Kondoh T, Nagai T *et al*. Haploinsufficiency of NSD1 causes Sotos syndrome. Nature genetics 2002; 30: 365–366.
32. Al-Mulla N, Belgaumi AF, Teebi A. Cancer in Sotos syndrome: report of a patient with acute myelocytic leukemia and review of the literature. J Pediatr Hematol Oncol 2004; 26: 204–208.
33. Martinez-Glez V, Lapunzina P. Sotos syndrome is associated with leukemia/lymphoma. Am J Med Genet A 20071; 143: 1244–1245.
34. Ziino O, Rondelli R, Micalizzi C, Luciani M, Conter V, Arico M. Acute lymphoblastic leukemia in children with associated genetic conditions other than Down's syndrome. The AIEOP experience. Haematologica 2006; 91: 139–140.
35. Brown J, Jawad M, Twigg SR, Saracoglu K, Sauerbrey A, Thomas AE *et al*. A cryptic t(5;11)(q35;p15.5) in 2 children with acute myeloid leukemia with apparently normal karyotypes, identified by a multiplex fluorescence *in situ* hybridization telomere assay. Blood 2002; 99: 2526–2531.
36. Henrich KO, Fischer M, Mertens D, Benner A, Wiedemeyer R, Brors B *et al*. Reduced expression of CAMTA1 correlates with adverse outcome in neuroblastoma patients. Clin Cancer Res 2006; 12: 131–138.
37. Kim MY, Yim SH, Kwon MS, Kim TM, Shin SH, Kang HM *et al*. Recurrent genomic alterations with impact on survival in colorectal cancer identified by genome-wide array comparative genomic hybridization. Gastroenterology 2006; 131: 1913–1924.
38. Mori N, Morosetti R, Mizoguchi H, Koeffler HP. Progression of myelodysplastic syndrome: allelic loss on chromosomal arm 1p. Br J Haematol 2003; 122: 226–230.
39. Mori N, Morosetti R, Spira S, Lee S, Ben-Yehuda D, Schiller G *et al*. Chromosome band 1p36 contains a putative tumor suppressor gene important in the evolution of chronic myelocytic leukemia. Blood 1998; 92: 3405–3409.
40. Melendez B, Cuadros M, Robledo M, Rivas C, Fernandez-Piqueras J, Martinez-Delgado B *et al*. Coincidental LOH regions in mouse and humans: evidence for novel tumor suppressor loci at 9q22-q34 in non-Hodgkin's lymphomas. Leuk Res 2003; 27: 627–633.
41. Bagchi A, Papazoglu C, Wu Y, Capurso D, Brodt M, Francis D *et al*. CHD5 is a tumor suppressor at human 1p36. Cell 2007; 128: 459–475.
42. Fischer A, Gessler M. Delta-Notch--and then? Protein interactions and proposed modes of repression by Hes and Hey bHLH factors. Nucleic Acids Res 2007; 35: 4583–4596.
43. Chinnaiyan AM, O'Rourke K, Yu GL, Lyons RH, Garg M, Duan DR *et al*. Signal transduction by DR3, a death domain-containing receptor related to TNFR-1 and CD95. Science (New York, NY 1996; 274: 990–992.

44. Cimmino A, Calin GA, Fabbri M, Iorio MV, Ferracin M, Shimizu M *et al*. miR-15 and miR-16 induce apoptosis by targeting BCL2. Proc Natl Acad Sci USA 2005; 102: 13944–13949.
45. Plantier I, Lai JL, Wattel E, Bauters F, Fenaux P. Inv(16) may be one of the only 'favorable' factors in acute myeloid leukemia: a report on 19 cases with prolonged follow-up. Leuk Res 1994; 18: 885–888.
46. Betts DR, Ammann RA, Hirt A, Hengartner H, Beck-Popovic M, Kuhne T *et al*. The prognostic significance of cytogenetic aberrations in childhood acute myeloid leukaemia. A study of the Swiss Paediatric Oncology Group (SPOG). Eur J Haematol 2007; 78: 468–476.
47. Filippova GN, Fagerlie S, Klenova EM, Myers C, Dehner Y, Goodwin G *et al*. An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. Mol Cell Biol 1996; 16: 2802–2813.
48. Inaba T, Murakami S, Oku N, Itoh K, Ura Y, Nakanishi S *et al*. Translocation between chromosomes 8q24 and 14q11 in T-cell acute lymphoblastic leukemia. Cancer Genet Cytogenet 1990; 49: 69–74.
49. Palomero T, Lim WK, Odom DT, Sulis ML, Real PJ, Margolin A *et al*. NOTCH1 directly regulates c-MYC and activates a feed-forward-loop transcriptional network promoting leukemic cell growth. Proc Natl Acad Sci USA 2006; 103: 18261–18266.
50. Weng AP, Millholland JM, Yashiro-Ohtani Y, Arcangeli ML, Lau A, Wai C *et al*. c-Myc is an important direct target of Notch1 in T-cell acute lymphoblastic leukemia/lymphoma. Genes Dev 2006; 20: 2096–2109.
51. Mullighan CG, Goorha S, Radtke I, Miller CB, Coustan-Smith E, Dalton JD *et al*. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature 2007; 446: 758–764.
52. Little M, Wells C. A clinical overview of WT1 gene mutations. Hum Mutat 1997; 9: 209–225.
53. King-Underwood L, Pritchard-Jones K. Wilms' tumor (WT1) gene mutations occur mainly in acute myeloid leukemia and may confer drug resistance. Blood 1998; 91: 2961–2968.
54. King-Underwood L, Renshaw J, Pritchard-Jones K. Mutations in the Wilms' tumor gene WT1 in leukemias. Blood 1996; 87: 2171–2179.
55. Miwa H, Beran M, Saunders GF. Expression of the Wilms' tumor gene (WT1) in human leukemias. Leukemia 1992; 6: 405–409.
56. Miyagi T, Ahuja H, Kubota T, Kubonishi I, Koeffler HP, Miyoshi I. Expression of the candidate Wilm's tumor gene, WT1, in human leukemia cells. Leukemia 1993; 7: 970–977.
57. Barragan E, Cervera J, Bolufer P, Ballester S, Martin G, Fernandez P *et al*. Prognostic implications of Wilms' tumor gene (WT1) expression in patients with de novo acute myeloid leukemia. Haematologica 2004; 89: 926–933.
58. Chiusa L, Francia di Celle P, Campisi P, Ceretto C, Marmont F, Pich A. Prognostic value of quantitative analysis of WT1 gene transcripts in adult acute lymphoblastic leukemia. Haematologica 2006; 91: 270–271

# CHAPTER 4

## NKL homeobox genes in leukemia

I. Homminga[1], R. Pieters[1] and J.P.P. Meijerink[1]

1. Department of Pediatric Oncology/Hematology, Erasmus MC/Sophia Children's Hospital, Rotterdam, The Netherlands

*Submitted*

**ABSTRACT**

NK-like (NKL) homeobox genes code for transcription factors which can act as key regulators in fundamental cellular processes. NKL genes have been implicated in divergent types of cancer. In this review we summarize the involvement of NKL genes in cancer and leukemia in particular. NKL genes can act as tumor suppressor genes as well as oncogenes, depending on tissue type. Aberrant expression of NKL genes is especially common in T-cell acute lymphoblastic leukemia (T-ALL). In T-ALL, eight NKL genes have been reported highly expressed in specific T-ALL subgroups and in ~30% of cases, high expression is caused by chromosomal rearrangement of one of 5 NKL genes. Most of these NKL genes are normally not expressed in T-cell development. We hypothesize that the NKL genes might share a similar downstream effect that promotes leukemogenesis, possibly due to mimicking a NKL gene that has a physiological role in early hematopoietic development, such as *HHEX*. All eight NKL genes posses a conserved Eh1 repressor motif, which plays an important role in regulating downstream targets in hematopoiesis and possibly in leukemogenesis as well. Identification of a potential common leukemogenic NKL downstream pathway will provide a promising subject for future studies.

# INTRODUCTION

## Homeobox genes

In 1894 William Bateson described a distinct kind of transformation within species, which he terms ‘homeosis’, derived from the Greek ‘homoiosis’, which means ‘the same’. He defined it as “a change of something into the likeness of something else”.(1) Famous examples are the mutant fruit fly *Antennapedia*, which grows a foot on the head where normally the antenna is located. Another example is a moth (*Zygaena*) which has an extra wing on the position of a hind leg. Later, the genes held responsible for these variations were called homeotic genes.(1) In 1983, two research groups independently identified a 180bp sequence that was conserved in all homeotic genes, denoted as the homeobox.(2, 3) The homeobox is strongly conserved among species and encodes for a DNA binding domain consisting of three alpha-helices(4) called the homeodomain. By now more than 200 human homeodomain proteins are known. Homeodomain proteins are transcription factors that are involved in many key processes such as body patterning, embryonic organogenesis and cell fate decisions. A mutation in one of these genes can have far fetching consequences and lead to homeotic transformations such as originally described by Bateson. The classification and nomenclature of the homeobox genes has varied substantially between research groups and over time. In 2007, Holland and others proposed a classification for all human homeobox genes based on their hypothetical evolutionary shared ancestry(5). Eleven homeobox gene classes are recognized that comprise a total of 102 homeobox gene families. The largest classes are the ANTP (analogous to the *Drosophila* a*ntennapedia (antp)* gene) and PRD (analogous to the *Drosophila paired* (*prd*) gene) classes. The ANTP class can be divided in the HOX-like (HOXL, including the *HOXA* and *HOXB* gene clusters) and the NKL (NK-like) subclasses. The involvement of HOXL genes in cancer and leukemia has been extensively reviewed in literature(6-11) while the NKL genes have received less attention. However, recent findings implicate an important role for various NKL genes in human cancer. Therefore this review will focus on the NKL homeobox genes, their role in cancer and leukemia in particular.

## NKL homeobox genes

NKL homeobox genes have important functions in cell fate specification and embryologic organ development and their expression is often cell type specific. They are key regulators in fundamental processes as differentiation, proliferation and apoptosis. The NKL subclass comprises 48 genes and 19 assumed pseudogenes

(table 1). All NKL genes have a similar homeodomain. The NKL genes are named after Nirenberg and Kim who identified NK1-NK4 genes in 1989 when searching for new genes containing homeobox sequences in *Drosophila melanogaster*(12). The human NKL genes can be divided in gene-families which consists of genes that are all derived from a single gene of the latest common ancestor of *Drosophila* and human(5) (table 1). In several cases, the gene names are somewhat misleading, frequently suggesting higher similarity than is actually the case. For instance, the human *NKX2-3, NKX2-5* and *NKX2-6* genes belong to the NK4 family and are most similar to the *Drosophila melanogaster* NK4 gene. *NKX2-1* and *NKX 2-4* resemble Drosophila *scarecrow* (*scro)* gene and comprise the Nk2.1 family, *NKX2-2* and *NKX2-8* are orthologous to Drosophila *ventral nervous system defective* (*vnd)* gene and part of the Nk2.2 family. Also, not all NKL genes have names that start with *NKX*, other genes such as *TLX1* and *TLX3* also belong to the NKL gene subclass (Table 1). The NKL genes are not strongly clustered on the chromosomes in contrast to the *HOXA-D* genes, though some NKL genes cluster in pairs. This linkage is conserved through species especially for *LBX1-TLX1, LBX2-TLX2, NKX2-6-NKX3-1* and *HMX2 –HMX3*(13) and to a lesser extend for *NKX2-1-NKX2-8* and *NKX2-2-NKX2-4(14)*. Initially, clustering is often the result of tandem duplications, however, the conservation of some of these pairs is likely due to shared regulatory regions.(13) In humans several NKL genes are loosely gathered over large areas on chromosome bands 2p13, 4p15-16, 5q35 and 10q23-26 (Table 1).

## NKL HOMEOBOX GENES IN CANCER

Aberrant expression of NKL homeobox genes may play an important role in oncogenesis. Different abnormalities of NKL genes have been associated with cancer (Table 2). For example, *NKX2-1* is amplified in 3-12% of primary lung adenocarcinomas(15, 16) and in 33% of lung cancer cell lines(15). However, amplification is significantly less frequent in other types of lung cancer, and deletion of *NKX2-1* has been described in squamous lung carcinomas(17). A recent study demonstrated a suppressive effect of *NKX2-1* expression on the progression

**Table 1. (page 93) NKL homeobox genes**

Each grey box corresponds to a separate branch in the phylogenetic tree as described by Holland et al(5) and therefore contains genes that are more related in regard to the homeodomain. Pseudogenes are left out of the table.

| Family | Drosophila homologue | Gene | Synonyms | Location | Eh-1 motif FxIxxIL |
|---|---|---|---|---|---|
| Nk2.1 | *scarecrow* | *NKX2-1* | *TITF1, TTF1, NKX2A, TEBP* | 14q13.3 | FSVSDIL |
| | *(scro)* | *NKX2-4* | *NKX2D* | 20p11.2 | FSVSDIL |
| Nk2.2 | *ventral nervous system* | *NKX2-2* | *NKX2B* | 20p11.2 | FSVKDIL |
| | *defective (vnd)/NK2* | *NKX2-8* | *NKX2H* | 14q13.3 | - |
| Nk3 | *bagpipe* | *NKX3-1* | *NKX3A, BAPX2* | 8p21.2 | FLIQDIL |
| | *(bap)* | *NKX3-2* | *BAPX1* | 4p15.33 | FSIQAIL |
| Nk4 | *tinman* | *NKX2-3* | *NKX2C, CSX3* | 10q24.2 | FSVKDIL |
| | *(tin)/NK4* | *NKX2-5* | *NKX2E, CSX1, CSX* | 5q35.1 | FSVKDIL |
| | | *NKX2-6* | *CSX2* | 8p21.2 | FSVKDIL |
| Nk5 | *H6-like-homeobox* | *HMX1* | *NKX5-3* | 4p16.1 | FLIENLL |
| | *(hmx)* | *HMX2* | *NKX5-2* | 10q26.1 | FTIQSIL |
| | | *HMX3* | *NKX5-1* | 10q26.1 | FSIKNLL |
| Nk6 | *Hgtx* | *NKX6-1* | *NKX6A* | 4q21.23 | HGINDIL |
| | | *NKX6-2* | *NKX6B* | 10q26.1 | HGISDIL |
| | | *NKX6-3* | | 8p11.21 | HGITDIL |
| Tlx | *C15* | *TLX1* | *HOX11, TCL3* | 10q24.3 | FGIDQIL |
| | | *TLX2* | *HOX11L1, NCX, ENX* | 2p13.1 | FGIDQIL |
| | | *TLX3* | *HOX11L2, RNX* | 5q35.1 | FGIDQIL |
| Msx | *drop* | *MSX1* | *HOX7* | 4p16.2 | - |
| | *(dr)* | *MSX2* | | 5q35.2 | - |
| Nanog | - | *NANOG* | *FLJ12581* | 12p13.3 | - |
| Dlx | *Distal-less* | *DLX1* | | 2q31.1 | - |
| | *(dll)* | *DLX2* | *TES1* | 2q31.1 | - |
| | | *DLX3* | | 17q21.3 | - |
| | | *DLX4* | *DLX7, BP1* | 17q21.3 | - |
| | | *DLX5* | | 7q21.3 | - |
| | | *DLX6* | | 7q21.3 | - |
| Dbx | *CG12361* | *DBX1* | | 11p15.1 | FGVNAIL |
| | | *DBX2* | | 12q12 | FLIENLL |
| Bsx | *Brain specific homeobox gene (bsh)* | *BSX* | *BSX1* | 11q24.1 | FFIEDIL |
| Barx | - | *BARX1* | | 9q22.32 | FMIEEIL |
| | | *BARX2* | | 11q24.3 | FMIDEIL |
| Barhl | *BarH1* | *BARHL1* | | 9q34.13 | FGIDSIL |
| | *BarH2* | *BARHL2* | | 1p22.2 | FGIDTIL |
| Lbx | *ladybird early (lbe)* | *LBX1* | *LBX1H* | 10q24.3 | FSIEDIL |
| | *ladybird late(lbl)* | *LBX2* | | 2p13.1 | LSIADIL |
| Hlx | *H2.0* | *HLX* | *HB24, HLX1* | 1q41 | FCIADIL |
| Hhex | *CG7056* | *HHEX* | *PRH, PRHX* | 10q23.3 | FYIEDIL |
| En | *Engrailed (en)* | *EN1* | | 2q14.2 | FFIDNIL |
| | *Invected (inv)* | *EN2* | | 7q36.3 | FFIDNIL |
| Vax | - | *VAX1* | | 10q26.1 | - |
| | | *VAX2* | | 2p13.3 | - |
| Ventx | - | *VENTX* | | 10q26.3 | - |
| Emx | *Empty Spiracles (ems)* | *EMX1* | | 2p13.2 | - |
| | *E5* | *EMX2* | | 10q26.1 | - |
| Noto | *CG18599* | *NOTO* | | 2p13.2 | FSVEAIL |
| Nk1 | *slouch* | *NKX1-1* | *HSPX153, HPX153* | 4p16.3 | FSVLDIL |
| | *(slou)* | *NKX1-2* | *C10orf121* | 10q26.1 | - |

of lung adenocarcinoma.(18) This suggests that *NKX2-1* can function as a tumor suppressor gene as well as an oncogene in lung cancer depending on the subtype, but can also have a dual function within a lung cancer subtype.(18) Also in theabsence of genetic hits, certain NKL homeobox genes have been associated with oncogenesis as crucial downstream targets of established oncogenes, or by effects on proliferation, differentiation or apoptosis (**Table 2**). For example, *NKX2-2* is a critical target of the EWS/FLI fusion protein in Ewing's sarcoma(19) and in clear cell sarcoma, a fusion of EWS/ATF-1 is responsible for *NKX6-1* upregulation(20). Hypermethylation of promoter regions of specific NKL genes has also been reported in human cancer, but in most cases their significance in oncogenesis remains poorly understood (Table 2). Promoter hypermethylation of NKL genes has also been interpreted as a sign of mitotic cell age possibly predisposing for certain types of cancer.(21, 22)

## NKL HOMEOBOX GENES IN LEUKEMIA

Leukemia can be divided in myeloid, B- or T-cell leukemia depending on the precursor cell that gave rise to the leukemia. In normal B- and T-cell development, immunoglobulin (heavy and lambda or kappa light chain) and T-cell receptors (TCR-alpha, -beta, -gamma or -delta) are rearranged to provide a wide diversity in antigen recognition. This process involves double-stranded DNA breaks, deletion, random addition of nucleic acids and ligation of DNA ends. Disregulation of this process can result in chromosomal translocations or inversions that may lead to the ectopic expression of oncogenes. Such oncogenic translocations are common in hematopoietic malignancies, in fact, more than 50% of lymphatic leukemias and lymphomas carry chromosomal translocations that involve the immunoglobulin or TCR genes(85). To date, many different oncogenes have been identified that are ectopically expressed as consequence of such translocations. Among these are homeobox genes such as *HOXA* and *HOXB* genes (for review of HOX genes in leukemia, see (7-10)). Thus far six NKL homeobox genes have been described to be involved in chromosomal translocations or inversions in leukemia and an additional four have been implicated in leukemogenesis by aberrant expression (Table 2).

**Table 2: (page 95) NKL aberrations associated with cancer**
In bold: NKL genes associated with leukemia

| Aberration type | Gene and cancer type |
|---|---|
| Chromosomal rearrangement | ***HHEX* in AML** (t(10;11)(q23;p15)(23))<br>***NKX2-1* in pediatric T-ALL** (inv(14)(q11.2q13), inv(14)(q13q32.33), t(7;14)(q34;q13))(24)<br>***NKX2-2* in pediatric T-ALL** (t(14;20)(q11;p11))(24)<br>***NKX2-5* in T-ALL(25, 26) and CLL(27)** (t(5;14)(q34;q32.2), t(5;14)(q35;q11))<br>***TLX1* in T-ALL** (t(10;14)(q24;q11),t(7;10)(q35;q24))<br>***TLX3* in T-ALL** (t(5;14)(q35;q32)(28-31) t(5;14)(q32;q11)(32), t(5;7)(q35;q21)(30), del(5)(q35.1)(33)) |
| Deletions | *BARX2* as part of a del(11)(q24a25) in ovarian carcinoma (34-36)<br>*EMX2* as part of a del(10)(q25.3q26.1) in endometrial cancer (37)<br>*NKX2-1* and *NKX2-8* in lung carcinoma (17)<br>*NKX2-3* as part of a del(10)(q23.31q24.33) in sporadic colorectal cancer(38) |
| Amplifications | *DLX4* in breast cancer (39)<br>*NKX2-1* and *NKX2-8* in lung adenocarcinoma(15, 16) |
| Mutations | *EMX2* in endometrial cancer(37) |
| SNP | ***HLX* increased risk of therapy related AML(40)**<br>*MSX1* increased risk of breast cancer (41)<br>*NKX3-1* increased risk of prostate cancer(42) |
| Overexpression | *BARX2* in estrogen dependent breast cancer cell lines(43)<br>***DLX4*** in prostate cancer(44) lung cancer(45), **promyelocytic leukemia(46), AML and T-ALL(47)** and breast cancer (48), *DLX5* in ovarian cancer(49) and lung cancer(50)<br>*LBX1* in breast cancer(51)<br>*MSX2* in pancreatic cancer(52-54)<br>*NKX2-2* in Ewing's sarcoma(19).<br>*NKX2-5* in ovarian yolk sac tumors (55)<br>*NKX3-1* in lobular breast carcinomas(56, 57), ***NKX3-1* in *TAL1* rearranged T-ALL(58, 59)**<br>*NKX6-1* in clear cell sarcoma(20)<br>***TLX2* in T-ALL(60)**<br>***VENTX* in AML(61)** |
| Downregulation | DLX4 in colorectal cancer(62)<br>HLX in colorectal cancer(62)<br>*MSX1* in cervical cancer(63)<br>*NKX2-3* in liver metastases of gastrointestinal carcinomas(64) |
| Promoter hypermethylation | *BARHL1* and *TLX2* in melanoma cell lines(65)<br>*DLX1* and *LBX1* in lung cancer(66)<br>***DLX1* and *DLX4* in chronic lymphocytic leukemia(67)**<br>***DLX2-4* in AML,** *DLX4* in breast cancer(68)<br>*EMX2* in lung carcinoma (69)<br>*EN1, LBX2, NKX2-4 and NKX2-5* in salivary gland carcinoma(70)<br>*HMX2* in colorectal cancer cell lines(71)<br>***MSX1* in T-ALL(72)** lung, breast, colon, prostate and gastric cancer(73, 74)<br>***MSX2* in ALL(75)**<br>*NKX2-2* and *DLX2* in luminal breast cancer(76)<br>*NKX2-3* in melanoma cell lines(77)<br>***NKX2-5* in MDS and acute leukemia(21),** *NKX2-5* in prostate cancer(22)<br>*NKX3-1* in prostate cancer (78)<br>***NKX6-1*** in cervical cancer(79, 80) , **ALL(81) and small B-cell lymphoma(82)**<br>*NKX6-1, BARHL2, NKX2-6, DLX2, EN1* and *NKX2-8* in astrocytomas(83)<br>*TLX3* in prostate cancer(84) |

*NKX2-1 and NKX2-2*

Our group recently reported on the *NKX2-1* and *NKX2-2* genes that are amplified (<1%) or involved in TCR or immunoglobulin driven chromosomal translocations or inversions in ~5% of pediatric T-cell acute lymphoblastic leukemia (T-ALL) patients(24). Repositioning of enhancer regions from T-cell receptor or immunoglobulin loci adjacent to *NKX2-1* or *NKX2-2* results in aberrant activation of these genes. *NKX2-1* and *NKX2-2* are very related homeobox genes forming gene families Nk2.1 and Nk2.2 together with *NKX2-4* and *NKX2-8*. Normally these genes are not expressed in T-cells or their precursors(86) and the mechanisms by which they can contribute to leukemia is not yet clear. T-ALL cases harboring these *NKX2-1* or *NKX2-2* rearrangements shared a gene expression signature with T-ALL patients harboring a translocation involving *TLX1*, another NKL gene. Of interest, various *TLX1* rearranged T-ALL cases express *NKX2-1*, albeit to a lower level compared to *NKX2-1* rearranged T-ALL cases. This signature was characterized by high expression of genes involved in proliferation(24). In most cases the lymphoblasts of these patients are arrested in the early cortical stage of thymocyte development(87-90).

*TLX1 (HOX11)*

The *TLX1* gene is upregulated by translocations in ~7% of pediatric and ~30% of adult T-ALL patients(29, 31, 91, 92). *TLX1* aberrations have a been associated with a relative good prognosis(91-94). Normally *TLX1* is not expressed in adult T-cells, thymocytes or hematopoietic stem cells (95, 96) (86) but plays a role in spleen and neural development(97-99). TLX1 is able to immortalize different hematopoietic cell lineages including T-cells(100-103), and over expression of TLX1 in transgenic mice results in lymphoid tumors with long latency(101, 104). TLX1 promotes proliferation and blocks differentiation of hematopoietic precursors, thereby contributing to leukemogenesis(100-102, 105-110), and *TLX1* deregulated T-ALL patients highly express genes that are involved in proliferation(24, 31). TLX1 enhances MYC protein levels by posttranscriptional regulation(111). It also has abrogating effects on cell cycle check points, for instance through down regulation of *CHEK1(104)* or inhibition of protein serine-threonine phosphatases PP1 and PP2A(110, 112). Additionally, TLX1 may increase genomic instability(104).

*TLX2 (HOX11L1)*

*TLX2* is normally not expressed in thymocytes(86) but was highly up regulated in a single T-ALL patient in a cohort of 92 patients(60). This patient co-clustered with *TLX3* rearranged cases in hierarchical cluster analysis based on micro-array gene-expression data(60), suggesting the existence of sporadic T-ALL cases where the *TLX2* gene plays an oncogenic role, that is homologous to *TLX3* rearranged cases. So far however, no translocation of this gene has been identified in this patient(60), therefore the role of *TLX2* in T-ALL is not yet clear.

*TLX3 (HOX11L2)*

About 20% of pediatric and 5% of adult T-ALL patients are characterized by *TLX3* rearrangements and ectopic expression mostly due to a cryptic translocation t(5;14)(q35;q32)(28, 29, 113-115) placing *TLX3* under the influence of the distal enhancer region downstream of *BCL11B*(28-31). In sporadic cases variant *TLX3* aberrations have been reported such as translocations to the *TRA/D@*(32) or *CDK6*(30) gene, a del(5)(q35.1)(33) or more complex rearrangements involving the 5q35region(30). Alike *TLX1* and *TLX2*, *TLX3* is not expressed in normal T-cell development. Leukemogenic modes of actions of *TLX3* have not been extensively studied. As *TLX1* and *TLX3* are closely related genes, a common oncogenic pathway is expected. This is supported by the fact that they can cluster together in gene-expression based hierarchical cluster analysis.(24, 31, 58) On the other hand, supervised gene-expressing profiling shows distinct profiles for *TLX1* and *TLX3* rearranged cases(60) and both groups have a different clinical outcome; *TLX1* rearranged patients have a better prognosis than *TLX3* rearranged.(31, 91-94, 113-116) In addition, *TLX3* cases are associated with the γδ-lineage and immature T-cell development whereas *TLX1* cases are associated with αβ-lineage commitment(87, 88, 114). *TLX1-3* chromosomal aberrations are not reported in other types of cancer.

*NKX2-5*

*NKX2-5* translocations to either *TRD@* or *BCL11B* sites are seen in sporadic T-ALL cases(25, 26). A single case of atypical B-cell chronic proliferative disorder has also been described to carry a *NKX2-5* translocation(27). *NKX2-5* is normally not expressed in thymocytes(26, 86) but involved in spleen and muscle formation. Different pathways that are targeted by NKX2-5 have been proposed to play a role in leukemogenesis. NKX2-5 activates NOTCH3(117) which can enhance

survival(118, 119) and NKX2-5 causes upregulation of miR-17-92 which may lead to increased proliferation(120). Additionally, NKX2-5 binds the promoter of *MEF2C* and activates *MEF2C* transcription in T-ALL cell lines(24, 121). Recently, we identified *MEF2C* as central oncogene in an immature T-ALL subgroup that shares characteristics with early T-cell precursors (ETP-ALL). In these patients, genetic aberrations were identified that target *MEF2C* or *MEF2C*-regulating transcription factors, including *NKX2-5*.(24) During early normal T-cell development, *MEF2C* is down regulated and ectopic *MEF2C* expression has been shown to provide a differentiation block(24). MEF2C can also inhibit apoptosis by repressing NR4A1/NUR77 which subsequently prevents transformation of BCL2 into a pro-apoptotic factor(121).

*HHEX*

ETP-ALL is characterized by early T-cell developmental arrest(24, 31, 122) and ectopic expression of *LMO2*, *LYL1* and the NKL homeobox gene *HHEX*(24, 31). We previously demonstrated that *LMO, LYL1* and *HHEX* are transcriptional targets of MEF2C.(24) *HHEX* may represent an important transcriptional target gene for this T-ALL subtype, as *HHEX* itself is sufficient to initiate self-renewal in thymocytes(123) and *Hex* can induce T-cell-derived lymphomas when over expressed in hematopoietic precursor cells in mice(124). *HHEX* is highly expressed in normal hematopoietic stem cells and down regulation is necessary for normal T-cell development(124, 125), whereas most other hematopoietic lineages maintain *HHEX* expression(126) (86). So far no genetic abnormalities of the *HHEX* gene itself have been found in human T-ALL, though FISH analysis for possible translocations involving the *HHEX* locus has been extensively performed in patients(24) as well as T-ALL cell lines.(117)

In AML a *NUP98/HHEX* fusion has been described due to a t(10;11)(q23;p15).(23) NUP98 is often involved in translocations that result in fusion genes and more than 20 different partner genes have been described. Among these are several other Antp homeobox genes such as *HOXA9(127, 128), HOXD11*(129) and *HOXC13*(130). For most NUP98-homeodomain fusion proteins transforming capacities have been demonstrated. The transforming activity seems to depend primarily on the NUP98 N-terminus and at least in part on an intact homeodomain.(23) (131)

*NKX3-1*

*NKX3-1* has been found over expressed in *TAL1/LMO* rearranged T-ALL(58, 59). Normally *TAL1* is only expressed in the early thymocyte differentiation stages (CD34+, CD1a-, CD4- and CD8-)(132) but it is aberrantly up regulated by interstitial deletions, translocations or unknown mechanisms in more than 40% T-ALL cases.(31, 91, 114, 133, 134) TAL1 can bind to many target genes. One of its target genes is the NKL homeobox gene *NKX3-1*. TAL1 binds to the *NKX3-1* promoter in a complex with LMO, Ldb1 and GATA3 and activates its transcription(59). NKX3-1 in turn seems to be required for T-ALL proliferation and accordingly, genes associated with NKX3-1 expression found by gene-expression analysis were involved in proliferation(59). NKX3-1 was shown to potentially down regulate microRNAs of the miR-17-92 cluster and *NKX3-1* and miR-17-92 expression were inversely correlated(59). However, T-ALL oncogenes *NKX2-5* and *TLX1* were reported to up regulate these miRNAs in T-ALL(120) in line with reported leukemogenic activity of this miRNA cluster in T-ALL models(135). Therefore the role of miR-17-92 suppression by NKX3-1 in T-ALL is not clear. Normally NKX3-1 is not expressed in adult tissues except in prostate and testis(136).

*HLX*

The NKL gene *HLX* has been suggested as a potential oncogene in AML and T-ALL. *HLX* is normally expressed in activated T-cells and early hematopoietic progenitors.(137, 138) Knock down of HLX in CD34+ bone marrow cells inhibits proliferation, while over expression of HLX impairs differentiation into mature hematopoietic lineages(138). HLX promotes proliferation in the T-ALL cell line Jurkat(139) and HLX stably transfected Jurkat cells produce tumors in mice.(140) High levels of HLX have been demonstrated in AML patient samples(141) and a SNP in the 3'-UTR of *HLX* is associated with increased risk of therapy related AML(40). No genetic aberrations of *HLX* have been found in T-ALL or AML. *HLX* is not associated with other types of cancer as oncogene or tumor suppressor gene.

*VENTX*

*VENTX* is highly expressed in a small portion of acute myeloid leukemia patients, especially those with translocation t(8,21) or a normal karyotype(61), but whether it actually has a role in leukemogenesis is not clear. There are inconsistencies in the

reports on the expression levels of *VENTX* in normal hematopoietic lineages. Healthy CD34+ early hematopoietic cells express VENTX at low to undetectable levels while mature myeloid cells highly express *VENTX*.(61, 86) Some reports describe *VENTX* expression in mature B- and T-cells(142), while others have reported low to absent expression in these cell types(61) (86). Enforced expression of VENTX in CD34+ progenitor cells impairs B- and T-cell development but promotes the development of myeloid cells(61). *VENTX* knockdown in AML cell lines impairs proliferation, suggesting a possible oncogenic role for *VENTX* in AML or a role in promoting myeloid phenotype over lymphoid phenotype in preleukemic or leukemic clones(61). In contrast, in chronic lymphoid leukemia, *VENTX* has been suggested as a potential tumor suppressor gene(142) that functions by inducing cell senescence through the activation of p53 and $p16^{ink4a}$.(143)

*DLX2, DLX3* and *DLX4*

In pediatric precursor B-ALL patients carrying an MLL-AF4 translocation, decreased levels of *DLX2*,*3* and *4*(144) have been reported, possibly due to promoter hypermethylation of these genes.(145) This points to a tumor suppressor like role of these genes in this leukemia type. However, the role of these specific NKL genes in oncogenesis is not clear, as extensive promoter hypermethylation of many genes was recently demonstrated in MLL-AF4 translocated infant B-ALL(146). *DLX2* and *DLX4* are normally highly expressed in healthy B-cell progenitors and down regulated during maturation.

**NKL AS ONCOGENE OR TUMOR SUPPRESSOR GENE?**

Since homeobox genes have been initially found to be over expressed, it has been concluded that homeobox genes act as oncogenes. Nowadays, studies have also reported deletions and down regulation of homeobox genes as important oncogenic events. Hence NKL genes cannot be considered as ‘classic oncogenes’. Apparently, the same NKL gene can act as an oncogene or tumor suppressor gene depending on the cellular context. In general it can be stated that homeobox genes that are normally expressed in undifferentiated cells are up regulated in cancer, whereas homeobox genes that are normally expressed in differentiated tissues are down regulated in cancer(147).

## NKL OVEREXPRESSION AS A COMMON THEME IN T-ALL

It is remarkable that in T-ALL many different NKL genes are implied in leukemogenesis, especially compared to other types of cancer (Table 2). Approximately 30% of pediatric T-ALL cases harbor a genetic aberration involving a NKL gene (5% *NKX2-1*, 1% *NKX2-2*, 6% *TLX1*, 19% *TLX3*, 1% *NKX2-5*). The percentage of cases that over express NKL genes is even higher, as *TAL1* rearranged cases over express *NKX3-1* and immature T-ALL cases over express *HHEX*. In addition, high expression of some NKL genes, such as *NKX2-1*, is observed in some cases where no genetic aberration was identified. Seven out of eight NKL homeobox genes that are association with T-ALL (*NKX2-1*, *NKX2-2*, *NKX2-5*, *NKX3-1*, *TLX1*, *TLX2* and *TLX3*) are part of a separate branch in the phylogenetic tree of the NKL homeobox genes and therefore have a similar homeodomain(5) (boxed in table 1). All these seven genes are not expressed in normal T-cell development. The other gene, *HHEX*, is part of a separate phylogenetic tree branch and is expressed in early hematopoietic stages and down regulated upon T-cell development. Two different but compatible lineage dependency models have been proposed for oncogenes, i.e. the "oncogene addiction" model and the "lineage-survival model". As the oncogene addiction model states that an oncogene will provide a tumor-specific gain-of-function, the lineage-survival model poses that tumors may become dependent on survival pathways that are already present in the precursor cells of the specific cell lineage. The seven *NKX* and *TLX* genes mentioned above might be involved in downstream pathways that are normally not involved in early T-cell development, but are nonetheless beneficial for survival of thymocytes, in line with the 'oncogene addiction model'. On the other hand, these genes might mimic NKL genes that play a role in normal hematopoietic development, thereby making use of existing pathways, in line with a 'lineage survival model'. *HHEX(108)* and *MSX2*(117) have both been proposed as candidate genes that might be mimicked by ectopically expressed NKL genes. *HHEX* is highly expressed in hematopoietic progenitors and down regulated upon T-cell differentiation. *HHEX* has also been implicated in leukemogenesis (see above), which makes *HHEX* a more likely candidate to be mimicked by NKL genes in T-ALL.

## NKLs MODES OF ACTION

NKLs have been implied in different processes essential for oncogenesis especially differentiation, proliferation and apoptosis (see above). The exact mechanisms by

which these processes are regulated remains poorly understood. In general, NKL genes function predominantly as transcriptional repressors, though activating properties have also been described(108, 148, 149). More than half (32/48) of the NKL homeodomain proteins contain a conserved Engrailed homogy 1 (Eh1) motif (FxIxxIL, whereby x can be any amino acid) (150, 151) at their N-terminal (defined here as having at least 3 out of 4 conserved amino acids present at the correct position at the N-terminal side, Table 1). The Eh1 domain functions as a strong transcriptional repressor. It can interact with Transducin-like Enhancer-of-split (TLE) co-repressor proteins(152) and these proteins can induce transcriptional repression by recruiting histone deacetylases(153, 154). For *HHEX* and *TLX1*, it has been shown that the interaction with TLE proteins is important in regulating downstream targets in hematopoietic lineages as well as in T-ALL(111, 155). Binding of TLE can result in repression of transcription of NKL targets, but it can also act as a competitive substrate for TLE proteins relieving other factors from TLE-mediated repression and activation of transcription(156). For example, TLX1 expression enhances NOTCH1 signaling in a T-ALL cell line, partly by sequestering TLE from the HES1-TLE-mediated transcriptional repression complex(111, 157). In line with this hypothesis are studies that have identified direct down regulation of TLE proteins by promoter hypermethylation and deletions in acute myeloid leukemia(158, 159). All NKL genes associated with T-ALL have an Eh1-like motif at their N-terminal. It is therefore tempting to speculate that transcriptional repression through the Eh1 motif and TLE scavenging may be general mechanisms by which NKLs promote leukemogenesis in T-ALL. Besides TLE, GATA proteins also interact with NKL proteins. Together they can activate target genes in muscle and lung tissue(160-164). What could be important downstream factors of NKL genes, or proteins affected by NKL genes in leukemogenesis? In general, gene-expression analyses have shown enrichment of genes involved in proliferation in *TLX1*, *TLX3*, *NKX2-1/2-2* and *NKX3-1* rearranged T-ALL patients(24, 31, 58, 59). Besides a general profile, specific genes or miRNAs have been identified for each of these NKL genes. *NOTCH3* has been identified as a recurrent target of NKL proteins TLX1, MSX2 and NKX2-5(117). NOTCH3 over expression induces T-cell lymphomas in mice(118) and is highly expressed in all thirty T-ALL cases examined in a study by Bellavia and others(165) posing a possible important NKL downstream target.

## SUMMARY AND FUTURE IMPLICATIONS

NKLs are implicated in divergent types of cancer and can function as oncogene or tumor suppressor gene, depending on the tissue type. Oncogenes are down regulated during tissue development in normal tissue, whereas tumor suppressor genes are normally up regulated during differentiation. Many different NKLs are involved in T-ALL compared to other types of cancer. The over expression of eight different NKL genes has been associated with T-ALL covering the majority of pediatric T-ALL cases. Most of these NKLs are normally not expressed in T-cell development. This suggests a potential similar downstream effect that promotes leukemogenesis in T-cell progenitors, possibly due to mimicking of a NKL gene that is expressed in early hematopoietic development, like *HHEX*. As all eight NKL genes posses a conserved Eh1 repressor motif, this might also play an important role. A common downstream pathway of NKLs might prove difficult to be discovered, especially when most gene-expression analyses are focused on comparisons between subgroups in T-ALL. To elucidate a potential common role of NKL genes in T-ALL, comparisons with normal thymocytes subsets in combination with ChIP-on-chip or chip-seq analysis and functional knock-down and knock-in experiments will be essential.

## CONFLICT OF INTEREST

Authors have no conflicts of interest to disclose.

## REFERENCES


1. Gehring WJ. Master Control Genes in Development and Evolution: The Homeobox Story. first ed. New Haven: Yale University Press; 1998.
2. Garber RL, Kuroiwa A, Gehring WJ. Genomic and cDNA clones of the homeotic locus Antennapedia in Drosophila. The EMBO journal. 1983;2(11):2027-2036.
3. Scott MP, Weiner AJ, Hazelrigg TI, Polisky BA, Pirrotta V, Scalenghe F, Kaufman TC. The molecular organization of the Antennapedia locus of Drosophila. Cell. 1983 Dec;35(3 Pt 2):763-776.
4. Qian YQ, Billeter M, Otting G, Muller M, Gehring WJ, Wuthrich K. The structure of the Antennapedia homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors. Cell. 1989 Nov 3;59(3):573-580.
5. Holland PW, Booth HA, Bruford EA. Classification and nomenclature of all human homeobox genes. BMC biology. 2007;5:47.
6. Shah N, Sukumar S. The Hox genes and their roles in oncogenesis. Nat Rev Cancer. 2010 May;10(5):361-371.
7. McGonigle GJ, Lappin TR, Thompson A. Grappling with the HOX network in hematopoiesis and leukemia. Front Biosci. 2008;13:4297-4308.
8. Argiropoulos B, Humphries RK. Hox genes in hematopoiesis and leukemogenesis. Oncogene. 2007 Oct 15;26(47):6766-6776.
9. Abramovich C, Humphries RK. Hox regulation of normal and leukemic hematopoietic stem cells. Curr Opin Hematol. 2005 May;12(3):210-216.
10. Abramovich C, Pineault N, Ohta H, Humphries RK. Hox genes: from leukemia to hematopoietic stem cell expansion. Annals of the New York Academy of Sciences. 2005 Jun;1044:109-116.
11. Grier DG, Thompson A, Kwasniewska A, McGonigle GJ, Halliday HL, Lappin TR. The pathophysiology of HOX genes and their role in cancer. J Pathol. 2005 Jan;205(2):154-171.
12. Kim Y, Nirenberg M. Drosophila NK-homeobox genes. Proceedings of the National Academy of Sciences of the United States of America. 1989 Oct;86(20):7716-7720.
13. Wotton KR, Weierud FK, Juarez-Morales JL, Alvares LE, Dietrich S, Lewis KE. Conservation of gene linkage in dispersed vertebrate NK homeobox clusters. Development genes and evolution. 2009 Oct;219(9-10):481-496.
14. Wang CC, Brodnicki T, Copeland NG, Jenkins NA, Harvey RP. Conserved linkage of NK-2 homeobox gene pairs Nkx2-2/2-4 and Nkx2-1/2-9 in mammals. Mamm Genome. 2000 Jun;11(6):466-468.
15. Kwei KA, Kim YH, Girard L, Kao J, Pacyna-Gengelbach M, Salari K, Lee J, Choi YL, Sato M, Wang P, Hernandez-Boussard T, Gazdar AF, Petersen I, Minna JD, Pollack JR. Genomic profiling identifies TITF1 as a lineage-specific oncogene amplified in lung cancer. Oncogene. 2008 Jun 5;27(25):3635-3640.
16. Weir BA, Woo MS, Getz G, Perner S, Ding L, Beroukhim R, Lin WM, Province MA, Kraja A, Johnson LA, Shah K, Sato M, Thomas RK, Barletta JA, Borecki IB, Broderick S, Chang AC, Chiang DY, Chirieac LR, Cho J, Fujii Y, Gazdar AF, Giordano T, Greulich H, Hanna M, Johnson BE, Kris MG, Lash A, Lin L, Lindeman N, Mardis ER, McPherson JD, Minna JD, Morgan MB, Nadel M, Orringer MB, Osborne JR, Ozenberger B, Ramos AH, Robinson J, Roth JA, Rusch

V, Sasaki H, Shepherd F, Sougnez C, Spitz MR, Tsao MS, Twomey D, Verhaak RG, Weinstock GM, Wheeler DA, Winckler W, Yoshizawa A, Yu S, Zakowski MF, Zhang Q, Beer DG, Wistuba, II, Watson MA, Garraway LA, Ladanyi M, Travis WD, Pao W, Rubin MA, Gabriel SB, Gibbs RA, Varmus HE, Wilson RK, Lander ES, Meyerson M. Characterizing the cancer genome in lung adenocarcinoma. Nature. 2007 Dec 6;450(7171):893-898.

17. Harris T, Pan Q, Sironi J, Lutz D, Tian J, Sapkar J, Perez-Soler R, Keller S, Locker J. Both gene amplification and allelic loss occur at 14q13.3 in lung cancer. Clin Cancer Res. Dec 10.
18. Winslow MM, Dayton TL, Verhaak RG, Kim-Kiselak C, Snyder EL, Feldser DM, Hubbard DD, DuPage MJ, Whittaker CA, Hoersch S, Yoon S, Crowley D, Bronson RT, Chiang DY, Meyerson M, Jacks T. Suppression of lung adenocarcinoma progression by Nkx2-1. Nature. 2011 May 5;473(7345):101-104.
19. Smith R, Owen LA, Trem DJ, Wong JS, Whangbo JS, Golub TR, Lessnick SL. Expression profiling of EWS/FLI identifies NKX2.2 as a critical target gene in Ewing's sarcoma. Cancer cell. 2006 May;9(5):405-416.
20. Jishage M, Fujino T, Yamazaki Y, Kuroda H, Nakamura T. Identification of target genes for EWS/ATF-1 chimeric transcription factor. Oncogene. 2003 Jan 9;22(1):41-49.
21. Mossner M, Hopfer O, Nowak D, Baldus CD, Neumann U, Kmetsch A, Benlasfer O, John T, Perka C, Thiel E, Hofmann WK. Detection of differential mitotic cell age in bone marrow CD34(+) cells from patients with myelodysplastic syndrome and acute leukemia by analysis of an epigenetic molecular clock DNA signature. Experimental hematology. 2010 Aug;38(8):661-665.
22. Kwabi-Addo B, Chung W, Shen L, Ittmann M, Wheeler T, Jelinek J, Issa JP. Age-related DNA methylation changes in normal human prostate tissues. Clin Cancer Res. 2007 Jul 1;13(13):3796-3802.
23. Jankovic D, Gorello P, Liu T, Ehret S, La Starza R, Desjobert C, Baty F, Brutsche M, Jayaraman PS, Santoro A, Mecucci C, Schwaller J. Leukemogenic mechanisms and targets of a NUP98/HHEX fusion in acute myeloid leukemia. Blood. 2008 Jun 15;111(12):5672-5682.
24. Homminga I, Pieters R, Langerak AW, de Rooi JJ, Stubbs A, Verstegen M, Vuerhard M, Buijs-Gladdines J, Kooi C, Klous P, van Vlierberghe P, Ferrando AA, Cayuela JM, Verhaaf B, Beverloo HB, Horstmann M, de Haas V, Wiekmeijer AS, Pike-Overzet K, Staal FJ, de Laat W, Soulier J, Sigaux F, Meijerink JP. Integrated Transcript and Genome Analyses Reveal NKX2-1 and MEF2C as Potential Oncogenes in T Cell Acute Lymphoblastic Leukemia. Cancer cell. 2011 Apr 12;19(4):484-497.
25. Nagel S, Kaufmann M, Drexler HG, MacLeod RA. The cardiac homeobox gene NKX2-5 is deregulated by juxtaposition with BCL11B in pediatric T-ALL cell lines via a novel t(5;14)(q35.1;q32.2). Cancer research. 2003 Sep 1;63(17):5329-5334.
26. Przybylski GK, Dik WA, Grabarczyk P, Wanzeck J, Chudobska P, Jankowski K, von Bergh A, van Dongen JJ, Schmidt CA, Langerak AW. The effect of a novel recombination between the homeobox gene NKX2-5 and the TRD locus in T-cell

acute lymphoblastic leukemia on activation of the NKX2-5 gene. Haematologica. 2006 Mar;91(3):317-321.

27. Su X, Della-Valle V, Delabesse E, Azgui Z, Berger R, Merle-Beral H, Bernard OA, Nguyen-Khac F. Transcriptional activation of the cardiac homeobox gene CSX1/NKX2-5 in a B-cell chronic lymphoproliferative disorder. Haematologica. 2008 Jul;93(7):1081-1085.
28. Bernard OA, Busson-LeConiat M, Ballerini P, Mauchauffe M, Della Valle V, Monni R, Nguyen Khac F, Mercher T, Penard-Lacronique V, Pasturaud P, Gressin L, Heilig R, Daniel MT, Lessard M, Berger R. A new recurrent and specific cryptic translocation, t(5;14)(q35;q32), is associated with expression of the Hox11L2 gene in T acute lymphoblastic leukemia. Leukemia. 2001 Oct;15(10):1495-1504.
29. Berger R, Dastugue N, Busson M, Van Den Akker J, Perot C, Ballerini P, Hagemeijer A, Michaux L, Charrin C, Pages MP, Mugneret F, Andrieux J, Talmant P, Helias C, Mauvieux L, Lafage-Pochitaloff M, Mozziconacci MJ, Cornillet-Lefebvre P, Radford I, Asnafi V, Bilhou-Nabera C, Nguyen Khac F, Leonard C, Speleman F, Poppe B, Bastard C, Taviaux S, Quilichini B, Herens C, Gregoire MJ, Cave H, Bernard OA. t(5;14)/HOX11L2-positive T-cell acute lymphoblastic leukemia. A collaborative study of the Groupe Francais de Cytogenetique Hematologique (GFCH). Leukemia. 2003 Sep;17(9):1851-1857.
30. Su XY, Busson M, Della Valle V, Ballerini P, Dastugue N, Talmant P, Ferrando AA, Baudry-Bluteau D, Romana S, Berger R, Bernard OA. Various types of rearrangements target TLX3 locus in T-cell acute lymphoblastic leukemia. Genes, chromosomes & cancer. 2004 Nov;41(3):243-249.
31. Ferrando AA, Neuberg DS, Staunton J, Loh ML, Huard C, Raimondi SC, Behm FG, Pui CH, Downing JR, Gilliland DG, Lander ES, Golub TR, Look AT. Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. Cancer cell. 2002 Feb;1(1):75-87.
32. Hansen-Hagge TE, Schafer M, Kiyoi H, Morris SW, Whitlock JA, Koch P, Bohlmann I, Mahotka C, Bartram CR, Janssen JW. Disruption of the RanBP17/Hox11L2 region by recombination with the TCRdelta locus in acute lymphoblastic leukemias with t(5;14)(q34;q11). Leukemia. 2002 Nov;16(11):2205-2212.
33. Van Vlierberghe P, Homminga I, Zuurbier L, Gladdines-Buijs J, van Wering ER, Horstmann M, Beverloo HB, Pieters R, Meijerink JP. Cooperative genetic defects in TLX3 rearranged pediatric T-ALL. Leukemia. 2008 Apr;22(4):762-770.
34. Sellar GC, Li L, Watt KP, Nelkin BD, Rabiasz GJ, Stronach EA, Miller EP, Porteous DJ, Smyth JF, Gabra H. BARX2 induces cadherin 6 expression and is a functional suppressor of ovarian cancer progression. Cancer research. 2001 Oct 1;61(19):6977-6981.
35. Davis M, Hitchcock A, Foulkes WD, Campbell IG. Refinement of two chromosome 11q regions of loss of heterozygosity in ovarian cancer. Cancer research. 1996 Feb 15;56(4):741-744.
36. Gabra H, Watson JE, Taylor KJ, Mackay J, Leonard RC, Steel CM, Porteous DJ, Smyth JF. Definition and refinement of a region of loss of heterozygosity at

11q23.3-q24.3 in epithelial ovarian cancer associated with poor prognosis. Cancer research. 1996 Mar 1;56(5):950-954.
37. Noonan FC, Mutch DG, Ann Mallon M, Goodfellow PJ. Characterization of the homeodomain gene EMX2: sequence conservation, expression analysis, and a search for mutations in endometrial cancers. Genomics. 2001 Aug;76(1-3):37-44.
38. Wang X, Zbou C, Qiu G, Fan J, Tang H, Peng Z. Screening of new tumor suppressor genes in sporadic colorectal cancer patients. Hepato-gastroenterology. 2008 Nov-Dec;55(88):2039-2044.
39. Cavalli LR, Man YG, Schwartz AM, Rone JD, Zhang Y, Urban CA, Lima RS, Haddad BR, Berg PE. Amplification of the BP1 homeobox gene in breast cancer. Cancer genetics and cytogenetics. 2008 Nov;187(1):19-24.
40. Jawad M, Seedhouse CH, Russell N, Plumb M. Polymorphisms in human homeobox HLX1 and DNA repair RAD51 genes increase the risk of therapy-related acute myeloid leukemia. Blood. 2006 Dec 1;108(12):3916-3918.
41. Sliwinski T, Synowiec E, Czarny P, Gomulak P, Forma E, Morawiec Z, Morawiec J, Dziki L, Wasylecka M, Blasiak J. The c.469+46_56del mutation in the homeobox MSX1 gene--a novel risk factor in breast cancer? Cancer Epidemiol. 2010 Oct;34(5):652-655.
42. Zheng SL, Ju JH, Chang BL, Ortner E, Sun J, Isaacs SD, Sun J, Wiley KE, Liu W, Zemedkun M, Walsh PC, Ferretti J, Gruschus J, Isaacs WB, Gelmann EP, Xu J. Germ-line mutation of NKX3.1 cosegregates with hereditary prostate cancer and alters the homeodomain structure and function. Cancer research. 2006 Jan 1;66(1):69-77.
43. Stevens TA, Meech R. BARX2 and estrogen receptor-alpha (ESR1) coordinately regulate the production of alternatively spliced ESR1 isoforms and control breast cancer cell growth and invasion. Oncogene. 2006 Aug 31;25(39):5426-5435.
44. Schwartz AM, Man YG, Rezaei MK, Simmens SJ, Berg PE. BP1, a homeoprotein, is significantly expressed in prostate adenocarcinoma and is concordant with prostatic intraepithelial neoplasia. Mod Pathol. 2009 Jan;22(1):1-6.
45. Yu M, Wan Y, Zou Q. Prognostic significance of BP1 mRNA expression level in patients with non-small cell lung cancer. Clin Biochem. 2008 Jul;41(10-11):824-830.
46. Awwad RT, Do K, Stevenson H, Fu SW, Lo-Coco F, Costello M, Campbell CL, Berg PE. Overexpression of BP1, a homeobox gene, is associated with resistance to all-trans retinoic acid in acute promyelocytic leukemia cells. Ann Hematol. 2008 Mar;87(3):195-203.
47. Haga SB, Fu S, Karp JE, Ross DD, Williams DM, Hankins WD, Behm F, Ruscetti FW, Chang M, Smith BD, Becton D, Raimondi SC, Berg PE. BP1, a new homeobox gene, is frequently expressed in acute leukemias. Leukemia. 2000 Nov;14(11):1867-1875.
48. Fu SW, Schwartz A, Stevenson H, Pinzone JJ, Davenport GJ, Orenstein JM, Gutierrez P, Simmens SJ, Abraham J, Poola I, Stephan DA, Berg PE. Correlation of expression of BP1, a homeobox gene, with estrogen receptor status in breast cancer. Breast Cancer Res. 2003;5(4):R82-87.

