Detection and characterization of PAX5 aberrations in childhood acute lymphoblastic leukemia

Dissertation

angestrebter akademischer Grad

Doktorin der Naturwissenschaften (Dr. rer.nat.)

Verfasserin / Verfasser: Karin Nebral
Matrikel-Nummer: 9604391
Dissertationsgebiet (lt. Studienblatt): 442 Anthropologie
Betreuerin / Betreuer: Univ.-Prof. Dr. Oskar A. Haas

Wien, am 04. November 2008
<table>
<thead>
<tr>
<th>Contents</th>
<th>pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbreviations</td>
<td>3</td>
</tr>
<tr>
<td>Abstract</td>
<td>6</td>
</tr>
<tr>
<td>Chapter 1</td>
<td>8</td>
</tr>
<tr>
<td>1. Introduction</td>
<td>8</td>
</tr>
<tr>
<td>1.1. Hematopoiesis</td>
<td>8</td>
</tr>
<tr>
<td>1.2. Lymphoid development</td>
<td>10</td>
</tr>
<tr>
<td>1.2.1. B-cell development and the role of PAX5</td>
<td>11</td>
</tr>
<tr>
<td>1.3. Acute lymphoblastic leukemia</td>
<td>15</td>
</tr>
<tr>
<td>1.4. The PAX gene family in oncogenesis</td>
<td>17</td>
</tr>
<tr>
<td>1.5. PAX5 aberrations in B-cell malignancies</td>
<td>20</td>
</tr>
<tr>
<td>1.5.1. PAX5 aberrations in lymphoma</td>
<td>20</td>
</tr>
<tr>
<td>1.5.2. PAX5 aberrations in acute lymphoblastic leukemia</td>
<td>21</td>
</tr>
<tr>
<td>1.5.2.1. Deletions and Mutations</td>
<td>21</td>
</tr>
<tr>
<td>1.5.2.2. PAX5 fusions in ALL</td>
<td>23</td>
</tr>
<tr>
<td>Chapter 2</td>
<td>25</td>
</tr>
<tr>
<td>Identification of PML as novel PAX5 fusion partner in childhood acute</td>
<td></td>
</tr>
<tr>
<td>lymphoblastic leukemia</td>
<td></td>
</tr>
<tr>
<td>Chapter 3</td>
<td>39</td>
</tr>
<tr>
<td>Incidence and diversity of PAX5 fusion genes in childhood acute</td>
<td></td>
</tr>
<tr>
<td>lymphoblastic leukemia</td>
<td></td>
</tr>
<tr>
<td>Chapter 4</td>
<td>65</td>
</tr>
<tr>
<td>Monoallelic loss and frequent mutation of the second allele of PAX5 in</td>
<td></td>
</tr>
<tr>
<td>dic(9;20) childhood acute lymphoblastic leukemia</td>
<td></td>
</tr>
<tr>
<td>Chapter 5</td>
<td>74</td>
</tr>
<tr>
<td>5. Additional Methods and Results</td>
<td>74</td>
</tr>
<tr>
<td>5.1. Establishment and validation of automated FISH screening</td>
<td>74</td>
</tr>
<tr>
<td>5.1.1. Analysis of control samples</td>
<td>80</td>
</tr>
<tr>
<td>5.1.2. Validation of the FISH assay in PAX5-rearranged samples</td>
<td>82</td>
</tr>
<tr>
<td>5.1.3. Novel PAX5 positive cases - Metafer 4 Metacyte analysis</td>
<td>83</td>
</tr>
</tbody>
</table>
5.2. Detailed analysis of selected cases
   5.2.1. PAX5 rearranged cases
   5.2.2. PAX5 deletions

<table>
<thead>
<tr>
<th>Chapter 6</th>
<th>101</th>
</tr>
</thead>
<tbody>
<tr>
<td>6. Discussion</td>
<td>101</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 7</th>
<th>106</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. References</td>
<td>106</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 8</th>
<th>113</th>
</tr>
</thead>
<tbody>
<tr>
<td>8. Appendix</td>
<td>113</td>
</tr>
</tbody>
</table>

| 8.1. ETV6-NCOA2 fusion defines a new entity of T-lymphoid/myeloid progenitor acute leukemia | 113 |
| 8.2. Additional FISH clones | 121 |
| 8.3. Additional PCR primers | 123 |

<table>
<thead>
<tr>
<th>Deutsche Zusammenfassung</th>
<th>125</th>
</tr>
</thead>
<tbody>
<tr>
<td>Curriculum Vitae</td>
<td>127</td>
</tr>
<tr>
<td>Danksagung</td>
<td>129</td>
</tr>
</tbody>
</table>
ABBREVIATIONS

ALL  acute lymphoblastic leukemia
BAC  bacterial artificial chromosome
BCP  B-cell precursor
BCR  B-cell receptor
CLP  common lymphoid progenitor
CML  chronic myelogenous leukemia
CMLP common myelo-lymphoid progenitor
CMP  common myeloid progenitor
CY3  cyanine 3
DLBCL diffuse large B-cell lymphoma
EFS  event-free survival
ELP  earliest lymphocyte progenitor
FAB French-American-British
FISH fluorescence in situ hybridization
FITC fluorescein isothiocyanate
GC B-cell germinal center B-cells
GMP  granulocyte/macrophage progenitor
HLH helix–loop–helix
HSC  hematopoietic stem cell
Ig  immunoglobulin
LIN  lineage
LMPP lymphoid-primed multipotent progenitor
LSK LIN⁺SCA1⁺KIT⁺
MkEP  megakaryocyte/erythroid progenitor
MPP  multipotent progenitor
RACE rapid amplification of cDNA ends
RT-PCR reverse-transcription polymerase chain reaction
SAGA Spt–Ada–Gcn5 acetyltransferase
shRNA short hairpin RNA
SNP  single nucleotide polymorphism
SSP DNA salmon sperm DNA
Human and mouse gene nomenclature according to the HUGO Gene Nomenclature Committee (http://www.genenames.org) and the Mouse Genome Informatics (http://www.informatics.jax.org/), respectively. Mouse protein nomenclature according to The Universal Protein Resource (UniProt) (http://www.uniprot.org/uniprot).

Human gene symbols are **ITALICIZED**, with all letters UPPERCASE, whereas mouse gene symbols are written in *Italics* with the first letter capitalized followed by lowercase letters. Protein designations are the same as the gene symbol, but NOT ITALICIZED and all letters UPPERCASE. However, to differentiate between human and mouse proteins they are often indicated in UPPERCASE and Lowercase, respectively [e.g. PAX5 (human) and Pax5 (mouse)].

**Genes (human)**

- **ABL1**: c-abl oncogene 1, receptor tyrosine kinase
- **ATXN1 (SCA1)**: ataxin 1
- **AUTS2**: autism susceptibility candidate 2
- **BCL11B**: B-cell CLL/lymphoma 11B (zinc finger protein)
- **BCR**: breakpoint cluster region
- **BDNF**: brain-derived neurotrophic factor
- **BLNK**: B-cell linker
- **BRD1**: bromodomain containing 1
- **BTG1**: B-cell translocation gene 1, anti-proliferative
- **C20orf112**: chromosome 20 open reading frame 112
- **CCR2**: chemokine (C-C motif) receptor 2
- **CD19**: CD19 molecule
- **CD22**: CD22 molecule
- **CD28**: CD28 molecule
- **CD72**: CD72 molecule
- **CD79A**: CD79a molecule, immunoglobulin-associated alpha
- **CD79B**: CD79b molecule, immunoglobulin-associated beta
- **CDKN2A**: cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)
- **CDKN2B**: cyclin-dependent kinase inhibitor 2B (p15, inhibits CDK4)
- **CEBPA (C/EBPα)**: CCAAT/enhancer binding protein (C/EBP), alpha
- **CR2 (CD21)**: complement component (3d/Epstein Barr virus) receptor 2
- **CREBBP (CBP)**: CREB binding protein
- **CSF1R (GM-CSFRα)**: colony stimulating factor 1 receptor
- **DACH1**: dachshund homolog 1
- **EBF1**: early B-cell factor 1
- **ELN**: elastin
- **ERG**: v-ets erythroblastosis virus E26 oncogene homolog (avian)
- **ETS1**: v-ets erythroblastosis virus E26 oncogene homolog 1 (avian)
- **ETV6 (TEL)**: ets variant gene 6 (TEL oncogene)
- **FLT3**: FMS-related tyrosine kinase 3
- **FOXO1 (FKHR)**: forkhead box O1
- **FOXO1**: forkhead box P1
- **GATA1**: GATA binding protein 1 (globin transcription factor 1)
- **GATA3**: GATA binding protein 3
- **HIPK1**: homeodomain interacting protein kinase 1
- **IGH@ (Igκ)**: immunoglobulin heavy locus
- **IGJ**: immunoglobulin J polypeptide, linker protein for immunoglobulin alpha and mu polypeptides
- **IGK@ (Igκ)**: immunoglobulin kappa locus
- **IgH**: immunoglobulin lambda-like polypeptide 1
- **IKZF1 (Ikaros)**: IKAROS family zinc finger 1 (Ikaros)
- **IKZF3 (Aiolos)**: IKAROS family zinc finger 3 (Aiolos)
- **IL7**: interleukin 7
- **IL7R (IL-7Rα)**: interleukin 7 receptor
- **IRF4**: interferon regulatory factor 4
<table>
<thead>
<tr>
<th>Gene/Protein</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRF8</td>
<td>interferon regulatory factor 8</td>
</tr>
<tr>
<td>JAK2</td>
<td>janus kinase 2</td>
</tr>
<tr>
<td>KIT (C-Kit, CD117)</td>
<td>v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog</td>
</tr>
<tr>
<td>LEF1</td>
<td>lymphoid enhancer-binding factor 1</td>
</tr>
<tr>
<td>LMO1</td>
<td>LIM domain only 1 (rhombotin 1)</td>
</tr>
<tr>
<td>LMO2</td>
<td>LIM domain only 2 (rhombotin-like 1)</td>
</tr>
<tr>
<td>MLL</td>
<td>myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)</td>
</tr>
<tr>
<td>MYC (c-Myc)</td>
<td>v-myb myelocytomatosis viral oncogene homolog (avian)</td>
</tr>
<tr>
<td>NOTCH1</td>
<td>Notch homolog 1, translocation-associated (Drosophila)</td>
</tr>
<tr>
<td>NUP98</td>
<td>nucleoporin 98kDa</td>
</tr>
<tr>
<td>PAX5</td>
<td>paired box 5</td>
</tr>
<tr>
<td>PBX1</td>
<td>pre-B-cell leukemia homeobox 1</td>
</tr>
<tr>
<td>PLCG2 (PLCγ2)</td>
<td>phospholipase C, gamma 2 (phosphatidylinositol-specific)</td>
</tr>
<tr>
<td>PML</td>
<td>promyelocytic leukemia</td>
</tr>
<tr>
<td>POM121</td>
<td>POM121 membrane glycoprotein</td>
</tr>
<tr>
<td>PPARγ (PPARγ1)</td>
<td>peroxisome proliferator-activated receptor gamma</td>
</tr>
<tr>
<td>PRDM1 (BLIMP1)</td>
<td>PR domain containing 1, with ZNF domain</td>
</tr>
<tr>
<td>PRKCB (PKCβ)</td>
<td>protein kinase C, beta</td>
</tr>
<tr>
<td>protein: BSAP, B-cell lineage specific activator protein</td>
<td></td>
</tr>
<tr>
<td>RAG1</td>
<td>recombination-activation gene 1</td>
</tr>
<tr>
<td>RAG2</td>
<td>recombination activating gene 2</td>
</tr>
<tr>
<td>RB1</td>
<td>retinoblastoma 1</td>
</tr>
<tr>
<td>RCSD1</td>
<td>RCSD domain containing 1</td>
</tr>
<tr>
<td>RUNX1 (AML1)</td>
<td>runt-related transcription factor 1</td>
</tr>
<tr>
<td>RUNX1T1 (ETO)</td>
<td>runt-related transcription factor 1; translocated to, 1 (cyclin D-related)</td>
</tr>
<tr>
<td>SCA1</td>
<td>stem cell antigen 1</td>
</tr>
<tr>
<td>SCL (TAL1)</td>
<td>T-cell acute lymphocytic leukemia 1</td>
</tr>
<tr>
<td>SPI1 (PU.1)</td>
<td>spleen focus forming virus (SFFV) proviral integration oncogene sp1</td>
</tr>
<tr>
<td>SPIB</td>
<td>Spi-B transcription factor (Spi-1/PU.1 related)</td>
</tr>
<tr>
<td>STAT5</td>
<td>signal transducer and activator of transcription 5A</td>
</tr>
<tr>
<td>TCF3 (E2A)</td>
<td>transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47)</td>
</tr>
<tr>
<td>THY-1 (CD90)</td>
<td>Thy-1 cell surface antigen</td>
</tr>
<tr>
<td>TP53 (p53)</td>
<td>tumor protein p53</td>
</tr>
<tr>
<td>VCAM1</td>
<td>vascular cell-adhesion molecule 1</td>
</tr>
<tr>
<td>VPREB1 (VpreB)</td>
<td>pre-B lymphocyte gene 1</td>
</tr>
<tr>
<td>WT1</td>
<td>Wilms tumor 1</td>
</tr>
<tr>
<td>XBP1</td>
<td>X-box binding protein 1</td>
</tr>
<tr>
<td>ZCCHC7</td>
<td>zinc finger, CCHC domain containing 7</td>
</tr>
<tr>
<td>ZNF521</td>
<td>zinc finger protein 521</td>
</tr>
</tbody>
</table>

**Mouse genes & proteins**

<table>
<thead>
<tr>
<th>Gene/Protein</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ets1</td>
<td>E26 avian leukemia oncogene 1, 5' domain</td>
</tr>
<tr>
<td>protein: Protein C-ets-1</td>
<td></td>
</tr>
<tr>
<td>Igh</td>
<td>immunoglobulin heavy chain complex</td>
</tr>
<tr>
<td>Ikzf1 (Ikaros)</td>
<td>IKAROS family zinc finger 1</td>
</tr>
<tr>
<td>Ikzf3 (Aiolos)</td>
<td>IKAROS family zinc finger 3</td>
</tr>
<tr>
<td>Plcg2</td>
<td>phospholipase C, gamma 2</td>
</tr>
<tr>
<td>Prkcb (Prkcb1)</td>
<td>protein kinase C, beta</td>
</tr>
<tr>
<td>Sfpi1 (PU.1)</td>
<td>SFFV proviral integration 1</td>
</tr>
<tr>
<td>Tcfe2a (E2A, TCF3)</td>
<td>transcription factor E2a</td>
</tr>
<tr>
<td>protein: Transcription factor E2-alpha</td>
<td></td>
</tr>
</tbody>
</table>
ABSTRACT

PAX5, a master regulator of B-cell development, was recently shown to be involved in several B-cell malignancy-associated genetic alterations, such as point mutations, deletions, and, of particular interest in the context of my work, also gene rearrangements. In B-cell non-Hodgkin's-Lymphoma with a t(9;14)(p13;q32), for instance, PAX5 is juxtaposed to the IGH@ locus, which results in an inappropriate over-expression of PAX5. In B-cell precursor acute lymphoblastic leukemia (BCP-ALL), on the other hand, PAX5 can fuse to several different partner genes such as FOXP1 (3p13), AUTS2 (7q11), ELN (7q11), ETV6 (12p13), ZNF521 (18q11), and C20orf112 (20q11), thereby generating fusion transcripts that encode chimeric proteins. Therefore, the aim of this study was to screen childhood ALL samples for PAX5 rearrangements and to determine their incidence and the types of PAX5 gene fusions in a systematic and population-based fashion.

To identify all potential PAX5-affecting breakpoints, including even those that result in juxtaposition of PAX5 under the regulatory elements of a partner gene, a novel dual-color split-apart fluorescence in situ hybridization (FISH) assay with BAC clones flanking the PAX5 gene was employed. All samples with suspicious FISH patterns were further analyzed with PAX5 exon-specific cosmid clones. In order to facilitate high-throughput screening, interphase FISH analysis was performed using an automated spot counting system (Metafer4-Metacyte, Metasystems). Novel fusion partners were identified by FISH, 3'- or 5'-RACE (Rapid Amplification of cDNA ends) and the presence of specific hybrid transcripts was verified by RT-PCR (Reverse Transcription-PCR) and sequence analysis.

A PAX5 rearrangement-indicating FISH pattern was observed in 10 (2.2%) of 446 children with de novo ALL registered in the Austrian ALL-BFM 2000 and Interfant-99 studies. Out of these 10 patients, one case was previously shown to harbor a PAX5-ETV6 fusion, in one the recently described PAX5-C20orf112 gene fusion was found and, despite all efforts, the fusion partner remained unidentified in another one. However, in seven cases we succeeded to identify six new PAX5 in-frame fusions with HIPK1 (1p13), POM121 (7q11), JAK2 (9p24), DACH1 (13q21), PML (15q24), or BRD1 (22q13.33). Apart from two PAX5-JAK2-positive cases, every other fusion gene was non-recurring. Moreover, at least in childhood ALL we did not find any evidence for PAX5 activating translocations. Given the large variety of recovered PAX5 fusion partners, our custom-made dual-color/two-step FISH screening approach has proven to be an appropriate and efficient tool for the reliable detection of PAX5 gene fusions.

The results of this study show that PAX5 rearrangements occur with a frequency of approximately 2.5% exclusively in BCP-ALL and fuse PAX5 to a broad range of different partner genes comprising transcription factors, structural proteins, and even a tyrosine kinase. All hypothetical fusion proteins retain at least the PAX5 paired DNA-binding domain,
which is joined to the C-terminal region or even the entire protein of the fusion partner. Thus, all PAX5 chimeric proteins are predicted to retain the ability to bind to PAX5 target genes suggesting that they act as aberrant transcription factors, which may antagonize intrinsic PAX5 transcriptional activity, and hence, may contribute to the pathogenesis of BCP-ALL.
CHAPTER 1

1. INTRODUCTION

1.1. Hematopoiesis

Hematopoiesis, the formation of the blood cellular components, is initiated in the fetal liver and postnatal bone marrow and all types of blood cells derive from hematopoietic stem cells (HSCs). Extensive proliferation occurs during differentiation into all lineages, which comprise the myeloid (including macrophages, granulocytes and polymorphonuclear cells such as neutrophils, basophils and eosinophils), the erythroid-megakaryocyte (including erythrocyte and platelets derived from megakaryocytes), and the lymphoid (B-cells, T-cells, natural killer cells, and dendritic cells) lineage (Fig. 1) (Katsura, 2002).

Figure 1. Hematopoietic cell classifications. Figure taken from Katsura, 2002.
The lineage differentiation processes during hematopoiesis are still under debate and currently various models are proposed, which almost all are based on experimental evidence of mice studies. The classic paradigm implies that blood-cell formation proceeds along an ordered pathway with binary decision steps, and that a single route is given for each major cell type. However, recent findings suggest that more dynamic alternative developmental pathways are generating myeloid and lymphoid cells.

Stem cells and all early progenitors are found within the LIN- SCA1+ KIThi subset (LSK subset) and are defined by the capacity of self-renewal and the capability of differentiation into all hematopoietic cell lineages. In one of the models, HSCs are subdivided into long-term HSCs (Thy-1low Flt3-), short-term HSCs (Thy-1low Flt3+), and multipotent progenitors (MPPs) (Thy-1 Flt3+), which differentiate into common lymphoid progenitors (CLP) as well as common myeloid progenitors (CMP) that generate the granulocytic-macrophage (GM) and the megakaryocytic-erythroid (MegE) lineages (Fig. 2A) (Kondo et al., 1997; Akashi et al., 2000; Laiosa et al., 2006). Short-term HSCs and MPPs sustain the full lympho-myeloid lineage potential of long-term HSCs, but have reduced self renewal capacity, which coincides with Flt3 expression. This concept implicates that the first lineage commitment step of adult HSCs results in an immediate and complete separation of myelopoiesis and lymphopoiesis (Fig. 2A) (Adolfsson et al., 2005; Welner et al., 2008).

In contrast, Katsura and co-workers detected bipotential myeloid/T-cell (p-MT) and myeloid/B-cell (p-MB) progenitor stages and proposed a model, in which T- and B-cells arise from a multipotent myeloid/T/B-cell progenitor and are produced through intermediate p-MT and p-MB stages. The existence of a common myelo-lymphoid progenitor (CMLP or p-MTP) also indicates that the erythroid potential is shut off at an early stage before branching towards T and B progenitors (Katsura, 2002). Furthermore, Adolfsson et al demonstrated that LSK Flt3+ HSCs sustain granulocyte, monocyte, and B- and T-cell potentials but lack a megakaryocytic-erythroid lineage potential (lymphoid-primed multipotent progenitors, LMPPs) revising the generally accepted concept of hematopoiesis (Fig. 2B) (Adolfsson et al., 2005). A summary of the currently existing models and a potential composite model as well as the expression of several key factors during lineage differentiation (described below) is shown in Figure 2.

Although at low levels, hematopoietic multipotential progenitors (MPPs) and HSCs promiscuously express genes of disparate lineages. This phenomenon, which is termed lineage priming, suggests that the fate of immature cells is not predetermined and that lineage selection extinguishes alternative potentials (Miyamoto et al., 2002; Laiosa et al., 2006; Mansson et al., 2007; Orkin & Zon, 2008). The coexistence of different transcriptional programs in progenitor cells, followed by the stepwise extinction of all except for one of them, is therefore a defining feature of the hematopoietic system (Laiosa et al., 2006). It was shown,
that hematopoietic progenitors as well as differentiated cells can be redirected into other lineages upon ectopic cytokine signaling or forced expression of lineage-specific transcription factors (Laiosa et al., 2006), which also challenges the model of an unidirectional differentiation process.

The important transcription factors implicated in early hematopoietic development and formation of HSCs include GATA-1, SPI1 (PU.1), CEBPA (C/EBPα), NOTCH1, and GATA-3 (Laiosa et al., 2006) as well as RUNX1, MLL, SCL (TAL-1), LMO2, and ETV6 (Orkin & Zon, 2008). Moreover, many genes commonly expressed in T- and B-cells are not active in HSCs, thus, HSCs more closely resemble myeloid than lymphoid precursors (Laiosa et al., 2006; Welner et al., 2008).

Figure 2. Models for Hematopoietic Stem Cell and Blood Lineage Commitment (Adolfsson et al., 2005). (A) Model supported through the identification of CMPs and CLPs (Kondo et al., 1997; Akashi et al., 2000). (B) Alternative model by Adolfsson et al., 2005. (C) Composite model incorporating experimental evidence for models (A) and (B). LT-HSC, long-term hematopoietic stem cell; ST-HSC, short-term hematopoietic stem cell; MPP, multipotent progenitor; LMPP, lymphoid-primed multipotent progenitor; CLP, common lymphoid progenitor; CMP, common myeloid progenitor; GMP, granulocyte/macrophage progenitor; MkEP, megakaryocyte/erythroid progenitor; B, B-cell; T, T-cell. Figure taken from Adolfsson et al., 2005.

1.2. Lymphoid development

One of the key players for initiation of lymphoid development is FLT3, which is progressively upregulated from long-term HSCs towards MPPs, accompanied by a loss of VCAM1 expression. In addition to high FLT3 expression, RAG1 is activated in LMPPs characterizing the emergence of the earliest lymphocyte progenitors (ELPs) (Welner et al., 2008). ELPs are the precursors of CLPs, which give rise to B- and T-lymphocytes, natural killer cells, and dendritic cells. In early lymphoid development important factors are IKZF1 (Ikaros) and SPI1.
Ikaros-deficient mice lack ELPs, but also alterations in the myeloid lineages have been observed, which might be explained by the expression of distinct Ikaros isoforms in different lineages and cooperation with the lymphoid-restricted Ikaros family member Aiolos (Laiosa et al., 2006; Welner et al., 2008). Ikaros may activate Flt3 and repress CSF1R (GM-CSFRα) and, thus, promote lymphoid cell fate. In contrast, loss of SPI1 (PU.1) inhibits B lineage and myelomonocytic cell formation as well as T-cell and dendritic cell formation, and PU.1-deficient mice lack expression of IL7R (IL-7Rα) and EBF1 (Laiosa et al., 2006).

GATA-3 and NOTCH1 drive lymphoid progenitors versus a T lineage differentiation, whereas entry of CLPs into the B-cell lineage critically depends on signaling of the IL7-receptor (IL7R), as well as expression of the transcription factors TCF3 (E2A), EBF1 and PAX5 (Laiosa et al., 2006).

1.2.1. B-cell development and the role of PAX5

The hallmark of B-cell development is the stepwise expression and assembly of components of the functional receptor for antigen, the B-cell receptor (BCR). Assembly of the pre-BCR requires the rearrangement of immunoglobulin heavy chain (IGH@) genes, which proceeds in two steps: (1) diversity (D) and joining (J) segments are assembled (D-J rearrangements), and (2) variable regions are joined to D-J segments to create mature V(D)J joints (Maier & Hagman, 2002). Additionally, the pre-BCR complex is composed of the surrogate light chains Igll1 (λ5) and VPREB1 (VpreB) as well as of the signal-transducing proteins CD79A (Igα) and CD79B (Igβ) (Schebesta et al., 2002).

The helix-loop-helix transcription factors TCF3 (E2A) and EBF1 coordinately regulate B-cell specific genes, such as components of the BCR, and in the absence of TCF3 (E2A) or EBF1 B-cell differentiation is blocked at the uncommitted pre-pro-B-cell stage (the first identifiable B-cell-specified progenitor stage arising from CLPs) (O’Riordan & Grosschedl, 1999; Nutt & Kee, 2007). Moreover, E2A-deficient CLPs and pre-pro-B-cells fail to undergo D_JH rearrangements at the Igh and V_{κ}-J_κ recombination at the Igk (Igκ) locus (Kwon et al., 2008). E2A activates Ebf1, which in turn activates Pax5 and, thus, promotes the B-cell transcriptional program (Nutt & Kee, 2007). Furthermore, E2A is required to maintain the expression of Ebf1, Pax5, and the B-cell gene program in pro-B-cells, whereas it is largely dispensable for the formation and function of mature B-lymphocytes and plasma cells (Kwon et al., 2008).

Ebf1-deficient mice lack expression of most B-cell genes including Cd79a (mb-1, Igα), Cd79b (Igβ, B29), Igll1(λ5), Vpreb1 (VpreB), and Rag1, and do not undergo any Igh gene recombinations (Lin & Grosschedl, 1995). Ebf1 is controlled through two promoters, a distal
(α) and a proximal (β). The Ebf1α promoter is regulated by TCF3 (E2A) and STAT5 (which is activated by IL7R signaling) and, intriguingly, also contains an EBF1 binding site suggesting an autoregulatory function for EBF1. The Ebf1β promoter is controlled by Ets1, PU.1 and Pax5, whose expression itself is dependent on EBF1 (Roessler et al, 2007). Thus, EBF1 regulates its own expression directly through induction of the Ebf1α promoter and indirectly through upregulation of Pax5 (Nutt & Kee, 2007; Roessler et al, 2007). The synergistic activity of E2A and EBF1 is required for B-cell specification, and both indirectly control B-cell development by induction of the B-cell commitment factor Pax5 (Nutt & Kee, 2007). However, activation of the B-cell specific transcription program is not sufficient to commit early progenitors to the B-lymphoid lineage without Pax5.

In the absence of Pax5, B-cell development is arrested at the early pro-B (pre-BI) cell stage, which is characterized by the expression of early B-cell markers and a markedly reduced frequency of VH-to-DHJH immunoglobulin rearrangements (Nutt et al, 1997). Pax5−/− pro-B-cells are not committed to the B-lymphoid lineage yet, they can be cultivated indefinitely in vitro in the presence of interleukin 7 and stroma but are unable to differentiate into mature B-cells. Upon stimulation with the appropriate cytokines Pax5−/− pro-B-cells can be differentiated into a broad spectrum of hematopoietic cell types in vitro and only restoration of Pax5 expression suppresses the multilineage potential of these cells (Fig. 3) (Nutt et al, 1999a). Conditional Pax5 inactivation in committed pro-B-cells reverts lineage commitment and results in retrodifferentiation of B-lymphocytes to an uncommitted progenitor stage (Fig. 3), and these cells can completely restore thymocyte development in vivo in Rag2-deficient mice (Mikkola et al, 2002). Upon conditional Pax5 deletion in mice mature B-cells from peripheral lymphoid organs are capable to dedifferentiate in vivo back to early uncommitted progenitors in the bone marrow, which rescues T lymphopoiesis in the thymus of T-cell-deficient mice (Cobaleda et al, 2007a).

Stages of B-cell development, the corresponding status of V(D)J recombination, and the approximate points at which B-cell lymphopoiesis is arrested in mice upon deletion of key regulatory factors are summarized in Figure 4.
Figure 3. B-cell lineage commitment by Pax5 (Cobaleda et al, 2007b). Uncommitted Pax5\(^{-/-}\) pro-B-cells are able to differentiate into several hematopoietic cell types. The cytokines required for \textit{in vitro} differentiation are indicated. Conditional Pax5 deletion (\(\Delta\) Pax5) results in retrodifferentiation of B-lymphocytes to an uncommitted progenitor cell stage. During terminal plasma cell differentiation Pax5 is physiologically downregulated. OPGL, osteoprotegerin ligand (also known as RANKL or TRANCE); ST2, stromal ST2 cells; TCR, T-cell receptor. Figure taken from Cobaleda et al, 2007.

