PhD Program in Translational and Molecular Medicine

DIMET

University of Milano-Bicocca

School of Medicine and School of Science

Impact of *PAX5* alterations on gene expression and signaling pathways in ALL

Coordinator: Prof. Andrea Biondi

Tutor: Prof. Andrea Biondi

Co-Tutor: Dr. Giovanni Cazzaniga

Dr. Valeria Cazzaniga

Matr. No. 062965

XXVI Cycle

Academic Year 2012-2013

Table of Contents

Chapter 1

General introduction

1. Molecular basis of leukemia

Acute lymphoblastic leukemia (ALL) is a hematological tumor, characterized by the malignant expansion of immature cells from lymphoid lineages,[1] or, as recently proposed, from a population of transformed hematopoietic stem/progenitor cells (leukemic stem cells).[2] The disease is characterized by the clonal proliferation and accumulation of malignant blast cells in the bone marrow and peripheral blood.[3]

ALL is the most common pediatric malignancy with a peak at 2±5 years of age, comprising 25% of cancers occurring before 15 years of age and 19% among those less than 20 years old.[3,4] Indeed, it occurs at an incidence of up to 30 childhood cases per one million populations per year in the USA. Although 80% of children are cured, relapse ALL remains a leading cause of childhood morbidity and mortality.[5]

ALL, like cancer in general, probably arises from interactions between exogenous or endogenous exposures, genetic (inherited) susceptibility and chance.[6] The challenge is to identify the relevant exposures and inherited genetic variants and decipher how and when these factors contribute to the multistep natural history of ALL from initiation (usually in utero) through the largely covert evolution to overt disease.[7] In the last years, many and many factors have been candidate to contribute to childhood disease through epidemiological and case-control studies,[8] such as ionizing radiations, electromagnetic fields, chemicals, infections and constitutional genetic defects.[1]

Approximately 85% of diagnosed ALL cases are a result of the expansion of B-cell precursors (BCP-ALL), and 15% correspond to T-cell precursors aberrancies (T-ALL). It is possible to further sub-classify the BCP-ALL in early pre-B cell or pro-B, common, pre-B cell, and B-cell ALL, according to the corresponding differentiation markers expressed by the leukemic cells.[1]

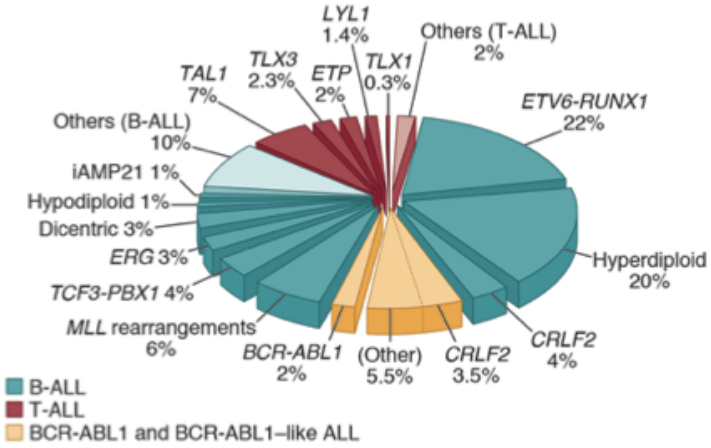


Figure 1. Frequency of cytogenetic subtypes of pediatric ALL.[9]

In particular, the differentiative stages of leukemia are identified on the basis of the expressed antigens on the blast:

a) Early pre-B (o Pro-B): CD19+ CD10- HLA-DR+ CD7- SmIg-
b) CALL (common): CD19+ CD10+ HLA-DR+ CD7- SmIg-
c) Pre-B: CD19+ CD10+ HLA-DR+ CD7- SmIg+ (κ- e λ-)
d) B: CD19+ CD10+/- HLA-DR+ CD7- SmIg+ (κ+ o λ+)
e) T: CD19- CD7+ CD3+.

Most childhood BCP-ALL may be sub-classified by the presence of either gross or submicroscopic genetic alterations.

Approximately 75% of B-ALL cases exhibit aneuploidy or a recurring gross chromosomal rearrangement. These rearrangements commonly perturb genes encoding regulators of hematopoiesis, tumor suppressors, oncogenes, or tyrosine kinases but commonly require additional genetic hits to establish the full leukemic phenotype.

In the last five years, genome-wide profiling using microarrays, candidate gene, and second-generation sequencing have provided a number of key insights into the genetic basis of ALL. These studies have identified new subtypes of ALL and have uncovered recurring submicroscopic genetic alterations in known ALL subtypes. These include loss-of-function mutations involving genes regulating lymphoid development that contribute to the arrest in maturation characteristic of B-ALL, mutations that inactivate

tumor suppressor and cell cycle regulatory proteins, and mutations that drive cytokine receptor and/or kinase signaling. Thus, concomitant lesions disrupting hematopoietic development and tumor suppression as well as driving signaling and proliferation are hallmarks of many ALL subtypes.[9]

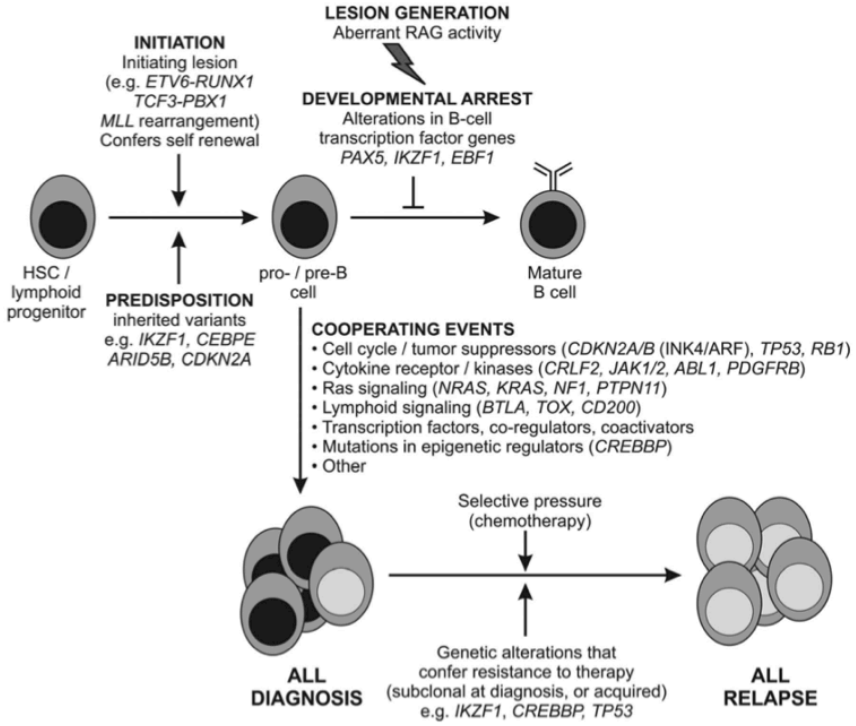


Figure 2. Proposed schema for the role of genetic alterations in the pathogenesis of BCP-ALL.[5]

Herein, we report on *PAX5* fusion genes, and especially on *PAX5/ETV6*, as new example of aberrant transcription factor in BCP-ALL.

Abnormalities of the short arm of chromosome 9 (9p) have been described in approximately 10% of cases, in childhood ALL, with a higher incidence in T-ALL.[3] The majority of 9p abnormalities results in a deletion of this arm, usually including the cell cycle regulatory genes p14, p15 and p16. Deletions of p16 have been detected by molecular analysis and FISH in approximately 80% of childhood T-ALL and 20% of common-pre-B ALL. Abnormalities of 9p have been associated to an adverse risk factor in B-lineage, but not T-lineage ALL, except in the case of homozygous form of this deletion.

2. B cell development

The hematopoietic system constantly generates a large number of specialized cell types from pluripotent hematopoietic stem cells (PHSCs), which have self-renewal potential and give rise to different progenitors with a more restricted differentiation capacity.[10,11] Each developmental step is a binary commitment switch, so at each of these steps, the cell can take alternative directions.

One of the earliest differentiated precursors is the multipotential progenitor (MPP), which is at the bifurcation between the myeloid and lymphoid lineages: MPPs can differentiate into common myeloid progenitors (CMPs) or into the recently identified early lymphoid progenitors (ELPs). The latter starts to express recombination-activating gene 1 (*Rag1*) and *Rag2* and initiate rearrangement at the immunoglobulin heavy chain (IgH) locus. ELPs can further differentiate into thymic precursors of the T-cell lineage (early T-cell-lineage progenitors, ETPs) or into bone-marrow common lymphoid progenitors (CLPs), which are lymphoid restricted and can generate B cells, T cells, dendritic cells (DCs) and natural killer (NK) cells.[12]

Expression of the B-cell marker B220 by a subset of CLPs (known as pro-B cells or CLP-2s) coincides with their entry into the B-cell-differentiation pathway. The next step can be identified by expression of CD19 and completion of IgH diversity (DH)-to-joining (JH) gene-segment rearrangement by pre-BI cells. The IgH locus then continues to rearrange its variable (V)-region gene segments until productive VH–DJH alleles are generated in large pre-BII cells. These cells cease to express *Rag1* and *Rag2*, and they display the product of the rearranged IgH gene at the cell surface; there, it assembles with the surrogate immunoglobulin light chains (IgLs), VpreB and λ5, together with the signaling molecules Igα (which is encoded by the *MB-1* gene) and Igβ (which is encoded by the *B29* gene) to form the pre-B-cell receptor (pre-BCR). Expression of the pre-BCR is a crucial checkpoint in early B-cell development, at which the functionality of the heavy chain is monitored. Signaling through the pre-BCR allows for allelic exclusion of the IgH locus and stimulates a burst of proliferative clonal expansion of large pre-BII cells, which is followed by re-expression of RAGs and rearrangement at the IgL locus in small pre-BII cells.[12]

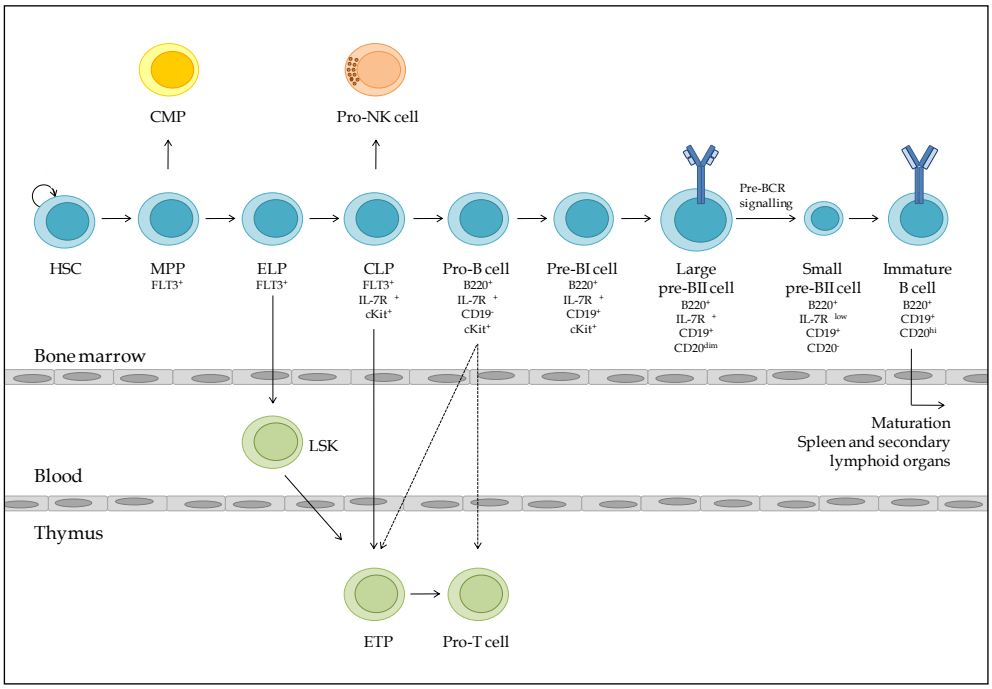


Figure 3. Schematic diagram of B cell development.[12]

During normal development, appearance of the assembled BCR at the cell surface defines the immature B-cell stage, at which cells are tested by a secondary immunoglobulin gene rearrangement, which is known as receptor editing, otherwise they are eliminated or inactivated by apoptosis or anergy, respectively.