49. Tan Y, Cheung M, Pei J, Menges CW, Godwin AK, Testa JR. Upregulation of DLX5 promotes ovarian cancer cell proliferation by enhancing IRS-2-AKT signaling. Cancer research. 2010 Nov 15;70(22):9197-9206.
50. Kato T, Sato N, Takano A, Miyamoto M, Nishimura H, Tsuchiya E, Kondo S, Nakamura Y, Daigo Y. Activation of placenta-specific transcription factor distal-less homeobox 5 predicts clinical outcome in primary lung cancer patients. Clin Cancer Res. 2008 Apr 15;14(8):2363-2370.
51. Yu M, Smolen GA, Zhang J, Wittner B, Schott BJ, Brachtel E, Ramaswamy S, Maheswaran S, Haber DA. A developmentally regulated inducer of EMT, LBX1, contributes to breast cancer progression. Genes & development. 2009 Aug 1;23(15):1737-1742.
52. Satoh K, Hamada S, Kanno A, Hirota M, Umino J, Ito H, Masamune A, Egawa S, Unno M, Shimosegawa T. Expression of MSX2 predicts malignancy of branch duct intraductal papillary mucinous neoplasm of the pancreas. J Gastroenterol. 2010 Jul;45(7):763-770.
53. Satoh K, Hamada S, Kanno A, Ishida K, Ito H, Hirota M, Masamune A, Egawa S, Unno M, Shimosegawa T. Evaluation of MSX2 mRNA in brush cytology specimens distinguished pancreatic carcinoma from chronic pancreatitis. Cancer Sci. 2011 Jan;102(1):157-161.
54. Satoh K, Hamada S, Kimura K, Kanno A, Hirota M, Umino J, Fujibuchi W, Masamune A, Tanaka N, Miura K, Egawa S, Motoi F, Unno M, Vonderhaar BK, Shimosegawa T. Up-regulation of MSX2 enhances the malignant phenotype and is associated with twist 1 expression in human pancreatic cancer cells. Am J Pathol. 2008 Apr;172(4):926-939.
55. Shibata K, Kajiyama H, Yamamoto E, Terauchi M, Ino K, Nawa A, Kikkawa F. Establishment and characterization of an ovarian yolk sac tumor cell line reveals possible involvement of Nkx2.5 in tumor development. Oncology. 2008;74(1-2):104-111.
56. Turashvili G, Bouchal J, Burkadze G, Kolar Z. Differentiation of tumours of ductal and lobular origin: I. Proteomics of invasive ductal and lobular breast carcinomas. Biomedical papers of the Medical Faculty of the University Palacky, Olomouc, Czechoslovakia. 2005 Jun;149(1):57-62.
57. Gelmann EP, Bowen C, Bubendorf L. Expression of NKX3.1 in normal and malignant tissues. The Prostate. 2003 May 1;55(2):111-117.
58. Soulier J, Clappier E, Cayuela JM, Regnault A, Garcia-Peydro M, Dombret H, Baruchel A, Toribio ML, Sigaux F. HOXA genes are included in genetic and biologic networks defining human acute T-cell leukemia (T-ALL). Blood. 2005 Jul 1;106(1):274-286.
59. Kusy S, Gerby B, Goardon N, Gault N, Ferri F, Gerard D, Armstrong F, Ballerini P, Cayuela JM, Baruchel A, Pflumio F, Romeo PH. NKX3.1 is a direct TAL1 target gene that mediates proliferation of TAL1-expressing human T cell acute lymphoblastic leukemia. The Journal of experimental medicine. Sep 27;207(10):2141-2156.
60. Van Vlierberghe P, van Grotel M, Tchinda J, Lee C, Beverloo HB, van der Spek PJ, Stubbs A, Cools J, Nagata K, Fornerod M, Buijs-Gladdines J, Horstmann M, van Wering ER, Soulier J, Pieters R, Meijerink JP. The recurrent SET-NUP214

fusion as a new HOXA activation mechanism in pediatric T-cell acute lymphoblastic leukemia. Blood. 2008 May 1;111(9):4668-4680.
61. Rawat VP, Arseni N, Ahmed F, Mulaw MA, Thoene S, Heilmeier B, Sadlon T, D'Andrea RJ, Hiddemann W, Bohlander SK, Buske C, Feuring-Buske M. The vent-like homeobox gene VENTX promotes human myeloid differentiation and is highly expressed in acute myeloid leukemia. Proceedings of the National Academy of Sciences of the United States of America. 2010 Sep 28;107(39):16946-16951.
62. Hollington P, Neufing P, Kalionis B, Waring P, Bentel J, Wattchow D, Tilley WD. Expression and localization of homeodomain proteins DLX4, HB9 and HB24 in malignant and benign human colorectal tissues. Anticancer Res. 2004 Mar-Apr;24(2B):955-962.
63. Shim C, Zhang W, Rhee CH, Lee JH. Profiling of differentially expressed genes in human primary cervical cancer by complementary DNA expression array. Clin Cancer Res. 1998 Dec;4(12):3045-3050.
64. Leja J, Essaghir A, Essand M, Wester K, Oberg K, Totterman TH, Lloyd R, Vasmatzis G, Demoulin JB, Giandomenico V. Novel markers for enterochromaffin cells and gastrointestinal neuroendocrine carcinomas. Mod Pathol. 2009 Feb;22(2):261-272.
65. Furuta J, Nobeyama Y, Umebayashi Y, Otsuka F, Kikuchi K, Ushijima T. Silencing of Peroxiredoxin 2 and aberrant methylation of 33 CpG islands in putative promoter regions in human malignant melanomas. Cancer research. 2006 Jun 15;66(12):6080-6086.
66. Rauch T, Li H, Wu X, Pfeifer GP. MIRA-assisted microarray analysis, a new technology for the determination of DNA methylation patterns, identifies frequent methylation of homeodomain-containing genes in lung cancer cells. Cancer research. 2006 Aug 15;66(16):7939-7947.
67. Tong WG, Wierda WG, Lin E, Kuang SQ, Bekele BN, Estrov Z, Wei Y, Yang H, Keating MJ, Garcia-Manero G. Genome-wide DNA methylation profiling of chronic lymphocytic leukemia allows identification of epigenetically repressed molecular pathways with clinical impact. Epigenetics. 2010 Aug 27;5(6).
68. Miyamoto K, Fukutomi T, Akashi-Tanaka S, Hasegawa T, Asahara T, Sugimura T, Ushijima T. Identification of 20 genes aberrantly methylated in human breast cancers. International journal of cancer. 2005 Sep 1;116(3):407-414.
69. Okamoto J, Hirata T, Chen Z, Zhou HM, Mikami I, Li H, Yagui-Beltran A, Johansson M, Coussens LM, Clement G, Shi Y, Zhang F, Koizumi K, Shimizu K, Jablons D, He B. EMX2 is epigenetically silenced and suppresses growth in human lung cancer. Oncogene. 2010 Nov 4;29(44):5969-5975.
70. Bell A, Bell D, Weber RS, El-Naggar AK. CpG Island Methylation Profiling in Human Salivary Gland Adenoid Cystic Carcinoma. Cancer. 2011 Jan 11.
71. Jin B, Yao B, Li JL, Fields CR, Delmas AL, Liu C, Robertson KD. DNMT1 and DNMT3B modulate distinct polycomb-mediated histone modifications in colon cancer. Cancer research. 2009 Sep 15;69(18):7412-7421.
72. Dunwell TL, Hesson LB, Pavlova T, Zabarovska V, Kashuba V, Catchpoole D, Chiaramonte R, Brini AT, Griffiths M, Maher ER, Zabarovsky E, Latif F. Epigenetic analysis of childhood acute lymphoblastic leukemia. Epigenetics. 2009 Apr 1;4(3):185-193.

73. Shames DS, Girard L, Gao B, Sato M, Lewis CM, Shivapurkar N, Jiang A, Perou CM, Kim YH, Pollack JR, Fong KM, Lam CL, Wong M, Shyr Y, Nanda R, Olopade OI, Gerald W, Euhus DM, Shay JW, Gazdar AF, Minna JD. A genome-wide screen for promoter methylation in lung cancer identifies novel methylation markers for multiple malignancies. PLoS Med. 2006 Dec;3(12):e486.
74. Yamashita S, Tsujino Y, Moriguchi K, Tatematsu M, Ushijima T. Chemical genomic screening for methylation-silenced genes in gastric cancer cell lines using 5-aza-2'-deoxycytidine treatment and oligonucleotide microarray. Cancer Sci. 2006 Jan;97(1):64-71.
75. Kuang SQ, Tong WG, Yang H, Lin W, Lee MK, Fang ZH, Wei Y, Jelinek J, Issa JP, Garcia-Manero G. Genome-wide identification of aberrantly methylated promoter associated CpG islands in acute lymphocytic leukemia. Leukemia. 2008 Aug;22(8):1529-1538.
76. Kamalakaran S, Varadan V, Giercksky Russnes HE, Levy D, Kendall J, Janevski A, Riggs M, Banerjee N, Synnestvedt M, Schlichting E, Karesen R, Shama Prasada K, Rotti H, Rao R, Rao L, Eric Tang MH, Satyamoorthy K, Lucito R, Wigler M, Dimitrova N, Naume B, Borresen-Dale AL, Hicks JB. DNA methylation patterns in luminal breast cancers differ from non-luminal subtypes and can identify relapse risk independent of other clinical variables. Mol Oncol. 2011 Feb;5(1):77-92.
77. Tellez CS, Shen L, Estecio MR, Jelinek J, Gershenwald JE, Issa JP. CpG island methylation profiling in human melanoma cell lines. Melanoma research. 2009 Jun;19(3):146-155.
78. Asatiani E, Huang WX, Wang A, Rodriguez Ortner E, Cavalli LR, Haddad BR, Gelmann EP. Deletion, methylation, and expression of the NKX3.1 suppressor gene in primary human prostate cancer. Cancer research. 2005 Feb 15;65(4):1164-1173.
79. Lai HC, Lin YW, Huang TH, Yan P, Huang RL, Wang HC, Liu J, Chan MW, Chu TY, Sun CA, Chang CC, Yu MH. Identification of novel DNA methylation markers in cervical cancer. International journal of cancer. 2008 Jul 1;123(1):161-167.
80. Lai HC, Lin YW, Huang RL, Chung MT, Wang HC, Liao YP, Su PH, Liu YL, Yu MH. Quantitative DNA methylation analysis detects cervical intraepithelial neoplasms type 3 and worse. Cancer. Sep 15;116(18):4266-4274.
81. Taylor KH, Pena-Hernandez KE, Davis JW, Arthur GL, Duff DJ, Shi H, Rahmatpanah FB, Sjahputera O, Caldwell CW. Large-scale CpG methylation analysis identifies novel candidate genes and reveals methylation hotspots in acute lymphoblastic leukemia. Cancer research. 2007 Mar 15;67(6):2617-2625.
82. Rahmatpanah FB, Carstens S, Guo J, Sjahputera O, Taylor KH, Duff D, Shi H, Davis JW, Hooshmand SI, Chitma-Matsiga R, Caldwell CW. Differential DNA methylation patterns of small B-cell lymphoma subclasses with different clinical behavior. Leukemia. 2006 Oct;20(10):1855-1862.
83. Wu X, Rauch TA, Zhong X, Bennett WP, Latif F, Krex D, Pfeifer GP. CpG island hypermethylation in human astrocytomas. Cancer research. Apr 1;70(7):2718-2727.

84. Wang Y, Hayakawa J, Long F, Yu Q, Cho AH, Rondeau G, Welsh J, Mittal S, De Belle I, Adamson E, McClelland M, Mercola D. "Promoter array" studies identify cohorts of genes directly regulated by methylation, copy number change, or transcription factor binding in human cancer cells. Annals of the New York Academy of Sciences. 2005 Nov;1058:162-185.
85. Zhang Y, Rowley JD. Chromatin structural elements and chromosomal translocations in leukemia. DNA Repair (Amst). 2006 Sep 8;5(9-10):1282-1297.
86. Novershtern N, Subramanian A, Lawton LN, Mak RH, Haining WN, McConkey ME, Habib N, Yosef N, Chang CY, Shay T, Frampton GM, Drake AC, Leskov I, Nilsson B, Preffer F, Dombkowski D, Evans JW, Liefeld T, Smutko JS, Chen J, Friedman N, Young RA, Golub TR, Regev A, Ebert BL. Densely interconnected transcriptional circuits control cell states in human hematopoiesis. Cell. 2011 Jan 21;144(2):296-309.
87. Asnafi V, Beldjord K, Libura M, Villarese P, Millien C, Ballerini P, Kuhlein E, Lafage-Pochitaloff M, Delabesse E, Bernard O, Macintyre E. Age-related phenotypic and oncogenic differences in T-cell acute lymphoblastic leukemias may reflect thymic atrophy. Blood. 2004 Dec 15;104(13):4173-4180.
88. van Grotel M, Meijerink JP, van Wering ER, Langerak AW, Beverloo HB, Buijs-Gladdines JG, Burger NB, Passier M, van Lieshout EM, Kamps WA, Veerman AJ, van Noesel MM, Pieters R. Prognostic significance of molecular-cytogenetic abnormalities in pediatric T-ALL is not explained by immunophenotypic differences. Leukemia. 2008 Jan;22(1):124-131.
89. Ludwig WD, Harbott J, Bartram CR, Komischke B, Sperling C, Teichmann JV, Seibt-Jung H, Notter M, Odenwald E, Nehmer A, et al. Incidence and prognostic significance of immunophenotypic subgroups in childhood acute lymphoblastic leukemia: experience of the BFM study 86. Recent Results Cancer Res. 1993;131:269-282.
90. Pullen J, Shuster JJ, Link M, Borowitz M, Amylon M, Carroll AJ, Land V, Look AT, McIntyre B, Camitta B. Significance of commonly used prognostic factors differs for children with T cell acute lymphocytic leukemia (ALL), as compared to those with B-precursor ALL. A Pediatric Oncology Group (POG) study. Leukemia. 1999 Nov;13(11):1696-1707.
91. Ferrando AA, Neuberg DS, Dodge RK, Paietta E, Larson RA, Wiernik PH, Rowe JM, Caligiuri MA, Bloomfield CD, Look AT. Prognostic importance of TLX1 (HOX11) oncogene expression in adults with T-cell acute lymphoblastic leukaemia. Lancet. 2004 Feb 14;363(9408):535-536.
92. Kees UR, Heerema NA, Kumar R, Watt PM, Baker DL, La MK, Uckun FM, Sather HN. Expression of HOX11 in childhood T-lineage acute lymphoblastic leukaemia can occur in the absence of cytogenetic aberration at 10q24: a study from the Children's Cancer Group (CCG). Leukemia. 2003 May;17(5):887-893.
93. Cave H, Suciu S, Preudhomme C, Poppe B, Robert A, Uyttebroeck A, Malet M, Boutard P, Benoit Y, Mauvieux L, Lutz P, Mechinaud F, Grardel N, Mazingue F, Dupont M, Margueritte G, Pages MP, Bertrand Y, Plouvier E, Brunie G, Bastard C, Plantaz D, Vande Velde I, Hagemeijer A, Speleman F, Lessard M, Otten J, Vilmer E, Dastugue N. Clinical significance of HOX11L2 expression linked to t(5;14)(q35;q32), of HOX11 expression, and of SIL-TAL fusion in childhood T-

cell malignancies: results of EORTC studies 58881 and 58951. Blood. 2004 Jan 15;103(2):442-450.
94. Schneider NR, Carroll AJ, Shuster JJ, Pullen DJ, Link MP, Borowitz MJ, Camitta BM, Katz JA, Amylon MD. New recurring cytogenetic abnormalities and association of blast cell karyotypes with prognosis in childhood T-cell acute lymphoblastic leukemia: a pediatric oncology group report of 343 cases. Blood. 2000 Oct 1;96(7):2543-2549.
95. Salvati PD, Ranford PR, Ford J, Kees UR. HOX11 expression in pediatric acute lymphoblastic leukemia is associated with T-cell phenotype. Oncogene. 1995 Oct 5;11(7):1333-1338.
96. Yamamoto H, Hatano M, Iitsuka Y, Mahyar NS, Yamamoto M, Tokuhisa T. Two forms of Hox11 a T cell leukemia oncogene, are expressed in fetal spleen but not in primary lymphocytes. Mol Immunol. 1995 Nov;32(16):1177-1182.
97. Kanzler B, Dear TN. Hox11 acts cell autonomously in spleen development and its absence results in altered cell fate of mesenchymal spleen precursors. Dev Biol. 2001 Jun 1;234(1):231-243.
98. Cheng L, Arata A, Mizuguchi R, Qian Y, Karunaratne A, Gray PA, Arata S, Shirasawa S, Bouchard M, Luo P, Chen CL, Busslinger M, Goulding M, Onimaru H, Ma Q. Tlx3 and Tlx1 are post-mitotic selector genes determining glutamatergic over GABAergic cell fates. Nat Neurosci. 2004 May;7(5):510-517.
99. Holland PW, Takahashi T. The evolution of homeobox genes: Implications for the study of brain development. Brain Res Bull. 2005 Sep 15;66(4-6):484-490.
100. Hawley RG, Fong AZ, Reis MD, Zhang N, Lu M, Hawley TS. Transforming function of the HOX11/TCL3 homeobox gene. Cancer research. 1997 Jan 15;57(2):337-345.
101. Hough MR, Reis MD, Singaraja R, Bryce DM, Kamel-Reid S, Dardick I, Breitman ML, Dube ID. A model for spontaneous B-lineage lymphomas in IgHmu-HOX11 transgenic mice. Proceedings of the National Academy of Sciences of the United States of America. 1998 Nov 10;95(23):13853-13858.
102. Keller G, Wall C, Fong AZ, Hawley TS, Hawley RG. Overexpression of HOX11 leads to the immortalization of embryonic precursors with both primitive and definitive hematopoietic potential. Blood. 1998 Aug 1;92(3):877-887.
103. Hawley RG, Fong AZ, Lu M, Hawley TS. The HOX11 homeobox-containing gene of human leukemia immortalizes murine hematopoietic precursors. Oncogene. 1994 Jan;9(1):1-12.
104. De Keersmaecker K, Real PJ, Gatta GD, Palomero T, Sulis ML, Tosello V, Van Vlierberghe P, Barnes K, Castillo M, Sole X, Hadler M, Lenz J, Aplan PD, Kelliher M, Kee BL, Pandolfi PP, Kappes D, Gounari F, Petrie H, Van der Meulen J, Speleman F, Paietta E, Racevskis J, Wiernik PH, Rowe JM, Soulier J, Avran D, Cave H, Dastugue N, Raimondi S, Meijerink JP, Cordon-Cardo C, Califano A, Ferrando AA. The TLX1 oncogene drives aneuploidy in T cell transformation. Nat Med. 2010 Nov;16(11):1321-1327.
105. Greene WK, Ford J, Dixon D, Tilbrook PA, Watt PM, Klinken SP, Kees UR. Enforced expression of HOX11 is associated with an immature phenotype in J2E erythroid cells. Br J Haematol. 2002 Sep;118(3):909-917.

106. Riz I, Hawley TS, Johnston H, Hawley RG. Role of TLX1 in T-cell acute lymphoblastic leukaemia pathogenesis. Br J Haematol. 2009 Apr;145(1):140-143.
107. Riz I, Akimov SS, Eaker SS, Baxter KK, Lee HJ, Marino-Ramirez L, Landsman D, Hawley TS, Hawley RG. TLX1/HOX11-induced hematopoietic differentiation blockade. Oncogene. 2007 Jun 14;26(28):4115-4123.
108. Dixon DN, Izon DJ, Dagger S, Callow MJ, Taplin RH, Kees UR, Greene WK. TLX1/HOX11 transcription factor inhibits differentiation and promotes a non-haemopoietic phenotype in murine bone marrow cells. Br J Haematol. 2007 Jul;138(1):54-67.
109. Owens BM, Hawley TS, Spain LM, Kerkel KA, Hawley RG. TLX1/HOX11-mediated disruption of primary thymocyte differentiation prior to the CD4+CD8+ double-positive stage. Br J Haematol. 2006 Jan;132(2):216-229.
110. Kawabe T, Muslin AJ, Korsmeyer SJ. HOX11 interacts with protein phosphatases PP2A and PP1 and disrupts a G2/M cell-cycle checkpoint. Nature. 1997 Jan 30;385(6615):454-458.
111. Riz I, Hawley TS, Luu TV, Lee NH, Hawley RG. TLX1 and NOTCH coregulate transcription in T cell acute lymphoblastic leukemia cells. Mol Cancer. 2010;9:181.
112. Riz I, Hawley RG. G1/S transcriptional networks modulated by the HOX11/TLX1 oncogene of T-cell acute lymphoblastic leukemia. Oncogene. 2005 Aug 25;24(36):5561-5575.
113. Mauvieux L, Leymarie V, Helias C, Perrusson N, Falkenrodt A, Lioure B, Lutz P, Lessard M. High incidence of Hox11L2 expression in children with T-ALL. Leukemia. 2002 Dec;16(12):2417-2422.
114. van Grotel M, Meijerink JP, Beverloo HB, Langerak AW, Buys-Gladdines JG, Schneider P, Poulsen TS, den Boer ML, Horstmann M, Kamps WA, Veerman AJ, van Wering ER, van Noesel MM, Pieters R. The outcome of molecular-cytogenetic subgroups in pediatric T-cell acute lymphoblastic leukemia: a retrospective study of patients treated according to DCOG or COALL protocols. Haematologica. 2006 Sep;91(9):1212-1221.
115. Ballerini P, Blaise A, Busson-Le Coniat M, Su XY, Zucman-Rossi J, Adam M, van den Akker J, Perot C, Pellegrino B, Landman-Parker J, Douay L, Berger R, Bernard OA. HOX11L2 expression defines a clinical subtype of pediatric T-ALL associated with poor prognosis. Blood. 2002 Aug 1;100(3):991-997.
116. Baak U, Gokbuget N, Orawa H, Schwartz S, Hoelzer D, Thiel E, Burmeister T. Thymic adult T-cell acute lymphoblastic leukemia stratified in standard- and high-risk group by aberrant HOX11L2 expression: experience of the German multicenter ALL study group. Leukemia. 2008 Jun;22(6):1154-1160.
117. Nagel S, Venturini L, Przybylski GK, Grabarczyk P, Meyer C, Kaufmann M, Battmer K, Schmidt CA, Drexler HG, Scherr M, Macleod RA. NK-like homeodomain proteins activate NOTCH3-signaling in leukemic T-cells. BMC cancer. 2009;9:371.
118. Bellavia D, Campese AF, Alesse E, Vacca A, Felli MP, Balestri A, Stoppacciaro A, Tiveron C, Tatangelo L, Giovarelli M, Gaetano C, Ruco L, Hoffman ES, Hayday AC, Lendahl U, Frati L, Gulino A, Screpanti I. Constitutive activation of

NF-kappaB and T-cell leukemia/lymphoma in Notch3 transgenic mice. The EMBO journal. 2000 Jul 3;19(13):3337-3348.
119. Doucas H, Mann CD, Sutton CD, Garcea G, Neal CP, Berry DP, Manson MM. Expression of nuclear Notch3 in pancreatic adenocarcinomas is associated with adverse clinical features, and correlates with the expression of STAT3 and phosphorylated Akt. J Surg Oncol. 2008 Jan 1;97(1):63-68.
120. Nagel S, Venturini L, Przybylski GK, Grabarczyk P, Schmidt CA, Meyer C, Drexler HG, Macleod RA, Scherr M. Activation of miR-17-92 by NK-like homeodomain proteins suppresses apoptosis via reduction of E2F1 in T-cell acute lymphoblastic leukemia. Leukemia & lymphoma. 2009 Jan;50(1):101-108.
121. Nagel S, Meyer C, Quentmeier H, Kaufmann M, Drexler HG, MacLeod RA. MEF2C is activated by multiple mechanisms in a subset of T-acute lymphoblastic leukemia cell lines. Leukemia. 2008 Mar;22(3):600-607.
122. Coustan-Smith E, Mullighan CG, Onciu M, Behm FG, Raimondi SC, Pei D, Cheng C, Su X, Rubnitz JE, Basso G, Biondi A, Pui CH, Downing JR, Campana D. Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia. The lancet oncology. 2009 Feb;10(2):147-156.
123. McCormack MP, Young LF, Vasudevan S, de Graaf CA, Codrington R, Rabbitts TH, Jane SM, Curtis DJ. The Lmo2 oncogene initiates leukemia in mice by inducing thymocyte self-renewal. Science. 2010 Feb 12;327(5967):879-883.
124. George A, Morse HC, 3rd, Justice MJ. The homeobox gene Hex induces T-cell-derived lymphomas when overexpressed in hematopoietic precursor cells. Oncogene. 2003 Oct 2;22(43):6764-6773.
125. Mack DL, Leibowitz DS, Cooper S, Ramsey H, Broxmeyer HE, Hromas R. Down-regulation of the myeloid homeobox protein Hex is essential for normal T-cell development. Immunology. 2002 Dec;107(4):444-451.
126. Manfioletti G, Gattei V, Buratti E, Rustighi A, De Iuliis A, Aldinucci D, Goodwin GH, Pinto A. Differential expression of a novel proline-rich homeobox gene (Prh) in human hematolymphopoietic cells. Blood. 1995 Mar 1;85(5):1237-1245.
127. Borrow J, Shearman AM, Stanton VP, Jr., Becher R, Collins T, Williams AJ, Dube I, Katz F, Kwong YL, Morris C, Ohyashiki K, Toyama K, Rowley J, Housman DE. The t(7;11)(p15;p15) translocation in acute myeloid leukaemia fuses the genes for nucleoporin NUP98 and class I homeoprotein HOXA9. Nature genetics. 1996 Feb;12(2):159-167.
128. Nakamura T, Largaespada DA, Lee MP, Johnson LA, Ohyashiki K, Toyama K, Chen SJ, Willman CL, Chen IM, Feinberg AP, Jenkins NA, Copeland NG, Shaughnessy JD, Jr. Fusion of the nucleoporin gene NUP98 to HOXA9 by the chromosome translocation t(7;11)(p15;p15) in human myeloid leukaemia. Nature genetics. 1996 Feb;12(2):154-158.
129. Taketani T, Taki T, Shibuya N, Ito E, Kitazawa J, Terui K, Hayashi Y. The HOXD11 gene is fused to the NUP98 gene in acute myeloid leukemia with t(2;11)(q31;p15). Cancer research. 2002 Jan 1;62(1):33-37.
130. Panagopoulos I, Isaksson M, Billstrom R, Strombeck B, Mitelman F, Johansson B. Fusion of the NUP98 gene and the homeobox gene HOXC13 in acute myeloid leukemia with t(11;12)(p15;q13). Genes, chromosomes & cancer. 2003 Jan;36(1):107-112.

131. Yassin ER, Sarma NJ, Abdul-Nabi AM, Dombrowski J, Han Y, Takeda A, Yaseen NR. Dissection of the transformation of primary human hematopoietic cells by the oncogene NUP98-HOXA9. PloS one. 2009;4(8):e6719.
132. Herblot S, Steff AM, Hugo P, Aplan PD, Hoang T. SCL and LMO1 alter thymocyte differentiation: inhibition of E2A-HEB function and pre-T alpha chain expression. Nat Immunol. 2000 Aug;1(2):138-144.
133. Bash RO, Hall S, Timmons CF, Crist WM, Amylon M, Smith RG, Baer R. Does activation of the TAL1 gene occur in a majority of patients with T-cell acute lymphoblastic leukemia? A pediatric oncology group study. Blood. 1995 Jul 15;86(2):666-676.
134. Armstrong SA, Look AT. Molecular genetics of acute lymphoblastic leukemia. J Clin Oncol. 2005 Sep 10;23(26):6306-6315.
135. Mavrakis KJ, Wolfe AL, Oricchio E, Palomero T, de Keersmaecker K, McJunkin K, Zuber J, James T, Khan AA, Leslie CS, Parker JS, Paddison PJ, Tam W, Ferrando A, Wendel HG. Genome-wide RNA-mediated interference screen identifies miR-19 targets in Notch-induced T-cell acute lymphoblastic leukaemia. Nature cell biology. 2010 Apr;12(4):372-379.
136. Korkmaz KS, Korkmaz CG, Ragnhildstveit E, Kizildag S, Pretlow TG, Saatcioglu F. Full-length cDNA sequence and genomic organization of human NKX3A - alternative forms and regulation by both androgens and estrogens. Gene. 2000 Dec 30;260(1-2):25-36.
137. Deguchi Y, Kehrl JH. Selective expression of two homeobox genes in CD34-positive cells from human bone marrow. Blood. 1991 Jul 15;78(2):323-328.
138. Deguchi Y, Thevenin C, Kehrl JH. Stable expression of HB24, a diverged human homeobox gene, in T lymphocytes induces genes involved in T cell activation and growth. The Journal of biological chemistry. 1992 Apr 25;267(12):8222-8229.
139. Deguchi Y, Moroney JF, Wilson GL, Fox CH, Winter HS, Kehrl JH. Cloning of a human homeobox gene that resembles a diverged Drosophila homeobox gene and is expressed in activated lymphocytes. New Biol. 1991 Apr;3(4):353-363.
140. Deguchi Y, Kehrl JH. High level expression of the homeobox gene HB24 in a human T-cell line confers the ability to form tumors in nude mice. Cancer research. 1993 Jan 15;53(2):373-377.
141. Deguchi Y, Kirschenbaum A, Kehrl JH. A diverged homeobox gene is involved in the proliferation and lineage commitment of human hematopoietic progenitors and highly expressed in acute myelogenous leukemia. Blood. 1992 Jun 1;79(11):2841-2848.
142. Gao H, Le Y, Wu X, Silberstein LE, Giese RW, Zhu Z. VentX, a novel lymphoid-enhancing factor/T-cell factor-associated transcription repressor, is a putative tumor suppressor. Cancer research. 2010 Jan 1;70(1):202-211.
143. Wu X, Gao H, Ke W, Hager M, Xiao S, Freeman MR, Zhu Z. VentX trans-activates p53 and p16ink4a to regulate cellular senescence. The Journal of biological chemistry. 2011 Feb 16.
144. Ferrari N, Palmisano GL, Paleari L, Basso G, Mangioni M, Fidanza V, Albini A, Croce CM, Levi G, Brigati C. DLX genes as targets of ALL-1: DLX 2,3,4 down-regulation in t(4;11) acute lymphoblastic leukemias. J Leukoc Biol. 2003 Aug;74(2):302-305.

145. Campo Dell'Orto M, Banelli B, Giarin E, Accordi B, Trentin L, Romani M, te Kronnie G, Basso G. Down-regulation of DLX3 expression in MLL-AF4 childhood lymphoblastic leukemias is mediated by promoter region hypermethylation. Oncology reports. 2007 Aug;18(2):417-423.
146. Stumpel DJ, Schneider P, van Roon EH, Boer JM, de Lorenzo P, Valsecchi MG, de Menezes RX, Pieters R, Stam RW. Specific promoter methylation identifies different subgroups of MLL-rearranged infant acute lymphoblastic leukemia, influences clinical outcome, and provides therapeutic options. Blood. 2009 Dec 24;114(27):5490-5498.
147. Abate-Shen C. Deregulated homeobox gene expression in cancer: cause or consequence? Nat Rev Cancer. 2002 Oct;2(10):777-785.
148. Greene WK, Bahn S, Masson N, Rabbitts TH. The T-cell oncogenic protein HOX11 activates Aldh1 expression in NIH 3T3 cells but represses its expression in mouse spleen development. Molecular and cellular biology. 1998 Dec;18(12):7030-7037.
149. Muhr J, Andersson E, Persson M, Jessell TM, Ericson J. Groucho-mediated transcriptional repression establishes progenitor cell pattern and neuronal fate in the ventral neural tube. Cell. 2001 Mar 23;104(6):861-873.
150. Smith ST, Jaynes JB. A conserved region of engrailed, shared among all en-, gsc-, Nk1-, Nk2- and msh-class homeoproteins, mediates active transcriptional repression in vivo. Development (Cambridge, England). 1996 Oct;122(10):3141-3150.
151. Lints TJ, Parsons LM, Hartley L, Lyons I, Harvey RP. Nkx-2.5: a novel murine homeobox gene expressed in early heart progenitor cells and their myogenic descendants. Development (Cambridge, England). 1993 Nov;119(3):969.
152. Desjobert C, Noy P, Swingler T, Williams H, Gaston K, Jayaraman PS. The PRH/Hex repressor protein causes nuclear retention of Groucho/TLE co-repressors. Biochem J. 2009 Jan 1;417(1):121-132.
153. Chen G, Fernandez J, Mische S, Courey AJ. A functional interaction between the histone deacetylase Rpd3 and the corepressor groucho in Drosophila development. Genes & development. 1999 Sep 1;13(17):2218-2230.
154. Choi CY, Kim YH, Kwon HJ, Kim Y. The homeodomain protein NK-3 recruits Groucho and a histone deacetylase complex to repress transcription. The Journal of biological chemistry. 1999 Nov 19;274(47):33194-33197.
155. Swingler TE, Bess KL, Yao J, Stifani S, Jayaraman PS. The proline-rich homeodomain protein recruits members of the Groucho/Transducin-like enhancer of split protein family to co-repress transcription in hematopoietic cells. The Journal of biological chemistry. 2004 Aug 13;279(33):34938-34947.
156. Rice KL, Kees UR, Greene WK. Transcriptional regulation of FHL1 by TLX1/HOX11 is dosage, cell-type and promoter context-dependent. Biochem Biophys Res Commun. 2008 Mar 14;367(3):707-713.
157. Riz I, Lee HJ, Baxter KK, Behnam R, Hawley TS, Hawley RG. Transcriptional activation by TLX1/HOX11 involves Gro/TLE corepressors. Biochem Biophys Res Commun. 2009 Mar 6;380(2):361-365.
158. Dayyani F, Wang J, Yeh JR, Ahn EY, Tobey E, Zhang DE, Bernstein ID, Peterson RT, Sweetser DA. Loss of TLE1 and TLE4 from the del(9q) commonly deleted

region in AML cooperates with AML1-ETO to affect myeloid cell proliferation and survival. Blood. 2008 Apr 15;111(8):4338-4347.

159. Fraga MF, Berdasco M, Ballestar E, Ropero S, Lopez-Nieva P, Lopez-Serra L, Martin-Subero JI, Calasanz MJ, Lopez de Silanes I, Setien F, Casado S, Fernandez AF, Siebert R, Stifani S, Esteller M. Epigenetic inactivation of the Groucho homologue gene TLE1 in hematologic malignancies. Cancer research. 2008 Jun 1;68(11):4116-4122.
160. Liu C, Glasser SW, Wan H, Whitsett JA. GATA-6 and thyroid transcription factor-1 directly interact and regulate surfactant protein-C gene expression. The Journal of biological chemistry. 2002 Feb 8;277(6):4519-4525.
161. Nishida W, Nakamura M, Mori S, Takahashi M, Ohkawa Y, Tadokoro S, Yoshida K, Hiwada K, Hayashi K, Sobue K. A triad of serum response factor and the GATA and NK families governs the transcription of smooth and cardiac muscle genes. The Journal of biological chemistry. 2002 Mar 1;277(9):7308-7317.
162. Sepulveda JL, Belaguli N, Nigam V, Chen CY, Nemer M, Schwartz RJ. GATA-4 and Nkx-2.5 coactivate Nkx-2 DNA binding targets: role for regulating early cardiac gene expression. Molecular and cellular biology. 1998 Jun;18(6):3405-3415.
163. Sepulveda JL, Vlahopoulos S, Iyer D, Belaguli N, Schwartz RJ. Combinatorial expression of GATA4, Nkx2-5, and serum response factor directs early cardiac gene activity. The Journal of biological chemistry. 2002 Jul 12;277(28):25775-25782.
164. Weidenfeld J, Shu W, Zhang L, Millar SE, Morrisey EE. The WNT7b promoter is regulated by TTF-1, GATA6, and Foxa2 in lung epithelium. The Journal of biological chemistry. 2002 Jun 7;277(23):21061-21070.
165. Bellavia D, Campese AF, Checquolo S, Balestri A, Biondi A, Cazzaniga G, Lendahl U, Fehling HJ, Hayday AC, Frati L, von Boehmer H, Gulino A, Screpanti I. Combined expression of pTalpha and Notch3 in T cell leukemia identifies the requirement of preTCR for leukemogenesis. Proceedings of the National Academy of Sciences of the United States of America. 2002 Mar 19;99(6):3788-3793.

# CHAPTER 5

## *NOTCH1* and/or *FBXW7* mutations predict for initial good prednisone response but not for improved outcome in pediatric T-cell acute lymphoblastic leukemia patients treated on DCOG or COALL protocols

Linda Zuurbier[1], Irene Homminga[1], Valerie Calvert[2], Mariël L. te Winkel[1], Jessica G.C.A.M. Buijs-Gladdines[1], Clarissa Kooi[1], Willem K. Smits[1], Edwin Sonneveld[3], Anjo J.P. Veerman,[3] Willem A. Kamps,[3,4] Martin Horstmann[5,6], Emanuel F. Petricoin III[2,7], Rob Pieters[1], and Jules P.P. Meijerink[1]

From the [1]Department of Pediatric Oncology/Hematology, Erasmus University Medical Center-Sophia Children's Hospital, Rotterdam, the Netherlands; [2]Center for Applied Proteomics and Molecular Medicine, George Mason University, Manassas, VA, USA; the [3]Dutch Childhood Oncology Group (DCOG), the Hague, the Netherlands; the [4]Department of Pediatric Oncology, University of Groningen-Beatrix Children's Hospital, Groningen, the Netherlands; the [5]German Cooperative Study Group for Childhood Acute Lymphoblastic Leukemia (COALL), Hamburg, Germany; [6]the Research Institute Children's Cancer Center Hamburg, Clinic of Pediatric Hematology and Oncology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; [7]NCI-FDA Clinical Proteomics Program, Food and Drug Administration, Bethesda, MD, USA.

*Leukemia, 2010: 24; 2014-2022*

**ABSTRACT**

Aberrant activation of the NOTCH1 pathway by inactivating and activating mutations in *NOTCH1* or *FBXW7* is a frequent phenomenon in T-cell acute lymphoblastic leukemia (T-ALL). We retrospectively investigated the relevance of *NOTCH1/FBXW7* mutations for pediatric T-ALL patients enrolled on Dutch Childhood Oncology Group (DCOG) ALL7/8 or ALL9 or the German Co-Operative Study Group for Childhood Acute Lymphoblastic Leukemia study (COALL-97) protocols. NOTCH1-activating mutations were identified in 63% of patients. *NOTCH1* mutations affected the heterodimerization, the juxtamembrane and/or the PEST domains, but not the RBP-J-κ-associated module, the ankyrin repeats or the transactivation domain. Reverse-phase protein microarray data confirmed that *NOTCH1 and FBXW7* mutations resulted in increased intracellular NOTCH1 levels in primary T-ALL biopsies. Based on microarray expression analysis, *NOTCH1/FBXW7* mutations were associated with activation of NOTCH1 direct target genes including *HES1*, *DTX1*, *NOTCH3*, *PTCRA* but not *cMYC*. *NOTCH1/FBXW7* mutations were associated with *TLX3* rearrangements, but were less frequently identified in *TAL1*- or *LMO2*-rearranged cases. NOTCH1-activating mutations were less frequently associated with mature T-cell developmental stage. Mutations were associated with a good initial *in vivo* prednisone response, but were not associated with a superior outcome in the DCOG and COALL cohorts. Comparing our data with other studies, we conclude that the prognostic significance for *NOTCH1/FBXW7* mutations is not consistent and may depend on the treatment protocol given.

## INTRODUCTION

T-cell acute lymphoblastic leukemia (T-ALL) accounts for approximately 10–15% of all leukemias in children. Despite improved therapy, still 30% of these cases relapse and ultimately die.[1, 2]

Various chromosomal aberrations are known in T-ALL and some have been associated with prognosis.[3, 4, 5] *NOTCH1* may be important for T-ALL pathogenesis and was initially identified as part of rare t(7;9) translocations.[6, 7] A role for NOTCH1 is now more clear as nearly 60% of T-ALL cases have *NOTCH1* mutations affecting the heterodimerization (HD), the juxtamembrane domain (JM) or the proline, glutamic acid, serine, threonine-rich (PEST) domains.[8, 9] HD or JM mutations result in ligand-independent proteolytical cleavages (reviewed in Grabher *et al.*[10]), resulting in the release of intracellular NOTCH1 (ICN). ICN is a transcription factor that regulates differentiation and proliferation through the activation of various target genes including *cMYC*, *HES1* and *PTCRA*.[10, 11, 12]

As an alternative NOTCH1 activation mechanism, inactivating mutations in the F-Box WD40 domain containing protein 7 gene (*FBXW7*) were identified in 8–30% of T-ALL patients.[13, 14, 15, 16] FBXW7 is part of the E3 ubiquitin ligase complex that controls the turnover of various proteins including ICN. FBXW7 interacts with phosphodegron domains located in the PEST domain of ICN. Therefore, inactivating mutations in *FBXW7* or loss of the phosphodegron domains through truncating *NOTCH1* PEST mutations both result in the stabilization of ICN in the nucleus. Mutations in *FBXW7* and *NOTCH1* PEST mutations are mutually exclusive,[13, 14, 16] indicating that they seem to exert an equivalent oncogenic effect.

Mutations in *NOTCH1* or *FBXW7* may have prognostic relevance in T-ALL. Breit *et al.*[17] reported that *NOTCH1* mutant pediatric patients in the German ALL-BFM 2000 study show a good *in vivo* prednisone response and have an improved event-free survival (EFS). In contrast, Zhu *et al.*[18] published an unfavorable outcome for *NOTCH1*-mutated adult T-ALL patients, but not for pediatric patients. We could not confirm a favorable prognostic effect for *NOTCH1*-mutated pediatric T-ALL patients treated on Dutch Childhood Oncology Group (DCOG) protocols,[19] and this was confirmed by children treated on POG protocols for which no relation was identified between the presence of *NOTCH1* mutations and relapse.[20] These initial studies investigated the relevance for *NOTCH1* HD and PEST mutations,[17, 18, 19] but did not include *NOTCH1* JM mutations or *FBXW7* mutations. We now extended our initial study by examining the prognostic effect of *NOTCH1* and *FBXW7* mutations in 141 pediatric T-ALL patients treated on DCOG or German Co-

Operative Study Group for Childhood Acute Lymphoblastic Leukemia study (COALL-97) protocols. The functional consequences of *NOTCH1/FBXW7* mutations in relation to ICN levels and activation of target genes in primary leukemia samples were investigated.

## PATIENTS/MATERIALS AND METHODS

### Patient samples

This study comprised 146 primary pediatric T-ALL patients, of which 72 were treated on DCOG protocols ALL-7/8,[21, 22] (*n*=30) or ALL-9 (*n*=42).[23] This cohort had a median follow-up of 67 months, and included 51 male and 21 female patients. As the overall disease-free survival for patients treated on these DCOG protocols are comparable, these patients will be analyzed as one cohort as carried out before.[19, 24] Of these, 70 patients were part of our previous study.[19] For ALL7/8 patients, *in vivo* prednisone response was monitored at day 8 following 7 days of BFM-like prednisone monotherapy and one intrathecal dose of methotrexate. A clearance to less than 1000 blasts per μl blood at day 8 was considered as an initial prednisone good response (PGR). In total, 74 patients were enrolled in the German COALL-97 protocol[19] with a median follow-up of 52 months. This cohort included 49 male and 25 female patients. The patients' parents or legal guardians provided informed consent to use leftover diagnostic biopsies for research in accordance with the institutional review board and the Declaration of Helsinki Principles. Isolation of leukemia cells from blood or bone marrow samples has been described before,[25] and all samples contained >90% of leukemic blasts. Clinical and immunophenotypic data were supplied by both study centers. Classification into T-cell development stages was based on EGIL criteria:[26] pro-/pre- ($CD7^+$, $CD2^+$ and/or $CD5^+$ and/or $CD8^+$), cortical ($CD1^+$) or mature T-cell stage ($sCD3^+/CD1^-$).

### Genomic DNA and RNA extraction

Isolation of genomic DNA and RNA from $5 \times 10^6$ leukemic cells using the TRIzol reagent (Invitrogen, Breda, the Netherlands) and copy DNA synthesis were carried out as described before.[24, 25]

### Mutational detection

*NOTCH1* exons 25–34 were screened for mutations that include all relevant domains (Supplementary Table S1). For *FBXW7*, the F-box and WD40 domains (exon 5, exons 7–11) were amplified, covering all *FBXW7* mutations as reported so

far. PCR reactions were carried out as described before.[19] Primers are shown in Table 1. PCR products were sequenced using the BigDye Terminator version 3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) on a 3130 DNA Analyzer (Applied Biosystems).

**Identification of recurrent rearrangements by FISH, RQ-PCR or array-CGH**
*SIL-TAL*, *CALM-AF10* or rearrangements of *LMO2*, *TLX1*, *TLX3*, *TAL1*, *CALM-AF10, SET-NUP214, HOXA* or *MLL* were determined with fluorescence *in-situ* hybridization (FISH) as previously described.[24, 27, 28] *NOTCH1* translocations were detected using bacterial artificial chromosomes clones RP11-769N4, RP11-1008C19, RP11-83N9 and RP11-662J2 covering both sides adjacent to the *NOTCH1* locus. Bacterial artificial chromosomes were obtained from BAC/PAC Resource Center (Children's Hospital, Oakland, CA, USA). Expression levels of *TLX1*, *TLX3*, *TAL1*, *LMO2* or *HOXA* or *CALM-AF10* and *SET-NUP214* fusion products were measured relative to the expression of glyceraldehyde-3-phosphate dehydrogenase as described before.[27, 28] Array-CGH analysis was performed as previously described,[27] on the human genome CGH Microarray 105 or 400K dual arrays (Agilent Technologies, Santa Clara, CA, USA), which consists of ~105 000 or ~400 000 60-mer oligonucleotide probes that span both coding and noncoding sequences with an average spatial resolution of ~15 or 5 kb. Microarray images were analyzed using feature extraction software (version 8.1; Agilent Technologies) and the data were subsequently imported into array-CGH analytics software version 3.1.28 (Agilent Technologies).

**Gene expression array analysis**
RNA integrity testing, copy DNA and copy-copy RNA (ccRNA) syntheses, hybridization and washing to Human Genome U133 plus2.0 microarrays (Affymetrix, Santa-Clara, CA, USA), extraction of probeset intensities from CEL-files and normalization with RMA or VSN methods were performed as described before.[28] Differentially expressed genes between *NOTCH1* mutant versus wild-type T-ALL patients were determined by Wilcoxon statistics and corrected for multiple testing error[29] using the Bioconductor package 'Multtest' in *R*. Heatmaps based on the TOP50 most significant differentially expressed genes were performed in Dchip software, Harvard University, Boston, MA, USA.[30] Microarray data are available at http://www.ncbi.nlm.nih.gov.proxy-ub.rug.nl/geo/.

**Reverse-phase protein microarray analysis and western blot**

Reverse-phase protein microarray construction and analysis was performed essentially as previously described.[31, 32] To isolate proteins from $10 \times 10^6$ leukemic cells, lysis was performed in 20 µl tissue protein extraction reagent (Pierce Biotechnology, Rockford, IL, USA) with 300 nM NaCl, 1 mM orthovanadate and protease inhibitors. Cells were incubated at 4 °C for 20 min and subsequently centrifuged at 10,000 r.p.m. for 5 min in an Eppendorf centrifuge. Supernatants were stored at −80 °C before printing on the microarrays. Lysates were diluted to 1.0 mg/ml protein concentration and mixed 1:1 with 2 × SDS Tris-glycine buffer (Invitrogen) containing 5% 2-mercaptoethanol (Sigma, Zwijndrecht, the Netherlands) (FC=0.5 mg/ml). Lysates were spotted at a concentration of 0.5 µg/µl (neat spot) and 0.125 µg/µl in duplicate with 350 µm pins on glass-backed nitrocellulose coated array slides (FAST slides; Whatman, Kent, UK) using an Aushon Biosystems 2470 (Aushon Biosystems, Billerica, MA, USA). Printed slides were stored at −20 °C or directly used. The first of each 25 slides printed were subjected to Sypro Ruby Protein Blot staining (Invitrogen) to determine total protein amount. These slides were visualized on a NovaRay CCD fluorescent scanner (Alpha Innotech, San Leandro, CA, USA). The remaining slides were used for staining with a specific antibody. Before this, slides were incubated with 1 × Reblot (Chemicon, Temecula, CA, USA) for 15 min and subsequently washed with phosphate-buffered saline twice. This was continued with a blocking procedure for 5 h using 1 gr I-Block (Applied Biosystems) diluted in 500 ml phosphate-buffered saline with 0.5% Tween 20. Slides were stained with an automated slide stainer (Dako, Glostrup, Denmark) according to the manufacturer's instructions using the Autostainer catalyzed signal amplification kit (Dako). In each staining run, a negative control slide was stained with the secondary antibody only for background subtraction. Briefly, endogenous biotin was blocked for 10 min with the biotin blocking kit (Dako), followed by application of protein block for 5 min; primary antibodies were diluted in antibody diluent and incubated on slides for 30 min and biotinylated secondary antibodies were incubated for 15 min. Signal amplification involved incubation with a streptavidin/biotin/peroxidase complex provided in the catalyzed signal amplification kit for 15 min, and amplification reagent (biotinyl-tyramide/hydrogen peroxide, streptavidin/peroxidase) for 15 min each. A signal is generated using streptavidin-conjugated IRDye680 (LI-COR Biosciences, Lincoln, NE, USA). Slides were allowed to air dry following development. Stained slides were scanned individually on the NovaRay scanner (Alpha Innotech) and files

were saved in TIF format in Photoshop 7.0. All slides were subsequently analyzed with the MicroVigene version 2.8.1.0 program (VigeneTech, Carlisle, MA, USA). To screen for ICN protein levels, we have used and optimized the conditions for the ICN Val1744 antiserum (catalog no. 2421; Cell Signaling Technology, Beverly MA, USA). Slides were scanned in a NovaRay scanner (Alpha Innotech) and analyzed with the MicroVigene version 2.8.1.0 program (VigeneTech). For western blot validation,[28] protein loading was validated by staining for Actin (Sigma, catalog no. 2547)

**Statistics**

Statistics were performed using SPSS 15.0 software (SPSS Inc., Chicago, IL, USA). The Pearson's $\chi^2$-test or the Fisher's exact test was used to test differences in the distribution of nominal data as indicated. Statistical significance for continuous distributed data was tested using the Mann—Whitney's *U*-test. Differences between patient populations in EFS and relapse-free survival (RFS) were tested by using the log-rank test. For RFS, an event is defined as relapse or nonresponse toward induction therapy at day 56 (COALL) or at start of consolidation therapy (DCOG). An event for EFS is defined as relapse, nonresponse toward induction therapy, death in remission because of toxicity or development of a secondary malignancy.

***NOTCH1* and/or *FBXW7* mutations in pediatric T-ALL patients**

Bone marrow or blood DNA samples for 146 primary T-ALL patients were analyzed for *NOTCH1* (exons 25–34) and/or *FBXW7* mutations (exons 5, 7–11) and 141 samples were successfully amplified and sequenced. The locations of mutations in specific NOTCH1 or FBXW7 domains are shown in Figure 1a.

*NOTCH1* and *FBXW7* mutations in pediatric T-ALL patients. (**a**) Schematic representation of identified mutations in the heterodimerization (HD), juxtamembrane (JM) and PEST domains in NOTCH1 and in the WD40-repeats of FBXW7. Missense mutations are indicated by an open triangle, a silent mutation is indicated by a filled gray triangle and nonsense mutations are indicated by a filled black triangle. (**b**) The distribution of *NOTCH1* and *FBXW7* mutation types in the DCOG and COALL cohorts.

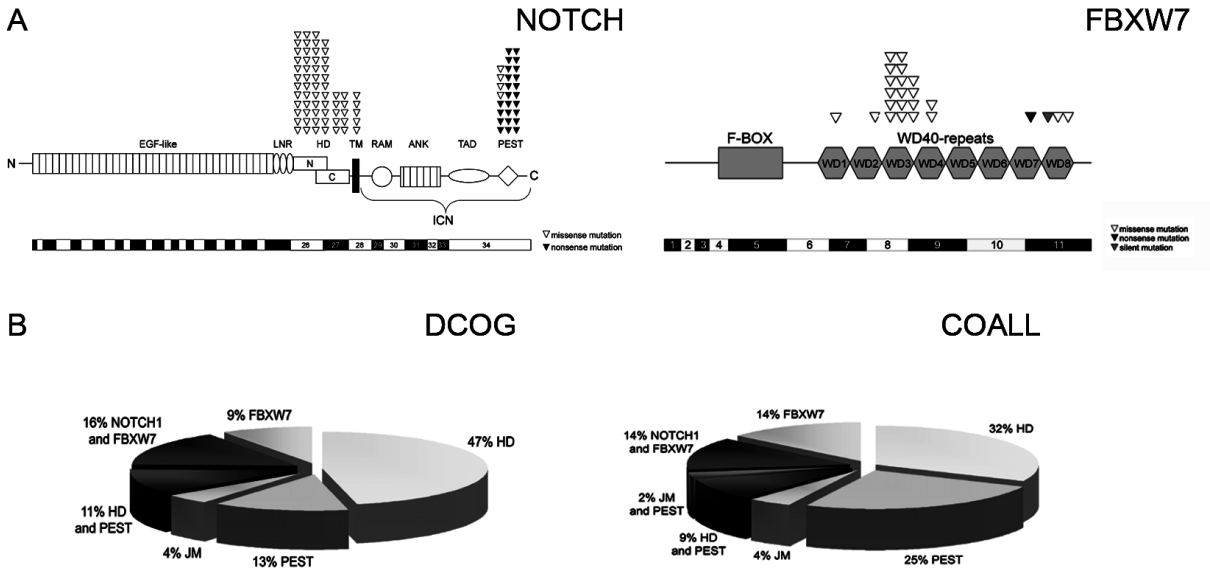



**Figure 1. NOTCH1 and FBXW7 mutations in pediatric T-ALL patients.** (A) Schematic representation of identified mutations in the heterodimerization (HD), juxtamembrane (JM) and PEST domains in NOTCH1 and in the WD40-repeats of FBXW7. Missense mutations are indicated by an open triangle, a silent mutation is indicated by a filled grey triangle, and nonsense mutations are indicated by a filled black triangle (B) The distribution of NOTCH1 and FBXW7 mutation types in the DCOG and COALL cohorts.