Figure 4. B-cell developmental stages (Hagman & Lukin, 2005). Progressive stages of B-cell lymphopoiesis are shown. Cell designations are indicated below, the status of V(D)J recombination is shown within each cell type and characteristic cell surface markers are depicted above each cell. The approximate points at which B-cell lymphopoiesis is arrested in PU.1\(^{+/−}\), TCF3 (E2A\(^{+/−}\)), EBF1 (EBF\(^{+/−}\)) and Pax5\(^{-/-}\) mice are designated above the cells. HSC, hematopoietic stem cell; MLP, multi-lineage progenitor; CLP, common lymphoid progenitor, Pro-B, pro-B-cell; Pre-B, pre-B-cell. Figure taken from Hagman et al, 2005.
At the molecular level, Pax5 fulfills a dual role by activating B-cell-specific genes and simultaneously repressing lineage-inappropriate genes to initiate B-lineage differentiation (Nutt et al, 1999a). In this context, the Pax5 paired domain functions as bipartite DNA-binding region, which binds to the distinct half-site of the degenerate Pax5 recognition sequence (Czerny et al, 1993; Garvie et al, 2001). Transcriptional regulation of Pax5 target genes is determined by the interaction of distinct partner proteins with the central and C-terminal protein interaction motifs of Pax5 (Cobaleda et al, 2007b). The partial homeodomain associates with the TATA-binding protein of the basal transcription machinery, while the transactivation domain regulates gene transcription most likely by interacting with histone acetyltransferases such as the coactivator CREBBP (CBP) or SAGA complex. In contrast, corepressors of the Groucho protein family, which are part of a larger histone deacetylase complex, convert Pax5 from a transcriptional activator to a repressor by binding to the octapeptide motif (Cobaleda et al, 2007b).

Pax5 target gene activation plays an essential role in controlling signal transduction from the pre-BCR and BCR, which constitute important checkpoints in B-cell development (Schebesta et al, 2007). Pax5 promotes V_{H}-DJ_{H} recombination at the IgH locus (Fuxa et al, 2004), it activates expression of Cd79a (Igα) (Fitzsimmons et al, 1996), Cd19 (Kozmik et al, 1992; Nutt et al, 1998), Cr2 (Cd21) (Horcher et al, 2001), Cd72 (Ying et al, 1998; Horcher et al, 2001), and Blnk (SLP-65) (Schebesta et al, 2002) and, thus, facilitates pre-BCR signaling. Comprehensive gene expression analysis identified additional Pax5-activated genes, which are implicated in the control of signaling from the pre-BCR on the cell surface to transcription in the nucleus at multiple levels (Schebesta et al, 2007). Moreover, it was shown that Pax5 regulates genes involved in B-cell adhesion and migration, such as cell-surface receptors and intracellular signal transducers, which leads to a remodeling of the actin cytoskeleton. Pax5 also activates a number of transcription factors involved in B-cell differentiation, including lkzf3 (Aiolos), Spib, Irf4, Irf8, Lef1, and Ebf1 suggesting that Pax5 activity initiates a downstream transcriptional cascade that reinforces the B-cell program (Schebesta et al, 2007; Pridans et al, 2008).

Besides the activation of B-cell specific genes Pax5 concurrently represses lineage-inappropriate genes, which become reactivated in Pax5^{-/-} pro-B-cells (Nutt et al, 1999a; Delogu et al, 2006). Global transcriptional profiling of Pax5^{-/-} pro-B-cells identified >100 genes repressed by Pax5 including genes implicated in cell-cell communication, cell adhesion, migration, nuclear processes, and cellular metabolism at B-cell commitment (Delogu et al, 2006). Importantly, many genes repressed by Pax5 are normally expressed in non-B-cell lineages, which underlines the lineage promiscuity of Pax5^{-/-} pro-B-cells. As an example, Pax5 represses cell surface receptors Csf1r (M-CSFR) and Notch1, associated with macrophage and T-cell development, respectively, rendering committed B-lymphocytes
unresponsive to lineage-inappropriate signals (Nutt et al, 1999a; Souabni et al, 2002; Nutt & Kee, 2007). Furthermore, Pax5 represses genes that are required to maintain stem cell or multipotent progenitor fate such as Atxn1 (Sca1) and Flt3 (Delogu et al, 2006).

Terminal differentiation of mature B-cells into antibody-secreting cells is antigen-driven and represents a crucial component of the immune response. The physiological downregulation of Pax5 upon antigen stimulation followed by the reactivation of Pax5-repressed plasma cell-specific genes including Xbp1, Igj, Cd28, Ccr2, and Prdm1 (Blimp1) seems to be the initial event in plasma cell differentiation (Delogu et al, 2006; Kallies et al, 2007). Final plasma cell differentiation requires the expression of functional Prdm1 (Blimp1), which represses Pax5 and, hence, the B-cell program by a feedback mechanism (Kallies et al, 2007). Thus, comprehensive analysis of just a small number of key transcription factors involved in B-cell differentiation has revealed that the transcriptional network controlling B-cell specification and commitment is not a simple linear cascade but involves multiple combinatorial inputs and feedback loops (Nutt & Kee, 2007).

1.3. Acute lymphoblastic leukemia

Acute lymphoblastic leukemia (ALL) is the most common leukemia in children accounting for approximately 80% of pediatric cases (Martinez-Climent, 1997). In ALL, the B- or T-cell lineage can be affected, and in children the most common immunophenotype is B-cell precursor (BCP) ALL, followed by T-cell, and mature B-cell ALL. The group of BCP-ALL has a good prognosis with an event-free survival (EFS) of approximately 80% in children between 1-18 years, whereas in infants (children <1 year of age) the outcome is worse with an EFS of 28-45% (Armstrong & Look, 2005; Pieters et al, 2007). About 15% of childhood ALL cases are diagnosed with T-ALL and this group of patients was historically linked to a poor prognosis (Aifantis et al, 2008). However, owing to the application of intensive chemotherapy regimens, nowadays also in T-ALL cure rates of approximately 75% are achieved. Nevertheless, in about 25% of the patients treatment failure occurs and the outcome of these patients remains dismal (Goldberg et al, 2003; Einsiedel et al, 2005).

Recurrent genetic abnormalities are a hallmark of acute leukemia and provide insights into the molecular mechanisms of leukemogenesis (Armstrong & Look, 2005). The most frequent targets of genetic alterations involved in hematological disorders are genes controlling transcription and tyrosine kinases (Mitelman et al, 2004). The large variety of genetic alterations includes point mutations and deletions, but the main genetic characteristics of acute leukemia are translocations and numerical chromosome imbalances resulting in hyper- or hypodiploidy (Fig. 5). These chromosomally defined subtypes also show distinctive patterns of global gene expression in microarray analysis (Greaves & Wiemels, 2003).
Chromosome translocations either result in inappropriate expression of an oncogene by juxtaposition of the entire coding sequence under constitutive activated regulatory elements of a partner gene or more commonly in leukemia in the formation of a chimeric fusion gene with novel properties. An increasing number of promiscuous genes (e.g. MLL, ETV6 or NUP98) that recombine with numerous different partner genes have been identified and, thus, the number of fusion genes exceeds the number of affected genes. So far, in ALL 1139 balanced aberrations have been described, 155 of which are recurrent resulting in 82 distinct gene fusions (Mitelman et al, 2007).

In childhood BCP-ALL the most common genetic rearrangement is the t(12;21)(p13;q22), which fuses ETV6 to RUNX1 and is present in about 25% of BCP-ALL cases (Fig. 5) (Armstrong & Look, 2005). Other commonly found chromosomal aberrations in childhood ALL are the t(1;19)(q23;p13)/E2A-PBX1 (TCF3-PBX1), the t(9;22)(q34;q11)/BCR-ABL1, and hyperdiploidy (presence of >46 chromosomes), which is often associated with a FLT3 mutation (Fig. 5) (Armstrong & Look, 2005). Rearrangements of the MLL gene occur in up to 80% of infant ALL and are associated with a pro-B ALL phenotype (Attarbaschi et al, 2006; O'Neil & Look, 2007; Pieters et al, 2007).

Recently, pediatric ALL was further characterized by genome-wide analyses using high-resolution SNP arrays and DNA sequencing, which uncovered that in about 40% of BCP-ALL genes implicated in B-cell development and differentiation are targets of mutations, deletions or structural rearrangements (Mullighan et al, 2007a). These genes comprised IKZF1 (Ikaros), IKZF3 (Aiolos), LEF1, EBF1, TCF3 (E2A), and PAX5 (Kuiper et al, 2007; Mullighan et al, 2007a). In addition to microdeletions in transcription factors involved in B-lineage
development, recurrent deletion of *BTG1* a negative effector of B-cell proliferation was observed. Moreover, other genes frequently affected by copy number losses were those controlling G1/S cell cycle progression (e.g. *CDKN2A*, *CDKN2B*, and *RB1*), and such deletions were detected in 54% of BCP-ALL and 86% of T-ALL, respectively (Kuiper et al, 2007).

Many pediatric leukemias originate *in utero*, which was demonstrated by sequence analysis of the unique genomic breakpoints of chromosomal rearrangements of concordant leukemia in monozygotic twins, and from screening of Guthrie cards (Fasching et al, 2000; Panzer-Grumayer et al, 2002; Greaves & Wiemels, 2003). In contrast to BCP-ALL, analysis of neonatal blood spots for leukemia specific rearrangements showed that most T-ALL cases are more likely initiated postnatally (Fischer et al, 2007). Interestingly, the *ETV6-RUNX1* and the *RUNX1-RUNX1T1* (*AML1-ETO*) fusion genes could be detected 100 times more often in blood samples from healthy newborns as the risk of the corresponding leukemia (Mori et al, 2002).

In transgenic mouse models it was shown that particular fusion genes can initiate, but seldom complete leukemogenesis and, thus, require cooperating mutations similar to the Knudson two-step model for non-inheritable pediatric solid tumors (Knudson, 1992; Bernardin et al, 2002; Greaves & Wiemels, 2003; Tsuzuki et al, 2004). For instance, *ETV6-RUNX1* positive leukemia is often associated with additional genetic changes, such as deletions of the second *ETV6* allele, *PAX5*, *CDKN2A*, or *BTG1*, which may represent one of the secondary rate-limiting hits in leukemogenesis (Greaves & Wiemels, 2003; Mullighan et al, 2007a; Tsuzuki et al, 2007). In a recent study of monochorionic twins – one twin had BCP-ALL while the other one was healthy – a tumor-propagating *ETV6-RUNX1* positive cell population could be isolated from both twins. The *ETV6-RUNX1* positive leukemic blasts of the twin diagnosed with full-blown BCP-ALL showed a deletion of the second *ETV6* allele, whereas the *ETV6-RUNX1* positive cells of the healthy twin harbored one intact copy of *ETV6* (Hong et al, 2008). These data strongly support the notion that inactivation of the second unrearranged *ETV6* allele indeed represents a crucial cooperating mutation (Hong et al, 2008).

### 1.4. The PAX gene family in oncogenesis

The mammalian paired box or PAX transcription factor family comprises nine members and is characterized by a highly conserved paired box DNA binding domain. Mouse and human *PAX* genes are classified into four paralogous groups according to the presence of two additional motifs, namely a conserved octapeptide and a complete or truncated homeodomain (Fig. 6) (Bouchard et al, 2003). Additionally, all members of the *PAX* gene family contain regions rich in proline, serine, and threonine residues (PST-rich) at the...
C-terminal end (Fig. 6), which comprises a transactivation and an inhibitory domain, a characteristic feature of inducible transcription factors.

![Diagram](image)

**Figure 6. Structure and classification of mammalian Pax proteins (Bouchard et al., 2003).** Pax proteins are classified according to their protein domains. PD, paired domain; HD, homeodomain; OP, conserved octapeptide, TAD, transactivation domain; PST, proline-serine-threonine. Figure taken from Bouchard et al., 2003.

The PAX family has been conserved throughout metazoan evolution and controls tissue development and differentiation processes during embryonic development including proliferation, stem-cell self-renewal, apoptosis, cell migration and invasion within a variety of cell lineages (Barr, 1997; Robson et al., 2006). The important developmental role of the PAX genes is further emphasized by the association of mutations with heritable murine and human developmental defects (Barr, 1997). Moreover, the cellular activities controlled by the PAX genes are also fundamental targets for the development of neoplasia (Barr, 1997). Therefore, the PAX genes are intriguing candidates to contribute to tumorigenesis in specific cell lineages, and indeed, some members of the PAX gene family have been shown to be involved in tumor development (Robson et al., 2006).

**PAX3** and **PAX7** are expressed during early neural and myogenic development. In alveolar rhabdomyosarcoma (ARMS), a pediatric soft tissue tumor related to the striated muscle lineage, the PAX3 and PAX7 genes are fused to the FOXO1 (FKHR) gene through the translocations t(2;13)(q35;q14) and t(1;13)(p36;q14), respectively (Galili et al., 1993; Davis et al., 1994). The t(2;13)(q35;q14)/PAX3-FKHR (PAX3-FOXO1) rearrangement is the most prevalent finding in ARMS detected in about 70% of cases, whereas the t(1;13)(p36;q14) is found in a smaller subset of cases. The chimeric fusion proteins contain the PAX3/PAX7 paired domain, the octapeptide and the homeodomain fused to the FOXO1 (FKHR) transcriptional activation domain. The PAX3-FOXO1 and PAX7-FOXO1 proteins function as transcription factors that activate genes containing a PAX3/PAX7 DNA-binding site in a more potent manner than the corresponding wild-type proteins (Barr, 2001). Moreover, both fusions are consistently overexpressed relative to the respective wild-type PAX transcripts.
Overexpression of the PAX7-FOXO1 fusion results mainly from \textit{in vivo} amplification of the fusion gene, whereas in case of PAX3-FOXO1 the increase in transcriptional rate is copy number independent (Barr, 2001).

Furthermore, alternative tumor-specific PAX3 and PAX7 isoforms are predominantly expressed in Ewing's sarcoma and in embryonal rhabdomyosarcoma, and melanoma cell lines, respectively (Barr et al, 1999). In squamos cell lung carcinoma PAX7 is frequently amplified (Racz et al, 2000) and in N-type neuroblastoma cell lines (with high \textit{N-\textit{MYC}} expression and/or amplification) two isoforms of PAX3 are expressed at abnormally high levels (Harris et al, 2002; Wang et al, 2008).

PAX2 is required for kidney, eye and ear development and in the mammary glands, whereas PAX8 controls thyroid development, but is also expressed during kidney organogenesis (Bouchard et al, 2003; Robson et al, 2006). In thyroid follicular carcinomas PAX8 is fused to PPARG (PPAR\textgamma1) as a result of the translocation t(2;3)(q13;p25) generating a chimeric fusion gene that contains the paired box, octapeptide, and partial homeodomain of PAX8 fused to the entire PPARG (PPAR\textgamma1) protein including all nuclear receptor domains (Kroll et al, 2000). The PAX8-PPARG (PAX8-PPAR\textgamma1) fusion is expressed at higher levels as endogenous PPARG (PPAR\textgamma1) and it might act in a dominant negative manner over wild-type PPARG (PPAR\textgamma1), which is implicated in growth inhibition and promotes differentiation of cancer cell lines (Kroll et al, 2000). Moreover, in carcinomas of the kidney, prostate, breast, ovary, and in blastemal tissues in Wilms' tumor and Kaposi sarcoma unattenuated tumor-associated expression of PAX2 and/or PAX8 was observed (Robson et al, 2006).

PAX5 expression is critically required for very early brain development but transcriptionally downregulated before birth. In astrocytoma, glioblastoma, medullablastoma, small cell lung cancer (neural-crest derived tumor) and N-type neuroblastoma (a malignant subset), but not in S-type cells (a benign subset), PAX5 is ectopically expressed during tumor development (Stuart et al, 1995b; Baumann Kubetzko et al, 2004; Robson et al, 2006). Moreover, it was shown that PAX5 expression levels correlate with or promote neoplastic tumor growth in astrocytoma and neuroblastoma, respectively (Stuart et al, 1995b; Baumann Kubetzko et al, 2004). Furthermore, in astrocytoma the expression of PAX5 is inversely proportional to the expression of \textit{TP53} (p53), whose transcriptional activity can be repressed by binding of PAX5 to the \textit{TP53} (p53) promoter (Stuart et al, 1995a).

The involvement of PAX5 in B-cell malignancies is described in section 1.5.

PAX6, a master regulator of eye development, is frequently expressed in brain, breast and other cancer cell lines (Muratovska et al, 2003). However, in glioblastoma a tumor-suppressor function for PAX6 was suggested and also in malignant astrocytic gliomas high levels of PAX6 expression correlate with improved prognosis (Robson et al, 2006).
**PAX9** expression was also widely detected in cancer cell lines (Muratovska *et al.*, 2003). On the other hand decreased **PAX9** expression correlates with increasing malignancy of oesophageal carcinomas and epithelial dysplasia (Robson *et al.*, 2006).

In summary, **PAX** genes seem to exhibit a pivotal role in the oncogenesis of several tumors arising from those tissues, in which they exert a developmental stage-dependent function during embryogenesis and differentiation (Lang *et al.*, 2007; Wang *et al.*, 2008). Because of their normal function in development, it is assumed that re-expression of **PAX** genes in malignant neoplasms promotes tumor development and progression by increasing proliferation and motility, while inhibiting apoptosis (Baumann Kubetzko *et al.*, 2004).

1.5. **PAX5** aberrations in B-cell malignancies

1.5.1. **PAX5** aberrations in lymphoma

In non-Hodgkin lymphoma **PAX5** is involved in the rare translocation t(9;14)(p13;q32) resulting in a juxtaposition of the intact **PAX5** coding sequence to regulatory elements of the **IGH@** locus, which leads to inappropriate **PAX5** expression (Busslinger *et al.*, 1996; Iida *et al.*, 1996). The breakpoints within the **PAX5** and **IGH@** loci are variable and, thus, the deregulation of **PAX5** occurs in two different ways: (1) In the KIS-1 cell line, which was established from a patient with diffuse large-cell lymphoma, the breakpoint at 9p13 occurs 1.8kb upstream of exon 1A of **PAX5** and juxtaposes **PAX5** in a head-to-head orientation in close proximity to the potent **IGH@** gene Eµ enhancer (Busslinger *et al.*, 1996). Thus, deregulation of **PAX5** transcription is caused by enhancer insertion. (2) In variant translocations, which were cloned from a patient with lymphoplasmacytoid lymphoma (Iida *et al.*, 1996) and a splenic marginal zone lymphoma (Morrison *et al.*, 1998), the breakpoints arise in the non-coding sequence and the 3' region of exon 1B of **PAX5**, respectively. In both cases the rearrangement translocates the **PAX5** gene into the Sµ region of the **IGH@** gene in a head-to-head position, which leads to replacement of the **PAX5** promoters by an antisense promoter of the Sµ region.

In a transgenic mouse model a Pax5 minigene was inserted into the Igh locus to mimic the t(9;14)(p13;q32)/**PAX5-IGH@** rearrangement (Souabni *et al.*, 2007). This knock-in mouse corresponds to a germline rather than a somatic mutation and therefore, curiously, the mice developed T-cell lymphomas. Nevertheless, this data identified Pax5 as a potent oncogene, in that ectopic Pax5 expression interferes with normal T-cell development and deregulated the T-cell transcription program (Souabni *et al.*, 2007). Conversely, in mice experiments biallelic Pax5 deletion in mature B-cells resulted in the development of aggressive lymphoma, which were by gene expression analysis characterized as progenitor cell tumors.
Moreover, in respect to their expression profile of Pax5 target genes these cells were indistinguishable from Pax5<sup>-/-</sup> pro-B-cells (Cobaleda <i>et al</i>, 2007a).

In follicular lymphoma, PAX5 and MYC are the only transcription factors consistently overexpressed as compared to their putative normal counterparts, germinal center B-cells (Husson <i>et al</i>, 2002). Furthermore, in diffuse large B-cell lymphoma (DLBCL) and to a lesser extent in Burkitt and follicular lymphomas, PAX5 showed a high frequency of somatic hypermutations, which cluster downstream of both transcription initiation sites, predominantly around exon 1B (Pasqualucci <i>et al</i>, 2001). Such hypermutations of PAX5 have not been detected in normal germinal-center B-cells, naïve B-cells and control fibroblasts and, thus, may cause PAX5 malfunction in these diseases (Pasqualucci <i>et al</i>, 2001). However, as the alternatively transcribed exon 1A and the second PAX5 allele mainly escape somatic hypermutations, the role of these hypermutations for lymphoma formation is doubtful (Cobaleda <i>et al</i>, 2007b).

Recent work also addressed the role of PAX5 during lymphomagenesis (Cozma <i>et al</i>, 2007). It was shown that activation of Pax5 significantly upregulated components of the BCR signaling, such as Cd79a, Cd19, Blnk, PKCβ, and PLCγ2, and that this activation stimulated tumor growth. Moreover, knock down of Pax5 expression in DLBCL cell lines by sh-RNA decreased the growth rate of these cell lines. The contribution of Pax5 to neoplastic growth appears to correlate with its ability to maintain expression of BCR components. Thus, interference with BCR signaling downstream of Cd79a either by overexpression of Cd22 or by pharmacological inhibition may represent a therapeutic option (Cozma <i>et al</i>, 2007).

### 1.5.2. PAX5 aberrations in acute lymphoblastic leukemia

#### 1.5.2.1. Deletions and Mutations

By genome-wide SNP array analysis of childhood ALL frequent deletions of the PAX5 gene were detected (Kawamata <i>et al</i>, 2007; Kuiper <i>et al</i>, 2007; Mullighan <i>et al</i>, 2007a). Deletions were found in about 30% of all cases and comprised focal intragenic PAX5 deletions (13%), broader deletions involving PAX5 and a variable number of flanking genes (3.6%), large 9p deletions including the 3' portion of PAX5 (2.6%), and deletions of the whole 9p arm or complete loss of one chromosome 9 (9.9%) (Fig. 7) (Mullighan <i>et al</i>, 2007a). The focal PAX5 deletions affect only a subset of PAX5 exons (Fig. 7) resulting in expression of internally deleted transcripts. These transcripts, the so-called hypermorphic alleles, encode proteins lacking either the PAX5 paired domain and/or transcriptional regulatory domains or lead to truncated mutants.
In several cases also PAX5 point mutations, which mainly comprise frameshift, splice site or missense mutations and clustered in exons encoding the paired domain or the transcriptional regulatory domains were found (Fig. 8). Mutations affecting the paired domain are suggested to impair the DNA-binding function of PAX5 and mutations in the transactivation domain may alter transcriptional regulatory functions. Indeed, an impaired function of PAX5 mutants was shown in luciferase-based reporter assays and a reduced DNA-binding activity for PAX5 variants with mutations or deletions of the paired-domain was demonstrated. However, in leukemic blasts no correlation between PAX5 mutation status and CD19 and CD79A expression was observed (Mullighan et al, 2007a).

Intriguingly, the type and frequency of PAX5 aberrations varied among the genetic subtypes of ALL: all hypodiploid cases showed loss of one PAX5 allele and about 50% of the cases harbored concomitant mutations in the second PAX5 allele, whereas 28% of ETV6-RUNX1 positive cases displayed focal mono-allelic deletions but lacked PAX5 mutations (Mullighan et al, 2007a).
The high frequency of PAX5 deletions in ETV6-RUNX1 BCP-ALL suggests that they may represent 'second hit' mutations cooperating in the pathogenesis of this leukemia subtype.

In addition, recent analysis of BCR-ABL1 positive pediatric and adult ALL by SNP microarrays identified IKZF1 (Ikaros) deletions in >80% of cases and in about 50% deletions of PAX5 and CDKN2A, which mostly coincided with IKZF1 deletions (Mullighan et al, 2008). Strikingly, these deletions were found in BCR-ABL1 positive ALL and CML blast crisis, but were not detected in chronic-phase CML (Mullighan et al, 2008). Together these data indicate that alterations in Ikaros possibly in conjunction with haploinsufficiency of PAX5 may contribute to the arrested B-lymphoid maturation in BCR-ABL1 positive ALL.

1.5.2.2. PAX5 fusions in ALL

In 2001 the first chimeric PAX5 fusion gene in a case of ALL with a t(9;12)(q11;p13) resulting in a PAX5-ETV6 fusion was identified (Cazzaniga et al, 2001). Subsequently, it was shown that the PAX5-ETV6 fusion defines the cytogenetic entity dic(9;12)(p13;p13), which occurs in about 1% of childhood ALL (Strehl et al, 2003). The PAX5-ETV6 rearrangement fuses exon 4 of PAX5 to exon 3 of ETV6, thus, the fusion protein contains the PAX5 DNA-binding paired domain and the HLH and ETS-binding domains of ETV6. The PAX5-ETV6 chimeric protein most likely acts as an aberrant transcription factor, probably as transcriptional repressor, which recruits transcriptional cofactors through the ETV6 regulatory elements (Fazio et al, 2008). In addition, PAX5-ETV6 functions as strong competitive inhibitor of wild-type PAX5 in co-transfection and co-transduction experiments (Mullighan et al, 2007a).

Six years after the first description of this PAX5 rearrangement in ALL, several other PAX5 fusion partners were identified, namely AUTS2, C20orf112 (Kawamata et al, 2008), ELN (Bousquet et al, 2007), FOXP1, and ZNF521 (Mullighan et al, 2007a). All breakpoints described in these novel PAX5 fusions occurred within PAX5 intron 5 or downstream of it fusing at least the PAX5 DNA-binding paired domain to the C-terminal region or nearly the entire protein of the fusion partner. It is assumed, that PAX5 fusion proteins act as transcriptional repressors, which antagonize the PAX5 activity provided by the normal wild-type allele (Cobaleda et al, 2007b).
CHAPTER 2

IDENTIFICATION OF PML AS NOVEL PAX5 FUSION PARTNER IN CHILDHOOD ACUTE LYMPHOBLASTIC LEUKEMIA

Karin Nebral¹, Margit König¹, Lana Harder², Reiner Siebert², Oskar Haas³, Sabine Strehl¹

British Journal of Haematology. 2007;139:269-274.

¹CCRI, Children’s Cancer Research Institute, St. Anna Kinderkrebsforschung, Kinderspitalgasse 6, A-1090 Vienna, Austria
²Institute of Human Genetics, University Hospital Schleswig-Holstein, Campus Kiel, Schwanenweg 24, D-24105 Kiel, Germany
³St. Anna Children's Hospital, Kinderspitalgasse 6, A-1090 Vienna, Austria
Summary

*PAX5* encodes the B-cell lineage specific activator protein (BSAP) and is required for B-cell development and maintenance. In B-cell precursor acute lymphoblastic leukemia (ALL), *PAX5* is involved in several chromosome translocations that fuse the N-terminal paired DNA-binding domain of *PAX5* with the C-terminal regulatory sequences of ETV6, FOXP1, ZNF521 or ELN. Herein, we describe the identification of a novel recurrent t(9;15)(p13;q24) in two cases of childhood ALL, which results in an in-frame fusion of *PAX5* to the promyelocytic leukemia (*PML*) gene. The putative *PAX5-PML* fusion gene encodes a chimeric protein that retains the paired domain, the octapeptid and the partial homeodomain of *PAX5*, and virtually the whole PML protein. The steadily increasing number of *PAX5* rearrangements suggests that *PAX5* is not only crucial for B-cell lymphopoiesis but also for the development of B-cell malignancies.

*Key Words*: *PAX5*, *PML*, fusion transcript, childhood acute lymphoblastic leukemia, t(9;15)(p13;q24).
Introduction

PAX5 is a member of the paired box (PAX) family, a group of nine highly conserved transcription factors that are implicated in brain development and organogenesis (Bouchard et al, 2003). PAX5 encodes the only PAX protein expressed within the hematopoietic system, the B-cell lineage specific activator protein (BSAP) that is required for B-cell commitment and maintenance. At the molecular level, Pax5 fulfills a dual role by activating B-cell specific genes and simultaneously repressing lineage-inappropriate genes (Nutt et al, 1999).

In the bone marrow, Pax5 is exclusively expressed from the pro-B to the mature B-cell stage and is down regulated during terminal differentiation into plasma cells (Nutt et al, 1998). In the absence of Pax5 in homozygous mutant mice, B-cell development is arrested at an early pro-B (pre-BI) cell stage (Urbanek et al, 1994; Nutt et al, 1999). Expression of Pax5 is also essential for maintaining B-cell identity as upon conditional inactivation using a CD19-driven Cre-loxP system that allows for tissue-specific deletion of Pax5 (Horcher et al, 2001), committed pro-B cells with a restricted B-lymphoid fate convert into progenitors with multilineage potential (Mikkola et al, 2002). Restoration of Pax5 expression suppresses the hematopoietic pluripotency of Pax5−/− pro-B cells while simultaneously promoting their development to mature B-cells. (Nutt et al, 1999) Together, these data render Pax5 as the critical B-lineage commitment factor (Cobaleda et al, 2007).

Chromosomal translocations affecting PAX5 have been described in different types of B-cell malignancies. The t(9;14)(p13;q32) translocation, which is mainly associated with B-cell non-Hodgkin-Lymphoma (B-NHL) results in the juxtaposition of PAX5 to the immunoglobulin heavy-chain (IGH@) locus, and thus, brings PAX5 under the control of potent enhancers or promoters from the IGH@ locus leading to elevated PAX5 expression (Busslinger et al, 1996; Morrison et al, 1998). In B-cell precursor acute lymphoblastic leukemia (ALL) PAX5 rearrangements involve several different partner genes including ETV6 (12p13), FOXP1 (3p14), ZNF521 (18q11), and ELN (7q11) (Cazzaniga et al, 2001; Strehl et al, 2003; Bousquet et al, 2007; Mullighan et al, 2007).