After they have successfully passed this examination, immature B cells leave the bone marrow and enter the periphery, where they seed the spleen and further differentiate through three transitional stages T1, T2 and T3, on the basis of expression of various cell-surface markers. Immature splenic B cells can be distinguished from their mature counterparts by their cell surface marker expression, their short half-life (2–4 days) and their sensitivity to apoptosis induced by antibodies specific for IgM. Transitional B cells can be negatively selected in the periphery, which is therefore a crucial checkpoint for the generation of mature B cells. The tumor necrosis factor family member B cell activating factor (BAFF; also known as BLYS) and its receptor have a crucial role in regulating the transition of immature B cells to mature B cells. About 10% of mature splenic B cells are marginal zone B cells. Marginal zone B cells are found only in the spleen. They are strategically positioned at the blood–lymphoid tissue interface, where they can initiate a fast and vigorous antibody response to blood-borne pathogens.

Transitional, follicular and even memory B cells can give rise to marginal zone B cells and/or be recruited into this compartment. In mice, most mature B cells are follicular B cells. These cells are mainly responsible for generating humoral immune responses to protein antigens. With the help of T cells, they form germinal centers. Germinal center B cells proliferate rapidly; undergo somatic hypermutation of their immunoglobulin variable gene segments and isotype-switch recombination of immunoglobulin genes. At about day ten after immunization, the germinal center reaction reaches its peak. Subsequently, germinal centers slowly vanish, and memory B cells and effector plasma cells are generated. The mechanisms and factors that guide some germinal center B cells to become memory B cells and others to become plasma cells are not yet clear.[12]

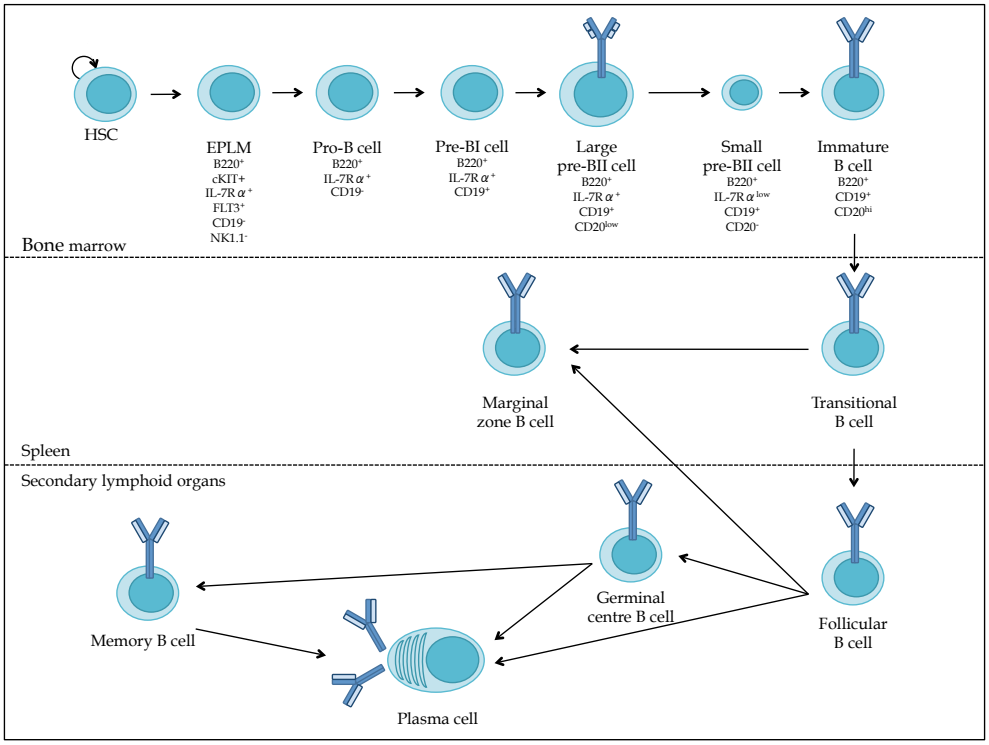


Figure 4. Representation of B cell development from HSC to mature B cell.

2.1. Transcription factors in B cell development

B lymphopoiesis has emerged as one of the leading models for studies of lineage specification (induction of a lineage specific gene expression program) and commitment (repression of alternative gene expression programs), and transcription factors play a

central role in this process; indeed, B cell development, maturation and function are coordinated by a battery of transcription factors and signal transduction molecules that regulate the sequential execution of the different steps.[11,13]

Developmental control of early B lymphopoiesis is exerted by a regulatory network of key transcription factors that include PU.1 (an ets-family member), Ikaros, Bcl11a (a zinc finger transcription factor), E2A (a helixloop-helix protein), EBF (early B cell factor) and PAX5.[14]

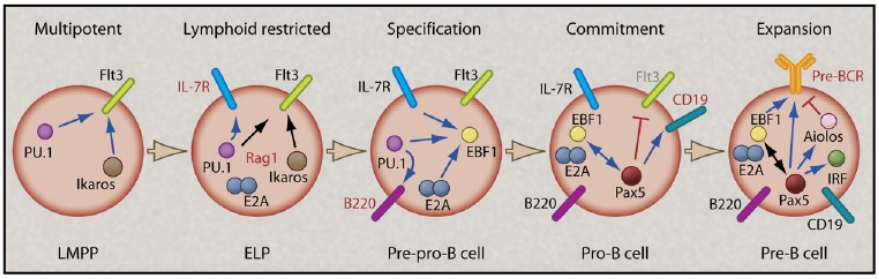


Figure 5. Multistep model of B cell development and transcription factors involved.[13]

As reviewed by Hagman and Lukin 2005,[15] the Ets family transcription factor **PU.1** (encoded by *sfpi-1* in mice) controls the inception of B cell development, in a dose-dependent manner, with low versus high expression of PU.1 favoring the generation of B cells versus macrophages, respectively. PU.1 is essential for the generation of B cells, but only facilitates this developmental outcome when expressed at reduced concentrations. Indeed, when expressed at high concentrations in early progenitors, PU.1 might block B cell specific gene expression by inhibiting histone acetylation and by recruiting DNA methyl transferases 3a and 3b to methylate gene targets. Conversely, at low concentrations PU.1 might prime modifications of the chromatin to facilitate transcriptional activation as cell differentiation proceeds.

Whereas PU.1 appears to be crucial for both myeloid and lymphoid progenitors, **Ikaros** controls the development of lymphoid progenitors. In the absence of **Bcl11a**, early B lymphopoiesis is blocked at the CLP stage before the pre-pro-B stage.[14]

But in particular, three transcription factors have been found to be essential for the differentiation of CLPs into specified pro-B cells: transcription factor E2A, early B cell factor (EBF; also known as OLF1) and paired box protein 5 (PAX5; also known as BSAP). Absence of any one of these factors leads to an early block in B-cell development at the

pro-B cell or pre-B cell stage. These three factors seem to work in collaboration, and together, they form a master control switch for engaging B cell differentiation.[11]

E2A and EBF are considered primary B cell fate determinants and coordinately activate the expression of B cell specific genes (e.g. both can initiate the remodeling of MB-1 promoter chromatin).[11,14]

The ***Tcfe2a*** gene encodes the basic helix–loop–helix (bHLH) proteins E12 and E47, also known as E2A proteins; they bind E box elements (50-CANNTG-30) within regulatory modules of many genes expressed in early (Ig enhancers and mb-1) and late (AID) stage B cells. **Ebf1** activates target genes through its carboxy-terminal domain, but other sequences in EBF contribute to transcriptional activation as well, and its expression correlates to expression of many components of the pre-BCR and BCR. In the absence of E2A or EBF, B cell development is blocked at early progenitor stages of development; however, EBF can activate the B cell lineage program in absence of E2A or PU.1, but conversely, in the absence of EBF, B cell development was not rescued by enforced expression of PAX5.[16]

The ability of EBF to mediate activation of the B cell program suggests that it has properties of a 'pioneer' factor (i.e. a protein capable of initiating the activation of transcriptionally quiescent genes).[15]

However, the mere activation of the B lymphocyte transcription program is not sufficient to commit B cell progenitors to the B lymphoid lineage in the absence of the paired domain protein Pax5.[14]

2.2. Checkpoints during normal B cell development

During early B cell development, there are multiple checkpoints that maintain the quality of B cells generated at every stage. These checkpoints ensure that early B cells do not go astray and undergo leukemic transformation. A loss of checkpoint controls by normal differentiation or by mutagenesis of important signaling molecules is deleterious to a B cell and increases the probability of leukemia initiation.

The first checkpoint comes into play when multi-lineage progenitors commit to the B lineage.[17] ***Pax5*** is the guardian at this checkpoint and it carries out the above function by repressing B lineage 'inappropriate' genes and simultaneously activating B lineage specific genes.[18]

Even after commitment to B lineage, there are multiple checkpoints that check the quality of B cells produced at every stage of differentiation. These checkpoints are important because they ensure that the B cell repertoire is only made up of cells with functional immunoglobulin rearrangements that are specific to foreign antigens and do not react to self.

The **μ heavy chain** of the pre-B cell receptor represents the second checkpoint.[19] This checkpoint filters out and eliminates cells that have not successfully rearranged their heavy chain, thereby preventing the accumulation of deleterious cells in the B cell repertoire. At this checkpoint two opposing processes occur: one serves to eliminate pre-B cells with non-functional V(D)J rearrangements (negative selection) and the other allows the survival and proliferation of pre-B cells carrying functional heavy chain rearrangements. The interplay between BACH2 and BCL6 maintains the balance between negative selection and survival respectively. A disruption of the balance between the two processes leads to leukemic transformation of the pre-B cell.[20,21].

The third and final checkpoint comes into the picture following the completion of negative selection. This checkpoint is governed by one of the signaling components of pre-BCR, **Blnk**. Blnk induces differentiation of pre-B cells that have survived and successfully passed the pre-BCR checkpoint control. Loss of *Blnk* leads to leukemic transformation of pre-B cells by blocking their differentiation from pre-BI to pre-BII cells.[22,23]

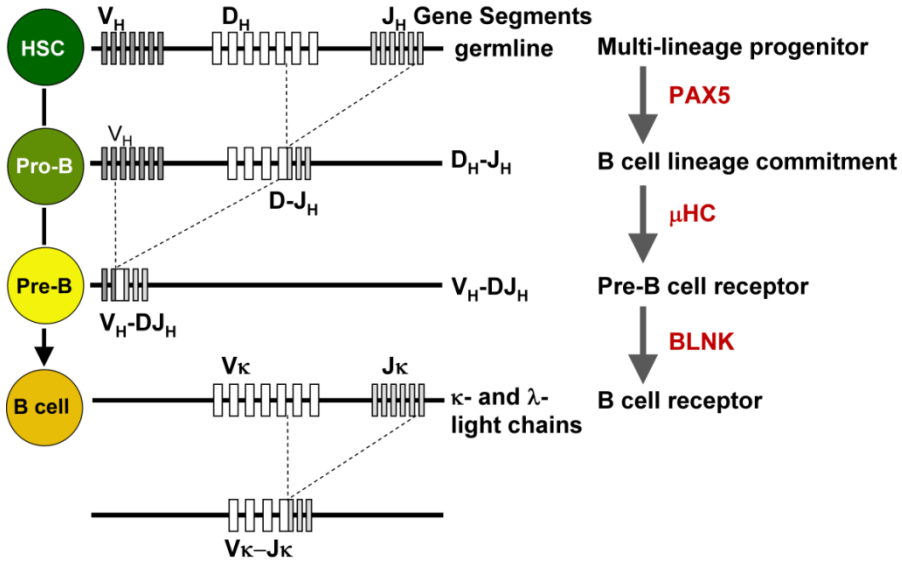


Figure 6. Checkpoints during B cell development. Image adapted from Swaminathan (2013).[24]

In summary, disruption of checkpoints at any stage of early B cell development predisposes the cell to leukemic transformation. Therefore, study of checkpoint signaling mechanisms are crucial to understand the process of leukemogenesis.

2.3. Signaling pathways in precursor B cells

During B cell development, the cell fate is controlled by two different signaling pathways: the one downstream the IL7R, and the one downstream the pre-BCR. They have a central role in controlling the survival, proliferation and differentiation of B cells.

The **IL7R** is composed of the common γ chain (IL2Rγ) and IL-7Rα (IL7Rα); the latter is also a component of the thymic-stromal-derived lymphopoietin receptor (TSLPR). The increased expression of the α chain of the IL7R demarks the passage from CLPs to LMMPs.