Heterozygous mutations in *NOTCH1* were detected in 79 out of 141 cases (56%), whereas 23 T-ALL patients (16%) harbored a point mutation in *FBXW7*. In total, 89 patients (63%) contained *NOTCH1* and/or *FBXW7* mutations. In total, 35 patients (39%) had a missense mutation or an in-frame insertion/deletion in the HD-domain of *NOTCH1*, whereas 9 (10%) and 13 (15%) patients harbored a combination of HD and PEST or *FBXW7* mutations, respectively. Seventeen patients (19%) had a single *NOTCH1* PEST mutation and ten (11%) had a single *FBXW7* mutation (Figure 1b). We confirmed that *NOTCH1* PEST domain mutations and *FBXW7* mutations were nearly mutual exclusive,[14, 16] but one patient carried a *FBXW7* and a *NOTCH1* PEST mutation. Five patients had a mutation in the JM domain of *NOTCH1* (5.6%) of which one also had a *NOTCH1* HD mutation. It is not known whether these JM and HD mutations occurred in *cis* or affected different alleles.

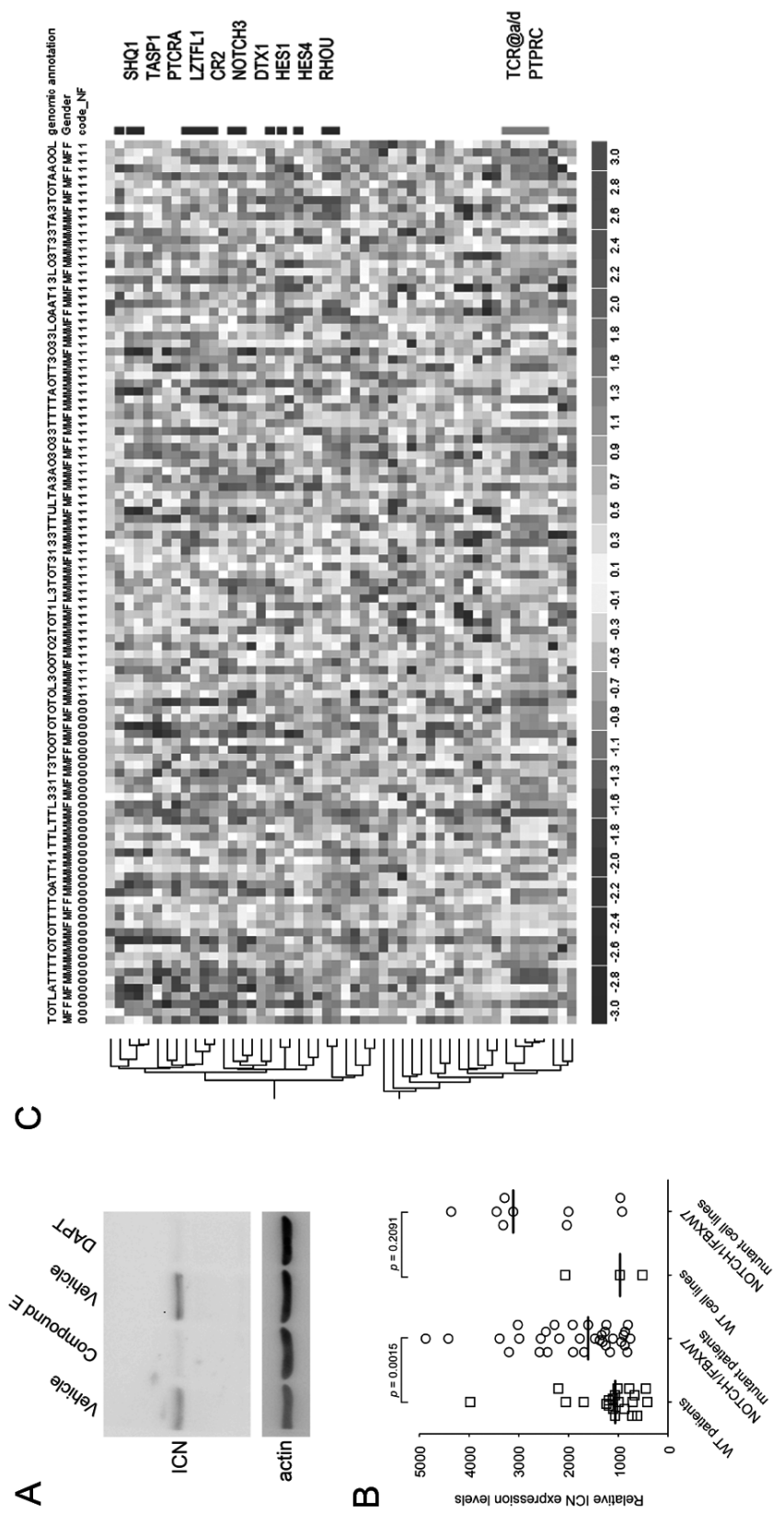
In total, 66 *NOTCH1* mutations were found and 10 HD and 9 PEST mutations were not reported before to the best of our knowledge (Supplementary Figure S1). Ten *FBXW7* point mutations were found, five of which have not been

observed before in T-ALL (Supplementary Figure S2). These are H379L in exon 7, R465P in exon 8 and K622STOP, G687V and E693K in exon 11. The E693K mutation was previously identified in a gastric carcinoma patient.[33]

***NOTCH1* and/or *FBXW7* mutations activate ICN and downstream target genes in primary T-ALL samples**

As published for T-ALL cell lines,[8, 9, 11, 12, 14, 16] we demonstrated by using reverse-phase protein microarrays that *NOTCH1* and/or *FBXW7* mutations result in enhanced levels of ICN in primary T-ALL cells. The specificity of the NOTCH1 antibody was validated on the T-ALL cell line HPB-ALL, and ICN detection was lost on treatment with a γ-secretase inhibitor (Figure 2a). *NOTCH1* and/or *FBXW7*-mutated patients showed about twofold higher ICN levels compared with wild-type patients (Figure 2b, $P$=0.0015). Strikingly, four wild-type patients also showed high ICN levels despite the absence of *NOTCH1* and/or *FBXW7* mutations (Figure 2b). Subsequent FISH and array-CGH analyses ruled out potential *NOTCH1* translocations or other chromosomal *NOTCH1* rearrangements in these four patients (data not shown).

**Figure 2. (page 126) NOTCH1/FBXW7 mutations activate the NOTCH1 pathway in primary T-ALL patient biopsies.** (A) Western blot analysis of lysates from the HPBALL T-ALL cell line which is NOTCH1-mutated. Treatment for 96 hrs with γ-secretase inhibitors including compound E (100 nM) or DAPT (5 μM), results in loss of activated intracellular NOTCH1 expression (ICN). Actin was used as loading control. (B) NOTCH1 ICN levels in wild-type and NOTCH1 and/or FBXW7-mutated T-ALL patients and T-ALL cell lines analyzed with Reverse-phase Protein microarray. NOTCH1/FBXW7 wild-type patient samples with high ICN levels are marked by an asterisk (C) Heatmap showing the TOP50 most differentially expressed genes between NOTCH1 and/or FBXW7 mutant patients versus wild-type patients. NOTCH1 direct target genes are indicated. Annotations indicated are genetic rearrangements, Gender and NOTCH1/FBXW7 mutation status. Genetic rearrangements indicated are: T, TAL1 or SIL-TAL1; L, LMO1 or LMO2 (includes del(11)(p12p13)); A, HOXA-activated (includes cases with SET-NUP214; CALM-AF10 or Inv(7)(p15q34)); 1, TLX1; 2, TLX2; 3, TLX3; O, Other; U, Aberration unknown. Gender is indicated F, Female or M, Male. NOTCH1/FBXW7 mutation status is indicated 0, wild-type and 1, NOTCH1 and/or FBXW7-mutated; NOTCH1/FBXW7 wild-type patients with high ICN levels are marked by an asterisk; NOTCH1/FBXW7 wild-type patients having a NOTCH1 signature that cluster with NOTCH1-activated patients based upon hierarchical clustering based on the TOP50 probeset are indicated with a filled triangle.


A
Vehicle
Compound E
Vehicle
DAPT
ICN
actin
B
Relative ICN expression levels
5000
4000
3000
2000
1000
0
p = 0.0015
p = 0.2091
WT patients
NOTCH1/FBXW7 mutant patients
WT cell lines
NOTCH1/FBXW7 mutant cell lines
C
SHQ1
TASP1
PTCRA
LZTFL1
CR2
NOTCH3
DTX1
HES1
HES4
RHOU
TCR@a/d
PTPRC
genomic annotation
Gender
code_NF
-3.0 -2.8 -2.6 -2.4 -2.2 -2.0 -1.8 -1.6 -1.3 -1.1 -0.9 -0.7 -0.5 -0.3 -0.1 0.1 0.3 0.5 0.7 0.9 1.1 1.3 1.6 1.8 2.0 2.2 2.4 2.6 2.8 3.0

We investigated whether *NOTCH1/FBXW7* mutations would result in the activation of specific genes. Expression array data[28, 34] were available for 111 T-ALL patients with a known *NOTCH1/FBXW7* mutation status. The TOP50 most significant and differentially expressed genes (probesets) between *NOTCH1*/*FBXW7* mutant and wild-type patients comprised previous published and validated NOTCH1 direct target genes including *HES1*, *HES4*, *DTX1*, *PTCRA*, *NOTCH3*, *PTPRC*, *CR2*, *LZTFL1*, *TASP1*, *SHQ1* and *RHOU* (Figure 2c).[10, 11] Although *cMYC* is a NOTCH1 target gene in T-ALL cell lines, this gene did not appear in our TOP50 nor TOP200 gene lists (not shown). Eight wild-type patients also seemed to express genes from this NOTCH1 signature (Figure2c; data not shown). For six out of these eight patients for which ICN levels were available, two patients were among the four wild-type cases having the highest ICN protein levels. Similarly to these two cases having a NOTCH1 signature and high ICN levels, none of the remaining six patients with a NOTCH1 signature carried *NOTCH1* translocations or alternative chromosomal abnormalities based on FISH and array-CGH results (data not shown).

### *NOTCH1*/*FBXW7* mutations in relation to clinical, immunophenotypic and cytogenetic parameters

We did not observe a relationship between *NOTCH1/FBXW7* mutations with gender, age or white blood cell counts (Table 1). For 23 patients, the *in vivo* prednisone response was known. NOTCH1-activated patients were correlated with a good *in vivo* prednisone response as 14 out of 16 patients with an initial PGR contained *NOTCH1* mutations, in contrast to 2 out of 7 cases with a poor response ($P$=0.01). This observation was stronger by including *FBXW7* data where 15 out of 16 cases with a PGR had a *NOTCH1*/*FBXW7* mutation in contrast to only 2 out of 7 prednisone poor response cases ($P$=0.003, Table 1). Classification into T-cell development stages on EGIL criteria[26] revealed that *NOTCH1*/*FBXW7* mutations were less frequently identified in mature T-ALL cases ($P$=0.05, Table 1). In relation to molecular cytogenetic data, *NOTCH1/FBXW7* mutations were identified in all cytogenetic T-ALL subgroups (Table 1). Considering *TAL1*- or *LMO2*-rearranged cases as a single *TAL/LMO* entity based on their identical expression profiles,[28, 34] and including an additional 19 *TALLMO*-like patients with a *TAL/LMO* signature that lack *TAL1* or *LMO2* rearrangements,[28] *NOTCH1* mutations were less frequent. Only 25 out of 60 *TALLMO* patients (42%) had a *NOTCH1* mutation ($P$=0.002, Table 1). This remained significant when including

*FBXW7* mutations as only 30 out of 60 cases (50%) had a *NOTCH1*/*FBXW7* mutation (*P*=0.004). *NOTCH1* mutations were more prevalent in *TLX3*-rearranged cases, in which 21 out of 27 cases (86%) had a *NOTCH1* mutation (*P*=0.02). This remained significant when taking *FBXW7* mutations into account (*P*=0.01).

**Prognostic relevance of *NOTCH1* and/or *FBXW7* mutations**

We then investigated the relevance of *NOTCH1* and/or *FBXW7* mutations in relation to treatment outcome. For the DCOG cohort, mutations in *NOTCH1* and/or *FBXW7* tended toward poor treatment outcome. The 5-year EFS rates for patients with *NOTCH1* mutations only compared with wild-type patients were 57±8% versus 76±8% (*P*=0.08) for the DCOG cohort but 63±8 versus 64±10% for the COALL cohort (*P*=0.99, Figures 3a and b). Inclusion of *FBXW7* mutations resulted in 5-year EFS rates of 58±7 versus 74±9% (*P*=0.16) for the DCOG cohort and 63±8 versus 68±10% for the COALL cohort (*P*=0.90; data not shown).

Events in both cohorts are summarized in Table 2. *NOTCH1* mutations tended toward a lower RFS in the DCOG (*P*=0.068) and COALL cohorts (*P*=0.094) with 5-year RFS of 83±7 versus 62±8% for the DCOG cohort and 89±6 versus 67±8% for the COALL cohort for wild-type and *NOTCH1*-mutated patients, respectively (Figures 3c and d). These trends became less evident when including *FBXW7* mutation data, with an RFS of 82±8 versus 62±8% in the DCOG cohort (*P*=0.101) and an RFS of 86±7 versus 70±8% in the COALL cohort (*P*=0.23) for wild-type or *NOTCH1*-activated patients (data not shown).

**Table 1. The distribution of wild-type and NOTCH1 mutations within categorized subgroups of T-ALL patients**

| Clinical | Comparison | *NOTCH1* mutation | | | | | *FBXW7* mutation | | | | | *NOTCH1/FBXW7* mutations | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | WT n | % | Mut n | % | p-value | WT n | % | Mut n | % | p-value | WT n | % | Mut n | % | p-value |
| gender | Male-Female distr. | | | | | 1.0 | | | | | 0.81 | | | | | 1 |
| | Male | 24 | *43* | 55 | *57* | | 81 | *84* | 16 | *16* | | 35 | *36* | 62 | *64* | |
| | Female | 19 | *43* | 25 | *57* | | 36 | *82* | 8 | *18* | | 16 | *36* | 28 | *64* | |
| Age (years) | Age distribution | 7.2 | | 7.8 | | 0.30* | 7.5 | | 8.2 | | 0.29* | 7 | | 7.8 | | 0.09* |
| WBC (x$10^9$ cells/liter) | WBC distribution | 131 | | 119 | | 0.69* | 126 | | 109 | | 0.28* | 134 | | 110 | | 0.41* |
| In vivo prednisone response | PGR or PPR | | | | | 0.01 | | | | | NE | | | | | .003 |
| | PGR | 2 | | 14 | | | | | | | | 1 | | 15 | | |
| | PPR | 5 | | 2 | | | | | | | | 5 | | 2 | | |
| **Immunophenotype** | | | | | | | | | | | | | | | | |
| Pre-T/Pro-T | Pre-T/Pro-T vs other | 16 | *41* | 23 | *59* | 0.81 | 34 | *87* | 5 | *13* | 0.42# | 13 | *33* | 26 | *67* | 0.68 |
| Cortical T | Cortical T vs other | 21 | *36* | 37 | *64* | 0.19 | 47 | *81* | 11 | *19* | 0.58 | 16 | *28* | 42 | *72* | 0.08 |
| Mature T | Mature T vs other | 21 | *54* | 18 | *46* | 0.18 | 32 | *83* | 7 | *17* | 0.89 | 20 | *51* | 19 | *49* | 0.05 |
| **Cytogenetics** | | | | | | | | | | | | | | | | |
| TAL1 | TAL1+ vs TAL1- | 13 | *50* | 13 | *50* | 0.51 | 23 | *88* | 3 | *12* | 0.57# | 13 | *50* | 13 | *50* | 0.12 |
| LMO2 | LMO2+ vs LMO2- | 7 | *50* | 7 | *50* | 0.78 | 11 | *79* | 3 | *21* | 0.71# | 5 | *36* | 9 | *64* | 1# |
| TAL/LMO | TAL/LMO+ vs | 35 | *58* | 25 | *42* | 0.002 | 51 | *85* | 9 | *15* | 0.66 | 30 | *50* | 30 | *50* | 0.004 |
| TLX3 | TLX3+ vs TLX3- | 6 | *22* | 21 | *78* | 0.02 | 19 | *70* | 8 | *30* | 0.08 | 4 | *15* | 23 | *85* | 0.01# |
| TLX1 | TLX1+ vs TLX1- | 3 | *43* | 4 | *57* | 1# | 7 | *100* | 0 | *0* | 0.60# | 3 | *43* | 4 | *57* | 0.70# |
| HOXA | HOXA+ vs HOXA- | 4 | *31* | 9 | *69* | 0.39# | 10 | *77* | 3 | *23* | 0.46# | 3 | *23* | 10 | *77* | 0.38# |
| Others | Others vs cytogenetic annotated | 13 | *39* | 20 | *61* | 0.69# | 29 | *88* | 4 | *12* | 0.60# | 11 | *33* | 22 | *67* | 0.84 |

Only significant *p*-values are indicated. #, *p*-values calculated by the Fisher exact method; *, *p*-values calculated by the Mann-Whitney-U method. WBC, white blood cell count; PGR, prednisone good response; PPR, prednisone poor response, NE, not evaluable.

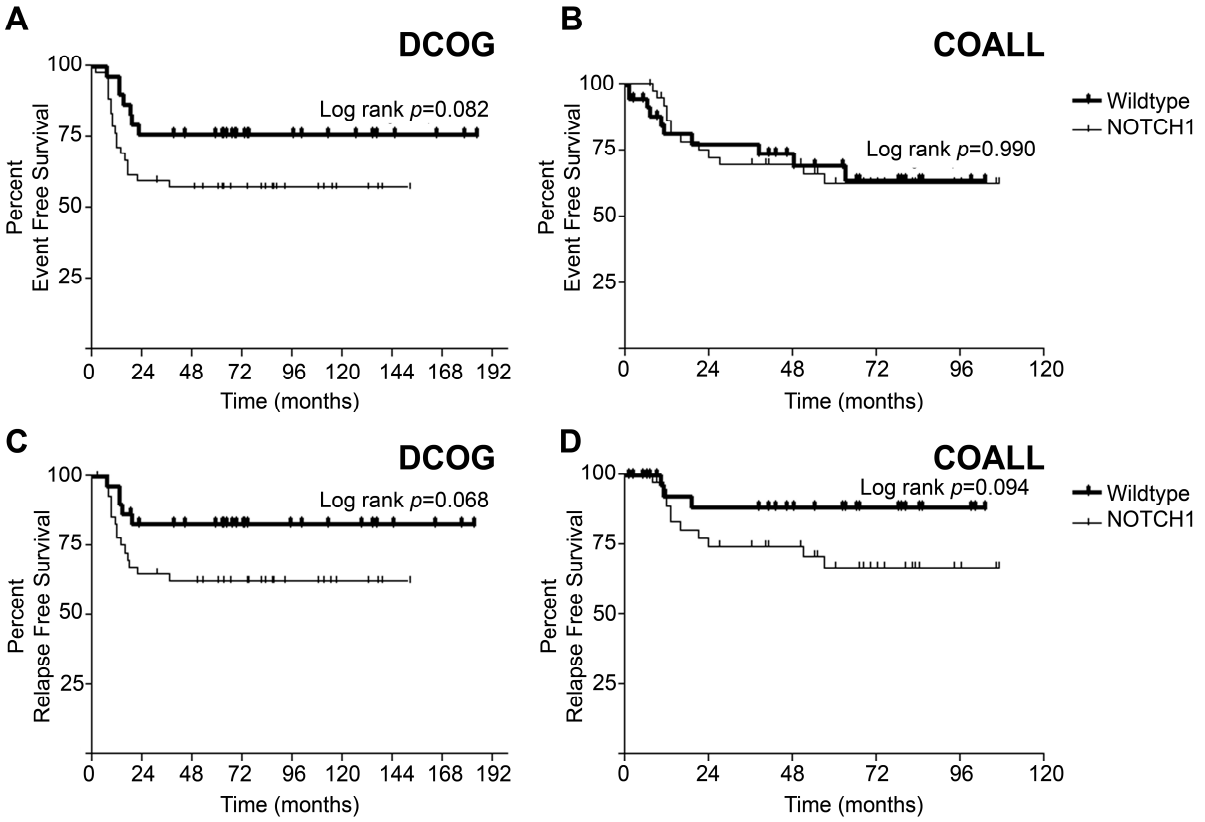



**Figure 3. NOTCH1-activating mutations have no prognostic implication in pediatric T-ALL.** Event-free survival (EFS) (**a**, **b**) and Relapse-free survival (RFS) (**c**, **d**) for the pediatric T-ALL patients treated on DCOG protocols (**a**, **c**) or the COALL protocol (**b**, **d**). Patients carrying *NOTCH1* and/or *FBXW7* mutations and wild-type patients have been indicated.

We also investigated the effect for specific *NOTCH1* and/or *FBXW7* mutations on the activation of downstream target genes and outcome. As reported by the group of Pear and co-workers, specific *NOTCH1* mutations or combinations of *NOTCH1*/*FBXW7* mutations may have strong NOTCH1-activating effects, whereas others may only have modest activating effects.[35] For this, we distinguished weak NOTCH1-activating mutations, that is, *NOTCH1* HD or PEST mutations or *FBXW7* mutations, and strong NOTCH1-activating mutations, that is, *NOTCH1* JM mutations or combinations of *NOTCH1* HD mutations with PEST mutations or FBXW7 mutations. Although ICN protein levels were significantly higher for *NOTCH1*- and/or *FBXW7* -mutated cases versus wild-type cases, there was relation to the types of NOTCH1-activating mutations investigated (Supplementary Figure S3). To investigate differential activation of downstream target genes between

patients with weak or strong NOTCH1-activating mutations, we first calculated the most significantly and differently expressed genes (probesets) between patients

**Table 2. Events in DCOG and COALL cohorts**

| | DCOG | | COALL | |
|---|---|---|---|---|
| | WT *n*=29 | Mut *n*=42 | WT *n*=35 | Mut *n*=39 |
| events | 7 | 18 | 10 | 13 |
| relapse* | 5 | 15 | 3 | 11 |
| CNS relapse | 2 | 2 | 1 | 4 |
| Toxic death | 1 | 2 | 5 | 3 |
| second malignancy | 1 | 1 | 2 | 1 |

WT, wild-type; Mut, mutant for NOTCH1; *, includes CNS relapse

with strong NOTCH1-activating mutations versus wild-type patients, which again revealed mostly bonafide NOTCH1 target genes. However, these genes were expressed at intermediate levels for patients having weak NOTCH1-activating mutations (Supplementary Figure S4), indicating that these types of mutations indeed differ in their potential to activate downstream target genes in primary leukemic samples. Distinction between these types of mutations may also have prognostic significance as patients from the DCOG cohort with strong NOTCH1-activating mutations had a significant poor outcome relative to wild-type patients ($P$=0.012) as well as to patients carrying weak NOTCH1-activating mutations ($P$=0.048) (Supplementary Figure S5a). However, this observation could not be substantiated for COALL-97 T-ALL patients (Supplementary Figure S5b). We also investigated whether ICN protein levels itself had prognostic significance. As 55 out of 66 patients for whom ICN protein levels were available were treated on the COALL cohort, we divided these patients into quartiles and determined their RFS and EFS rates. However, no relationship between ICN protein levels and RFS or EFS was present ($P$=0.98 and 0.97, respectively).

## DISCUSSION

Activation of NOTCH1 as a consequence of activating *NOTCH1* mutations or inactivating *FBXW7* mutations is a frequent phenomenon in T-ALL.[8] We screened for *NOTCH1* and *FBXW7* mutations in 141 pediatric T-ALL patient samples and identified *NOTCH1* mutations in 56% and *FBXW7* mutations in 16% of the

patients. In total, 63% of the patients had an aberrantly activated NOTCH1 pathway due to mutations. In line with previous studies,[14, 16] we observed that *NOTCH1* PEST domain mutations and *FBXW7* mutations occurred in a mutually exclusive manner with the exception of one patient. This patient had a nonsense mutation in *FBXW7* in contrast to missense mutations that are normally observed in *FBXW7*-mutated patients. This implies that mutant FBXW7 but not truncated FBXW7 proteins exert a dominant-negative effect in the E3-ubiquitin ligase complex. Interestingly, Park *et al.*[15] also discovered a nonsense mutation due to a 5 bp insertion in *FBXW7* in combination with a *NOTCH1* PEST mutation in a non-Hodgkin's lymphoma patient.

The frequency of NOTCH1-activating mutations is in line with other studies also comprising adult T-ALL patient series.[36, 37] In adult studies, *NOTCH1* and *FBXW7* mutations were identified in 60–62% and 18–24% of the T-ALL patients, respectively. This indicates that the oncogenic role for *NOTCH1/FBXW7* during T-cell oncogenesis remains conserved over age. We did not find evidence for mutations outside the *NOTCH1* HD, JM and PEST domains in any of the 141 pediatric T-ALL patients indicating that reported mutations in the LNR region, the RAM, ANK and TAD domains are very rare.[18, 36, 38]

We found that *NOTCH1* and *FBXW7* mutations resulted in increased levels of cleaved NOTCH1 (ICN) in primary leukemia cells and was associated with the activation of NOTCH1 target genes,[10, 11, 12] including *HES1*, *HES4*, *DTX1*, *PTCRA*, *NOTCH3*, *PTPRC*, *CR2*, *LZTFL1*, *TASP1* and *RHOU*. This confirms that the mutations manifest functionally at the protein level in patient samples. We identified 10 patients that lacked *NOTCH1* and/or *FBXW7* mutations that either expressed high levels of ICN or that expressed NOTCH1 target genes. As we did not find chromosomal translocations or other types of rearrangements involving the *NOTCH1* locus, this implies that additional mutation mechanisms in *NOTCH1* or directly downstream regulatory genes must exist that so far been left unnoticed in T-ALL. Although *cMYC* was identified as a prominent NOTCH1 target in T-ALL cell lines,[11, 12] it was not identified as target gene in primary samples. However, two cases expressed ectopic cMYC levels due to a t(8;14)(q24;q11) translocation, which were both wild-type for *NOTCH* and *FBXW7*, supporting a role for *MYC* as NOTCH1 target. Further research will be required to establish whether cMYC is generally upregulated by means of other oncogenic mechanisms in addition to activated *NOTCH1* in primary samples and therefore left undetected, or that the expression of cMYC is rapidly lost on isolation of primary leukemic cells.

*NOTCH1/FBXW7* mutations were identified at a lower frequency in T-ALL cases with a mature immunophenotype. This may explain the low incidence of *NOTCH1/FBXW7* mutations in the *TAL/LMO* subgroup because *TAL1* rearrangements, which are the most recurrent abnormality in this subgroup, are associated with a mature T-cell development arrest.[24, 39] This is an interesting finding and suggests that the oncogenic role of *NOTCH1* is less prominent in T-ALL cases arrested at a relative mature T-cell developmental stage. Interestingly, *NOTCH1/FBXW7* mutations were identified at a higher frequency in *TLX3*-rearranged T-ALL. The oncogenic activation of NOTCH1 thus far has been regarded as one of the earliest acquired abnormalities in a preleukemic progenitor cell that therefore becomes committed to the T-ALL.[10, 40] In this perspective, our data indicate that the importance of deregulated NOTCH1 as initiating event during T-cell oncogenesis depends on additional collaborating events like *TLX3* or *TAL1* rearrangements. It also suggests that the oncogenic program that is followed by T-ALL cases that eventually arrest at the mature development stage may be less dependent on NOTCH1. Whether NOTCH1-activating mutations represent truly initiating leukemic events or not needs to be established, as evidence is emerging that NOTCH1 activation in some T-ALL cases may have occurred as a secondary event which may be acquired or lost at relapse.[41]

In the study of Breit *et al.*,[17] *NOTCH1* mutations were associated with an initial PGR and a significantly lower minimal residual disease content at day 78. Our study supports this association with initial prednisone response for *NOTCH1/FBXW7* mutant patients. This association is also validated for patients of the EORTC-CLG study. In that study, NOTCH1-activating mutations were also associated with reduced minimal residual disease during therapy.[42] The association for NOTCH1-activating mutations with PGR seems to be in contrast with the finding that γ-secretase inhibitors can sensitize for glucocorticoids in glucocorticoid-resistant cells.[43] It may be that the NOTCH pathway has opposing effects in the glucocorticoid response in responsive against resistance patients, but it now seems clear that activation of NOTCH1 by mutations does not drive glucocorticoid resistance. Further research will be required to clarify this seeming contradiction.

*NOTCH1* mutations are not associated with a superior outcome for patients treated on the BFM-like DCOG protocols or the COALL-97 protocol. The survival rate of NOTCH1-activated patients was actually less than for wild-type patients. Separating patients carrying strong NOTCH1-activating mutations from those with

weak NOTCH1-activating mutations[35] or patients that were wildtype showed a significant poor outcome for patients having strong NOTCH1-activating mutations in the DCOG cohort. This could not be reproduced for T-ALL patients treated on the German COALL-97 protocol. In the accompanying article of Clappier *et al.*,[42] NOTCH1-activating mutations did not predict improved outcome for patients treated on the BFM-derived EORTC-CLG protocols either. These observations are in contrast to the findings by the BFM study group.[17] In the accompanying article of Kox *et al.*,[44] this finding is now validated in an extended series comprising 301 pediatric T-ALL patients treated on the ALL-BFM 2000 protocol. A favorable prognostic effect of *NOTCH1* and/or *FBXW7* mutations was also identified in a recent study by Park *et al.*,[15] although the overall incidence of identified *NOTCH1* mutations was only 31%. No favorable outcome of *NOTCH1* and/or *FBXW7*-mutated cases has been observed neither for adult T-ALL patients treated on GMALL 05/93 and 06/99 multicenter protocols,[45] nor for patients treated on the MRC UKALLXII/ECOG E2993[37] or LALA-94[36] protocols. A significant association with improved outcome for NOTCH1-activating mutations has only been observed for adult T-ALL patients treated on the GRAALL-2003 multicenter protocol.[36] These results indicate that the prognostic effect of *NOTCH1/FBXW7* mutations may strongly depend on the treatment protocol given.

Compared with the ALL-BFM-2000 protocol, the DCOG ALL-7/8 protocol in general showed an inferior outcome.[22] Although both protocols are highly related, part of the patients treated on the DCOG ALL-7/8 cohort received less chemotherapy and none of them received prophylactic cranial irradiation, except for patients with initial central nervous system involvement. NOTCH1-activating mutations may provoke central nervous system relapse because of the activation of the CCR7 chemokine.[46] This study therefore predicts that NOTCH1-activating mutations would result in increased risk for central nervous system relapse through the CCL19-CCR7 axis in the absence of cranial irradiation. However, the numbers of central nervous system relapses in our cohorts were too low to substantiate this notion. In addition, neither the *CCR7* gene nor its ligand *CCL19* was identified as significantly differentially expressed genes that were activated in *NOTCH1/FBXW7*-mutated T-ALL patients based on our microarray expression data set (data not shown). As our patient biopsies were all obtained from peripheral blood or bone marrow samples, we cannot exclude that these genes are only upregulated in malignant blasts in the context of a neuronal environment. Cranial radiation may contribute to the differences in prognostic value for

NOTCH1-activating mutations between the DCOG and ALL-BFM-2000 cohorts, but this does not apply for the COALL-97 cohort that includes cranial irradiation. Therefore, other differences among treatment protocols seem important.

In conclusion, *NOTCH1/FBXW7* mutations that activate the NOTCH1 pathway are identified in >60% pediatric T-ALL patients and result in elevated ICN levels and activation of NOTCH1 target genes. Mutations were more often found in association with *TLX3*-rearranged T-ALL, but were less frequently identified in *TAL/LMO* T-ALL patients and T-ALL patients with a mature T-cell phenotype. *NOTCH1/FBXW7* mutations predict for an initial PGR, which does not translate into a superior outcome of T-ALL on DCOG ALL-7/8, ALL-9 or COALL-97 protocols.

**Author contributions**

LZ designed experiments, performed research and wrote the article; IH performed research and wrote the article; VC performed RPMA analysis; MLW performed research; JB-G performed *NOTCH1* and *FBXW7* mutation analysis; CK performed western blot analysis; WS prepared samples for RPMA analysis; ES, AJPV, WK and MH provided patient samples and clinical and immunophenotypic data; EP supervised study and wrote the article; RP designed and supervised study and wrote the article; JPPM was principal investigator, designed and supervised the study and wrote the article.

**Conflict of interest**

The authors declare no conflict of interest.

**Acknowledgements**

LZ and WKS were financed by the Stichting Kinderen Kankervrij (KiKa; Grant no. KiKa 2007-012). IH was financed by the Dutch Cancer Society Dutch Cancer Society (KWF-EMCR 2006-3500) CK was financed by KiKa (Grant no. KiKa 2008-029). We thank the German Jose Carreras Leukemia Foundation (Grant no. SP 04/03 to MH).

Supplementary Information available in the Supplementary Chapter of this thesis.

**REFERENCES**


1. Pieters R, Carroll WL. Biology and treatment of acute lymphoblastic leukemia. Pediatr Clin North Am 2008; 55: 1–20, ix.
2. Pui CH, Evans WE. Treatment of acute lymphoblastic leukemia. N Engl J Med 2006; 354: 166–178
3. Cave H, Suciu S, Preudhomme C, Poppe B, Robert A, Uyttebroeck A *et al*. Clinical significance of HOX11L2 expression linked to t(5;14)(q35;q32), of HOX11 expression, and of SIL-TAL fusion in childhood T-cell malignancies: results of EORTC studies 58881 and 58951. Blood 2004; 103: 442–450.
4. Meijerink JP, den Boer ML, Pieters R. New genetic abnormalities and treatment response in acute lymphoblastic leukemia. Semin Hematol 2009; 46: 16–23.
5. Van Vlierberghe P, Pieters R, Beverloo HB, Meijerink JP. Molecular-genetic insights in paediatric T-cell acute lymphoblastic leukaemia. Br J Haematol 2008; 143: 153–168.
6. Ellisen LW, Bird J, West DC, Soreng AL, Reynolds TC, Smith SD *et al*. TAN-1, the human homolog of the Drosophila notch gene, is broken by chromosomal translocations in T lymphoblastic neoplasms. Cell 1991; 66: 649–661.
7. Suzuki S, Nagel S, Schneider B, Chen S, Kaufmann M, Uozumi K *et al*. A second NOTCH1 chromosome rearrangement: t(9;14)(q34.3;q11.2) in T-cell neoplasia. Leukemia 2009; 23: 1003–1006.
8. Weng AP, Ferrando AA, Lee W, Morris JPt, Silverman LB, Sanchez-Irizarry C *et al*. Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. Science 2004; 306: 269–271.
9. Sulis ML, Williams O, Palomero T, Tosello V, Pallikuppam S, Real PJ *et al*. NOTCH1 extracellular juxtamembrane expansion mutations in T-ALL. Blood 2008; 112: 733–740.
10. Grabher C, von Boehmer H, Look AT. Notch 1 activation in the molecular pathogenesis of T-cell acute lymphoblastic leukaemia. Nat Rev Cancer 2006; 6: 347–359.
11. Palomero T, Lim WK, Odom DT, Sulis ML, Real PJ, Margolin A *et al*. NOTCH1 directly regulates c-MYC and activates a feed-forward-loop transcriptional network promoting leukemic cell growth. Proc Natl Acad Sci USA 2006; 103: 18261–18266.
12. Weng AP, Millholland JM, Yashiro-Ohtani Y, Arcangeli ML, Lau A, Wai C *et al*. c-Myc is an important direct target of Notch1 in T-cell acute lymphoblastic leukemia/lymphoma. Genes Dev 2006; 20: 2096–2109.
13. Malyukova A, Dohda T, von der Lehr N, Akhoondi S, Corcoran M, Heyman M *et al*. The tumor suppressor gene hCDC4 is frequently mutated in human T-cell acute lymphoblastic leukemia with functional consequences for Notch signaling. Cancer Res 2007; 67: 5611–5616.
14. O'Neil J, Grim J, Strack P, Rao S, Tibbitts D, Winter C *et al*. FBW7 mutations in leukemic cells mediate NOTCH pathway activation and resistance to gamma-secretase inhibitors. J Exp Med 2007; 204: 1813–1824.
15. Park MJ, Taki T, Oda M, Watanabe T, Yumura-Yagi K, Kobayashi R *et al*. FBXW7 and NOTCH1 mutations in childhood T cell acute lymphoblastic

leukaemia and T cell non-Hodgkin lymphoma. Br J Haematol 2009; 145: 198–206.
16. Thompson BJ, Buonamici S, Sulis ML, Palomero T, Vilimas T, Basso G *et al*. The SCFFBW7 ubiquitin ligase complex as a tumor suppressor in T cell leukemia. J Exp Med 2007; 204: 1825–1835.
17. Breit S, Stanulla M, Flohr T, Schrappe M, Ludwig WD, Tolle G *et al*. Activating NOTCH1 mutations predict favorable early treatment response and long-term outcome in childhood precursor T-cell lymphoblastic leukemia. Blood 2006; 108: 1151–1157.
18. Zhu YM, Zhao WL, Fu JF, Shi JY, Pan Q, Hu J *et al*. NOTCH1 mutations in T-cell acute lymphoblastic leukemia: prognostic significance and implication in multifactorial leukemogenesis. Clin Cancer Res 2006; 12: 3043–3049.
19. van Grotel M, Meijerink JP, van Wering ER, Langerak AW, Beverloo HB, Buijs-Gladdines JG *et al*. Prognostic significance of molecular-cytogenetic abnormalities in pediatric T-ALL is not explained by immunophenotypic differences. Leukemia 2008; 22: 124–131.
20. Larson Gedman A, Chen Q, Kugel Desmoulin S, Ge Y, LaFiura K, Haska CL *et al*. The impact of NOTCH1, FBW7 and PTEN mutations on prognosis and downstream signaling in pediatric T-cell acute lymphoblastic leukemia: a report from the Children′s Oncology Group. Leukemia 2009; 23: 1417–1425.
21. Kamps WA, Bokkerink JP, Hahlen K, Hermans J, Riehm H, Gadner H *et al*. Intensive treatment of children with acute lymphoblastic leukemia according to ALL-BFM-86 without cranial radiotherapy: results of Dutch Childhood Leukemia Study Group Protocol ALL-7 (1988–1991). Blood 1999; 94: 1226–1236.
22. Kamps WA, Bokkerink JP, Hakvoort-Cammel FG, Veerman AJ, Weening RS, van Wering ER *et al*. BFM-oriented treatment for children with acute lymphoblastic leukemia without cranial irradiation and treatment reduction for standard risk patients: results of DCLSG protocol ALL-8 (1991–1996). Leukemia 2002; 16: 1099–1111.
23. Veerman AJ, Kamps WA, van den Berg H, van den Berg E, Bokkerink JP, Bruin MC *et al*. Dexamethasone-based therapy for childhood acute lymphoblastic leukaemia: results of the prospective Dutch Childhood Oncology Group (DCOG) protocol ALL-9 (1997–2004). Lancet Oncol 2009; 10: 957–966.
24. van Grotel M, Meijerink JP, Beverloo HB, Langerak AW, Buys-Gladdines JG, Schneider P *et al*. The outcome of molecular-cytogenetic subgroups in pediatric T-cell acute lymphoblastic leukemia: a retrospective study of patients treated according to DCOG or COALL protocols. Haematologica 2006; 91: 1212–1221.
25. Stam RW, den Boer ML, Meijerink JP, Ebus ME, Peters GJ, Noordhuis P *et al*. Differential mRNA expression of Ara-C-metabolizing enzymes explains Ara-C sensitivity in MLL gene-rearranged infant acute lymphoblastic leukemia. Blood 2003; 101: 1270–1276.
26. Bene MC, Castoldi G, Knapp W, Ludwig WD, Matutes E, Orfao A *et al*. Proposals for the immunological classification of acute leukemias. European Group for the Immunological Characterization of Leukemias (EGIL). Leukemia 1995; 9: 1783–1786.

27. Van Vlierberghe P, van Grotel M, Beverloo HB, Lee C, Helgason T, Buijs-Gladdines J *et al*. The cryptic chromosomal deletion del(11)(p12p13) as a new activation mechanism of LMO2 in pediatric T-cell acute lymphoblastic leukemia. Blood 2006; 108: 3520–3529.
28. Van Vlierberghe P, van Grotel M, Tchinda J, Lee C, Beverloo HB, van der Spek PJ *et al*. The recurrent SET-NUP214 fusion as a new HOXA activation mechanism in pediatric T-cell acute lymphoblastic leukemia. Blood 2008; 111: 4668–4680.
29. Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. Stat Med 1990; 9: 811–818.
30. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. Proc Natl Acad Sci USA 2001; 98: 31–36.
31. Paweletz CP, Charboneau L, Bichsel VE, Simone NL, Chen T, Gillespie JW *et al*. Reverse-phase protein microarrays which capture disease progression show activation of pro-survival pathways at the cancer invasion front. Oncogene 2001; 20: 1981–1989.
32. Petricoin 3rd EF, Espina V, Araujo RP, Midura B, Yeung C, Wan X *et al*. Phosphoprotein pathway mapping: Akt/mammalian target of rapamycin activation is negatively associated with childhood rhabdomyosarcoma survival. Cancer Res 2007; 67: 3431–3440.
33. Lee JW, Soung YH, Kim HJ, Park WS, Nam SW, Kim SH *et al*. Mutational analysis of the hCDC4 gene in gastric carcinomas. Eur J Cancer 2006; 42: 2369–2373.
34. Homminga I, langerak AW, De Laat W, de Rooij JJ, Stubbs A, Buijs-Gladdines J *et al*. NKX2-1 and MEF2C oncogenes delineate novel subtypes in T-cell acute lymphoblastic leukemia. (in preparation).
35. Chiang MY, Xu L, Shestova O, Histen G, L′Heureux S, Romany C *et al*. Leukemia-associated NOTCH1 alleles are weak tumor initiators but accelerate K-ras-initiated leukemia. J Clin Invest 2008; 118: 3181–3194.
36. Asnafi V, Buzyn A, Le Noir S, Baleydier F, Simon A, Beldjord K *et al*. NOTCH1/FBXW7 mutation identifies a large subgroup with favorable outcome in adult T-cell acute lymphoblastic leukemia (T-ALL): a Group for Research on Adult Acute Lymphoblastic Leukemia (GRAALL) study. Blood 2009; 113: 3918–3924.
37. Mansour MR, Sulis ML, Duke V, Foroni L, Jenkinson S, Koo K *et al*. Prognostic implications of NOTCH1 and FBXW7 mutations in adults with T-cell acute lymphoblastic leukemia treated on the MRC UKALLXII/ECOG E2993 protocol. J Clin Oncol 2009; 27: 4352–4356.
38. Gordon WR, Roy M, Vardar-Ulu D, Garfinkel M, Mansour MR, Aster JC *et al*. Structure of the Notch1-negative regulatory region: implications for normal activation and pathogenic signaling in T-ALL. Blood 2009; 113: 4381–4390.
39. Asnafi V, Radford-Weiss I, Dastugue N, Bayle C, Leboeuf D, Charrin C *et al*. CALM-AF10 is a common fusion transcript in T-ALL and is specific to the TCRgammadelta lineage. Blood 2003; 102: 1000–1006.

40. Eguchi-Ishimae M, Eguchi M, Kempski H, Greaves M. NOTCH1 mutation can be an early, prenatal genetic event in T-ALL. Blood 2008; 111: 376–378.
41. Mansour MR, Duke V, Foroni L, Patel B, Allen CG, Ancliff PJ *et al*. Notch-1 mutations are secondary events in some patients with T-cell acute lymphoblastic leukemia. Clin Cancer Res 2007; 13: 6964–6969.
42. Clappier E, Collette S, Grardel N, Girard S, Suarez L, Brunie G *et al*. Prognostic significance of NOTCH1 and FBXW7 mutations in childhood T-cell acute leukemia: results from the EORTC Children Leukemia Group. Leukemia 2010, (this issue).
43. Real PJ, Tosello V, Palomero T, Castillo M, Hernando E, de Stanchina E *et al*. Gamma-secretase inhibitors reverse glucocorticoid resistance in T cell acute lymphoblastic leukemia. Nat Med 2009; 15: 50–58.
44. Kox C, Zimmermann M, Stanulla M, Leible S, Schrappe M, Ludwig WD *et al*. The favorable effect of activating *NOTCH1* receptor mutations on long-term outcome in T-ALL can be seperated from NOTCH pathway activation by *FBXW7* loss of function. Leukemia 2010, (this issue).
45. Baldus CD, Thibaut J, Goekbuget N, Stroux A, Schlee C, Mossner M *et al*. Prognostic implications of NOTCH1 and FBXW7 mutations in adult acute T-lymphoblastic leukemia. Haematologica 2009; 94: 1383–1390.
46. Buonamici S, Trimarchi T, Ruocco MG, Reavie L, Cathelin S, Mar BG *et al*. CCR7 signalling as an essential regulator of CNS infiltration in T-cell leukaemia. Nature 2009; 459: 1000–1004.

# CHAPTER 6

## In vitro efficacy of forodesine and nelarabine (ara-G) in pediatric leukemia

Irene Homminga[1], C. Michel Zwaan[1], Chantal Y. Manz[2], Cynthia Parker[3], Shanta Bantia[3]; Willem Korstiaan Smits[1], Fiona Higginbotham[2]; Rob Pieters[1] and Jules P.P. Meijerink[1]

[1]Department of Pediatric Hemato-Oncology, Erasmus MC Rotterdam – Sophia Children's Hospital, Rotterdam, the Netherlands;
[2]Mundipharma International Limited, Cambridge, England;
[3]Department of Biological Sciences, Biocryst Pharmaceutical, Inc., Birmingham, USA.

*Blood, 2011: 118; 2184-2190*

**ABSTRACT**

Forodesine and nelarabine (the pro-drug of ara-G) are two nucleoside analogues with promising anti-leukemic activity. To better understand which pediatric patients might benefit from forodesine or nelarabine (ara-G) therapy, we investigated the *in-vitro* sensitivity to these drugs in 96 diagnostic pediatric leukemia patient samples and the mRNA expression levels of different enzymes involved in nucleoside metabolism. Forodesine and ara-G cytotoxicities were higher in T-cell acute lymphoblastic leukemia (T-ALL) samples than in B-cell precursor (BCP-) ALL and acute myeloid leukemia (AML) samples. Resistance to forodesine did not preclude ara-G sensitivity and vice versa, indicating that both drugs rely on different resistance mechanisms. Differences in sensitivity could be partly explained by significantly higher accumulation of intracellular dGTP in forodesine sensitive samples compared with resistant samples, and higher mRNA levels of *dGK* but not *dCK*. The mRNA levels of the transporters *ENT1* and *ENT2* were higher in ara-G sensitive than resistant samples. We conclude that especially T-ALL, but also BCP-ALL pediatric patients may benefit from forodesine or nelarabine (ara-G) treatment.

## INTRODUCTION

Leukemia is the most common childhood malignancy, and the general incidence in both adults and children to develop acute lymphoblastic leukemia (ALL) or acute myeloid leukemia (AML) is approximately 1 per 100,000 and 2-3 per 100,000, respectively. Although overall cure rates have been improved over the last decades, still about 20% of children with acute lymphoblastic leukemia (ALL) and 40% of children with acute myeloid leukemia (AML) eventually die from their disease[1,2]. In adults, the prognosis is worse with a survival below 60% in ALL[3] and 50% in AML[4], indicating that there is still a great need for better therapy. Currently, purine nucleosides analogues are in clinical trials for different types of leukemia including clofarabine, forodesine (BCX-1777/Immucillin H) and nelarabine (506U78/Arranon/Atriance) the latter being the prodrug for 9-β-D-arabinofuranosylguanine (ara-G).

Forodesine is a non-cleavable inosine analogue developed to bind and inhibit the purine nucleoside phosphorylase (PNP) enzyme[5]. PNP normally degrades excess of intracellular deoxyguanosine (dGuo) into guanosine and deoxyribose-1-phosphate through phosphorylysis. dGuo is continuously produced in the body as the result of DNA degradation during cellular turnover. Inhibition of PNP by forodesine results in the intracellular accumulation of dGuo. DGuo is rapidly phosphorylated to deoxyguanosine triphosphate (dGTP) in the purine salvage pathway leading to dGTP accumulation[6,7]. High intracellular levels of dGTP cause cell death through mechanisms that are still not fully understood, but which may likely involve imbalance in the deoxynucleotide pool and/or inhibition of ribonucleotide reductase[8] resulting inhibition of DNA synthesis and/or by activation of a p53-induced cell cycle arrest and apoptosis[9]. Whereas most nucleoside analogues depend on DNA incorporation to exert their toxic effect, this is not the case for forodesine. T-cells seem to be especially sensitive to PNP inhibition as severe combined immunodeficient (SCID) patients with PNP deficiency have increased plasma levels of dGuo[10,11] and a severe depletion of T-cells compared to other cell types[12,13]. In contrast to SCID however, severe opportunistic infections are not seen in treatment with forodesine, as there seems to be a selective toxicity towards leukemic cells[7].

Ara-G is an arabinosylguanine analogue that is resistant to PNP mediated phosphorylysis. Accumulated intracellular ara-G is rapidly converted to ara-GTP which results in cell death through inhibition of ribonucleotide reductase and incorporation of ara- GTP in the DNA which blocks further DNA synthesis[14,15]. In

contrast to various other arabinonucleoside compounds including ara-C, selective T-cell toxicity has only been demonstrated for ara-G[14-17]. However, the use of ara-G is limited due to its poor water solubility. Therefore nelarabine, a pro-drug of ara-G that is eight fold more water soluble[18], is used in clinical settings. *In-vivo* nelarabine is rapidly converted into ara-G through demethoxylation by adenosine deaminase.

Forodesine has been tested in clinical phase I/II trials in relapsed or refractory patients with T-cell ALL or lymphoblastic lymphoma[7,19], BCP-ALL and chronic lymphocytic leukemia[20] (reviewed in [21]). Forodesine treatment resulted in an overall response in 32% of the T-cell leukemia patients, with 21% of the having a complete response. Forodesine administration resulted in an increase in plasma dGuo and intracellular dGTP levels. Adverse affects were mild, with only grade 3 thrombocytopenia and leukopenia[19,21]. For BCP-ALL patients forodesine treatment resulted in complete responses in 17% of the patients[21].

Nelarabine, the pro-drug of ara-G, has been tested in clinical phase I/II trials in adults[22] and children[23,24] with refractory or relapsed T-ALL or T-cell lymphoblastic lymphoma (T-LBL) and is an approved drug for T-cell disease in both the US and Europe. Thirty-one percent of adult T-ALL and T-LBL patients achieved a complete remission with an overall response rate in 41 % of the patients. Median disease-free survival (DFS) and overall survival (OS) were 20 weeks, with 28 percent of the patients surviving 1 year. Principal toxicity was a grade 3 or 4 neutropenia and thrombocytopenia[22]. For pediatric T-ALL patients at first relapse, complete responses were documented for 55% of the patients. For patients in second relapse or for patients with extramedullary relapses, response rates ranged from 14-33 %. However, 18 % of the patients had a ≥grade 3 neurologic adverse event[23].

To better predict which patients might benefit from forodesine or nelarabine treatment, we investigated the *in-vitro* sensitivity to forodesine or ara-G in pediatric ALL and AML diagnostic patient samples. Forodesine toxicity was investigated in relation to intracellular accumulation of dGTP levels. We also investigated potential mechanisms that may be responsible for differences in drug sensitivity among patient samples. To this end, we measured mRNA expression levels of proteins that are involved in purine metabolism and uptake (**Figure 1**). In addition, we tested whether forodesine had a synergistic or antagonistic effect with 7 commonly used drugs in leukemia treatment.

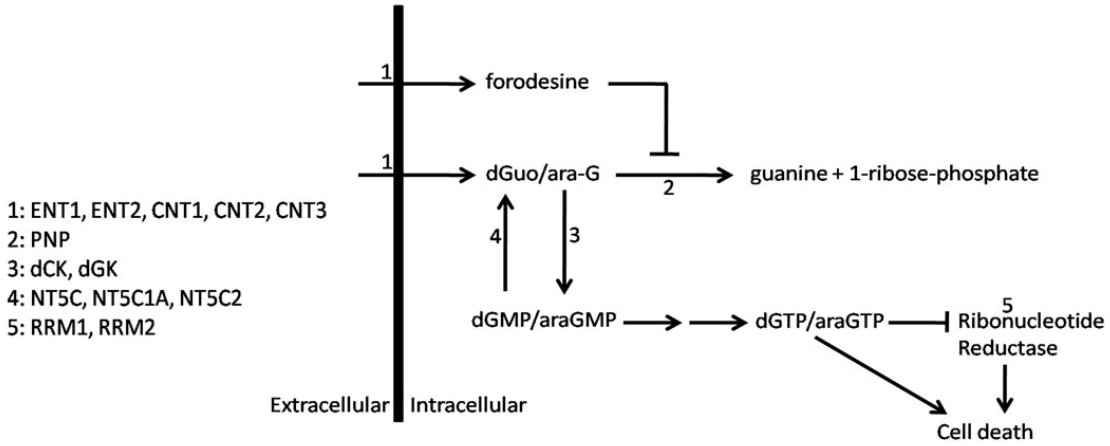



**Figure 1. Purine metabolism overview.** Schematic overview of main enzymes and transporters involved in purine conversion and uptake. ENT1-2*:* equilibrative nucleoside transporter 1-2, CNT1-3: concentrative nucleoside transporter 1-3, PNP: purine nucleoside phosphorylase dCK: deoxycytidine kinase, dGK: deoxyguanine kinase, NT5C: cytosolic 5' nucleotidase 1A, NT5C1A: cytosolic 5' nucleotidase 1A, NT5C2: cytosolic 5' nucleotidase RRM1 and RRM2: ribonucleotide reductase subunit 1 and subunit 2.