PAX5 consists of a N-terminal paired domain, which is a bipartite DNA-binding region and a C-terminal proline-serine-threonine-rich region that harbors a transactivation domain. The central region contains an octapetide capable of recruiting members of the Groucho proteins, a family of transcriptional corepressors that are required for many developmental processes, including lateral inhibition, segmentation, sex determination, dorsal/ventral pattern formation, terminal pattern formation and eye development (Chen & Courey, 2000; Eberhard et al, 2000), and a partial homeodomain functioning as a protein-protein interaction motif (Bouchard et al, 2003). The evolutionary highly conserved paired box domain that is shared by all PAX genes is retained in every fusion protein involving other members of the PAX family, namely PAX3, PAX7, and PAX8, which are affected by tumor-specific translocations.
in alveolar rhabdomyosarcoma and thyroid follicular carcinoma (Kroll et al, 2000; Barr, 2001). Each of the PAX5 chimeric genes identified so far also encodes a fusion protein that maintains the paired-box DNA-binding domain that is fused to the DNA-binding and transcriptional regulatory domains of the partner protein (Cazzaniga et al, 2001; Strehl et al, 2003; Bousquet et al, 2007; Mullighan et al, 2007). PAX5 fusion proteins may contribute to leukemogenesis by acting as constitutive repressor, and thus interfering with normal PAX5 function (Cobaleda et al, 2007).

This study identified a novel recurrent t(9;15)(p13;q24) in two cases of childhood B-ALL, which results in the fusion of the 5’ region of PAX5 to almost the entire promyelocytic leukemia (PML) gene.
Material and Methods

Case History

Patient 1, a 9-months-old infant suffering from continuous fever, otorrhoe and hepatomegaly was diagnosed with ALL. The bone marrow (BM) showed 99% blast cell infiltration and 63% in the peripheral blood (PB). Immunophenotyping was performed on BM cells by means of flow cytometry with a panel of monoclonal antibodies. The blast cells were positive for CD19, CD79a, CD10, CD22, CD34, TdT, and HLA-DR typical for B-II-ALL (common-ALL). Cytogenetic analysis revealed a 46,XX,add(9)(p13)[8]/46,XX[14] karyotype. The patient was treated according to the ALL - Berlin-Frankfurt-Münster (BFM) 2000 Interfant protocol and is in complete remission more than 4 years from diagnosis.

Patient 2 was diagnosed with ALL at 19.5 months of age, following a 1-month period of pneumonic complaints. Immunophenotyping revealed a B-II-ALL positive for CD19, CD10, CD22, TdT, and HLA-DR. Treatment was performed according to the ALL-BFM 86 protocol (Reiter et al, 1994) and the patient achieved remission after 40 d, but relapsed 2 years and 3 months after diagnosis. At relapse biopsy of the testis showed infiltration with ALL blast cells and in the BM 45% of lymphoblasts were detected. Cytogenetic analysis showed a 46,XY,t(9;15)(p21;q25)[12] karyotype in the testis and a 46,XY,t(9;15)(p21;q25)[3]/46,XY,add(1)(p?)[3] in the BM. The patient was treated with high-dose methotrexate, but died of progressive disease.

Conventional and molecular cytogenetics

Cytogenetic analysis was performed using standard methods and karyotypes were described according to the International System for Human Cytogenetic Nomenclature (Shaffer & Tommerup, 2005). PAX5 rearrangements were detected using exon-specific cosmids cos-hPAX5-1 (exons 2-5) and cos-hPAX5-3 (exons 9-10) (Busslinger et al, 1996). In addition, the LSI PML-RARA dual-color, dual-fusion-translocation probe (Vysis, Downers Grove, IL, USA) and the BAC clone RP11-2M12 (The Sanger Institute, Cambridge, United Kingdom) encompassing the whole PML gene were used. To ensure analysis of abnormal metaphases, whole chromosome painting (WCP) probes were combined with gene-specific probes. The 24-color-fluorescence in situ hybridization (FISH) analysis was performed with the Spectra Vysion probe (Vysis, Downers Grove, IL, USA). Probes were differentially labeled by nick translation either with digoxigenin-11-dUTP or biotin-16-dUTP (Roche Diagnostics, Vienna, Austria). Slides for FISH were prepared from the methanol/acetic acid-fixed cell suspension used for cytogenetic analysis and FISH was performed as previously described (Konig et al, 2002). Samples were evaluated using an Axioplan fluorescence microscope (Zeiss, Vienna, Austria) equipped with the appropriate filter sets. Images were
taken with a CCD camera (Photometrix, Tucson, AZ) using the IPLabs software (Vysis, Downers Grove, IL, USA).

**Reverse-Transcription-PCR analysis**

Total RNA was isolated from cryopreserved mononuclear cells (MNCs) of the BM obtained from patient 1 at diagnosis, and from methanol/acetic acid-fixed cell suspension of BM cells obtained from patient 2 at relapse as previously described (Strehl et al, 2001). RNA extraction was performed using the peqGOLD Total RNA kit (peqLab, Biotechnologie GmbH, Erlangen, Germany) according to the manufacturer’s recommendations. RNA was reverse transcribed with 200 Units Moloney-murine leukaemia virus (M-MLV) reverse transcriptase (Invitrogen, Lofer, Austria) and 100 pmol random hexamers (GE Healthcare, Vienna, Austria) at 42°C for 1 h. Reverse-transcription polymerase chain reaction (RT-PCR) were performed using Hot Start Taq polymerase (Qiagen, Vienna, Austria) according to the manufacturer’s instructions, with annealing at 61-63°C for 30 s, elongation at 72°C for 30-45 s for 30-40 cycles. Primer sequences are listed in Table 1.

**Results**

**Conventional and molecular cytogenetics**

Cytogenetic analysis of patient 1 showed a 46,XX,add(9)(p13). To determine the chromosomal origin of the extra material on 9p, 24-color FISH was performed. This analysis revealed a t(9;15) (Fig 1B, inset), and subsequent FISH using the PAX5 exon-specific cosmids detected a separation of the probes (Fig 1A) suggesting involvement of PAX5 in the t(9;15).

To narrow down the precise breakpoint at 15q, FISH-based chromosome walking with various locus-specific BAC clones located at 15q24-25 was performed (data not shown). Hybridization of the PML-specific BAC RP11-2M12 in combination with a whole chromosome painting probe specific for chromosome 9 to ensure hybridization of aberrant metaphases resulted in a split signal of PML (Fig 1B) providing compelling evidence that PAX5 was fused to PML. Thus, the karyotype was refined as t(9;15)(p13;q24).

In the second patient (patient 2), cytogenetic analysis revealed a similar aberration, namely a t(9;15)(p21;q25). Also in this case, FISH analysis using the PML-RARA and the PML-specific BAC probes showed disruption of PML. Subsequent hybridization with the PAX5 exon-specific cosmids showed a deletion of the PAX5 3’-end (data not shown) indicating that this patient also displayed a PAX5-PML fusion, associated with a deletion of the PAX5 3’-end.
RT-PCR analysis
Fusion gene-specific RT-PCR experiments using primers located in PAX5 exon 5 and PML exon 2 detected chimeric PAX5-PML transcripts in both patients (Fig 2A, left), but not in normal peripheral blood used as a negative control. Sequence analyses identified exactly the same in-frame fusion between PAX5 exon 6 and PML exon 2 in both patients (Fig 2B and C). The PML protein consists of a RING domain followed by additional zinc fingers (B-boxes) and an α-helical coiled-coil motif (collectively referred to as RBCC domain) (Bernardi & Pandolfi, 2003). Thus, the putative PAX5-PML chimeric fusion protein consists of the paired domain, the octapeptide and the partial homeodomain of PAX5, and almost the entire PML protein lacking only the 5' proline-rich region (Fig 2D). Amplification of the reciprocal PML-PAX5 fusion transcript using different primer combinations failed, which indicate that the PAX5-PML fusion is responsible for leukemogenesis. These data are in concordance with the deletion of the PAX5 3'-end detected in patient 2 by FISH analysis. According to the FISH data, the second alleles of both genes involved in the translocation were retained and expression of normal PAX5 and PML transcripts was verified by RT-PCR (Fig 1A, right; analysis was performed only for patient 1, because there was a lack of material for patient 2).

Discussion
In this study, we report the identification of a novel recurrent t(9;15)(p13;q24), which results in a fusion of the B-cell specific transcription factor PAX5 and the PML gene. The putative PAX5-PML fusion protein fuses the paired domain, the octapeptide, and the homeodomain of PAX5 to almost the entire PML protein. Lack of reciprocal transcripts strongly suggest that PAX5-PML and not PML-PAX5 is responsible for leukemogenesis.

The second known translocation involving PML, the t(15;17)(q22;q21), fuses PML to the retinoic acid receptor alpha (RARA), and is the genetic hallmark of acute promyelocytic leukemia (APL). The PML-RARA protein functions as an aberrant retinoid receptor with altered DNA-binding properties as compared to wild-type RARA and acts as a constitutive transcriptional repressor of RARA target genes (Lo-Coco & Ammatuna, 2006). As approximately 20-30% of APL lack expression of the reciprocal RARA-PML, PML-RARA must be the chimeric protein that is critical for the development of APL (Melnick & Licht, 1999; Lo-Coco & Ammatuna, 2006).

PML is detected in the nucleus in multiprotein complexes termed PML nuclear bodies (NBs), which are implicated in the regulation of transcription, apoptosis, DNA repair, control of genomic stability, tumor suppression, cellular senescence, and anti-viral response (Zhong et al, 2000; Dellaire & Bazett-Jones, 2004). In this respect, a specific feature of PML-RARA positive cells is the delocalization of PML from the PML-NBs to a microspeckled nuclear pattern and relocalization of PML to the NBs upon ATRA treatment (Melnick & Licht, 1999)
In the PAX5-PML fusion protein almost the entire PML is retained, and thus the chimeric fusion protein might heterodimerize with normal PML resulting in impaired PML function. In this regard, the PML-NBs play a key role in the regulation and functional activation of a number of proapoptotic/tumor suppressive transcription factors (Bernardi & Pandolfi, 2003). Thus, one might speculate that impairment of PML by the PAX5-PML fusion may exert a survival advantage by interfering with cellular apoptotic programs. To date it remains elusive whether PAX5-PML has per se transforming potential or additional mutations are required for the development of overt leukemia. Thus, impairment of PML, which is known to be involved in genome stability (Bernardi & Pandolfi, 2003) may facilitate the accumulation of additional mutations.

On the other hand, PAX5, as a master regulator of B-cell development, is indispensable for B-lineage commitment and continuous expression is required to maintain B-cell fate (Nutt et al, 2001; Mikkola et al, 2002; Busslinger 2004). A common feature of most PAX5 chimeric proteins described to date, is the fusion of the paired box DNA-binding domain of PAX5 with C-terminal regulatory sequences of a second transcription factor implicated in B-cell development or hematopoietic malignancy. Thus, the fusion proteins are predicted to retain the ability to bind to PAX5 transcriptional targets without providing normal transcriptional regulatory functions (Cobaleda et al, 2007; Mullighan et al, 2007). Indeed, in transient transfection assays PAX5-ETV6 and PAX5-FOXP1 competitively inhibit the transcriptional activation of wild-type PAX5 (Mullighan et al, 2007). However, no transcriptional regulatory function for the PAX5 partner ELN, which encodes an extracellular matrix protein that is the main component of elastic fibers, has been demonstrated. Nevertheless, the PAX5-ELN fusion protein also acts in a dominant-negative manner over PAX5 in in vitro CD19 reporter gene assays (Bousquet et al, 2007). Thus, it is highly likely that the PAX5-PML chimera also operates as an aberrant transcription factor exerting a repressor activity antagonizing normal PAX5 function.

Yet, analyses of the effects of PAX5-ELN on PAX5 endogenous targets resulted in conflicting data, as in DG75 (Burkitt lymphoma) transfected cells, the expression of the PAX5 target genes BLNK, CD79A and LEF1 was downregulated, whereas CD19 and BLK remained unaffected. In contrast, PAX5-ELN leukemic pre-B-cells showed exactly the opposite expression pattern. These data suggest that, in case of PAX5 fusions, regulation of PAX5 target gene transcription may be cellular context-dependent (Bousquet et al, 2007).

Although PAX5 fusions seem to account for just about 2% of childhood ALL (Mullighan et al, 2007; unpublished observation) and the functional consequences of all PAX5 chimaeric proteins need to be elucidated in more detail, the increasing number of PAX5-involving rearrangements renders this critical B-cell-specific transcription factor not only crucial for
normal B-cell lymphopoiesis but it may also be considered as a major player in leukaemogenesis.

**Acknowledgements**

We thank Meinrad Busslinger (IMP, Vienna, Austria) for kindly providing the PAX5 cosmid probes, and Claudia Becher (Campus Kiel, Kiel, Germany) and Elisabeth Lang (St. Anna Kinderkrebsforschung, Vienna, Austria) for excellent technical assistance.

*Grant support:* This work was supported by the Austrian Ministry of Science and Research (GEN-AU II, GZ 200.136/1-VI/1/2005), the St. Anna Kinderkrebsforschung and the Kinderkrebs-Initiative Buchholz/Holm-Seppensen.

**References**


### Table 1. Oligonucleotide Primer Sequences

<table>
<thead>
<tr>
<th>Primer</th>
<th>Oligonucleotide sequence</th>
<th>Direction</th>
<th>Gene/Exon&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Transcript</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAX5ex5-F1</td>
<td>TACTCCATCAGCGGCATCC</td>
<td>sense</td>
<td>PAX5/5</td>
<td>PAX5-PML</td>
</tr>
<tr>
<td>PMLex2-R1</td>
<td>CGCACTCAAGACACCAGAAG</td>
<td>antisense</td>
<td>PML/2</td>
<td></td>
</tr>
<tr>
<td>PAX5ex5-F1</td>
<td>TACTCCATCAGCGGCATCC</td>
<td>sense</td>
<td>PAX5/5</td>
<td>PAX5</td>
</tr>
<tr>
<td>PAX5ex7-R1</td>
<td>GGCTTCTATGTCGTCCAG</td>
<td>antisense</td>
<td>PAX5/7</td>
<td></td>
</tr>
<tr>
<td>PMLex1-F1</td>
<td>CTCAAGGGACTCAGCCAACTG</td>
<td>sense</td>
<td>PML/1</td>
<td>PML</td>
</tr>
<tr>
<td>PMLex2-R1</td>
<td>CGCACTCAAGACACCAGAAG</td>
<td>antisense</td>
<td>PML/2</td>
<td></td>
</tr>
</tbody>
</table>

<sup>1</sup>Exon nomenclature according to the Ensembl Genome Browser exon information (http://www.ensembl.org/).
Figure 1. Fluorescence *in situ* hybridization (FISH) analysis of patient 1.  
(A) Metaphase hybridized with exon-specific PAX5 probes cos-hPAX5-1 (red) and cos-hPAX5-3 (green) showing a disruption of PAX5. Arrows indicate the normal chromosome 9 (orange), the derivative chromosome 9 (red) and the derivative chromosome 15 (green).  
(B) 24-color FISH analysis showing the t(9;15) (inset). Metaphase analysis using probes WCP 9 (red) and RP11-2M12 (green) encompassing the whole promyelocytic leukemia (PML) gene and displaying a disruption of PML. The filled arrow is pointing to derivative chromosome 15, the arrowhead to derivative chromosome 9, and the open arrow to the normal chromosome 15.
Figure 2. Reverse-transcription polymerase chain reaction (RT-PCR) analysis.

(A) RT-PCR of the PAX5-PML fusion transcript (left) and the normal PML and PAX5 alleles. M, molecular weight marker, 100 bp ladder; lane 1, patient 1; lane 2, patient 2; lane 3, negative control (left). Lanes 1, patient 1; lanes 2, normal control; lanes 3, negative control (right). (B) Sequence analysis of the PAX5-PML transcript showing a fusion of PAX5 exon 6 to PML exon 2. (C) Partial nucleotide and amino acid sequence of the PAX5-PML chimera. (D) Schematic representation of the PAX5 and PML wild-type proteins and the putative chimeric PAX5-PML protein. PD, paired domain; O, conserved octapeptide; HD, homeobox domain; TD, transactivation domain; ID, inhibitory domain; P, proline-rich sequence; RING, ‘really interesting new gene’ finger; BBox, B box zinc fingers; COILED, coiled-coil domain; NLS, nuclear localization signal; S/P-rich, serine/proline-rich region;
CHAPTER 3

INCIDENCE AND DIVERSITY OF PAX5 FUSION GENES IN CHILDHOOD ACUTE LYMPHOBLASTIC LEUKEMIA

Karin Nebral\textsuperscript{1}, Dagmar Denk\textsuperscript{1}, Andishe Attarbaschi\textsuperscript{2}, Margit König\textsuperscript{1}, Georg Mann\textsuperscript{2}, Oskar A. Haas\textsuperscript{2}, and Sabine Strehl\textsuperscript{1}


\textsuperscript{1}CCRI, Children's Cancer Research Institute, St. Anna Kinderkrebsforschung, Vienna, Austria;
\textsuperscript{2}Department of Pediatric Hematology and Oncology, St. Anna Children's Hospital, Vienna, Austria
Abstract

*PAX5*, a master regulator of B-cell development, was recently shown to be involved in several leukemia-associated rearrangements, which result in fusion genes encoding chimeric proteins that antagonize PAX5 transcriptional activity. In a population-based fluorescence *in situ* hybridization (FISH) screening study of 446 childhood acute lymphoblastic leukemia (ALL) patients we now show that *PAX5* rearrangements occur at an incidence of about 2.5% of B-cell precursor leukemia (BCP-ALL). Identification of several novel *PAX5* partner genes including *POM121*, *BRD1*, *DACH1*, *HIPK1*, and *JAK2* brings the number of distinct *PAX5* in-frame fusions to at least twelve. Our data show that these not only comprise transcription factors but also structural proteins and genes involved in signal transduction, which at least in part have not been implicated in tumorigenesis.

Keywords

*PAX5* rearrangements, fusion genes, FISH screening, B-cell precursor acute lymphoblastic leukemia
Introduction

The transcription factor *Pax5* encodes the B-cell lineage specific activator protein (BSAP) and is a master regulator of B-cell development. Within the hematopoietic system, *Pax5* is exclusively expressed in the B-lymphoid lineage and is required for progression beyond the pro-B cell stage.1 *Pax5* is not only indispensable for B-lineage commitment, but its continuous expression is also essential to maintain this fate.1-3 During B-cell lineage commitment, *Pax5* fulfils a dual role by repressing B-lineage inappropriate genes and simultaneously activating B-lineage–specific genes.4-6 Loss of *Pax5* expression in pro-B cells by conditional gene inactivation arrests B-cell development at an early pro-B cell stage and reverts committed B-cell precursors with a restricted B-lymphoid potential to progenitors with self-renewal capacity and hematopoietic pluripotency.2,4 The restoration of *Pax5* expression suppresses the multilineage potential of *Pax5*-/- pro-B cells while simultaneously promoting their differentiation to mature B-cells.4

The essential role of *Pax5* for proper B-cell development renders *PAX5* also an intriguing candidate to be involved in B-cell neoplasia. The t(9;14)(p13;q32) found in a subset of B-cell non-Hodgkin's lymphoma juxtaposes the intact coding sequence of *PAX5* under the control of the *IGH@* locus, leading to inappropriate expression of *PAX5*.7,8 Recently, it was also shown that deletion, amplification, point mutation and structural rearrangements in genes encoding regulators of B-lymphocyte development and differentiation occur in 40% of B-progenitor ALL.9 Amongst the affected genes, *PAX5* was the most frequent target of somatic mutation, being altered in about 32% of the cases.9 In ALL *PAX5* is involved in different translocations that result in fusion genes encoding chimeric proteins with novel functions. So far, *ETV6*,10,11 *ELN*,12 *FOXP1*,9 *ZNF521*,9 *PML*,13 *AUTS2*,14 and *C20orf112*14 were identified as *PAX5* fusion partners.

In this population-based FISH screening study of 446 consecutive childhood ALL cases, we determined that in B-cell precursor ALL (BCP-ALL) *PAX5* rearrangements occur at an incidence of about 2.5%. The subsequent identification of the partner genes discovered five novel in-frame *PAX5* fusions to *HIPK1*, *POM121*, *JAK2*, *DACH1*, and *BRD1*, a set of genes with diverse functions including not only transcription factors but also structural proteins and even a tyrosine kinase.

Materials and Methods

Patients

Between June 1999 and December 2007, 486 infants and children with *de novo* acute lymphoblastic leukemia (ALL) were registered in the Austrian ALL-BFM 2000 (n=475) and Interfant-9915 (n=11) studies. The 486 patients also include those only registered but
subsequently for various reasons not treated accordingly. The patients comprised n=65 with T-ALL and n=414 with B-cell leukemia including all immunophenotypes: pro-B ALL (n=20), cALL (n=268), pre-B ALL (n=118), mature ALL (n=5); and (n=3) without specific classification. Further, the study included 4 biphenotypic and one NK-cell leukemia, as well as two ALLs, which were not analyzed in detail. Based on successful routine diagnostic work-up by cytogenetics, FISH and RT-PCR approaches, the n=421 ALLs (excluding only the n=65 T-ALL cases) consisted of 126/400 (31.5%) with high hyperdiploidy (> 50 chromosomes), 38/400 (9.5%) with low hyperdiploidy (47-50 chromosomes), 18/400 (4.5%) with hypodiploidy (< 46 chromosomes) and 113/421 (26.8%) ETV6-RUNX1, 12/418 (2.9%) TCF3-PBX1 (E2A-PBX1), 7/421 (1.7%) BCR-ABL1, and 11/421 (2.6%) MLL positive cases. The entire patient cohort consisted of 272 male and 214 female patients, and the age distribution ranged from 0.18 – 19.3 years (median = 5.6 years). Written informed consent was obtained from the patients, their parents or their legal guardians that surplus material not required for diagnostic purposes may be used for accompanying cancer research projects. Out of these 486 patients from 454 (93.4%) sufficient material for FISH analysis was available. FISH analysis was successful in 446 (98.2%) of the analyzed cases, while in 8 (1.8%) poor quality of the fixed cells and, thus, insufficient hybridization efficiency precluded unambiguous evaluation of the FISH pattern.

**FISH approach for the detection of PAX5 rearrangements**

PAX5 rearrangements were detected using PAX5 flanking BAC clones RP11-12P15 and RP11-22011 (obtained from Pieter de Jong, BACPAC Resources, Children's Hospital and Research Center Oakland, CA, USA). Direct involvement of PAX5 was verified using the exon-specific cosmids cos-hPAX5-1 (exons 2-5) and cos-hPAX5-3 (exons 9-10)\(^7\). For metaphase analysis also RP11-465P6 and RP11-84P7 were applied (obtained from M. Rocchi, Department of Cytogenetics, University of Bari, Bari, Italy). Exact clone positions relative to PAX5 are illustrated in Supplementary Figure S1.

BAC and cosmid DNA was isolated using the PSI-Clone BAC DNA Kit (emp Biotech, Berlin, Germany) and the QIAPrep Spin Miniprep Kit (Qiagen, Vienna, Austria), respectively. BAC and cosmid DNA was then amplified with the TempliPhi Amplification Kit (GE Healthcare, Vienna, Austria) according to the manufacturer’s recommendations. Slides for FISH were prepared from the methanol/acetic acid-fixed cell suspensions used for cytogenetic analysis and incubated in Nonidet P40 (Sigma-Aldrich, Vienna Austria) 0.4% / 2xSSC at 37°C for 1 hour and then immediately dehydrated through an ascending ethanol series followed by 3 minutes of pepsin digestion. FISH was essentially performed as previously described.\(^{16}\) Metaphase images were acquired with a Zeiss Axioplan 2 Imaging fluorescence microscope (Zeiss, Göttingen, Germany) equipped with appropriate filter sets using a Zeiss AxioCam
MRm CCD camera and the Isis version 5.0 SR-6 FISH Imaging System (MetaSystems, Altluessheim, Germany). High-throughput automated interphase FISH spot counting was accomplished using the Axioplan 2 microscope coupled to the Metafer4-Metacyte system, Version V 3.1.122 (MetaSystems). Following automated analyses of 300-400 nuclei per case, each sample was manually reevaluated. Based on the analysis of normal controls and leukemia samples with normal PAX5 status, separation of differentially labeled clones was considered when the distance between the signals was > 10 pixel.

**Identification of PAX5 fusion partners**

PAX5 fusion partners were identified by rapid amplification of cDNA ends (RACE) or fluorescence *in situ* hybridization (FISH) analysis.

**RACE.** Total RNA was extracted from mononuclear cells isolated from bone marrow using the PeqGOLD total RNA Kit (Peqlab Biotechnology, Erlangen, Germany) including an on column DNAse digestion step. 300 ng – 2 µg of total RNA were reverse transcribed using the AMV Reverse Transcriptase and the cDNA synthesis primer provided with the Marathon Kit (Takara Bio Europe/Clontech, Saint-Germain-en-Laye, France) followed by second-strand cDNA synthesis according to the manufacturer’s instructions. Ligation of the Marathon cDNA Adaptor to the ds-cDNA was performed overnight at 16°C. Appropriately diluted adapter-ligated ds-cDNA was amplified using the PAX5 gene-specific forward primer PAX5ex2-3-F1 (5’-TCTTGGCAGGTATTATGAGACAGGAAG-3’) or the reverse primer PAX5ex6-7-R1 (5’-TGGCTGAATACCTCTGTGTCTGTGCT-3’) and the adaptor primer AP1 (5’-CCATCCTAATACGACTCACTATAGGGC-3’). A nested PCR reaction was done with the PAX5-specific primers PAX5ex3-F2 (5’-CAGAGCGGGTGTGTGACAATGAC-3’) or PAX5ex6-R1 (5’-CTGCTGCTGTGTGAACATGAC-3’) and the AP2 (5’-ACTCCTATAGGGCTGAGCGGC-3’) universal primer. PCRs were carried out with a T3000 thermocycler (Biometra, Göttingen, Germany) using the following cycling parameters: 95°C initial denaturation for 1 minute; 5 cycles of 94°C for 15 seconds, 72°C for 5-8 minutes; 5 cycles of 94°C for 15 seconds, 70°C for 5-8 minutes; 25 cycles of 94°C for 15 seconds, 68°C for 5-8 minutes.

RACE products were cut out from the gels, extracted using the PeqGOLD Gel Extraction Kit (Peqlab Biotechnology) and directly sequenced or cloned into the pGEM-T Easy vector (Promega, Mannheim, Germany) and sequenced. Sequencing was performed by Eurofins MWG Operon (Ebersberg, Germany).

**FISH.** The PAX5-C20orf112 fusion was detected using RP11-431F4 and RP11-465P6 (M. Rocchi) encompassing the complete PAX5 gene (Supplementary Figure S1) in combination with RP5-1184F4 (Welcome Trust Sanger Institute; http://www.sanger.ac.uk), which spans the C20orf112 locus.
Reverse Transcription-PCR Analysis

Total RNA was reverse transcribed using 2µg of random hexamers (Amersham) and 200 units Moloney-murine leukemia virus (M-MLV) reverse transcriptase (Invitrogen, Lofer, Austria) at 42°C for 60 minutes. RT-PCR reactions were carried out using Hot Start Taq (Qiagen): initial denaturation at 95°C for 14 minutes, 38-42 cycles at 95°C for 30 seconds, at 60-64°C for 30 seconds, at 72°C for 30-90 seconds, followed by a final elongation at 72°C for 7 minutes. Amplification of full length PAX5 fusion transcripts was accomplished either in one round or in two consecutive nested PCR reactions with Finnzymes Phusion™ Hot Start High-Fidelity DNA Polymerase (Biozym Scientific GmbH, Vienna, Austria). All primer sequences and the combinations used are provided in Supplementary Table 1. PCR products were either cut out from the gels and extracted with the PeqGOLD Gel Extraction Kit (Peqlab Biotechnology) or directly purified using the QiaQuick PCR Purification kit (Qiagen) and sequenced by Eurofins MWG Operon.

Western blotting

Appropriate material for Western blot analysis was only available from one case with a PAX5 rearrangement. The KIS-1 cell line served as PAX5 wild-type positive control. Protein was extracted with standard lyses buffer in the presence of protease inhibitor (Roche). Total proteins were separated by SDS-PAGE on NuPAGE 4-12% Bis-Tris gels (Invitrogen) and transferred to nitrocellulose membrane using the XCell SureLock and the XCell II Blot module (Invitrogen) according to the manufacturer's instructions. After blocking with blocking reagent (Roche) membranes were incubated with an anti-N-terminal PAX5 antibody (ab12000, Abcam, Cambridge, UK). Following incubation with a secondary antibody bands were visualized using the LI-COR Odyssey system (LI-COR Biosciences GmbH, Bad Homburg, Germany). Membrans were stripped with 1% SDS and 25 mM Glycine pH 2 and reincubated with anti-GAPDH antibody (6C5; Santa Cruz Biotechnology Inc., Santa Cruz, CA) and an appropriate secondary antibody.