The absence of the IL7 signal in mice that have a targeted deletion of either the gene encoding IL7 or the genes encoding components of IL7R results in an arrest of B cell development at the pro-B cell stage and impaired V to DJ recombination. In these mice, only a few B cells mature and populate the periphery, supporting the hypothesis that IL7 is required for proper cell proliferation during early B cell development. Enhanced IL7R signaling results in the proliferation of pro- and pre-B cells in the bone marrow, as well as the migration of these populations to the spleen, lymph nodes and blood. Of note, most of the mice that have enhanced IL7R signaling develop pro- and pre-B-cell tumors or bipotent B-cell and myeloid-cell tumors.[25]

The IL7R pathway begins by the binding of the IL7 to its cognate receptor, which basally is completely inactive. Cytokine binding induces a conformational change that induces receptor dimerization. As a consequence, JAK3, pre-bound to the cytoplasmic region of the receptor, become activated, leading to cross-phosphorylation and receptor tyrosine phosphorylation, thus creating binding sites for STAT5. Phosphorylation leads to STAT5 homodimerization, translocation to and retention in the nucleus, where it acts as a transcription factor. Additionally, STAT5 can be activated by other kinases, such as SRC family kinase.[26]

Physiologically, signaling by cytokine receptors is transient. The termination of the signal is due to:

a) internalization and degradation of the receptor by the lysosome and proteasome pathways, which require ubiquitination of intracellular lysines mediated by CBL proteins;[27]
b) phosphatase recruitment to the receptor, thereby dephosphorylating JAK;[28]
c) recruitment of negative regulators, such as SOCS proteins;[29]
d) interactions of STAT5 proteins with protein inhibitors (PIAS), which induce sumoylation of several transcription factors and form signaling platforms localized to subnuclear regions.[26,30]

This pathway has a central role in the signaling of cytokines by regulating cell proliferation, survival, differentiation and immune response. Constitutively active cytokine receptors lead to permanent signaling. Thus, constitutively active STAT5 in the nucleus overcomes negative regulatory mechanisms, eventually leading to oncogenesis.[26]

Several experiments indicate that there is intensive crosstalk between the signaling pathways that are triggered by the pre-BCR and IL7R. Expression of the pre-BCR by pro-B cells upregulates IL7R expression on the cell surface, thereby increasing the responsiveness of these cells to IL7. This results in the selective expansion of pre-BCR positive cells in conditions in which IL7 concentration is low. At the same time, however, the pre-BCR also induces the differentiation of pre-B cells into a state in which they have reduced IL7 responsiveness, which probably limits the expansion of pre-B cells during B-cell development *in vivo.*[25,31]

The **pre-BCR** is formed by the coupling of μ heavy chain with surrogate light chain components VpreB (*Vpreb1*) and λ5 (*Igll1*). The pre-BCR is present on the surface of the pre-BI cells. The extracellular part of the pre-BCR is linked to two signaling chains, Igα and Igβ, which contain an immunoreceptor tyrosyne-based activation motifs (ITAM). The pre-BCR serves as an immunological synapse where the Igα and Igβ signaling chains serve as docking sites to assemble and activate the components of the pre-BCR signaling cascade.

The pre-BCR has a dual function: it is responsible for both proliferation and survival, and growth arrest and differentiation. Therefore, it serves as a checkpoint in early B cell development, it decides whether the cell should survive for further differentiation or undergo apoptosis. If the pre-BCR does not carry a functional heavy chain

rearrangement, it is eliminated by apoptosis. This phenomenon is termed as negative selection. If the rearrangement is functional, the cell survives and the pre-BCR signals for proliferation and differentiation.[32]

One of the main proteins that docks at the immunological synapse and mediates the dual function of the pre-BCR is SYK. It induces proliferation and survival by activating a number of downstream survival pathways, like the PI3K signaling cascade.[33-35] SYK also subsequently performs the opposing function of growth arrest and differentiation by phosphorylating and activating BLNK.[22,36,37] Phosphorylated BLNK recruits BTK which in turn activated PLCγ2.[36,38-40] PLCγ2 generates diacylglycerol and IP3 by hydrolysis of PI(4,5)P2, which serve as second messengers for activation of PKC and Ca^{2+} release fro cytoplasmic stores. BLNK triggers the process of κ light chain rearrangement by activating its downstream effectors.[41,42] Moreover, it blocks the proliferation signal emanating from SYK and thus shifts the cell from proliferation to differentiation mode. The exact mechanism by which this switch occurs is still unknown. BLNK inactivates the PI3K signaling pathway and thereby activates FoxO factors. The latter transcriptionally activate the RAG enzymes and allow for immunoglobulin light chain rearrangement.[25] *Blnk*$^{-/-}$pre-B cells are arrested at the pre-BI stage of development and are constantly proliferating.[22]

The pre-BCR has also been implicating in inducing the mutator enzyme AICDA through *Btk.* [43] This implies that the pre-BCR can mimic the signaling of a mature BCR in the context of infection or inflammation. The role played by inflammations and/or infections in inducing leukemogenesis at the pre-BI cell transition is still poorly understood.

Taken together, the signaling molecules at the pre-BCR checkpoint maintain the balance between negative selection of non-functional B cells (apoptosis) and leukemic transformation.

3. PAX5: the guardian of B cell identity and function

The *PAX5* gene is a member of the paired box (*PAX*) gene family of transcription factors and it is essential for B lymphoid lineage commitment.[44-47] PAX5 fulfills a unique function by controlling the identity of B lymphocytes throughout B cell development from the pro-B to the mature B cell stage.[10,11]

3.1. *PAX5* gene

PAX5 or the B cell specific activator protein (BSAP) was independently discovered as a DNA-binding protein with the same DNA sequence specificity as the sea urchin transcription factor TSAP. Biochemical purification and cDNA cloning showed that BSAP is encoded by the *PAX5*geneand is expressed in B lymphocytes, the developing CNS, and adult testis.[18]

The human *PAX5* gene is located on chromosome 9p13 and it is organized in 10 exons and encodes for a transcript of 8536bp, giving rise to a coding sequence of 1176 bp which encodes for a protein of 391 residues, completely homologous to *Mus musculus PAX5* gene.[48]

3.2. PAX5 protein structure

PAX5 is a homeodomain protein, which is a member of a class of transcription factors that contains a DNA-binding domain with homology to the *Drosophila melanogaster* homeodomain regulatory proteins. This DNA-binding domain contains a helix–turn–helix motif, which binds to a distinct half-site of the PAX5 recognition sequence in adjacent major grooves of the DNA helix.[18]

The defining feature of the PAX protein family is the conserved 'paired' domain, which functions as a bipartite DNA-binding region consisting of an N- and a C-terminal subdomain. The bipartite nature of the paired domain is responsible for its degenerate consensus recognition sequence, as each half-site independently contributes to the overall affinity of a given binding site.[49]

The transcriptional activity of PAX5 at other gene regulatory elements is, however, determined by the interaction of distinct partner proteins with central and C-terminal protein interaction motifs of PAX5. The partial homeodomain of PAX5 associates with

the TATA-binding protein of the basal transcription machinery, while a potent C-terminal transactivation domain regulates gene transcription most likely by interacting with histone acetyltransferases (HAT) such as the coactivator CBP or SAGA complex. In contrast, corepressors of the Groucho protein family, which are part of a larger histone deacetylase (HDAC) complex, are able to convert PAX5 from a transcriptional activator to a repressor by binding to a conserved octapeptide motif of PAX5.[18]

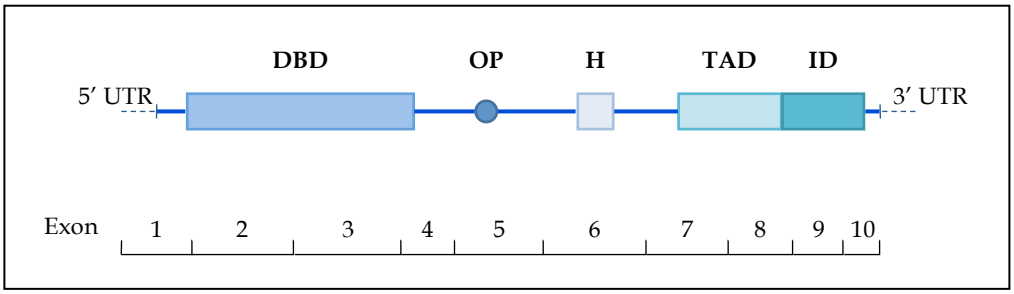


Figure 7. PAX5 structure, showing functional domains and corresponding exons.[12]

3.3. PAX5 role in B cell development

Pax5 is a transcription factor essential for B cell development that localizes in the nucleus.[11,18]

The *Pax5*$^{-/-}$ mouse model had been fundamental to study its role in B cell development and allowed to demonstrate that Pax5 fulfills a dual role in B lineage commitment by activating B cell specific genes while simultaneously repressing lineage-inappropriate genes.[45]

Mutations in the *Pax5* gene leads to a complete block in B-cell differentiation, which is immediately downstream of the block that is seen in the absence of E2A or EBF: in *Pax5*$^{-/-}$ B cells, expression of E2A and EBF is normal, and rearrangement at the IgH locus has already been initiated.[11] Although DJ gene segments are rearranged normally in these cells, rearrangement of VH gene segments is severely impaired. Pax5 is required not only for rearrangement of VH gene segments and for expression of genes that are required for progression to the pre-B cell stage but also for commitment to and maintenance of the B cell differentiation pathway.

Remarkably, *Pax5*$^{-/-}$ pre-BI cells posses an extraordinary developmental plasticity showing hematopoietic stem cell features such as multipotency and self renew. Both *in*

vitro and *in vivo*, in the absence of Pax5, pre-BI cells show an amazing degree of plasticity and can differentiate *in vitro*, after stimulation with the appropriate cytokines, into macrophages, osteoclasts, DCs, granulocytes or NK cells; while *in vivo* they develop into all the lineages as described *in vitro* but even into T cells and erythrocytes.[50]

In addition, conditional *Pax5* deletion in mice allows mature B cells from peripheral lymphoid organs to dedifferentiate *in vivo* back to early uncommitted progenitors in the bone marrow, which rescued T lymphopoiesis in the thymus of T cell deficient mice. These B cell derived T lymphocytes carried not only immunoglobulin heavy- and light-chain gene rearrangements but also participated as functional T cells in immune reactions.[51]

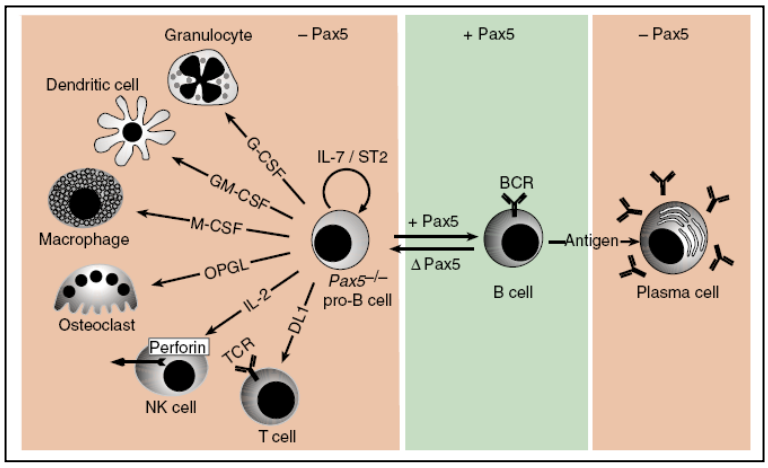


Figure 8. B cell lineage commitment by Pax5. The uncommitted *Pax5*$^{-/-}$pro-B cells are able to differentiate into several hematopoietic cell types either *in vitro* in the presence of the indicated cytokines or *in vivo* after transplantation into recipient mice. Conditional *Pax5* deletion (ΔPax5) results in retrodifferentiation of B lymphocytes to an uncommitted progenitor cell stage.[51]

Moreover, conditional *Pax5* deletion in mature B cells leads to the development of aggressive lymphomas in mice, suggesting that the complete loss of *Pax5* in late B cells could initiate lymphoma development and uncovered an extraordinary plasticity of mature peripheral B cells despite their advanced differentiation stage.[51]

In conclusion, *Pax5* loss o 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.[51]

3.4. PAX5 target genes

The transcription factor PAX5 represses B lineage inappropriate genes and activates B cell specific genes in B lymphocytes, indeed it functions both as a transcriptional activator and as a repressor on different target genes.[52,53]

On one hand, Pax5 represses several genes involved in many biological activities including receptor signaling, cell adhesion, migration, transcriptional control, and cellular metabolism at B cell commitment. Indeed, *Pax5* is responsible for the repression of genes leading to other lineages determination, such as *Flt3* and *Sca1* (progenitor cells);[54] *Csf1r*, *Ramp1*, *Cd33* (myeloid cells);[44] *Notch1* and *Lck* (T cells);[46] *Sdc1* and *Prdm1* (plasma cells).[53] These lineage inappropriate genes require continuous Pax5 activity for their repression, as they are reactivated in committed pro-B cells and mature B cells following conditional Pax5 deletion.[18,53]

On the other hand, the Pax5 activated genes encode for key regulatory and structural proteins involved in B cell signaling, adhesion, migration, antigen presentation, and germinal center B cell formation, thus revealing a complex regulatory network that is activated by Pax5 to control B cell development and function. In particular, Pax5 has been implicated in controlling the signaling from the pre-BCR on the cell surface to transcription in the nucleus at multiple levels, regulating numerous genes such as *Cd19*, *Cd79a*/Iga, *Blnk*, *Vpreb3*, *Igk* and *Bach2*.[20,55-59]