## MATERIALS & METHODS

### Patient Material

Fresh or viably frozen bone marrow or peripheral blood samples from a total of 96 de novo, untreated pediatric acute leukemia patients were used, comprising 36 T-ALL, 43 BCP-ALL and 17 AML samples. All samples were tested for forodesine cytotoxicity, whereas additional assays were performed on the same samples based on the availability of material. The patients' parents or legal guardians provided informed consent to use leftover diagnostic patient biopsies for research in accordance with the Institutional Review Board of the Erasmus MC Rotterdam and in accordance with the Declaration of Helsinki. Leukemic cells were isolated and enriched as previously described[25]. All resulting samples contained ≥90% leukemic cells, as determined morphologically by May-Grünwald-Giemsa-stained cytospins (Merck, Darmstadt, Germany) and were viably frozen in liquid nitrogen as described earlier[25].

**Cell lines**

T-ALL cell lines (CCRF-CEM, LOUCY, BE-13, MOLT-4, PEER, KARPAS-45, MOLT-3, JURKAT, HPB-ALL, PF-382) were purchased from the German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany), and cultured under recommended conditions.

**Assessment of PNP inhibition by forodesine (dGuo measurements)**

The efficacy of forodesine to inhibit phosphorylysis of dGuo into guanosine and deoxyribose-1-phosphate by PNP was assessed in 4 pediatric T-ALL and 2 pediatric BCP-ALL patient samples. For this, the decrease in dGuo concentration was measured over time in the supernatant of cell cultures that were treated with varying concentrations of forodesine. Cells were cultured in RPMI 1640 Dutch modification without L-glutamine, 20% fetal calf serum, 2 mM L-glutamine (Invitrogen), 5 μg/ml insulin, 5 μg/ml transferrin, 5 ng/ml sodium selenite (ITS media supplement; Sigma, St Louis MO, USA), 100 IU/ml penicillin, 100 μg/ml streptomycin, 0.125 μg/ml fungizone and 0.2 mg/ml gentamycin (Invitrogen) at a concentration of $1.6 \cdot 10^6$ cells/ml. Forodesine (provided by Mundipharma Research Ltd) was added to final concentrations of 1, 3 or 10 μM; or replaced by $dH_2O$ in the control. dGuo (Sigma) was added to all cultures to a final concentration of 10 μM. Cells were plated in triplicate in 96 well plates (Bioplastics, Landgraaf, the Netherlands) for each condition ($320 \cdot 10^3$ cells/well). After 0, 4, 24, 48 and 96 hours, cells were pelleted by centrifugation and the supernatant was collected for dGuo measurement and stored at -80°C until further analysis. dGuo levels were analyzed by high-performance liquid chromatography (HPLC or LC) with tandem mass spectrometry detection (MS/MS) as previously described.[19] Briefly, dGuo was extracted from the supernatant using a Waters Oasis "HLB" affinity solid phase extraction (SPE) cartridge. The mass of dGuo plus H+ (268.1 m/z) was monitored in quadrupole one (Q1). The dGuo product ion 157.0 m/z was monitored in quadrupole three (Q3). The concentrations of dGuo were determined by weighted (1/x) quadratic regression analysis of peak areas produced from the standard curve.

***In-vitro* forodesine, ara-G and ara-C cytotoxicity (MTT assay)**

Forodesine (36 T-ALL, 43 BCP-ALL and 17 AML samples), ara-G (28 T-ALL, 35 BCP-ALL and 17 AML samples) and ara-C (28 T-ALL samples) cytotoxicities were determined using the MTT assay as described previously.[26] Ara-G is the

active metabolite of the pro-drug nelarabine. We measured cell viability in the presence of 1 μM forodesine and 6 concentrations (0.01, 0.1, 1, 3, 10 and 50 μM) of dGuo, following an incubation period of four days. As control, samples were incubated with the same range of dGuo concentrations in the absence of forodesine. Additional controls were 1 μM forodesine in the absence of dGuo, and vehicle only. dGuo is added to the culture to mimic the natural variable presence of dGuo in the blood, as this compound mediates forodesine cytotoxicity. For ara-G (Carbosynth Limited, Berkshire, UK) the following concentrations were used: ara-G 0.01, 0.1, 1, 3, 10, 50 μM. The concentrations used in the MTT assay for ara-C were: 0.01, 0.04, 0.16, 0.625, 2.5 and 10 μM.

**Combination cytotoxicity assay**

Using the MTT assay as previously described[26], we screened for potential antagonistic or synergistic effects in forodesine mediated cytotoxicity for 7 compounds that are used in ALL treatment, comprising ara-C, ara-G, 6MP (Sigma Aldrich, St. Louis, USA), asparaginase (Medac, Augusta, USA), daunorubicin (cerubidine®, Sanofi-aventis, Bridgewater, USA), prednisolone (BUFA BV, Uitgeest, the Netherlands), and vincristine (TEVA pharmachemie, Haarlem, the Netherlands). Four to 9 T-ALL and 6 to 8 BCP-ALL pediatric patient samples were tested for each drug combination. Prior to this, the median concentration that is lethal to 10% (LC10) and to 30% (LC30) of cells were determined for dGuo in the presence of 1 μM forodesine on the basis of *in-vitro* forodesine cytotoxicity assay results (see above) for 10 T-ALL and 10 BCP-ALL patient samples. The T-ALL and BCP-ALL median LC10 or LC30 concentrations were used in the combination assay for T-ALL and BCP-ALL samples, respectively. Forodesine (1μM) and the median LC10 or LC30 concentrations of dGuo were then combined with a range of each of the 7 drugs (ara-C: 0.01, 0.04, 0.16, 0.625, 2.50, 10.0 μM, ara-G: 0.01, 0.10, 1.0, 3.0, 10, 50 μM, 6-mercaptopurine (6MP): 0.016, 0.031, 0.063, 0.125, 0.50, 1.0 mg/ml, asparaginase: 0.003, 0.016, 0.08, 0.40, 2.0, 10.0 IE/ml, daunorubicin: 0.002, 0.008, 0.031, 0.125, 0.5, 2.0 μg/ml, prednisolone: 0.008, 0.06, 0.49, 3.9, 31.3, 250 μg/ml and vincristine: 0.05, 0.20, 0.78, 3.1, 12.5, 50.0 μg/ml). The controls were: 1 μM forodesine in combination with the median LC10 or LC30 value of dGuo. Previous experiments on T-ALL cell lines (JURKAT, HPB-ALL, LOUCY and PF-382) showed no effect of addition of the median LC30 values of dGuo on the cytotoxicity of the 7 drugs in the absence of forodesine (data not shown). Since 6MP solutions give a background signal in the MTT assay,

varying concentrations of 6MP in culture medium were included as an additional control. For each patient and each concentration of compound tested, a hypothetical maximal additive effect of either LC10 or LC30 forodesine/dGuo treatment in combination with the other compound was calculated by the following formula: ((100-A) X B/100) +A, where A and B are the percentages of cell death caused by each compound individually. We performed a t-test to analyze for each drug concentration whether the median calculated hypothetical values were significantly different from the actual measured median values obtained by combining the drugs, i.e whether the results differed significantly from the hypothetical maximum additive effect. When a significant difference was observed, we performed another t-test to analyze whether the median cell survival measured with drug only increased significantly by addition of forodesine/dGuo, i.e. whether an antagonistic effect was present.

**dGTP measurement**

Accumulation of dGTP was calculated using a polymerase assay as previously described[27] in 22 T-ALL, 6 BCP-ALL and 2 AML samples. Ten million cells were cultured for 24 hours in 5 ml culture medium (see above) in the presence of 3 μM dGuo and 1 μM forodesine. The control reaction comprised 3 μM dGuo. Proliferation and apoptosis were measured with Trypan-blue staining and counting in a Bürker-Türk counting chamber. Cells were washed twice with PBS and spinned down by centrifugation. The cell pellet was resuspended in 1 ml 60% methanol (-20°C) and stored at -20°C. The samples were centrifuged and supernatants were dried in a TurboVap. Dried extracts were stored at -20°C until further analysis. Extracts were suspended in 25 μl buffer (20 mM Hepes-NaOH, pH 7.3; 2 mM $MgCl_2$) and 20 μl was used in the assay. dGTP standards were used at 0, 0.5, 1, 5, 10, and 50 pmol. Reactions contained 20 μl of extract or standard, 100 mM Hepes-NaOH (pH 7.3), 10 mM $MgCl_2$, 50 nM primer, 2.5 μM [3H]-dATP, 0.5 U Klenow Exo-Free DNA Polymerase I, and $dH_2O$ to 100 μl final volume. Reactions were incubated in U-bottom 96-well tissue culture plates at RT for 1 hour. Samples were harvested onto Whatman DE81 DEAE cellulose paper using a Packard cell harvester, washed three times with 5% $Na_2PO_4$, once with $dH_2O$, once with 95% ethanol and then air-dried and counted on a Packard Matrix-9600 beta counter. A standard curve was generated (cpm vs. dGTP concentration) for each experiment and the amounts of dGTP present in the extracts were calculated using the standard curve.

**Real time Quantitative Polymerase Chain Reaction (RQ-PCR)**

cDNA was available for 25 T-ALL samples, 24 BCP-ALL samples, and 1 AML patient sample. RNA extraction and cDNA synthesis were performed as previously described[25]. RQ-PCR reactions were performed in 1x DyNAmo™ HS SYBR® Green mastermix (Finnzymes, Espoo, Finland), 1x ROX™ (Finnzymes), 8.3 pmol forward primer, 8.3 pmol reverse primer, 20ng cDNA and 4mM $MgCl_2$ in a final volume of 27.5 μl. RQ-PCR was performed on a 9700HT Fast Real-Time PCR system (Applied Biosystems, Foster City, CA, USA) starting with DNA polymerase heat activation at 95°C for 10 minutes, followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. A melting curve was recorded during a heating step from 25°C to 95°C during a 10 minute period. We performed cycle threshold analysis for each reaction using SDS2.3 analysis software (Applied Biosystems) and expression levels were quantified relative to the endogenous housekeeping gene glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) using the ΔCt-method.[28] All reactions were performed in duplicate. Primer sequences for deoxycytidine kinase (*dCK)*, cytosolic 5' nucleotidase 1A *(PNI/NT5C/P5N2)*, equilibrative nucleoside transporter 1 *(ENT1/SLC29A1)*, ribonucleotide reductase subunit 1 *(RRM1)* and subunit 2 *(RRM2)* and *GAPDH* have been described elsewhere[25,29]. Other primer combinations are listed in **supplementary Table 1**. cDNA of a T-ALL cell line pool (CCRF-CEM, LOUCY, BE-13, MOLT-4, PEER, KARPAS-45, MOLT-3 and JURKAT) was used as positive control for these targets.

**Statistical analysis**

Differences in the distribution of continuous variables were analyzed using the Mann-Whitney U test. Analyses of proportional differences were performed by Chi-square test or Fisher exact test. Student's t-test was used to analyze whether differences in cell survival differed significantly from zero. Statistical tests were performed at a two-tailed significance level of 0.05.

## RESULTS

### *In-vitro* forodesine and ara-G cytotoxicity levels

To explore the efficacy of purine nucleosides analogues as a potential therapeutic drug for ALL, we tested *in-vitro* toxicity levels of forodesine and ara-G on pediatric ALL and AML samples. Forodesine toxicity depends on the plasma availability of dGuo and its conversion into dGTP, and we first tested the ability of forodesine to block the degradation of dGuo into guanosine and deoxyribose-1-phosphate by PNP. These measurements were performed in the presence of dGuo and increasing forodesine concentrations. Without forodesine, dGuo levels in the culture media are rapidly being depleted as consequence of PNP-mediated degradation to nearly undetectable levels within 24 hrs in 5 out of 6 patient samples. For all samples tested, 1 μM of forodesine was sufficient to block PNP activity (**supplementary figure 1**) resulting in the complete stabilization of dGuo levels in the culture supernatants. This dose of forodesine was then chosen in subsequent cellular cytotoxicity experiments.

We then measured the cellular toxicity to 1 μM forodesine in 96 pediatric primary leukemia samples in the presence of varying concentrations of dGuo (**Figures 2A and 2B**). In our assay, dGuo itself elicited no cellular toxicity up to concentrations of 10 μM as it is rapidly being degraded by PNP (data not shown). One μM of forodesine in the absence of dGuo had no effect on survival (data not shown). However, in the presence of forodesine and subsequent blockage of PNP activity, T-ALL samples were more sensitive to dGuo levels (median LC50 = 1.6μM dGuo) than BCP-ALL (median LC50 = 8.8μM dGuo, $p=0.001$) and AML (median LC50 >10μM, $p<0.001$) samples (**Figure 2A**). Only one out of 17 AML samples reached an LC50 in our assay.

Ara-G cytotoxicity was measured in 28 T-ALL, 35 BCP-ALL and 17 AML pediatric patient samples. Again, T-ALL samples were most sensitive to treatment (median LC50 = 20.5μM) compared to BCP-ALL (median LC50 >50μM, $p<0.001$) or AML (median LC50 = 45.8μM, $p=0.012$) samples.(**Figure 2B**).

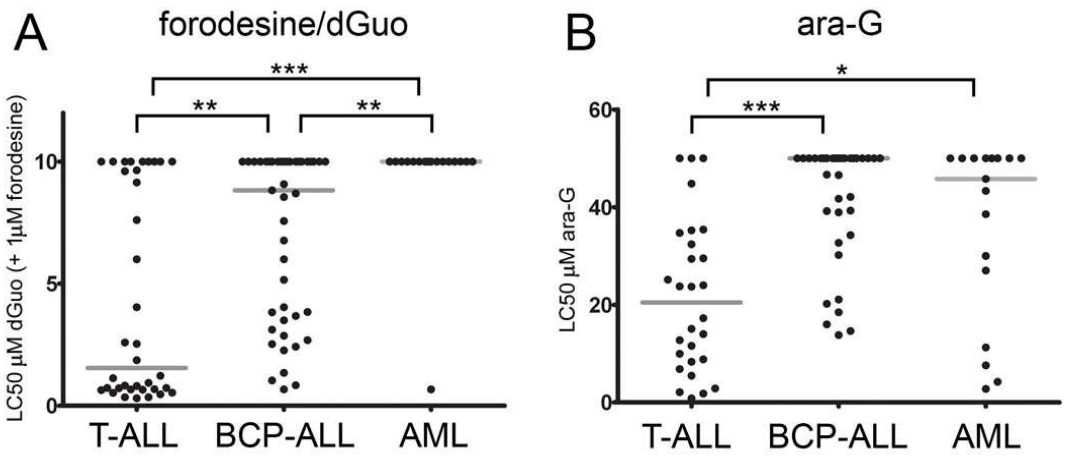



**Figure 2. Forodesine/dGuo and ara-G sensitivity in pediatric leukemia**. (A) LC50 values for forodesine/dGuo for T-ALL, BCP-ALL and AML leukemia samples. When no LC50 was reached, a value of 10µM was assigned. (B) LC50 values for ara-G for T-ALL, BCP-ALL and AML leukemias. When no LC50 was reached a value of 50µM was assigned. Median LC50 values are indicated by grey horizontal lines. Significance levels are indicated by asterisks: *,$p<0.05$; **,$p<0.01$; ***,$p<0.001$.

As conversion of dGuo and ara-G rely on the same enzymatic pathways, we investigated potential cross-resistance towards dGuo/forodesine and ara-G in T-ALL patient samples. Patient that require drug concentrations higher than 10µM of dGuo (at 1µM of forodesine) or 50µM of ara-G as LC50 values in our assay were regarded as resistant. We did not find any correlation between dGuo/forodesine and ara-G cytotoxicities, nor between dGuo/forodesine and the pyrimidine equivalent of the ara-G drug, i.e. ara-C (**Figure 3A and 3B**). For T-ALL patients, 2 out of 3 samples that were resistant to ara-G were sensitive to forodesine/dGuo exposure whereas 6 out of 7 forodesine/dGuo resistant samples remained sensitive for ara-G. For all patient samples tested, 10 out of 30 ara-G resistant samples remained sensitive to forodesine/dGuo exposure and 19 out of 39 forodesine/dGuo resistant samples were still sensitive to ara-G exposure. Therefore, resistance to ara-G exposure did not preclude sensitivity to forodesine/dGuo exposure and vice versa, and suggests that the modes of cytotoxicity or resistance between forodesine and ara-C or ara-G are different. In contrast, LC50 values for ara-C and ara-G cytotoxicities strongly correlated ($p<0.001$, $R=0.72$; **Figure 3C**), indicating that the cytotoxic mechanisms are the same for ara-G and ara-C compounds.

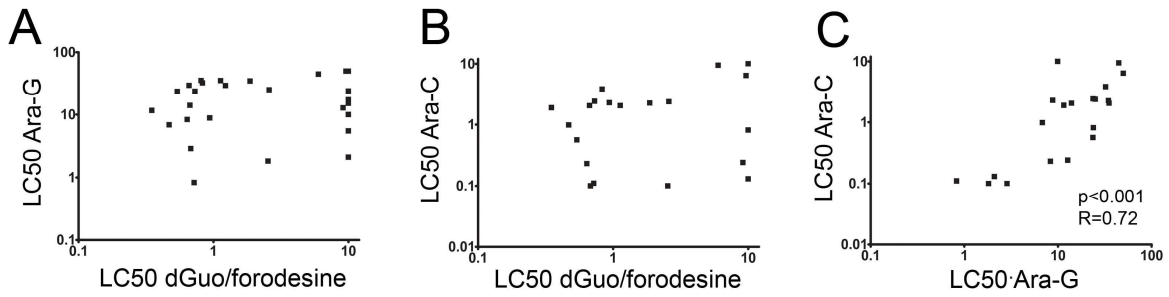



**Figure 3. Relation between LC50 values of forodesine/dGuo, ara-G and ara-C in T-ALL**. Relationship between (**A**) ara-G and forodesine/dGuo LC50 values, (**B**) between ara-C and forodesine/dGuo LC50 values and (**C**) between ara-G and ara-C LC50 values. LC50 values for ara-G and ara-C were available for 28 and 21 T-ALL patients respectively.

## dGTP accumulation

To investigate whether differences in forodesine sensitivity levels could be attributed to differences in intracellular accumulation of dGTP, we analysed dGTP levels among patient samples in the absence or presence of forodesine. After 24 hours, no significant differences were found in proliferation rate or the number of apoptotic cells between forodesine/dGuo-treated or dGuo-treated control cells (not shown). Without blocking PNP activity, T-ALL patient samples accumulated higher basal intracellular dGTP levels within 24 hours than BCP-ALL samples (p=0.004) (**Figure 4A),** so BCP-ALL cells may have a higher intrinsic ability to degrade dGuo levels than T-ALL cells or have a slower conversion rate of dGuo into dGTP). Upon blockage of PNP by forodesine, total intracellular dGTP levels increased 10 to 100-fold within 24 hours (**Figure 4B**).

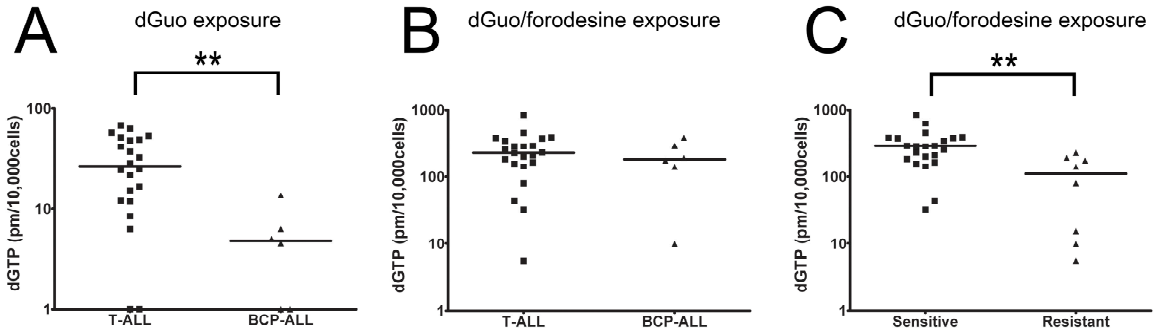



**Figure 4. dGTP accumulation.** (A) Basal dGTP levels after 24 hrs of 10 µM of dGuo exposure and (B) dGTP accumulation after 24 hrs of 10 µM dGuo and 1 µM forodesine exposure in 22 T-ALL and 6 BCP-ALL patient samples. Undetectable dGTP levels have been assigned a value of 1. (C) Intracellular dGTP levels after 24 hrs 10 µM dGuo and 1

μM forodesine exposure in forodesine-sensitive versus resistant patients (22 T-ALL, 6 BCP-ALL and 2 AML samples). Horizontal lines represent median values. ** p<0.01.

No difference was observed between T-ALL and BCP-ALL samples indicating that both ALL types are equally efficient to convert dGuo into dGTP. Intracellular dGTP accumulation was significantly higher for forodesine sensitive cells than for resistant cells (p=0.001, **Figure 4C**). So, resistant patients may convert less dGuo into dGTP or resistant patients more efficiently consume (toxic) dGTP levels.

**Gene expression**

To find potential explanations for differences in forodesine or ara-G sensitivity levels, we determined mRNA expression levels of different transporters and enzymes that are involved in the purine metabolism (**Figure 1**). Of the 13 genes investigated, 4 genes (*CNT1, CNT2, CNT3, NT5C1A*) were expressed at low to undetectable levels in most of our patient samples and were therefore excluded from further analyses. *ENT1* and *ENT2* were both expressed at higher levels in T-ALL samples than in BCP-ALL samples (p=0.007 and p=0.036 respectively) while levels of the nucleotidase *NT5C2*/*PNT5* and *PNP* were expressed at lower levels (p=0.016 and p<0.001, respectively; **Figure 5**).

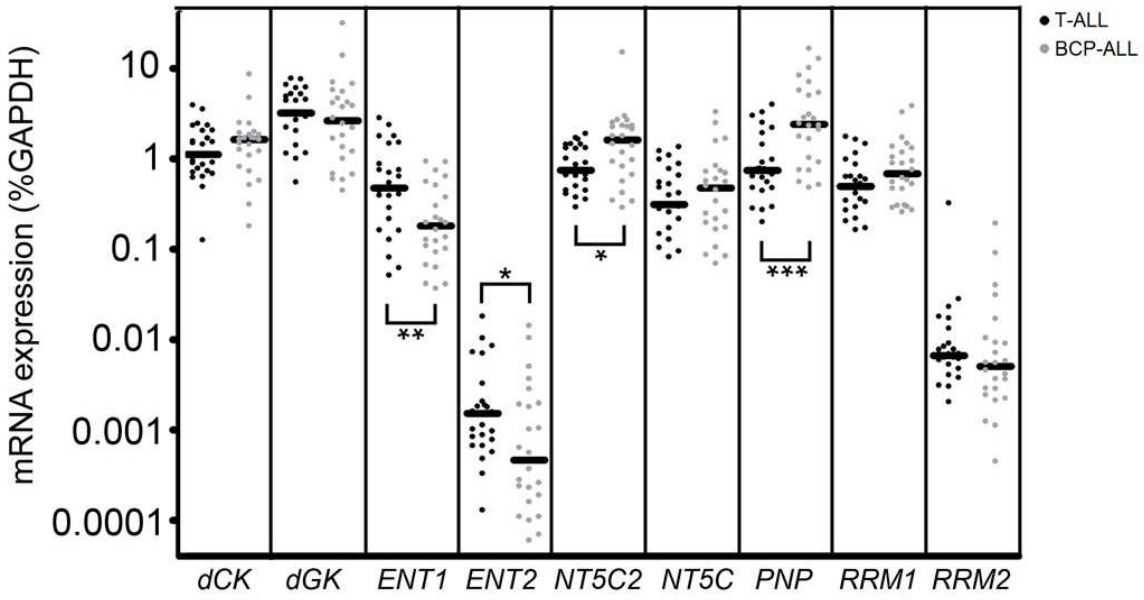



**Figure 5. Gene expression in leukemia subtypes**. mRNA expression of 9 genes in T-ALL and BCP-ALL patients. Each dot represents a measurement in one patient sample. cDNA was available for 25 T-ALL samples, 24 BCP-ALL samples. * p<0.05, ** p<0.01. ***p<0.001.

Patient samples sensitive to forodesine/dGuo expressed higher levels of *dGK* (p=0.039; **Figure 6A**), and may more efficiently convert dGuo into dGMP as a first activation step in the conversion of dGuo into dGTP. *ENT1* and *ENT2* levels were significantly higher in ara-G sensitive patients than in resistant patients (p=0.010 and p=0.009, respectively; **Figure 6B**) permitting a higher uptake of ara-G. *ENT1* expression levels strongly correlated with ara-G sensitivity levels (p=0.005 R=-0.503). Also for T-ALL samples, we found a correlation between *ENT1* levels and ara-C sensitivity (p=0.011 R=-0.60). Strikingly, *ENT1* and *ENT2* levels were not related to forodesine sensitivity, indicating that cellular uptake of forodesine may be facilitated by another transporter.

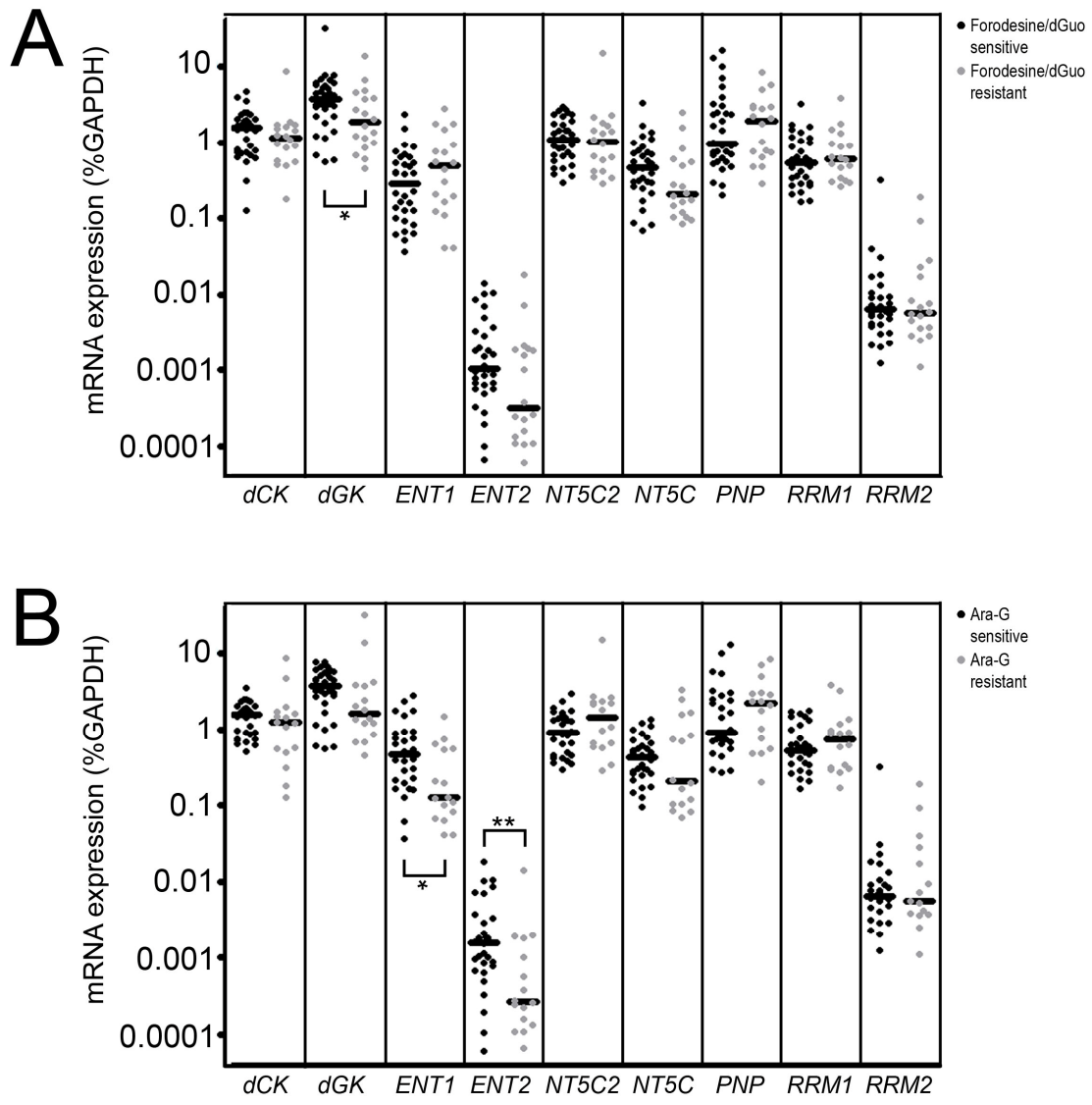


**Figure 6. (page 154) Gene expression in relation to forodesine/dGuo or ara-G sensitivity**. mRNA expression of 9 genes in forodesine/dGuo (A) or ara-G (B) sensitive and resistant patient samples. Each dot represents a measurement in one patient sample. * $p<0.05$, ** $p<0.01$.

## Combination studies

In the treatment of leukemia, multiple drugs are administered simultaneously or administered sequentially. It is therefore important to test for drug interactions. To this end we explored the presence of synergistic, additive or antagonistic effects between forodesine/dGuo and 7 other compounds that are currently used in ALL treatment protocols. Leukemic cells were incubated with a concentration range of these 7 compounds with or without the LC10 or LC30 cytotoxic dGuo concentrations (0.02 μM and 0.48 μM for T-ALL and 0.5 μM and 3.5 μM for BCP-ALL, respectively) in the presence of 1 μM of forodesine. As controls, samples were incubated with LC10 or LC30 concentrations of dGuo with 1 μM forodesine only. For prednisone, vincristine, and asparaginase, no significant synergistic or antagonistic effects were found in combination with forodesine/dGuo. To our surprise, no antagonism was observed between forodesine/dGuo and ara-G or ara-C despite the fact that these drugs depend on the same enzymes of the guanosine salvage pathway. For daunorubicin, addition of LC10 forodesine/dGuo levels resulted in an increase of cellular viability, both for T-ALL as well as for BCP-ALL samples (42% vs 61% for T-ALL ($p=0.009$) and 45% vs 66% for BCP-ALL ($p=0.018$). This effect was only observed at a daunorubicin concentration of 0.125 μg/ml (**Supplementary Figure 2A-B**), but not at other daunorubicin concentrations. Also, no antagonistic effect was measured for any of the daunorubicin concentrations combined with the LC30 forodesine/dGuo level. For various concentrations of 6MP combined with the LC10 or LC30 concentrations forodesine/dGuo, synergistic toxicity was observed for T-ALL samples (**Supplementary Figure 2C, D**).

**DISCUSSION**

In this study, we have demonstrated selective toxicity of forodesine/dGuo treatment for pediatric T-ALL compared to BCP-ALL and AML samples. The median forodesine/dGuo LC50 value was more than 5 fold lower for T-ALL than for BCP-ALL samples. Only one out of 17 AML patients reached an LC50 below 10µM. This patient was also a Down syndrome patient, a syndrome known to display increased sensitive to a wide range of drugs and these patients are highly susceptibility towards toxic side effects[30,31]. High sensitivity of pediatric T-ALL patients towards forodesine/dGuo exposure is in line with expectations, as natural occurring PNP deficiency is known to result in T-cell lymphopenia[12,13], and provided the rationale to develop PNP inhibitors for treatment of T-cell malignancies. Forodesine is a very potent inhibitor of PNP that inhibits PNP activity in the picomolar range in biochemical experiments[32]. Cytotoxic effects of forodesine were shown on T-ALL cell lines before[7], and a clinical response has been documented in a phase I trial for advanced T-cell malignancies[19]. Our *in-vitro* studies indicated that 1µM of forodesine is sufficient to inhibit PNP activity in a cellular system, which is well within clinical achievable plasma concentrations. Steady-state forodesine levels that range between 4 to 8 µM were documented in the plasma of patients following intravenous infusion of 40 mg/m$^2$ of forodesine[19]. In this clinical phase I trial, elevated dGuo levels up to 34 µM in plasma were documented. As the median LC50 dGuo levels (in the presence of 1 µM of forodesine) in our study for forodesine responsive T-ALL samples was estimated on 1.6 µM (range 0.31-10 µM), this indicates that forodesine may be a promising compound in future clinical trials for nearly 75 percent of pediatric T-ALL patients.

In the present study, we demonstrate that nearly half of all BCP-ALL patient samples responded to dGuo/forodesine with dGuo LC50 values that ranged between 0.67 and 10 µM. Again, this is well within clinical achievable plasma dGuo levels following forodesine infusion, suggesting that forodesine treatment may be effective for nearly 50 percent of BCP-ALL samples.

Selective T-cell toxicity was also demonstrated for the arabinoguanosine derivative compound ara-G. Primary T-ALL patient samples had a median LC50 value of 20.5 µM ara-G whereas about half of BCP-ALL or AML samples did not reach an LC50 within the limits of our assay. T-cell selective toxicity of ara-G is in line with previous studies[14-17,33], and nelarabine is an approved drug for T-cell malignancies[23]. One of the explanations for selective T-cell toxicity by

forodesine/dGuo or ara-G treatment is the finding that T-ALL samples express less *PNP*, which is in line with our previous finding that T-ALL cells have lower PNP activity compared to BCP-ALL cells[34]. Also, the expression of cytosolic purine 5-prime nucleotidase *NT5C2* was lower in T-ALL cells than in BCP-ALL cells, so T-ALL cells have a reduced capacity to revert phosphorylation of dGuo. The expression of the equilibrative nucleoside transporters *ENT1* and *ENT2* was higher for T-ALL than for BCP-ALL cells, possibly resulting in enhanced cellular uptake of dGuo and ara-G. Lower expression levels of *PNP* and *NT5C2* but higher expression of *ENT1* and *ENT2* transporters in T-ALL cells are in line with our finding of higher basal intracellular dGTP levels after exposure to dGuo in T-ALL patient samples than in BCP-ALL samples. However, following inhibition of PNP activity by forodesine, both responding T-ALL and BCP-ALL samples seem equally efficient to accumulate comparable levels of intracellular dGTP. So, differential sensitivity for T-ALL and B-ALL cells towards forodesine may not be due to differences in the dGuo to dGTP activation steps in the purine salvage pathway, but may be due to differential cytotoxic effects of accumulated dGTP levels on ribonucleotide reductase activity and inhibition of DNA synthesis, or intrinsic differences in the apoptotic thresholds between T-cell and B-cells.

Although dGuo mediated toxicity through forodesine and ara-G toxicity depends on stepwise phosphorylation steps in the purine salvage pathway, no relationship could be demonstrated between forodesine/dGuo sensitivity and ara-G sensitivity. This was further supported by the fact that resistance to ara-G exposure did not preclude sensitivity for forodesine/dGuo or vice versa. In contrast, sensitivity levels towards ara-G strongly correlated with ara-C sensitivity levels. Despite the fact that T-ALL samples have different expression levels of enzymes and transporters that favour preferential phosphorylation of dGuo or ara-G in T-ALL cells compared to BCP-ALL cells, our results imply that toxicity levels for both compounds are determined by different components in the purine salvage pathway. For this, dCK has been suggested as an important and rate-limiting factor in the phosphorylation of pyrimidine and purine deoxynucleosides[9] that has been associated with ara-C resistance[35-38] or relapse[39,40]. However, we did not observe differences in *dCK* expression levels between forodesine/dGuo sensitive and resistant patients, nor between ara-G sensitive and resistant patients. In our previous study on infant BCP-ALL, a 2 fold lower expression in *dCK* levels was identified despite a 3.3 fold higher sensitivity levels towards ara-C compared to non-infant ALL patients[25]. This indicates that dCK is not a major contributer to

ara-C, ara-G or forodesine/dGuo toxicity, even when non-physiological high levels of deoxycytidine can block ara-G toxicity[14,17].

We observed significant differences in the mitochondrial deoxyguanosine kinase (*dGK*) expression levels between forodesine/dGuo sensitive and resistant patient samples, but not between ara-G sensitive and resistant patients. This finding is completely in line with previous findings by Gandhi and co-workers who demonstrated that dGuo is predominantly phosphorylated by dGK but not by dCK, whereas ara-G can be phosphorylated by both enzymes with dGK as preferential enzyme at limiting ara-G concentrations[41]. Ara-G resistance could be associated with significant lower expression levels of the *ENT1* and *ENT2* transporters. These transporters have been shown important for the import of ara-C[42], and elevated *ENT1* levels have been reported to explain the high ara-C sensitivity of infant ALL, and a strong correlation was observed between *ENT1* expression levels and ara-C sensitivity[25]. Lower *ENT1* expression levels have been related to ara-C resistance in childhood AML[29]. Previous work by Huang et al[43] on the T-ALL cell line CCRF-CEM demonstrated that while the cellular uptake of forodesine was dependent on ENT1 and ENT2, forodesine toxicity was not. This is in agreement with our data, and *ENT1* and *ENT2* expression levels were not related to forodesine toxicity levels. These data therefore suggest that forodesine import and subsequent PNP inhibition seems not limited in leukemia cells but may depend on the import and activation of dGuo. Import of dGuo has been reported to occur via concentrative nucleoside transporters[43]. Although observations as described above may contribute to forodesine/dGuo or araG resistance, exact resistance mechanisms are not yet clear. For CLL blasts, forodesine/dGuo effectiveness has been related to basal levels of MCL1 and BIM, elevated phospho-dCK to dCK ratios following treatment, and induction of p73 that may upregulate BIM via the FOXO1 and FOXO3A transcription factors[6]. A recent study provided an alternative mechanism of forodesine resistance as marrow stromal cells were shown to antagonize forodesine-enforced apoptosis in CLL cells[44].

The combination cytotoxicity assays revealed no antagonistic or synergistic effect of forodesine/dGuo combined with prednisone, vincristine or asparaginase. For daunorubicin we observed an antagonistic effect, but only at a single concentration combined with the LC10, but not with the concentration of LC30 forodesine/dGuo. We found no antagonistic effect for forodesine/dGuo with either the purine analogue ara-G, nor with the pyrimidine analogue ara-C. Moreover forodesine/dGuo had a synergistic effect in T-ALL with another purine analogue,

6MP, at multiple concentrations combined with the LC10 and LC30 dGuo/forodesine concentrations. The molecular basis of these differences in combined effects remains elusive.

We conclude that forodesine and ara-G have cytotoxic effects on T-ALL and to a lesser extent on BCP-ALL cells *in-vitro* and could therefore have potential beneficial therapeutic effects in both types of leukemia, possibly in a combined therapy approach. In AML patients forodesine treatment is expected to result in little response. Our study gives no indication of clear antagonistic effects of forodesine/dGuo when combined with any of the 7 drugs as currently used in leukemia therapy.

**ACKNOWLEDGEMENTS**

I.H. is financed by the Dutch Cancer Society (KWF-EMCR 2006-3500). W.K.S. is financed by the Stichting Kinderen Kankervrij (KiKa; Grant no. KiKa 2008-029).This work was supported by research funding from Mundipharma International Ltd. to J.P.P M and C.M.Z.

**AUTHORSHIP CONTRIBUTIONS**

I.H. wrote manuscript and performed experiments, C.M.Z. wrote manuscript and designed experiments, C.Y.M., F. H. and S.B. designed experiments, C.P. and W.K.S. performed experiments, R.P designed experiments and wrote manuscript, J.P.P.M. was principal investigator, designed study and wrote manuscript.

**CONFLICT OF INTEREST DISCLOSURES**

C.P. and S.B. are employees of Biocryst Pharmaceuticals Inc. C.Y.M an F.H. are employees of Mundipharma International Ltd. J.P.P.M and C.M.Z. received research funding from Mundipharma International Ltd for this study. The remaining authors declare no competing financial interests.

## REFERENCES


1. Pui CH, Evans WE. Treatment of acute lymphoblastic leukemia. *N Engl J Med*. 2006;354(2):166-178.
2. Kaspers GJ, Zwaan CM. Pediatric acute myeloid leukemia: towards high-quality cure of all patients. *Haematologica*. 2007;92(11):1519-1532.
3. Bassan R, Hoelzer D. Modern therapy of acute lymphoblastic leukemia. *J Clin Oncol*. 2011;29(5):532-543.
4. Burnett A, Wetzler M, Lowenberg B. Therapeutic advances in acute myeloid leukemia. *J Clin Oncol*. 2011;29(5):487-494.
5. Miles RW, Tyler PC, Furneaux RH, Bagdassarian CK, Schramm VL. One-third-the-sites transition-state inhibitors for purine nucleoside phosphorylase. *Biochemistry*. 1998;37(24):8615-8621.
6. Alonso R, Lopez-Guerra M, Upshaw R, et al. Forodesine has high antitumor activity in chronic lymphocytic leukemia and activates p53-independent mitochondrial apoptosis by induction of p73 and BIM. *Blood*. 2009;114(8):1563-1575.
7. Bantia S, Ananth SL, Parker CD, Horn LL, Upshaw R. Mechanism of inhibition of T-acute lymphoblastic leukemia cells by PNP inhibitor--BCX-1777. *Int Immunopharmacol*. 2003;3(6):879-887.
8. Kicska GA, Long L, Horig H, et al. Immucillin H, a powerful transition-state analog inhibitor of purine nucleoside phosphorylase, selectively inhibits human T lymphocytes. *Proc Natl Acad Sci U S A*. 2001;98(8):4593-4598.
9. Balakrishnan K, Nimmanapalli R, Ravandi F, Keating MJ, Gandhi V. Forodesine, an inhibitor of purine nucleoside phosphorylase, induces apoptosis in chronic lymphocytic leukemia cells. *Blood*. 2006;108(7):2392-2398.
10. Carson DA, Kaye J, Matsumoto S, Seegmiller JE, Thompson L. Biochemical basis for the enhanced toxicity of deoxyribonucleosides toward malignant human T cell lines. *Proc Natl Acad Sci U S A*. 1979;76(5):2430-2433.
11. Carson DA, Kaye J, Seegmiller JE. Lymphospecific toxicity in adenosine deaminase deficiency and purine nucleoside phosphorylase deficiency: possible role of nucleoside kinase(s). *Proc Natl Acad Sci U S A*. 1977;74(12):5677-5681.
12. Cohen A, Gudas LJ, Ammann AJ, Staal GE, Martin DW, Jr. Deoxyguanosine triphosphate as a possible toxic metabolite in the immunodeficiency associated with purine nucleoside phosphorylase deficiency. *J Clin Invest*. 1978;61(5):1405-1409.
13. Giblett ER, Ammann AJ, Wara DW, Sandman R, Diamond LK. Nucleoside-phosphorylase deficiency in a child with severely defective T-cell immunity and normal B-cell immunity. *Lancet*. 1975;1(7914):1010-1013.
14. Cohen A, Lee JW, Gelfand EW. Selective toxicity of deoxyguanosine and arabinosyl guanine for T-leukemic cells. *Blood*. 1983;61(4):660-666.
15. Verhoef V, Fridland A. Metabolic basis of arabinonucleoside selectivity for human leukemic T- and B-lymphoblasts. *Cancer Res*. 1985;45(8):3646-3650.
16. Shewach DS, Daddona PE, Ashcraft E, Mitchell BS. Metabolism and selective cytotoxicity of 9-beta-D-arabinofuranosylguanine in human lymphoblasts. *Cancer Res*. 1985;45(3):1008-1014.

17. Ullman B, Martin DW, Jr. Specific cytotoxicity of arabinosylguanine toward cultured T lymphoblasts. *J Clin Invest*. 1984;74(3):951-955.
18. Lambe CU, Averett DR, Paff MT, Reardon JE, Wilson JG, Krenitsky TA. 2-Amino-6-methoxypurine arabinoside: an agent for T-cell malignancies. *Cancer Res*. 1995;55(15):3352-3356.
19. Gandhi V, Kilpatrick JM, Plunkett W, et al. A proof-of-principle pharmacokinetic, pharmacodynamic, and clinical study with purine nucleoside phosphorylase inhibitor immucillin-H (BCX-1777, forodesine). *Blood*. 2005;106(13):4253-4260.
20. Balakrishnan K, Verma D, O'Brien S, et al. Phase 2 and pharmacodynamic study of oral forodesine in patients with advanced, fludarabine-treated chronic lymphocytic leukemia. *Blood*. 2010;116(6):886-892.
21. Al-Kali A, Gandhi V, Ayoubi M, Keating M, Ravandi F. Forodesine: review of preclinical and clinical data. *Future Oncol*. 2010;6(8):1211-1217.
22. DeAngelo DJ, Yu D, Johnson JL, et al. Nelarabine induces complete remissions in adults with relapsed or refractory T-lineage acute lymphoblastic leukemia or lymphoblastic lymphoma: Cancer and Leukemia Group B study 19801. *Blood*. 2007;109(12):5136-5142.
23. Berg SL, Blaney SM, Devidas M, et al. Phase II study of nelarabine (compound 506U78) in children and young adults with refractory T-cell malignancies: a report from the Children's Oncology Group. *J Clin Oncol*. 2005;23(15):3376-3382.
24. Kurtzberg J, Ernst TJ, Keating MJ, et al. Phase I study of 506U78 administered on a consecutive 5-day schedule in children and adults with refractory hematologic malignancies. *J Clin Oncol*. 2005;23(15):3396-3403.
25. Stam RW, den Boer ML, Meijerink JP, et al. Differential mRNA expression of Ara-C-metabolizing enzymes explains Ara-C sensitivity in MLL gene-rearranged infant acute lymphoblastic leukemia. *Blood*. 2003;101(4):1270-1276.
26. Den Boer ML, Harms DO, Pieters R, et al. Patient stratification based on prednisolone-vincristine-asparaginase resistance profiles in children with acute lymphoblastic leukemia. *J Clin Oncol*. 2003;21(17):3262-3268.
27. Sherman PA, Fyfe JA. Enzymatic assay for deoxyribonucleoside triphosphates using synthetic oligonucleotides as template primers. *Anal Biochem*. 1989;180(2):222-226.
28. Meijerink J, Mandigers C, van de Locht L, Tonnissen E, Goodsaid F, Raemaekers J. A novel method to compensate for different amplification efficiencies between patient DNA samples in quantitative real-time PCR. *J Mol Diagn*. 2001;3(2):55-61.
29. Hubeek I, Stam RW, Peters GJ, et al. The human equilibrative nucleoside transporter 1 mediates in vitro cytarabine sensitivity in childhood acute myeloid leukaemia. *Br J Cancer*. 2005;93(12):1388-1394.
30. Zwaan CM, Kaspers GJ, Pieters R, et al. Different drug sensitivity profiles of acute myeloid and lymphoblastic leukemia and normal peripheral blood mononuclear cells in children with and without Down syndrome. *Blood*. 2002;99(1):245-251.
31. Zwaan MC, Reinhardt D, Hitzler J, Vyas P. Acute leukemias in children with Down syndrome. *Pediatr Clin North Am*. 2008;55(1):53-70, x.

32. Bantia S, Miller PJ, Parker CD, et al. Purine nucleoside phosphorylase inhibitor BCX-1777 (Immucillin-H)--a novel potent and orally active immunosuppressive agent. *Int Immunopharmacol*. 2001;1(6):1199-1210.
33. Shewach DS, Mitchell BS. Differential metabolism of 9-beta-D-arabinofuranosylguanine in human leukemic cells. *Cancer Res.* 1989;49(23):6498-6502.
34. Pieters R, Huismans DR, Loonen AH, et al. Adenosine deaminase and purine nucleoside phosphorylase in childhood lymphoblastic leukemia: relation with differentiation stage, in vitro drug resistance and clinical prognosis. *Leukemia*. 1992;6(5):375-380.
35. Dumontet C, Fabianowska-Majewska K, Mantincic D, et al. Common resistance mechanisms to deoxynucleoside analogues in variants of the human erythroleukaemic line K562. *Br J Haematol*. 1999;106(1):78-85.
36. Kawasaki H, Shindou K, Higashigawa M, et al. Deoxycytidine kinase mRNA levels in leukemia cells with competitive polymerase chain reaction assay. *Leuk Res*. 1996;20(8):677-682.
37. Owens JK, Shewach DS, Ullman B, Mitchell BS. Resistance to 1-beta-D-arabinofuranosylcytosine in human T-lymphoblasts mediated by mutations within the deoxycytidine kinase gene. *Cancer Res*. 1992;52(9):2389-2393.
38. Stegmann AP, Honders MW, Kester MG, Landegent JE, Willemze R. Role of deoxycytidine kinase in an in vitro model for AraC- and DAC-resistance: substrate-enzyme interactions with deoxycytidine, 1-beta-D-arabinofuranosylcytosine and 5-aza-2'-deoxycytidine. *Leukemia*. 1993;7(7):1005-1011.
39. Kakihara T, Fukuda T, Tanaka A, et al. Expression of deoxycytidine kinase (dCK) gene in leukemic cells in childhood: decreased expression of dCK gene in relapsed leukemia. *Leuk Lymphoma*. 1998;31(3-4):405-409.
40. Stammler G, Zintl F, Sauerbrey A, Volm M. Deoxycytidine kinase mRNA expression in childhood acute lymphoblastic leukemia. *Anticancer Drugs*. 1997;8(5):517-521.
41. Rodriguez CO, Jr., Mitchell BS, Ayres M, Eriksson S, Gandhi V. Arabinosylguanine is phosphorylated by both cytoplasmic deoxycytidine kinase and mitochondrial deoxyguanosine kinase. *Cancer Res*. 2002;62(11):3100-3105.
42. White JC, Rathmell JP, Capizzi RL. Membrane transport influences the rate of accumulation of cytosine arabinoside in human leukemia cells. *J Clin Invest*. 1987;79(2):380-387.
43. Huang M, Wang Y, Gu J, et al. Determinants of sensitivity of human T-cell leukemia CCRF-CEM cells to immucillin-H. *Leuk Res*. 2008;32(8):1268-1278.
44. Balakrishnan K, Burger JA, Quiroga MP, et al. Influence of bone marrow stromal microenvironment on forodesine-induced responses in CLL primary cells. *Blood*. 2010;116(7):1083-1091.

# Summary, general discussion and future perspectives

# SUMMARY, GENERAL DISCUSSION and FUTURE PERSPECTIVES

## Solving pieces ot the 'type-A-puzzle'

At the beginning of this project, no oncogenic rearrangement or major driving aberration had been identified in ~40% of pediatric T-ALL patients. We denoted these cases as "unknown" T-ALL cases. In the **first chapter** of this thesis we aimed to identify novel aberrations in these patients by combined genome-wide expression analyses and detailed molecular-cytogenetic analyses. Supervised gene-expression profiling led to the identification of two novel T-ALL subgroups that comprised more than half of these unknown cases. These novel subgroups each had a distinct expression signature and each subgroup comprised approximately 10% of T-ALL cases (Figure 1).

The first group, which we denoted the proliferative cluster, was characterized by high expression of genes involved in cell proliferation and ectopic expression of the NK-like transcription factor *NKX2-1*. In unsupervised analyses, these cases formed a single cluster that also comprised most of the *TLX1* rearranged cases. These cases were characterized by an immunophenotype that corresponded to a T-cell development arrest at the cortical stage. This analysis implies a possible common pathobiology for *TLX1* and *NKX2-1* aberrations, or simply a similar preference to occur in cells arrested in the cortical stage. In 7 out of 12 patients of this cluster, we found novel chromosomal rearrangements involving the *NKX2-1* gene or its homologue *NKX2-2*. These aberrations included translocations and inversions involving *NKX2-1* or *NKX2-2* and the T-cell receptor genes or the *IGH@* locus and resulted in ectopic upregulation of *NKX2-1* or *NKX2-2*.

The second group, which we called the immature cluster, was characterized by an immature immunophenotype and high expression of the transcription factor *MEF2C*. In this group we identified novel genetic rearrangements that targeted the transcriptional activation of *MEF2C*.

Beside the proliferative and immature cluster, most of the other unknown T-ALL cases clustered together with *TAL1* and *LMO2* rearranged patients (Figure 1). In the *TAL/LMO*-like patients, rearrangements were identified in genes that are homologous to *TAL1* or *LMO2*, such as *LMO1*[1], *LMO3*[2], *TAL2*[1] and *LYL1* (**Chapter 2**). Therefore at the end of this project, 99% of pediatric T-ALL patients can now be classified into 6 T-ALL subgroups (see **Figure 1**), and in more than

half of the 40% of the unknown T-ALL cases, novel type A abnormalities have been identified.

With the advances in whole-genome sequencing and the development of bioinformatic tools to facilitate data processing, it will become possible in the near future to analyse the whole genome of a leukemic patient sample, and thereby simultaneously map mutations and breakpoint regions of deletions, amplifications, translocations and inversions, at the single basepair level.[3] The costs for these applications are still decreasing, making it perhaps the best suitable tool to identify additional type A and type B aberrations in the future.

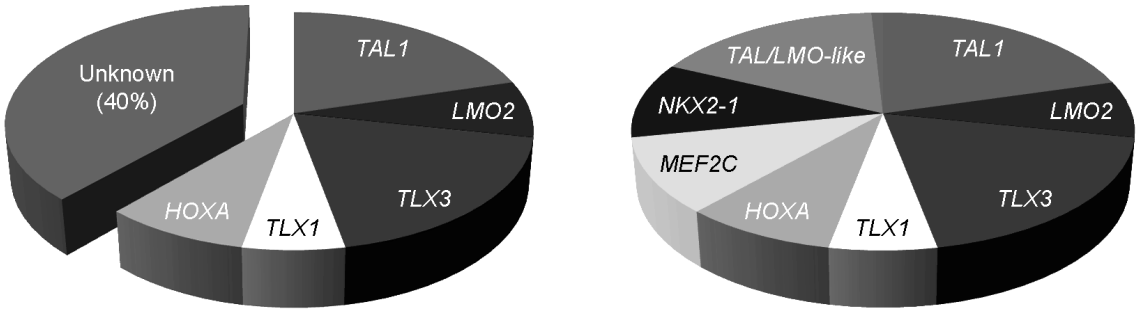



**Figure 1**: Schematic overview of classification of 117 pediatric T-ALL patients according to gene-expression profile (corresponding to type A aberration). The left pie-chart refers to 4 major genetic groups that were known at the start of this thesis (*TAL1* and *LMO2* form the *TAL/LMO* subgroup). The right pie-chart demonstrates how unknown cases can be assigned to one of 2 novel T-ALL subgroups, or the TAL/LMO like group that clusters with *TAL1* and *LMO2* rearranged cases.