Gene and exon nomenclature

The gene nomenclature throughout this manuscript follows that approved by the human genome nomenclature committee HUGO (http://www.genenames.org/). Nucleotide reference sequences used for primer design and the description of the of the novel PAX5 partner genes were the following: PAX5, NM_016734; HIPK1, NM_198268; POM121, NM_172020 and OTTHUMT00000252020 POM121-001 (the reference sequence differs from the latter by lack of exon 4 described in Ensembl and the sequences have alternative 3' ends); JAK2, NM_004972; DACH1, NM_080759; BRD1, NM_014577; C20orf112, NM_080616; (National Center of Biotechnology Information [NCBI]).
Results

**FISH screening for PAX5 rearrangements**

Interphase FISH analysis of 446 childhood ALL samples detected FISH patterns suggestive for PAX5 rearrangements in 10 cases. PAX5 rearrangements were exclusively found in B-cell precursor ALL (BCP-ALL) and, thus, the overall frequency in childhood ALL was 2.2% whereas the incidence in BCP-ALL was 2.6%.

In 7 cases (1.6%) the dual-color split-apart assay with PAX5-flanking BAC clones detected a separation of the signals suggesting the presence of a PAX5 rearrangement. In 5 cases (1.1%) deletions of the 3' clone and in 15 (3.4%) of the 5' clone were observed. One single case displayed an additional 3' signal and further analysis confirmed a duplication encompassing the PAX5 3'-end and flanking sequences.

All aberrant cases were further analyzed using PAX5 gene-specific cosmids, which proved the direct involvement of PAX5 in all seven cases showing a split FISH pattern with the BAC clones. Out of the 15 cases with a 5' BAC clone deletion 8 were ETV6-RUNX1 positive and were not further analyzed. One case was not analyzable, three displayed no PAX5 aberration, and three a 5’ deletion suggesting focal PAX5 deletions. In 2 out of the 5 cases with a 3’ BAC clone deletion one entire copy of PAX5 was absent, whereas in three cases a 3’ internal deletion was confirmed indicating a PAX5 fusion associated with a 3’ deletion.

**Identification of PAX5 fusion partners**

In all 10 cases that displayed a FISH pattern suggestive for the presence of a PAX5 fusion (Table 1) we attempted to identify the respective partner gene. In two of the cases, one each with a split FISH pattern and a PAX5 3’ deletion we have previously identified PML (case 1) and ETV6 (case 9) as fusion partners, respectively (Table 1). To determine whether in any of the cases one of the known PAX5 partners was involved PAX5 fusion gene-specific FISH and/or RT-PCR experiments were performed first, and as soon as we were able to unravel a novel fusion partner all cases were retrospectively analyzed.

**Identification of PAX5 fusion partners in PAX5-rearranged leukemia**

In cases 2 and 4, only a minor but significant percentage of the cells (20.8% and 19.3%) displayed a split FISH signal pattern (Table 1). Subsequent evaluation of metaphases clearly showed that the PAX5-flanking probes were separated and both located on chromosome 9p suggesting an inversion event (Figure 1A).

3’ RACE using RNA isolated from the bone marrow obtained at relapse of case 4 resulted in an approximately 2 kb amplification product (data not shown) and sequence analysis
revealed fusion of PAX5 exon 5 to JAK2 exon 19 (Figure 1D). RT-PCR analysis of the diagnostic samples of both cases confirmed the presence of the same PAX5-JAK2 transcripts (Figure 1B). Amplification of the reciprocal JAK2-PAX5 fusion showed multiple splice variants that either included all respective JAK2 exons or lacked exon 18 or exons 17 and 18 (Figure 1C,E). The splice variant containing all JAK2 exons and that lacking exons 17 and 18 resulted in open reading frames. Using primers in the respective first and last coding exons, both PAX5-JAK2 and JAK2-PAX5 full length transcripts could be amplified, which apart from the splice variants described above did not lack any other exons (Figure 3A and B, and data not shown). Western blot analysis revealed expression of the predicted size mutant PAX5 protein as well as wild type PAX5 (Figure 3D).

The putative PAX5-JAK2 chimeric protein contains the paired domain (PD) and the octapeptide (OP) domain of PAX5 and the JAK homology (JH) 1 kinase domain of JAK2. The full length hypothetical reciprocal JAK2-PAX5 fusion protein consists of the JAK2 kinase domains JH2-JH7 fused to the PAX5 homeodomain (HD) and the transactivation (TA) and inhibitory (ID) domains (Figure 4B). The shorter JAK2-PAX5 in-frame isoform would lack the JH2 domain.

In case 3, 3' RACE and direct sequencing of a PCR product revealed fusion of PAX5 exon 5 to the noncoding region of BRD1 exon 1. Subsequent fusion gene-specific RT-PCR verified the data obtained by RACE (Figure 2A). PAX5 exon 5 was joined to 14 bp of the noncoding exon 1 of BRD1 resulting in a putative chimeric protein consisting of the PAX5 PD and the octapeptide domain fused to 4 miscellaneous amino acids and the entire BRD1 protein, which contains highly conserved domains, such as an amino-terminal plant homeodomain (PHD) zinc finger and a bromodomain (Figure 4C). Insufficiency of material prevented amplification of the full length fusion transcript, whose coding region would have an estimated length of 3795 bp.

In case 5, FISH analysis clearly showed a PAX5 rearrangement (Figure 2Bi-ii) but 3' RACE failed to identify the PAX5 fusion partner. However, cloning and sequencing of 5' RACE products revealed that one clone encompassed PAX5 exon 6 fused to the noncoding exon 4 of POM121 (data not shown). These data prompted us to perform PAX5-POM121-specific RT-PCR experiments and indeed PAX5-POM121 transcripts could be amplified (Figure 2Biii). Sequence analysis showed that PAX5 exon 5 was fused to 112 bp of genomic DNA derived from chromosome 12 followed by POM121 exon 5 (Figure 2B) suggesting a complex rearrangement between chromosomes 7, 9, and 12. Exons 1-4 of POM121 are non-coding and the translational start codon is located in exon 5. Nevertheless the insertion of the 112 bp genomic DNA resulted in a complete open reading frame and a putative fusion protein consisting of the PD and the octapeptide domain of PAX5 joined to 88 amino acids neither homologous to PAX5 nor POM121, and the entire POM121 protein (Figure 4D). Owing to the
lack of appropriate material the full length coding transcript, which would have a calculated size of 3867 bp could not be amplified and, thus, it remains elusive whether all exons are retained or different splice variants of PAX5-POM121 are expressed.

In case 6, FISH experiments showed that the 5'-end of PAX5 was located on a der(9) chromosome whereas the 3'-end was translocated to 14q32 (data not shown). Further thorough FISH analysis confirmed the presence of complex rearrangements involving at least chromosomes 9, 3, 11, and 12 with insertion of chromosome 11p material into 9p (data not shown) indicating fusion of PAX5 with a gene located on 11p. However, though 20 BAC clones, which encompassed 11p13-15 were hybridized none of them showed a co-localization with PAX5 and, thus, this strategy failed to identify any candidate gene. Consequently, also in this case 3' and 5' RACE was performed but despite extensive efforts we were unable to identify the PAX5 partner gene.

In case 7, FISH analysis showed a split signal for the PAX5-specific clones and 1-2 additional 5' signals (Figure 2Ci). Further FISH analysis again indicated complex aberrations involving several chromosomes, in particular 1p, which was translocated to 9p (data not shown). Cloning and sequencing of an approximately 1.8 kb 3' RACE product suggested involvement of the HIPK1 gene. Fusion gene-specific RT-PCR experiments with primers located in exons 5 and exons 9/10 of PAX5 and HIPK1, respectively, confirmed the RACE data (Figure 2Cii). Amplification of the fusion transcript with primers located in exons 1 and 5 of PAX5 and at the junction of the last coding exons 15/16 of HIPK1 showed that PAX5 exons 1-5 were consistently present, but HIPK1 C-terminal exons were alternatively spliced and, thus, several variants are expressed (data not shown). Reciprocal HIPK1-PAX5 transcripts could not be detected.

The putative PAX5-HIPK1 chimeric protein encoded by the notional full length transcript consists of the PD and the octapeptide domain of PAX5 fused to a part of the homeodomain-interacting domain (ID), the Proline-, Glutamic acid-, Serine-, Threonine-rich (PEST), and the tyrosine/histidine-rich (YH) domains of HIPK1 (Figure 4E).

Identification of PAX5 fusion partners in cases with PAX5 3' deletions

In three out of the 446 childhood ALL patients FISH analysis with PAX5-specific cosmid clones displayed 3' deletions of PAX5 also indicating the presence of PAX5 rearrangements. In this respect, the dic(9;12)/PAX5-ETV6 aberration results in loss of the PAX5 3'-end,11 and the PAX5 fusion partners FOXP1, ZNF521, AUTS2, and C20orf112 were detected based on array CGH data that only permit the delineation of unbalanced genetic alterations.9,14

In case 8, fusion gene-specific FISH and RT-PCR assays for the known PAX5 rearrangements revealed that in this case PAX5 was fused to C20orf11214 (Figure 2D). Sequence analysis of the PAX5-C20orf112 transcripts showed an in-frame fusion of PAX5
Chapter 3

**PAX5 fusion genes in childhood ALL**

exon 8 to **C20orf112** exon 12. Amplification of the full length coding chimeric transcript showed no alternative splicing of any of the exons (data not shown). At least in this case, the PAX5-C20orf112 fusion displays the most 3’ breakpoint within PAX5 described so far, which joins almost the entire PAX5 protein including the PD, the octapeptide, the homeodomain (HD) and parts of the transactivation domain (TA) to the C-terminal end of C20orf112. Owing to the opposite transcriptional orientations of **PAX5** (centromere-telomere) and **C20orf112** (telomere-centromere), the generation of a functional fusion gene requires a complex genetic rearrangement or the formation of a dicentric chromosome.

In case 10, interphase FISH analysis showed deletion of the PAX5 3’-end (Figure 2Ei) and direct sequencing of one of several 3’ RACE PCR products revealed an in-frame fusion of **PAX5** exon 5 to **DACH1** exon 5 (Figure 2Eii). Subsequent RT-PCR experiments using a forward primer located in exon 5 of **PAX5** and a set of reverse primers in exons 5, 8, 9, and 12 of **DACH1** showed the formerly described alternative splicing of **DACH1** skipping exons 4, 5, 6, or 7 or a combination thereof.\(^{17}\) Fusion of **PAX5** to a **DACH1** isoform that was only detected in spleen could not be verified. Sequencing of RT-PCR products amplified with primers located in the first and last coding exon of the respective genes consistently showed the presence of **PAX5** exons 1-5 and confirmed splicing of **DACH1** (Figure 3C). Also in this case the centromere-telomere orientation of **PAX5** and the opposite telomere-centromere transcriptional direction of **DACH1** suggested a more complex rearrangement rather than a simple reciprocal translocation. The PAX5-DACH1 putative consensus fusion protein consisted of the PD and the octapeptide domain and the C-terminal conserved DD2 domain of **DACH1** (Figure 4F).

**Discussion**

Using a FISH approach for the detection of **PAX5** rearrangements, we performed a population-based screening of 446 consecutive childhood ALL cases and identified 10 (2.6%) BCP-ALL patients with a **PAX5** rearrangement. All **PAX5** fusion positive cases were negative for the most common genetic aberrations found in childhood ALL (**ETV6-RUNX1**, **BCR-ABL1**, **TCF3-PBX1**, **MLL-AF4**) and, thus, in contrast to **PAX5** deletions, they are most likely distinctive primary genetic events.

**PAX5** rearrangements were particularly associated with a common ALL phenotype. Except for the **PAX5-POM121** positive case, all patients showed a good response to prednisone according to the ALL-BFM 2000 or Interfant-99\(^{15}\) protocol. Based on prednisone response and MRD risk stratification, all patients were treated with the respective therapy regimen and 9/10 patients are in first complete remission 6-84 months from diagnosis. Only one of the two **PAX5-JAK2** positive patients with a pre-B phenotype relapsed 2.5 years after initial diagnosis.
but after recommencing therapy has achieved a second complete remission (Supplementary Table 2).

So far, the majority of PAX5 rearrangements was detected by high-resolution single nucleotide polymorphism (SNP) array analysis,\textsuperscript{9,14} a technology which, however, only allows for the detection of unbalanced aberrations and precludes the identification of balanced reciprocal translocations or inversions. Therefore, we took an alternative approach and developed a FISH screening assay that permits the unambiguous detection of all PAX5 rearrangements independent of their balanced or unbalanced nature, including even those that would result in juxtaposition of PAX5 under the regulatory elements of a partner gene as seen in the PAX5-IGH@ translocation.\textsuperscript{7,8} It is interesting to note that for the most PAX5 rearrangements were found in cases with either normal or complex karyotypes, a fact that prevents detection by conventional cytogenetics and emphasizes their often cryptic nature.

Further, at least in childhood ALL, we did not find any evidence for PAX5 activating translocations.

Two of the ten PAX5-rearranged cases have been previously reported to harbor a PAX5-ETV6\textsuperscript{11} and a PAX5-PML\textsuperscript{13} aberration, and one showed the recently described PAX5-C20orf112 fusion.\textsuperscript{14} However, we identified five hitherto unknown PAX5 fusion partners, namely HIPK1, POM121, JAK2, DACH1, and BRD1 bringing the number of distinct PAX5 chimera to at least twelve (Table 2). Similar to all previously described PAX5 rearrangements, the majority of the fusion transcripts encode putative novel transcription factors, which consist of at least the amino-terminal paired DNA-binding domain and in most instances also the octapeptide of PAX5, and C-terminal regulatory sequences of a second transcription factor. However, involvement of ELN and POM121 as structural proteins and the tyrosine kinase JAK2 are remarkable as neither of these genes is directly implicated in transcriptional regulation.

The Janus kinase (JAK) family currently comprises four human members JAK1, JAK2, JAK3, and TYK2, which are receptor associated protein tyrosine kinases and are of critical importance for cytokine-mediated signal transduction.\textsuperscript{18,19} Somatically acquired activating mutations in JAK2 were recently reported to play a central role in the pathogenesis of myeloproliferative disorders.\textsuperscript{20} Further, JAK2 fusions with ETV6, BCR, PCM1, and SSBP2 were described in a variety of hematopoietic malignancies.\textsuperscript{21,22} The transforming potential of the previously described JAK2 fusion proteins has been attributed to the cytokine-independent constitutive activation of JAK2, mediated by motifs of the partner gene that serve as dimerization/oligomerization interfaces.\textsuperscript{21} For example, for the three distinct ETV6-JAK2 fusions, which are potent activators of STAT5, transformation is strictly dependent on the ETV6 pointed (PNT) self-association domain.\textsuperscript{23} However, there is no evidence that the PAX5 domains retained in the PAX5-JAK2 protein are capable to mediate dimerization. In
contrast, both the DNA-binding domain and the nuclear localization signal (NLS) of PAX5 are retained in the PAX5-JAK2 fusion protein, which is suggestive of a nuclear localization, whereas JAK2-PAX5 may reside in the cytoplasm.

BRD1, BRPF1, MLLT6, and MLLT10 belong to a small evolutionary conserved family of putative nuclear transcription factors, which share a highly homologous cysteine-rich region containing an amino-terminal PHD finger motif. Two members of this family are involved in myeloid leukemia-associated rearrangements, namely MLL-MLLT6 and MLL-MLLT10. However, the respective chimeric proteins differ considerably from PAX5-BRD1 in that the conserved PHD domains of MLLT6 and MLLT10 are lost, whereas the entire BRD1 protein is fused to PAX5. Even though the actual function of BRD1 itself remains elusive, the presence of a PHD-bromodomain module, which is frequently found in chromatin-associated proteins, strongly indicates a role in chromatin remodelling and epigenetic regulation of gene transcription.

Considering that PAX5 has the capability to activate and suppress large sets of genes, this potential feature of BRD1 supports the notion of the PAX5-BRD1 chimera to modulate transcriptional activity.

HIPK1 belongs to the homeodomain-interacting protein kinase (HIPK) family, whose currently four members (HIPK1-4) are nuclear serine/threonine kinases that are primarily localized in the nucleus. The HIPKs were originally identified as nuclear protein kinases that function as corepressors for various homeodomain-containing transcription factors but recently were also shown to interact with other proteins involved in apoptosis and signal transduction in a cellular localization-dependent manner. HIPK1 physically interacts with and promotes phosphorylation of e.g. TP53, DAXX, EP300, and RUNX1. HIPK1 also regulates the nuclear export of DAXX and both proteins collaborate in transcriptional regulation, a functional aspect of HIPK1, which is further substantiated by its modulation of TP53 activity. On the other hand, in the cytoplasm HIPKs appear to transduce signals by death receptors and to induce MAP3K5 dependent apoptosis.

Owing to the multiple functions of HIPK1 it is intricate to ascribe any potential specific function to the PAX5-HIPK1 fusion protein, however, the most likely one is also transcriptional regulation.

DACH1 is a human homologue of the Drosophila dachshund (dac) gene, which is a key regulator of cell fate determination during eye, leg, and brain development in the fly. Members of the dachshund family of nuclear proteins encode highly conserved putative transcription factors, which contribute to the fundamental mechanisms of morphogenesis.

The DACH1 protein contains two domains (DD1, Dachbox N-domain and DD2, Dachbox C-domain or EYA domain), which are highly conserved from Drosophila to human. DACH1 functions as a transcriptional repressor of TGF-β-signaling in breast and ovarian cancer. DACH1 is also a physiological regulator of endogenous JUN function, inhibiting JUN and JUN target gene expression, as well as a CCND1 repressor. Although target gene
repression by DACH1 requires the DD1 domain it is tempting to speculate that the PAX5-DACH1 chimeric protein, in spite of the fact that it lacks this conserved domain, may act as transcriptional repressor of PAX5 activated target genes.

PAX5 fusion partners, however, not only comprise transcription factors but also structural proteins such as ELN\textsuperscript{12} and POM121. POM121 is one of the two integral pore membrane proteins that were identified as specific components of nuclear pore complexes of higher eukaryotes.\textsuperscript{45} Both pore membrane proteins, NUP210\textsuperscript{46} and POM121\textsuperscript{45} have been proposed, although controversially discussed, to play important roles in nuclear pore complex (NPC) formation and anchoring the peripheral nucleoporins (NUPs) to the nuclear membrane.\textsuperscript{47,48}

The NPC is composed of multiple copies of about 30 different NUPs,\textsuperscript{49} and so far, only two NUPs, namely NUP98\textsuperscript{50,51} and NUP214,\textsuperscript{52-54} were found involved in leukemia-associated translocations, which makes POM121 only the third component of the nuclear envelope implicated in leukemogenesis.

All PAX5 fusion proteins contain the PAX5 DNA-binding domain and, thus, are predicted to retain the ability to bind to PAX5 transcriptional targets, but no longer provide normal transcriptional regulatory functions.\textsuperscript{55} PAX5-ETV6, PAX5-FOXP1, and also PAX5-ELN indeed competitively inhibit the transcriptional activity of PAX5 suggesting that PAX5 fusions act as constitutive repressors to antagonize PAX5 function provided by the second, wild-type PAX5 allele.\textsuperscript{9,12,56} Comprehensive studies are now required to elucidate whether all PAX5 chimera in fact operate as aberrant transcription factors that impair the finely tuned PAX5 target gene transcriptional network.

Acknowledgements

This work was supported by a grant of the Austrian Ministry of Science and Research (GEN-AU II, GZ 200.136/1-VI/1/2005) (to S.S.) and the St. Anna Kinderkrebsforschung. We thank Meinrad Busslinger (IMP, Vienna, Austria) for kindly providing the PAX5 cosmid probes and Tilman Johannes (MetaSystems, Altlussheim, Germany) for assistance with the Metafer4-Metacyte system. We would also like to thank all those people, who perform the routine diagnostic work-up and consistently provide the basis for our research. Further, we are indebted to Dasa Janousek for the efficient clinical data management and analysis.
References


Table 1. Summary of all PAX5-rearranged childhood ALL cases registered in the Austrian ALL-BFM 2000 and Interfant-99 studies between June 1999 and December 2007.

<table>
<thead>
<tr>
<th>Case</th>
<th>PAX5 FISH pattern</th>
<th>% positive cells</th>
<th>PAX5 fusion</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>split</td>
<td>49.1</td>
<td>PAX5-PML</td>
<td>Nebral et al\textsuperscript{13}</td>
</tr>
<tr>
<td>2</td>
<td>split</td>
<td>20.8</td>
<td>PAX5-JAK2</td>
<td>this work</td>
</tr>
<tr>
<td>3</td>
<td>split</td>
<td>30.0</td>
<td>PAX5-BRD1</td>
<td>this work</td>
</tr>
<tr>
<td>4</td>
<td>split</td>
<td>19.3</td>
<td>PAX5-JAK2</td>
<td>this work</td>
</tr>
<tr>
<td>5</td>
<td>split</td>
<td>59.7</td>
<td>PAX5-POM121</td>
<td>this work</td>
</tr>
<tr>
<td>6</td>
<td>split</td>
<td>45.8</td>
<td>PAX5-?</td>
<td>this work</td>
</tr>
<tr>
<td>7</td>
<td>split</td>
<td>61.7</td>
<td>PAX5-HIPK1</td>
<td>this work</td>
</tr>
<tr>
<td>8</td>
<td>3' deletion</td>
<td>85.0</td>
<td>PAX5-C20orf112</td>
<td>this work</td>
</tr>
<tr>
<td>9</td>
<td>3' deletion</td>
<td>ND</td>
<td>PAX5-ETV6</td>
<td>Strehl et al\textsuperscript{11}</td>
</tr>
<tr>
<td>10</td>
<td>3' deletion</td>
<td>78.2</td>
<td>PAX5-DACH1</td>
<td>this work</td>
</tr>
</tbody>
</table>

ND, not determined
Table 2. PAX5 fusion partners in B-cell precursor ALL.

<table>
<thead>
<tr>
<th>Partner gene</th>
<th>Chr.</th>
<th>Localization/Function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIPK1 homeodomain interacting protein kinase 1</td>
<td>1p13</td>
<td>nuclear (speckles) transcription regulation</td>
<td>this work</td>
</tr>
<tr>
<td>FOXP1 forkhead box P1</td>
<td>3p13</td>
<td>nuclear transcription factor</td>
<td>Mullighan et al&lt;sup&gt;9&lt;/sup&gt;</td>
</tr>
<tr>
<td>POM121 membrane glycoprotein</td>
<td>7q11</td>
<td>nuclear pore membrane assembly of nuclear envelope</td>
<td>this work</td>
</tr>
<tr>
<td>ELN elasitin</td>
<td>7q11</td>
<td>extracellular matrix, elastic fibers structural protein</td>
<td>Bousquet et al&lt;sup&gt;12&lt;/sup&gt;</td>
</tr>
<tr>
<td>AUTS2 autism susceptibility candidate 2</td>
<td>7q11</td>
<td>intracellular unknown</td>
<td>Kawamata et al&lt;sup&gt;14&lt;/sup&gt;</td>
</tr>
<tr>
<td>JAK2 Janus kinase 2</td>
<td>9p24</td>
<td>cytoplasmic tyrosine kinase, receptor signaling</td>
<td>this work</td>
</tr>
<tr>
<td>ETV6 ets variant gene 6 (TEL oncogene)</td>
<td>12p13</td>
<td>nuclear transcription repressor</td>
<td>Cazzaniga et al&lt;sup&gt;10&lt;/sup&gt;</td>
</tr>
<tr>
<td>DACH1 dachshund homolog 1</td>
<td>13q24</td>
<td>nuclear transcription factor</td>
<td>Strehl et al&lt;sup&gt;11&lt;/sup&gt;</td>
</tr>
<tr>
<td>PML promyelocytic leukemia</td>
<td>15q21</td>
<td>nuclear, PML bodies multiple functions</td>
<td>Nebral et al&lt;sup&gt;13&lt;/sup&gt;</td>
</tr>
<tr>
<td>ZNF521 zinc finger protein 521</td>
<td>18q11</td>
<td>nuclear transcription factor</td>
<td>Mullighan et al&lt;sup&gt;9&lt;/sup&gt;</td>
</tr>
<tr>
<td>C20orf112 chromosome 20 open reading frame 112</td>
<td>20q11</td>
<td>unknown</td>
<td>Kawamata et al&lt;sup&gt;14&lt;/sup&gt;</td>
</tr>
<tr>
<td>BRD1 bromodomain containing 1</td>
<td>22q13</td>
<td>nuclear putative transcription factor</td>
<td>this work</td>
</tr>
</tbody>
</table>
Figure 1. PAX5-JAK2 rearrangement. (A) FISH with BAC clones RP11-220I1 (CY3) and RP11-12P15 (FITC) showing a split signal. Arrows indicate the normal chromosome 9 (black) and the derivative chromosome 9 (white). (B) RT-PCR analysis of the PAX5-JAK2 fusion transcript. Lane 1, patient 4; lane 2, patient 2; lane 3, patient 6; lane 4, normal control. Case 6 was negative for the PAX5-JAK2 fusion. (C) RT-PCR analysis of the JAK2-PAX5 transcript showing multiple splice variants. Lane 1, patient 4; lane 2, patient 2; lane 3, normal control; lane 4, negative control. (D-E) Sequence chromatograms of PAX5-JAK2 and JAK2-PAX5 transcripts. (D) fusion of PAX5 exon 5 to JAK2 exon 19 and (E) JAK2 exon 18 to PAX5 exon 6. M, molecular weight marker, 100 bp ladder (Promega).
Figure 2. FISH, RT-PCR, RACE, and sequence analyses of PAX5-rearranged patients. (A) Case 3. RT-PCR and sequence analysis showing fusion of PAX5 exon 5 to BRD1 exon 1. Lane 1, patient 3; lane 2, normal control; lane 3, negative control. (B) Case 5. (Bi) Representative interphase nuclei displaying a PAX5 split signal using RP11-220I1 (CY3) and RP11-12P15 (FITC); (Bii) metaphase showing a PAX5 split signal using RP11-465P6 (FITC) and RP11-84P7 located downstream of PAX5 (CY3). Arrows indicate the normal chromosome 9 (black) and the derivative chromosomes (white). (Biii) RT-PCR showing the PAX5-POM121 fusion transcript. Lane 1, patient 5; lane 2, negative control. Sequence chromatograms presenting fusion of PAX5 exon 5 to 112 bp derived from chromosome 12 followed by POM121 exon 5. (C) Case 7. (Ci) Representative aberrant interphase nuclei hybridized with RP11-220I1 (CY3) and RP11-12P15 (FITC). (Cii) RT-PCR of PAX5-HIPK1. Lane 1, patient 7; lane 2, negative control, and sequence chromatogram showing fusion of PAX5 exon 5 to HIPK1 exon 9. (D) Case 8. RT-PCR and sequence chromatograms demonstrating fusion of PAX5 exon 8 to C20orf112 exon 8. Lane 1, patient 3; lane 2, normal control; lane 3, negative control. (E) Case 10. (Ei) Representative interphase nuclei hybridized with RP11-220I1 (CY3) and RP11-12P15 (FITC) displaying a PAX5 3’-end deletion and (Eii) RT-PCR showing PAX5-DACH1 transcripts. Lane 1, patient 10; lane 2, normal control; lane 3, negative control, and sequence chromatogram presenting fusion of PAX5 exon 5 to DACH1 exon 5. M, molecular weight marker, 100 bp ladder (Promega).
Figure 3. Amplification of full length fusion transcripts and Western blot analysis. (A-C) Amplification of full length chimeric transcripts. (A) PAX5-JAK2 and (B) reciprocal JAK2-PAX5 transcripts of case 4. Lane 1, patient; lane 2, negative control. (Owing to the small size differences of the splice variants these are not distinguishable from the full length transcript.) (C) Case 10. PAX5-DACH1 full length fusion transcript and several splice variants are expressed. M,peqGOLD Ladder-Mix (Peqlab Biotechnology). (D) Western blot of primary leukemic blasts of case 4 using an anti-N-terminal PAX5 antibody. Lane 1, KIS-1 cell line; lane 2, PAX5-JAK2 positive case. Asterisk indicates the mutant PAX5-JAK2 fusion protein. GAPDH served as loading control. M, SeeBlue® Plus2 Pre-Stained Standard (Invitrogen), sizes in kDa are indicated.
Figure 4. Schematic representation of PAX5, partner wild-type and the putative chimeric PAX5 fusion proteins.