Altogether, target genes that are activated by Pax5 encode for essential components of the pre-B- and B-cell receptor (pre-BCR and BCR, respectively) signaling pathways; and when Pax5 acts as a transcriptional repressor, its function is to limit lineage choices that differ from B-cells.[14] Continuous repression by PAX5 is known to be required, because reactivation of these repressed targets is observed following conditional deletion of PAX5 in pro-B and mature B cells. Surprisingly, even at the transition to the plasma cell stage, when PAX5 is physiologically lost during terminal differentiation, repressed genes are reactivated and contribute to the plasma cell transcriptional program.[14]

3.5. Oncogenic role of *PAX5* in hematological malignancies

In hematological malignancies, *PAX5* has been first described as involved in lymphomas. In diffuse large B cell lymphomas (DLBCL), the exon 1B of *PAX5* is target of

misdirected class-switch recombination and somatic hypermutations (SHM), which are essential for affinity maturation of immunoglobulins in germinal center B cells, but potentially can be misdirected to generate oncogenic mutations or chromosomal translocations involved in lymphomagenesis.[18,60] In non-Hodgkin lymphoma, the t(9;14) translocation brings the complete coding sequences of *PAX5* under the control of strong enhancers or promoters from the IGH locus, leading to increased expression of the translocated PAX5 allele. The regulatory IGH-PAX5 translocation is likely to contribute to tumor formation by perturbing the PAX5-dependent gene expression program owing to increased PAX5 transcription in B cells or failed PAX5 repression at the onset of plasma cell differentiation.[61,62]

More recently, a new genetic lesion has been identified on 9p, affecting the PAX5 gene in about 30% of childhood BCP-ALL (B Cell Precursor ALL) cases,[63] which can be considered one of the most common alterations. In addition to deletion, the *PAX5* gene has been reported as a recurrent target of mutation (about 7%) and translocation (2-3%), both in adult and in childhood BCP-ALL cases with similar incidence.[63,64]

Deletion is the most frequent *PAX5* aberrancy, occurring in about 25% of patients. Three types of deletions have been found:

a) wide-range deletions, involving the full length *PAX5* gene and flanking genes, or even the whole short arm of chromosome 9, extended over *CDKN2A-2B*;
b) focal deletions involving a subset of *PAX5* exons extending to its 3' region, thus leading to a prematurely truncated gene/protein;
c) focal deletions involving only a subset of internal *PAX5* exons, determining different PAX5 isoforms lacking functional domains, such as the DNA binding domain, the octamer or the transcriptional regulatory domain.

In childhood ALL, deletions have been reported as monoallelic in 53/192 of the cases (27.6%), and among them 25 were focal (13%), while the complete PAX5 was lost in 28 cases (14.6%); only 3/192 (1.6%) were biallelic.[63] In addition, in adult BCP-ALL, PAX5 has been found as target of exclusively monoallelic deletions in 27/117 cases (23.1%).[64] Both in adult and childhood cases, *PAX5* deletions seem to be secondary events, since they are frequently associated with other lesions, such as ETV6/AML1,[63] BCR/ABL1 or TCF3/PBX1.[64-67]

Similarly, aberrant splicing variants have been described as alternative mechanisms of deletion, both in adults and in children. Different isoforms have been described, the most frequent lacking either only exon 2, causing frame shift and premature stop; or exon 5, corresponding to the octamer domain; or both exons 8 and 9 which encode the transactivation domain. In leukemic blast cells, these variants are more abundant compared to full length wild type PAX5, and are predicted to code for less functional or even completely non-functional PAX5 proteins.[68]

Point mutations have been found in about 7% of both adult and pediatric cases. In childhood leukemia cases, point mutations have been reported as hemizygous in 14/192 of the cases (7.3%) while only 1 case as homozygous (1/192, 0.5%).[63] A recent study reported a higher rate of PAX5 point mutations compared to the previous investigations. Point mutations were found in 9/50 (18%) adults and in 14/50 (28%) children, showing novel sites of mutations scattered along all the exons.[68]

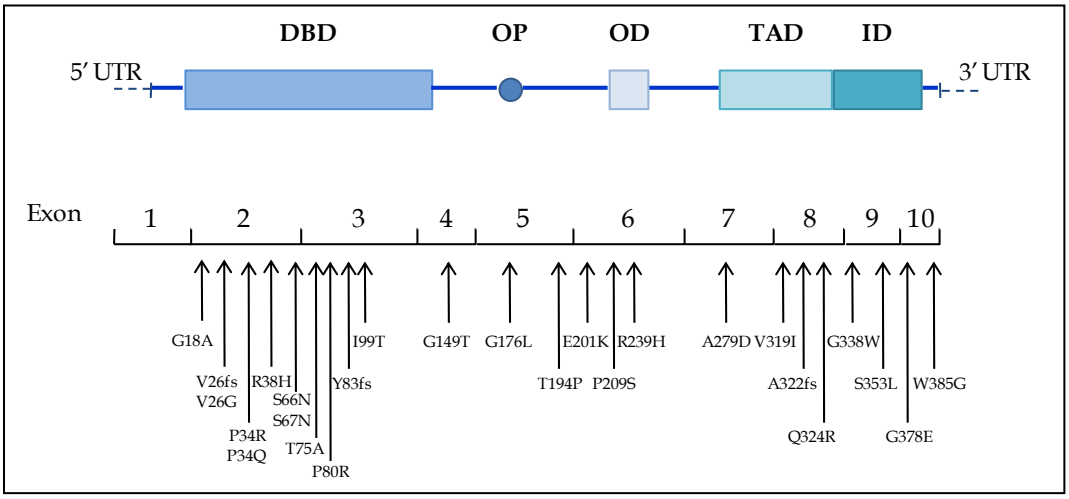


Figure 9. PAX5 point mutations.[12]

Modeling studies using the PAX5 crystal structure suggested that point mutation should either impair DNA binding, or alter transcriptional regulation, or cause frame shift, splice site, or missense mutations. A single case with an exon 1B frame shift mutation resulted in a prematurely truncated ten-residue peptide.

Altogether, somatically acquired point mutations are hemizygous; they result in reduced expression of *PAX5* mRNA, and lost or altered DNA-binding or transcriptional

regulatory function. Therefore they generate hypomorphic *PAX5* alleles with reduced gene function that could lead to haploinsufficiency.[63,69]

Very recently, a new heterozygous germline variant, c.547G>A (p.Gly183Ser), has been described in three different families.[70,71] This point mutation affects the octapeptide domain of *PAX5* and was found to segregate with disease in three unrelated kindreds (one of Puerto Rican, one of Afro-American and the other of Ashkenazi Jewish origin) with autosomal dominant B-ALL. Leukemic cells from all affected individuals in all the families exhibited 9p deletion, with loss of heterozygousity and retention of the mutant *PAX5* allele at 9p13. Functional and gene expression analysis of the *PAX5* mutation demonstrated that it had a significantly reduced transcriptional activity, thus extending the role of *PAX5* alterations in the pathogenesis of BCP-ALL and implicating *PAX5* in a new syndrome of susceptibility to pre-B neoplasia.[70,71]

Amplifications of the *PAX5* gene are less common and include:

a) partial amplification, which covers a subset of internal exons (in particular from exon 2 to 5/6) (our unpublished data);
b) complete amplification, which targets the full sequence of the gene.[63,64] It has been proposed that focal amplification of *PAX5*exons 2–5 is predicted to abolish expression of normal *PAX5* from the amplified allele, although further research is required to clarify the importance of this rare genetic lesion.

Epigenetic mechanisms have been investigated to explain lower *PAX5* gene expression, in absence of classical rearrangements (e.g. point mutations and deletions).[63] However, although high-levels of methylation of the *PAX5*promoter were detected in T-ALL, only a minimal methylation level was observed in BCP-ALL, independently by *PAX5*mutational status.[72]

In ALL, PAX5 is involved in a growing number of **chromosomal translocations** with several partner genes, resulting in *in frame* fusion genes, with an estimated frequency of 2-3% in pediatric BCP-ALL.[63,73]

Since the first case encoding for PAX5/ETV6 was identified in 2001, [74] as many as 21 known partner genes have been identified in several chromosomes, although chromosomes 7, 12 and 20 are the most frequently involved, with different genes on each chromosome.

Interestingly, PAX5 genetic lesions are typically associated to dicentric chromosomes, although the breakpoint on the short arm of chromosome 9 is very heterogeneous.[69] PAX5 was rearranged in 54 out of 110 pediatric patients carrying a dicentric chromosome: 26/38 of the cases carrying the dic(9;12)(p11~13;p13), 24/59 positive for the dic(9;20)(p11~13;q11) and in 4/13 of the cases harboring the dic(7;9)(p11;p11~13).[69] Dicentric chromosomes can coexist with established chromosomal changes, for example, dic(7;9) is found in association with t(9;22)(q34;q11) (*BCR*/*ABL1* fusion), and dic(9;12) occurs with t(12;21)(p13;q22) (*ETV6*/*RUNX1* fusion), suggesting that these events may cooperate. A more recent study confirmed the involvement of the *PAX5* gene in dicentric chromosome events in adult and childhood BCP-ALL.[75] In this study, 40 patients (14 adult and 26 pediatric) harbored a dicentric chromosome involving chromosome 9 with various different partners (such as, 7, 8, 12, 15, 16, 17, 20) and as much as 90% of the cases involved PAX5 alterations.

The various fusion genes can be classified in different sub-groups according to the molecular function of the *PAX5* partner gene, which can encode either for a transcription factor (e.g. *ETV6/TEL*, *PML*, *FOXP1*, *ZNF521/EVI3*, *BRD1*, *DACH1*), for proteins related to transcription regulation (*HIPK1*, *NCOR1*), structural proteins (*ELN*, *POM121*, *LOC39027*, *KIF3B*), kinases (*JAK2*, *TAOK1*), carriers of molecules (*SLCO1B3*) or co-activator proteins (*ASXL1*). The implicated genes encode less frequently for molecules of unknown function (e.g. *C20orf112*, *AUTS2*, *GOLGA6*). Among this plethora of fusion genes, *PAX5/ETV6* originating from the translocation t(9;12) or dic(9;12) is the most common.[12]

A common feature of PAX5 translocations is that they result in the fusion of the 5' N-terminal DNA-binding domain of PAX5 (PAX5-DBD) with the 3' C-terminal sequences of the partner gene, whose domains therefore substitute PAX5 regulatory domains. The breakpoints are mostly located after exon 5 of *PAX5,* or after exon 4, at the end of the DBD region. Cases with a more distal PAX5 breakpoint are less common and retain additional domains, including the highly conserved octamer domain. As a further feature, *PAX5* translocations are unbalanced,[69] and therefore the reciprocal fusion gene (5' partner gene to 3' *PAX5*) is not preserved, with the sole known exception of the translocation, t(9;9)(p13;p24), encoding for *PAX5/JAK2.*[73] *ETV6/TEL*, *ELN* and *FOXP1*

are recurrent partner genes, independent of patient age, as well as translocation events giving rise to PAX5-truncated isoforms.

Furthermore, even though in some cases the translocation events lead only to a truncated form of PAX5, it can be speculated that the role of PAX5 in this setting can be different from deletions or point mutations. The fusion protein could retain the functional activities of the partner genes and therefore it may alter the molecular function of the normal counterparts, affecting both the partner gene and wild type PAX5 functions. Indeed, many of the *PAX5* fusion genes have a dominant negative role on normal PAX5; in fact, in leukemic blasts, expression analysis revealed that its transcriptional targets (e.g. *EBF1*, *FLT3*, *ATP9A*, *ALDH1A1*) are repressed in dicentric cases, indicating a reduced activity of wild type PAX5.[69] Despite these data, CD19 and CD79a expression (direct PAX5 targets) did not correlate with PAX5 altered status.[63]

Comprehensively, several fundamental questions on PAX5 involvement in leukemogenesis are still unresolved. Two points remain to be understood:

a) whether these genetic alterations are responsible for the disruption of PAX5 functions, generating a haploinsufficiency setting;
b) whether these lesions are more generally responsible for the B cell development block, through deregulation of PAX5 control, which is crucial for normal B lymphopoiesis.

A more dominant role for fusion genes cannot be excluded, in which a gain of function could lead to cell transformation. *In vitro* and *in vivo* studies will be helpful to better elucidate the role of fusion genes.