## Type A aberrations: mechanism of action

Besides identifying type A aberrations, it is important to determine how these aberrations contribute to leukemogenesis, since these insights will help to develop future targeted therapeutic approaches. For most type A aberrations, these mechanism are not yet clear, though in general they are considered to block T-cell differentiation. This is supported by the fact that type A aberrations are associated with arrest in distinct maturation stages[4] (reviewed by Meijerink[5]). In addition to this, effects on other processes such as proliferation and apoptosis have also been described (partly reviewed in **Chapter 4**). In **Chapter 1**, we identified *MEF2C* and *NKX2-1* as two novel type A oncogenes. We have shown that upregulation of these genes increased colony formation in a cellular assay, indicating effects on

proliferation. We demonstrated that ectopic *MEF2C* expression inhibits differentiation and that part of the gene-expression signature of the immature cluster is caused by MEF2C. This indicates that the transcription factor MEF2C has wide ranging effects on many different downstream genes, which are involved in proliferation as well as differentiation. In the future, the mechanisms of action of MEF2C and NKX2-1 will be further investigated. We have initiated Chip-on-chip experiments on patient and cell line material, which aims to identify direct transcriptional target genes that are controlled by MEF2C and NKX2-1. The effect of the over expression of *MEF2C* and *NKX2-1* on thymocyte differentiation will be addressed using the stromal-support culturing system OP9-DL1[6], which facilitates the differentiation of hematopoietic stem cells into the T-cell lineage. The next step is to study *NKX2-1*, *MEF2C* and other type A oncogenes in conditional knock-down and knock-in experiments *in-vitro* as well as in xenograft mouse models. These model systems are currently in development in our lab. This will enable the investigation of many aspects ranging from downstream targets to the effect of potential drugs.

**NKL overexpression as common theme in T-ALL**

So far we have focused on different subgroups in T-ALL, but in **Chapter 4** we have reviewed the role of NK-like (NKL) homeobox genes and their participation in T-ALL. More than half of pediatric T-ALL samples have over expression of NKL genes. Most of these NKLs are not expressed in normal T-cell development. These NKL genes could have a similar downstream effect that promotes leukemogenesis in T cell progenitors, possibly due to mimicking of a NKL gene that has a role in hematopoietic development, like *HHEX*. All NKL genes that are implicated in T-ALL posses a conserved Eh1 repressor motif. Future research has to be conducted to determine whether this repressor motif may facilitate a leukemogenic effect and whether there is evidence for a common downstream pathway of these NKL genes, or whether these genes each fulfill specific pathologic roles. To identify a common downstream pathway of NKL homeobox genes, one cannot compare T-ALL samples with each other. Healthy thymocyte subpopulations reflecting different maturation stages could be compared to their malignant (T-ALL) counterparts, and in this way T-ALL specific genes may be identified, that are common in all T-ALL subgroups. ChIP-on-chip experiments could be used to find a set of genes that is targeted by multiple NKL genes and that could be associated with T-ALL. If a common pathway for these NKL genes is

identified, this may offer therapeutic opportunities for a large part of pediatric T-ALL patients.

**A multitude of hits model?**

In T-ALL, we have discussed the existence of type A and B aberrations; however, a two-hit model is not adequate to explain leukemogenesis in T-ALL. Often more than just two genetic aberrations are present at once in leukemic cells, and not all aberrations have been identified yet.

Interestingly, several aberrations seem to have a 'preference' to occur together in T-ALL. For example, in **Chapter 3** we identified a recurrent del(5)(q35) that was exclusively found in combination with *TLX3* rearranged patients. In **Chapter 5** we show that *NOTCH1* mutations are found significantly more frequent in *TLX3* rearranged patients than in other subtypes, and relatively less frequent within the *TAL/LMO* subgroup. Several reasons may be given for these associations. One aberration could be the consequence of the other. In the case of the *TLX3* accompanying aberration del(5)(q35), the aberration is very close to the *TLX3* locus and might be part of a more complex rearrangement involving multiple sites. Another possibility is that both aberrations act synergistically in T-cell pathogenesis. A third possibility is that both aberrations need a similar maturation stage to exert an oncogenic effect. In **Chapter 2**, we identified two type A aberrations in a single T-ALL patient: a *LYL1* translocation and a small deletion near *LMO2*. Usually type A aberrations are mutually exclusive, but within the *TAL/LMO* subgroup this might be slightly different. Combinations of *TAL* and *LMO* aberrations in a single leukemic cell have been described before[1] and several papers have reported strong synergism between these genes in leukemia development in mice.[7-10] LMO2 and TAL1 have been shown to participate in the same transcription complex and impact the same pathways. This is also reflected by the fact that *TAL1/2* and *LMO1/2/3* rearranged leukemias have similar gene-expression profiles. 'Double' TAL/LMO hits may occur in specific cases, e.g. the small deletion near *LMO2* as found in our *LYL1* translocated case was also identified in a patient that carried an additional *TAL1* aberration. This small LMO2 deletion may 'need' the help of an additional hit to be truly leukemogenic. However, based on the current data this remains speculation and future identification of additional T-ALL samples with double *TAL/LMO* hits will provide more insight into this matter.

Recent pilot experiments of whole genome sequencing in our group have given a glimpse of the extent of additional mutations in pediatric T-ALL samples. In addition, several aberrations are only found in subclonal populations, making the picture even more complex. So, even when part of the puzzle seems to be solved, the reality is getting more complex as many more genetic aberrations remain to be identified, passenger as well as driving mutations, perhaps unique ones in individual patients. Besides the genetic aberrations that play their part in leukemogenesis, it is likely that other types of aberrations may contribute to leukemogenesis as well. In precursor B-cell ALL, miRNAs and epigenetic changes have been identified that may play a role in leukemogenesis.[11-13] Also in T-ALL, more is becoming known of these types of aberrations[11,14-16] and as these research fields are further progressing, it is becoming feasible and well within our reach to identify novel aberration in T-ALL that might lead to the identification of drugable targets.

**Classifying T-ALL**

Stratification of patients based upon the assessment of their prognosis can lead to adjusted treatment protocols to improve survival. E.g. patients harboring Philadephia chromosome have a very poor prognosis and therefore receive treatment according to a specific protocol (EsPhALL protocol). Also, infant precursor B-ALL with MLL-rearrangements have benefited from a dedicated infant-ALL international protocol (Interfant). Other B-lineage and T-lineage ALL patients are stratified into standard, intermediate and high risk treatment arms of a protocol, such as the current protocols of the Dutch Children Oncology Group (DCOG), based on their early response to chemotherapy, often measured by PCR techniques that quantify so-called minimal residual disease.

In recent years, classification based on maturation stage, i.e. the EGIL classification[17] and more recently the TCR classification[18], has shown prognostic relevance. The immature (pro- or pre-T) groups have a relative poor prognosis[19-22] and for the immature group defined by the TCR classification, similar results have been found.[23] Cortical T-ALL, defined by EGIL by the presence of CD1a, has been associated with a good outcome[24-26].

The identification of genetic aberrations and gene-expression profiling has now led to the possibility of classifying T-ALL based on expression-signatures and genetic aberrations, which might prove to be better classifiers or an addition to those already used. Genetic classification has been shown to have prognostic

significance, though results have not been consistent and might also be influenced by differences in treatment protocols. For example, in **Chapter 5** we demonstrated that in pediatric T-ALL, NOTCH1 activating mutations are associated with a good initial *in vivo* prednisone response, but not with a favourable long-term outcome, in fact, a trend towards a poor outcome was visible. Though a study by Clappier et al showed similar results[27], other groups found a favorable prognostic effect of NOTCH1 activating mutations in their cohorts.[28-30] In adult T-ALL similar discrepancies in results have been reported, as two studies report no effect on outcome[31,32] and one study observed a good outcome for *NOTCH1* mutated patients treated according to GRAALL-2003 protocol, but not for patients treated according to LALA-94 protocol.[33] This last study demonstrates that the impact of a genetic aberration on prognosis could depend on the treatment protocol, which is very interesting, as it implies that some protocols are better or less suited for T-ALL patients harboring certain abnormalities.

Another example of a possible T-ALL classifier is the early T-cell progenitor T-ALL signature (ETP T-ALL). This recently identified T-ALL subtype shares many characteristics with normal early thymocytes based upon gene-expression as well as immunophenotypic markers.[34] It was characterized by a very poor prognosis which was confirmed within the same study in an independent T-ALL cohort[34]. These data have now led to the implementation of an intensified treatment protocol for ETP T-ALL patients in the St Jude Children's Research Hospital, Memphis, USA. In **Chapter 1** of this thesis we identified an immature cluster which can largely be predicted by the ETP-ALL signature as described.[34] Thus far, we, we did not find a poor prognosis for this subgroup of immature patients (Chapter 1). The explanation for this discrepancy could lie in the limited number of patients used in our cohort, or perhaps in differences in treatment protocols, and indicates that additional validation in other patient cohorts is required.

So, a big challenge for future research is to determine the prognostic relevance of T-ALL aberrations in different and continuously changing treatment protocols. Therefore continuous research on prognostic relevance of T-ALL aberrations, gene-signatures and other potential classifiers is important. This will enable future meta-analysis that will help in determining the best T-ALL classifiers.

**Future therapy**

In **Chapter 6**, we have investigated two drugs (forodesine and nelarabine) *in-vitro* on pediatric ALL and AML samples. Both drugs target the same enzyme: purine nucleoside phosphorylase (PNP). In humans, a natural occurring deficiency of this enzyme causes a specific T-cell depletion, and therefore these drugs are potentially suitable to treat T-ALL. Indeed we found that T-ALL samples were most sensitive to these drugs *in-vitro*, whereas still half of pediatric precursor B-cell acute lymphoblastic leukemia (BCP-ALL) samples also responded to these drugs. The differences in sensitivity are likely multifactorial, but could partly be explained by differences in dGTP accumulation (which causes cell death) and *dGK* levels (and enzyme that leads to dGTP formation) for forodesine, and differences in *ENT1* and *ENT2* levels (transporters that could transport the drugs into the cell) for nelarabine/ara-G sensitivity. Also, *PNP* and *NT5C2* levels were lower in T-ALL samples compared with B-ALL samples whereas *ENT1* and *ENT2* levels were higher. Although both drugs target similar pathways, resistance to one drug did not preclude sensitivity to the other, which is important information for future treatment strategies.

The above treatment was designed to specifically target T-cells, but not leukemic cells or specific aberrations found in T-ALL. As more and more of the pathology of T-ALL is unraveled, more potential treatment targets are becoming available. Specific genetic hits (*LMO*, *NOTCH*, *MEF2C*, *NKX*) could be targeted, or perhaps a more general subgroup specific pathway e.g. one that activated by *TAL1/2/LYL1* and *LMO1/2/3*. Though for *in-vitro* testing, targets can be specifically knocked down by for example siRNAs or shRNA constructs, in vivo there is no vehicle available to deliver such compounds intact at the right place. To find drugs that can reverse a specific expression profile *in-vitro*, a free online tool is available called 'Connectivity mapping'.[35] This is an easy-to-use method that compares a list of differentially expressed genes (provided by the researcher, for instance MEF2C knock in mice compared to normal mice) to a database that comprises many gene-signature changes related to drug exposure to multiple cell-lines. Perhaps by using such approaches, potential 'signature targeting' drugs can be identified.

**Conclusion**

To conclude, this thesis has elucidated two new genetic abnormalities involving *NKX2-1/2-2* and *MEF2C* and provided additional insight into (common or co-operating) genetic rearrangements and gene-expression profiles in T-ALL, as well as the prognostic relevance of certain T-ALL genetic aberrations and the *in-vitro* sensitivity of ALL samples to two potential new drugs. As our knowledge of genetic aberrations in T-ALL is improving, an important next step will be a further investigation of the leukemogenic mechanisms of these genetic aberrations, using *in-vitro* but also *in-vivo* models. These models could also be used to screen for potential new therapeutics to provide more efficient and more targeted therapy. Then hopefully, the rapid technological advances in genome research will be followed by advances in the treatment of children with T-ALL.

## REFERENCES


1. Van Vlierberghe P, van Grotel M, Tchinda J, et al. The recurrent SET-NUP214 fusion as a new HOXA activation mechanism in pediatric T-cell acute lymphoblastic leukemia. Blood. 2008;111:4668-4680.
2. Simonis M, Klous P, Homminga I, et al. High-resolution identification of balanced and complex chromosomal rearrangements by 4C technology. Nat Methods. 2009;6:837-842.
3. Welch JS, Westervelt P, Ding L, et al. Use of whole-genome sequencing to diagnose a cryptic fusion oncogene. Jama. 2011;305:1577-1584.
4. van Grotel M, Meijerink JP, Beverloo HB, et al. The outcome of molecular-cytogenetic subgroups in pediatric T-cell acute lymphoblastic leukemia: a retrospective study of patients treated according to DCOG or COALL protocols. Haematologica. 2006;91:1212-1221.
5. Meijerink JP. Genetic rearrangements in relation to immunophenotype and outcome in T-cell acute lymphoblastic leukaemia. Best Pract Res Clin Haematol. 2010;23:307-318.
6. Schmitt TM, de Pooter RF, Gronski MA, Cho SK, Ohashi PS, Zuniga-Pflucker JC. Induction of T cell development and establishment of T cell competence from embryonic stem cells differentiated in vitro. Nat Immunol. 2004;5:410-417.
7. Wadman I, Li J, Bash RO, et al. Specific in vivo association between the bHLH and LIM proteins implicated in human T cell leukemia. Embo J. 1994;13:4831-4839.
8. Larson RC, Lavenir I, Larson TA, et al. Protein dimerization between Lmo2 (Rbtn2) and Tal1 alters thymocyte development and potentiates T cell tumorigenesis in transgenic mice. Embo J. 1996;15:1021-1027.
9. Draheim KM, Hermance N, Yang Y, Arous E, Calvo J, Kelliher MA. A DNA-binding mutant of TAL1 cooperates with LMO2 to cause T cell leukemia in mice. Oncogene. 2010;30:1252-1260.
10. Aplan PD, Jones CA, Chervinsky DS, et al. An scl gene product lacking the transactivation domain induces bony abnormalities and cooperates with LMO1 to generate T-cell malignancies in transgenic mice. Embo J. 1997;16:2408-2419.
11. Schotte D, Lange-Turenhout EA, Stumpel DJ, et al. Expression of miR-196b is not exclusively MLL-driven but is especially linked to activation of HOXA genes in pediatric acute lymphoblastic leukemia. Haematologica. 2010;95:1675-1682.
12. Stumpel DJ, Schneider P, van Roon EH, et al. Specific promoter methylation identifies different subgroups of MLL-rearranged infant acute lymphoblastic leukemia, influences clinical outcome, and provides therapeutic options. Blood. 2009;114:5490-5498.
13. Stumpel DJ, Schotte D, Lange-Turenhout EA, et al. Hypermethylation of specific microRNA genes in MLL-rearranged infant acute lymphoblastic leukemia: major matters at a micro scale. Leukemia. 2011;25:429-439.
14. Schotte D, Chau JC, Sylvester G, et al. Identification of new microRNA genes and aberrant microRNA profiles in childhood acute lymphoblastic leukemia. Leukemia. 2009;23:313-322.

15. Schotte D, Pieters R, Den Boer ML. MicroRNAs in acute leukemia: from biological players to clinical contributors. Leukemia. 2011.
16. Schotte D, Moqadam FA, Lange-Turenhout EA, et al. Discovery of new microRNAs by small RNAome deep sequencing in childhood acute lymphoblastic leukemia. Leukemia. 2011;25:1389-1399.
17. Bene MC, Castoldi G, Knapp W, et al. Proposals for the immunological classification of acute leukemias. European Group for the Immunological Characterization of Leukemias (EGIL). Leukemia. 1995;9:1783-1786.
18. Asnafi V, Beldjord K, Boulanger E, et al. Analysis of TCR, pT alpha, and RAG-1 in T-acute lymphoblastic leukemias improves understanding of early human T-lymphoid lineage commitment. Blood. 2003;101:2693-2703.
19. Garand R, Voisin S, Papin S, et al. Characteristics of pro-T ALL subgroups: comparison with late T-ALL. The Groupe d'Etude Immunologique des Leucemies. Leukemia. 1993;7:161-167.
20. Thiel E, Kranz BR, Raghavachar A, et al. Prethymic phenotype and genotype of pre-T (CD7+/ER-)-cell leukemia and its clinical significance within adult acute lymphoblastic leukemia. Blood. 1989;73:1247-1258.
21. Vitale A, Guarini A, Ariola C, et al. Adult T-cell acute lymphoblastic leukemia: biologic profile at presentation and correlation with response to induction treatment in patients enrolled in the GIMEMA LAL 0496 protocol. Blood. 2006;107:473-479.
22. Uckun FM, Gaynon P, Sather H, et al. Clinical features and treatment outcome of children with biphenotypic CD2+ CD19+ acute lymphoblastic leukemia: a Children's Cancer Group study. Blood. 1997;89:2488-2493.
23. Asnafi V, Buzyn A, Thomas X, et al. Impact of TCR status and genotype on outcome in adult T-cell acute lymphoblastic leukemia: a LALA-94 study. Blood. 2005;105:3072-3078.
24. Marks DI, Paietta EM, Moorman AV, et al. T-cell acute lymphoblastic leukemia in adults: clinical features, immunophenotype, cytogenetics, and outcome from the large randomized prospective trial (UKALL XII/ECOG 2993). Blood. 2009;114:5136-5145.
25. Pui CH, Behm FG, Crist WM. Clinical and biologic relevance of immunologic marker studies in childhood acute lymphoblastic leukemia. Blood. 1993;82:343-362.
26. Niehues T, Kapaun P, Harms DO, et al. A classification based on T cell selection-related phenotypes identifies a subgroup of childhood T-ALL with favorable outcome in the COALL studies. Leukemia. 1999;13:614-617.
27. Clappier E, Collette S, Grardel N, et al. NOTCH1 and FBXW7 mutations have a favorable impact on early response to treatment, but not on outcome, in children with T-cell acute lymphoblastic leukemia (T-ALL) treated on EORTC trials 58881 and 58951. Leukemia. 2010;24:2023-2031.
28. 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;108:1151-1157.
29. Kox C, Zimmermann M, Stanulla M, et al. The favorable effect of activating NOTCH1 receptor mutations on long-term outcome in T-ALL patients treated on

the ALL-BFM 2000 protocol can be separated from FBXW7 loss of function. Leukemia. 2010;24:2005-2013.
30. Park MJ, Taki T, Oda M, et al. FBXW7 and NOTCH1 mutations in childhood T cell acute lymphoblastic leukaemia and T cell non-Hodgkin lymphoma. Br J Haematol. 2009;145:198-206.
31. Mansour MR, Sulis ML, Duke V, et al. Prognostic implications of NOTCH1 and FBXW7 mutations in adults with T-cell acute lymphoblastic leukemia treated on the MRC UKALLXII/ECOG E2993 protocol. J Clin Oncol. 2009;27:4352-4356.
32. Baldus CD, Thibaut J, Goekbuget N, et al. Prognostic implications of NOTCH1 and FBXW7 mutations in adult acute T-lymphoblastic leukemia. Haematologica. 2009;94:1383-1390.
33. Asnafi V, Buzyn A, Le Noir S, et al. NOTCH1/FBXW7 mutation identifies a large subgroup with favorable outcome in adult T-cell acute lymphoblastic leukemia (T-ALL): a Group for Research on Adult Acute Lymphoblastic Leukemia (GRAALL) study. Blood. 2009;113:3918-3924.
34. Coustan-Smith E, Mullighan CG, Onciu M, et al. Early T-cell precursor leukaemia: a subtype of very high-risk acute lymphoblastic leukaemia. Lancet Oncol. 2009;10:147-156.
35. Lamb J, Crawford ED, Peck D, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science. 2006;313:1929-1935.

# Nederlandse samenvatting voor de leek

# NEDERLANDSE SAMENVATTING VOOR DE LEEK

## Leukemie

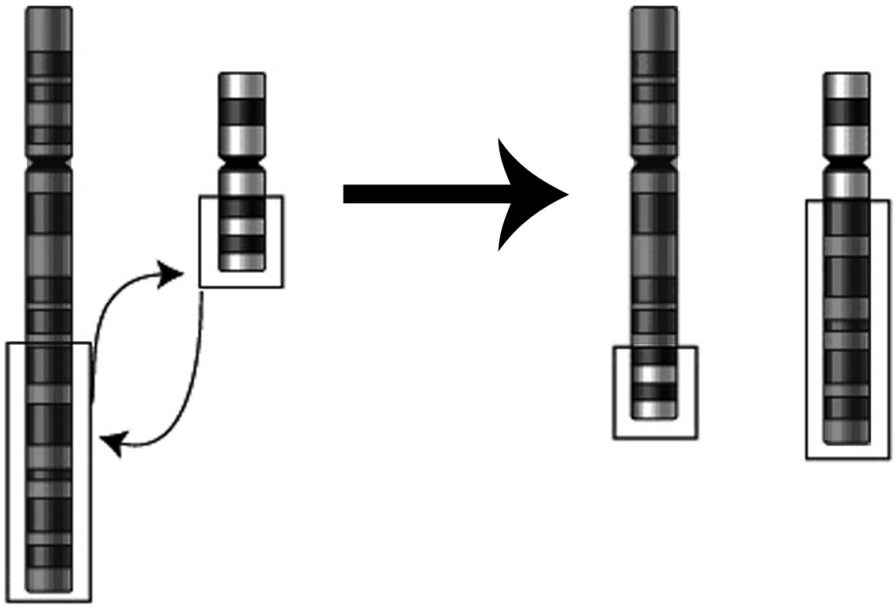
Witte bloedcellen (leukocyten) spelen een belangrijke rol in ons afweer systeem. Ze ontstaan uit stamcellen in het beenmerg. De stamcellen rijpen via voorloper stadia uit tot verschillende soorten leukocyten; lymfatische leukocyten (B- en T-cellen) en myeloïde leukocyten (granulocyten en monocyten). B- en T-cellen zijn met name belangrijk voor ons specifieke (verworven) immuunsysteem. B-cellen rijpen in het **b**eenmerg en T-cel voorloper cellen ontwikkelen zich eerst in het beenmerg en rijpen daarna uit in de **t**hymus (zwezerik). De voorlopers van bloedcellen kunnen door DNA mutaties ongecontroleerd gaan delen, er is dan sprake van leukemie. De leukemie cellen verdringen de aanmaak van gezonde cellen in het beenmerg waardoor bloedarmoede (te weinig rode bloedcellen), bloedingen (te weinig bloedplaatjes) of infecties (te weinig gezonde leukocyten) kunnen ontstaan. Men spreekt van T-cel acute lymfatische leukemie bij ongecontroleerde groei van T-cel voorloper cellen, en over deze aandoening bij kinderen gaat dit proefschrift.

T-cel acute lymfatische leukemie (T-ALL) wordt elk jaar in Nederland bij ongeveer 20 kinderen gediagnosticeerd. Als de kinderen niet behandeld worden, overlijden zij aan deze aandoening. Intensieve behandeling bestaat uit een combinatie van chemotherapeutica, die echter meerdere directe bijwerkingen hebben zoals haarverlies en verminderde weerstand en soms ook complicaties op latere leeftijd, zoals hartfalen en ontwikkeling van andere vormen van kanker. Ondanks verbeterde behandeling in de laatste decennia overlijdt nog steeds ongeveer 30% van de kinderen aan deze ziekte. Een beter inzicht in het ontstaan van deze ziekte kan bijdragen aan de ontwikkeling van nieuwe therapieën. Daarom is het belangrijk om afwijkingen in DNA, die ten grondslag liggen aan de ontwikkeling van T-ALL, in kaart te brengen. In de toekomst zouden we op basis van dit inzicht de behandeling direct kunnen richten op deze afwijkingen. Op deze manier blijven de normale cellen zo veel mogelijk gespaard, dit wordt ook wel ‘targeted therapy’ genoemd. Zo bestrijden we niet alleen de leukemie, maar beperken we ook de bijwerkingen op de lange en korte termijn.

Aan het begin van dit onderzoek waren reeds meerdere DNA afwijkingen (oncogenen) bekend die een rol spelen in het ontstaan van T-ALL. Een deel van de T-ALL patiënten bleek echter geen afwijkingen aan een van deze bekende oncogenen te hebben. Het voornaamste doel van dit onderzoek was daarom het ontdekken van genetische afwijkingen en hun werkingsmechanismen in deze patiënten. Ook wilden we weten of deze afwijkingen voorspellend zijn voor de wijze waarop deze patiënten reageren op therapie.

**Wat zijn DNA afwijkingen?**

Een mens heeft 46 chromosomen in al zijn cellen. Dit zijn lange ketens van dubbelstrengs DNA (**D**eoxyribo**N**ucleic **A**cid). DNA bestaat uit vier verschillende bouwstenen ook wel nucleotiden genaamd: guanine (G), cytosine (C), adenine (A) en thymine (T). In ons DNA ligt de informatie opgeslagen die elke cel nodig heeft om te functioneren. Het grootste deel van de informatie ligt gecodeerd in kleine pakketjes DNA, ook wel genen genoemd. Elk gen bevat de informatie voor de aanmaak van een bepaald eiwit. Ons DNA bevat ongeveer 25.000 verschillende genen. DNA afwijkingen die gerelateerd zijn aan kanker ontregelen vaak een bepaald gen. De meeste vormen van kanker, inclusief leukemie, worden niet overgeërfd van vader of moeder, maar ontstaan door nieuwe afwijkingen in slechts een enkele cel in het lichaam. Er zijn verschillende soorten afwijkingen; er kan bijvoorbeeld een stuk DNA ontbreken (ook wel een deletie genoemd), of er is een stuk vermenigvuldigd (ook wel een amplificatie genoemd). Stukken DNA kunnen ook breken en weer verkeerd aan elkaar gezet worden, dit noemen we een translocatie (zie figuur 1). DNA afwijkingen kunnen meerdere genen beslaan, echter ze kunnen ook heel klein zijn. Er kan slechts één of enkele bouwstenen gedeleteerd worden, of vervangen door andere bouwstenen (deletie/insertie mutaties), of één enkele bouwsteen in het DNA (A, G, C of T) is veranderd door een andere bouwsteen (puntmutatie).



*Figuur 1: Schematisch weergave van een translocatie.*
*Links zijn twee chromosomen weergegeven. Er treedt een breuk op in zowel het grote, donker grijze chromosoom, als het kleine, licht grijze chromosoom. De onderste delen van de chromosomen worden verkeerd hersteld zodat ze verwisseld worden en de rechter situatie ontstaat. Genen die op of vlakbij de breukpunten van translocaties liggen raken hierdoor ontregeld.*

**HOOFDSTUK 1: nieuw ontdekte genetische afwijkingen in T-ALL: *MEF2C* en *NKX2-1***

Binnen T-ALL onderscheiden we verschillende typen genetische afwijkingen die we type A en type B afwijkingen genoemd hebben. Voorgaand onderzoek heeft duidelijk gemaakt dat type A afwijkingen gekoppeld zijn met specifieke rijpingsstadia van leukemische voorloper T-cellen. Ook blijken T-cel leukemieën met dezelfde type A mutaties een specifiek patroon van genen te hebben die aan- of uitstaan (dit patroon wordt ook wel een gen-expressie profiel genoemd). Type A mutaties lijken daarmee dus specifieke T-cel leukemie subgroepen te bepalen. Hierbij blijkt elke subgroep geblokkeerd te staan in een specifiek onrijp ontwikkelingsstadium en ook een karakteristiek gen-expresssie profiel te hebben. Binnen deze verschillende T-cel leukemie groepen kunnen dezelfde type B afwijkingen voorkomen, echter soms is er wel een duidelijke hogere frequentie van bepaalde type B mutaties in specifieke (type A) T-cel leukemie subgroepen. Bij aanvang van dit promotietraject was voor ongeveer 40% van de T-ALL patiënten geen type A afwijking bekend. In hoofdstuk 1 van dit proefschrift hebben we gezocht naar nieuwe type A afwijkingen in deze patiënten. We hebben hiervoor gebruik gemaakt van de gen-expressie data van 117 kinderen met een T-cel leukemie. We hebben bewijs gevonden voor het bestaan van 2 nieuwe T-ALL subgroepen waarvoor geen type A mutatie bekend was. Gebruik makend van een aantal verschillende moleculaire technieken vonden we in een van deze nieuwe subgroepen afwijkingen aan het *NKX2-1* of *NKX2-2* gen, dat hierdoor geactiveerd was. In de andere groep vonden we verschillende genetische afwijkingen die ervoor zorgden dat het *MEF2C* gen hoog aangezet werd. Ook bleek dit de meest onrijpe T-cel leukemie subgroep te zijn. We hebben vervolgens aangetoond dat hoge expressie van *NKX2-1/NKX2-2* of *MEF2C* cellen er toe aanzet om hard te gaan delen en de cellen belet zich verder te ontwikkelen richting een rijpe T-cel. Voor MEF2C werd tevens aangetoond dat het andere genen kan aanzetten zoals o.a. *LMO2* en *LYL1*. Dit zijn twee reeds bekende genen die soms betrokken zijn bij andere T-cel leukemie patiënten.

**HOOFDSTUK 2: *LYL1* afwijkingen lijke op *TAL1* & *LMO2* afwijkingen**

In hoofdstuk 2 van dit proefschrift wordt één T-ALL patiënt beschreven die een zeldzame afwijking, een translocatie, heeft van het *LYL1* gen. Door deze translocatie wordt het *LYL1* gen gekoppeld aan een T-cel receptor gen op een ander chromosoom, hierdoor raakt het *LYL1* gen ontregeld en komt hoog tot expressie. Het gen-expressie profiel van deze patient kwam overeen met een andere bekende T-cel leukemie subgroep die met name gekenmerkt is door afwijkingen aan *TAL1* en/of *LMO2* oncogenen. TAL1 en LMO2 zijn beide regeleiwitten die specifieke genen aan- of uit kunnen zetten. De hoge mate van gelijkenis tussen TAL1 en LYL1, wat ook een regeleiwit is, verklaart waarom deze leukemie een *TAL/LMO* gelijkend expressie profiel heeft.

**HOOFDSTUK 3**: **extra DNA afwijkingen bij *TLX3* afwijkingen**
Een andere T-cel leukemie subgroep, waar ongeveer 20% van de kinderen met T-ALL toe behoort, wordt gekenmerkt door veranderingen in het *TLX3* oncogen. In hoofdstuk 3 hebben we bekeken of er naast *TLX3* afwijkingen ook andere afwijkingen aanwezig waren in deze leukemie subgroep. Dit hebben we gedaan met behulp van een methode waarmee we een totaal overzicht verkregen van alle deleties (verlies van kopieën) en amplificaties (extra kopieën) van het leukemisch DNA, ook wel de array Comparative Genomic Hybridization (array-CGH) techniek genoemd. In leukemisch DNA van een kwart van de patiënten in deze TLX3 subgroep werden extra deleties aan de tip van chromosoom 5 aangetoond, die niet bleken voor te komen in andere T-cel leukemie subgroepen. Mogelijk speelt een verlies van deze chromosomale regio een rol in het ontstaan van deze leukemie, maar het verantwoordelijk gen kon niet worden aangewezen.

**HOOFDSTUK 4: NK-like homeobox genen in T-ALL**
Diverse oncogenen zoals *NKX2-1/NKX2-2*, *NKX2-5* (zoals beschreven in hoofdstuk 1), alsook *TLX1* en *TLX3,* zijn verwante genen die behoren tot de NK-like (NKL) gen familie. Ongeveer 30% van de T-ALL patiënten heeft een DNA afwijking waarbij een NKL gen betrokken is. In hoofdstuk 4 wordt een literatuur overzicht gegeven van genetische afwijkingen aan NKL homeobox genen zoals gevonden in kanker bij de mens. Behalve afwijkingen aan verschillende NKL genen in T-ALL subgroepen blijken ook meerdere NKL genen hoog tot expressie te worden gebracht door andere oncogenen. De meerderheid van de T-ALL gevallen blijkt hoge expressie van een NKL gen te hebben. Dit suggereert een belangrijke rol voor NKL genen in het ontstaan van T-cel leukemie.

**HOOFDSTUK 5: *NOTCH1* & *FBXW7* afwijkingen als voorspellers voor prognose**
Ongeveer 60% van de T-ALL patiënten heeft een afwijking in het *NOTCH1* of *FBXW7* gen. Dit zijn type B afwijkingen waarbij het NOTCH1 eiwit ontregeld raakt en als regulator veel andere genen aan- of uitzet. In hoofdstuk 5 hebben we gekeken of *NOTCH1* en *FBXW7* afwijkingen de behandelingsuitkomst van T-ALL patiënten voorspelt. Deze afwijkingen blijken een goede eerste reactie op therapie met prednison te voorspellen, maar ze blijken geen voorspellende waarde te hebben voor uiteindelijke overleving van kinderen met T-ALL.

**HOOFDSTUK 6: forodesine en ara-G gevoeligheid in T-ALL**

In dit hoofdstuk is de effectiviteit van twee nieuwe geneesmiddelen, Forodesine en ara-G, in het laboratorium getest op leukemie cellen van patiënten met T-cel ALL, B-cel ALL alsook patiënten met een acute myeloide leukemie (AML). Het bleek dat met name T-cel leukemie cellen gevoelig waren voor beide middelen, alsook de helft van alle B-cel leukemieën. AML cellen bleken uiterst ongevoelig. Om beter te begrijpen waarom cellen gevoelig of ongevoelig voor deze middelen waren, is gekeken naar het expressie niveau van diverse genen, die coderen voor enzymen of transporteiwitten die betrokken zijn bij het verwerken en opnemen van deze geneesmiddelen. T-ALL bleek meer transport eiwitten tot expressie te brengen (*ENT1* en *ENT2*). Daarnaast hadden leukemie cellen die gevoelig waren voor Forodesine een hogere expressie van een enzym (*dGK*) dat ervoor zorgt dat Forodesine wordt omgezet naar een actieve vorm. Het bleek dat resistentie tegen het ene middel niet betekende dat deze cellen dan ook resistent waren voor het andere middel. Momenteel lopen er klinische trials met deze medicijnen bij kinderen en volwassenen met T-cel leukemieën.

**CONCLUSIE**

Het onderzoek zoals beschreven in dit proefschrift heeft nieuwe type A afwijkingen in kaart gebracht. Ongeveer een kwart van alle T-cel leukemieën bij kinderen behoort tot een van deze nieuwe subgroepen (MEF2C of NKX2-1/NKX2-2). Afwijkingen aan NK-like homeobox genen zoals *NKX2-1*, *NKX2-2*, *NKX2-5*, *TLX1* en *TLX3* komen bij ongeveer een derde van alle T-cel leukemieën bij kinderen voor. Andere oncogenen blijken ook NK-like genen te ontregelen, en de meerderheid van alle T-cel leukemieën wordt gekenmerkt door hoge expressie van een van deze genen. NK-like homeobox genen lijken daarmee dus veel belangrijker voor T-ALL dan tot dusver bekend. NOTCH1-activerende mutaties (type B) komen bij ongeveer 60 procent van alle T-cel leukemieën voor. In tegenstelling tot resultaten uit andere studies, blijkt dat NOTCH1-activerende mutaties in het Nederlandse DCOG (ALL-7, -8 en -9) alsook het Duitse COALL-97 cohort geen betere overleving te voorspellen. Leukemische cellen van kinderen met een T-cel leukemie blijken zeer gevoelig voor de nieuwe medicijnen Ara-G (Nelarabine) en Forodesine. Ondanks overeenkomende werkingsmechanismen blijkt dat resistentie tegen het ene middel gevoeligheid voor het andere middel niet uitsluit.

Het identificeren van type A afwijkingen, alsook het verder in kaart brengen van bekende en nieuwe type B mutaties is belangrijk, en staat aan de wieg van nieuw onderzoek waarbij verder gekeken gaat worden hoe ontregeling van deze genen kan leiden tot het ontstaan van een T-cel leukemie. Ook zullen nieuwe technieken gebruikt gaan worden zoals ‘whole genome sequencing’ waarbij de complete DNA code van leukemische cellen wordt gecontroleerd om nog andere verborgen genmutaties op te sporen. Betere kennis van deze afwijkingen en hun werkingsmechanismen zal hopelijk leiden tot het ontwikkelen van betere middelen

die leukemie cellen bij de bron kunnen uitschakelen. Dit kan leiden tot een verdere verbetering van het genezingspercentage van deze leukemie bij kinderen en mogelijk ook bij volwassenen. Deze gerichte werking van nieuwe middelen zal mogelijk ook leiden tot minder schade aan normale bloedcellen, weefsels en organen, waardoor de kans afneemt dat patiënten in de toekomst toch last krijgen van bijeffecten.

# About the author

## CURRICULUM VITAE

Irene Homminga werd op 23 januari 1979 geboren te Groningen, waar zij van 1991 tot 1997 het gymnasium doorliep aan het Praedinius Gymnasium. Hierna studeerde zij Wijsbegeerte en Geneeskunde aan de Rijksuniversiteit Groningen. In 2002 behaalde zij haar bachelor Wijsgeer van de Sociale en Praktische Wetenschappen. In 2004 deed zij haar co-schappen in het kader van de geneeskunde studie in het ziekenhuis in Almelo, en een wetenschappelijke stage op de afdeling oogheelkunde van het Leids Universitair Medisch Centrum. In deze stage deed zij onderzoek naar *in-vitro* effecten van geneesmiddelen op uvea melanoom cellijnen, en werd haar enthousiasme voor wetenschappelijk onderzoek gewekt. Na het behalen van haar artsenbul in 2006, is zij dan ook gestart met een promotieonderzoek in het Erasmus Medisch Centrum te Rotterdam. Op de afdeling Kinderoncologie/hematologie (hoofd prof. Rob Pieters) deed zij van 2006 tot 2011 onderzoek onder leiding van dr. Jules Meijerink. Dit promotie traject betrof onderzoek naar de identificatie en klinische relevantie van nieuwe genetische afwijkingen in kinderen met T-cel acute lymfatische leukemie (T-ALL) alsook onderzoek naar nieuwe therapeutische middelen voor de behandeling van T-ALL zoals beschreven in dit proefschrift. Momenteel is Irene werkzaam als post-doc en fertiliteitsarts in het Universitair Medische Centrum Groningen waar zij zowel onderzoek als patiëntenzorg combineert. Irene is getrouwd met Hans Klück, en heeft twee gezonde zonen Max en Henk.

## LIST OF PUBLICATIONS

(Authored and co-authored by Irene Homminga)

**Homminga I**, Zwaan CM, Manz CY, Parker C, Bantia S, Smits WK, Higginbotham F, Pieters R, Meijerink JP. In vitro efficacy of forodesine and nelarabine (ara-G) in pediatric leukemia. *Blood*. 2011 Aug 25;118(8):2184-90.

**Homminga I**, Pieters R, Langerak AW, de Rooi JJ, Stubbs A, Verstegen M, Vuerhard M, Buijs-Gladdines J, Kooi C, Klous P, van Vlierberghe P, Ferrando AA, Cayuela JM, Verhaaf B, Beverloo HB, Horstmann M, de Haas V, Wiekmeijer AS, Pike-Overzet K, Staal FJ, de Laat W, Soulier J, Sigaux F, Meijerink JP.Integrated transcript and genome analyses reveal NKX2-1 and MEF2C as potential oncogenes in T cell acute lymphoblastic leukemia. *Cancer Cell*. 2011 Apr 12;19(4):484-97.

Zuurbier L, **Homminga I**, Calvert V, te Winkel ML, Buijs-Gladdines JG, Kooi C, Smits WK, Sonneveld E, Veerman AJ, Kamps WA, Horstmann M, Petricoin EF 3rd, Pieters R, Meijerink JP.NOTCH1 and/or FBXW7 mutations predict for initial good prednisone response but not for improved outcome in pediatric T-cell acute lymphoblastic leukemia patients treated on DCOG or COALL protocols. *Leukemia*. 2010 Dec;24(12):2014-22.

Simonis M, Klous P, **Homminga I**, Galjaard RJ, Rijkers EJ, Grosveld F, Meijerink JP, de Laat W. High-resolution identification of balanced and complex chromosomal rearrangements by 4C technology. *Nat Methods*. 2009 Nov;6(11):837-42.

Real PJ, Tosello V, Palomero T, Castillo M, Hernando E, de Stanchina E, Sulis ML, Barnes K, Sawai C, **Homminga I**, Meijerink J, Aifantis I, Basso G, Cordon-Cardo C, Ai W, Ferrando A.Gamma-secretase inhibitors reverse glucocorticoid resistance in T cell acute lymphoblastic leukemia. *Nat Med*. 2009 Jan;15(1):50-8.

El Filali M, **Homminga I**, Maat W, van der Velden PA, Jager MJ. Triamcinolone acetonide and anecortave acetate do not stimulate uveal melanoma cell growth. *Mol Vis*. 2008 Sep 24;14:1752-9.

Van Vlierberghe P, **Homminga I**, Zuurbier L, Gladdines-Buijs J, van Wering ER, Horstmann M, Beverloo HB, Pieters R, Meijerink JP.Cooperative genetic defects in TLX3 rearranged pediatric T-ALL. *Leukemia.* 2008 Apr;22(4):762-70.

Smid HG, de Witte MR, **Homminga I**, van den Bosch RJ.Sustained and transient attention in the continuous performance task. *J Clin Exp Neuropsychol*. 2006 Aug;28(6):859-83.

**PhD PORTFOLIO**

# Summary of PhD training and teaching

| | |
|---|---|
| Name PhD student: Irene Homminga<br>Erasmus MC Department:<br>Pediatric Oncology/Hematology<br>Research School: Molecular Medicine | PhD period: 2006 - 2011<br>Promotor: Dr. R. Pieters<br>Supervisor: Dr. J.P.P. Meijerink |

| **1. PhD training** | | |
|---|---|---|
| | **Year** | **Workload (ECTS)** |
| **General courses** | | |
| - Biomedical English Writing and Communication | 2010 | 3 |
| - Statistics | 2009 | 6 |
| - Basic Translational Techniques | 2006 | 1 |
| **Presentations International conferences** | | |
| - Oral presentation & Poster presentation, American Society of Haematology (ASH) annual meeting, Orlando, USA | 2010<br>2009 | 2<br>1 |
| - Poster presentation, ASH, New Orleans, USA | 2008 | 1 |
| - Poster presentation, ASH, San Francisco, USA | | |
| **Other** | | |
| - AIO committee, organising research meetings | 2010 | 2 |
| **2. Teaching** | | |
| | **Year** | **Workload (ECTS)** |
| **Supervising student internships** | | |
| - HLO student, last year 9 month internship | 2009 | 5 |
| - HLO student, last year 7 month internship | 2008 | 4 |

# Dankwoord

**DANKWOORD**

Het onderzoek dat beschreven staat in dit proefschrift was nooit tot stand gekomen zonder de hulp van een heleboel collega's, instellingen en naasten, die ik hieronder graag wil bedanken.

Ten eerste bedank ik de KWF kankerbestrijding voor het financieel mogelijk maken van het onderzoek. Ook wil ik het SKION bedanken voor het beschikbaar stellen van materiaal en informatie, bedankt Valerie de Haas en Edwin Sonneveld. Cooperation from the CoALL group from Germany also was indispensable in this research, thank you very much Dr. Horstmann.

Rob, bedankt voor je scherpe en krachtige analyses op vele momenten tijdens mijn promotietraject. Ik heb veel bewondering voor de manier waarop je elke keer snel en efficiënt met beperkte informatie precies de goede (soms wel heel lastige) vragen wist te stellen.

Jules, ik ben je heel erg dankbaar voor je geweldige begeleiding de afgelopen 5 jaren! Je bent een onderzoeker in hart en nieren en je hebt me ongelofelijk veel geleerd. Jouw visie in het onderzoek en harde werken hebben ervoor gezorgd dat ik in een bevoorrechte positie aan mijn promotie kon beginnen. Naast je passie voor het onderzoek ben je ook nog eens persoonlijk begaan met je 'T-ALL'ers. Iets wat ook weer naar voren kwam op de momenten dat ik je moest vertellen dat ik (weer) zwanger was. Ondanks dat dit onderzoekstechnisch op zijn minst niet handig was, was je oprecht blij voor me. Ook de laagdrempeligheid waarmee je benaderbaar bent, was voor mij heel prettig. Ik kon werkelijk overal met je over praten en discussiëren, maar ook met je lachen (je herinnert je vast wel de avonturen van T-rex in Orlando). Ik ben dankbaar en prijs mezelf gelukkig dat ik bij zo'n top onderzoeker en goed mens mijn promotietraject heb mogen volgen!

Dear Francois Sigaux and Jean Soulier, I would like to thank you very much for the fruitful cooperation on our paper on *MEF2C* and *NKX2-1* and also for the opportunity to visit your hospital in Paris (including the nice café around the corner). Merci beaucoup!

Wouter, Marieke en Petra, heel erg bedankt voor de prettige samenwerking! Wouter, ik heb je heldere uitleg over 4C en je duidelijke en 'recht door zee' manier van communiceren heel erg gewaardeerd. Petra, bedankt voor al je geduld terwijl je mij de 4C techniek hebt aangeleerd. Marieke en Petra, bedankt voor alle keren dat jullie me hielpen als ik weer eens om het hoekje kwam kijken om wat te vragen.

Bedankt Ton en Brenda. Brenda, toen ik bij jou de LM-PCR kwam leren wist ik net hoe ik een pipet vast moest houden. Ook jij had veel geduld met me (moet ik dit dan bovenin, in het midden of onderin het epje pipetteren?), heel erg bedankt! Ton, bedankt voor de prettige samenwerking en de verhelderende gesprekken.

Berna, ik heb je aanwezigheid bij het werkoverleg heel erg gewaardeerd. Wanneer we afdwaalden van de kern van de zaak, bracht jij ons altijd weer terug. Bedankt voor je enthousiasme, je bereidheid op ieder moment te helpen en mee te denken over oplossingen voor problemen (ik heb denk ik nog nooit een 'nee' van je gehoord). En niet te vergeten de gezellige praatjes tussendoor!

Michel, bedankt voor de prettige samenwerking in het onderzoek naar de geneesmiddelen forodesine en ara-G.

Beste Monique, bedankt voor de interesse die je altijd toonde als we elkaar op de gang of in de lift tegenkwamen, en ook voor de keren dat je me zo snel hebt geholpen met vragen tussendoor!

Beste Ronald, ten eerste bedankt voor het beste appeltaartrecept ooit. Maar vooral ook voor je inspirerende vragen en antwoorden! Na jouw presentaties werd mijn wetenschappelijk enthousiasme altijd weer een tandje hoger gezet.

Shanta Bantia, Chantal Meinz, Fiona Higginbotham and Jim MacDonald-Clink, I would like to thank you and your companies (BioCryst & Mundipharma) very much for the opportunity to study forodesine in our laboratory. Shanta Bantia and Cynthia Parker, thank you very much for you efforts in determining dGTP and dGuo levels in all the samples from our lab!

Johan de Rooi, Andrew Stubbs en Peter van der Spek, heel erg bedankt voor de prettige samenwerking!

Beste Anna-Sophie Wiekmeijer, Karin Pike-Overzet en Frank Staal, bedankt voor jullie medewerking bij het onderzoek naar de effecten van MEF2C en NKX2-1 op T-cel ontwikkeling.

Graag wil ik ook alle mensen van de T-ALL groep bedanken! Beste Pieter, toen ik in de T-ALL groep begon was jij al bijna klaar. De eerste weken tijdens het werkoverleg had ik geen idee waar je het over had, daarna werd ik wel een beetje zenuwachtig. Als dit de standaard was had ik wel een probleem. Hoe deed jij toch zoveel werk, had je zoveel goede ideeën, elke maandagochtend weer? Ik denk dat je een groot onderzoeker bent en ben blij met je te hebben samengewerkt. Je liet regelmatig je werk vallen als ik wat kwam vragen en loste het meteen voor me op. Ik wil je ook erg bedanken voor het wetenschappelijk werk dat je hebt gedaan.

Zonder dit werk had dit proefschrift niet bestaan. Martine, we zijn elkaar net misgelopen, maar ook jou wil ik bedanken voor al de betrouwbare data en analyses waar ik op heb kunnen bouwen in het onderzoek. Jessica, met een engelengeduld heb je me de eerste maanden bij de hand genomen en de 'basics' van laboratorium onderzoek geleerd, en meer! Heel erg bedankt voor alles, het inwerken en alle werkzaamheden die je voor de verschillende projecten hebt verricht. Ik wens jou, Paul, Daan en Timo alle geluk! Wilco, bedankt voor je vrolijkheid en positiviteit! En niet te vergeten de honderden PCRs voor het forodesine project. Ik stel voor dat als we met pensioen zijn, we samen met Marcel een nieuw wow guild beginnen. Heb er nu al zin in! (wel Horde hoor). Clarissa, de rust die je uitstraalt tijdens je werk is heel prettig. Bedankt voor alle blots, FISH'es en de rest! Linda, wat was het heerlijk om samen in de trein naar huis de beslommeringen van de dag nog eens door te nemen. We hoefden niets uit te leggen, want we zaten precies in hetzelfde schuitje. Ik heb het erg gemist toen ik naar Rotterdam was verhuisd. Ik wens je heel veel succes met de laatste loodjes en hoop dat we allebei veel mooie ideeën hebben in onze toekomstige wetenschappelijke carrière! Beste Maartje, bedankt voor je gezelligheid en je behulpzaamheid, vooral als ik weer eens array-CGH data op jouw computer wilde bekijken. Monique, bedankt voor de introductie in de wereld van de virussen! Kirsten, je bent een heerlijke collega. Sinds je komst is de gemiddelde tijd dat er per dag gelachen wordt exponentieel toegenomen. Ik vind het jammer maar zo kort te hebben samengewerkt. Bedankt voor de gezelligheid! Laura, je was er kort maar krachtig, bedankt voor de fijne samenwerking! All the new people in the T-ALL group, I wish you all the best!

Beste Mattis en Marije, bedankt dat ik jullie in je afstudeerproject heb mogen begeleiden. Jullie waren beiden erg prettig om mee samen te werken. Ik zat in een luxe positie toen jullie de 4C en ChIP-on-chip methoden konden overnemen. Heel erg bedankt voor al jullie inzet (in weer en wind en sneeuw)! Ik wens jullie beiden een mooie carrière, met wat minder reistijd dan tijdens jullie stage!

Beste Iris, bijna tegelijk begonnen, bijna hetzelfde project, tegelijkertijd zwanger, bijna tegelijkertijd ons eerste kindje, bijna tegelijkertijd de tweede, dezelfde crèche, beide om 17:30 weg om de kinderen te halen, en allemaal een abonnement op Blijdorp. Ik vond het heerlijk om dit alles met je te kunnen delen, onze frustraties, maar ook de leuke dingen! Ook wetenschappelijke ideeën werden op de fiets naar huis besproken. Ik vind je een geweldige wetenschapper, altijd dacht je net een stap verder. Ik ga je missen in Groningen, maar ik hoop dat we elkaar ondanks een druk leven nog vaak zullen spreken en zien!

Beste Diana, je was bijna 5 jaar lang mijn ‘buurvrouw’. Bedankt voor je rust, nuchterheid en humor. Beste Ines, bedankt voor alle spelletjes Pictionary en het lekkere eten! Beste Jill, bedankt voor je oprechte interesse. Heel veel succes en plezier als postdoc! Beste Susan, als ik aan jou denk, denk ik aan Brabantse gezelligheid, bedankt dat ik (en Hans, Max en Henk) hier van hebben mogen genieten met jullie! Kamergenoten Farhad, Floor, Eva, Jenny, Janet, Trudy, Stefanie en Lieke, bedankt voor alle gezelligheid!

Alle analisten, heel erg bedankt voor het draaiende houden van het lab en alle hulp. En alle andere onco’s, we zijn inmiddels te groot geworden om iedereen op te noemen, bedankt voor alle hulp en gezelligheid!

Beste Ingrid, bedankt voor alle regeldingen tussendoor (bv. even ergens stikstof gas regelen), en het samen organiseren van de taart- en koekjeswedstrijden! Ook wil ik graag de mensen van de andere labs bedanken. Theo, bedankt voor alle theorieën en feiten die meestal in de verste verte niets met ons werk te maken hadden, maar daarom niet minder interessant waren! Marcel, Dicky, Janneke, Peng, Marieke, Sylvias, Nanda, Roelien en Ram en ook iedereen die ik niet met name noem, allemaal heel erg bedankt!

Tot slot wil ik mijn familie bedanken. Lieve Nico en Ineke, bedankt dat jullie altijd een luisterend oor hebben, en in alle rust al mijn problemen serieus nemen, maar ook zo goed kunnen relativeren. Mathijs, bedankt voor de hulp bij dit boekje!

Lieve Hans, als ik in het weekend of ’s avonds nog even naar het lab moest kreeg ik zelden een enthousiaste reactie, en dat is eigenlijk maar goed ook, ik ben ook graag bij jou! Bedankt voor al je steun, vooral in de laatste maanden die toch wat meer stress meebrachten dat ik van tevoren had ingeschat. Ook veel dank voor het vermaken van onze lieve ‘mannen’, als ik bijvoorbeeld weer achter de computer zat en je ze soms letterlijk van me af moest plukken. Ik ben blij dat ik je gevonden heb, houd heel veel van jou!

Lieve Max en Henk, bedankt dat jullie er zijn!