Wild-type proteins are always depicted below the PAX5 chimera. (A) PAX5 wild-type protein. PD, paired domain; OP, octapeptide domain; HD, homeodomain; TA, transactivation domain; I, inhibitory domain; NLS, nuclear localization signal. (B) PAX5-JAK2 and the reciprocal full-length JAK2-PAX5 fusion proteins. JH1-7, JAK homology domains 1-7. (C) BRD1 wild-type and PAX5-BRD1 fusion protein with the insertion of 4 miscellaneous amino acids. PHD, plant homeodomain zinc finger domain; BROMO, bromodomain; PWWP, proline-tryptophan-tryptophan-proline motif. (D) POM121 wild-type and PAX5-POM121 fusion protein, insertion of the 88 novel amino acids encoded by the 112 bp of genomic DNA derived from chromosome 12 and the normally untranslated region of POM121 exon 5 is depicted. TRANS, potential transmembrane domain; FXFG, repetitive FXFG pentapeptide motif; according to Ensembl POM121-001. (E) HIPK1 wild-type and PAX5-HIPK1 fusion protein. KINASE, protein kinase domain; ID, homeodomain-interacting domain; PEST, Prolin-, Glutamic acid-, Serine-, Threonine-rich sequence; YH, tyrosine/histidine-rich motif. (F) DACH1 wild-type and PAX5-DACH1 fusion protein. DD1, Dachbox N-domain; DD2, Dachbox C-domain.
### Supplementary Table 1. Oligonucleotide Primer Sequences

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Direction</th>
<th>Gene/Exon</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAX5ex5-F1</td>
<td>TACTCCATCAGCGGGAATCC</td>
<td>sense</td>
<td>PAX5/5</td>
</tr>
<tr>
<td>PAX5ex6-R2</td>
<td>CTGCTGCTGTGTAACAGTC</td>
<td>antisense</td>
<td>PAX5/6</td>
</tr>
<tr>
<td>PAX5ex7-R1</td>
<td>GCCCTTCATGTCGACTCCAG</td>
<td>antisense</td>
<td>PAX5/7</td>
</tr>
<tr>
<td>PAX5ex10non-R1</td>
<td>AGTCCCTGGAAGGAGAGAGG</td>
<td>antisense</td>
<td>PAX5/10</td>
</tr>
<tr>
<td>P1-PAX5ex1-F1</td>
<td>ATGGATTTAGAAATATCCGACT</td>
<td>sense</td>
<td>PAX5/1</td>
</tr>
<tr>
<td>P1-PAX5ex10-R1</td>
<td>TCAGTGACGCTAGCCGAGCAG</td>
<td>antisense</td>
<td>PAX5/10</td>
</tr>
<tr>
<td>JAK2ex3-F1</td>
<td>AAGACTCTGCATGGGATGCA</td>
<td>sense</td>
<td>JAK2/3</td>
</tr>
<tr>
<td>JAK2ex20-21-R1</td>
<td>TACGCCGACCAGACCTGTAGG</td>
<td>antisense</td>
<td>JAK2/20-21</td>
</tr>
<tr>
<td>JAK2-forward</td>
<td>ACGGTCAAATGGAATGCAACTGAGAACA</td>
<td>sense</td>
<td>JAK2/13</td>
</tr>
<tr>
<td>P1-JAK2ex3-F1</td>
<td>ATGGGAATGGCCATGCC</td>
<td>sense</td>
<td>JAK2/3</td>
</tr>
<tr>
<td>P1-JAK2ex25-R1</td>
<td>TCATCCAGCCAATCTATCCG</td>
<td>antisense</td>
<td>JAK2/25</td>
</tr>
<tr>
<td>BRD1ex1-R4</td>
<td>TCTCGAAGCGGTCATCCAG</td>
<td>antisense</td>
<td>BRD1/1</td>
</tr>
<tr>
<td>POM121ex13-R3</td>
<td>GCAGGCAAGGTTAAGGTAATG</td>
<td>antisense</td>
<td>POM121/13</td>
</tr>
<tr>
<td>POM121-ex3-4-F1</td>
<td>GTCCAGCGCTCAGCCTGCTTC</td>
<td>sense</td>
<td>POM121/3-4</td>
</tr>
<tr>
<td>HIPK1ex9-10-R1</td>
<td>CTTCCCTGCGTGAATCCGT</td>
<td>antisense</td>
<td>HIPK1/9-10</td>
</tr>
<tr>
<td>HIPK1ex15-16-R1</td>
<td>TGCTGGTTCTGCTGATAGGTG</td>
<td>antisense</td>
<td>HIPK1/15-16</td>
</tr>
<tr>
<td>C20orf112ex8-R1</td>
<td>AGCAGGGAAGGCGGAGCAGACTC</td>
<td>antisense</td>
<td>C20orf112/8</td>
</tr>
<tr>
<td>DACH1ex5-R1</td>
<td>GTGTTCTCAGCTGCTATTGC</td>
<td>antisense</td>
<td>DACH1/5</td>
</tr>
<tr>
<td>DACH1ex9-R1</td>
<td>GCCAAGCCTGTCTCAAGCTGTTT</td>
<td>antisense</td>
<td>DACH1/9</td>
</tr>
<tr>
<td>P1-DACH1ex12-R1</td>
<td>TCAGTACATGACAGTGGTTTCAATACAG</td>
<td>antisense</td>
<td>DACH1/12</td>
</tr>
</tbody>
</table>
Detection of Translocations by Reverse Transcription-PCR Analysis

Fusion transcripts were amplified using the following primer combinations.

PAX5-JAK2: PAX5ex5-F1 and JAK2ex20-21-R1.
JAK2-PAX5: JAK2-forward and PAX5ex6-R2.

PAX5-BRD1: PAX5ex5-F1 and BRD1ex1-R4.

PAX5-POM121: PAX5ex5-F1 and POM121ex13-R3.

POM121-PAX5: POM121ex3-4-F1 and PAX5ex7-R1.
PAX5-HIPK1: PAX5ex5-F1 and HIPK1ex9-10-R1; PAX5ex5-F1 and HIPK1ex15-16-R1.
PAX5-C20orf112: PAX5ex5-F1 and C20orf112ex8-R1.
PAX5-DACH1: PAX5ex5-F1 and DACH1ex5-R1; PAX5ex5-F1 and DACH1ex9-R1.

Detection of full length fusion transcripts by Reverse Transcription-PCR Analysis

PAX5-DACH1: P1-PAX5ex1-F1 and P1-DACH1ex12-R1
PAX5-JAK2: P1-PAX5ex1-F1 and P1-JAK2ex25-R1

JAK2-PAX5: JAK2ex3-F1 and PAX5ex10non-R1; P1-JAK2ex3-F1 and P1-PAX5ex10-R1
(nested PCR reaction)
Supplementary Figure S1.

(A) **PAX5 FISH screening assay.** Adapted screenshot from the UCSC Genome Browser on Human March 2006 Assembly ([www.genome.ucsc.edu](http://www.genome.ucsc.edu)), which displays the localization of the used PAX5 BAC clones RP11-12P15 and RP11-22011, and the cosmid clones coshPAX5-1 and coshPAX5-3.

(B) **FISH clones - PAX5 locus.** Adapted screenshot from the UCSC Genome Browser on Human March 2006 Assembly ([www.genome.ucsc.edu](http://www.genome.ucsc.edu)) showing approximately 800 kb of chromosomal region at 9p13.2. The utilized FISH clones RP11-84P7, RP11-431F4 and RP11-465P6 are depicted.
**Supplementary Table 2.**

Clinical and immunophenotypic characteristics of the PAX5-rearranged cases.

<table>
<thead>
<tr>
<th>Case</th>
<th>Age (yrs)/Sex</th>
<th>Phenoype</th>
<th>Clinical status</th>
<th>Prednisone response</th>
<th>MRD risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.8/F</td>
<td>cALL</td>
<td>CR 84 mo +</td>
<td>good</td>
<td>ND*</td>
</tr>
<tr>
<td>2</td>
<td>6.9/F</td>
<td>cALL</td>
<td>CR 65 mo +</td>
<td>good</td>
<td>IR</td>
</tr>
<tr>
<td>3</td>
<td>2.2/M</td>
<td>cALL</td>
<td>CR 44 mo +</td>
<td>good</td>
<td>SR</td>
</tr>
<tr>
<td>4</td>
<td>9.6/M</td>
<td>pre-B ALL</td>
<td>1st CR 29 mo 2nd CR 10 mo +</td>
<td>good</td>
<td>IR</td>
</tr>
<tr>
<td>5</td>
<td>2.0/M</td>
<td>pre-B ALL</td>
<td>CR 20 mo +</td>
<td>poor</td>
<td>HR</td>
</tr>
<tr>
<td>6</td>
<td>3.4/M</td>
<td>cALL</td>
<td>CR 11 mo +</td>
<td>good</td>
<td>IR</td>
</tr>
<tr>
<td>7</td>
<td>3.4/M</td>
<td>cALL</td>
<td>CR 6 mo +</td>
<td>good</td>
<td>IR</td>
</tr>
<tr>
<td>8</td>
<td>1.8/M</td>
<td>cALL, CD117+</td>
<td>CR 71 mo +</td>
<td>good</td>
<td>IR</td>
</tr>
<tr>
<td>9</td>
<td>12.0/F</td>
<td>cALL</td>
<td>CR 71 mo +</td>
<td>good</td>
<td>IR</td>
</tr>
<tr>
<td>10</td>
<td>4.8/M</td>
<td>cALL, My+</td>
<td>CR 23 mo +</td>
<td>good</td>
<td>IR</td>
</tr>
</tbody>
</table>

* Interfant-99, standard risk

# Good prednisone response: less than 1000/µl peripheral blasts after a 7-day prephase with prednisone and one intrathecal dose of methotrexate on day 1 according to the ALL-BFM 2000 or Interfant-99 protocol.

CR, complete remission
HR, high risk
IR, intermediate risk
mo, months
MRD, minimal residual disease
ND, not done
SR, standard risk
yrs, years
CHAPTER 4

MONOALLELIC LOSS AND FREQUENT MUTATION OF THE SECOND ALLELE OF PAX5 IN DIC(9;20) CHILDHOOD ACUTE LYMPHOBLASTIC LEUKEMIA

Daniela Krehan, Karin Nebral, Margit König, Dagmar Denk, Sabine Strehl

in preparation

CCRI, Children's Cancer Research Institute, St. Anna Kinderkrebsforschung, Vienna, Austria
Abstract

The presence of a dic(9;20) is a characteristic abnormality found in roughly 1.5% of childhood B-cell precursor acute lymphoblastic leukemia (BCP-ALL). Although in about 40% of the cases this is the sole structural rearrangement and, thus, considered as the primary leukemogenic event little is known about the underlying molecular genetic lesions. In order to determine whether PAX5 is implicated in the leukemogenesis of this genetically poorly characterized leukemia subtype, several cases were analyzed for their PAX5 status. Our studies show that dic(9;20) leukemia is consistently associated with deletion of one copy of PAX5 and frequent concomitant mutation of the retained allele.
Introduction

Approximately 1.5% of childhood B-cell precursor acute lymphoblastic leukemia (BCP-ALL) is associated with the presence of a dic(9;20)(p13;q11).\(^1\)\(^2\) In roughly 40% of the cases the dic(9;20) is the sole karyotypic change suggesting that this alteration is the primary leukemogenic event.\(^1\)\(^2\) However, so far all efforts to delineate the gene(s) involved in this specific rearrangement have failed. This is mainly due to the heterogeneity of the breakpoints and, thus, neither a single affected gene on either chromosome nor the generation of a fusion gene has been observed.\(^3\)\(^4\) Owing to the fact that the dic(9;20) aberration consistently results in partial monosomy 9p but not always 20q\(^1\)\(^3\) the crucial gene(s) are more likely located at 9p. Accordingly, a dic(9;20) is frequently associated with hetero- or homozygous deletion of CDKN2A.\(^1\)\(^3\)

The transcription factor PAX5, located at 9p13, encodes the B-cell lineage specific activator protein (BSAP) and is a master regulator of B-cell development.\(^5\)\(^6\) Recently, it was shown that hypodiploid ALL has one null PAX5 allele and a significant proportion of cases harbor point mutations in the second allele.\(^7\)

Here we show that dic(9;20), leukemia in general accompanied by hypodiploidy,\(^8\) which, however, can be masked by nonrandom gains of chromosomes\(^1\) is associated with consistent loss of one PAX5 copy and frequent mutation of the retained allele.

Patients, Material and Methods

Patients

Seven cases with de novo ALL were selected based on cytogenetic evidence of a dic(9;20)\(^2\) (Table 1). Written informed consent that surplus material not required for diagnostic purposes may be used for research purposes was obtained from the patients’ parents or their legal guardians.

FISH analysis

Samples were hybridized with: centromere-specific probes for chromosomes 9 (D9Z5; Oncor/Qbiogene, Heidelberg, Germany) and 20 (D20Z1; M. Rocchi, Department of Cytogenetics, University of Bari, Italy); PAX5 flanking BAC clones RP11-12P15 and RP11-220I1 (Pieter de Jong, BACPAC Resources, Children's Hospital and Research Center Oakland, CA, USA); PAX5 exon-specific cosmids cos-hPAX5-1 (exons 2-5) and cos-hPAX5-3 (exons 9-10);\(^9\) Vysis LSI p16 (9p21) SpectrumOrange/CEP9 Spectrum Green (Abbott Molecular, Vienna Austria). FISH was performed as described.\(^10\)
PAX5 mutation analysis

Mutations in all coding exons and flanking intronic sequences of PAX5 were identified by direct sequencing (Eurofins MWG Operon, Ebersberg, Germany) of PCR amplified genomic DNA. Primer sequences and PCR conditions are available upon request.

Identification of PAX5 splice variants

PAX5 isoforms were amplified using primers located in (i) exons 1A and 10 (PAX5ex1A-F2 5'-CCCTGTCCATTCCATCAAGTCC-3', PAX5ex10-R1 5'-TCACCCTCAATAGGTGCGCA TCAG-3'), (ii) exons 1A and 6 (PAX5ex1A-F2, PAX5ex6-R1 5'-CTCCCCGCATCTCTGCTT CC-3'), (iii) exons 5 and 10 (PAX5ex5-F2 5'-CGGCATCTGGGATGCAAC-3', PAX5ex10-R2 5'-
CGGTCTCATGGGCTCTCTGG-3'). RT-PCR reactions were carried out using Hot Start Taq (Qiagen, Vienna, Austria); cycling conditions: 95°C for 14 min; 35–40 cycles at 95°C for 30 sec, at 66°C for 30 sec, at 72°C for 2 min; 72°C for 7 min.

Results and Discussion

PAX5 and CDKN2A status in dic(9;20) leukemia

FISH analysis confirmed the presence of a dic(9;20) in all 7 cases (Figure 1A) and showed that one PAX5 allele was consistently deleted (Figure 1B). Sequencing of all coding exons of the retained allele revealed mutations in PAX5 in 4/7 (57%) of the cases (Figure 1D, Table 1). These mutations affected amongst PAX genes highly conserved residues11,12 and substituted either threonines with alanines (T75A, exon 3; T311A, exon 8) or proline with arginine (P80R, exon 3). The T75A and the P80R mutations affect the PAX5 paired DNA-binding domain, and the P80R mutation was reported to have a significant impact on protein structure.7 The novel T311A substitution is located in the transactivation domain of PAX5 in proximity to a P321 frame-shift mutation described earlier.7,11

Evaluation of the expression patterns of PAX5 isoforms showed that in one of the non-mutated cases (case 7) no full-length transcript was present (Figure 1Ei and 1Eiii; lanes 7). All other cases expressed full-length PAX5 and different splice variants in variable patterns and levels (Figure 1E). Direct sequencing of several transcripts revealed that these corresponded mainly to the known PAX5 splice variants PAX5∆8, PAX5∆9, PAX5∆7/8, and PAX5∆7/8/9.13,14 Of note, in the case that harbored the T311A exon 8 mutation, the transcript lacking exon 8 was predominantly expressed (Figure 1Ei and 1Eiii; lanes 4). Further, a novel PAX5 splice variant skipping exon 2 (Figure Eii; lane 5), which so far was only described for murine Pax515,16 was identified.

While multiple PAX5 isoforms were observed in B-cells of normal healthy donors14 also a possible association with childhood leukemia was reported.13 Further studies are required to
resolve whether the differential expression of PAX5 alternatively spliced transcripts is leukemia-associated or reflects physiological stages of B-cell development.

FISH showed either a heterozygous (n=3), a homozygous (n=2), or a mixed pattern of homo- and heterozygous deletion (n=1) of CDKN2A (Figure 1Ci; Table 1). However, one sample (case 3) had two CDKN2A copies (Figure 1Cii). CDKN2A was lost from the dic(9;20) and the second signal was located on a marker chromosome (data not shown) suggesting a more complex rearrangements that results in a slightly different dic(9;20), nevertheless indicating that loss of CDKN2A is not a consistent feature of cytogenetically defined dic(9;20) leukemia. In conclusion, we show that dic(9;20) childhood BCP-ALL or at least a subtype of this purely cytogenetically defined entity is consistently association with deletion of one PAX5 allele and frequent mutation of the second allele. A certain impairment of PAX5 function in combination with loss or mutation of other genes such as CDKN2A may play a crucial role in the development of this leukemia.

Acknowledgements

This work was supported by a grant of the Austrian National Bank (Project No. 12547) (to S.S.) and the St. Anna Kinderkrebsforschung. We thank Meinrad Busslinger (IMP, Vienna, Austria) for kindly providing the PAX5 cosmid probes and Maximilian-Otto Kauer for his help with statistical analysis.

Authorship

D.K., K.N., and M.K. performed the experiments; D.D. analyzed data; S.S. designed and supervised the project and wrote the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.
References


Table 1. Cytogenetics, FISH and mutation analysis of dic(9;20) leukemia.

<table>
<thead>
<tr>
<th>Case</th>
<th>Age (yrs)/Sex</th>
<th>Phenotype</th>
<th>Cytogenetics</th>
<th>PAX5 mutation</th>
<th>CDKN2A deletion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.6/F</td>
<td>pre-B ALL</td>
<td>47,XX, dic(9;20)(p11-13;q11),+10,+21[8]/48,idem,+X[6]/49,idem,+del(X)(q23),+8?[2]</td>
<td>WT</td>
<td>23% hetero</td>
</tr>
<tr>
<td>2</td>
<td>9.9/F</td>
<td>pro-B ALL</td>
<td>46,XX, dic(9;20)(p13;q11),+21[17]/46,idem,del(2)(p23?)[3]</td>
<td>p.P80R</td>
<td>40% hetero/50% homo</td>
</tr>
<tr>
<td>3</td>
<td>9.6/M</td>
<td>pre-B ALL</td>
<td>46-47,XY,+7[4],der(9)dic(9;20)(p11;q11),add(12)(q24),-20[7],+21,add(22)(q13)[cp12]</td>
<td>p.T311A</td>
<td>normal</td>
</tr>
<tr>
<td>4</td>
<td>4.0/F</td>
<td>pre-B ALL</td>
<td>45,XX,dic(9;20)(p11;q11),-20[15]/45,idem,del(12)(p12)[9]</td>
<td>WT</td>
<td>90% hetero</td>
</tr>
<tr>
<td>5</td>
<td>2.5/F</td>
<td>pre-B ALL</td>
<td>45,XX,dic(9;20)(p11-13;q11)[7]</td>
<td>p.T75A</td>
<td>42% hetero</td>
</tr>
<tr>
<td>6</td>
<td>4.4/F</td>
<td>cALL</td>
<td>45,XX,del(1)(q32),t(1;17)(q32;q25),add(2)(p25),-20[10]</td>
<td>WT</td>
<td>60% homo</td>
</tr>
</tbody>
</table>

cALL, common ALL
WT, wild-type
hetero, heterozygous
homo, homozygous
Figure 1.
Figure 1. Analysis of dic(9;20) leukemia.

(A-C) FISH analysis. Images were taken using an Axioplan fluorescent microscope (Zeiss, Göttingen, Germany) equipped with the appropriate filter sets for DAPI, FITC, and Cy3, fitted with a Plan-Neofluar 100x/1.3 oil immersion objective and a CCD camera (CH250, Photometrix LTD, Tucson, AZ) using the IPLabs software (Vysis, Inc., Stuttgart, Germany). (A) Hybridization with chromosome 9 (Cy3) and 20 (FITC) centromere-specific probes showing colocalization of the signals confirming the presence of a dic(9;20); (Ai) case 3 and (Aii) case 5. (B) Hybridization with PAX5 exon-specific cosmid clones cos-hPAX5-1 (FITC) and cos-hPAX5-3 (Cy3). (Bi) Metaphase of case 7 and (Bii) representative interphase nuclei of case 4 both displaying deletion of one PAX5 copy. (C) (Ci-i) Hybridization with CDKN2A-specific probe LSI p16 (9p21) (SpectrumOrange)/CEP9 (Spectrum Green) showing the mixed pattern of hetero- and homozygous loss of CDKN2A in case 2 (Ci) and the normal FISH pattern observed in case 3 (Cii). (D) Mutation analysis of leukemic blast (bottom chromatograms) and corresponding remission samples (top chromatograms). (Di-iii) Exon 3 (encoding the C-terminal portion of the PAX5 paired DNA-binding domain) mutations, (Di-ii) C>G P80R and (Diii) A>G T75A. (Div) A>G T311A exon 8 (encoding part of the PAX5 transactivation domain) mutation. (E) Expression of PAX5 splice variants. (Ei) Amplification of full-length PAX5 using primers located in exons 1A and 10. Lane 1, REH cell line; lane 2, case 1; lane 3, case 2; lane 4, case 3; lane 5, case 4; lane 7, case 6. (Eii) N-terminal isoforms amplified with primers located in exons 1A and 6. Lane 1, REH cell line; lane 2, case 1; lane 3, case 2; lane 4, case 3; lane 5, case 4; lane 6, case 5; lane 7, case 6; lane 8, case 7. (Eiii) C-terminal isoforms amplified with primers located in exons 5 and 10. Lane 1, REH cell line; lane 2, case 1; lane 3, case 2; lane 4, case 3; lane 5, case 4; lane 6, case 5; lane 7, case 6; lane 8, case 7; lane 9, negative control. M, molecular weight marker, 100 bp ladder (Promega, Mannheim, Germany).
CHAPTER 5

5. ADDITIONAL METHODS AND RESULTS

5.1. Establishment and validation of automated FISH screening

FISH screening of large numbers of samples requires highly standardized approaches including reliable and robust FISH protocols, and evaluation and data management procedures. For automated FISH analysis, FISH signals have to be highly specific and homogenous, and high hybridization efficiencies have to be achieved. Thus, both the FISH probe and the slide preparation were adjusted to the requirements of automated analysis.

_TempliPhi™ DNA amplification_

For preparation of large-scale FISH probes the TempliPhi™ 100 Amplification Kit (GE Healthcare, Vienna, Austria) was utilized, which produces microgram quantities of DNA from nanogram amounts of starting material. The TempliPhi™ DNA Amplification Kit uses rolling circle amplification (RCA) catalyzed by bacteriophage Phi29 DNA polymerase.

One µl (10ng/µl; quantified using the Hoefer DyNA Quant™ 200 fluorometer (GE Healthcare) of highly purified BAC DNA was transferred to 50µl of TempliPhi sample buffer, followed by denaturation at 93°C for 3min. Then 50µl of TempliPhi premix consisting of 50µl reaction buffer complemented with 2µl of enzyme mix were added to each denatured sample. Following incubation at 30°C for about 20hrs, the Phi29 DNA polymerase was heat-inactivated at 65°C for 10min and the amplification products were stored at 4°C or -20°C until further usage. Amplification of cosmid DNA was essentially performed as for BAC DNA, but only 1/10 of the reaction volume was required and the incubation time was reduced to 12-18hrs. However, in some instances the reaction volume was scaled-up to obtain higher amounts of DNA.

The amplified DNA was precipitated with 3 volumes of 95% ethanol (VWR International GmbH, Vienna, Austria) and 3M Na-Aacetate (VWR International GmbH), washed with 500µl of 70% ethanol (VWR International GmbH) and resuspended in 25-50µl of sterile water. After DNA amplification the quality of the BAC and cosmid DNA was checked on 2% agarose gels by gel electrophoreses and the DNA concentration was quantified by UV spectrophotometry.

In general, from approximately 10ng of DNA between 3-15µg BAC or cosmid DNA was obtained. The amount of input material was a critical factor for the successful amplification of high-quality DNA.
**Probe preparation**

In order to standardize probe preparation always 1µg of BAC or cosmid DNA was labeled by incorporation of digoxigenin-11-dUTP or biotin-16-dUTP (Roche Diagnostics, Vienna, Austria) nucleotides in a nick-translation reaction.

The labeled DNA was ethanol (VWR International GmbH) precipitated together with COT-1 and SSP DNA (Invitrogen GesmbH, Lofer, Austria), afterwards resuspended in hybridization solution [60% Formamide (VWR International GmbH), 1% Triton-X (VWR International GmbH), 2xSCC] and dissolved by incubation in a thermomixer at 45°C shaking at 500rpm for about one hour. These FISH probes were then denatured and pre-annealed to avoid cross-hybridization with repetitive sequences and were stored at -20°C, ready for use.

**Sample and slide preparation**

Methanol/acetic acid fixed cell suspensions of patient samples originally used for cytogenetic analysis were dropped on test slides to control the cell density. To allow for automated FISH scanning of a slide within a reasonable time-frame (10hrs - 14hrs) and to ensure appropriate hybridization efficiency, the cell density had to be relatively high but individual cells clearly separated. To achieve equal cell densities for all patient samples each sample was fixed with a freshly prepared appropriate amount of cold fixative [methanol (VWR International GmbH):acetic acid (VWR International GmbH) = 3:1]. Three µl of the fixed cell suspensions were dropped on frozen 3-well slides, each sample on one individual well, and incubated in Nonidet P40 (Sigma-Aldrich, Vienna, Austria) 0.4% / 2xSSC at 37°C for 1hr. After dehydration through an ascending ethanol series followed by 3min of pepsin digestion FISH was essentially performed as previously described (Konig et al, 2002).

**Automated fluorescence in situ hybridization - Metafer4-Metacyte (Metasystems)**

For high-throughput analysis and the objective evaluation of FISH patterns a standardized automated spot counting system including 3-dimensional distance measurements between the signals captured in different color channels was used. In particular, when split-apart FISH probes on the genomic level are separated by a larger distance (PAX5 clones RP11-220I1 and RP11-12P15 are separated by approximately 250kb), it is difficult to distinguish a separation of the probes from a mere slightly larger gap between signals (false-positives) dependent on the genomic distance by manual scoring (for details see 5.1.1.).

The Metafer4-Metacyte (MetaSystems, Altlussheim, Germany) slide scanning system is based on a motorized Zeiss Axioplan 2 Imaging fluorescence microscope (Zeiss, Göttingen, Germany) equipped with a motorized slide scanning stage and a Zeiss AxioCam MRm CCD camera. Both the microscope and the scanning stage are controlled by the special Metafer software package.
The system automatically scans up to 8 slides, takes images in different focal planes and enumerates the numbers of signals per cells (Fig. 9). Further analysis is conducted on the acquired digital images according to preset software parameters. These parameters have to be defined and saved in a so-called classifier. As the system allows to define a broad spectrum of different parameters, it is highly flexible and can be utilized for an almost unlimited number of assays. However, to obtain optimal results, the system has to be trained and adjusted to the individual requirements of each FISH assay.

Figure 9. Metafer4-Metacyte user interface. Screenshot of the Metafer4-Metacyte user interface, which consists of a menu bar for selection of the operation mode (top), an image area for live image display (upper left), the gallery (upper right) showing the acquired interphase nuclei, and command buttons with quick links to the most important functions (lower right). Further, current slide data including the scanned area and feature diagrams (lower part) as well as slide-specific data of one whole scan (bottom) are displayed.

The classifier parameters depend on the objective lens, the cell type and the fluorochromes used for the detection of the FISH probe. Parameters how the system captures the image fields in the different color channels are set in ’Capture’, e.g. an automatic exposure mode with a maximum integration time for ~2s and 2.9s for channels CY3 and FITC, respectively,
were defined, and the number and the distance of focus planes were set at 5 and 0.75 µm, respectively (Fig. 10). Further, parameters for an automatic exposure mask and for image processing operations were adjusted. In the classifier group ‘Cell Selection’ the parameters define how the system selects the interphase cells to be analyzed. These parameters strongly depend on the cell type and have to be optimized by training. For example, values for minimum and maximum nucleus area were 35 µm² and 500 µm², respectively, and for the maximum aspect ratio 1.7 was used. This setting defines the ratio of the nucleus diameters along the long and short principle axis to discriminate round objects from more elongated ones. If a cell is accepted by the cell selection procedure, a sub-image of the cell from the captured field image is created, which is the basis for further processing and measurements, and therefore parameters for image processing operations were sparingly defined.

![Figure 10. Metafer4-Metacyte Classifier Setup. Screenshot of the 'Capture' submenu of the classifier is shown as an example. The parameters, which have to be defined within this group are displayed.](image)

Some of the most important parameters in our assay were set in the ‘Features’, in which the definitions for the measurements and the spot counting applied to the cell images during scanning are adjusted.
The system measures the 3-dimensional distances between spots of different colors and defines the 1st and the 2nd smallest distance between them (Fig. 11). The parameter for 'Spot Fusion Count' was set at ≤10 Pixel meaning that 2 signals of different colors are counted as fusion signal if the distance is ≤10 Pixel (about 1.6µm). In cells lacking a value for the 2nd smallest distance, for example due to a deletion event, the value for the 2nd smallest distance was arbitrarily set at 100 Pixel (for details see 5.1.1.).

Figure 11. Distance Measurement. Schematic representation of a normal FISH pattern (left) showing a normal 2nd smallest distance and an abnormal FISH pattern (right) with a larger 2nd smallest distance.

The spot counting parameters also included an absolute spot area, which was set at 0.15µm², a minimum spot distance (1.4µm) and a relative minimum spot intensity (30%) within one color channel. The parameters were the same in both channels, CY3 and FITC. Additionally, a relative maximum spot area compared to the whole nucleus was set at 80/1000 units and the size of the area, which should be scanned during the search process was defined by the number of search fields around a certain coordinate (18x18 in the PAX5 classifier).
Several parameters for the 'Gallery', and finally the output form were defined. A special 'MetaCyte Report', which included for example information fields in terms of the analyzed sample and the hybridization procedure was created. Moreover, in this report histograms of the first and second smallest distances and display fields for a selection of the processed cells were displayed.