4. ETV6: a versatile player in leukemogenesis

The *TEL/ETV6* gene belongs to the *Ets*-transcription factor family,[76] and it is a sequence-specific transcriptional repressor of *Ets*-binding site-driven transcription [77]. It was first named as *Translocation Ets Leukemia* Gene and later renamed to ETV6 (*ets* variant gene 6, where *ets* is derived from E-26 transforming specific).[78]

The *ETV6* gene is required for hematopoiesis within the bone marrow[79-81] and it is frequently involved in several translocations occurring in hematological malignancies, both of myeloid and lymphoid origins, leading to many different fusion genes.[76,78]

4.1. Cloning of *ETV6*

The *ETV6* gene, located at chromosomal band 12p13, was first isolated by virtue of its recombination with the gene encoding the β chain of the platelet derived growth factor receptor (PDGFRβ), following the t(5;12)(q33;p13) of chronic myelomonocytic leukemia (CMML).[82]

The *ETV6* gene is organized in 8 exons and the transcript length is 5989 bp, which encodes for a protein of 452 residues. *ETV6* contains two alternative translational start codons (position 1 and position 43) leading to the expression of two isoforms of *ETV6*, originating two proteins of 57 kda and 50 kda, respectively.[76]

4.2. ETV6 protein structure

Normal *ETV6* products are nuclear phosphor-proteins that exist as several species. *Ets*-transcription factor family members are conserved among mammalian species and share 88 aminoacids (aa) region of homology: the ETS domain, located at the carboxy terminus, responsible for nuclear localization and sequence specific DNA-binding activity, as well as protein–protein interaction. Indeed, ETS proteins bind to a GGAA/T core motif flanked by 5 ± 8 nucleotides of surrounding sequences that contribute the specificity of individual family members. Transcriptional activation by ETS proteins usually involves interaction with other transcription factors to achieve transcriptional regulation and/or target gene specificity.[76]

The other evolutionarily conserved domain in the 652aa of ETV6 is the N-terminally located pointed or sterile alpha motif (SAM) domain whose 3D structure has recently

been elucidated. This domain is also called HLH domain and is even more highly conserved in evolution and found in many *ets* family member. It has been found in yeast proteins and has been shown to be involved in homo- and hetero-dimerization in transcription factors and in signal transducing proteins (e.g. of the MAPK pathway).[78]

Nucleotide sequence analysis of a murine Etv6 cDNA predicted a larger Etv6 product in mouse (485 aa) than in human (452 aa), with a high similarity between the two species. As expected, the identity between the two proteins is particularly high within the pointed (97%) and the ETS domain (100%). The identity is lower in the central part of the protein and, since transcriptional regulatory domains are known to be weakly conserved between species, this region might be responsible for the transcriptional regulation by *Etv6*, similarly to other ETS proteins members.[76]

The murine Etv6 protein possesses a 37 aa C-terminal extension compared to human ETV6. It is acidic both in human and murine ETV6 proteins and could possess transcriptional regulation properties. Alignment of the ETS domain of human ETV6 with consensus ETS domain sequences shows several specific features that are conserved in murine Etv6.[76]

4.3. ETV6 in hematopoietic development

The *Etv6* gene is widely expressed throughout embryonic development and in the adult. *Etv6*$^{-/-}$ mice are embryonic lethal and die between E10.5-11.5 with defective yolk salc angiogenesis and intra-embryonic apoptosis of mesenchymal and neural cells.[79] Two-thirds of *Etv6*-deficient yolk sacs at E9.5 lack vitelline vessels yet possess capillaries, indicative of normal vasculogenesis. Vitelline vessels regress by E10.5 in the remaining *Etv6*$^{-/-}$ yolk sacs. Hematopoiesis at the yolk sac stage, however, appears unaffected in *Etv6*$^{-/-}$ embryos. *Etv6* is required for maintenance of the developing vascular network in the yolk sac and for survival of selected neural and mesenchymal cells within the embryo proper.

The same group addressed the question of a possible requirement of ETV6 in adult hematopoiesis, generating mouse chimeras with *Etv6*$^{-/-}$ ES cells.[80] Although not required for the intrinsic proliferation and/or differentiation of adult type hematopoietic lineages in the yolk sac and fetal liver, *Etv6* is essential for the establishment of hematopoiesis of all lineages in the bone marrow. This defect is

manifest within the first week of postnatal life, suggesting a critical role for *Etv6* in the normal transition of hematopoietic activity from fetal liver to bone marrow. This might reflect an inability of *Etv6*$^{-/-}$ hematopoietic stem cells or progenitors to migrate or home to the bone marrow or, more likely, the failure of these cells to respond appropriately and/or survive within the bone marrow microenvironment.

These data establish Etv6 as the first transcription factor required specifically for hematopoiesis within the bone marrow, as opposed to other sites of hematopoietic activity during development. In conclusion, the inability of *Etv6*$^{-/-}$ HSCs (or progenitors) to stably colonize the bone marrow might reflect a role of Etv6 in adhesion or adhesion-mediated cellular responses (both mechanisms are defective in Etv6$^{-/-}$ HSCs). Moreover, Etv6 may mediate aspects of adhesion-independent cellular responses to the bone-marrow microenvironment.

4.4 ETV6 target genes

At molecular level, ETV6 acts as a sequence-specific transcriptional repressor. ETV6-repressive activity depends on two autonomous transcriptional repression domains. The first maps to the N-terminus in the pointed domain, in a region rich in proline residues (20% proline between amino acid residues 171 and 285), as commonly found in other transcriptional repressor. The second is located closer to the C-terminus region, in the central domain, between the pointed and the ets domains.[78]

Moreover, the ETS domain also mediates protein-protein interactions with unrelated factors either on its own or in combination with an adjacent domain. Some of these interactions are rather promiscuous with several ETS domains being able to interact with the same partner, whereas others are highly specific.[77]

Repression mediated by the central domain of ETV6 can be relieved by inhibiting histone deacetylases (HDACs). Co-immunoprecipitation studies showed that this domain interacts with corepressors that recruit HDACs (N-Cor, mSin3 and SMRT). Repression by the pointed domain is most likely mediated through interaction with the polycomb group protein L3MBTL (the human homolog of the Drosophila lethal(3)malignant brain tumor protein), which also contains a pointed domain.[78]

ETV6 is a ubiquitously expressed transcription factor of the ETS family with very few known targets. In particular, *Etv6* has been described to repress the transcription of

stromyelisin/Mmp3 (a gene encoding for a member of the matrix metalloproteinases); [83] *Bcl-XL* (a gene encoding for an anti-apoptotic factor, belonging to Bcl-2 pathway);[84] *Tcl1* (a gene encoding for the oncogene T-cell lymphoma 1); [85] *Ptger4, Sphk1*, *Tp53* and *Vegf.*[86] The categories of the modulated genes (e.g. cell adhesion/cellular matrix, cell growth/proliferation, apoptosis and angiogenesis) suggested a pleiotropic activity of ETV6.

4.5. Oncogenic role of *ETV6* in B cell malignancies

Deletions at chromosome 12p12-13 are observed in 26–47% of childhood pre-B ALL cases, suggesting the presence of a tumor suppressor gene,[87,88] as supported by genetic and functional evidences, both *in vitro* and *in vivo.*[78] For instance, ETV6 expression is able to abrogate growth in soft-agar of Ras-transformed NIH3T3 fibroblast,[83] and its over-expression induces an increase in apoptosis in serum-starved NIH3T3 cells through Bcl-XL pathway.[84]

ETV6 is involved in a high number of translocations with several fusion partners, which can be grouped into two families according to the function of the partner genes.

a) The protein kinases (PTK), where ETV6/PTK fusion structure is characterized by the pointed domain of ETV6 in the N-terminal half and tyrosine kinase domain of the partner gene in the C-terminal half of the fusion protein. The fusions of *ETV6* to *PDGFRβ*, *ABL1*, *ABL2*, *JAK2*, *NTRK3*, *FGFR3* and *SYK* have been described.[89] From a mechanistic point of view, the ETV6/PTK fusions are well understood: the pointed/SAM domain of ETV6 serves as a dimerization module for the fusion protein, thus leading to the constitutive activation of the PTK domain, autophosphorylation of the fusion protein as well as phosphorylation of cellular proteins. The ETV6/PTK fusions are found in a surprisingly wide spectrum of hematological and even non-hematological malignancies which include CML, AML, early pre-BALL, T-ALL, atypical CML, myelodysplastic syndrome, peripheral T-cell lymphoma and in some rare solid tumors.[78]

b) Fusions with other transcription factors give origin to aberrant transcription factors, through the joining of their regulatory domains and/or the DNA binding domain. Usually this type of fusions is not able to transform the cells *per se*, but they require additional hits. The most common fusion gene is *ETV6/RUNX1*, originated by the translocation t(12;21)(p13;q22), which is the most recurrent translocation among the childhood BCP-ALL with a rate equal to 25%.[1] One interesting feature of this translocation is that the second allele on the short arm of the non-rearranged chromosome 12p is frequently deleted.[90] Moreover in cases not carrying deletions in the normal short arm of chromosome 12, there is no expression of the wild type *ETV6* allele, supporting a very important role for loss of function of *ETV6* as a secondary event in ALL[90] Other transcription factors involved in ETV6/TX fusions are: *MN1*, *ARNT*, *EVI1*, *BTL*, *CDX2*, *HLXB9*, *MDS2*, *STL*, *PAX5* and others.[89]

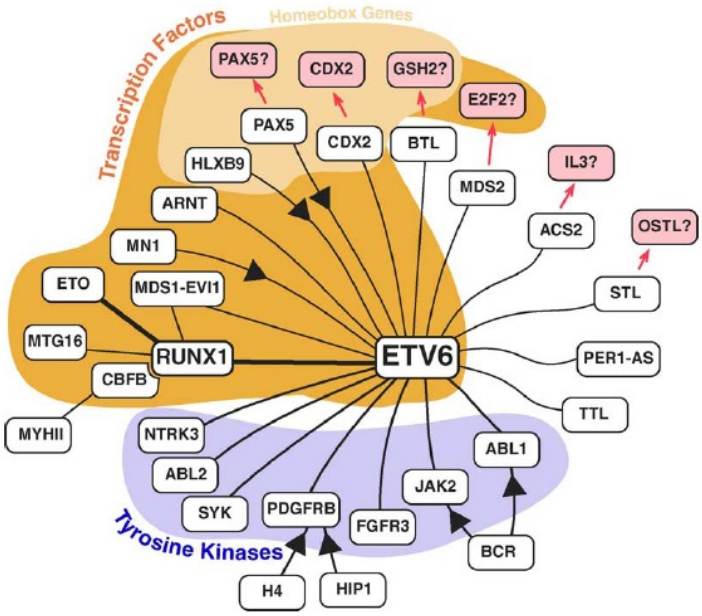


Figure 10. Fusion gene network of ETV6.[78]

5. PAX5/ETV6

Our group first described the generation of the *PAX5/ETV6* fusion gene, as a result of the translocation t(9;12)(q11;p13) in a patient affected by acute lymphoblastic leukemia (ALL) with BCP-lineage phenotype.[74] The novel fusion transcript resulted from joining the 5' end, containing the NH2-terminal region of the *PAX5* gene to the almost whole sequence of the *ETV6* gene. After the first patient, many other groups described the *PAX5/ETV6* fusion gene as a result either of the translocation t(9;12) or the dicentric dic(9;12), thus identifying *PAX5/ETV6* as the most recurrent *PAX5* fusion gene.[63,69,75,89,91,92]

The molecular and functional roles of PAX5/ETV6 are still poorly understood. Our group showed that PAX5/ETV6 protein acts as an aberrant transcription factor with repressor function, recruiting mSin3A, down-regulating *B220*, *CD19*, *BLNK*, *MB-1*, *FLT3*, and μ-heavy chain expression, thus suggesting a block on B-cell differentiation. Moreover, the fusion protein enhances cell migration towards CXCL12, with the overexpression of CXCR4. The presence of the fusion gene overcomes IL7 withdrawal and interferes with TGFβ1 pathway, inducing resistance and conferring cells an advantage in proliferation and survival.[93]

6. Aims of the study

Recently, PAX5 has been identified as a frequent target of mutation in BCP-ALL (around 30%). Among them 25% of the cases report deletions, 7% point mutations and 2-3% translocations that juxtapose PAX5 to several partner genes.[63]

To dissect the real frequency of *PAX5* alterations in the Italian Cohort of BCP-ALL pediatric cases, we screened patients enrolled into AIEOP-IBFM-ALL2000 protocol, identifying alterations in the *PAX5* locus, including deletions and translocations. To this purpose, we used different techniques in order to:

a) select the patients carrying alterations on the chromosome 9p by cytogenetic analysis (e.g. karyotype analysis);
b) describe the exact locus of the alterations and, in case of translocations, identify the partner genes;
c) detect additional events, providing further elements to characterize *PAX5* alterations as driver or passenger lesions.