# Color figures

INTRODUCTION

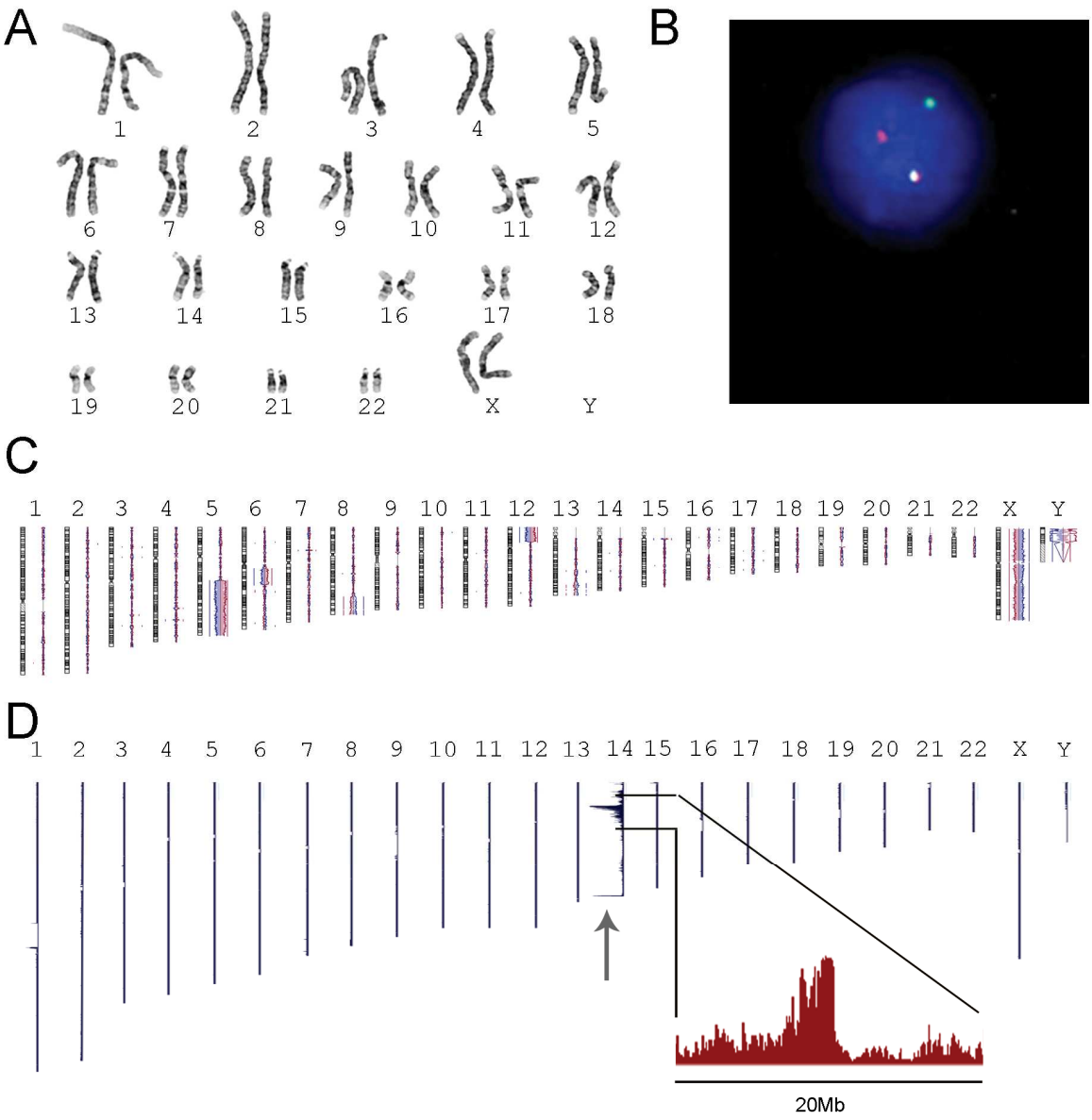



**Figure 2: Results for different types of genetic analyses used in this thesis.** A, Sorted human chromosomes in a karyotypic analysis. Each chromosome has a unique banding pattern. B, FISH analysis showing a red/green fusion signal (normal chromosome) and separate red and green signals for both derivative chromosomes of a translocation. C, ArrayCGH results for all chromosomes for one patient sample. A red signal to the right indicates a deletion (chromosome 5, 6, 12 and Y), a blue signal to the right an amplification (chromosome 8 and X). D, 4C results for all chromosomes for one viewpoint (*NKX2-1*) on chromosome 14 (upper signal and inset). An additional signal is present on the lower tip (grey arrow) of chromosome 14 indicating an inversion of chromosome 14 (in this case between *NKX2-1* and the *IgH@* locus).

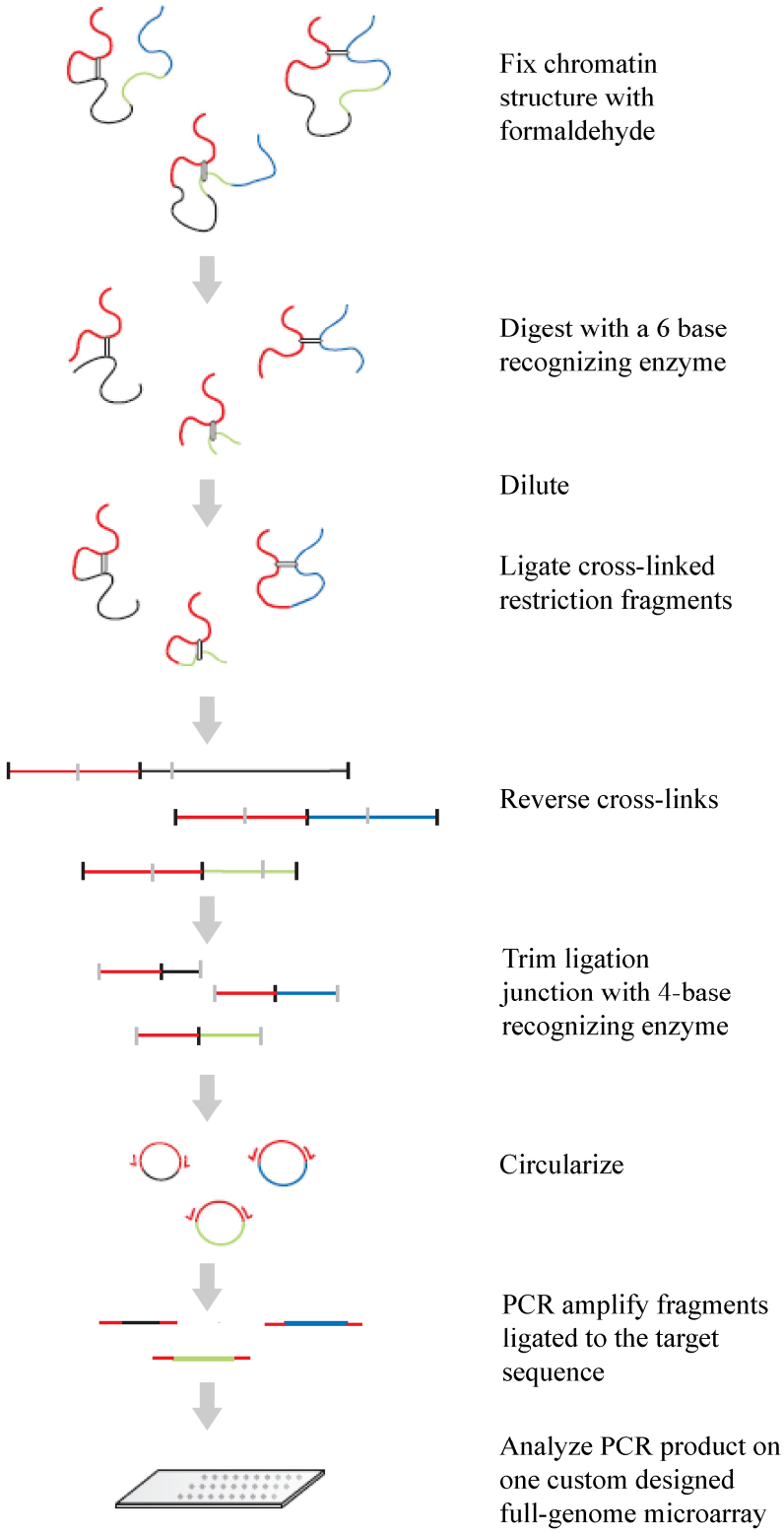



**Figure 3: Outline of 4C-technology.**

# CHAPTER 1

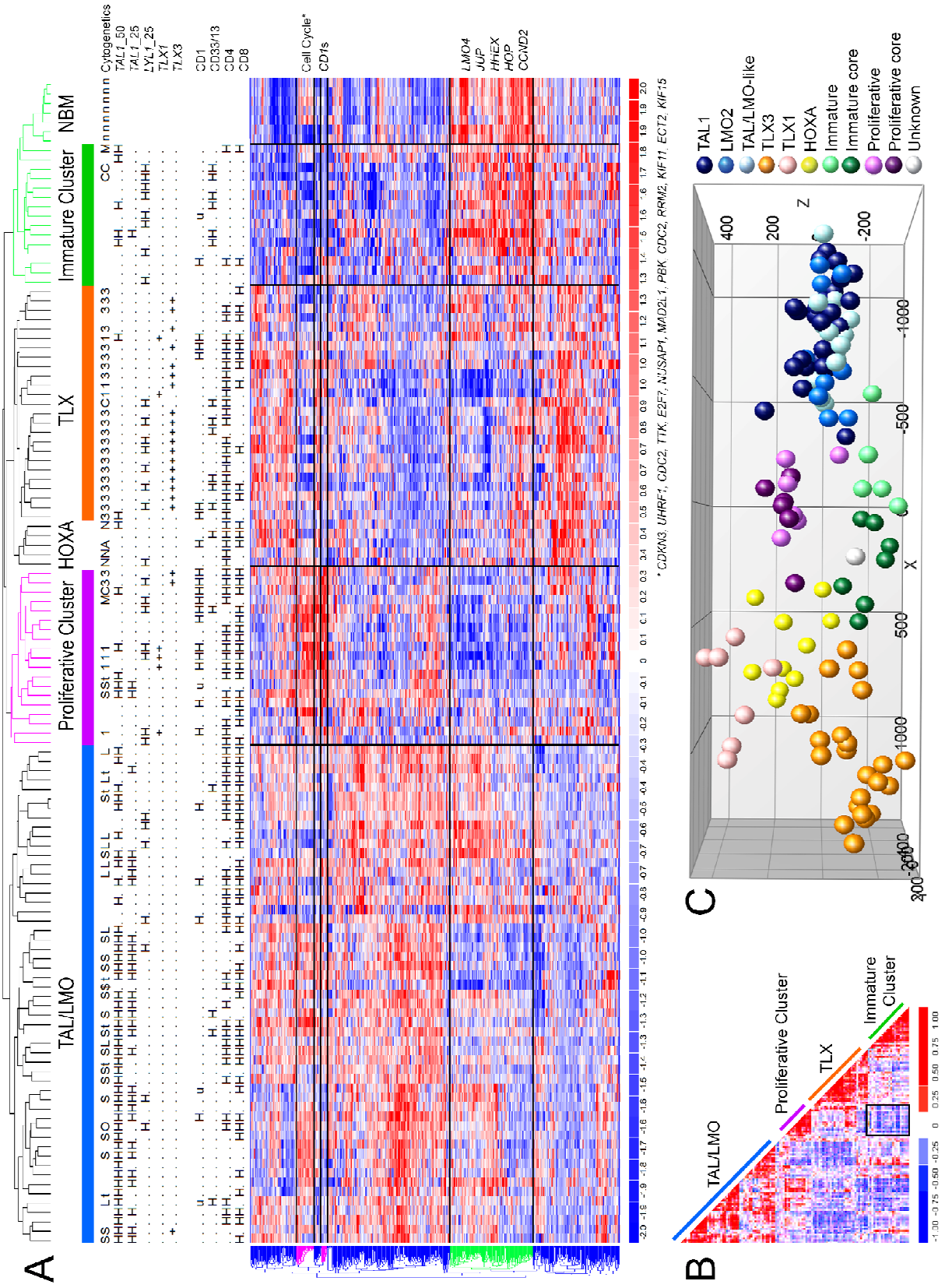

A
TAL/LMO
Proliferative Cluster
HOXA
TLX
Immature Cluster
NBM
Cytogenetics
TAL1_50
TAL1_25
LYL1_25
TLX1
TLX3
CD1
CD33/13
CD4
CD8
Cell Cycle*
CD1s
LMO4
JUP
HHEX
HOP
CCND2
*CDKN3, UHRF1, CDC2, TTK, E2F7, NUSAP1, MAD2L1, PBK, CDC2, RRM2, KIF11, ECT2, KIF15
B
TAL/LMO
Proliferative Cluster
TLX
Immature Cluster
C
TAL1
LMO2
TAL/LMO-like
TLX3
TLX1
HOXA
Immature
Immature core
Proliferative
Proliferative core
Unknown
X
Z

**Figure 1. (page 200) Identification of 2 Entities in Pediatric T-ALL That Lack Known Driving Oncogenic Hits.** (**A**) Unsupervised hierarchical cluster analysis by the average linkage method in dCHIP based on 435 probesets (**Table S3**) for RMA-solo (Soulier et al., 2005) normalized U133 plus 2 Affymetrix data from 117 pediatric T-ALL samples and 7 normal bone marrow controls. Cytogenetic rearrangements indicated are: S, *SIL-TAL1*; T, *TAL1*; t, *TAL2*; O, *LMO1*; L, *LMO2* (includes del(11)(p12p13)); $, *TAL2*/*LMO1*; N, *SET-NUP214*; C, *CALM-AF10*; M, *MYB*; A, Inv(7)(p15q34); 1, *TLX1*; 3, *TLX3*; n, normal bone marrow controls. The $50^{th}$ and/or the $25^{th}$ percentiles of samples with the highest *TAL1* or *LYL1* expression, positivity for *TLX1* and *TLX3* expression as measured by RQ-PCR, and expression of the immunophenotypic markers CD13 and/or CD33, CD4 or CD8 are indicate; u, no data available. (**B**) Pearson correlation plot for the patient samples belonging to the 4 unsupervised *TAL/LMO*, *TLX*, proliferative and immature clusters. (**C**) Principal component analysis of pediatric T-ALL patients based upon the top 100 most significant differentially expressed probesets among major T-ALL subgroups (i.e. *TAL1*/*LMO2*, *HOXA*, *TLX1*, and *TLX3* (**Table S3**)). The immature cluster (12 cases) and the proliferative cluster (12 cases) are indicated by green and purple dots, respectively. Samples repeatedly assigned to the proliferative or immature clusters (i.e. the core samples) in multiple unsupervised analyses on RMA-solo (**Figure 1A**), RMA or VSN normalized datasets (not shown) or the supervised cluster analysis (**Figure 1C**) are visualized by dark green or purple dots. See also **Figure S1** and **Tables S1-S4**.

**Figure 3. (page 202) *NKX2-1* and *NKX2-2* Rearrangements in Proliferative Cluster Patient Samples.** (**A**) 4C-results obtained from *NKX2-1* or *TRB@* viewpoints (VP). Position of *TRA@*, *NKX2-1* and *IGH@* loci are shown by grey vertical bars. 4C-results for a normal control are shown in grey. Higher magnifications of the reciprocal breakpoint regions are given in **Figure S3**. (**B**) Validation of *NKX2-1* rearrangements by FISH. Schematic positions of FISH probes are shown. (**C**) Schematic representation of the der(7) breakpoint region and breakpoint sequence of the unbalanced t(7;14)(q34;q13) for patient #9989. (**D**) Visualization of a single copy *NKX2-1* amplification (green box) in patient #2702 as identified by array-CGH. (**E**) Schematic representation of t(14;20)(q11;20p11) breakpoint regions and cloned breakpoint sequences for the *NKX2-2* rearranged patient #10138. (**F**) NKX2-1 protein expression in representative proliferative cluster and immature cluster patient samples as shown by western blot. Actin was used as loading control.

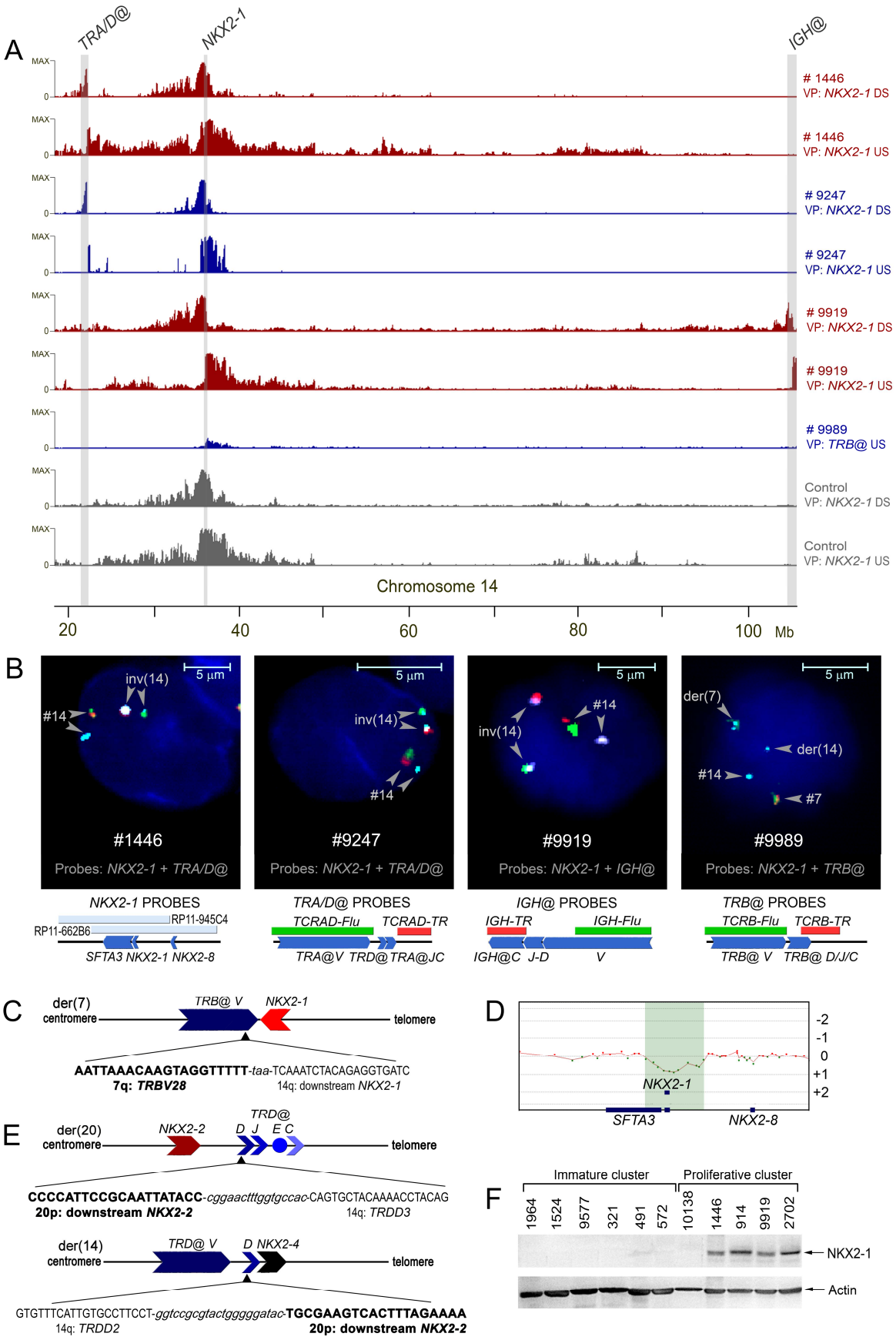
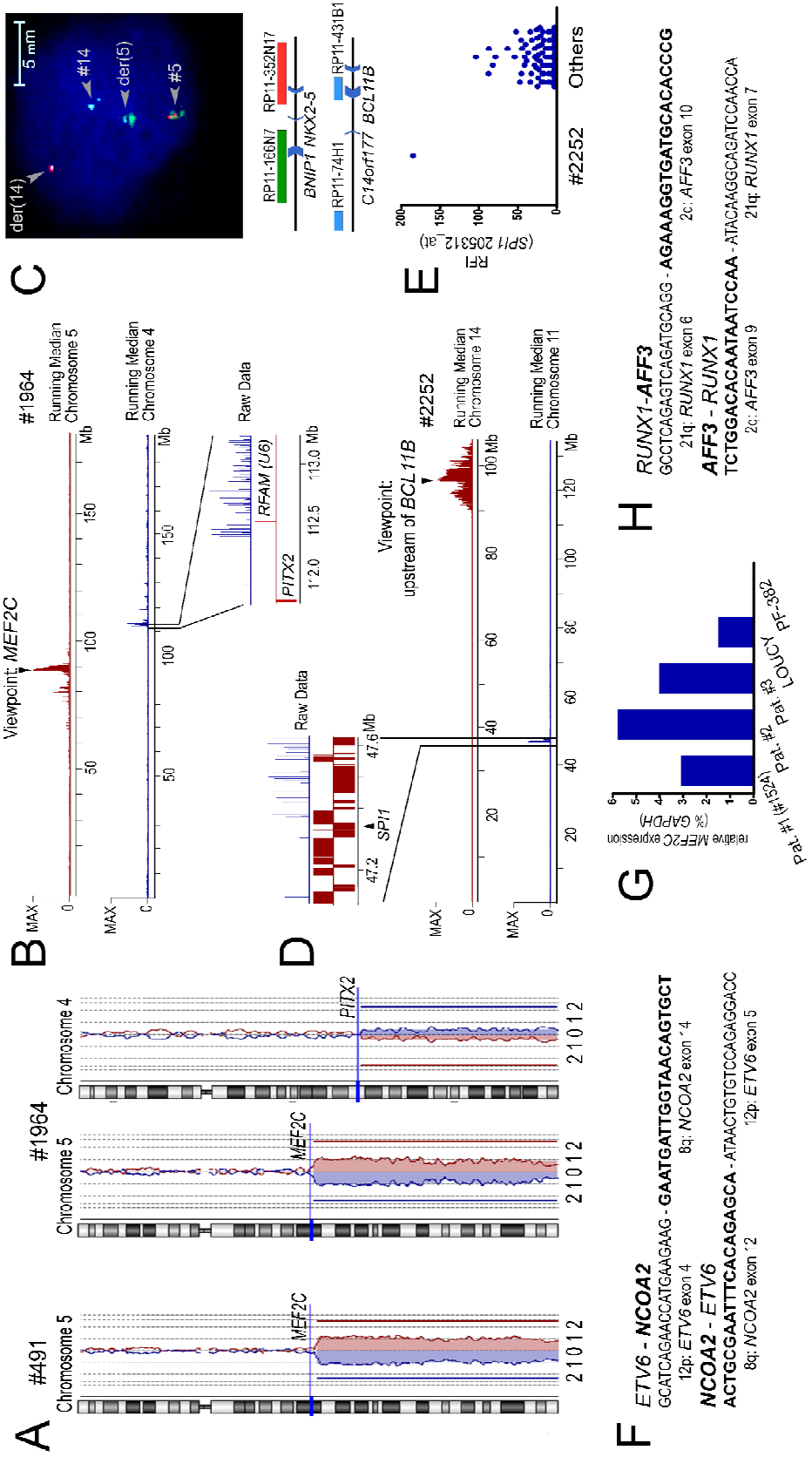

A
TRA/D@
NKX2-1
IGH@
# 1446 VP: NKX2-1 DS
# 1446 VP: NKX2-1 US
# 9247 VP: NKX2-1 DS
# 9247 VP: NKX2-1 US
# 9919 VP: NKX2-1 DS
# 9919 VP: NKX2-1 US
# 9989 VP: TRB@ US
Control VP: NKX2-1 DS
Control VP: NKX2-1 US
Chromosome 14
20
40
60
80
100 Mb
B
inv(14)
#14
5 µm
#1446
Probes: NKX2-1 + TRA/D@
#9247
Probes: NKX2-1 + TRA/D@
#9919
Probes: NKX2-1 + IGH@
der(7)
der(14)
#7
#9989
Probes: NKX2-1 + TRB@
NKX2-1 PROBES
RP11-662B6
RP11-945C4
SFTA3 NKX2-1 NKX2-8
TRA/D@ PROBES
TCRAD-Flu
TCRAD-TR
TRA@V TRD@ TRA@JC
IGH@ PROBES
IGH-TR
IGH-Flu
IGH@C J-D V
TRB@ PROBES
TCRB-Flu
TCRB-TR
TRB@ V TRB@ D/J/C
C
der(7)
centromere
TRB@ V
NKX2-1
telomere
AATTAAACAAGTAGGTTTTT-taa-TCAAATCTACAGAGGTGATC
7q: TRBV28
14q: downstream NKX2-1
D
-2
-1
0
+1
+2
NKX2-1
SFTA3
NKX2-8
E
der(20)
centromere
NKX2-2
TRD@
D J E C
telomere
CCCCATTCCGCAATTATACC-cggaactttggtgccac-CAGTGCTACAAAACCTACAG
20p: downstream NKX2-2
14q: TRDD3
der(14)
centromere
TRD@ V
D NKX2-4
telomere
GTGTTTCATTGTGCCTTCCT-ggtccgcgtactgggggatac-TGCGAAGTCACTTTAGAAAA
14q: TRDD2
20p: downstream NKX2-2
F
Immature cluster
Proliferative cluster
1964
1524
9577
321
491
572
10138
1446
914
9919
2702
NKX2-1
Actin

**Figure 4. (page 204) *MEF2C* Activating Rearrangements for Immature Cluster Samples.** (**A**) Array-CGH results for chromosomes 4 and/or 5 for patients #491 and #1964. Blue and red tracings represent dye swopped experiments. Positions of *MEF2C* and *PITX2* have been indicated. (**B**) Visualization of an unbalanced chromosomal translocation t(4;5)(q26;q14) for patient #1964 by 4C-analysis. The *MEF2C* VP is indicated by an arrow. Running median of probeset intensities for chromosome 5 and 4 are indicated in red and blue, respectively. (**C**) Validation of a chromosomal translocation between *NKX2-5* and *BCL11B* in patient #9577 by FISH. Schematic positions of FISH probes are shown. (**D**) Identification of the t(11;14)(p11.2;q32.2) chromosomal translocation between *SPI1* and *BCL11B* in patient #2252 by 4C. The VP is positioned ~0.6 Mb upstream of *BCL11*B as indicated by an arrow. (**E**) Ectopic *SPI1* expression in patient #2252 compared to 116 additional T-ALL patient samples. Raw fluorescent intensities of probeset 205312_at are shown. (**F**) Cloned fusion areas for reciprocal *ETV6-NCOA2* and *NCOA2-ETV6* fusion transcripts in patient #1524. (**G**) Relative *MEF2C* expression by RQ-PCR in 3 selected *ETV6-NCOA2* rearranged T-ALL patients (Pat. #1-3). Cell lines LOUCY and PF382 are positive and negative controls for *MEF2C* expression, respectively. (**H**) Cloned fusion areas for reciprocal *RUNX1-AFF3* and *AFF3-RUNX1* fusion transcripts for patient #572. See also **Figure S4**.


A
#491
Chromosome 5
MEF2C
2 1 0 1 2
#1964
Chromosome 5
MEF2C
Chromosome 4
PITX2
B
#1964
Viewpoint: MEF2C
Running Median Chromosome 5
Running Median Chromosome 4
Raw Data
RFAM (U6)
PITX2
C
der(14)
5 mm
#14
der(5)
#5
RP11-166N7
RP11-352N17
BNIP1 NKX2-5
RP11-74H1
RP11-431B1
C14orf177 BCL11B
D
Raw Data
47.6 Mb
SPI1
47.2
Viewpoint: upstream of BCL11B
#2252
Running Median Chromosome 14
Running Median Chromosome 11
E
#2252
Others
RT (SPI1 205312_at)
F
ETV6 - NCOA2
GCATCAGAACCATGAAGAAG - GAATGATTGGTAACAGTGCT
12p: ETV6 exon 4
8q: NCOA2 exon 14
NCOA2 - ETV6
ACTGCGAATTTCACAGAGCA - ATAACTGTGTCCAGAGGACC
8q: NCOA2 exon 12
12p: ETV6 exon 5
G
relative MEF2C expression (% GAPDH)
Pat. #1 (#1524)
Pat. #2
Pat. #3
LOUCY
PF-382
H
RUNX1-AFF3
GCCTCAGAGTCAGATGCAGG - AGAAAGGTGATGCACACCCG
21q: RUNX1 exon 6
2c: AFF3 exon 10
AFF3 - RUNX1
TCTGGACACAATAATCCAA - ATACAAGGCAGATCCAACCA
2c: AFF3 exon 9
21q: RUNX1 exon 7

## CHAPTER 2

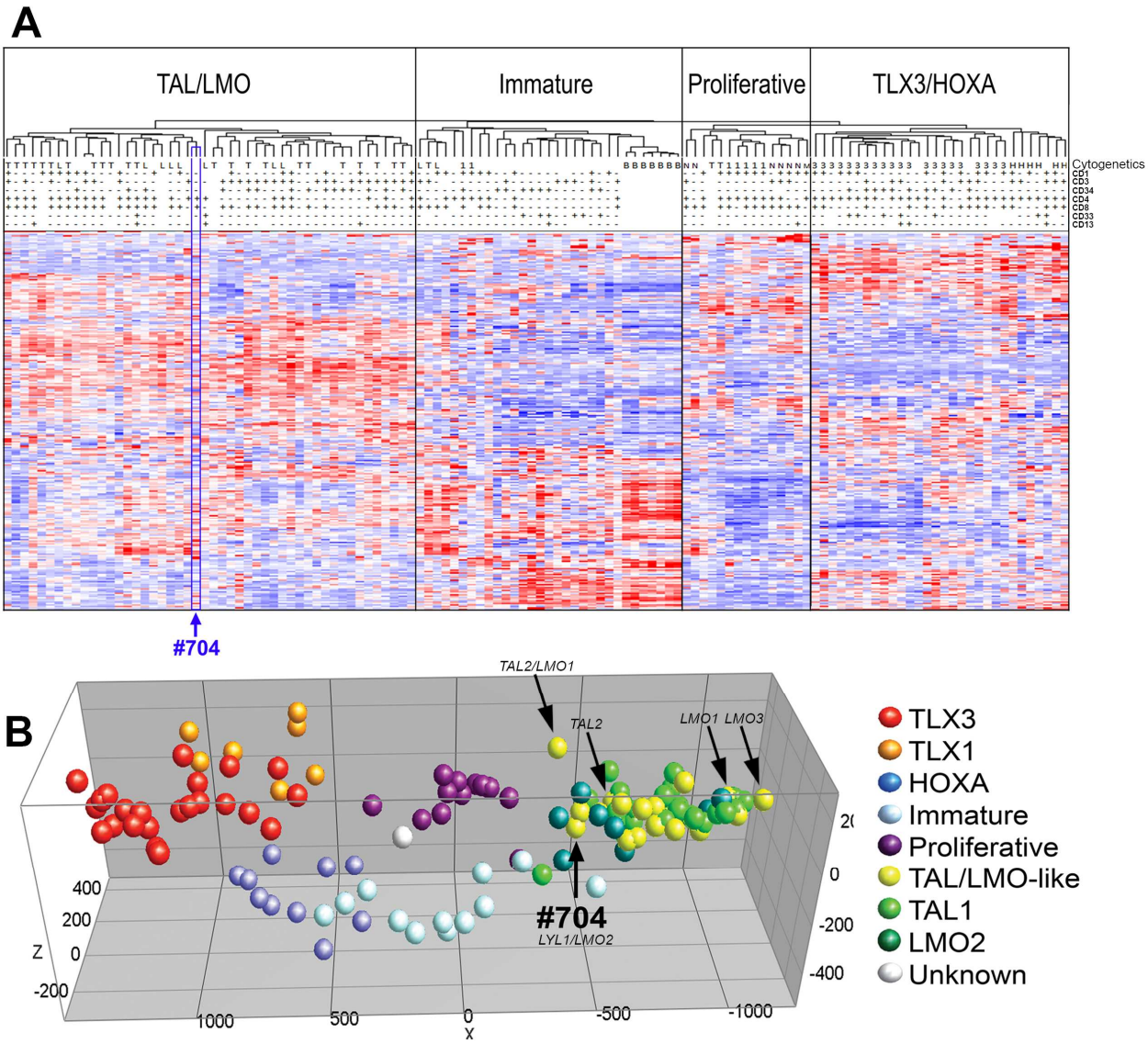



**Figure 1. Unsupervised and supervised hierarchical clustering of 117 pediatric T-ALL samples and 7 normal bone marrow samples and *LYL1* expression of unsupervised subgroups.** (A) Unsupervised hierarchical clustering of 117 pediatric T-ALL samples and 7 normal bone marrow samples (horizontal axis), according to microarray gene-expression (genes on vertical axis, gene names not shown)(1). Red corresponds to high expression, blue to low expression. CD surface markers are shown as present (>25%, "+"), absent (<25% "-") or not performed (blanc). Complete immunophenotype for #704: CD1-, CD2+, CD3-, CD4+, CD5-, CD7+, CD8+, cytoplasmatic CD3+, CD33-, CD14-, CD34-, CD71+, HLA_DR-, TDT+. Cytogentic abnormalities are annotated as follows: T: *SIL-TAL* deletion or *TAL1* translocation, L: *LMO2* translocation/deletion, 1: *TLX1* translocation, 3: *TLX3* translocation, B: normal bone marrow, N: *NKX2-1* translocation/inversion/duplication, M: *MYB* translocation, H: *HOXA* activating aberration (*CALM-AF10, SET-NUP, HOXA* inversion). Patient #704 is highlighted by a blue box. (B) Principal component analysis of supervised analyses of gene-expression data of 117 pediatric T-ALL samples(1). The position of the yellow dots representing *LMO1, TAL2, LMO3, TAL2/LMO1* rearranged cases and sample #704 (LYL1/LMO2) are indicated by arrows.

## CHAPTER 3

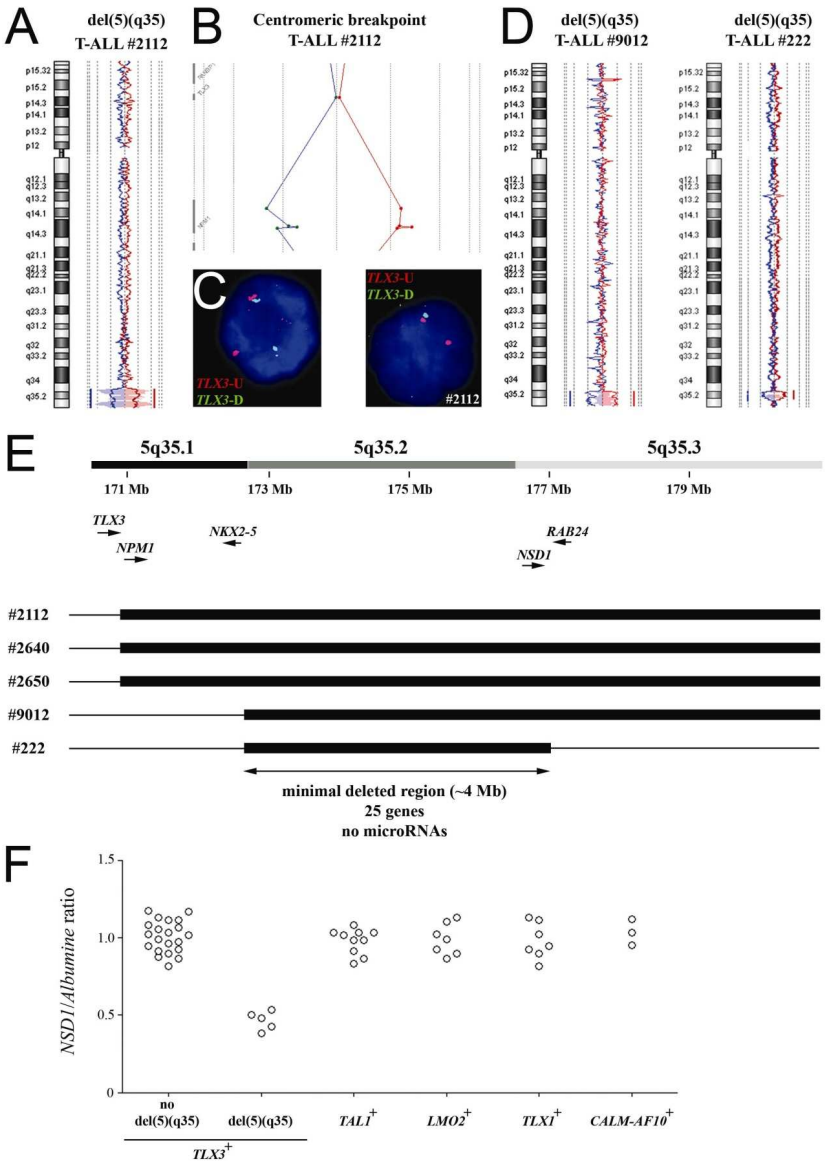



**Figure 1. The recurrent cryptic deletion, del(5)(q35), in *TLX3* rearranged pediatric T-cell acute lymphoblastic leukemia (T-ALL).** (**a**) Chromosome 5 ideogram and corresponding oligo microarray-based comparative genome hybridization (array-CGH) plot of case DNA:control DNA ratios (blue tracing) versus the dye-swap experiment (red tracing) for T-ALL cases 2112. Hybridization signals around the -2X or +2X lines represent loss of the corresponding region in the case DNA. (**b**) Detailed analysis of the centromeric breakpoint of the deletion in case 2112. (**c**) Dual-color fluorescence *in situ* hybridization (FISH) analysis on interphase cells of case 9858 (left panel) and case 2640 (right panel) using the *TLX3-U* (Red) and *TLX3-D* (green) translocation probe set. Case 9858 showed a split signal, indicative for a *TLX3* translocation, whereas case 2640 showed loss of the *TLX3-D* (green) signal. (**d**) Similar chromosome 5 ideograms as in (**a**) for T-ALL cases 9012 and 222. (**e**) Schematic overview of the minimal deleted region on chromosomal band 5q35 for the 5 TLX3

rearranged T-ALL cases showing a del(5)(q35). Depicted genome positions and gene locations are based on the UCSC Genome Browser at http://genome.ucsc.edu/. (**f**) Quantitative PCR analysis of *NSD1*, present in the minimal deleted region, on 26 TLX3 rearranged T-ALL cases and 27 TLX3 negative cases.

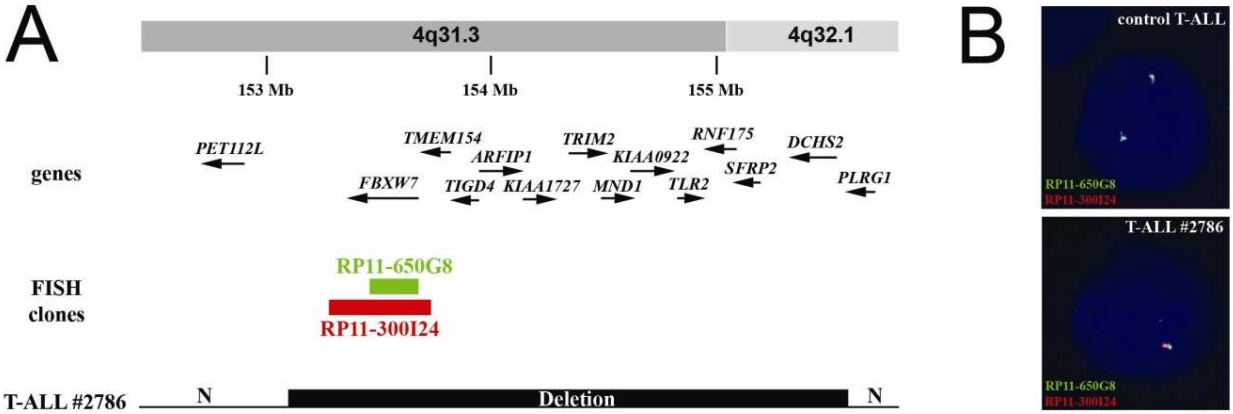
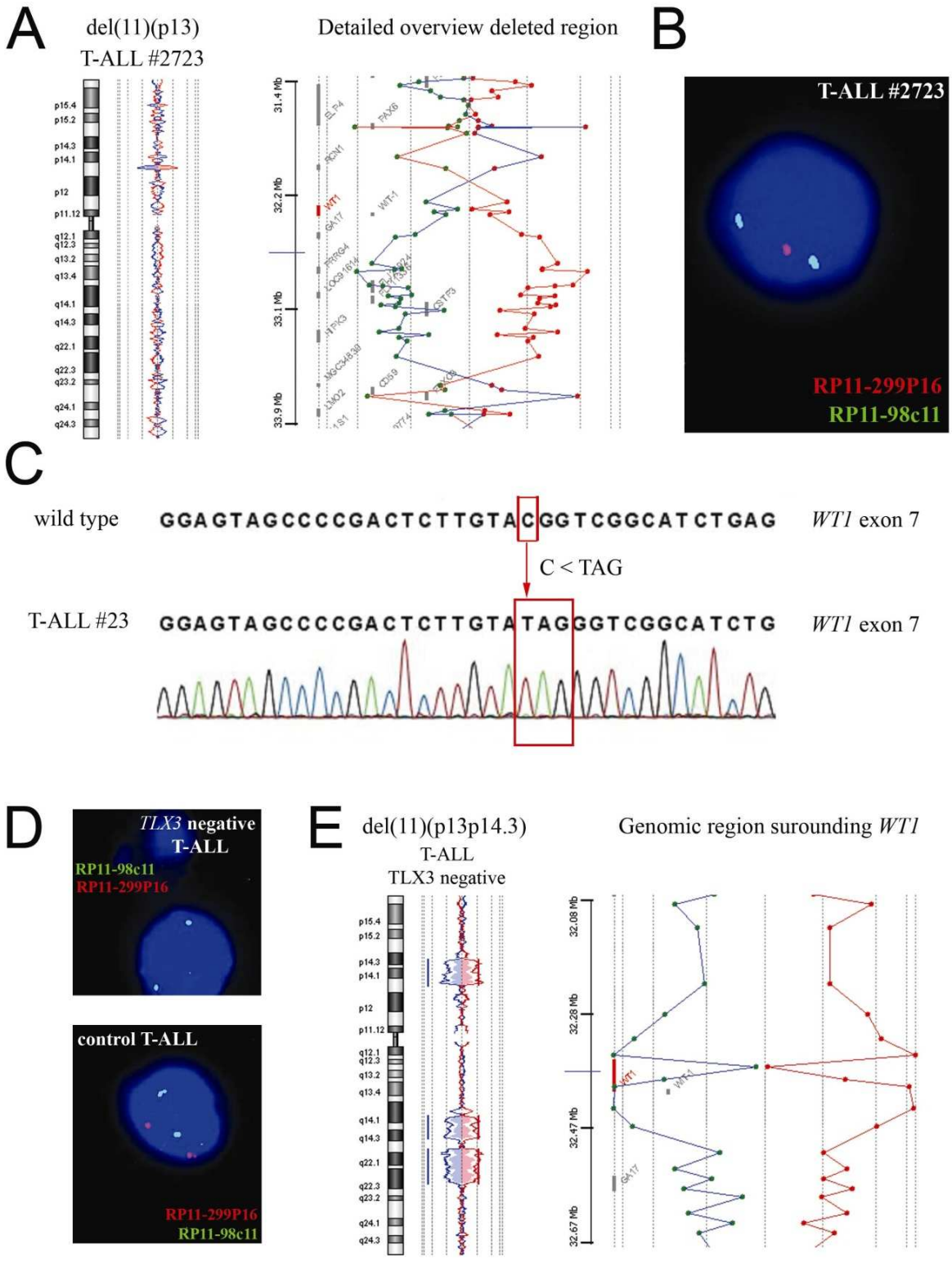



**Figure 2. *FBXW7* deletion in pediatric T-cell acute lymphoblastic leukemia (T-ALL).** (**a**) Schematic overview of the chromosomal deletion, del(4)(q31.3q32.1), as detected in case 2786. Genomic positions of genes situated in this chromosomal region and bacterial artificial chromosome (BAC) clones used for fluorescence *in situ* hybridization (FISH) analysis are depicted. (**b**) FISH analysis using RP11-650G8 (green) and RP11-300I24 (red) confirms the presence of the del(4)(q31.3q32.1) in case 2786.

**Figure 3. (page 208) Wilms' tumor 1 (*WT1*) inactivation in pediatric T-cell acute lymphoblastic leukemia (T-ALL).** (**a**) Chromosome 11 ideogram and oligo microarray-based comparative genome hybridization (array-CGH) plot for the deletion, del(11)(p13), as detected in case 2723 (left panel). The right panel shows a detailed overview of the deleted region for this 11p13 deletion. (**b**) Fluorescence *in situ* hybridization (FISH) analysis using RP11-98C11 (green) and RP11-299P16 (red, covering *WT1*) confirms the presence of the del(11)(p13) in case 2723. (**c**) Sequence analysis shows a truncating *WT1* exon 7 mutation on the remaining allele of case 2723. (**d**) Similar FISH analysis as in (**b**) on TLX3 wild-type T-ALL cases identified one additional case showing a biallelic *WT1* deletion. (**e**) Array-CGH analysis confirmed the presence of a large monoallelic deletion, del(11)(p13p14.3), in combination with an additional loss of the genomic region surrounding the *WT1* gene on the other allele.


A
del(11)(p13)
T-ALL #2723
Detailed overview deleted region
B
T-ALL #2723
RP11-299P16
RP11-98c11
C
wild type
GGAGTAGCCCCGACTCTTGTACGGTCGGCATCTGAG
WT1 exon 7
C < TAG
T-ALL #23
GGAGTAGCCCCGACTCTTGTATAGGGTCGGCATCTG
WT1 exon 7
D
TLX3 negative
T-ALL
RP11-98c11
RP11-299P16
control T-ALL
RP11-299P16
RP11-98c11
E
del(11)(p13p14.3)
T-ALL
TLX3 negative
Genomic region surounding WT1

CHAPTER 5

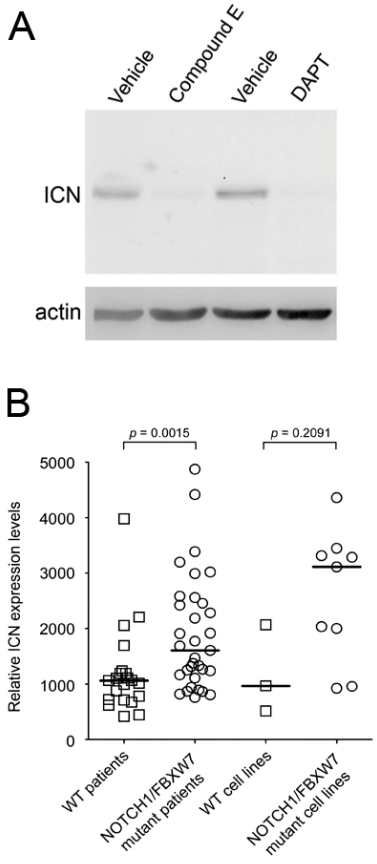
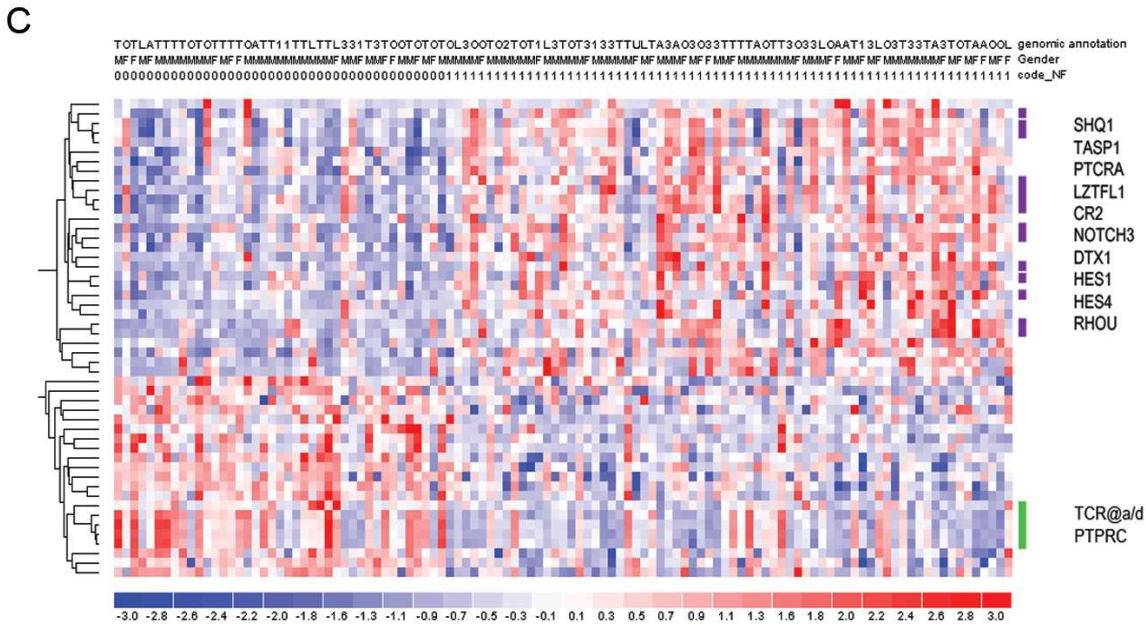



**Figure 2. NOTCH1/FBXW7 mutations activate the NOTCH1 pathway in primary T-ALL patient biopsies.** (A) Western blot analysis of lysates from the HPBALL T-ALL cell line which is NOTCH1-mutated. Treatment for 96 hrs with γ-secretase inhibitors including compound E (100 nM) or DAPT (5 μM), results in loss of activated intracellular NOTCH1 expression (ICN). Actin was used as loading control. (B) NOTCH1 ICN levels in wild-type and NOTCH1 and/or FBXW7-mutated T-ALL patients and T-ALL cell lines analyzed with Reverse-phase Protein microarray. NOTCH1/FBXW7 wild-type patient samples with high ICN levels are marked by an asterisk (C) Heatmap showing the TOP50 most differentially expressed genes between NOTCH1 and/or FBXW7 mutant patients versus wild-type patients. NOTCH1 direct target genes are indicated. Annotations indicated are genetic rearrangements, Gender and NOTCH1/FBXW7 mutation status. Genetic rearrangements indicated are: T, TAL1 or SIL-TAL1; L, LMO1 or LMO2 (includes del(11)(p12p13)); A, HOXA-activated (includes cases with SET-NUP214; CALM-AF10 or Inv(7)(p15q34)); 1, TLX1; 2, TLX2; 3, TLX3; O, Other; U, Aberration unknown. Gender is indicated F, Female or M, Male. NOTCH1/FBXW7 mutation status is indicated 0, wild-type and 1, NOTCH1 and/or FBXW7-mutated; NOTCH1/FBXW7 wild-type patients with high ICN levels are marked by an asterisk; NOTCH1/FBXW7 wild-type patients having a NOTCH1 signature that cluster with NOTCH1-activated patients based upon hierarchical clustering based on the TOP50 probeset are indicated with a filled triangle.



# Supplementary Data

# CHAPTER 1

**(Supplementary Tables S2-S5: online available.)**

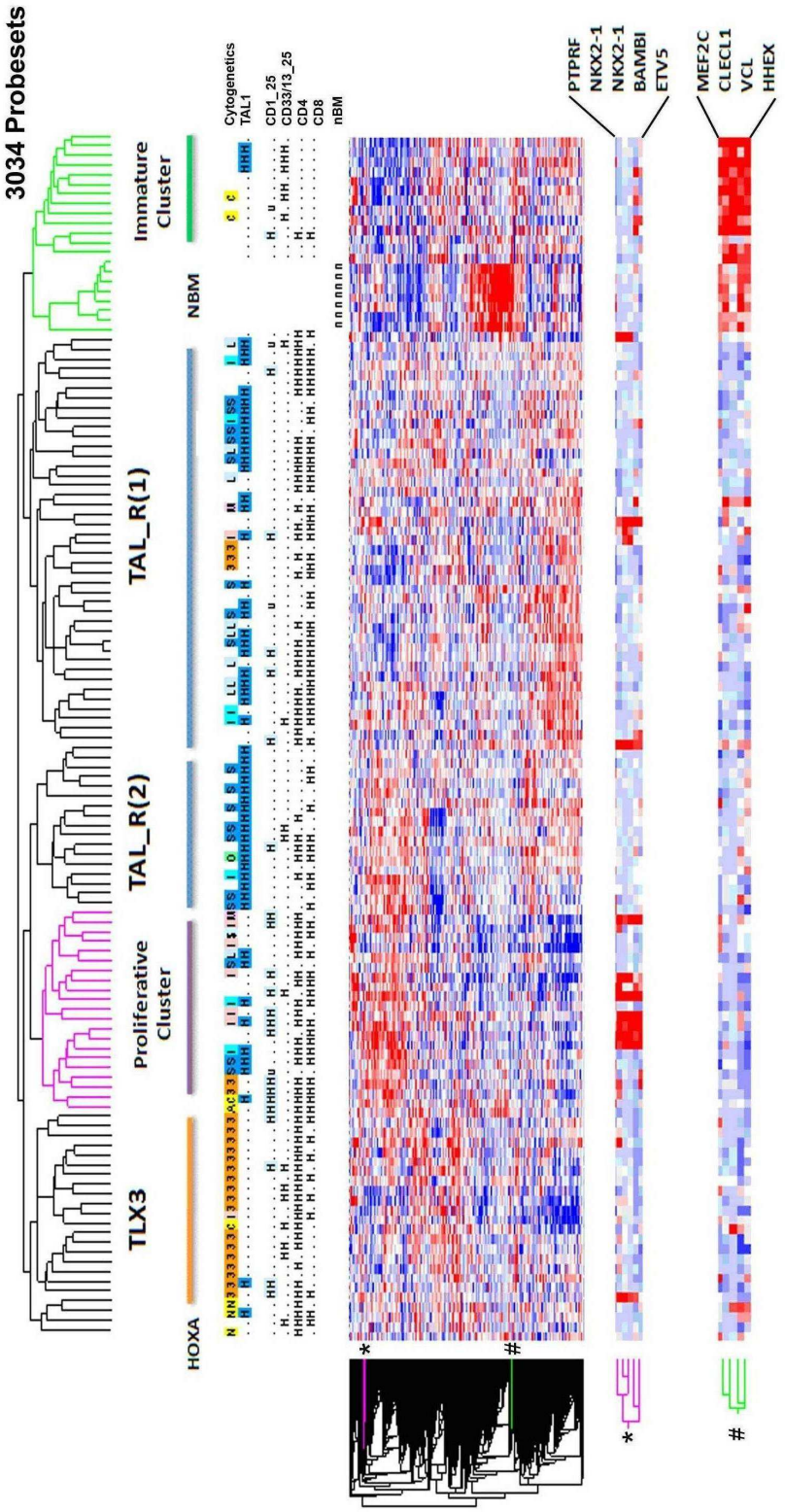



**Figure S1 Relates to Figure 1**. Unsupervised clustering of 117 pediatric T-ALL patient biopsies and 7 normal bone marrow controls on RMA-solo normalized data, using CV range of 0.9-10 resulting in 3034 probesets. Genetic rearrangements indicated are: S, SIL-TAL1; T, TAL1; t, TAL2; O, LMO1; L, LMO2 (includes del(11)(p12p13)); $, TAL2/LMO1; N, SET-NUP214; C, CALM-AF10; M, MYB; A, Inv(7)(p15q34); 1, TLX1; 3, TLX3; n, normal bone marrow controls. The 50 and 25 percentiles of samples with highest TAL1 and CD1 expression as measured by RQ-PCR, respectively, and samples positive for CD13 and/or CD33, CD4 or CD8 are indicated; u, no data available. Insets refer to probesets highly expressed in the proliferative cluster (*) or the immature cluster (#).