All parameters defined in the PAX5 'Metafer4 Metacyte Classifier' are listed in the Appendix.

**Evaluation**

All patient samples were automatically analyzed followed by manual evaluation of each individual cell captured and displayed in the cell gallery. Evaluation is facilitated by options such as cell sorting according to different parameters, changing of signal thresholds, and relocation of individual cells for visual inspection. This procedure was in particular required to reject false-positive cells and to correct erroneous spot counts.

**Figure 12. Metafer4-Metacyte Features Setup.** As an example, the feature list of channel 2 (FITC) of the submenu 'Features' is displayed with the spot counting parameters highlighted.
5.1.1. Analysis of control samples

The classifier for the PAX5 split-apart FISH assay was established using mononuclear cells isolated from peripheral blood of 3 healthy controls and blast cells of 2 different leukemic bone marrow samples that showed a normal karyotyp and a normal PAX5 status.

Representative mononuclear cells of a control hybridized with the PAX5 split-apart FISH assay are shown in Figure 13A. These examples illustrate the difficulty to unambiguously distinguish a fusion signal from a split signal pattern by visual observation. In Figure 13B typical histograms of the 1st and 2nd smallest distances found in normal controls are shown. In 8 different hybridization experiments of 3 control samples second smallest distances between 0 and 28 Pixel were observed with a mean of 4.8 Pixel and a standard deviation of 0.56. Moreover, in at least 96% of cells the 2nd smallest distances were ≤10 Pixel, thus, the spot fusion count parameter was set accordingly in the PAX5 classifier.

![Figure 13. Distance measurements Metafer4-Metacyte (Metasystems). (A) Representative interphase nuclei of a normal control hybridized with PAX5-flanking BAC clones RP11-220I1 (5'-end-specific; red) and RP11-12P15 (3'-end-specific; green) showing a normal FISH pattern (left lower corner of each picture: total number of red and green signals; right lower corner: number of fusion signals) and different 1st and 2nd smallest distances (shown in the upper right corner of each picture). Interphase pictures were taken from the Metafer4-Metacyte cell gallery. (B) Representative histograms that illustrate the distribution of the 1st and 2nd smallest distances in a normal control.](image-url)
The data obtained by the automated FISH analysis were summarized in a "MetaCyte Report" and as an example one of a normal control is shown in Figure 14.

Figure 14. MetaCyte Report of a normal control. Summary of the results obtained by hybridization with PAX5-flanking BAC clones RP11-22011 (5'-end-specific; red) and RP11-12P15 (3'-end-specific; green) to a control sample. The different classes of FISH patterns and the number of enumerated cells are shown: 305 interphase nuclei were captured, 29 of them were rejected upon evaluation, and 276 were finally analyzed. R, red; G, green; F, fusion signal.
5.1.2. Validation of the FISH assay in PAX5-rearranged samples

As positive controls the KIS-1 cell line and a patient previously identified as PAX5-rearranged were used to validate the FISH assay.

The KIS-1 cell line harbors a t(9;14)(p13;q32)/PAX5-IGH@ translocation and the breakpoint is located upstream of PAX5 exon 1A. Cytogenetic analysis revealed further aberrations and a very complex karyotype with additional marker chromosomes, which is in concordance with published data (George et al., 2005). Representative interphase nuclei of the KIS-1 cell line showing a separation of the PAX5 FISH probes and 3-5 additional PAX5 3'-end-specific signals are depicted in Figure 15A. Examples of interphase nuclei of the PAX5-rearranged patient that also display a separation of the PAX5 FISH probes are shown in Figure 16A. The histograms of the positive controls illustrate the broader distribution of the 2nd smallest distances, which correspond to the separated PAX5 FISH probes (Fig 15B and 16B).

Figure 15. KIS-1 cell line as positive control for the Metafer4-Metacyte automated FISH analysis and distance measurements. (A) Representative interphase nuclei of the KIS-1 cell line hybridized with PAX5-flanking BAC clones RP11-22011 (5'-end-specific; red) and RP11-12P15 (3'-end-specific; green); (left lower corner of each picture: total number of red and green signals; right lower corner: number of fusion signals) with different 1st and 2nd smallest distances (shown in the upper right corner of each picture). Images of interphase nuclei were taken from the Metafer4-Metacyte cell gallery. (B) Histograms that show the distribution of the distances between the FISH signals.
5.1.3. Novel PAX5 positive cases - Metafer 4 Metacyte analysis

Ten patients showed a FISH pattern suggestive for a PAX5 rearrangement. As an example, the Metafer4-Metacyte Report of a representative case is shown in Figure 17. In this case, 350 interphase nuclei were captured, 92 of these were rejected, thus, 258 were included in the statistical analysis, and a separation of the PAX5 FISH probes was found in 59.7% of cells. The corresponding distribution of the 2nd smallest distances between the FISH signals is also illustrated in the 'MetaCyte Report'.
Figure 17. Representative MetaCyte Report. The results of the automated FISH analysis using PAX5-flanking BAC clones RP11-2201I (5'-end-specific; red) and RP11-12P15 (3'-end-specific; green) of a PAX5-rearranged case are shown. The significant aberrant FISH pattern is boxed in red. The histograms of the 1st (first diagram) and 2nd smallest (second diagram) distances are depicted at the bottom. R, red; G, green; F, fusion signal.
Distances between FISH signals in PAX5 positive patients

In all positive patients that showed a separation of the PAX5 FISH probes the distributions of 2\textsuperscript{nd} smallest distances of the FISH signals were analyzed in detail (Table 1). All cells, in which the 2\textsuperscript{nd} smallest distance was set to 100, e.g. cells with a deletion event, were excluded from this analysis. The percentage of cells that showed 2\textsuperscript{nd} smallest distances >10 Pixel representing cells with an aberrant pattern ranged from about 24-71%. The differences between the 2\textsuperscript{nd} smallest distances of all positive patients as compared to those of normal controls were highly significant (p-values $10^{-10}$ - $10^{-58}$ by Student's T-test).

Table 1. FISH signal distances in PAX5 positive cases

<table>
<thead>
<tr>
<th>Case No*</th>
<th>2\textsuperscript{nd} smallest distance (mean)</th>
<th>2\textsuperscript{nd} smallest distance (standard deviation)</th>
<th>Distance values (min. - max.)</th>
<th>% of cells with distance &gt;10 Px</th>
<th>No of cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>8,33</td>
<td>4,14</td>
<td>1-23</td>
<td>23,60</td>
<td>212</td>
</tr>
<tr>
<td>3</td>
<td>10,46</td>
<td>7,93</td>
<td>1-41</td>
<td>36,40</td>
<td>228</td>
</tr>
<tr>
<td>4</td>
<td>8,11</td>
<td>6,86</td>
<td>1-46</td>
<td>25,10</td>
<td>219</td>
</tr>
<tr>
<td>5</td>
<td>14,17</td>
<td>6,57</td>
<td>1-36</td>
<td>70,50</td>
<td>210</td>
</tr>
<tr>
<td>6</td>
<td>12,98</td>
<td>8,44</td>
<td>2-50</td>
<td>53,80</td>
<td>182</td>
</tr>
<tr>
<td>7</td>
<td>13,10</td>
<td>6,45</td>
<td>1-33</td>
<td>66,30</td>
<td>270</td>
</tr>
</tbody>
</table>

*For case descriptions see Nebral et al, in press.
5.2. Detailed analysis of selected cases

PAX5 rearrangements were observed either in cases with normal or complex karyotypes. In order to further characterize these aberrations, additional PAX5 gene-specific probes and clones located adjacent to PAX5 were hybridized in various combinations to metaphase and/or interphase nuclei (Fig. 18). Several whole chromosome painting probes, and sets of probes hybridizing to the potentially affected chromosome regions in the individual cases were also applied.

Further, particularly, in cases with focal/partial PAX5 deletions FISH clones located in proximity to the PAX5 locus were used to elucidate the extent of these deletions (Fig. 18). The additional FISH clones are listed in Table 1 in the Appendix.

Figure 18. FISH clones located at 9p13. Adapted screenshot from the UCSC Genome Browser on Human March 2006 Assembly (www.genome.ucsc.edu) with the utilized FISH clones depicted in red. The PAX5 gene is transcribed from centromere to telomere (5' - 3').
5.2.1. PAX5 rearranged cases

Selected PAX5 positive cases were analyzed in further detail. The case numbers correspond to those in Chapter 3: "Incidence and diversity of PAX5 fusion genes in childhood acute lymphoblastic leukemia", Nebral et al, page 39.

Case 5

In this case, cytogenetic analysis showed a 46,XY,del(7)(q22q33)?,del(9)(q22?),del(12)(p11)[8] karyotype suggesting complex rearrangements with involvement of chromosomes 7, 9, and 12, but 9p13 was not affected. Thus, conventional chromosome banding did not provide any evidence for involvement of PAX5. However, FISH analysis indicated a PAX5 rearrangement and on the molecular level PAX5 exon 5 was fused to POM121 exon 5 with insertion of genomic material from chromosome 12 at the breakpoint. Subsequent hybridization of PAX5-aberrant metaphases with whole chromosome painting probes for chromosomes 7 and 9 confirmed involvement of these chromosomes in a complex at least 3-way translocation (Fig. 19B). Moreover, this analysis revealed that the 3'-end of PAX5 was located on a der(7;9) chromosome, whereas the 5'-end was translocated to a der(7), most probably harboring the PAX5-POM121 fusion (Fig. 19).

Figure 19. FISH analysis of case 5. (A) Metaphase hybridized with the PAX5 5'-end encompassing clone RP11-456P6 (green) and RP11-84P7 (red), located telomeric (3') of PAX5, showing a split signal pattern. (B) The same metaphase was analyzed with whole chromosome painting probes for chromosomes 9 (red) and 7 (green). Arrows indicate the normal chromosomes 9 (red) and 7 (green), the der(9) (white) and der(7) (short yellow), and the der(7;9) (yellow).
Case 6
In case 6, about 46% of interphase nuclei showed a separation of the PAX5-flanking BAC clones and rearrangement of PAX5 was confirmed by hybridization with the PAX5-specific cosmids. Subsequent fusion gene-specific RT-PCR and/or FISH analysis for all known PAX5 partner genes did not detect any specific chimeric transcript and no abnormal FISH pattern (Fig. 20) was observed. All Oligonucleotide primer sequences and the FISH clones used for these analyses are listed in Table 2 and Table 1, respectively, in the Appendix.

Figure 20. Examples of the fusion gene-specific FISH analysis of case 6. (A) Interphase nuclei hybridized with ELN 5'-end RP11-148M21 (red) and ELN 3'-end RP11-349P21 (green) specific clones displaying a normal FISH signal pattern. (B) Analysis of interphase nuclei with PAX5-spanning clones RP11-465P6 and RP11-431F4 (green) in combination with AUTS2-spanning clones RP11-243F5 and RP11-88H22 (red) which showed a PAX5 split signal and two normal signals for AUTS2.

Hybridization of metaphases with whole chromosome painting probes for chromosomes 9 and 14 showed that the PAX5 5'-end was located on a der(9) chromosome whereas the 3'-end was translocated to 14q32 (Fig. 21A). Participation of IGH@ and BCL11B located at 14q32 was excluded by FISH analysis. Yet, 24-color FISH suggested involvement of chromosomes 3, 9, 11, and 12 in complex rearrangements and further FISH analysis revealed translocation of chromosome 9 material to 3q (Fig. 21Aii) and 3q material to 9q (Fig. 21B and data not shown), as well as insertion of 11p material into both 9p and 14q (Fig. 21C). Based on these data we suspected that a novel PAX5 fusion partner was located at 11p and in order to narrow down the potential breakpoint between LMO1 and WT1, 20 BAC clones (see Appendix, Table 1 and Fig. 22), which encompassed 11p13-15 were hybridized to metaphases in combination with PAX5-specific probes. However, although all clones were properly either observed on the der(9) or the der(14) chromosomes (examples are shown in Fig 23) neither of them showed a co-localization with PAX5 in interphase nuclei. Finally, one single BAC clone, RP11-52H5, located at the 5'-end of the BDNF gene, was deleted. This gene consists of only 2 exons but RT-PCR experiments using primers PAX5ex5-F1 and BDNFex2-R1 or BDNFex2-R2 did not detect any specific fusion transcripts.
Figure 21. FISH analysis of case 6. (A) Metaphase hybridized with the PAX5 locus-specific BAC clones RP11-220I1 (red) and RP11-12P15 (green) (Ai) and with whole chromosome painting probes for chromosome 9 (red) and 14 (green) (Aii). Arrows indicate the normal chromosome 9 (yellow), the der(9) (red) and the der(14) (green). Red arrowhead points to the der(3). (Bi) Metaphase that shows a PAX5 split signal pattern using PAX5-spanning clones RP11-465P6 and RP11-431F4 (green) and two normal TOP2B-specific (3p24) signals with clone RP11-659P16 (red). Arrows denote the normal chromosome 9 (yellow), the der(9) (red) and the der(14) (green). (Bii) Same metaphase as in (Bi) displaying a translocation of the 3q-subtelomere region using clone 196F4 (red) and normal signals for 3p with 3p-arm-specific painting probe (green). Arrows point to the der(9) (red) and the der(3) (green) chromosomes and the normal chromosome 3 (yellow). (Ci) Hybridization of a metaphase with PAX5-spanning clones RP11-465P6 and RP11-431F4 (red) and the WT1-specific (11p13) clone RP11-74J1 (green) shows a disruption of PAX5. Arrow denotes the der(14). (Cii) Metaphase displaying a split signal with PAX5-spanning clones RP11-465P6 and RP11-431F4 (red) and signals for LMO1 (11p15) using RP11-379P15 clone (green). Arrow indicates the der(9).
Figure 22. Predicted breakpoint region at chromosome 11p15.4-p13. Adapted screenshot from the UCSC Genome Browser on Human March 2006 Assembly (www.genome.ucsc.edu) displaying all genes located within this region. Genes detected by locus-specific FISH clones (see Appendix, Table 1) are encircled and the localization of the signals on the respective derivative chromosomes is indicated by the arrows.
Figure 23. Analysis of case 6 with FISH probes located at 11p. PAX5-spanning clones RP11-465P6 and RP11-431F4 (green) were applied to metaphases in combination with 11p-specific clones. In all metaphases shown the arrows indicate the der(9) (yellow), the normal chromosomes 9 (red) and 11 (white), and the yellow arrowhead the der(14). (Ai) Hybridization with PAX5-spanning clones (green) in combination with MRVI-specific clone RP11-58H20 (red). (Aii) NELL1-specific clones RP11-3E17 and RP11-116O9 (red) in combination with E2F8-specific clone RP11-428C19 (green). All these 11p-specific clones were localized on the der(9). (B) The ELP4 5'-end specific clone RP5-1137O17 (red) was observed on the der(14) (Bii), whereas the LGR4 3'-end-specific clone RP11-426P16 (red) was located on the der(9) (Bii). (C) Metaphase hybridized with LUZP2 5'-end RP11-372B5 (red) (Ci) and BDNF 5'-end-specific clone RP11-52H5 (red) (Cii) shows the localization of LUZP2 5'-end on der(9) and deletion of the BDNF 5'-end encompassing clone.
As the FISH strategy was unsuccessful to delineate the PAX5 partner gene, rapid amplification of cDNA ends (RACE) was performed, but also all RACE experiments using various primer combinations and PCR conditions, which should permit amplification of any fusion product resulting from every possible breakpoint within PAX5 (Fig. 24 and Table 2 of Appendix) failed to identify the fusion partner.

**Figure 24. PAX5 RACE assay.** A schematic structure of the PAX5 gene is depicted and all known breakpoints of PAX5 fusions are indicated (arrows). Localizations of the different gene-specific RACE primers used in the first round of amplification are shown (arrowheads).

**Case 7**

Case 7 showed a PAX5-split signal in about 62% of cells with additional 1-2 PAX5 5'-end signals. Cytogenetics revealed a particularly complex karyotype: 46,XY,add(1)(q44)[3], der(1)t(1;?)(p31;?)add(1)(q44)[3],-5,-8,del(9)(p13),del(11)(q23),+2mar[6cp]/46,XY[14]. Analysis of aberrant metaphases by FISH showed that the genes ABL1 (9q34) (Fig. 25A) and RCSD1 (1q24) (Fig. 25C) were located on a derivative chromosome, which carried the PAX5 5'-end signal strongly suggesting the presence of a dic(1;9)(p?;p13) (Fig. 25). Moreover, additional derivative chromosomes with PAX5 signals were observed but not further analyzed. In concordance with the data obtained by FISH, RACE identified HIPK1 located at 1p13 as novel PAX5 fusion partner. Owing to the opposite transcriptional orientation of PAX5 (centromere-telomere) and HIPK1 (telomere-centromere) the formation of a dic(1;9)(p13;p13) as seen in this case or a complex rearrangement is required to generate a functional fusion gene.
Figure 25. FISH analysis of case 7. (A) Metaphase hybridized with probes specific for the *BCR* (green) and *ABL1* (red) genes located at 22q11 and 9q34, respectively. (B) Same metaphase as in A hybridized with *PAX5* flanking clones RP11-220I1 (red) and RP11-12P15 (green) clones. Arrows indicate the normal chromosome 9 (black) and the der(9) (red). (C) Metaphase hybridized with *RCSD1* 5'-end (red) and 3'-end-specific (green) probes located at 1q24. Arrows point to the normal chromosome 1 (black) and the dic(1;9). (D) Same metaphase as in C hybridized with *PAX5* 5'-end RP11-220I1 (red) and *PAX5* 3'-end RP11-12P15 (green) specific clones. Arrows denote the normal chromosome 9 (black) and the dic(1;9) (red).
5.2.2. PAX5 deletions

Recently, it was shown that apart from PAX5 rearrangements, PAX5 is also a target of other somatic mutations in particular monoallelic deletions, which in a significant proportion of the cases are focal and just affecting a few exons (Mullighan et al, 2007a). Although our PAX5 FISH assay was not specifically designed for the detection of PAX5 deletions per se, nevertheless a high percentage of the cases (about 10%) displayed monoallelic loss of the gene. Moreover, in approximately 4% of the non-rearranged cases focal deletions of the PAX5 locus affecting either the 5’ or the 3’ region were detected. These cases were further analyzed using a panel of FISH probes, whose signal patterns reflected the size and the extent of the deletions. Examples of the most frequent patterns are described in the following.
Pattern No. 1

In some cases hybridization with the PAX5-flanking BAC clones revealed a deletion of the 3’-clone RP11-12P15 in a significant proportion of the interphase nuclei (Fig. 26Ai), whereas after hybridization with the PAX5-specific cosmids a deletion of one entire copy of PAX5 was observed (Fig. 26Aii). Further FISH analysis using PAX5-spanning probes RP11-465P6 and RP11-431F in combination with RP11-84P7 showed that the deletion encompassed also the region telomeric of PAX5 (Fig. 26B). Moreover, a small second signal of the PAX5-spanning clones was observed, probably reflecting partial retention of the 5’-end clone RP11-465P6. This clone extends beyond the 5’-end of PAX5 suggesting that the deletion starts in PAX5 exon 1 or just upstream of exon 1 encompassing at least clone RP11-84P7 (Fig. 26).

Figure 26. PAX5 deletion pattern no. 1. (Ai) Metaphase and interphase nuclei hybridized with FISH probes RP11-220I1 (red) and RP11-12P15 (green) displaying deletion of one signal of RP11-12P15. (Aii) Hybridization of interphase nuclei with cos-hPAX5-1 (green) and cos-hPAX5-3 (red) that shows a deletion of one entire copy of PAX5. (B) Interphase and metaphase analysis with PAX5-spanning clones RP11-465P6 and RP11-431F4 (green) in combination with RP11-84P7 (red) demonstrate deletion of the region telomeric of PAX5. Arrows indicate the normal chromosome 9 (yellow) and the chromosome 9 with the interstitial deletion (red) showing a smaller signal resulting from partial retention of clone RP11-465P6. The schematic maps show the utilized FISH clones.
Pattern No. 2

Some cases showed a deletion of the PAX5-flanking clone RP11-220I1 (Fig. 27A), whereas hybridization of the PAX5-specific cosmid probes detected two normal copies of PAX5 (Fig. 27B). Moreover, FISH analysis using the clone RP11-397D12 together with PAX5-spanning probes demonstrated two normal signals (Fig. 27D) suggesting the presence of a focal deletion encompassing only the region, which is spanned by RP11-220I1 (Fig. 27). Of note, this pattern was also observed in ETV6-RUNX1 positive cases.

**Figure 27. PAX5 deletion pattern no. 2.** Interphase nuclei hybridized with (A) PAX5-flanking clones RP11-220I1 (red) and RP11-12P15 (green) displaying a deletion of RP11-220I1 and (B) PAX5-specific probes cos-hPAX-1 (green) and cos-hPAX5-3 (red) showing a normal pattern. (D) Hybridization of interphase nuclei with PAX5 probes RP11-465P6 and RP11-431F4 (green) in combination with RP11-397D12 (red), which demonstrates normal signals. Schematic representation of the localization of the utilized FISH clones are shown in the lower panels.
Pattern No. 3

A third different pattern was also observed in both ETV6-RUNX1 positive and negative cases. FISH analysis with PAX5 BAC clones RP11-220I1 and RP11-12P12 showed a deletion of the 5’-end flanking clone RP11-220I1 (Fig. 28Ai and Aii) and also a deletion of the PAX5 5’-specific probe cos-hPAX5-1 (Fig. 28Aiii). In the example shown, hybridization was done on ETV6-RUNX1 positive metaphases to ensure analysis of aberrant cells and to verify that the PAX5 deletion concurs in the same leukemic clone (Fig. 28A and B).

FISH was also performed using the PAX5-spanning clones in combination with RP11-397D12, which showed either complete loss of the clone or a smaller but clearly visible second signal suggesting a partial deletion of the clone (Fig. 28C). This pattern may on the one hand depend on the hybridization efficiency or indicate the presence of two different subclones. Of note, the remaining 3’-part of PAX5 could not be detected with the PAX5-spanning clones, which possible results from a weaker hybridization efficiency of clone RP11-431F4. Together, this type of deletion encompasses only the 5’-end of PAX5 and extends at least beyond the ZCCHC7 gene (Fig. 28C).
Figure 28. FISH analysis of PAX5 deletion pattern 3. (Ai-ii) Hybridization of interphase nuclei with PAX5 RP11-220I1 (red) and RP11-12P15 (green) showing a deletion of the 5’-end flanking clone RP11-220I1. (Aiii) Metaphase hybridized with PAX5 cos-hPAX5-1 (green) and cos-hPAX5-3 (red) that display a PAX5 5’-end deletion. (B) The LSI TEL (green)-AML1 (red) ES Dual Color Translocation probe (Abbott Molecular) was applied on the same metaphase as in (A). Arrows point to the der(12), which harbors the ETV6-RUNX1 (TEL-AML1) fusion (yellow), and to the normal chromosomes 12 (green) and 21 (red). (C) Metaphase and interphase nuclei analysis with the PAX5-spanning probes RP11-465P6 and RP11-431F4 (green) in combination with RP11-397D12 (red) that revealed either no, a partial or a complete deletion of clone RP11-397D12. (A and C) Arrows indicate the normal (yellow) and the derivative chromosomes 9 (red). Schematic maps depict the applied FISH clones.
Pattern No. 4

The most frequently observed deletion pattern was a complete deletion of *PAX5* that occurred in approximately 10% of the analyzed patients and in nearly all cases was associated with a cytogenetically detectable deletion of 9p. These cases were only analyzed with the *PAX5*-flanking clones RP11-220I1 and RP11-12P15.

In Figure 29 an example is shown that displays a deletion of one entire copy of *PAX5*.

![Figure 29. FISH analysis of PAX5 deletion pattern 4. Interphase and metaphase nuclei hybridized with PAX5-flanking probes RP11-220I1 (red) and RP11-12P15 (green), which display a deletion of one copy of PAX5.](image)
A unique case

In one single patient cytogenetic analysis showed del(9)(p22) suggesting a deletion telomeric of PAX5. However, 54% of the interphase nuclei displayed only one signal for the PAX5 3'-flanking BAC clone and no other signals were observed (Fig. 30A). Hybridization with the PAX5 cosmid probes demonstrated deletion of one entire copy of PAX5 and one apparently intact allele (Fig. 30B). Therefore, FISH analysis indicated loss of one PAX5 allele accompanied by a deletion of the region upstream of PAX5 encompassing ZCCHC7 on the second chromosome.

Figure 30. A special deletion pattern. (Ai-ii) Examples of interphase nuclei hybridized with RP11-12P15 (green) and RP11-22011 (red) that show only one PAX5 3'-end RP11-12P15 signal. (B) Metaphase analyzed with PAX5 cos-hPAX5-1 (green) and cos-hPAX5-3 (red) displaying a deletion of one PAX5 allele. Arrow points to the “normal” chromosome 9. The schematic map at the bottom shows the localization of the applied FISH clones.
CHAPTER 6

6. DISCUSSION

The discovery of the PAX5 fusion genes

The aim of the presented research project was to determine the overall incidence of PAX5 rearrangements and the whole spectrum of fusion partners in childhood ALL in a population-based survey. At the time this study was initiated only the PAX5-IGH@ activating translocation (Busslinger et al, 1996; Iida et al, 1996) and one single PAX5 fusion gene, namely PAX5-ETV6 were known (Cazzaniga et al, 2001; Strehl et al, 2003).

During the time this study was still ongoing, genome-wide analysis using high-resolution SNP arrays revealed chromosomal imbalances resulting in PAX5 fusions in 2.1% (Mullighan et al, 2007a) and 2% (Kawamata et al, 2008) of childhood B-cell precursor ALL (BCP-ALL). However, our study still represents the first population-based study using a FISH assay, which allows for the detection of all possible PAX5 rearrangements irrespective of chromosomal imbalances. Our data corroborate the finding that PAX5 fusions occur exclusively in BCP-ALL and at an incidence of about 2.5%.

Intriguingly, PAX5 can fuse to a multitude of different partner genes encoding transcription factors, structural proteins, a tyrosine kinase, and some proteins of still unknown function. All breakpoints within PAX5 occur in intron 5 or downstream (except for the PAX5-ETV6 rearrangement) (Fig. 31) fusing at least the PAX5 DNA-binding paired domain to the C-terminal region or even the entire protein encoded by the partner gene (Bousquet et al, 2007; Mullighan et al, 2007a; Nebral et al, 2007; Kawamata et al, 2008; Nebral et al, in press). Therefore, all PAX5 chimeric proteins are predicted to retain the ability to bind to PAX5 target genes, but would no longer provide normal transcriptional regulatory functions.
The relevance of PAX5 deletions

The second major finding of the SNP array studies was the frequent deletion of regulators of B-cell development, amongst which PAX5 was affected in approximately 30% of BCP-ALL (Kuiper et al., 2007; Mullighan et al., 2007a). Intriguingly, PAX5 deletions were also detected in ETV6-RUNX1 positive BCP-ALL (Mullighan et al., 2007a; Parker et al., 2008) and may, thus, belong to the cooperating mutations required to complete leukemogenesis. Yet, PAX5 deletions occur only in about 30% of the cases and to some extent coincide with other genetic lesions such as deletions of the second ETV6 allele or CDKN2A (S. Strehl, K. Nebral, M. König et al., unpublished observation). Thus, whether PAX5 deletions indeed belong to those secondary genetic lesions required to convert an ETV6-RUNX1 positive preleukemic clone to overt leukemia remains to be proven.

Monoallelic loss of PAX5 is proposed to lead to haploinsufficiency of the BSAP protein contributing to the differentiation arrest seen in BCP-ALL (Kuiper et al., 2007; Mullighan et al., 2007a). This assumption is based on the initial observation that only a single Pax5 allele is transcribed during the earliest phase of B-cell commitment and that B-cell differentiation relies on the switch to biallelic expression (Nutt et al., 1999b). Consequently, loss of a wild-type allele would eliminate the ability to switch on biallelic transcription. However, recent investigations clearly show biallelic expression of Pax5 at all stages of B-cell development (Fuxa & Busslinger, 2007). The hypothesis that haploinsufficiency of PAX5 may contribute to a differentiation block (Kuiper et al., 2007; Mullighan et al., 2007a) is further challenged by the fact that in heterozygous Pax5<sup>+/−</sup> mice B-cell development is normal (Urbanek et al., 1994; Nutt et al., 1999b). Moreover, inactivation of one Pax5 allele in mature B-cells in
heterozygous Cd19-cre Pax5<sup>fl/+</sup> mice in the absence of other oncogenic lesions is not sufficient to induce tumor development. On the other hand, complete loss of Pax5 in late B-cells results in the development of aggressive progenitor cell lymphoma (Cobaleda et al, 2007a). Although some of the deletions in BCP-ALL are confined to PAX5, in several cases even to a few exons (focal deletions), many of them are broader and encompass a number of genes, whose concomitant deletion/haploinsufficiency may as well promote leukemogenesis.