To study the functional properties of *PAX5* fusion genes, we established an *in vitro* murine model of PAX5/ETV6 protein, as the most frequently reported example of PAX5 translocations. In particular, we aimed to widely characterize PAX5/ETV6 as an aberrant transcription factor, identifying the alterations induced in the gene expression profile and in the signaling pathway of primary populations of wt pre-BI cells co-cultured on OP9 BM stroma.

Finally, to investigate the potential role of *PAX5* as a tumor suppressor gene in ALL, we approached to study *PAX5* deletions, in a xenograft model of primary BCP-ALL cells transplanted in NSG mice. We focused the attention on the Ph+ subgroup, because of the high frequency of PAX5 deletions in BCR/ABL1 positive patients. To this purpose we used different approaches:

a) establishment of xenografts of 14 childhood BCP-ALL patients (Ph+ and Ph-, as control group);
b) ChIP sequencing in order to identify PAX5 target genes in leukemic settings;

c) analysis of the phosphorylation status;
d) generation of an *in vitro* inducible vector model, in order to better study the effect of PAX5 and PAX5/ETV6 in human leukemic cells.

Taken together, this study aimed to contribute to a better understanding of the role of *PAX5* lesions in BCP-ALL leukemia with a comprehensive approach, from patients samples to the establishment of murine *in vitro* and human *ex vivo* models.

References

1 Pui, C. H., Relling, M. V. & Downing, J. R. Acute lymphoblastic leukemia. *N Engl J Med* **350**, 1535-1548 (2004).

2 Armstrong, S. A. & Look, A. T. Molecular genetics of acute lymphoblastic leukemia. *J Clin Oncol.* **23**, 6306-6315 (2005).

3 Harrison, C. J. Acute lymphoblastic leukaemia. *Best Practice & Research Clinical Haematology* **14**, 593-607 (2001).

4 Hunger, S. P. *et al.* Children's Oncology Group's 2013 blueprint for research: acute lymphoblastic leukemia. *Pediatr Blood Cancer* **60**, 957-963, doi:10.1002/pbc.24420 (2013).

5 Mullighan, C. G. The molecular genetic makeup of acute lymphoblastic leukemia. *Hematology Am Soc Hematol Educ Program* **2012**, 389-396, doi:2012/1/389 [pii]10.1182/asheducation-2012.1.389 (2012).

6 Inaba, H., Greaves, M. & Mullighan, C. G. Acute lymphoblastic leukaemia. *Lancet* **381**, 1943-1955, doi:S0140-6736(12)62187-4 [pii]10.1016/S0140-6736(12)62187-4 (2013).

7 Greaves, M. F. & Wiemels, J. Origins of chromosome translocations in childhood leukaemia. *Nat Rev Cancer* **3**, 639-649 (2003).

8 Greaves, M. Infection, immune responses and the aetiology of childhood leukaemia. *Nat Rev Cancer* **6**, 193-203 (2006).

9 Mullighan, C. G. Molecular genetics of B-precursor acute lymphoblastic leukemia. *J Clin Invest* **122**, 3407-3415, doi:61203 [pii]10.1172/JCI61203 (2012).

10 Busslinger, M. Transcriptional control of early B cell development. *Annu. Rev. Immunol.* **22**, 55-79 (2004).

11 Matthias, P. & Rolink, A. G. Transcriptional Networks In Developing And Mature B Cells. *Nature Reviews* **5**, 497-508 (2005).

12 Fazio, G., Biondi, A. & Cazzaniga, G. *The role of PAX5 in ALL*. 1st edn, (InTech Open Access Publisher, 2011).

13 Nutt, S. L. & Kee, B. L. The Transcriptional Regulation of B Cell Lineage Commitment. *Immunity Reviews* **26**, 715-725 (2007).

14 Fuxa, M. & Skok, J. A. Transcriptional regulation in early B cell development. *Current Opinion in Immunology* **19**, 129-136 (2007).

15 Hagman, J. & Lukin, K. Early B-cell factor 'pioneers' the way for B-cell development. *Trends Immunol* **26**, 455-461, doi:S1471-4906(05)00180-8 [pii]10.1016/j.it.2005.07.001 (2005).

16 Hagman, J. & Lukin, K. Transcription factors drive B cell development. *Current Opinion in Immunology* **18**, 127-134 (2006).

17 Nutt, S. L., Heavey, B., Rolink, A. G. & Busslinger, M. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature* **401**, 556-562 (1999).

18 Cobaleda, C., Schebesta, A., Delogu, A. & Busslinger, M. Pax5: the guardian of B cell identity and function. *Nature Immunology* **8**, 463-470 (2007).

19 Hendriks, R. W. & Kersseboom, R. Involvement of SLP-65 and Btk in tumor suppression and malignant transformation of pre-B cells. *Semin Immunol* **18**, 67-76 (2006).

20 Swaminathan, S. *et al.* BACH2 mediates negative selection and p53-dependent tumor suppression at the pre-B cell receptor checkpoint. *Nat Med* **19**, 1014-1022, doi:nm.3247 [pii]10.1038/nm.3247 (2013).
21 Duy, C. *et al.* BCL6 is critical for the development of a diverse primary B cell repertoire. *J Exp Med* **207**, 1209-1221, doi:jem.20091299 [pii]10.1084/jem.20091299 (2010).
22 Jumaa, H. *et al.* Abnormal development and function of B lymphocytes in mice deficient for the signaling adaptor protein SLP-65. *Immunity* **11**, 547-554, doi:S1074-7613(00)80130-2 [pii] (1999).
23 Ta, V. B. *et al.* Malignant transformation of Slp65-deficient pre-B cells involves disruption of the Arf-Mdm2-p53 tumor suppressor pathway. *Blood* **115**, 1385-1393, doi:blood-2009-05-222166 [pii]10.1182/blood-2009-05-222166 (2010).
24 Swaminathan, S. Loss of checkpoint control in acute lymphoblastic leukemia. (2013).
25 Herzog, S., Reth, M. & Jumaa, H. Regulation of B-cell proliferation and differentiation by pre-B-cell receptor signalling. *Nat Rev Immunol* **9**, 195-205 (2009).
26 Vainchenker, W. & Constantinescu, S. N. JAK/STAT signaling in hematological malignancies. *Oncogene* **32**, 2601-2613, doi:onc2012347 [pii]10.1038/onc.2012.347 (2013).
27 Bacher, U. *et al.* Mutations of the TET2 and CBL genes: novel molecular markers in myeloid malignancies. *Ann Hematol* **89**, 643-652, doi:10.1007/s00277-010-0920-6 (2010).
28 Klingmuller, U., Lorenz, U., Cantley, L. C., Neel, B. G. & Lodish, H. F. Specific recruitment of SH-PTP1 to the erythropoietin receptor causes inactivation of JAK2 and termination of proliferative signals. *Cell* **80**, 729-738, doi:0092-8674(95)90351-8 [pii] (1995).
29 van de Geijn, G. J., Gits, J. & Touw, I. P. Distinct activities of suppressor of cytokine signaling (SOCS) proteins and involvement of the SOCS box in controlling G-CSF signaling. *J Leukoc Biol* **76**, 237-244, doi:10.1189/jlb.0104041jlb.0104041 [pii] (2004).
30 Palvimo, J. J. PIAS proteins as regulators of small ubiquitin-related modifier (SUMO) modifications and transcription. *Biochem Soc Trans* **35**, 1405-1408, doi:BST0351405 [pii]10.1042/BST0351405 (2007).
31 Ochiai, K. *et al.* A self-reinforcing regulatory network triggered by limiting IL-7 activates pre-BCR signaling and differentiation. *Nat Immunol* **13**, 300-307, doi:ni.2210 [pii]10.1038/ni.2210 (2012).
32 Rickert, R. C. New insights into pre-BCR and BCR signalling with relevance to B cell malignancies. *Nat Rev Immunol* **13**, 578-591, doi:nri3487 [pii]10.1038/nri3487 (2013).
33 Deane, J. A. & Fruman, D. A. Phosphoinositide 3-kinase: diverse roles in immune cell activation. *Annu Rev Immunol* **22**, 563-598, doi:10.1146/annurev.immunol.22.012703.104721 (2004).
34 Kanie, T. *et al.* TEL-Syk fusion constitutively activates PI3-K/Akt, MAPK and JAK2-independent STAT5 signal pathways. *Leukemia* **18**, 548-555, doi:10.1038/sj.leu.24032662403266 [pii] (2004).

35 Streubel, B., Vinatzer, U., Willheim, M., Raderer, M. & Chott, A. Novel t(5;9)(q33;q22) fuses ITK to SYK in unspecified peripheral T-cell lymphoma. *Leukemia* **20**, 313-318, doi:2404045 [pii]10.1038/sj.leu.2404045 (2006).
36 Chiu, C. W., Dalton, M., Ishiai, M., Kurosaki, T. & Chan, A. C. BLNK: molecular scaffolding through 'cis'-mediated organization of signaling proteins. *EMBO J* **21**, 6461-6472 (2002).
37 Guo, B., Kato, R. M., Garcia-Lloret, M., Wahl, M. I. & Rawlings, D. J. Engagement of the human pre-B cell receptor generates a lipid raft-dependent calcium signaling complex. *Immunity* **13**, 243-253, doi:S1074-7613(00)00024-8 [pii] (2000).
38 Hashimoto, S. *et al.* Identification of the SH2 domain binding protein of Bruton's tyrosine kinase as BLNK--functional significance of Btk-SH2 domain in B-cell antigen receptor-coupled calcium signaling. *Blood* **94**, 2357-2364 (1999).
39 Ishiai, M. *et al.* BLNK required for coupling Syk to PLC gamma 2 and Rac1-JNK in B cells. *Immunity* **10**, 117-125, doi:S1074-7613(00)80012-6 [pii] (1999).
40 Su, Y. W. *et al.* Interaction of SLP adaptors with the SH2 domain of Tec family kinases. *Eur J Immunol* **29**, 3702-3711, doi:10.1002/(SICI)1521-4141(199911)29:11<3702::AID-IMMU3702>3.0.CO;2-R [pii]10.1002/(SICI)1521-4141(199911)29:11<3702::AID-IMMU3702>3.0.CO;2-R (1999).
41 Yamamoto, M. *et al.* BASH-novel PKC-Raf-1 pathway of pre-BCR signaling induces kappa gene rearrangement. *Blood* **108**, 2703-2711, doi:blood-2006-05-024968 [pii]10.1182/blood-2006-05-024968 (2006).
42 Herzog, S. *et al.* SLP-65 regulates immunoglobulin light chain gene recombination through the PI(3)K-PKB-Foxo pathway. *Nat Immunol* **9**, 623-631, doi:ni.1616 [pii]10.1038/ni.1616 (2008).
43 Han, J. H. *et al.* Class switch recombination and somatic hypermutation in early mouse B cells are mediated by B cell and Toll-like receptors. *Immunity* **27**, 64-75, doi:S1074-7613(07)00330-5 [pii]10.1016/j.immuni.2007.05.018 (2007).
44 Morrison, A. M., Nutt, S. L., Thevenin, C., Rolink, A. & Busslinger, M. Loss- and gain-of-function mutations reveal an important role of BSAP (Pax-5) at the start and end of B cell differentiation. *Semin Immunol* **10**, 133-142 (1998).
45 Nutt, S. L. *et al.* Independent regulation of the two Pax5 alleles during B-cell development. *Nature Genetics* **21**, 390-395 (1999).
46 Souabni, A., Cobaleda, C., Schebesta, M. & Busslinger, M. Pax5 Promotes B Lymphopoiesis and Blocks T Cell Development by Repressing Notch1. *Immunity* **17**, 781-793 (2002).
47 Cotta, C. V., Zhang, Z., Kim, H. G. & Klug, C. A. Pax5 determines B- versus T-cell fate and does not block early myeloid-lineage development. *Blood* **101**, 4342-4346 (2003).
48 Ghia, P., ten Boekel, E., Rolink, A. G. & Melchers, F. B-cell development: a comparison between mouse and man. *Immunol Today* **19**, 480-485 (1998).
49 Czerny, T. B. M. DNA-Binding and Transactivation Properties of Pax-6: Three Amino Acids in the Paired Domain Are Responsible for the Different Sequence Recognition of Pax-6 and BSAP (Pax-5). *Molecular and Cellular Biology* **15**, 2858-2871 (1995).
50 Rolink, A. G., Schaniel, C., Bruno, L. & Melchers, F. In vitro and in vivo plasticity of Pax5-deficient pre-B I cells. *Immunol Lett* **82**, 35-40, doi:S0165247802000160 [pii] (2002).