Table S1 **Summary of patient samples clustering in the various cluster analyses and stability testing.**

| Sample name | Clustering (unsupervised selection of probsets) # unique genes | 435 filtered | 581 unfiltered | 889 unfiltered | 1343 unfiltered | 3034 unfiltered | Class prediction | | Clustering (supervised) | *NKX2-1* expression | *MEF2C* expression | Cytogenetic annotation | Core samples without previously identified cytogentic alteration |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Relates to: | | **Fig 1A Table S3** | **Table S3** | | | | **Table S2 (Analysis 1)** | **Table S2 (Analysis 2)** | **Fig 1C Table S3** | | | | |
| 167 | | I | I | I | I | I | I | I | I | | + | | I |
| 321 | | I | I | I | I | I | I | I | I | | + | | I |
| 491 | | I | I | I | I | I | I | I | I | | + | | I |
| 572 | | I | I | I | I | I | I | I | I | | + | | I |
| 1524 | | I | I | I | I | I | I | I | I | | + | | I |
| 1964 | | I | I | I | I | I | I | I | I | | + | | I |
| 10030 | | I | I | I | I | I | I | I | I | | + | | I |
| 1509 | | I | I | I | I | I | I | I | | | + | CALM-AF10 | |
| 2736 | | I | I | I | I | I | I | I | | | + | CALM-AF10 | |
| 9194 | | I | I | I | I | I | I | I | | | + | | |
| 2130 | | I | I | I | I | I | I | I | I | | | | |
| 2252 | | I | I | I | I | I | I | | I | | + | | |
| 2703 | | I | I | I | I | I | I | | I | | + | | |
| 9577 | | I | I | I | I | | I | | I | | + | | |
| 9105 | | I | I | I | I | | | | P | | | MYB | |
| 9226 | | | | | | | | | I | | | | |
| 914 | | P | P | P | P | P | P | P | P | + | | MYB | |
| 1446 | | P | P | P | P | P | P | P | P | + | | | P |
| 2702 | | P | P | P | P | P | P | P | P | + | | | P |
| 9247 | | P | P | P | P | P | P | P | P | + | | | P |
| 10138 | | P | P | P | P | P | P | P | P | +* | | | P |
| 9696 | | P | P | P | P | | | | P | | | | |
| 750 | | P | P | P | P | P | P | P | | | | SIL_TAL | |
| 2669 | | P | P | P | P | P | P | P | | | | SIL_TAL | |
| 8628 | | P | P | P | P | P | P | P | | | | TAL | |
| 9919 | | P | P | P | P | | P | | P | + | | | P |
| 9989 | | P | P | P | P | | P | P | P | + | | | P |
| 2113 | | P | P | | | | | | P | + | | | |
| 2691 | | P | P | P | P | P | P | P | | | | TLX3 | |
| 2737 | | P | P | P | P | P | P | P | | | | TLX1 | |
| 3028 | | P | P | P | P | P | P | P | | | | TLX3 | |
| 3044 | | P | | P | P | P | P | P | | | | CALM-AF10 | |
| 2781 | | P | P | P | P | P | P | P | | + | | TLX1 | |
| 2229 | | P | P | P | P | P | P | P | | + | | TLX1 | |
| 2792 | | P | | P | P | P | | | | + | | TLX1 | |
| 2641 | | | | | | | | | P | | | | |
| 9827 | | | | | | | | | P | | | | |

**NKX2-2* expression

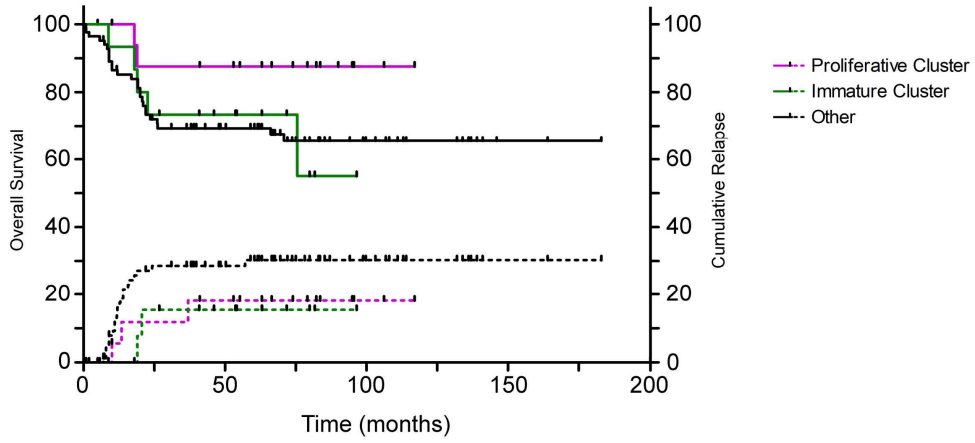



**Figure S2** Relates to Table 1. The immature cluster is not associated with poor survival. The overall survival curves (solid lines, scale indicated at the left axis) and cumulative relapse incidence (dotted lines, scale indicated at the right axis) for the unsupervised proliferative cluster (purple lines), the immature cluster (green lines) as well as for all other T-ALL cases (black lines). The survival of the proliferative cluster that has a 5yr OS of 88±8% seems higher compared to both other subgroups (5yr OS for the immature cluster = 73±11% and the 5yr OS for the remainder of the T-ALL cohort = 65±6%), but does not reach statistical significance (*P*=0.096).

**Table S6. Rearrangement status of various loci for immature cluster or proliferative cluster samples**

| Immature cluster | *TRB@* | *TRAD@* | *MYB* | *BCL11B* | *NKX2-1* | *LYL1* | *TEL* | *TEL/AML1* | *NCOA2* | *NKX2-5* | *MEF2C* | *FOXH1* | *HHEX* | *MN1* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **10030*** | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | | **GL** | GL/GL | GL/GL | | | | | GL/GL |
| **2703** | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | | GL/GL | GL/GL | GL/GL | GL/GL |
| **2130** | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | | GL/GL | GL/GL | GL/GL | GL/GL |
| **2252** | GL/GL | **GL/R** | GL/GL | GL/GL/**R** | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | | GL/GL | GL/GL | GL/GL | GL/GL |
| **167*** | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | | GL/GL | | GL/GL | GL/GL | GL/GL | GL/GL |
| **321*** | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | | **GL** | GL/GL | **GL/GL/GL** | GL/GL |
| **491*** | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | **GL** | GL/**R** | GL/GL | | GL/GL | **GL/GL/GL** | GL/GL | GL/GL |
| **572*** | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/**TR** | GL/GL | | GL/GL | | GL/GL | GL/GL |
| **1524*** | GL/GL | GL/GL | **GL/GL/(GL)** | | GL/GL | GL/GL | **GL** | | GL/**TR** | | | GL/GL | | GL/GL |
| **1964*** | GL/GL | GL/GL | GL/GL | | GL/GL | | | | GL/GL | | | | | GL/GL |
| **9577** | GL/GL | GL/GL | GL/GL | GL/**TR** | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/**TR** | GL/GL | GL/GL | GL/GL | GL/GL |
| **9226** | GL/GL | GL/GL | **GL/GL/(GL)** | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | | GL/GL | GL/GL | GL/GL | GL/GL |

| Proliferative cluster | *TRB@* | *TRAD@* | *MYB* | *BCL11B* | *NKX2-1* | *NKX2-2* |
|---|---|---|---|---|---|---|
| **9919**** | GL/GL | GL/GL | GL/GL | GL/GL | **GL/R** | GL/GL |
| **9247**** | GL/GL | **GL/R** | GL/GL | GL/GL | **GL/R** | GL/GL |
| **10138**** | GL/GL | **GL/TR** | GL/GL | | GL/GL | **GL/TR** |
| **914** | **GL/TR** | GL/GL | **GL/TR** | | GL/GL | GL/GL |
| **2113** | GL/GL | GL/GL | | | GL/GL | GL/GL |
| **2641** | GL/GL | **GL/GL/R** | GL/GL | **GL/GL/R** | **GL/GL/R** | GL/GL |
| **9989**** | **GL/R** | GL/GL | GL/GL | | **GL/GL/R** | GL/GL |
| **2702**** | GL/GL | **R** | GL/GL | **GL/TR** | GL/GL | GL/GL |
| **1446**** | GL/GL | **GL/R** | GL/GL | GL/GL | **GL/R** | GL/GL |
| **9105** | **GL/GL/GL/R** | **GL** | **R/TR** | | **GL** | GL/GL |
| **9696** | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL | GL/GL |
| **9827** | GL/GL | GL/GL | GL/GL | | GL/GL | GL/GL |

Relates to Table 2. *immature cluster core samples; **proliferative cluster core samples; GL, Germline; (GL), a subclone had 1 additional germline chromosome; TR, translocated; R, rearranged.

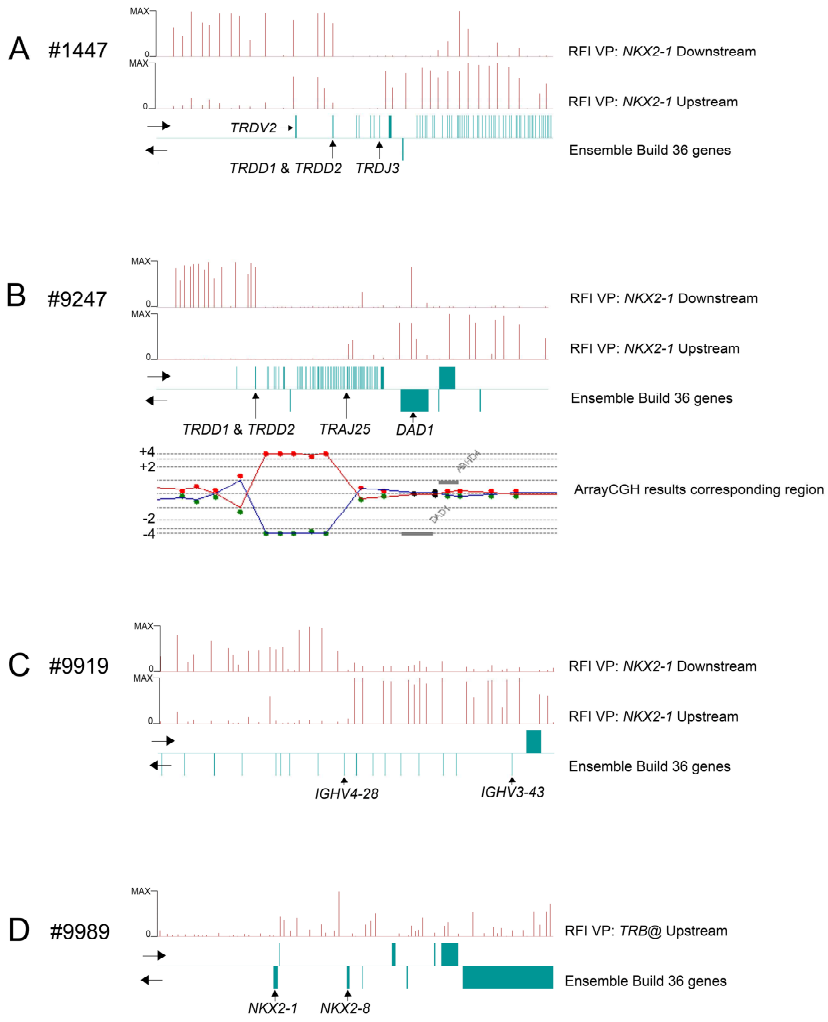



**Figure S3** Relates to Figure 3. High resolution 4C-technology results of *NKX2-1* breakpoints in *NKX2-1* rearranged T-ALL patient samples. Raw Fluorescent Intensities (RFI) for probes for the *TRAD@*, *IGH@* and *TRB@* that are being captured from viewpoints as indicated on the right. Viewpoints (VP) are the HindIII restriction fragments where the primers for 4C PCR are located. Each Bar represents a single HindIII fragment that is covered by a probe. Ensemble genes are shown below in green (build 36). The orientation of the genes is indicated by arrows at the left side. **(A)** For patient #1446, fragments centromeric of *TRDD1* and *TRDD2* are captured starting from the VP located downstream of *NKX2-1*, whereas fragments telomeric of *TRDJ3* are captured starting from the VP located upstream of *NKX2-1*. These data indicated that the breakpoint is located between *TRDD2* and *TRDJ3*. **(B)** In patient #9247 fragments centromeric of *TRDD1* and *TRDD2* are captured starting from the VP located downstream of *NKX2-1*, and fragments telomeric of *TRAJ25* are captured starting from the VP located upstream of *NKX2-1*. The large area between the fragments that are being captured from the VP up- and downstream of *NKX2-1* suggests a deletion of the intermediate area, that corresponds to array-CGH data for this sample as depicted in the lower row. For the array-CGH data, blue and red tracings represent dye swopped experiments for which each dot represents the log 2 scaled data for a single probeset. The genomic *TRAD@* reciprocal breakpoints are located between *TRDD2* and *TRAJ25* according to the 4C data. **(C)** In patient #9919 the breakpoint region can be confined to the area between 3 probes. The only gene segment situated in this area is *IGHV4-28*. **(D)** As patient #9989 has a deletion telomeric of the *TRB@* locus (according to FISH), 4C was started from the VP centromeric of *TRB@*. In this experiment, HinDIII fragments located telomeric of *NKX2-1* are captured. These results predict the location of the genomic breakpoint close to *NKX2-1*.

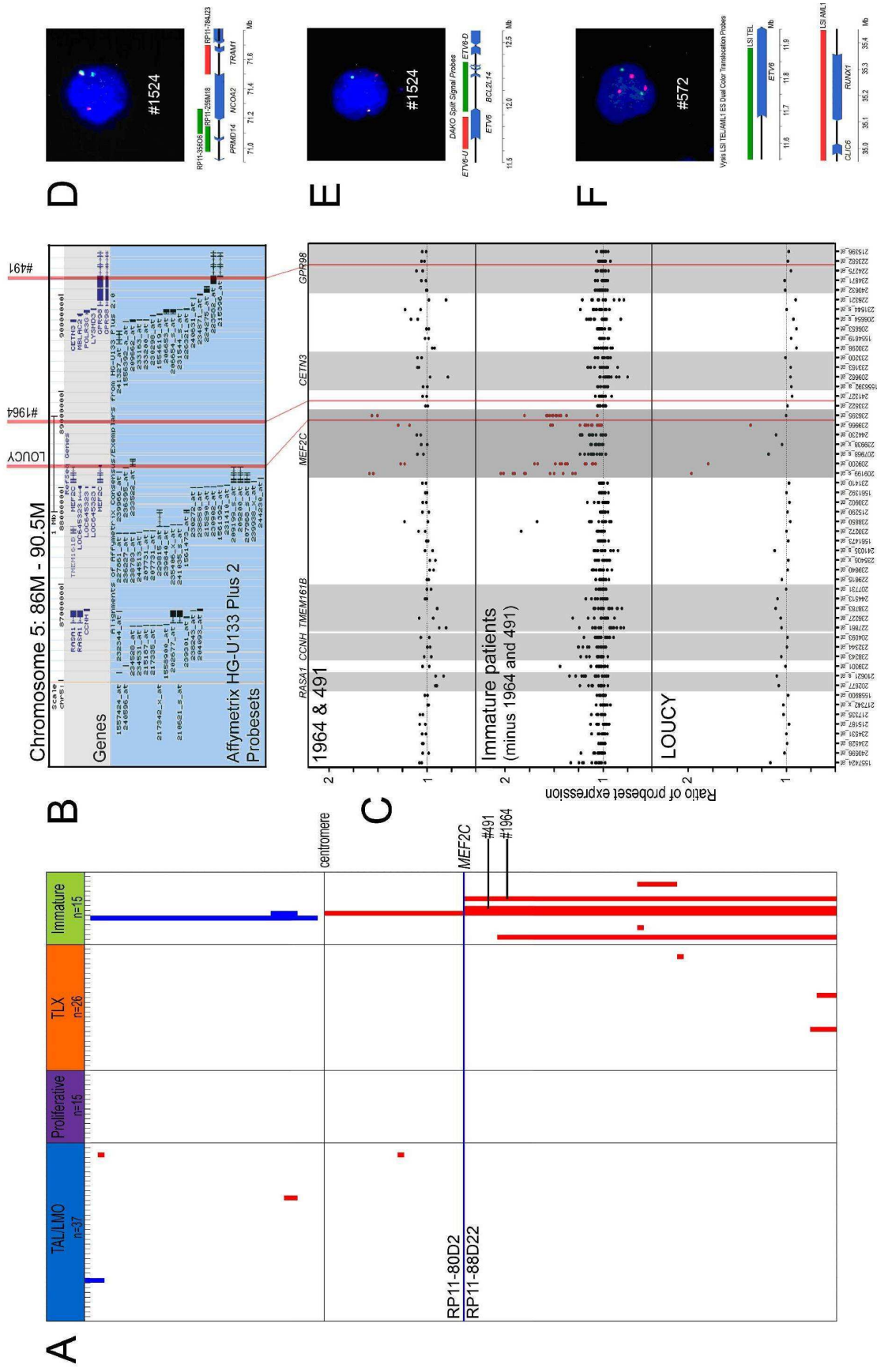

A
TAL/LMO n=37
Proliferative n=15
TLX n=26
Immature n=15
centromere
MEF2C
#491
#1964
RP11-80D2
RP11-88D22
B
Chromosome 5: 86M - 90.5M
LOUCY
#1964
#491
Genes
Affymetrix HG-U133 Plus 2 Probesets
C
1964 & 491
Immature patients (minus 1964 and 491)
LOUCY
RASA1
CCNH
TMEM161B
MEF2C
CETN3
GPR98
Ratio of probeset expression
D
#1524
RP11-35E06
RP11-259M18
RP11-784J23
PRMD14
NCOA2
TRAM1
E
#1524
DAKO Split Signal Probes
ETV6-U
ETV6
BCL2L14
ETV6-D
F
#572
Vysis LSI TEL/AML1 ES Dual Color Translocation Probes
LSI TEL
ETV6
LSI AML1
RUNX1
CLIC6

**Figure S4 (page 216)** Relates to Figure 4. Molecular-cytogenetic characterization of immature cluster cases. **(A)** Schematic illustration of array-CGH results for chromosome 5 for 92 pediatric T-ALL cases that cluster in the 4 unsupervised clusters (TAL/LMO, TLX, Immature and Proliferative, **Figure 1A**). Two patients (#491 and #1964) harbour a heterozygous 5q14-5qter deletion (shown in red) with breakpoints just telomeric (0.5-2 Mb) of the *MEF2C* gene (relative position is indicated). One copy gains (amplifications) are shown in blue. **(B)** Selective *MEF2C* activation as consequence of chromosomal rearrangements telomeric of *MEF2C*. Schematic representation of the 5q14 *MEF2C* chromosomal region. Location of genes and Affymetric probeset are indicated. The 5q14 chromosomal breakpoints for the T-ALL LOUCY cell line and T-ALL patient #1964 and #491 are shown by red lines. **(C)** Relative probeset intensities are given as ratio of probeset expression from immature patients or the cell line LOUCY compared to the median of T-ALL patients from all other subgroups combined or a panel of 17 cell lines (ALL-SIL, BE-13, CCRF-CEM, DND-41, HPB-ALL, HSB-2, JURKAT, KARPAS-45, KE-37, MOLT-3, MOLT-16, P12-ICHIKAWA, PF-382, RPMI-8402, SKW-3, SUPT1, TALL1), respectively. Only the *MEF2C* gene but not centromeric genes including *RASA1*, *CCNH*, *TMEM161B* are activated as consequence of these rearrangements. Split signal FISH analysis for *NCOA2* **(D)** and *ETV6* **(E)** loci in patient #1524. Both photographs show a split signal next to the normal fusion signal, indicating a translocation. **(F)** FISH analysis of *RUNX1* and *ETV6* loci in patient #572. FISH analysis shows a triple red signal in patient #572 corresponding to a translocation breakpoint in *RUNX1*. A schematic representation of probe positions is shown below the FISH photograph.

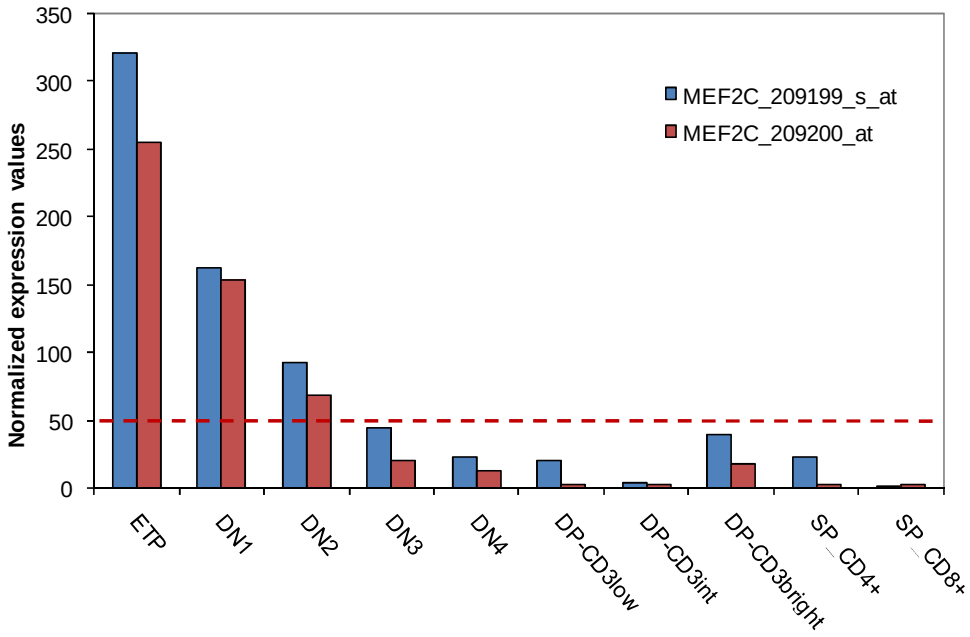



**Figure S5** Relates to Figure 6. Expression of *MEF2C* in normal human T-cell development. *MEF2C* (probeset 209100_s_at and 209200_at) expression levels relative to *ABL1* expression levels (probeset 202123_s_at) as extracted from the gene expression levels by microarrays for flow sorted thymic subsets (Soulier et al., 2005). Isolation and purification of thymic fractions have been described before ((Soulier et al., 2005), and references therein). The dashed line represents the positivity level. All values below this line were indicated as absent calls.

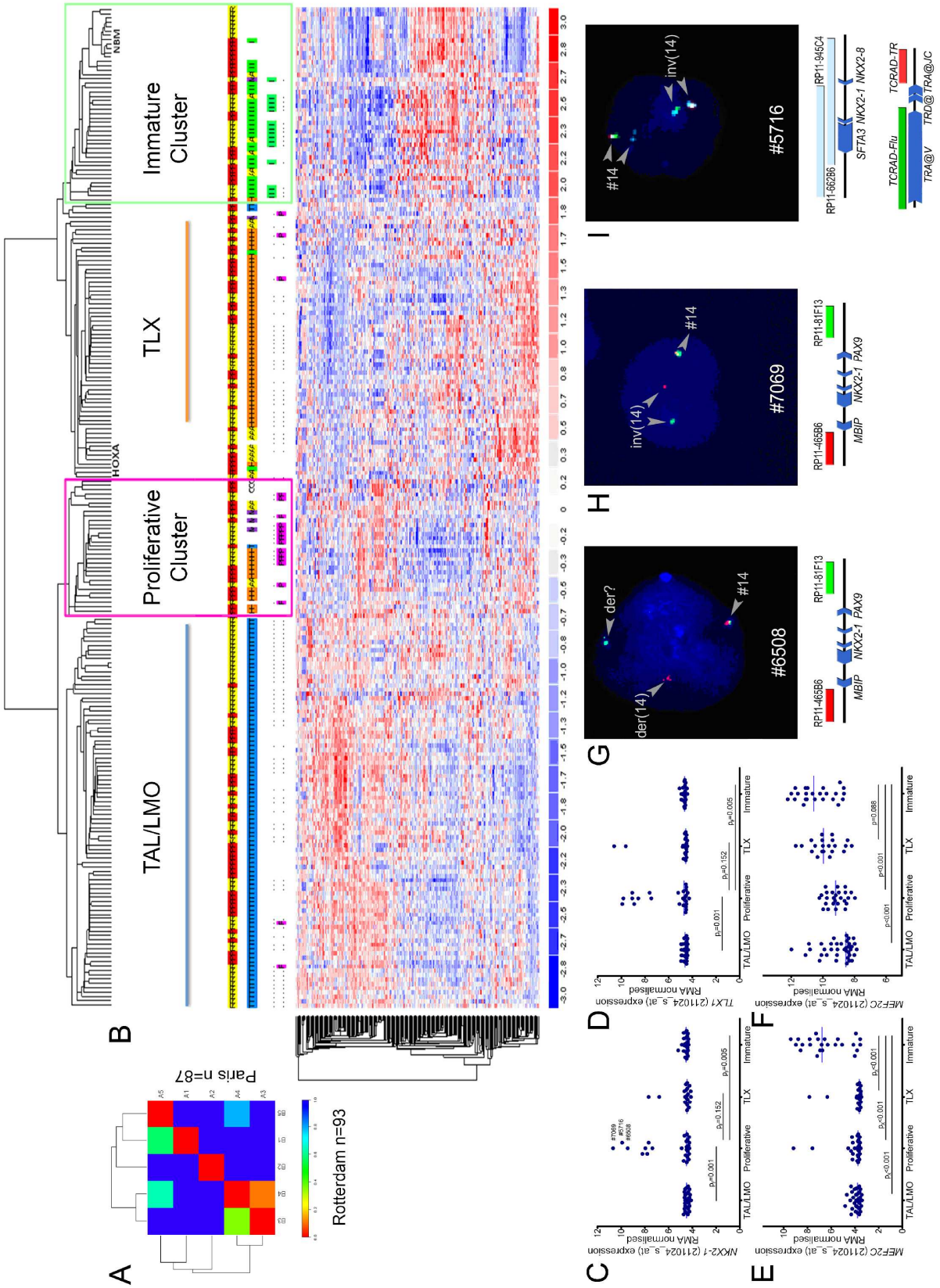

A
Paris n=87
Rotterdam n=93
B
TAL/LMO
Proliferative Cluster
TLX
Immature Cluster
HOXA
NBM
C
NKX2-1 (211024_s_at) expression
RMA normalised
p<0.001
p=0.152
p=0.005
TAL/LMO
Proliferative
TLX
Immature
D
TLX1 (211024_s_at) expression
RMA normalised
p=0.001
p=0.152
p=0.005
E
MEF2C (211024_s_at) expression
RMA normalised
p<0.001
p<0.001
p<0.001
F
MEF2C (211024_s_at) expression
RMA normalised
p<0.001
p=0.086
p<0.001
G
der?
#14
der(14)
#6508
RP11-466B6
RP11-81F13
MBIP
NKX2-1
PAX9
H
inv(14)
#14
#7069
RP11-466B6
RP11-81F13
MBIP
NKX2-1
PAX9
I
#14
inv(14)
#5716
RP11-66286
RP11-945C4
SFTA3
NKX2-1
NKX2-8
TCRAD-Flu
TRA@V
TCRAD-TR
TRD@TRA@JC

**Figure S6 (page 218)** Relates to Table 2. (**A)** Comparability of similar subgroups in the Paris and Rotterdam T-ALL datasets as visualized by the SUBMAP method. Shown is an SA matrix that reveals the comparability between the various subgroups in the Paris T-ALL dataset and the Rotterdam T-ALL dataset. For the Paris cohort the normal bone marrow controls (A1) and the T-ALL subgroups *TAL-R* (A2), *TLX1* (A3), *TLX3* (A4) and immature subgroup (A5) are shown versus the Rotterdam subgroups such as normal bone marrow controls (B1), *TAL/LMO* (B2), *TLX1* (B3), *TLX3* (B4) and immature subgroups (B5, as defined in **Figure 1**). (**B)** Identification of immature cluster and proliferative cluster samples in the combined Rotterdam and Paris T-ALL cohorts. Unsupervised cluster analysis in dChip for RMA-solo normalized datasets, after matching for identical probesets and adjustment for batch effects by the combat method (Johnson et al., 2007). Filtering criteria for the unsupervised analysis and annotations of the Rotterdam dataset have been described in the legend of **Figure 1**. The Paris dataset (107 T-ALL cases) has been described before (Clappier et al., 2007; Soulier et al., 2005). Other annotations used are: First line: yellow boxed cases, Rotterdam series; Red boxed cases, Paris series. Second line, consensus cytogenetic annotations: blue boxed cases, *TAL/LMO* rearrangements; orange boxed cases, *TLX3* rearranged; yellow boxed cases, *HOXA* activated cases; green boxed cases, immature cases; Third line: green boxed cases, Immature cluster samples from the Rotterdam series; Fourth line: purple boxed cases, proliferative cluster cases from the Rotterdam series; C, cell line. *NKX2-1* is expressed in 6 T-ALL cases from the Paris series, 4 of which belonging to the proliferative cluster. Fourteen out of 16 French cases previously annotated as immature fall in the combined immature cluster. *MEF2C* is expressed in 10 out of these 14 cases. **(C-I)** Molecular characterisation of the Rotterdam validation cohort. RMA normalized expression values for **(C)** *NKX2-1*, **(D)** *TLX1*, **(E)** *MEF2C* and **(F)** *LYL1* probesets for the PAM predicted unsupervised clusters in the Rotterdam cohort 2. Median values are indicated by lines. P-values for distribution of *NKX2-1*, *TLX1* and *MEF2C* were calculated by using the Chi-square ($P_x$) or the Fisher's exact test ($P_F$) assuming expression as positive when the normalized fluorescent intensity was >5, otherwise samples were considered as negative. Differences in the distribution of *LYL1* expression has been analyzed using the non-parametric Mann-Whitney-U test. The 3 *NKX2-1* rearranged cases (#7069, #5716 and #6508) as validated by FISH **(G-I)** are shown. FISH data for *NKX2-1* rearranged patients of the Rotterdam validation cohort. **(G)** Patient #6508 has a *NKX2-1* translocation. **(H)** Patient #7069 has a *NKX2-1* inversion. **(I)** Patient #5716 has a *NKX2-1/TRAD@* inversion on chromosome 14.

## Supplemental experimental Procedures

**Fluorescent *in-situ* hybridization (FISH).** FISH analysis was performed on thawed cytospin slides. The following probes were used: TCRβ, TCRα/δ, TCRγ and IgH split signal FISH DNA probes (Dako, Glostrup, Denmark), LSI TEL-AML1 ES dual color translocation probes (Abbot Molecular, Illinois), LSI ETV6 dual color break apart rearrangement probes (Abbot Molecular), following the manufacturer's instructions. For other loci we used home-labeled BAC clones (BAC/PAC Resource Center, Children's Hospital, Oakland, USA) as described before (Van Vlierberghe et al., 2008a). BAC clones are listed below.

**Microarray-based comparative genomic hybridization (array-CGH).** Human genome CGH 105A oligo microarrays (Agilent, Santa Clara, USA) were used according to the manufacturer's instructions. Slides were scanned in a 2565AA DNA microarray scanner (Agilent). Microarray

images were analyzed using feature extraction software (Agilent) and the data were subsequently imported into array-CGH analytics software (Agilent).

**Ligation mediated polymerase chain reaction (LM-PCR).** LM-PCR was performed as described before (Przybylski et al., 2005). We used primers located in the *Dδ3*, *Jδ1* and *Dβ1*, *Dβ2* gene segments of the *TRD@* and *TRB@* loci, respectively (van Dongen et al., 2003). To clone the *NKX2-2* der(20)chromosomal breakpoint region for patient #10138, we used a primer located telomeric of *NKX2-2*. Specific primers and adapter primer sequences used for LM-PCR are described in the table below.

**3' Rapid amplification of cDNA ends (3'-RACE).** 3'-RACE was performed from *RUNX1* for patient #571. For this, 1μg of RNA was incubated for 5 minutes at 37°C with 25pmol oligo-dT adapter primer in a volume of 15μl. cDNA synthesis was performed in a final volume of 25μl in the presence of 40U/μl RNAsin (Promega, Wisconsin, USA), 200nM of dNTPs, and 200U/μl MMLV-RT (Promega) in 1x RT-buffer at 42°C for 1 hr, followed by inactivation of MMLV-RT enzyme for 5 minutes at 95°C. One hundred microliters of $dH_2O$ were added, and cDNA was stored at –80°C. 3'RACE amplification was performed in a final volume of 50μl in the presence of 40ng of cDNA template, 1.5mM $MgCl_2$, 350nM dNTPs, 300nM adapter_1 primer, 300nM of RUNX1_1 primer, and 0,025U/μl IntegroTaq in 1x PCR buffer. After initial denaturation for 5' at 95°C, product was amplified for 14 touch-down cycles of 20'' at 95°C, 1' at annealing temperature (starting at 65°C which was lowered by a 0.5°C each consecutive cycle) and 1' at 72°C. Product was further amplified for another 31 cycles of 20'' at 95°C, 1' at 58°C and 1' at 72°C. One microliter of the PCR product was used for a nested PCR using the same reaction conditions and nested PCR primers. Product was amplified for 5' at 95°C, followed by 35 cycles of 20'' at 95°C, 1' at 60°C and 1' at 72°C. Total PCR product was separated on a 2% agarose gel, followed by excision of bands. DNA was isolated for sequencing using the QIAquick gel extraction kit (Qiagen, Venlo, Netherlands). Specific primers and adapter primer sequences used for 3'-RACE are described in the table below.

**DNA Sequencing.** Sequencing of PCR products according to the manufacturer's recommendations were performed in a 3130XL Genetic Analyzer (Applied Biosystems, Fostercity, USA).

**siRNA knockdown.** siRNA transfection was performed by electroporation as described before (Van Vlierberghe et al., 2008b). We used siRNAs directed against *NKX2-5* (siGenome on-target J-019795-05) or *MEF2C* (siGenome on-target J-009455-07) purchased at Dharmacon (Thermo Fisher Scientific, Waltham, USA). A FITC labeled random siRNA control (Eurogentec, Seraing, Belgium) was used to verify transfection efficiency.

***MEF2C* transfection.** Cell lines were cultured as preciously described (Van Vlierberghe et al., 2008b). A pCMV6 entry vector containing the *MEF2C* cDNA (TrueORF RC220584, Origene, Rockville, USA) was transfected into JURKAT cells by electroporation (350V, 10 ms rectangular pulse). Forty-eight hours after electroporation, cells were plated in 96-wells flat bottom plates (10,000 cells/well), and stable transfectants were selected on selective media containing Geneticin (2mg/ml, Invitrogen Life Technologies, Breda, The Netherlands). Stable transfected clones were validated for MEF2C expression by RQ-PCR and western blot.

**Chromatin Immunoprecipitation (ChIP).** Chromatin immunoprecipitation was performed as described before (Van Vlierberghe et al., 2008b). For each ChIP, 8 □g of antibody directed against NKX2-5 (ab54567, Abcam, Cambridge, United Kingdom) or MEF2C (D80C1, Cell Signaling, Danvers, USA) was used. RQ-PCR for the *MEF2C*, *HHEX*, and distal and proximal *LMO2* promotor regions were performed on immunoprecipitated DNA as described above. Primers sequences are described in the table below.

**Western blotting.** Western blotting was performed as described before (Van Vlierberghe et al., 2008b). The following antibodies were used: NKX2-5 (ab54567, Abcam, Cambridge, United Kingdom), NKX2-1 (sc-13040, Santa Cruz Biotechnology, Santa Cruz, USA), MEF2C (D80C1, Cell Signaling) and β-actin (A2547, Sigma).

**Flow Cytometry**. Expression of surface CD3 and TCRγ/δ was analyzed by flow cytometry using a FACS Calibur (Becton Dickinson, San Jose, CA, USA). Cells were washed with 1XPBS containing 0.1% BSA and stained with PerCP-labeled anti-CD3 (BD Pharmingen, San Diego, USA) and PE-labeled anti-TCR γ/δ (BD Pharmingen) antibodies after which two additional washes followed. For analysis 100,000 events were recorded.

**Colony Formation Assay.** NIH3T3 or BJ-EHT cells were cultured in DMEM + glutamax (Invitrogen) supplemented with 10% fetal calf serum, 2 mM L-glutamine (Invitrogen), 5 µg/ml insulin, 5 µg/ml transferrin, 5 ng/ml sodium selenite (ITS media supplement; Sigma, St Louis MO, USA), 100 IU/ml penicillin, 100 µg/ml streptomycin, 0.125 µg/ml fungizone and 0.2 mg/ml gentamycin (Invitrogen). Cells were transfected in 24 well plates ($1.5 \cdot 10^3$ cells/well) using a total of 1 µg total plasmid DNA and 1.5 µl Lipofectamine 2000 per well (Invitrogen) for 4 hours in 500 µl culture medium without antibiotics, after which medium was refreshed. Transfection was performed with 0.5µg of pCMV6 entry vector containing *MEF2C* (TrueORF RC220584, Origene) or *NKX2-1* (TrueORF RC217520, Origene) cDNA inserts, *MYC* or *RAS* expression vectors. All transfections with the constructs indicated were supplemented with empty pCDNA3 vector DNA (Invitrogen) to a total concentration of 1µg . The pCDNA3 vector served as empty vector control. After 24 hours, cells were plated in triplo at a densitiy of $3 \cdot 10^3$ cells/100mm plate. Following 8 days (NIH3T3) or 30 days (BJ-EHT) of culture, adherent cells were washed, fixed in 10% formaldehyde and dyed with 0.1% crystal violet in 20% methanol. Colonies with a diameter of 2 mm or larger were scored blind by 4 independent observers.

**Statistical analyses.** Asymmetric distribution of cytogenetic abnormalities or expression of immunohenotypic markers over unsupervised clusters was calculated by using the Chi-square statistical test. Statistical significance on mCD3 and TCRγδ expression in siRNA mediated knockdown of *MEF2C* were analyzed Student t-test. The p-values lower than p=0.05 were considered significant.

**Expression microarray data normalization.** For most analyses, Affymetrix CEL files were pre-processed using the RMA normalization method (Irizarry et al., 2003) followed by quantile normalization (Soulier et al., 2005)**.** To exclude clustering as consequence of the normalization method chosen, the dataset was also normalized using the VSN method (Huber et al., 2002), leading to consistent results.

**Unsupervised cluster analysis.** Unsupervised cluster analysis was performed in dChip software (Li and Wong, 2001) that calculates a coefficient of variation (CV) for each probeset in a dataset. Probesets with the largest variability in expression have the largest CVs. Selection of probesets occurs by setting upper and lower thresholds, with broader ranges resulting in a larger selection of probesets. Probesets were selected by the following criteria: CV between 1.2-10 (or as indicated in figure legends) for probesets expressed at minimal intensity values of 6 in ≥5 percent of the samples while masking for redundant probesets. The CV range of 1.2-10 resulted in the selection of 581 probesets (**Table S3**). Probesets reflecting contaminating erythroid or normal bone marrow cells in patient samples were removed, resulting in 435 probesets (**Table S3**) as used in **Figure 1A.** The influence of increasing number of probesets on the proliferative and immature clusters by expanding CV ranges in the unsupervized analysis was tested for unfiltered 581, 889, 1343 and 3034 probsets (**Table S1**) corresponding to CV ranges of 1.2-10, 1.1-10, 1.0-10 and 0.9-10, respectively. The general typology of clusters was conserved, as most cases initially assigned to prolerative or immature clusters (**Figure 1**) remained in these clusters (**Table S1**).

**Supervised cluster analysis based on Wilcoxon statistics.** For the supervised analysis, p-values for differentially expressed genes between T-ALL subgroups were calculated using a Wilcoxon statistical test ($p_{cox}$), and corrected for multiple testing error ($p_{fdr}$) according to the Hochberg and Benjaminin developed false discovery rate procedure (Hochberg and Benjamini, 1990). The Bioconductor package Multtest was used on VSN normalized data to determine the p-values and FDR corrected p–values for probesets that were significantly and differentially expressed for specific T-ALL genetic subgroups tested versus all other genetic subgroups, i.e. the *TAL1* subgroup (n=24; $p_{cox} < 2.21 \cdot 10^{-6}$ and

$p_{fdr}$<0.0012), the *LMO2* subgroup (n=9; $p_{cox}$=NS and $p_{fdr}$=NS), the combined *TAL/LMO* subgroup (n=33; $p_{cox}<1.47\cdot10^{-8}$ and $p_{fdr}<8.02\cdot10^{-6}$), the *TLX3* subgroup (n=22; $p_{cox}<1.24\cdot10^{-6}$ and $p_{fdr}$<0.00068), the *TLX1* subgroup (n=7; $p_{cox}$<0.00031 and $p_{fdr}$<0.053) and the *HOXA* subgroup (n=10; $p_{cox}<1.43\cdot10^{-5}$ and $p_{fdr}$<0.039). Using the combined top100s for these subgroups (324 probesets in total, **Table S3**), principal component analysis was performed using GeneMath XT 1.6.1. software (Applied Maths, Inc, Austin TX, USA). As displayed in **Figure 1C**, 24 cases formed two separate T-ALL subgroups that overlapped in great extend with samples consistently associated with the proliferative (P) and immature (I) clusters as identified by the unsupervised analyses (**Figure 1A**).

**Prediction of classes identified by unsupervised methods.** Stability of sample assignment towards the proliferative cluster (P), the immature cluster (I) or others (O) was tested by using various algorithms including Diagonal Linear Discriminant Analysis, 1-nearest neighbour, 3-nearest neighbour and nearest centroid in BRB tools version 3.7 (R. Simon & A.P. Lam). These analyses were performed on 12676 highest variable probesets as obtained in dChip (CV range 0.5-10). Normal bone marrow samples were excluded from the analysis. On average, results were optimal using 250 probesets (analysis 1). Using these parameters, sample #9105 and samples #9696, #2792, and #2113 were not predicted as immature cluster or proliferative cluster samples, respectively (**Table S2**). Repeating the procedure while leaving out these samples (analysis 2) identified another 4 cases (#9577, #2703 and #2252 (immature cluster) and #9919 (proliferative cluster) that no longer stably clustered into these clusters (**Table S2**). Repeating this procedure again while leaving out these 4 samples (analysis 3) did not further result in the identification of samples that were not stably assigned to the immature cluster, the proliferative cluster or other T-ALL subgroups (**Table S2**).

As an alternative strategy, PAM analysis (Tibshirani et al., 2002) was performed for the same classes (P, I and O) using an identical set of probesets (12676). Using a 10-fold cross validation, most stable clustering of patient samples was achieved by 100 probesets, in which proliferative cluster sample #9989 was classified as belonging to one of the other T-ALL subgroups. All immature cluster samples were correctly assigned to the immature cluster (**Table S2**).

**Outlier gene analysis for the immature and proliferative subgroups.**

Outlier gene analysis was performed by COPA statistics (Tomlins et al., 2005). For this, quantiles from 0.75 to 1.0 were analyzed. Normal bone marrow samples were excluded. The proliferative and immature clusters were tested separately whereby all other cases were considered as a single control group (e.g. when testing the proliferative cluster, all others samples including those belonging to the immature cluster were included in the control). Probesets for which maximum expression in the tested group was below 8 were excluded from the analysis. Permutation analysis was done on the top200 most significantly, differentially expressed probesets. Lists of the most significant probesets of outlier genes expressed in the proliferative or immature clusters are displayed in **Tables S5**.

For the proliferative cluster samples, COPA statistics identified outlier expression of the *NKX2-2* and *NKX2-1* genes as potential candidates for rearrangements (**Table S5**). In support of this, proliferative cluster sample #10138 uniquely expressed high levels of *NKX2-2* while *NKX2-1* expression was detected in 13 T-ALL samples in total of which 10 belonged to the proliferative cluster. *MEF2C* was identified as top ranking gene in the immature cluster that was strongly up-regulated in all samples (**Table S5**). *MN1* and *CLECL1* were also identified as top-ranking genes in this analysis. *MN1* has been described as a gene that is targeted by chromosomal alterations in the inv(16) M4EO AML subtype (Buijs et al., 2000; Grosveld, 2007).By using PAM, the 6 subgroups as identified by PCA analysis could robustly be reproduced based on a minimal set of 195 probesets (**Tables S5**). Nine out of 12 proliferative cluster samples were repeatedly assigned to the corresponding group based on only 2 probesets that both encoded for *NKX2-1*. Eleven of the immature samples were most robustly assigned to this group based on 13 probesets that included 3 *MEF2C* probesets (**Table S5**). This further confirms that *NKX2-1* and *MEF2C* may actually be novel oncogenes for the proliferative and immature cluster samples, respectively.

**Comparability of equivalent T-ALL subgroups between the French and the Rotterdam datasets: OrderedList method and Subclass mapping.** Comparability between both datasets was

tested by 2 different strategies, i.e. the OrderedList method (Lottaz et al., 2006) and the Submap method (Hoshida et al., 2007). For the OrderedList method, *TAL1* rearranged samples, *HOX*-rearranged samples (i.e. *TLX1* or *TLX3*) and immature T-ALL samples were mutually compared between the Rotterdam T-ALL and the French T-ALL datasets (Soulier et al., 2005). For this, genelists of differentially expressed genes for indicated subgroups in the French dataset were then compared to genelists of the equivalent subgroups in the Rotterdam dataset. Cytogenetically non-annotated samples from the French cohort (n=24) and 21 cases including all proliferative cluster samples from the Rotterdam cohort were left out of this analysis. For this comparison, we used the Bioconductor package OrderedList (Lottaz et al., 2006). Differentially expressed genes for these subgroups in both datasets were compared by z-statistics. Briefly, the OrderedList method relies on the definition of parameter α that defines the weighing scheme for each ranked probeset and how many ranks are taken into account in this comparison between 2 datasets. For this, the distribution of observed scores and random scores were evaluated to decide which parameter α results in the most reliable score. Observed scores are derived by drawing 80 percent of the samples of the subgroups for each cohort, whereas random scores (empirical p-value) are derived by random shuffling of samples in each cohort. The number of permutations for this analysis was 10000. The overlap of 2 score distributions is evaluated by a similarity-score. In both datasets, we compared the ordered genelists for the subgroups as described above. Cut-off's for genelists was set at 99 percent of the probesets that explained the overall similarity between comparable subgroups. For these subgroups, the optimal α-value was 0.115, 0.115 and 0.038, respectively (not shown). At these α-values, the percentage of overlapping genes between both datasets was 30-50 percent for each subgroup comparison (not shown). For Subclass mapping (SubMap), we compared lists of differentially expressed probesets for comparable or identical subgroups including *TAL_R* (Paris) versus *TAL/LMO* (Rotterdam), *TLX1*, *TLX3*, immature or normal bone marrow samples between both datasets (**Figure S6A**). As for the OrderedList method, only major subtypes were considered in this analysis as well. To briefly outline this method, the overlap in top ranking differentially expressed genes from comparable subgroups in both datasets are compared in a two-way comparison b Gene Set Enrichment Analysis (GSEA). These results provide enrichment scores (ES) for mutual comparisons. A nominal p-value is calculated for each subgroup in a dataset by randomly distributing subgroup labels in the other dataset. A 1000 permutations were performed to calculate the nominal p-value in each comparison. The 2 enrichment scores for the mutual subgroup comparisons are combined by using Fisher inverse chi-square statistics (*F*), and compared to a nominal p-value for F by randomly picking 10000 ES scores from null distributions in the 2-way Gene Set Enrichment Analyses. After correction for multiple testing, p-values for similarity between comparable subgroups in the datasets are summarized in a subclass association matrix.