**Hypomorphic alleles and splice variants**

Alternative splicing of pre-mRNA is a fundamental process that increases proteomic diversity and contributes to genetic variability. Increasing evidence substantiates that splicing defects (e.g. caused by inherited or somatic mutations in regulatory elements) not only account for inherited diseases susceptibility but also increase proteome complexity in cancer cells, and that alternative splicing may well be one of the basic causes for cancerogenesis (Kalnina et al, 2005; Venables, 2006; Skotheim & Nees, 2007). In this regard, also for several members of the PAX gene family a cancer-specific expression of certain isoforms has been observed (reviewed by (Barr, 1997; Robson et al, 2006; Lang et al, 2007; Wang et al, 2008)).

Several human PAX5 isoforms have been described that result either from transcription of two alternative promoters leading to expression of exon 1A or 1B containing transcripts (Busslinger et al, 1996), or from alternative splicing of exons that encode the C-terminal transactivation and inhibitory domains (Robichaud et al, 2004; Sekine et al, 2007). Expression of different PAX5 isoforms was observed in B-cells of normal healthy donors (Robichaud et al, 2004) but also a possible association with childhood leukemia was reported (Sadakane et al, 2007). In this respect, we have identified a novel human PAX5 isoform that skips exon 2, and which may specifically occur in BCP-ALL (Krehan et al, in preparation).

Therefore, in-depth studies are required to elucidate whether the differential expression of PAX5 alternatively spliced transcripts is of any pathological relevance or reflects physiological stages of B-cell development. In particular, imbalances in the expression of PAX5 isoforms may result in modulatory effects on downstream genes.

Deletions within the PAX5 gene, which are confined to a subset of internal exons result in the expression of so-called hypomorphic alleles, which resemble potential splice variants. However, such splice variants have so far not been detected in normal B-lymphocytes suggesting that they are specifically generated by the intragenic deletion events. Recently, similar focal deletions of IKZF1 (Mullighan et al, 2008) and ERG (Mullighan et al, 2007b) in BCR-ABL1 positive ALL and a novel BCP-ALL subtype with a unique gene expression profile, respectively, were identified. Thus, focal mono-allelic intragenetic deletions may represent a distinct mechanism to generate tumor-specific isoforms.
The potential oncogenic function of PAX5 mutants

Several lines of evidence suggest that the various PAX5 aberrations found in BCP-ALL result in an impairment of PAX5 activity rather than a complete loss of function.

One of the defining features of BCP-ALL blast cells is the expression of the cell surface antigens CD19 and/or CD79A and/or CD22, which were readily expressed in both the PAX5-rearranged cases and those with monoallelic PAX5 deletions. Even concomitant mutations of the second allele did not abolish expression of the CD19 and CD79A proteins. In leukemic blast cells, also on the transcriptional level no correlation between PAX5 mutation status and CD19 and CD79A expression was observed (Mullighan et al, 2007a; Kawamata et al, 2008). While the lack of CD19 responsiveness, whose expression is strictly PAX5-dependent, to PAX5 mutation remains elusive, transcription of CD79A is also initiated by EBF1 and its expression can be activated independently of PAX5 (Hagman & Lukin, 2005; Pongubala et al, 2008). Although gene expression profiling identified a differential gene expression signature between PAX5-deleted and PAX5 wild-type ETV6-RUNX1 BCP-ALL (Mullighan et al, 2007a), the differentially expressed genes included only a small subset of those regulated by PAX5.

**In vitro**, using a luc-CD19 reporter gene assay PAX5 mutants showed reduced transcriptional activity as compared to wild-type PAX5 (Bousquet et al, 2007; Mullighan et al, 2007a; Kawamata et al, 2008), and mutations that affect the PAX5 DNA-binding paired domain had a lower binding capacity to CD19 promoter sequences (Mullighan et al, 2007a). In contrast to PAX5 DNA-binding and internal deletion mutants, which have only a weak competitive activity (Mullighan et al, 2007a), PAX5 fusion genes, appear to act in a dominant-negative manner over wild-type PAX5 (Bousquet et al, 2007; Mullighan et al, 2007a; Fazio et al, 2008; Kawamata et al, 2008). Thus, PAX5 chimera are considered to function as aberrant transcription factors that antagonize PAX5 activity provided by the second, wild-type allele (Cobaleda et al, 2007b).

Recent data obtained by *in vitro* transfection experiments with the PAX5-C20orf112 fusion also suggest that the antagonizing function of the chimeric protein is confined to a subset of PAX5 target genes (Kawamata et al, 2008). This observation raises the question whether the various PAX5 chimera deregulate a common, or dependent on the moieties provided by the partner protein, distinct sets of target genes.

While *in vitro* studies are certainly required to determine the functional consequences of PAX5 mutants, only validation of the data in primary leukemic blast cells will prove their true impact. In this respect, transfection of the post-germinal-center B-cell line DG75 with PAX5-ELN resulted in downregulation of the PAX5 target genes BLNK, LEF1, and CD79A, but in leukemic pre-B cells these genes were not affected (Bousquet et al, 2007). These data suggest a cellular context-dependent impact of the PAX5 fusion proteins, and that the results
obtained by *in vitro* transfection experiments - in particular using lymphoma cells, which are derived from later stages of B-cell development - may not faithfully reflect their effect in the context of B-cell precursor leukemia.

The B-cell developmental stage-dependent function of Pax5 provides a possible explanation for this observation. While during the early stages of B-cell development Pax5 expression is required for B-cell development, its downregulation is pivotal for terminal plasma cell differentiation. Thus, reliant on the developmental stage of the B-cell both loss-of-function and gain-of-function mutation of PAX5 may disrupt B-cell homeostasis.

Indeed, in B-cell lymphoma derived from later stages of B-cell development ectopic expression of *PAX5* suggests a gain-of-function mutation (Thomas-Tikhonenko & Cozma, 2008). In these tumors, which mostly arise from germinal-center B-cells with a functional BCR, enforced *PAX5* expression results in ligand-independent BCR signaling (Cozma *et al.*, 2007). Thus, increased expression of *PAX5* in post-germinal center B-cells promotes tumor growth by perturbing the PAX5-dependent B-cell gene expression program or a block in terminal differentiation into plasma cells by failed PAX5 repression (Shaffer *et al.*, 2002; Cobaleda *et al.*, 2007b). In this cellular context, *PAX5* most likely exerts its oncogenic effects via the functional BCR (Cozma *et al.*, 2007; Thomas-Tikhonenko & Cozma, 2008). Loss of Pax5 in the context of strong BCR signaling results in forward differentiation of mature B-cells to plasma cells, whereas Pax5 inactivation in the absence of BCR signaling initiates the reversal of differentiation to uncommitted progenitors (Cobaleda *et al.*, 2007a).

In contrast, in BCP-ALL, which is derived from immature pro- or pre-B-cells PAX5 mutants are without exception loss-of-function mutations, which impair but not abolish PAX5 function. However, the impact of the PAX5 mutants on the finely tuned B-cell transcription network has yet to be elucidated *in vitro* and more importantly, *in vivo*.

**Concluding remarks**

Although the fundamental role of PAX5 for B-cell development was recognized almost two decades ago, until very recently its possible involvement in B-cell malignancy has been essentially neglected. Only now PAX5 also emerges as a major player in leukemogenesis and it will probably take at least another two decades before the potential oncogenic role of PAX5 mutation will be finally elucidated. Dissecting the impact of these peculiar mutants on the complex B-cell development regulatory transcriptional network with its multiple combinatorial inputs and feedback loops will be a challenging task.
CHAPTER 7

7. References


Krehan, D., Nebral, K., Konig, M. & Strehl, S. Monoallelic loss and frequent mutation of the second allele of PAX5 are genetic hallmarks of dic(9;20) childhood acute lymphoblastic leukemia (in preparation).


Nebral, K., Denk, D., Attarbaschi, A., Konig, M., Mann, G., Haas, O.A. & Strehl, S. Incidence and diversity of PAX5 fusion genes in childhood acute lymphoblastic leukaemia. *Leukemia (in press).*


8.1. *ETV6-NCOA2* fusion defines a new entity of T-lymphoid/myeloid progenitor acute leukemia

Sabine Strehl, Karin Nebral, Margit König, Jochen Harbott, Stephanie Struski, Bella Bielorai, Michel Lessard, Herbert Strobl, Martin Zimmermann, Oskar A. Haas, Shai Izraeli

ETV6-NCOA2: A Novel Fusion Gene in Acute Leukemia Associated with Coexpression of T-Lymphoid and Myeloid Markers and Frequent NOTCH1 Mutations

Sabine Strehl, Karin Nebral, Margit König, Jochen Harbott, Herbert Strobl, Richard Ratei, Stephanie Struski, Bella Bielorai, Michel Lessard, Martin Zimmermann, Oskar A. Haas, and Shai Izraeli

Abstract

Purpose: The ETV6 gene has been reported to be fused to a multitude of partner genes in various hematologic malignancies with 12p13 aberrations. Cytogenetic analysis of six cases of childhood acute lymphoblastic leukemia revealed a novel recurrent t(8:12) (q13:p13), suggesting involvement of ETV6.

Experimental Design: Fluorescence in situ hybridization was used to confirm the involvement of ETV6 in the t(8:12) (q13:p13) and reverse transcription-PCR was used to identify the ETV6 partner gene. Detailed immunologic characterization was done, and owing to their lineage promiscuity, the leukemic blast cells were analyzed for NOTCH1 mutations.

Results: We have identified a novel recurrent t(8:12) (q13:p13), which results in a fusion between the transcriptional repressor ETV6 (TEL) and the transcriptional coactivator NCOA2 (TIF2) in two cases of childhood leukemia expressing both T-lymphoid and myeloid antigens. The ETV6-NCOA2 transcript encodes a chimeric protein that consists of the pointed protein interaction motif of ETV6 that is fused to the COOH terminus of NCOA2, including the cyclic AMP–responsive element binding protein–binding protein (CBP) interaction and the AD2 activation domains. The absence of the reciprocal NCOA2-ETV6 transcript in one of the cases suggests that the ETV6-NCOA2 chimeric protein and not the reciprocal NCOA2-ETV6 is responsible for leukemogenesis. In addition, ETV6-NCOA2 leukemia shows a high frequency of heterozygotic activating NOTCH1 mutations, which disrupt the heterodimerization of the PEST-domain.

Conclusions: The ETV6-NCOA2 fusion may define a novel subgroup of acute leukemia with T-lymphoid and myeloid features, which is not associated with a high prevalence of NOTCH1 mutations.

Authors’ Affiliations: Children’s Cancer Research Institute, St. Anna Kinderkrebsforschung, Vienna, Vienna Competence Center; St. Anna Children’s Hospital, Vienna, Austria; Department of Pediatric Hematology and Oncology, Children’s University Hospital, Gissis, Germany; Department of Hematology, Oncology and Tumor Immunology Robert-Roessle-Clinic at the HELIOS Klinikum Berlin, Charité Medical School, Berlin, Germany; Laboratoire d’Hematologie, Hôpital de Hautpoups, Strasbourg, France; Pediatric Hemato-Oncology, Cancer Research Center, Sheba Medical Center, Tel Hashomer, Ramat Gan, Israel; Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel; and Department of Pediatric Hematology/Oncology, Children’s Hospital, Hannover Medical School, Hannover, Germany.

Grant support: Austrian Ministry of Science and Research grant GEN-AU #GZ 200136/1-VO-1/2005 (S. Strobl and O.A. Haas); St. Anna Kinderkrebsforschung; Israel Cancer Association; Israel Science Foundation (S. Izraeli); and Waxman Cancer Research Foundation (S. Izraeli).

Acute leukemia is subdivided into myeloid or lymphoid according to cytology and immunophenotyping. However, using these criteria, a minor proportion of acute leukemia is difficult to unambiguously assign because the blast populations do not allow classification in either a pure myeloid or lymphoid category (1). Such cases are designated as biphenotypic acute leukemia (BAL) or acute leukemia of ambiguous lineage.

In an effort to establish objective diagnostic criteria for defining such leukemias, several different immunologic classification and scoring systems, which are mostly based on the number and specificity of the lymphoid and myeloid markers expressed by the blast cells, have been introduced (1–4). Based on these categorizations, the prevalence of BAL has been determined to range from 2% to 5% in adult and childhood acute leukemia (5–7). Coexpression of myeloid and B-lymphoid antigens occurs in approximately 65% to 70%, whereas T-lymphoid and myeloid antigens account for 25% of all BAL and thus for ≤1% of acute leukemia (6).

As a basic principle, the classification of hematopoietic neoplasms should attempt to incorporate immunophenotypic, biologic, genetic, and clinical features to define specific disease
Human Cancer Biology

etties. However, due to the lack of specific genetic features, BCL is defined by the application of arbitrary criteria (3). Genetic lesions such as t(9;22)(q34;q11)/BCR-ABL1, t(4;11)(q21;q23)/MLL-AF9/AF11, and other 11q23 abnormalities seem to commonly concur with BCL of the myeloid/B-lymphoid type, whereas the specific genetic features of myeloid/T-lymphoid BCL remain widely elusive (6–10). Thus far, two NUP98 fusions (i.e., NUP98-RAPGDS1 and NUP98-ADD3J) seem to be associated with a subset of adult T-cell acute lymphoblastic leukemia (T-ALL) with variable expression of mature T-cell and myeloid markers (11–13). Furthermore, the PICALM-MLLT10 (CALM-AF10) fusion, which is mainly associated with immature T-ALL, has also been found in leukemia with a multilineage phenotype coexpressing T-cell and myeloid antigens (14). Recently, a specific subset of acute myeloid leukemia (AML) with a multilineage phenotype in T-cell and myeloid malignancies with coexpression of T-cell genes that is characterized by silencing of the CEBPA gene and recurring mutations in NOTCH1 has been described (15).

Activating NOTCH1 mutations that disrupt the heterodimerization and/or the PEST domains are present in >50% of childhood T-ALL but only in rare cases of AML, particularly in the context of lineage switch leukemia (15–18). In contrast to the well-established role of NOTCH1 in T-ALL pathogenesis (19), controversial results have been obtained with regard to NOTCH1 signaling in myeloid development. Enforced expression of constitutively activated NOTCH1 or treatment with NOTCH1 ligands results in the inhibition of granulocyte differentiation and an increase of immature precursors, suggesting a potential role of NOTCH1 signaling in the development of myeloid leukemia (20–23). However, NOTCH1 signaling has also been shown to irreversibly reduce the self-renewal capacity of multipotent progenitors and to induce multilineage myeloid differentiation (24).

We have identified a novel genetic subtype of acute leukemia with a recurrent t(8;12)(q13;p13), which fuses the ETV6 (TEL) and NCOA2 (TFI2) genes, and a high prevalence of NOTCH1 mutations. The combination of these two genetic lesions seems to be specifically associated with acute leukemia with a mixed T-lymphoid and myeloid immunophenotype.

Materials and Methods

Patients. Bone marrow samples from children with newly diagnosed acute leukemia who were enrolled on either International Berlin-Frankfurt-Münster (IFM) or European Organization for Research and Treatment of Cancer (EORTC) protocols were obtained after informed consent of the patients or their legal guardians. Cytogenetic analysis of patients with acute leukemia enrolled in the Austrian AML-BFM 98 (n = 53, collected between 1998 and 2004; successfully karyotyped, n = 48) and the AML-BFM 2000 (n = 432, including n = 49 T-ALLs, collected between 2000 and 2006; successfully karyotyped, n = 403) studies identified two cases carrying a t(8;12)(q10.1;p10.13). Subsequently, additional patient samples with this specific translocation were collected from other study centers: two from Germany and one each from Israel and France.

Conventional and molecular cytogenetics. Samples were processed according to standard cytogenetic techniques and karyotypes were described according to the International System for Human Cytogenetic Nomenclature (25). ETV6 rearrangements were detected using differentially labeled ETV6 exon-specific cosmids 179A6 (exon 1A), 50F4 ( intron 1 and exon 2), 50F4 (exon 2), 163E7 (exons 3-5), 54DS (exons 5-8), and 148B6 (exon 8; kindly provided by P. Marquen, Center for Human Genetics, University of Leuven, Leuven, Belgium, ref. 26).

Results

Patient characteristics. Clinical characteristics and cytogenetic data of the patients are summarized in Table 2. All patients displayed a t(8;12) with variable breakpoints that were difficult to assign accurately by conventional cytogenetics. In three cases (cases 1, 2, and 4), the t(8;12) was the sole karyotypic abnormality, whereas in two cases (cases 3 and 6) it occurred within complex karyotypes. In four of the patients, the morphology of the blast cells was consistent with ALL, whereas two cases were undifferentiated AML. Taking morphologic and immunologic criteria into consideration, patients were treated with the most appropriate therapy regimen. Irrespective of the treatment protocol, four of the patients (cases 1, 2, 3, and 5) are in complete remission 40 to 84 months after initial diagnosis. Case 4 was initially enrolled in the ALL-BFM 2002 study but responded very poorly to prednisone. Thus, the protocol was modified and more anti-AML drugs were administered and complete remission could be achieved in week 6 of therapy. In addition, this patient underwent stem cell transplantation from a matched sibling and is in continuous remission 17 months...
after diagnosis. Case 6 was treated according to the AML-BFM 93 protocol, relapsed 16 months after diagnosis, and died of progressive disease.

The patients analyzed were selected based on the availability of cytogenetic data in various study centers, and thus, it is difficult to estimate the overall frequency of the (t;8;12)/ETV6-NCOA2 rearrangement. However, deduced from the numbers of patients enrolled in the Austrian AML-BFM 98 and ALL-BFM 2000 studies, in which one positive patient each was found, the t(8;12) is a rare but nevertheless recurrent genetic aberration. This finding agrees with previously reported cytogenetic data in childhood T-ALL (29).

Immunophenotyping. The percentages of blast cells positive for a specific immunologic marker are given in Table 3. According to the EGIL criteria (2), all patients had T-lymphoid scores of ≥2, consistently expressed cCD3 and CD7, and additional T-cell–specific markers, such as CD2 and CD5, to a variable extent. Except for case 3, the blast cells lacked expression of mCD3. In this leukemia, most of the cells (>50%) were devoid of mCD3 expression and as such are defined as T-ll. As a consequence of the myeloid antigen expression pattern, the patients were assigned to different EGIL subtypes (Table 3). True BAL is only considered when the scores are greater than two points for both the lymphoid and the myeloid lineage (2). Nevertheless, the blast cells of each patient expressed at least one myeloid marker, either MPO (EGIL score 2: three cases), CD33 (EGIL score 1: five cases), or CD13 (EGIL score 1: two cases) in combination with T-cell markers (Fig. 1). In case 6, two main subpopulations (i.e., cyCD3<sup>+</sup>/MPO<sup>+</sup> and cyCD3<sup>+</sup>/MPO<sup>+</sup>) were observed (Fig. 1B) and a minor population expressing both markers. Together with the percentage of positive blast cells for the individual markers, this leukemia fulfills the EGIL criteria for BAL. In two of the samples, also expression of CD56, a marker specific for mature natural killer cells, was observed. The immature stage of the majority of the T-lymphoid/myeloid leukemias was supported by the double negativity for CD4/CD8 and, except for one case, the presence of CD34 (Table 3).

### Table 1. Oligonucleotide primer sequences

<table>
<thead>
<tr>
<th>Designation</th>
<th>Oligonucleotide sequence</th>
<th>Direction</th>
<th>Gene/exon*</th>
<th>Transcript</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETV6ex23-F1</td>
<td>TCCTGGCTGCTGACAAAGAGG</td>
<td>Sense</td>
<td>ETV6/3</td>
<td>ETV6-NCOA2</td>
</tr>
<tr>
<td>TIF2ex17-R1</td>
<td>CAACCCTGACCTGACCTTCCA</td>
<td>Antisense</td>
<td>NCOA2/17</td>
<td></td>
</tr>
<tr>
<td>TIF2ex14-F1</td>
<td>AGATGGGGCCACGAGAAGTCC</td>
<td>Sense</td>
<td>NCOA2/14</td>
<td></td>
</tr>
<tr>
<td>TIF2ex13-R1</td>
<td>TCACACCGAGAATGCTGGATATC</td>
<td>Antisense</td>
<td>NCOA2/13</td>
<td></td>
</tr>
<tr>
<td>ETv6ex6-R1</td>
<td>ACGACTGAGGACACTGCTTTG</td>
<td>Antisense</td>
<td>ETV6/6</td>
<td></td>
</tr>
<tr>
<td>TIF2ex11-F1</td>
<td>GAGGATGACTTCGACTGTTG</td>
<td>Sense</td>
<td>NCOA2/11</td>
<td></td>
</tr>
<tr>
<td>TIF2ex9-F1</td>
<td>CCATGAGAGGAGCAATGAACCC</td>
<td>Sense</td>
<td>NCOA2/9</td>
<td></td>
</tr>
<tr>
<td>TIF2ex7-F1</td>
<td>CAGTGTGGCTGCTTCGTGACCC</td>
<td>Sense</td>
<td>NCOA2/7</td>
<td></td>
</tr>
<tr>
<td>ETv6ex8-R1</td>
<td>CCGCTGAGGAGGGATGTT</td>
<td>Antisense</td>
<td>ETV6/8</td>
<td></td>
</tr>
<tr>
<td>NOTCH1ex25-F1</td>
<td>GCCCTCTTCACAGGGTTGG</td>
<td>Sense</td>
<td>NOTCH1/25</td>
<td>NOTCH1</td>
</tr>
<tr>
<td>NOTCH1ex26-R1</td>
<td>TTCTGCTCTGCTGACACTTTG</td>
<td>Antisense</td>
<td>NOTCH1/26</td>
<td>NOTCH1</td>
</tr>
<tr>
<td>NOTCH1ex34-F1</td>
<td>TCCCGTCTGCGAGACTCT</td>
<td>Sense</td>
<td>NOTCH1/34</td>
<td></td>
</tr>
<tr>
<td>NOTCH1ex34-R1</td>
<td>GCTCGCTCTCTGCCCTAGGG</td>
<td>Antisense</td>
<td>NOTCH1/34</td>
<td></td>
</tr>
<tr>
<td>NOTCH1ex34-F2</td>
<td>AGCGGGCTCTGGCCTGAGG</td>
<td>Sense</td>
<td>NOTCH1/34</td>
<td></td>
</tr>
<tr>
<td>NOTCH1ex34-R2</td>
<td>GCAACAGACGCCGCAAGG</td>
<td>Antisense</td>
<td>NOTCH1/34</td>
<td></td>
</tr>
</tbody>
</table>

*Exon nomenclature according to the Ensembl Genome Browser exon information (http://www.ensembl.org/).

### Table 2. Clinical characteristics and cytogenetic data of patients with coexpression of T-lymphoid and myeloid antigens

<table>
<thead>
<tr>
<th>Case</th>
<th>Age (y)</th>
<th>Sex</th>
<th>FAB</th>
<th>Cytogenetics</th>
<th>Treatment protocol</th>
<th>Last follow-up from diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.4</td>
<td>M</td>
<td>ALL</td>
<td>46.XY,t(8;12)[c13;p13][15]/46.XY[5]</td>
<td>ALL-BFM 2000</td>
<td>60 mo + CR</td>
</tr>
<tr>
<td>2</td>
<td>8.5</td>
<td>F</td>
<td>AML-M1</td>
<td>46.XX,t(8;12)[c10;p10][19]</td>
<td>AML-BFM 98</td>
<td>44 mo + CR</td>
</tr>
<tr>
<td>3</td>
<td>14.5</td>
<td>F</td>
<td>AML-M1</td>
<td>46.XX,t(8;12)[c13;p13],del(16)(p13)[7]/46.XX[12]</td>
<td>EORTC 98901</td>
<td>40 mo + CR</td>
</tr>
<tr>
<td>4</td>
<td>2.0</td>
<td>M</td>
<td>ALL</td>
<td>46.XY,t(8;12)[c12;p12][17]</td>
<td>ALL-IC-BFM 2002</td>
<td>17 mo + CR</td>
</tr>
<tr>
<td>5</td>
<td>2.5</td>
<td>M</td>
<td>AML-M1</td>
<td>47.XY,t(8;12)[c13;p11.2],[22</td>
<td>4]</td>
<td>AML-BFM 97</td>
</tr>
<tr>
<td>6</td>
<td>10.9</td>
<td>M</td>
<td>ALL</td>
<td>46.XY,t(8;12)[c21],-4.del(5)(q34),[23,del(5)[c23],del(6)[c21],t(8;12)[c11;p11],[del(11)[c11],del(13)[del(13)[q14],[72</td>
<td>21],+72</td>
<td>21][6]/ 46.XY[12]</td>
</tr>
</tbody>
</table>

Abbreviations: FAB, French-American-British; BFM, Berlin-Frankfurt-Munster; CR, complete remission; EORTC, European Organization for Research and Treatment of Cancer; DOD, dead of disease.

*Lost for follow-up.
### Table 3. Immunophenotyping of blast cells with ETV6-NCOA2 fusion and classification according to EGI1 criteria

<table>
<thead>
<tr>
<th>Case</th>
<th>CD34</th>
<th>CD4</th>
<th>CD8</th>
<th>CD1a</th>
<th>CD2</th>
<th>cyCD3</th>
<th>mCD3</th>
<th>CD5</th>
<th>CD7</th>
<th>TCR</th>
<th>TdT</th>
<th>MPO</th>
<th>CD13</th>
<th>CD33</th>
<th>CD65</th>
<th>CD117</th>
<th>CD56</th>
<th>EGIL</th>
</tr>
</thead>
<tbody>
<tr>
<td>a/b</td>
<td>g/d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>32</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>19</td>
<td>75</td>
<td>2</td>
<td>ND</td>
<td>ND</td>
<td>2</td>
<td>1</td>
<td>36</td>
<td>59</td>
<td>13</td>
<td>1</td>
<td>ND</td>
<td>T-1/I/Mt+y*</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>91</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>54</td>
<td>2</td>
<td>ND</td>
<td>ND</td>
<td>18</td>
<td>16</td>
<td>3</td>
<td>95</td>
<td>10</td>
<td>15</td>
<td>86</td>
<td>BAL</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>98</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>58</td>
<td>99</td>
<td>37</td>
<td>94</td>
<td>95</td>
<td>59</td>
<td>5</td>
<td>34</td>
<td>0</td>
<td>60</td>
<td>0</td>
<td>40</td>
<td>30</td>
<td>T-II/Mt+y*</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>74</td>
<td>74</td>
<td>14</td>
<td>8</td>
<td>7</td>
<td>9</td>
<td>14</td>
<td>94</td>
<td>34</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>ND</td>
<td>T-1/I/Mt+y*</td>
</tr>
<tr>
<td>5</td>
<td>84</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>45</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>75</td>
<td>ND</td>
<td>10</td>
<td>50</td>
<td>11</td>
<td>11</td>
<td>85</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>6</td>
<td>76</td>
<td>4</td>
<td>6</td>
<td>1</td>
<td>15</td>
<td>70</td>
<td>16</td>
<td>78</td>
<td>16</td>
<td>3</td>
<td>14</td>
<td>17</td>
<td>73</td>
<td>3</td>
<td>30</td>
<td>ND</td>
<td>ND</td>
<td>BAL</td>
</tr>
</tbody>
</table>

NOTE: Values in Table are expressed as percentage of positive blast cells. Abbreviation: ND, not determined. *Due to the lack of CD5 data discrimination between T-1 and T-2 stage not possible.

Identification of NCOA2 as ETV6 partner gene. FISH analysis with ETV6 exon-specific probes revealed disruption of this gene (Fig. 2A). To narrow down the breakpoint in E21FISH with various gene-specific probes located at 8q13.1-21.1 was carried out and the most likely breakpoint was determined to occur in chromosomal band 8q13.3 (data not shown). Out of the genes located in this chromosomal region, NCOA2 was considered a likely candidate gene because it has already been identified as partner gene of MYST3 (MOZ) in cases with an inv(8)(p11q13) (30). Fusion gene-specific reverse transcription-PCR experiments using primers located in exons 3 and 17 of ETV6 and NCOA2, respectively, led to the identification of chimeric ETV6-NCOA2 transcripts (Fig. 2B). Sequence analyses revealed an in-frame fusion between ETV6 exon 4 and NCOA2 exon 15 in five of the six cases (cases 1, 2, 3, 5, and 6; Fig. 2C). In case 5, also an alternatively spliced transcript lacking exon 16 of NCOA2 was detected. In the remaining case (no. 4), ETV6 exon 5 was fused in-frame to NCOA2 exon 14. Thus, the putative ETV6-NCOA2 consensus fusion protein consists of the ETV6 pointed domain (PNT) and the cyclic AMP-responsive element binding protein–binding protein (CBP) interaction (CID) and the AD2 (transactivation domain 2) domains of NCOA2 (Fig. 2D). The splice variant lacking NCOA2 exon 16 (case 5) would result in a partial deletion of the CID domain; however, the second transcript also encodes the putative consensus protein. Expression analysis for the reciprocal NCOA2-ETV6 identified two different chimeric transcripts that fused NCOA2 exon 14 to ETV6 exons 5 or 6, respectively, thus showing alternative splicing of ETV6 exon 5 (Fig. 2B). In case 4, despite all efforts using various primer combinations (see Table 1), no NCOA2-ETV6 transcripts could be amplified, suggesting that the ETV6-NCOA2 fusion gene is responsible for leukemogenesis.