51 Cobaleda, C., Jochum, W. & M., B. Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors. *Nature* **449**, 473-479 (2007).
52 Nutt, S. L., Morrison, A. M., Dorfler, P., Rolink, A. & Busslinger, M. Indentification of BSAP (Pax-5) target genes in early B-cell development by loss- and gain-of-function experiments. *The EMBO Journal* **17**, 2319-2333 (1998).
53 Delogu, A. *et al.* Gene repression by Pax5 in B cells is essential for blood cell homeostasis and is reversed in plasma cells. *Immunity* **24**, 269-281 (2006).
54 Holmes, M. L., Carotta, S., Corcoran, L. M. & Nutt, S. L. Repression of Flt3 by Pax5 is crucial for B-cell lineage commitment. *Genes Dev* **20**, 933-938 (2006).
55 Schebesta, A. *et al.* Transcription Factor Pax5 Activates the Chromatin of Key Genes Involved in B Cell Signaling, Adhesion, Migration, and Immune Function. *Immunity* **27**, 49-63 (2007).
56 Maier, H. *et al.* Requirements for selective recruitment of Ets proteins and activation of mb-1/Ig-a gene transcription by Pax-5 (BSAP). *Nucleic Acids Research* **31**, 5483-5489 (2003).
57 Schebesta, M., Pfeffer, P. L. & Busslinger, M. Control of Pre-BCR Signaling by Pax5-Dependent Activation of the BLNK Gene. *Immunity* **17**, 473-485 (2002).
58 Kozmik, Z., Wang, S., Dorfler, P., Adams, B. & Busslinger, M. The Promoter of the CD19 Gene Is a Target for the B-Cell-Specific Transcription Factor BSAP. *Molecular and Cellular Biology* **12**, 2662-2672 (1992).
59 Horcher, M., Souabni, A. & Busslinger, M. Pax5/BSAP Maintains the Identity of B Cells in Late B Lymphopoiesis. *Immunity* **14**, 779-790 (2001).
60 Busslinger, M., Klix, N., Pfeffer, P., Graninger, P. G. & Kozmik, Z. Deregulation of PAX-5 by translocation of the Em enhancer of the IgH locus adjacent to two alternative PAX-5 promoters in a diffuse large-cell lymphoma. *Proc. Natl. Acad. Sci. USA* **93**, 6129-6134 (1996).
61 Lida, S. *et al.* The t(9;14)(p13;q32) Chromosomal Translocation Associated With Lymphoplasmacytoid Lymphoma Involves the PAX-5 Gene. *Blood* **88**, 4110-4117 (1996).
62 Souabni, A., Jochum, W. & Busslinger, M. Oncogenic role of Pax5 in the T-lymphoid lineage upon ectopic expression from the immunoglobulin heavy-chain locus. *Blood* **109**, 281-289 (2007).
63 Mullighan, C. G. *et al.* Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* **446**, 758-764 (2007).
64 Familiades, J. *et al.* PAX5 mutations occur frequently in adult B-cell progenitor acute lymphoblastic leukemia and PAX5 haploinsufficiency is associated with BCR-ABL1 and TCF3-PBX1 fusion genes: a GRAALL study. *Leukemia* **23**, 1989-1998, doi:leu2009135 [pii]10.1038/leu.2009.135 (2009).
65 Paulsson, K. *et al.* Microdeletions are a general feature of adult and adolescent acute lymphoblastic leukemia: Unexpected similarities with pediatric disease. *Proc Natl Acad Sci U S A* **105**, 6708-6713, doi:0800408105 [pii]10.1073/pnas.0800408105 (2008).
66 Den Boer, M. L. *et al.* A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide classification study. *Lancet Oncol* **10**, 125-134, doi:S1470-2045(08)70339-5 [pii]10.1016/S1470-2045(08)70339-5 (2009).
67 Iacobucci, I. *et al.* The PAX5 gene is frequently rearranged in BCR-ABL1-positive acute lymphoblastic leukemia but is not associated with outcome. A report on

behalf of the GIMEMA Acute Leukemia Working Party. *Haematologica* **95**, 1683-1690, doi:haematol.2009.020792 [pii]10.3324/haematol.2009.020792 (2010).
68 Santoro, A. *et al.* Altered mRNA expression of PAX5 is a common event in acute lymphoblastic leukaemia. *Br J Haematol* **146**, 686-689 (2009).
69 An, Q. *et al.* Variable breakpoints target PAX5 in patients with dicentric chromosomes: a model for the basis of unbalanced translocations in cancer. *Proc Natl Acad Sci U S A* **105**, 17050-17054 (2008).
70 Shah, S. *et al.* A recurrent germline PAX5 mutation confers susceptibility to pre-B cell acute lymphoblastic leukemia. *Nat Genet* **45**, 1226-1231, doi:ng.2754 [pii]10.1038/ng.2754 (2013).
71 Auer, F. *et al.* Inherited susceptibility to pre B-ALL caused by germline transmission of PAX5 c.547G>A. *Leukemia*, doi:leu2013363 [pii]10.1038/leu.2013.363 (2013).
72 Paulsson, K. *et al.* Methylation of tumour suppressor gene promoters in the presence and absence of transcriptional silencing in high hyperdiploid acute lymphoblastic leukaemia. *Br J Haematol* **144**, 838-847, doi:BJH7523 [pii]10.1111/j.1365-2141.2008.07523.x (2009).
73 Nebral, K. *et al.* Incidence and diversity of PAX5 fusion genes in childhood acute lymphoblastic leukemia. *Leukemia* **23**, 134-143 (2009).
74 Cazzaniga, G. *et al.* The paired box domain gene PAX5 is fused to ETV6/TEL in an acute lymphoblastic leukemia case. *Cancer Res* **61**, 4666-4670 (2001).
75 Coyaud, E. *et al.* Wide diversity of PAX5 alterations in B-ALL: a Groupe Francophone de Cytogenetique Hematologique study. *Blood* **115**, 3089-3097, doi:blood-2009-07-234229 [pii]10.1182/blood-2009-07-234229 (2010).
76 Poirel, H. *et al.* The TEL gene products: nuclear phosphoproteins with DNA binding properties. *Oncogene* **14**, 349-357 (1997).
77 Lopez, R. G. *et al.* TEL is a sequence-specific transcriptional repressor. *J Biol Chem* **274**, 30132-30138 (1999).
78 Bohlander, S. K. ETV6: a versatile player in leukemogenesis. *Semin Cancer Biol* **15**, 162-174 (2005).
79 Wang, L. C. *et al.* Yolk sac angiogenic defect and intra-embryonic apoptosis in mice lacking the Ets-related factor TEL. *EMBO J* **16**, 4374-4383 (1997).
80 Wang, L. C. *et al.* The TEL/ETV6 gene is required specifically for hematopoiesis in the bone marrow. *Genes Dev* **12**, 2392-2402 (1998).
81 Hock, H. M. E. M. S. S. J. W. V. P. J. M. F. Y. O. S. H. Tel/Etv6 is an essential and selective regulator of adult hematopoietic stem cell survival. *GENES & DEVELOPMENT* **18**, 2336â€“2341 (2004).
82 Golub, T. R., Barker, G. F., Lovett, M. & Gilliland, D. G. Fusion of PDGF receptor beta to a novel ets-like gene, tel, in chronic myelomonocytic leukemia with t(5;12) chromosomal translocation. *Cell* **77**, 307-316 (1994).
83 Fenrick, R. *et al.* TEL, a putative tumor suppressor, modulates cell growth and cell morphology of ras-transformed cells while repressing the transcription of stromelysin-1. *Mol Cell Biol* **20**, 5828-5839 (2000).
84 Irvin, B. J. *et al.* TEL, a putative tumor suppressor, induces apoptosis and represses transcription of Bcl-XL. *J Biol Chem* **278**, 46378-46386 (2003).
85 Fears, S., Chakrabarti, S. R., Nucifora, G. & Rowley, J. D. Differential expression of TCL1 during pre-B-cell acute lymphoblastic leukemia progression. *Cancer Genet Cytogenet* **135**, 110-119 (2002).

86 Boily, G., Larose, J., Langlois, S. & Sinnett, D. Identification of transcripts modulated by ETV6 expression. *Br J Haematol* **136**, 48-62 (2007).
87 Montpetit, A. *et al.* Mutational and expression analysis of the chromosome 12p candidate tumor suppressor genes in pre-B acute lymphoblastic leukemia. *Leukemia* **18**, 1499-1504 (2004).
88 Raimondi, S. C. *et al.* 12p abnormalities and the TEL gene (ETV6) in childhood acute lymphoblastic leukemia. *Blood* **90**, 4559-4566 (1997).
89 Zhou, M. H. *et al.* Detection of ETV6 gene rearrangements in adult acute lymphoblastic leukemia. *Ann Hematol* **91**, 1235-1243, doi:10.1007/s00277-012-1431-4 (2012).
90 Patel, N. *et al.* Expression profile of wild-type ETV6 in childhood acute leukaemia. *Br J Haematol* **122**, 94-98 (2003).
91 Strehl, S., Konig, M., Dworzak, M. N., Kalwak, K. & Haas, O. A. PAX5/ETV6 fusion defines cytogenetic entity dic(9;12)(p13;p13). *Leukemia* **17**, 1121-1123 (2003).
92 Kawamata, N. *et al.* Molecular allelokaryotyping of pediatric acute lymphoblastic leukemias by high-resolution single nucleotide polymorphism oligonucleotide genomic microarray. *Blood* **111**, 776-784 (2008).
93 Fazio, G., Palmi, C., Rolink, A., Biondi, A. & Cazzaniga, G. PAX5/TEL acts as a transcriptional repressor causing down-modulation of CD19, enhances migration to CXCL12, and confers survival advantage in pre-BI cells. *Cancer Res* **68**, 181-189 (2008).

Chapter 2

Four novel fusion transcripts involving *PAX5* in B-cell precursor ALL

Four novel fusion transcripts involving *PAX5* in B-cell precursor ALL.

Manuscript

Grazia Fazio,[1]* Giulia Daniele,[2]* Valeria Cazzaniga,[1] Luciana Impera,[3] Marco Severgnini,[4] Ilaria Iacobucci,[5] Marta Galbiati,[1] Anna Leszl,[6] Ingrid Cifola,[4] Gianluca De Bellis,[4] Paola Bresciani,[7] Giovanni Martinelli,[5] Giuseppe Basso,[6] Andrea Biondi,[1] Clelia Tiziana Storlazzi[2°] and Giovanni Cazzaniga[1°].

[1]Centro RicercaTettamanti, Clinica Pediatrica, Dipartimento di Scienze della Salute, Università di Milano-Bicocca, Ospedale S. Gerardo/Fondazione MBBM, Monza, Italy; [2]Dipartimento di Biologia, Università degli Studi di Bari "Aldo Moro", Bari, Italy; [3]Dept. of Emergency and Organ Transplantation, Hematology - University of Bari, Bari, Italy; [4]Istituto di Tecnologie Biomediche, Consiglio Nazionale delle Ricerche, Milano, Italy; [5]Institute of Hematology "L. e A. Seragnoli", Dep. of Experimental, Diagnostic and Specialty Medicine, University of Bologna, Bologna, Italy; [6]Laboratorio di Oncoematologia, Dipartimento di Pediatria Università Padova, Padova, Italy; [7]Department of Medical and Surgical Sciences, Azienda Ospedaliero-Universitaria Policlinico, University of Modena and Reggio Emilia, Modena, Italy.

*These authors equally contributed to this work.

°These authors shared the senior authorship of this study.

Keywords: B-cell precursor ALL, *PAX5*, *CHFR*, fusion gene

Running Title: Four novel *PAX5* fusion transcripts

Corresponding author:
Andrea Biondi, MD
Pediatric Clinic
University of Milano-Bicocca
Hospital S.Gerardo/Fondazione MBBM
Via Pergolesi, 33
20900-Monza (IT)
phone: +39-039-2333513
abiondi.unimib@gmail.com

Abstract

PAX5, a transcription factor essential for B-cell development, has been found as a frequent target of abnormalities in B-cell precursor ALL cases, showing point mutations, deletions, as well as translocations with several partner genes. Performing a screening on BCP-ALL cases with 9p rearrangements, we identified four novel *PAX5* fusion partner genes. *PAX5* was found fused to an unidentified transcript isoform of *CHFR* in a child with a t(9;12)(p13;q34) translocation. In another pediatric patient, harboring a dic(9;12)(p13;p13) chromosome, *PAX5* was juxtaposed to *SOX5*. Furthermore, a *PAX*5/*POM121C* chimeric gene was detected in an additionalcase with a t(7;9)(q11;p13) translocation. A fourth *PAX5* partner gene was identified in an adult B-ALL case, which showed a deletion of the short arm of chromosome 9, leading to the *PAX5*/*MLLT3* fusion. Furthermore, we detected two pediatric patients harboring a translocation t(7;9)(q11.2;p13.2) with a *PAX5/AUTS2* fusion transcript, confirming it as a recurrent alteration in pediatric B-ALL. Copy Number Variation analysis of translocated samples showed that few significant cooperative genetic lesions are present in addition to the translocation event, suggesting that it might have a primary role in leukemogenesis.