**Validation of the immature cluster and the proliferative cluster in an independent T-ALL cohort, i.e. the French T-ALL U133A expression set.** The proliferative and immature clusters were validated in an independent validation cohort based upon expression signatures. For this, we used the microarray expression dataset of the French T-ALL cohort as previous described (Soulier et al., 2005). Annotations for this cohort were as described previously (Clappier et al., 2007; Soulier et al., 2005). As these T-ALL cases had been arrayed on the U133A Affymetrix expression arrays, we matched the Rotterdam U133 plus 2 dataset with the French U133A dataset for overlapping probesets resulting in 22676 overlapping probesets. The combined dataset was RMA-solo normalized and corrected for batch effects using the COMBAT method (Johnson et al., 2007). As expected, clustering of the two combined series revealed a strong batch effect which disappeared after applying the "Combat" method based on an empirical Bayes strategy (**Figure S6B**).

| **4C** | **Forward** | **Reverse** |
|---|---|---|
| *NKX2.1* upstream | 5'-TGAGACCCACCAACTACA-3' | 5'-ATTAGCCACCACTGAACC-3' |
| *NKX2.1* downstream | 5'-GTGTGAGGTCAGAAAGAAGA-3' | 5'-ACTTTCCACTGACACAACTC-3' |
| *BCL11B* upstream | 5'-TATCCCAAACTTTTACAACC-3' | 5'-CTTCCTGCAGTGAGTGTAC-3' |
| *MEF2C* located in gene | 5'-GAAAAGGGAGTTCCTGAGT-3' | 5'-GGCATTACCCTTGATGTAC-3' |
| *TCRβ* upstream | 5'-CCTTGATGTTTCTCCCTTTACC-3' | 5'-CATGAAGAAACGAGCACCC -3' |
| **PCR** | **Forward** | **Reverse** |
| #9989-genomic breakpoint | 5'-CTTTGGCAAAAAAAATGACTT-3' | 5-GAGGACCGAAATACCAAACTA-3' |
| **RT-PCR** | **Forward** | **Reverse** |
| *NCOA2* exon 11 | | 5'-GTGAGGGGCTGTTCATTT-3' |
| *NCOA2* exon 12 | 5'-AGCCCTGTCACACCTGTT-3' | |
| *NCOA2* exon 13 | 5'-ATGGGTAATCAAGGGATGATA'-3 | |
| *NCOA2* exon 14 | | 5'-TAGGCCGAGAAGCACTGT-3' |
| *NCOA2* exon 17 | | 5'-GGCGATGCTGAAGTTGA-3' |
| *NCOA2* exon 21 | | 5'-TGCTGCCCAAAGTGTG-3' |
| *ETV6* exon 1 | 5'-GCTGGAAGAAACTTCTTAAATGA-3' | |
| *ETV6* exon 3 | 5'-CTTTCGCTATCGATCTCCTC-3' | |
| *ETV6* exon 5 | 5'-TGCCCATTGGGAGAATAG-3' | 5'-GAGCGGTGCAACAGTTC-3' |
| *ETV6* exon 6 | | 5'-TCCCATCGGATGAAGTTT-3' |
| *ETV6* exon 7 | 5'-CCTGCGCCACTACTACAA-3' | |
| *RUNX1* exon 7 | | 5'-TGGGGATGGTTGGATCT-3' |
| *AFF3* exon 8 | 5'-CTGGCTTCCACCACTTTC-3' | |
| **RQ-PCR** | **Forward** | **Reverse** |
| *MEF2C* | 5'-GCGCTGATCATCTTCAAC-3' | 5'-CTTTGCCTGCTGATCATT-3' |
| *NKX2-1* | 5'-TACCAGGACACCATGAGG-3' | 5'-GTCGCTCCAGCTCGTAC-3' |
| *NKX2-5* | 5'-TATCCACGTGCCTACAGC-3' | 5'-TGCGTGGACGTGAGTT-3' |
| *ENO2* | 5'-CTGCCTGGTCCAAGTTC-3' | 5'- TCCTGAGCGATGACTCAC-3' |
| *HHEX* | 5'-ATCGACGCGCTAAATG-3' | 5'- ATGCCAATGCCAGTGG-3' |
| *PSCD4* | 5'-GCACGGGTCATCTTTTC-3' | 5'- CTTGCGCCCAATACAC-3' |
| *CHRNA3* | 5'-GGAGAGGCCGTCTCTG-3' | 5'- ACAGGCCGGATGATCT-3' |
| *PDK1* | 5'-AATGCTTGTGAAAAGACCTC-3' | 5'- CATCCTCAGCACTTTTGTC-3' |
| *TUSC3* | 5'-GCACCACCTCGAAACTATT-3' | 5'- TCTGTCCCCTCATCATAGTC-3' |
| *FAM46A* | 5'-ACTGCCTGTTGGACTTCTT-3' | 5'- TTTGCCACTGTTGTTTGAC-3' |
| *LYL1* | 5'-CGCTGCAACTCTC-3' | 5'-ACCAGGAAGCCGATGTA-3' |
| *LMO2* | 5'-TTGGGGACCGCTACTT-3' | 5'-ATGTCCTGTTCGCACACT-3' |
| **LM-PCR** | **First Primer** | **Nested Primer** |
| #10138-genomic breakpoint | 5'-GGGCAGTTGGGTGTTTCTT-3' | 5'-CGTTGCTTTTCCCATCTTTG-3' |
| **3'RACE** | **First Primer** | **Nested Primer** |
| *RUNX1* | 5'-GTCGGTCGAAGTGGAAGA-3' | 5'-AAGTCGCCACCTACCACA-3' |
| Adapter | 5'-TTCGCACGAGCAATTAG(T)17-3' | |
| Adapter Primers | 5'-TTCGCACGAGCAATTAG-3' | 5'-CGCACGAGCAATTAGTTT-3' |
| **ChIP** | **Forward** | **Reverse** |
| *MEF2C*-8000 exon 4/1a | 5'-ATGGCTTCAGAAGTCCTATG-3' | 5'-AGTGCCAAGTTCTCTGTTTC-3' |
| *HHEX* | 5'-CCGTTCATACAGGAAATCTT-3' | 5'-GGCTATCAGAAGTCGAGTGT-3' |
| *LMO2* distal promotor | 5'AGGTCCAATGTGAACTCAAT-3' | 5'-TGGTCTGGTGGTTAGCATA-3' |
| *LMO2* proximal promotor | 5'-GTGGGTTACTTTCTGCCTT-3' | 5'-CTCCTGGGGATTAGCAAT-3' |

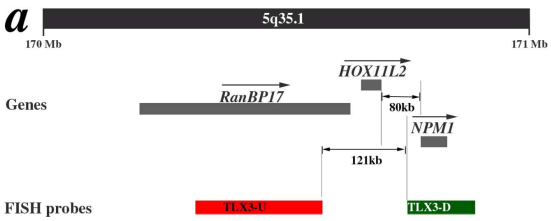
**BAC clones used for FISH analyses**

| | |
|---|---|
| *BCL11B* | RP11-431B1, RP11-74H1, RP11-68I8, RP11-889B13 |
| *FOXH1* | RP11-48I1, RP11-1143I12, RP11-1022M7 |
| *HHEX* | RP11-148C9, RP11-790D23 |
| *LYL1* | RP11-352L7, RP11-356L15 |
| *MEF2C* | RP11-749E19, RP11-845O10, RP11-467C24, |
| *MEF2C distal deletion* | RP11-236J15, RP11-1147F22 |
| *MN1* | RP11-46E17, RP11-1056M20, RP11,79G21 |
| *MYB* | RP11-378M4, RP11-104D9, RP11-141K5, RP11-937M14 |
| *MYC* | RP11-100H12, RP11-372A3 |
| *NCOA2* | RP11-259M18, RP11-356O6, RP11-784J23 |
| *NKX2-1* | RP11-945C4, RP11-662B6 or RP11-465B6, RP11-81F13 |
| *NKX2-2* | RP11-872K7, RP11-1065O2 |
| *NKX2-5* | RP11-166N7, RP11-352N17 |

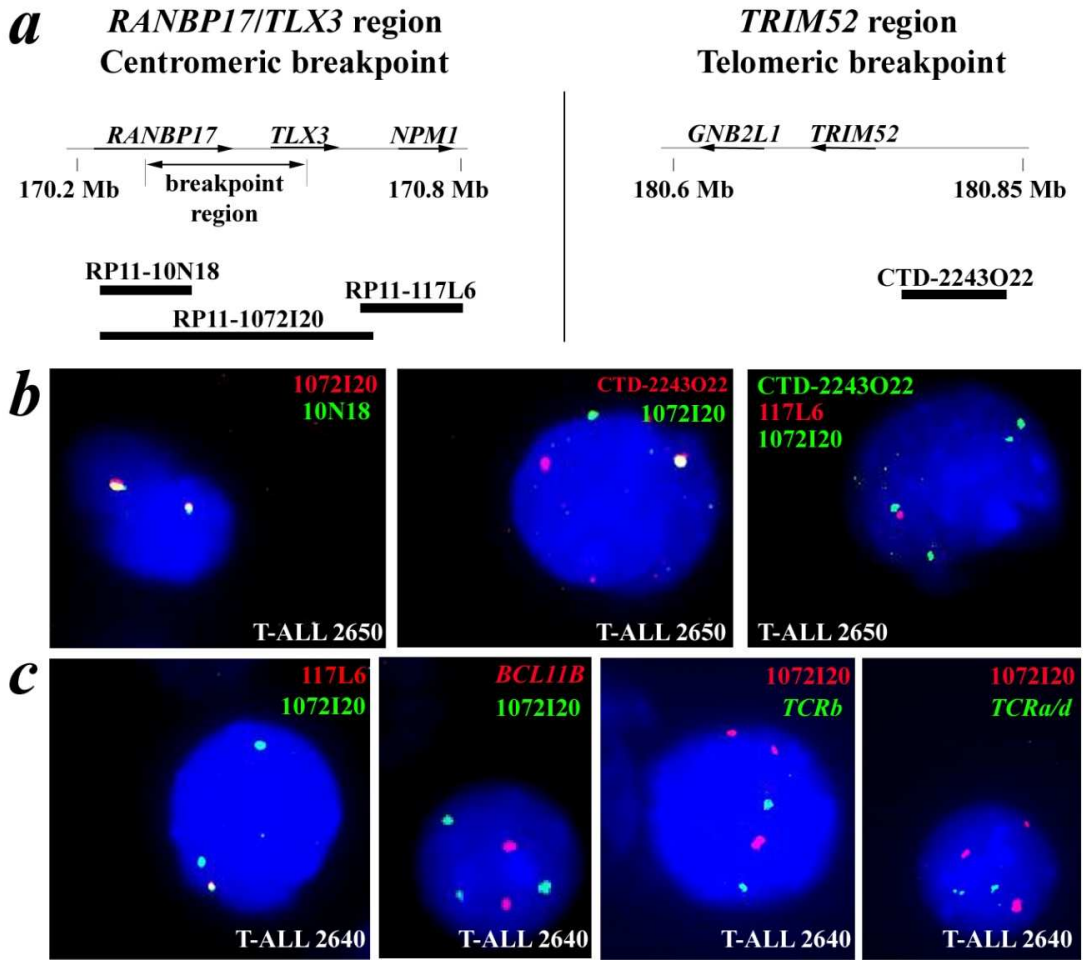
**Supplementary table 1. Combinations of genetic defects in *TLX3* rearranged T-ALL patients**

| | Known genetic defects | | | Novel recurrent lesions | | | | | Novel lesions identified in single *TLX3* rearranged patient samples | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ID | NOTCH1[A] | *p15/p16*[B] | *NUP214-ABL1* | del(5)(q35) | del(1)(p36.31) | del(13)(q14.3) | del(16)(q22.1) | del(19)(p13.2) | del(1)(p36.12p36.13) | del(1)(p34.2p34.3) | del(2)(p23.3p24.1) | del(2)(q37.1) | del(3)(q13.32q21.2) | del(3)(q26.2q26.31) | del(4)(q31.3q32.1) | del(6)(q25.1qter) | del(7)(q31.33q36.2) | amp(8) | del(9)(p24.1p24.2) | del(9)(q21.13q31.1) | del(10)(p15.2p15.3) | del(10)(p11.23p12.1) | del(11)(p13) | del(11)(q21q22.3) | del(12)(p13.1p13.2) | del(17)(q11.2) | del(20)(p12.3pter) |
| 634 | WT | -/- | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 720 | MUT; del P 1583 | +/+ | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 1179 | WT | -/- | + | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | + | - | - |
| 2738 | MUT; ins Q 1603 | -/- | - | - | + | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 2757 | MUT; ins 34 aa stop 2444 | -/- | + | - | + | - | - | - | + | + | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 2773 | MUT; V 1605 E | +/+ | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 2780 | MUT; L 1601 P | -/- | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | + | - |
| 2786 | MUT; MIFP 1616 ST | -/- | - | - | - | - | - | - | - | - | - | - | - | - | + | - | - | - | - | - | - | - | - | - | - | - |
| 2105 | ND | +/+ | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | + | - | - | - | - | - | - | - | - |
| 2100 | MUT; R 1599 P | -/- | - | - | - | - | + | - | - | - | - | - | - | - | - | - | - | - | + | - | - | - | - | - | - | - |
| 1946 | WT | +/+ | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 2112 | ND | -/- | + | + | + | + | + | - | - | - | - | + | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 2640 | MUT; L 1594 P | -/- | - | + | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 2650 | MUT; L 1586 P | +/+ | - | + | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 2723 | MUT; L 1679 P | -/- | - | - | - | + | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | + | - | - |
| 222 | MUT; L 1594 P; S 1598 I | -/- | - | + | - | - | - | + | - | - | - | - | - | - | - | + | - | - | - | + | - | - | - | - | - | - |
| 378 | MUT; L 1601 P | -/- | - | - | - | - | - | + | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 585 | MUT; ins IEVSIYRGL 2473 | -/- | + | - | + | - | - | - | - | - | - | - | + | + | - | - | - | - | - | - | - | - | - | - | - | - |
| 749 | WT | -/- | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 9012 | WT | -/- | - | + | - | - | + | - | - | - | + | - | - | - | - | - | + | - | - | - | - | + | - | - | - | + |
| 9858 | MUT; L 1575 P | -/- | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | + | - | - | - | + | - |

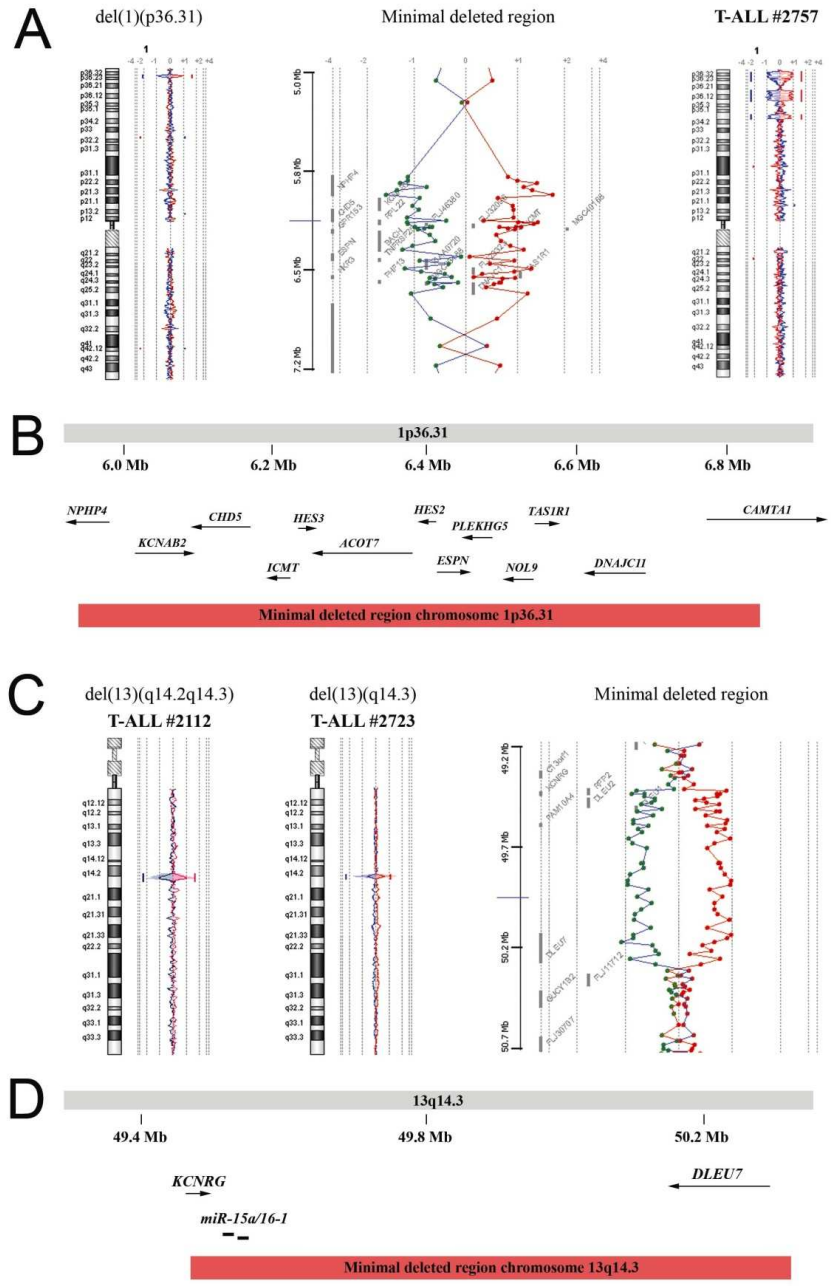
[A]As determined in van grotel et al. (24). [B] +/+, wildtype; +/-, heterozygous deletion; -/-, homozygous deletion of the *p15/p16* locus.



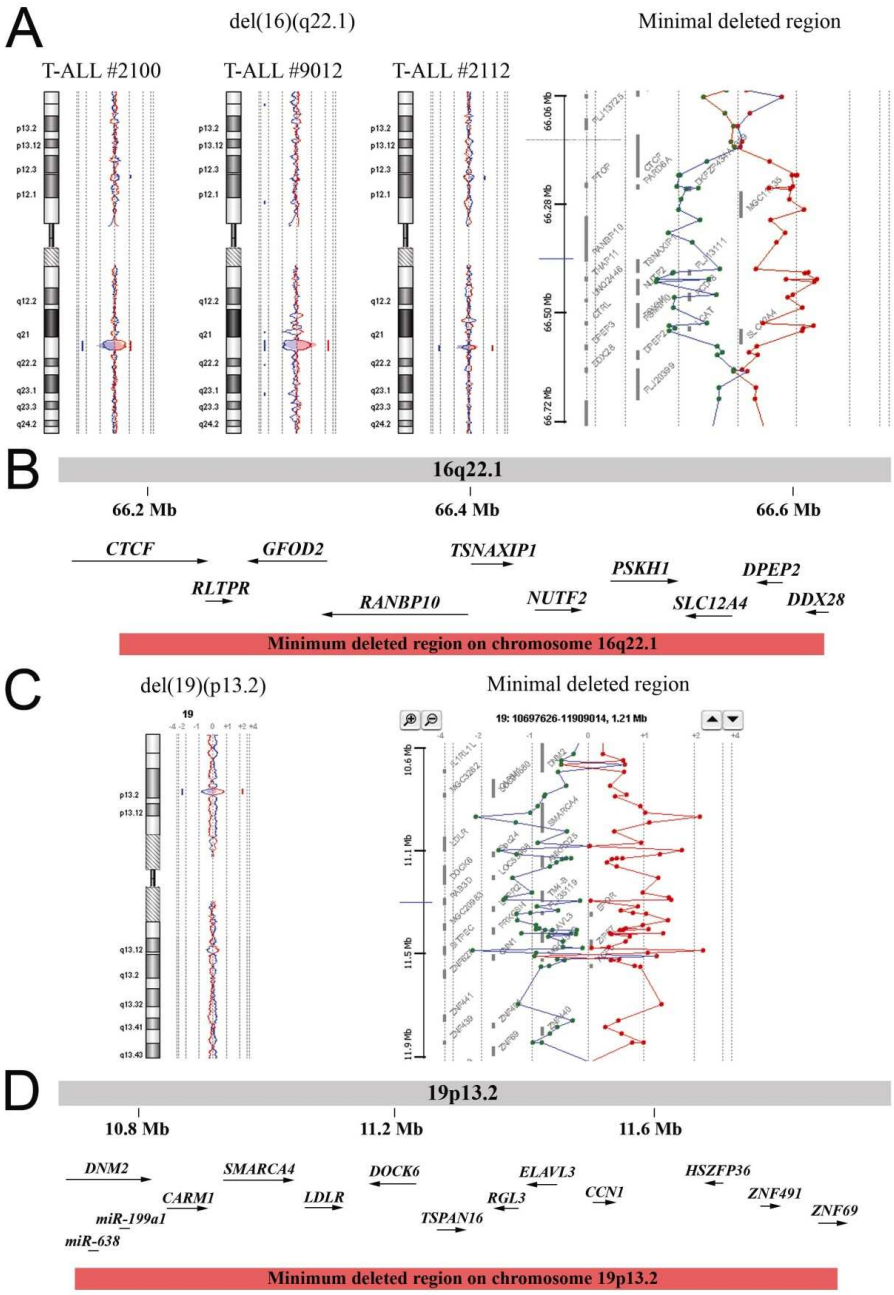

**Supplementary Figure 1. Genomic localization of the *TLX3-U/TLX3-D* translocation probeset. (a)** Overview of the genomic positions of the *TLX3* break apart probe set, at chromosome 5 band q35.1, used for FISH analysis. Specific genes located in this region are indicated. Depicted genome positions are based on the UCSC Genome Browser.



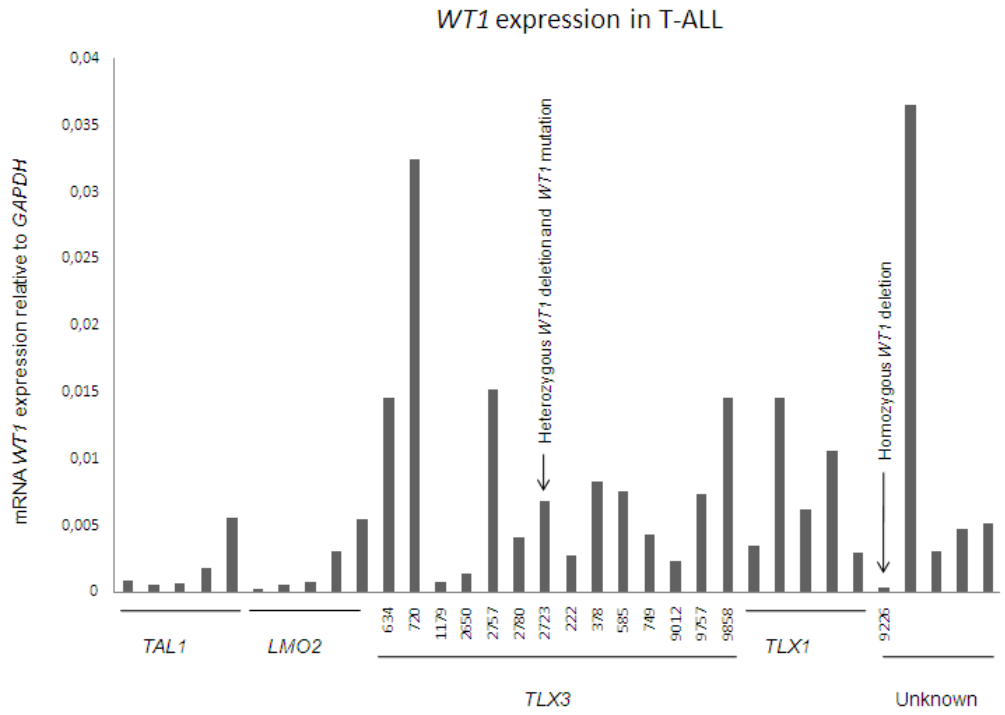

**Supplementary Figure 2. Further characterisation of the del(5)(q35) in T-ALL using FISH analysis. (a)** Overview of the genomic positions of BAC clones used for additional FISH analysis. **(b)** Further characterization of the deletion on 5q in patient 2650. FISH analysis using RP11-1072I20 and RP11-10N18 (left panel) excludes the presence of a *TLX3* translocation. FISH analysis using RP11-1072I20 and CTD-2243O22 (middle panel) revealed a fusion signal on the derivative chromosome 5. FISH analysis using RP11-1072I20, RP11-117L6 and CTD-2243O22 (right panel) revealed a fusion of the green signals and loss of the red signal. **(c)** Further characterization of the del(5)(q35.1) in patient 2640. FISH analysis revealed an additional RP11-1072I20 hybridization signal, indicative for partial *RANBP17*/*TLX3* amplification. FISH analysis showed no involvement of the *TCRα/δ*, *TCRβ* and *BCL11B* loci.




**Supplementary Figure 3. The recurrent deletions, del(1)(p36.31) and del(13)(q14.3), in *TLX3* rearranged pediatric T-ALL. (a)** Representative chromosome 1 ideogram and oligo array-CGH plot for the deletion, del(1)(p36.31), as detected in cases 2738, 2112 and 585 (left panel). Case 2757 (right panel) shows multiple deletions on the short arm of chromosome 1. The middle panel shows a detailed overview of the minimal deleted region for 1p36 deletions in T-ALL. **(b)** Schematic overview of the genes situated in the chromosomal region 1p36.31.
**(c)** Chromosome 13 ideograms and oligo array-CGH plots for the deletions, del(13)(q14.2q14.3) and del(13)(q14.3), as detected in cases 2112 (left panel) and 2723 (middle panel). The right panel shows a detailed overview of the minimal deleted region for 13q14 deletions in T-ALL. **(d)** Schematic overview of the genes and microRNAs situated in the chromosomal region 13q14.3.




**Supplementary Figure 4. The recurrent deletions, del(16)(q22.1) and del(19)(p13.2), in *TLX3* rearranged pediatric T-ALL. (a)** Chromosome 16 ideograms and oligo array-CGH plots for the deletions, del(16)(q22.1), as detected in cases 2100, 2112 and 9012 (left panel). The right panel shows a detailed overview of the minimal deleted region for 16q22.1 deletions in T-ALL. **(b)** Schematic overview of the genes situated in the chromosomal region 16q22.1. **(c)** Representative chromosome 19 ideogram and oligo array-CGH plot for the deletion, del(19)(p13.2), as detected in cases 222 and 378 (left panel). The right panel shows a detailed overview of the minimal deleted region for 19p13.2 deletions in T-ALL. **(d)** Schematic overview of the genes and microRNAs situated in the chromosomal region 19p13.2.




**Supplementary Figure 5. *WT1* expression analysis in T-ALL.** *WT1* mRNA expression data relative to *GAPDH*, based upon gene expression array data, which were available for a selection of patient samples. *WT1* expression levels are shown for 14 of the *TLX3* rearranged T-ALL cases and for 20 non *TLX3* rearranged T-ALL cases including *TAL1, TLX1, LMO2* rearranged patients and unknown cases.

## CHAPTER 5

Heterodimerization domain

```
HN1 1563  AEHVPERLAAGTLVVVVLMPPEQLRNSSFHFLRELSRVLHTNVVFKRDAHGQQMIFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKASLL 1654
2113      ------------P-------------------------------------------------------------------------------- (ref 1,2,3,4,5)
9858      ------------P-------------------------------------------------------------------------------- (ref 1,2,3,4,5)
2737      ----------------E---------------------------------------------------------------------------- (ref 4,5)
9696      ----------------E---------------------------------------------------------------------------- (ref 4,5)
2775      ----------------LMPPEQLRNSSFHFLRELSRVLHTNVVFKRDAHGQQMIFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKASLL  (ref 1,3,4,5,6)
2774      ----------------LMPPEQLRNSSFHFLRELSRVLHTNVVFKRDAHGQQMIFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKASLL  (ref 1,3,4,5,6)
2772      ----------------LMPPEQLRNSSFHFLRELSRVLHTNVVFKRDAHGQQMIFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKASLL  (ref 1,3,4,5,6)
9757      ----------------MPPEQLRNSSFHFLRELSRVLHTNVVFKRDAHGQQMIFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKASLL  (new)
10110     ----------------VLMPPEQLRNSSFHFLRELSRVLHTNVVFKRDAHGQQMIFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKASL (new)
2322      ----------------PPEQLMSSFHFLRELSRVLHTNVVFKRDAHGQQMIFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKASLL   (ref 5)
1949      ------------------PRYELPPEQLRNSSFHFLRELSRVLHTNVVFKRDAHGQQMIFPYYGREEELRKHPIKRAAEGWAAPDALLGQV (ref 4)
2749      --------------------EQLRNSSFHFLRELSRVLHTNVVFKRDAHGQQMIFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKASLL- (ref 4)
720       ------------------EQLRNSSFHFLRELSRVLHTNVVFKRDAHGQQMIFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKASLL- (ref 4)
1446      ------------------GPEQLRNSSFHFLRELSRVLHTNVVFKRDAHGQQMIFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKASL (new)
9919      --------------------LPREQLRNSSFHFLRELSRVLHTNVVFKRDAHGQQMIFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKA (new)
2130      ------------------PEQLRNSSFHFLRELSRVLHTNVVFKRDAHGQQMIFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKASL (new)
2229      ----------------------P---------------------------------------------------------------------- (ref 1,2,3,4,5,6)
2720      ----------------------P---------------------------------------------------------------------- (ref 1,2,3,4,5,6)
2650      ----------------------P---------------------------------------------------------------------- (ref 1,2,3,4,5,6)
2748      ----------------------P---------------------------------------------------------------------- (ref 1,2,3,4,5,6)
2793      ----------------------P---------------------------------------------------------------------- (ref 1,2,3,4,5,6)
1953      ---------------------Q----------------------------------------------------------------------- (ref 4)
2750      ------------------------------S---------------------------------------------------------------- (ref 1,2,3,4,5)
9247      ------------------------------S---------------------------------------------------------------- (ref 1,2,3,4,5)
2640      ----------------------------P------------------------------------------------------------------ (ref 1,2,3,4,5,6)
2847      ----------------------------P------------------------------------------------------------------ (ref 1,2,3,4,5,6)
2854      ----------------------------P------------------------------------------------------------------ (ref 1,2,3,4,5,6)
222       ----------------------------P---I-------------------------------------------------------------- (ref 1,2,3,4,5,6 / new)
2100      ------------------------------------P--------------------------------------------------------- (ref 1,2,3,4,5)
2509                                      QKVLHTNVVFKRDAHGQQMIFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKASL (new)
1701      ---------------------------------------P------------------------------------------------------ (ref 1,2,3,4,5)
2780      ---------------------------------------P------------------------------------------------------ (ref 1,2,3,4,5)
2116      ---------------------------------------P------------------------------------------------------ (ref 1,2,3,4,5)
2788      ---------------------------------------P------------------------------------------------------ (ref 1,2,3,4,5)


HN1 1563 AEHVPERLAAGTLVVVVLMPPEQLRNSSFHFLRELSRVLHTNVVFKRDAHGQQMIFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKASLL 1654
2105      ---------------------------------------P------------------------------------------------------ (ref 1,2,3,4,5)
378       ---------------------------------------P------------------------------------------------------ (ref 1,2,3,4,5)
9938      ---------------------------------------P------------------------------------------------------ (ref 1,2,3,4,5)
8577      ---------------------------------------P------------------------------------------------------ (ref 1,2,3,4,5)
8639      --------------------------------------Q------------------------------------------------------- (ref 2)
2738      ----------------------------------------QTNVVFKRDAHGQQMIFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKASL (ref 4)
2773      ------------------------------------------E----------------------------------------------------- (ref 4)
419       ------------------------------------------G----------------------------------------------------- (ref 4)
2117      ---------------------------------------------RSERDAHGQQMIFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKASL (ref 4)
1632      ----------------------------------------------PSYLRPGGSDTRPADDLPLLRPRGGAAQAPHQACRRGLGRT*      (new)
9421      ---------------------------------------------RDAHGQQMIFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKASLL  (new)
2786      ------------------------------------------------------STYYGREEELRKHPIKRAAEGWAAPDALLGQVKASLL   (ref 4)
2844      ------------------------------------------------------------------V---------------------- (ref 4)
```

Heterodimerization domain

```
HN1 1655 PGGSEGGRRRRELDPMDVRGSIVYLEIDNRQCVQASSQCFQSATDVAAFLGALASLGSLNIPYKIEAVQSETVEPPPPAQ1734
9577      -----------------I------------------------------------------------------------------  (ref 1,7)
1944      --------------------MF---------------------------------------------------------------  (ref 4)
8628      --------------------D----------------------------------------------------------------  (ref 1,2,5)
258       -------------------------P-----------------------------------------------------------  (ref 1,2,3,4,5)
2702      -------------------------P-----------------------------------------------------------  (ref 1,2,3,4,5)
2723      -------------------------P-----------------------------------------------------------  (ref 1,2,3,4,5)
2651      -------------------------Q-----------------------------------------------------------  (ref 3,4,5)
2691      -------------------------Q-----------------------------------------------------------  (ref 3,4,5)
491       ------------------------N------------------------------------------------------------  (ref 1,3)
1113      ----------------------------D-------------------------------------------------------  (new)
1955      ------------------------------------------------------------T---------------  (ref 4)
```

Juxtamembrane domain

```
HN1 1724 SETVEPPPPAQLHFMYVAAAAFVLLFFVGCGVLLSRKRRRQHGQLWFPEGFKVSEASKKKRREPLGEDSVGLK 1796
1570      --------------RTVEPPPPAQLHFMYVAAAAFVLLFFVGCGVLLSRKRRRQHGQLWFPEGFKVSEASKKK   (ref 8)
704       ----------------PPPPAQLHFMYVAAAAFVLLFFVGCGVLLSRKRRRQHGQLWFPEGFKVSEASKKKRR   (ref 8)
1946      -----------------GEPPPPAQLHFMYVAAAAFVLLFFVGCGVLLSRKRRRQHGQLWFPEGFKVSEASKK   (ref 8)
1950      -----------------AVEPPPPAQLHFMYVAAAAFVLLFFVGCGVLLSRKRRRQHGQLWFPEGFKVSEASK   (ref 8)
1815      -----------------EARQLHFMYVAAAAFVLLFFVGCGVLLSRKRRRQHGQLWFPEGFKVSEASKKKRR   (ref 8)
```

PEST domain

```
HN1 2299 VGGSTSLNGQCEWLSRLQSGMVPNQYNPLRGSVAPGPLSTQAPSLQHGMVGPLHSSLAASALSQMMSYQGLPSTRLATQPHLVQTGQVQPQN 2390
9194      ------------------------------------LPAAWHGHG*     (new)
9989      ------------------------------------LPAAWHGHG*     (new)
```


**Figure S1. *NOTCH1* mutations in pediatric T-ALL patients.** Amino acid changes in the HD, JM and PEST domains of NOTCH1, as a result of *NOTCH1* mutations, are listed for each patient. New mutations and the reference of each known mutation are indicated.

Known *NOTCH1* mutations as previous identified in the study of Weng *et al* (2004)[1], Breit *et al* (2006)[2], Zhu *et al* (2006)[3], Van Grotel *et al* (2008)[4], Asnafi *et al* (2009)[5], Park *et al* (2009)[6], Larson Gedman *et al* (2009)[7] and Suzuki *et al* (2009)[8].

References

1. Weng, A.P. et al. Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. *Science* 306, 269-71 (2004).
2. Breit, S. et al. Activating NOTCH1 mutations predict favorable early treatment response and long-term outcome in childhood precursor T-cell lymphoblastic leukemia. *Blood* 108, 1151-7 (2006).
3. Zhu, Y.M. et al. NOTCH1 mutations in T-cell acute lymphoblastic leukemia: prognostic significance and implication in multifactorial leukemogenesis. *Clin Cancer Res* 12, 3043-9 (2006).
4. van Grotel, M. et al. Prognostic significance of molecular-cytogenetic abnormalities in pediatric T-ALL is not explained by immunophenotypic differences. *Leukemia* 22, 124-31 (2008).
5. Asnafi, V. et al. NOTCH1/FBXW7 mutation identifies a large subgroup with favorable outcome in adult T-cell acute lymphoblastic leukemia (T-ALL): a Group for Research on Adult Acute Lymphoblastic Leukemia (GRAALL) study. *Blood* 113, 3918-24 (2009).
6. Park, M.J. et al. FBXW7 and NOTCH1 mutations in childhood T cell acute lymphoblastic leukaemia and T cell non-Hodgkin lymphoma. *Br J Haematol* 145, 198-206 (2009).
7. Larson Gedman, A. et al. The impact of NOTCH1, FBW7 and PTEN mutations on prognosis and downstream signaling in pediatric T-cell acute lymphoblastic leukemia: a report from the Children's Oncology Group. *Leukemia* 23, 1417-25 (2009).
8. Suzuki, S. et al. A second NOTCH1 chromosome rearrangement: t(9;14)(q34.3;q11.2) in T-cell neoplasia. *Leukemia* 23, 1003-6 (2009).

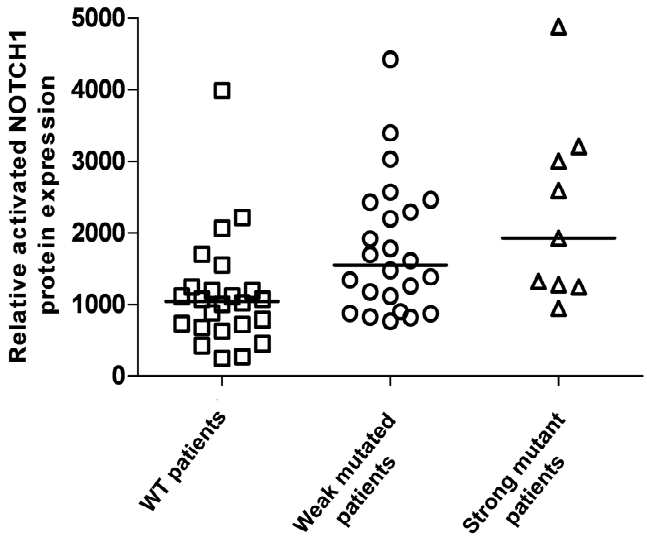



**Figure S3.** ICN levels in wild-type, weak NOTCH1-activated (single HD, single PEST, single FBXW7 mutation) and strong NOTCH1-activated (HD+PEST, HD+FBXW7, JM mutation) T-ALL patients analyzed with reverse-phase protein microarray.

```
WD40-repeats

   362 DTNWRRGELKSPKVLKGHDDHVITCLQFCGNRIVSGSDDNTLKVWSAVTGKCLRTLVGHTDGVWSSQMRDNIIISGSTD 440
3028   -----------------L------------------------------------------------------------ (new)
2691   ------------------------------------------------------------V----------------- (ref 39,50)

   441 RTLKVWNAETGECIHTLYGHTSTVRCMHLHEKRVVSGSRDATLRVWDIETGQCLHVLMGHVAAVRCVQYDGRRVVSGA  518
1179   ------------------------C------------------------------------------------------ (ref 1,2,3,4,5,6,7)
2650   ------------------------C------------------------------------------------------ (ref 1,2,3,4,5,6,7)
2698   ------------------------C------------------------------------------------------ (ref 1,2,3,4,5,6,7)
2720   ------------------------C------------------------------------------------------ (ref 1,2,3,4,5,6,7)
1179   ------------------------C------------------------------------------------------ (ref 1,2,3,4,5,6,7)
2650   ------------------------C------------------------------------------------------ (ref 1,2,3,4,5,6,7)
2698   ------------------------C------------------------------------------------------ (ref 1,2,3,4,5,6,7)
9928   ------------------------C------------------------------------------------------ (ref 1,2,3,4,5,6,7)
8639   ------------------------C------------------------------------------------------ (ref 1,2,3,4,5,6,7)
2780   ------------------------H------------------------------------------------------ (ref 1,2,3,4,5,7)
2788   ------------------------H------------------------------------------------------ (ref 1,2,3,4,5,7)
222    ------------------------H------------------------------------------------------ (ref 1,2,3,4,5,7)
8577   ------------------------H------------------------------------------------------ (ref 1,2,3,4,5,7)
9919   ------------------------H------------------------------------------------------ (ref 1,2,3,4,5,7)
9376   ------------------------P------------------------------------------------------ (new)
768    -------------------------------------Q------------------------------------- (ref 1,2,3,4,5,7)
2748   -------------------------------------Q------------------------------------- (ref 1,2,3,4,5,7)
2773   -------------------------------------Q------------------------------------- (ref 1,2,3,4,5,7)
9027   -------------------------------------Q------------------------------------- (ref 1,2,3,4,5,7)
9938   ------------------------------------------------------------C-------------- (ref 1,2,3,4,5,7)
3044   ------------------------------------------------------------C-------------- (ref 1,2,3,4,5,7)


   620 PNKHQSAVTCLQFNKNFVITSSDDGTVKLWDLKTGEFIRNLVTLESGGSGGVVWRIRASNTKLVCAVGSRNGTEETKLLVLDFDVDMK*  708
2854   --*          (new)
2460   ------------------------------------------------------------------------V---------------------- (new)
2782   ------------------------------------------------------------------------K--------------- (new)
```

**Figure S2. *FBXW7* mutations in pediatric T-ALL patients**. Amino acid changes in the WD40-repeats, as a result of *FBXW7* mutations, are listed for each patient. New mutations and the reference of each known mutation are indicated. Known FBXW7 mutations as previous identified in the studies of Thompson et al (2007)[1], O'Neil et al (2007)[2], Malyukova et al (2007)[3], Park et al (2009)[4], Asnafi et al (2009)[5], Larson et al (2009)[6] and Mansour et al (2009)[7].

References:


1. Thompson, B.J., et al., The SCFFBW7 ubiquitin ligase complex as a tumor suppressor in T cell leukemia. J Exp Med, 2007. 204(8): p. 1825-35.
2. O'Neil, J., et al., FBW7 mutations in leukemic cells mediate NOTCH pathway activation nd resistance to gamma-secretase inhibitors. J Exp Med, 2007. 204(8): p. 1813-24.
3. Malyukova, A., et al., The tumor suppressor gene hCDC4 is frequently mutated in human T-cell acute lymphoblastic leukemia with functional consequences for Notch signaling. Cancer Res, 2007. 67(12): p. 5611-6.
4. Park, M.J., et al., FBXW7 and NOTCH1 mutations in childhood T cell acute lymphoblastic leukaemia and T cell non-Hodgkin lymphoma. Br J Haematol, 2009. 145(2): p. 198-206.
5. Asnafi, V., et al., NOTCH1/FBXW7 mutation identifies a large subgroup with favorable outcome in adult T-cell acute lymphoblastic leukemia (T-ALL): a Group for Research on Adult Acute Lymphoblastic Leukemia (GRAALL) study. Blood, 2009. 113(17): p. 3918-24.
6. Larson Gedman, A., et al., The impact of NOTCH1, FBW7 and PTEN mutations on prognosis and downstream signaling in pediatric T-cell acute lymphoblastic leukemia: a report from the Children's Oncology Group. Leukemia, 2009. 23(8): p. 1417-25.
7. Mansour, M.R., et al., Prognostic implications of NOTCH1 and FBXW7 mutations in adults with T-cell acute lymphoblastic leukemia treated on the MRC UKALLXII/ECOG E2993 protocol. J Clin Oncol, 2009. 27(26): p. 4352-6.

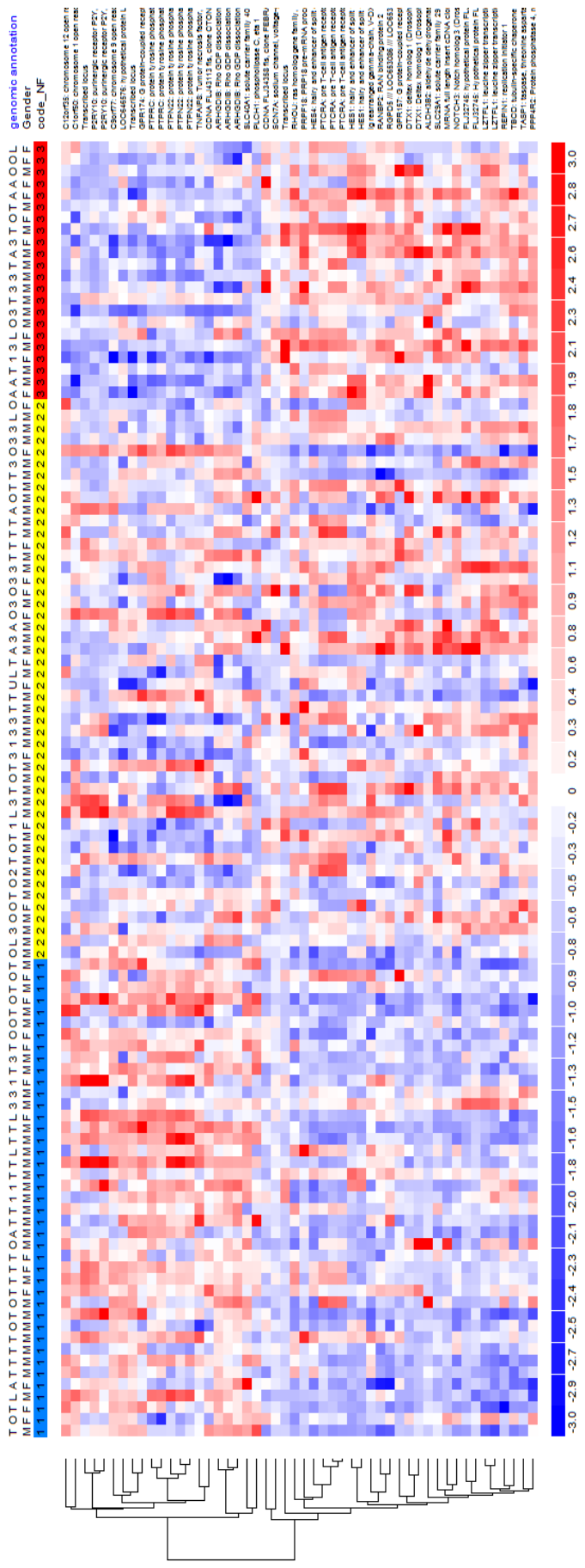



Figure S4. Heatmap showing the TOP50 most differentially expressed genes between patients with strong NOTCH1-activating mutations versus wild-type patients. NOTCH1 JM mutations or combinations of NOTCH1 HD mutations with PEST mutations or FBXW7 mutations were considered as strong NOTCH1-activating mutations. Annotations indicated are genetic rearrangements, gender and NOTCH1/FBXW7 mutation status. Genetic rearrangements indicated are: T, TAL1 or SIL-TAL1; L, LMO1 or LMO2 (includes del(11)(p12p13)); A, HOXA-activated (includes cases with SET-NUP214; CALM-AF10 or Inv(7)(p15q34)); 1, TLX1; 2, TLX2; 3, TLX3; O, Other; U, Aberration unknown. Gender is indicated F, Female or M, Male. NOTCH1/FBXW7 mutation status is indicated "1" in blue box for wild-type, "2" in yellow box for patients with weak NOTCH1-activating0 activating mutations and "3" in red box for patients having strong NOTCH1-activating mutations. mutations and "3" in red box for patients having strong NOTCH1-activating mutations.

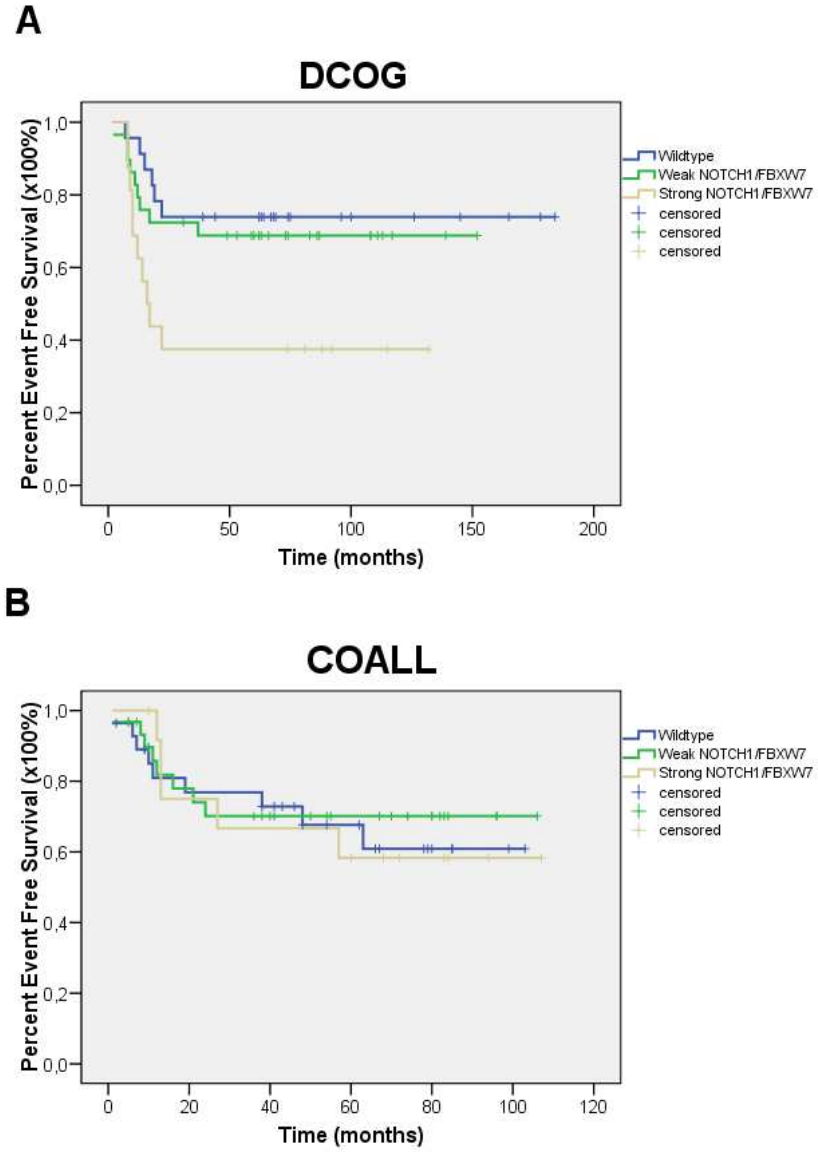
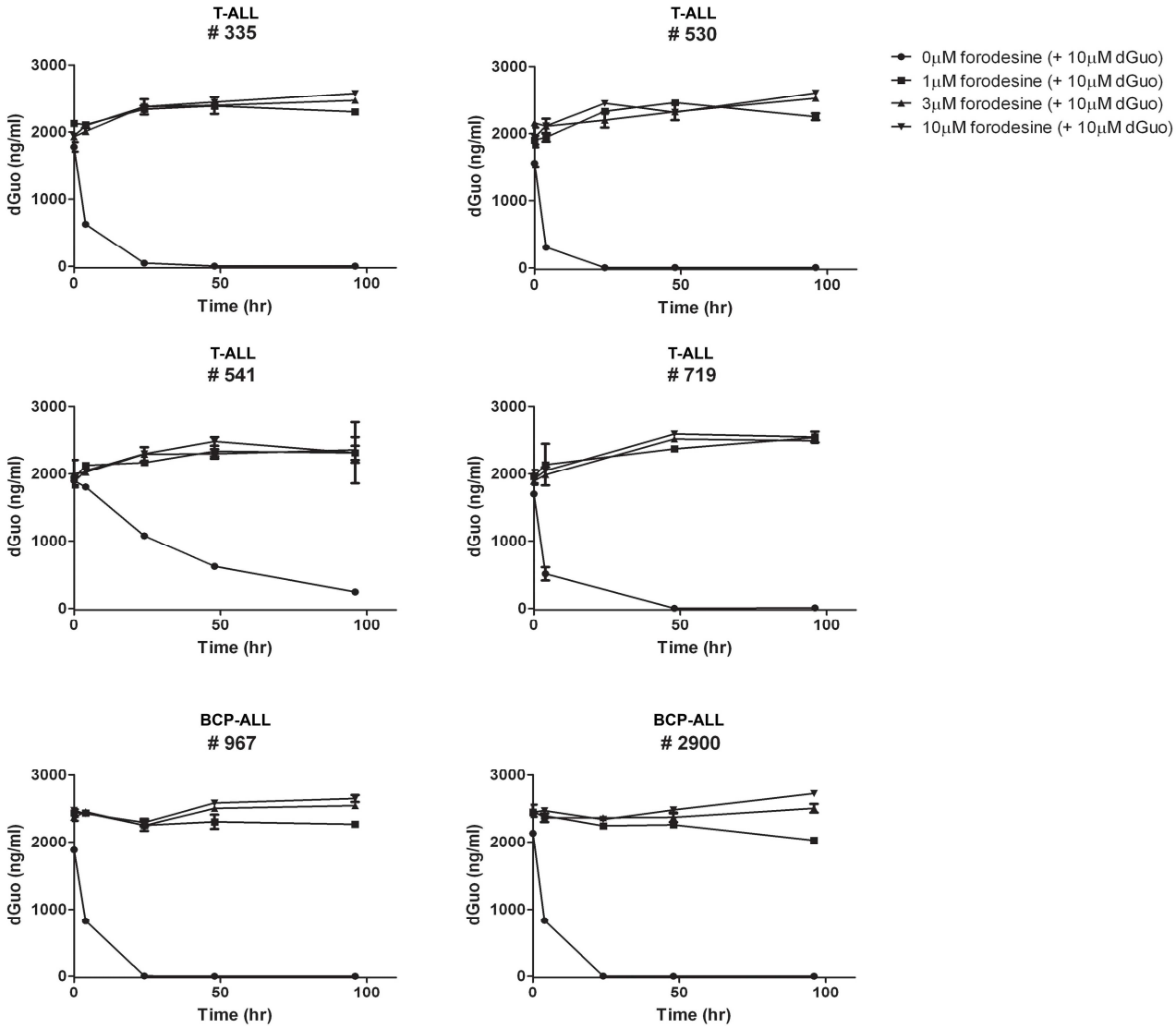



**Figure S5. The prognostic effect of *NOTCH1/FBXW7* mutations**. Weak NOTCH1-activating mutations were considered as *NOTCH1* HD, PEST or FBXW7 mutations, whereas strong NOTCH1-activating mutations were considered as *NOTCH1* JM mutations or NOTCH1 HD mutations in combination with PEST or *FBXW7* mutations. **A.** For DCOG T-ALL patients, strong NOTCH1-activating mutations are significantly associated with poor outcome, with *p*-values of $p=0.012$ and $p=0.048$, compared to patients without *NOTCH1*/*FBXW7* mutations (wild-type) or patients with weak NOTCH1-activating mutations, respectively. **B**. No significant association with poor outcome was observed for weak or strong-activated NOTCH1 T-ALL patients treated according to the COALL-97 protocol.

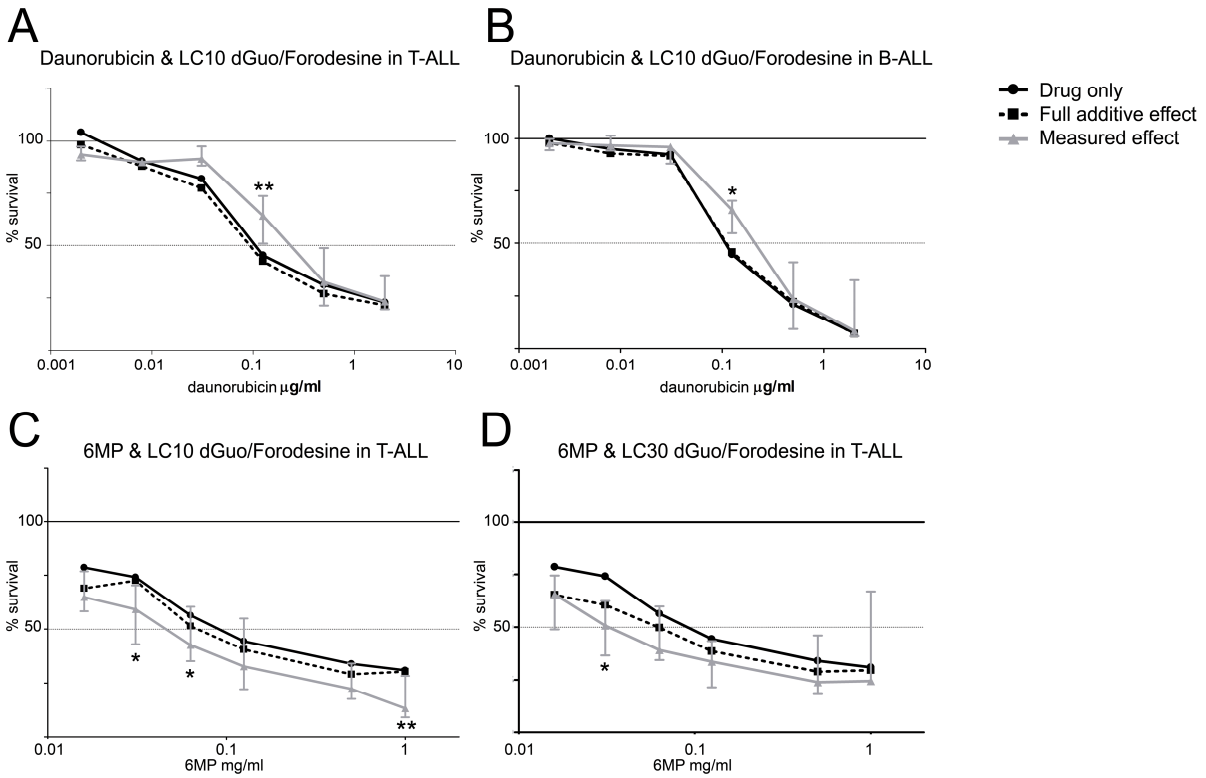
# CHAPTER 6

**Supplementary Table 1. Primer sequences used for RQ-PCR**

| Gene | Forward primer | Reverse primer |
|---|---|---|
| equilibrative nucleoside transporter 2 (*ENT2/SLC29A2)* | 5'-ACA GGG CAG CCT CTT C-3' | 5'-TAG CGG GCA AAC TTC A-3' |
| cytosolic 5' nucleotidase 2 *(NT5C2/NT5B/PNT5)* | 5'-TAT GCC TGC TAA CAT GGA T-3' | 5'-ACC AAG GGA CTC ATA CTC TG-3' |
| *PNP (NP*) | 5'-TCC CCG AAG TAC AGT GC-3' | 5'-GGG TTC TGA CCA CTG AAA C-3' |
| deoxyguanine kinase *(dGK/dGUOK)* | 5'-AGG CTC TGA TGA ACA TTC C-3' | 5'-AAC AAT GGC AAA GTC TAA CAA-3' |
| concentrative nucleoside transporter 1 *(CNT1/SLC28A1)* | 5'-TGG ATG CTG ACA GAA ACA-3' | 5'-CTC CAG CTG CTC CTG AT-3' |
| concentrative nucleoside transporter 2 *(CNT2/SLC28A2*) | 5'-AGC TGG GTT GAG GAG AAC-3' | 5'-AAG CTG GCG TGT GTT TT-3' |
| concentrative nucleoside transporter 3 (*CNT3/SLC28A3*) | 5'-CCC AGG TCC CTG TAA CA-3' | 5'- TGT GTG CTC CCT GCT T-3' |
| cytosolic 5' nucleotidase 1A (*NT5C1A/CN1A)* | 5'-GGA GGA AGC CAA GAT TTT-3' | 5'-CTG AAG GGT TCG TTC TCA-3' |




**Supplementary Figure 1. dGuo degradation for different concentrations of forodesine**. DGuo concentrations in culture supernatants at different time points (x-axis) following dGuo (1μM) or forodesine (0, 1, 3 or 10 μM) + dGuo (10μM) administration for 4 T-ALL and 2 BCP-ALL primary diagnostic patient samples.




**Supplementary Figure 2. Combination studies**. Median survival values for drug only (solid black line), the calculated full additive effect of the drug and the LC10 or LC30 of forodesine/dGuo (dotted black line) and the measured effect (solid grey line) on survival of the combination is depicted for daunorubicin and the LC10 of forodesine/dGuo in T-ALL samples (A) and BCP-ALL samples (B) and for 6MP for T-ALL samples and LC10 (C) or LC30 (D) values of forodesine/dGuo. Vertical grey lines represent interquartile ranges.