Detection of NOTCH1 mutations. The identification of activating NOTCH1 mutations not only in T-ALL but also in rare cases of myeloid leukemia (13–18) prompted us to analyze our samples for the presence of NOTCH1 mutations. Mutation analysis of the heterodimerization and PEST domains of NOTCH1 revealed heterozygous NOTCH1 mutations in four of five samples of which sufficient material was available. Mutations were detected in one sample in the heterodimerization domain, in two in the PEST domain, and in one in both domains. As previously described (18), heterodimerization domain mutations were missense, in-frame deletions and insertions, whereas the PEST domain mutations created premature termination codons (Table 4). The two heterodimerization domain mutations found in cases 2 and 3, and the PEST domain mutations of cases 1 and 6, have been previously described (18). A PEST domain mutation similar to that detected in case 2 has as well been reported (31).

### Discussion

In this study, we report six cases of childhood acute leukemia with a novel recurrent (8:12), which results in a fusion of the repressor gene ETV6 and the transcriptional coactivator NCOA2. The ETV6-NCOA2 fusion concurs with a high prevalence of NOTCH1 activating mutations and the coexpression

---

**Fig. 1.** Fluorescence-activated cell sorting data. A and B, expression of cyCD3 and cytoplasmic MPO (cyMPO) in a subpopulation of the blast cells of case 2 (A) and case 6 (B). C and D, expression of CD7 and CD33 in case 1 (C) and case 4 (D). PE, phycoerythrin; APC, allophycocyanin.
ETV6-NCOA2 Fusion in T-Lymphoid/Mylloid Leukemia

Fig. 2. ETV6-NCOA2 fusion gene. A. FISH analysis of case 1 using ETV6 exon-specific cosmids clones S094 ( exon 2; green signal) and 14BBB (exon 8; red signals), showing disruption of the gene. B. Reverse transcription-PCR analysis. M, 100-bp ladder (Promega); lanes 1 to 6, patient samples; lane 7, negative control. B1, reverse transcription-PCR analysis of the ETV6-NCOA2 fusion transcript using primers located in exon 3 of ETV6 (ETV6ex3-F) and exon 17 of NCOA2 (TF2x717-R). B2, reverse transcription-PCR analysis of the reciprocal NCOA2-ETV6 fusion transcript using primers located in exon 14 of NCOA2 (TF2x214-F) and exon 6 of ETV6 (ETV6ex6-R). The size differences are due to alternative splicing of ETV6 exon E and NCOA2 exons 4 and 15, respectively. D, predicted amino acid sequence; amino acid 154 of ETV6 is fused to amino acid 1010 of NCOA2. E, schematic representation of the putative consensus ETV6-NCOA2 chimeric protein.

of T-cell–specific antigens (cyCD3 and CD7) and at least one myeloid marker (i.e., MPO, CD33, CD13, and CD65).

As a result of the (t(6,12)) translocation, the PNT protein interaction domain of ETV6, which is involved in homodimerization and heterodimerization, is fused to the COOH terminus of NCOA2, including the CBP interaction domain and the AD2 activation domain. The same COOH-terminal domains of NCOA2 are also retained in the previously identified MYST-NCOA2 fusion protein, which is generated by an inv(6)(p11q13) (30, 32). The absence of the reciprocal NCOA2-ETV6 transcript in one of the cases, as well as the facts that MYST-NCOA2 is transforming and that the reciprocal NCOA2-MYST3 is not expressed, suggests that the ETV6-NCOA2 chimeric protein and not the reciprocal NCOA2-ETV6 is responsible for leukemogenesis. However, this finding needs to be confirmed in a larger cohort of ETV6-NCOA2–positive cases.

The transforming properties of MYST3-NCOA2 depend on the MYST3 MYST and the NCOA2 CBP interaction domains, whereas the MYST3 PHD and the putative AD2 acetyltransferase motifs of NCOA2 are not required for transformation (33). Expression of MYST3-NCOA2 correlates with a depletion of CBP from PML bodies and reduced cellular levels of CBP. Thus, MYST3-NCOA2 acts as a modulator of the transcriptional activity of CBP-dependent activators (34, 35). Future studies will determine whether the ETV6-NCOA2 chimeric protein has similar properties or may recruit CBP to ETV6 target genes resulting in their constitutive activation. In this respect, it is interesting to note that Ets pointed domains can also interact with the SRC1 interaction (CID) domain of CBP (36), suggesting a possible competition between ETV6-NCOA2 and normal CBP for CBP. Moreover, all NUP98 chimera, including those associated with a T/myeloid phenotype (11–13), retain the COOH-terminal FG repeats, which have the potential to bind CBP (37), indicating a probable involvement of CBP in these types of leukemia.

Bona fide transcriptional coactivators are rarely involved in leukemogenic translocations. Such fusion genes, which are expressed as a consequence of chromosomal translocations, include MLL-CREBBP/CBP (38–41), MLL-EP300/300 (41), MYST3-CREBBP (42), MYST3-EP300 (43, 44), MYST4/MORF-CREBBP (45), and MYST3-NCOA2 (30, 32). In contrast to MYST3-NCOA2 and to all other rearrangements involving transcriptional coactivators, which are exclusively associated with AML, the ETV6-NCOA2 irrespective of its occurrence in different morphologic and immunophenotypic leukemia subtypes determined by the FAB and EGIL classifications coincides with coexpression of T-lymphoid and myeloid markers.

Recently, a similar novel subtype of AML with coexpression of T-lymphoid markers characterized by CEBPA silencing through promoter hypermethylation and associated with frequent NOTCH1 mutations has been identified (15). This subgroup of acute leukemia showed several cytogenetic abnormalities, none of which was common to all cases, suggesting that the CEBPA/NOTCH1 and the ETV6-NCOA2/NOTCH1 leukemia represent distinct entities.

Mixed myeloid and lymphoid T-cell–specific or B-cell–specific leukemia has been previously described and may either represent an unphysiologic transformation-related anomaly or, alternatively, reflect the immunophenotypic features of a physiologic common myeloid-lymphoid progenitor cell (46). The concurrence of the ETV6-NCOA2 fusion with NOTCH1 activating mutations raises the question whether this combination of genetic lesions or ETV6-NCOA2 alone has instructive properties or at least one of them targets an uncommitted progenitor cell with a myeloid and T-lymphoid potential. In this context, the
presence of NOTCH1 mutations in lineage infidelity (17) as well as in AML with T-lymphoid features (15) suggests that these mutations may occur in a leukemic stem cell that precedes both myeloid and T-lineage commitment (17). Accumulating evidence indicates that a myeloid potential accompanies early stages of T (also B and erythroid) development and that T-cell (and B-cell) progenitors are most likely produced from a common myeloid-lymphoid progenitor through intermediate bipotent or even multipotent stages (47–52). On the other hand, some fusion proteins encoded by translocations impart, through ectopic reactivation of genes associated with self-renewal, leukemia stem cell properties to committed hematopoietic progenitors (53, 54).

Although at this point the sequental or simultaneous occurrence of the two genetic events, fusion of ETV6 to NCOA2 and mutation of NOTCH1, remains indeniable, it is tempting to speculate that mutation of NOTCH1 confers self-renewal capacity to early progenitors, which are then susceptible to the accumulation of additional genetic hits such as fusion genes. Vice versa, the ETV6-NCOA2 fusion might bestow transforming properties and the subsequent activation of NOTCH1 may directly differentiate toward the T-cell lineage. Ectopic expression of ETV6-NCOA2 will eventually reveal whether this chimeric protein reprograms progenitors or leads to expansion of a T/myeloid subset.

The eminent clinical question that derives from our findings is whether this type of leukemia should be treated as ALL or AML. The currently available rather limited information indicates that the affected patients have a decent prognostic outlook when treated with either ALL or AML therapy regimens. However, only the retrospective and prospective collection, evaluation, and comparison of treatment results of a larger number of such cases will provide a conclusive answer. The identification of the ETV6-NCOA2 gene fusion and its accompanying NOTCH1 mutations is thus the first essential step to achieve this goal.

References

25. Shaffer L, Marks F. ISCN 2005: an international
33. Dugaschi K, Aynon PM, Canepa M, et al. MOZ-TIF2-induced acute myeloid leukemia requires the MOZ nucleosome binding motif and TIF2-mediated recruitment of CBF.
## 8.2. Additional FISH clones

### Table 1. Additional FISH clones used for analysis

<table>
<thead>
<tr>
<th>Gene/Region</th>
<th>Locus</th>
<th>Clone</th>
<th>Accession Number/ Company</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSCD1-5'</td>
<td>01q24.2</td>
<td>RP3-455J7*</td>
<td>AL031733.3</td>
</tr>
<tr>
<td>RSCD1-3'</td>
<td>01q24.2</td>
<td>RP11-138P14*</td>
<td>AQ349162 and AQ385458 (BAC end pairs)</td>
</tr>
<tr>
<td>3p</td>
<td>03p</td>
<td>3p-arm-specific painting probe</td>
<td>NA</td>
</tr>
<tr>
<td>TOP2B</td>
<td>03p24.2</td>
<td>RP11-659P16‡</td>
<td>AC093416</td>
</tr>
<tr>
<td>3q-subtelomere</td>
<td>03q subtelomere</td>
<td>196F4#</td>
<td>NA</td>
</tr>
<tr>
<td>Chromosome 7</td>
<td>whole chromosome</td>
<td>pBS7</td>
<td>NA</td>
</tr>
<tr>
<td>AUTS2-5'</td>
<td>07q11.22</td>
<td>RP11-243F5*</td>
<td>AQ486680</td>
</tr>
<tr>
<td>AUTS2-3'</td>
<td>07q11.22</td>
<td>RP11-88H22*</td>
<td>AZ516257</td>
</tr>
<tr>
<td>ELN-5'</td>
<td>07q11.23</td>
<td>RP11-148M21*</td>
<td>AC093168</td>
</tr>
<tr>
<td>ELN-3'</td>
<td>07q11.23</td>
<td>RP11-349P21*</td>
<td>AQ528659</td>
</tr>
<tr>
<td>Chromosome 9</td>
<td>whole chromosome</td>
<td>pBS9</td>
<td>NA</td>
</tr>
<tr>
<td>5' of PAX5</td>
<td>09p13</td>
<td>RP11-397D12‡</td>
<td>AL158155.24</td>
</tr>
<tr>
<td>5' of PAX5</td>
<td>09p13</td>
<td>RP11-3J10‡</td>
<td>AL138752</td>
</tr>
<tr>
<td>3' of PAX5</td>
<td>09p13</td>
<td>RP11-84P7‡</td>
<td>AL161792</td>
</tr>
<tr>
<td>LSI BCR/ABL Dual Color, Dual Fusion Translocation Probe Set</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>09q34 / 22q11.2</td>
<td></td>
<td></td>
<td>Vysis (Abott)</td>
</tr>
<tr>
<td>ELP4-5'</td>
<td>11p13</td>
<td>RP5-1137O17*</td>
<td>AL133295</td>
</tr>
<tr>
<td>ELP4-3' and PAX6</td>
<td>11p13</td>
<td>RP11-307I15*</td>
<td>DX934410 and DX934411 (BAC end pairs)</td>
</tr>
<tr>
<td>WT1</td>
<td>11p13</td>
<td>RP11-74J1</td>
<td>Research Genetics - Invitrogen</td>
</tr>
<tr>
<td>BDNF-5'</td>
<td>11p14.1</td>
<td>RP11-52H5*</td>
<td>AQ083161 and AQ115784 (BAC end pairs)</td>
</tr>
<tr>
<td>KCNA4-3'</td>
<td>11p14.1</td>
<td>RP11-215H22*</td>
<td>AC021233.9</td>
</tr>
<tr>
<td>LGR4-3' and CCDC34</td>
<td>11p14.1</td>
<td>RP11-426P16*</td>
<td>AQ554326 and AQ554330 (BAC end pairs)</td>
</tr>
<tr>
<td>Gene</td>
<td>Chromosome</td>
<td>Location</td>
<td>Accession Numbers</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
<td>----------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>MUC15 and TMEM16C</td>
<td>11p14.2</td>
<td>RP11-283H3*</td>
<td>AC036114</td>
</tr>
<tr>
<td>LUZP2-3'</td>
<td>11p14.3</td>
<td>RP11-54J7*</td>
<td>AC087373</td>
</tr>
<tr>
<td>GAS2-5'</td>
<td>11p14.3</td>
<td>RP11-109H8*</td>
<td>AQ350699 and AQ323046 (BAC end pairs)</td>
</tr>
<tr>
<td>LUZP2-5'</td>
<td>11p14.3</td>
<td>RP11-372B5*</td>
<td>AC040968</td>
</tr>
<tr>
<td>NELL1-5'</td>
<td>11p15.1</td>
<td>RP11-3E17*</td>
<td>AC010811</td>
</tr>
<tr>
<td>NELL1-3'</td>
<td>11p15.1</td>
<td>RP11-116O9*</td>
<td>AC105190</td>
</tr>
<tr>
<td>E2F8</td>
<td>11p15.1</td>
<td>RP11-428C19*</td>
<td>AC009652</td>
</tr>
<tr>
<td>DKK3</td>
<td>11p15.3</td>
<td>RP11-313H5*</td>
<td>AC118656</td>
</tr>
<tr>
<td>GALNTL4-3'</td>
<td>11p15.3</td>
<td>RP11-75K2*</td>
<td>AQ267423</td>
</tr>
<tr>
<td>GALNTL4-5'</td>
<td>11p15.3</td>
<td>RP11-47D7*</td>
<td>AQ200362</td>
</tr>
<tr>
<td>LMO1</td>
<td>11p15.4</td>
<td>RP11-379P15*</td>
<td>AC091013</td>
</tr>
<tr>
<td>WEE1</td>
<td>11p15.4</td>
<td>RP11-16F15*</td>
<td>AC011979</td>
</tr>
<tr>
<td>SWAP70-3'</td>
<td>11p15.4</td>
<td>RP11-540A21*</td>
<td>AC026250.16</td>
</tr>
<tr>
<td>MRVI1</td>
<td>11p15.4</td>
<td>RP11-58H20*</td>
<td>AC009532</td>
</tr>
<tr>
<td>LSI® TEL/AML1</td>
<td>12p13 / 22q22</td>
<td>NA</td>
<td>Vysis (Abott)</td>
</tr>
<tr>
<td>Chromosome 14</td>
<td>whole chromosome</td>
<td>pBS14</td>
<td>NA</td>
</tr>
<tr>
<td>LSI® IGH DUAL</td>
<td>14q32</td>
<td>NA</td>
<td>Vysis (Abott)</td>
</tr>
<tr>
<td>break apart</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BCL11B-5'</td>
<td>14q32.2</td>
<td>RP11-431B1*</td>
<td>AC036222 and AC103702 (BAC end pairs)</td>
</tr>
<tr>
<td>BCL11B-3'</td>
<td>14q32.2</td>
<td>RP11-15E14*</td>
<td>AL352981</td>
</tr>
</tbody>
</table>

*obtained from The Wellcome Trust Sanger Institute; http://www.sanger.ac.uk, Hinxton, Cambridge, United Kingdom.

†obtained from M. Rocchi, Department of Cytogenetics, University of Bari, Bari, Italy.

‡obtained from L. Kearney, MRC Medical Research Council, John Radcliffe Hospital, Molecular Haematology Unit, Headington, Oxford, United Kingdom.

NA, not applicable
### 8.3 Additional PCR primers

#### Table 2. Additional oligonucleotide primer sequences

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5' - 3')</th>
<th>Direction</th>
<th>Gene/Exon¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDNFex2-R1</td>
<td>TTCTGGTCCTCATCCAACACG</td>
<td>antisense</td>
<td>BDNF/2</td>
</tr>
<tr>
<td>BDNFex2-R2</td>
<td>CGGCAACAAACCACAACATTAC</td>
<td>antisense</td>
<td>BDNF/2</td>
</tr>
<tr>
<td>ELNex16-R1</td>
<td>GCACGCCAGGAACACCAG</td>
<td>antisense</td>
<td>ELN/16</td>
</tr>
<tr>
<td>ELNex18-R1</td>
<td>GCCACCAGCCACTAAGC</td>
<td>antisense</td>
<td>ELN/18</td>
</tr>
<tr>
<td>ELNex6-R2</td>
<td>AGCAGCGTCAGCCACTCCAC</td>
<td>antisense</td>
<td>ELN/6</td>
</tr>
<tr>
<td>ETV6ex3-R1</td>
<td>CCTCTTTGTGTCAGCAGGGAGGAG</td>
<td>antisense</td>
<td>ETV6/3</td>
</tr>
<tr>
<td>FOXP1ex11-R1</td>
<td>TTGTTCCTGTGTTTCTCTGC</td>
<td>antisense</td>
<td>FOXP1/11</td>
</tr>
<tr>
<td>FOXP1ex14-R1</td>
<td>GGCAGCTTTTGTTCTGTAG</td>
<td>antisense</td>
<td>FOXP1/14</td>
</tr>
<tr>
<td>FOXP1ex8-R1</td>
<td>TGCTGGAGAGAGACCTGGAG</td>
<td>antisense</td>
<td>FOXP1/8</td>
</tr>
<tr>
<td>HIPK1ex8-F1</td>
<td>CTGCACCAGTTCTGGAGTTG</td>
<td>sense</td>
<td>HIPK1/8</td>
</tr>
<tr>
<td>PAX5ex3-F1</td>
<td>CCATGTTTGCTGGAGATCAG</td>
<td>sense</td>
<td>PAX5/3</td>
</tr>
<tr>
<td>PAX5ex5-F1</td>
<td>TACTCCATCAGGGGATCC</td>
<td>sense</td>
<td>PAX5/5</td>
</tr>
<tr>
<td>PAX5ex6-7R1</td>
<td>TGGCTGAATACTCTGTGGTCTGTC</td>
<td>antisense</td>
<td>PAX5/6-7</td>
</tr>
<tr>
<td>PAX5ex6-F1</td>
<td>CTGGACGGCTGGTGTGAGAGA</td>
<td>sense</td>
<td>PAX5/6</td>
</tr>
<tr>
<td>PAX5ex7-R1</td>
<td>GGCCTTCATGTGTCAGCC</td>
<td>antisense</td>
<td>PAX5/7</td>
</tr>
<tr>
<td>PAX5ex9-10-R1</td>
<td>GCTATAATAGTGAGGGGAGCCAAGCA</td>
<td>antisense</td>
<td>PAX5/9-10</td>
</tr>
<tr>
<td>PAX5ex9-R2</td>
<td>ACGAGGAAATCTGAGGGTGCTGT</td>
<td>antisense</td>
<td>PAX5/9</td>
</tr>
<tr>
<td>PMLex1-F2</td>
<td>TCAGCTTCTCTTCAGCCACTC</td>
<td>sense</td>
<td>PML/1</td>
</tr>
<tr>
<td>ZNF521ex4-R1</td>
<td>TGAGAGCGGTCTGGTTTCCCT</td>
<td>antisense</td>
<td>ZNF521/4</td>
</tr>
<tr>
<td>ZNF521ex7-R1</td>
<td>CTGCTGCAACTGTGGTTTGC</td>
<td>antisense</td>
<td>ZNF521/7</td>
</tr>
</tbody>
</table>

¹Exon nomenclature according to the Ensembl Genome Browser exon information (http://www.ensembl.org/).
Fusion transcripts were amplified using the following primer combinations.

**PAX5-BDNF:**
- PAX5ex5-F1 and BDNFex2-R1
- PAX5ex5-F1 and BDNFex2-R2

**PAX5-ELN:**
- PAX5ex3-F1 and ELNex6-R2 or ELNex16-R1
- PAX5ex5-F1 and ELNex18-R1

**PAX5-ETV6:**
- PAX5ex3-F1 and ETV6ex3-R1

**PAX5-FOXP1:**
- PAX5ex6-F1 and FOXP1ex8-R1
- PAX5ex3-F1 and FOXP1ex11-R1
- PAX5ex5-F1 and FOXP1ex14-R1

**PAX5-ZNF521:**
- PAX5ex3-F1 and ZNF521ex4-R1
- PAX5ex6-F1 and ZNF521ex7-R1

**HIPK1-PAX5:**
- HIPK1ex8-F1 and PAX5ex6-7-R1

**PML-PAX5:**
- PMLex1-F2 and PAX5ex7-R1

To identify novel **PAX5** fusion partner genes RACE was performed using various **PAX5** gene specific primers in combination with universal 5' or 3' primers provided by the Marathon Kit (Takara Bio Europe/Clontech, Saint-Germain-en-Laye, France) (Nebal et al, in press). Additional 5' RACE experiments were performed using PAX5ex9-10-R1 and AP1 for the first round of amplification, and PAX5ex9-R2 and AP2 for the nested second round of amplification. Alternatively, instead of a "touch down" PCR reaction a "one step" PCR was performed using the following cycling parameters: 95°C initial denaturation for 1min, 35 cycles of 94°C for 15sec, and 68°C for 5-8min, and final elongation at 68°C for 6min. The nested PCR was done for 30 cycles of 94°C for 15sec, and 68°C for 5-8min.
DEUTSCHE ZUSAMMENFASSUNG

PAX5 ist ein Transkriptionsfaktor der Paired Box Genfamilie, welcher sowohl für die Entwicklung von B-Zellen aus hämatopoietischen Vorläuferzellen als auch für deren Fortbestand als reife B-Zellen von essentieller Bedeutung ist. Vor kurzem wurde gezeigt, dass PAX5 in B-Zell-Neoplasien häufig involviert ist. Die entsprechenden genetischen Alterationen umfassen Punktmutationen und Deletionen und, von besonderem Interesse im Zusammenhang mit meiner Arbeit, auch Genarrangements. In B-Zell-Non-Hodgkin-Lymphomen mit einer t(9;14)(p13;q32), zum Beispiel, gelangt das PAX5 Gen durch die Translokation in die Nähe des IGH@ Locus, wodurch es zu einer Überexpression von PAX5 kommt. Im Gegensatz dazu wurden bei akuten lymphpatischen B-Vorläuferzell Leukämien ("B-cell-precursor"; BCP-ALL) Fusionen des PAX5 Gens mit FOXP1 (3p13), AUTS2 (7q11), ELM (7q11), ETV6 (12p13), ZNF521 (18q11) und C20orf112 (20q11) beschrieben, wobei die daraus resultierenden Fusionstranskripte für chimäre Proteine kodieren.


Zum Nachweis von potentiellen PAX5 Genarrangements wurden eigene Fluoreszenz in situ Hybridisierung (FISH) Assays entwickelt und die Interphase FISH Muster mithilfe der Metafer4-Metacyte Software (Metasystems) automatisch ausgewertet. Um alle im PAX5 Gen möglichen Bruchpunkte zu erfassen (auch jene, die zu einer veränderten Expression des intakten PAX5 Gens durch ein Partnergen führen) wurden primär das Gen flankierende zwei Farben FISH Sonden verwendet. Auffällige FISH Muster wurden weiter mit Gen-spezifischen Sonden abgeklärt. Anschließend wurden die Fusionspartner entweder mit FISH und/oder 3’- oder 5’-RACE (Rapid Amplification of cDNA ends) identifiziert und das Vorhandensein der chimären Transkripte mittels RT-PCR (Reverse Transcription-PCR) und Sequenziierung überprüft.

10 von 446 untersuchten Proben (2.2%) zeigten ein für ein PAX5 Rearrangement charakteristisches FISH Muster. Ein Fall mit einer PAX5-ETV6 Fusion war bereits aus Vorbefunden bekannt, bei einem weiteren wurde eine kürzlich in der Literatur beschriebene PAX5-C20orf112 Genfusion gefunden und bei einem Patienten konnte die vorliegende PAX5 Aberration trotz intensiver Analysen nicht vollständig aufgeklärt werden. Jedoch gelang es uns bei sieben Fällen sechs neue PAX5 Fusionspartner zu identifizieren: HIPK1 (1p13), POM121 (7q11), JAK2 (9p24), DACH1 (13q21), PML (15q24) und BRD1 (22q13.33). Mit Ausnahme von zwei PAX5-JAK2 positiven Fällen, traten alle anderen Fusionen jeweils nur
einmal auf. Weiters konnten, zumindest bei der ALL im Kindesalter, keine PAX5-aktivierenden Translokationen gefunden werden. Aufgrund der großen Heterogenität der Fusionspartner ist die von uns etablierte FISH Strategie eine verlässliche und sinnvolle Screeningmethode mit der alle potentiellen PAX5 Genfusionen systematisch erfasst werden können.

Aus dem Ergebnis meiner Studie kann man zusammenfassend ableiten, dass zirka 2.5% der BCP-ALL Fälle PAX5 Genfusionen aufweisen, welche jedoch eine Vielzahl von Partnergenen betreffen, die nicht nur Transkriptionsfaktoren sondern auch Strukturproteine und eine Tyrosinkinase, einschließen. Die hypothetischen Fusionsproteine bestehen in allen Fällen zumindest aus der PAX5-paired DNA-Bindungsdomäne, die mit der C-terminalen Region oder sogar dem gesamten Protein des jeweiligen Partners fusioniert. Die Struktur der PAX5 chimären Proteine legt nahe, dass diese zwar an PAX5 Zielgene binden können, aber keine normale transkriptionelle Regulation ausüben und dadurch der intrinsischen PAX5-Aktivität entgegenwirken, was möglicherweise bei der Pathogenese der BCP-ALL eine Rolle spielt.
CURRICULUM VITAE

PERSÖNLICHE DATEN

Name Karin Nebral
Geburtsdatum 28.10.1977
Geburtsort Gmünd, Niederösterreich, Österreich
Staatsangehörigkeit Österreich

AUSBILDUNG

1984-1988 Volksschule Schwarzenau
1988-1992 Sporthauptschule Waidhofen/Thaya
1992-1996 Bundesrealgymnasium Zwettl
06/1996 Reifeprüfung mit ausgezeichnetem Erfolg bestanden

STUDIUM

10/1996-05/2004 Diplomstudium der Biologie, Studienzweig Anthropologie, Universität Wien
09/2001-07/2002 Erasmusstipendium an der Universität Tor Vergata, Rom, Italien
10/2002-05/2004 Diplomarbeit am CCRI, St. Anna Kinderkrebsforschung, Wien
Gruppenleiter: Univ. Prof. Dr. Oskar A. Haas
Thema: "Detection and characterization of NUP98 rearrangements in leukemia"
Supervisor: Dr. Sabine Strehl
05/2004 2. Diplomprüfung mit Auszeichnung bestanden; Verleihung des akademischen Titels: Magistra der Naturwissenschaften (Mag. rer. nat.)
Gruppenleiterin: Dr. Sabine Strehl
Thema: "Detection and characterization of PAX5 aberrations in childhood acute lymphoblastic leukemia"
10/2005-2008 Doktoratsstudium Biologie, Studienzweig Anthropologie, Universität Wien

SONSTIGES

11.-12.9.2003 Schulungen: Metafer4 Metacyte System (Metasystems)
06/2008 Travel Grant: 13th Congress of the EHA (European Hematology Association), Kopenhagen, Dänemark


DANKSAGUNG

Ein großes DANKE an all jene, die mich während der letzten Jahre begleitet haben –
einerseits allen KollegInnen des CCRI, die mir immer mit Rat & Tat zur Seite standen und die
den Arbeitsalltag mit einer positiven sozialen Komponente erfüllten - andererseits allen
Freunden, die mit mir durch Ups & Downs gegangen sind, mir in den Phasen des Zweifels
zugehört haben und die mir gezeigt haben, dass es auch noch ein anderes Leben neben der
Arbeit gibt, das es zu genießen & zu leben heißt.

Einzelne Namen zu nennen ist schwierig (ich hoffe, es ist niemand beleidigt, der nicht
genannt wird):

Oskar, der mir die Tür ins CCRI geöffnet hat, dem ich diesen wissenschaftlichen Werdegang
verdanke & den ich für sein unglaubliches Wissen und sein Selbstverständnis für die
Bedeutung des Lernens und Lehrens bewundere und den ich auch privat sehr schätze.

Sabine, eine Betreuerin, die mensch kein zweites Mal findet! Sie hat mich immer mit großem
Engagement & einem unendlichen Wissen geleitet. Auch wenn es manchmal hart war, die
Dinge richtig zu machen, so war es das Endergebnis immer wert. Sie hat mich sowohl
wissenschaftlich als auch sozial geprägt und mich immer voll & ganz unterstützt, auch wenn
es mir mal nicht so gut ging.

Weiters allen meinen lieben Labor-KollegInnen für Hilfe im Laboralltag, für gemeinsame
Kaffeepausen, für das freundschaftliche und respektvolle Miteinander:

Anne, Barbara, Bettina, Bettina, Brigitte, Brigitte, Dagmar, Daniela, Eva, Gudrun, Margit,
Markus, Nicky, Ulli, Susanna, Trude
& allen aus dem Labor 7

für die Freundschaft, die angeregten Diskussionen & die Unterstützung bei Experimenten,
sowie die Hilfestellung bei klinischen Fragen: Andrea, Gerd, Ruth & Renate und Andishe.

sowie für methodische und geräte-technische Unterstützung, Tipps & Tricks ein Dank an das
gesamte CCRI!

Vielen Dank auch an mein PhD-Komitee: Ulrich Jäger, Berthold Streubel und Johannes
Schmid, die mich durch positives Feedback und konstruktive Diskussion immer sehr
motivierten.

Von meinen Freunden besonders Marc, Sascha, Erich, Susi & Dani.

Der größte Dank gehört meiner Familie, der Basis meines Lebens. Danke für das
Verständnis, die immerwährende Unterstützung, dass wir alle für einander da sind!

Ich widme diese Dissertation meinem Opa, der mich immer wieder fragte, wann ich nun
den Arbeitstag fertig sei, und der leider die Antwort auf diese Frage "jetzt!" nicht mehr miterleben
durfte.