INTRODUCTION

PAX5 belongs to the *PAX* gene family and encodes for a transcription factor essential for B lymphoid cell commitment.[1] It acts both as a transcriptional activator and repressor on different target genes involved in lineages development.[2,3] *PAX5* has been recently reported to be target of aberrancies (including point mutation, deletions, and translocations), in approximately 30% of pediatric patients affected by B-Cell Precursor Acute Lymphoblastic Leukemia (BCP-ALL),[4] the most frequent leukemia subset of children.[5] Since the first report of the *PAX5/ETV6* fusion in 2001,[6] an increasing number of translocations involving *PAX5* in BCP-ALL has been described, both in childhood and adult patients. As recently reviewed,[7]*PAX5* translocations are estimated to occur at an incidence of 2-3%,[8] with a variety of partner genes, encoding for transcription factors (*TEL*,[4,6,9-12] *PML*,[13] *FOXP1*[4,11]), kinases (*JAK2*[8]), structural proteins (*ELN*,[14,15] *POM121*[8]) or molecules of unknown function (*C20orf112*[11], *AUTS2*[11,16,17]).

The N-terminal DNA binding domain of PAX5 is invariably joined to the C-terminus of the partner protein, which substitutes the PAX5 regulatory domains. The reciprocal fusion product, joining the N-terminus of the partner protein to the C-terminus of PAX5 is not usually preserved, with the sole known exception of JAK2/PAX5.[8]

The high heterogeneity of fusion genes makes difficult the understanding of the contribution of each single *PAX5* partner gene and the identification of shared features. Moreover, the observation that translocations involving *PAX5* are mainly unbalanced[10] suggests that the fusion protein retaining the PAX5 DNA-Binding-Domain (DBD) is the driving lesion. However, it is not known whether it is *per se* sufficient for leukemia pathogenesis or if additional events are needed.

Herein, we analyzed 22 patients carrying *PAX5* alterations, such as deletions and translocations in 17 and 5 cases, respectively. Interestingly, in 4 out of the 6 fusion genes originated, we identified novel partners of *PAX5*: a novel transcript isoform of *CHFR*, *SOX5*, *POM121C*, and *MLLT3*. Moreover, we reported two further cases of *PAX5/AUTS2*.

Additionally, to the best of our knowledge, this is the first study characterizing *PAX5*-translocated cases by Copy Number Variation analysis in pediatric leukemia, aiming to disclose whether *PAX5* alterations might be part of a complex scenario of cooperating

genetic lesions involved in leukemogenesis or if they represent single genetic aberration events.

PATIENTS AND METHODS

Patient cohort. Among pediatric patients with ALL enrolled in the AIEOP-BFM ALL2000 protocol (for whom conventional or molecular cytogenetic analyses were available), 21 (19 BCP-ALL, and 2 T-ALL) presented visible abnormalities on the 9p chromosomal region, and were therefore considered for this study. In addition, an adult BCP-ALL patient was analyzed, as having a 9p abnormality in addition to the t(9;22) translocation. The main patients' features are reported in Table 1.

FISH Analysis. FISH analysis was performed on bone marrow metaphases from archival methanol:acetic acid-fixed chromosome suspensions, as detailed in Supplementary Information and previously described.[18] BAC and fosmid clones were derived from the University of California Santa Cruz (UCSC) database (release of March 2006, NCBI36/hg18), and previously tested on normal human metaphases.

RNA Isolation. Total RNA was extracted by the standard guanidinium thiocyanate-phenol-chloroform extraction method.[19]

Rapid Amplification of cDNA Ends (RACE) Analyses. 3' RACE PCR and 5' RACE PCR were performed using commercial assays (3´ RACE or 5' RACE System for Rapid Amplification of cDNA Ends, Life Technologies, Carlsbad, California, USA), according to the manufacturer's instructions. A primer specific for *PAX5* exon1A was used in combination with the commercial UAP primer (3' RACE); first strand cDNA synthesis was primed using a gene-specific antisense oligonucleotide on the *PAX5*-partner gene sequence (5' RACE, see Supplementary Table S1). Gene specific primers used in first and nested RACE PCRs are also listed in Supplementary Table S1. Both 3' and 5' RACE products were cloned into a T&A cloning vector (RBC Bioscience Corporation, NewTaipei City, Taiwan), and selected clones were directly sequenced in both directions, using vector-specific M13 oligonucleotides, using the ABI-3130 Genetic Analyzer instrument (Applied Biosystems). Alignment was carried out using the Basic Local Alignment Search Tool database (BLAST, www.blast.ncbi.nlm.nih.gov).

RT-PCR. One μg of total RNA was reversely transcribed with Superscript II reverse transcriptase (Life Technologies). To amplify the breakpoint region, RT-PCR was

performed on patient's cDNA, using a *PAX5* forward primer and the correspondent reverse primer of the specific partner gene. All the RT-PCR reactions were performed in the following conditions: one activation step (2' at 94°C), thirty-five cycles of amplification (30 s at 94°C, 30 s at 60°C, 60 s at 72°C), using TaqPlatinum polymerase (Life Technologies). RT-PCR primers and product length are reported in Supplementary Table S1.

Characterization of *CHFR* expression in normal tissues. To investigate the *CHFR* tissue-specific expression pattern, we tested cDNA from multiple tissues (bladder, brain, cervix, esophagus, heart, kidney, liver, lung, ovary, placenta, prostate, skeletal muscle, small intestine, spleen, testes, thymus, thyroid, trachea) (First Choice Total RNA Survey Panel, Catalog No. AM6000, Ambion, Milan, Italy), together with cDNA from fetal brain RNA, peripheral blood and bone marrow (Clontech, Saint-Germain-en-Laye, France). cDNA was obtained by reverse transcription using the QuantiTect Reverse Transcription Kit (Qiagen) and amplified using the Applied Biosystems Real-Time PCR System 7300 in the presence of 1X Platinum®SYBR Green qPCR SuperMix-UDG with ROX (Life Technologies), and 300nM of each sense and antisense primer. All measurements were performed at least in triplicate.

The PCR conditions were as it follows: 2 min at 50°C, 10 min at 95°C, followed by 40 cycles of 15 s at 95°C and 1 min at 60°C for all the primer pairs used. Melting curves were acquired and analyzed to control the specificity. Two primer pairs (corresponding to the *CHFR* wild-type and novel transcript isoform, respectively) were designed to amplify products of 65 and 133 bp, respectively. Primer sequences are reported in the Supplementary Table S1.The gene expression level was calculated using the relative quantification approach based on the ΔΔCt method. [20] In all the experiments, *GAPDH* gene was used as reference, and normal bone marrow as calibrator.

RNA-Seq analysis. Total RNA was checked for integrity and purity by capillary electrophoresis (Agilent Bioanalyser 2100, Agilent Technologies, Santa Clara, CA, USA), and 2 μg RNA was used to prepare a double-stranded cDNA library for high-throughput sequencing by Truseq RNA kit (Illumina, San Diego, CA, USA) and paired-end sequenced on a GAIIx instrument as described in Supplementary Information.

Copy Number Variation analysis. Five diagnostic samples (pt. no. 5, 7, 10, 15 and 19), as well as their corresponding remission samples, were genotyped by the Affymetrix®

Cytogenetics Whole Genome 2.7M Array (Affymetrix®, Santa Clara, CA), while patient n°22 has been analyzed by Genome-Wide Human SNP Array 6.0 (Affymetrix®). Affymetrix® GeneChip Command Console™ (AGCC) v3.1 generated .CEL files, and analyzed by Chromosome Analysis Suite software (Affymetrix®). Amplified and deleted regions with a minimal length of 100kbp and 50kbp respectively and involving a minimum number of 20 markers were considered.. Strechtes of Loss Of Heterozygousity were annotated if covering a minimal region of 2000kbp.

Conversely, constitutional copy number polymorphisms, based on the comparison of the tumor sample with its corresponding remission sample, were not annotated.

RESULTS

Identification of *PAX5* altered cases.

We selected 20 BCP-ALL and 2 T-ALL patients, as presenting conventional or molecular cytogenetic abnormalities involving the 9p13 chromosomal region, therefore potentially involving *PAX5*. By performing a detailed FISH analysis using BAC probes specific for *PAX5* (Supplementary Tables S2-S3), we found gene deletions and translocations in 17 and 5 cases, respectively (Table 1). FISH with BAC probes determined a large loss of 9p, including the *PAX5* gene (Supplementary Figure S1). A *PAX5* translocation was detected in 5 patients (Pts. 5, 7, 10, 15 and 19) (Figure 1). Interestingly, patient 22 had a 9p intra-chromosomal deletion, showing the proximal breakpoint within *PAX5*. The characterization of each novel fusion gene involving *PAX5* is described below.

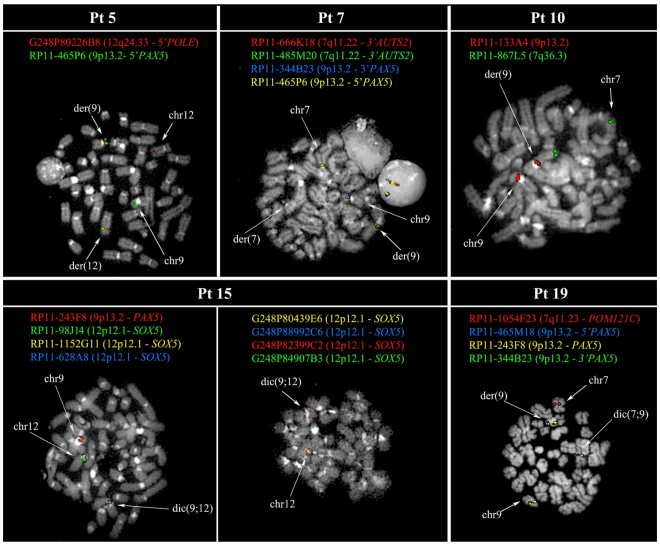


Figure 1. FISH analysis of *PAX5* translocations. BAC probes, shown on the top of each panel, are coloured as their corresponding signals in the merged image.

PAX5* is fused to a novel transcript isoform of *CHFR.

The cytogenetic analysis of patient 5 revealed a translocation between chromosomes 9 and 12, and then defined by the co-localization of RP11-465P6 and G248P80226B8 probes (Figure 1). RACE-PCR identified the 5' sequence of *PAX5* joined to a novel coding sequence of the *Checkpoint with fork-head associated and ring finger* (*CHFR)* gene on chromosome 12, herewith indicated as exons 19 and 20 in Figure 2A. More precisely, Blast and UCSC Genome Browser (release of Feb. 2009, GRCh37/hg19) analyses associated the unknown coding sequence to two ESTs on chromosome 12 genomic contig, at 133,408,802-133,408,929 and at 133,409,527-133,409,822 genomic locations, respectively. The two EST sequences, 296 bp and 128 bp, were separated by 599 bp and located after the 3' UTR region of the *CHFRCH-001* transcript, the longest one encoded by *CHFR* (Figure 2A). Several additional alternative transcripts of *CHFR* gene are described (Supplementary Figure S2A). By RACE and RT-PCR, we were able to demonstrate that the two unknown sequence regions we identified (named as exon 19 and 20) corresponded to additional exons (after exon 18 and 3'UTR of *CHFR-001* and before alternative exon 6 of *CHFR-028*). The three different transcript isoforms (*CH-001*, *CH-028* and the newly identified) used alternative 3'UTR (defined as 3' UTR (A), (B) and (C) in Figure 2B). We therefore propose to name this new alternative *CHFR* transcript as *CH-029.* RACE-PCR identified two alternative fusion transcripts, joining either *PAX5* exons 1-4 to *CHFR* ex19-20 (transcript 1,952 nt), or *PAX5* ex1 to *CHFR* ex19-20 (transcript 2, 522nt) (Figure 2B). As expected, the *PAX5/CHFR* transcripts 1 and 2 were exclusive of patient 5, being not present in TOM1 and REH negative control cell lines (Figure 2C). Conversely, both patient 5 and TOM1 expressed the alternative CH-029 as well as the exons 19-20 (Supplementary Figure S2B), the longest transcript *CHFR CH-001,* and the wild type *PAX5* (Supplementary Figure S2C). The *CH-029* was detected together with *CH-001* in several leukemia and lymphoma cell lines (BCP-ALL: TOM1, REH; AML: HL-60, THP1; histiocytic lymphoma: U937; CML: K562), as well as in the 293T kidney fibroblasts cell line (Supplementary Figure S2D).

Interestingly, although the chromosomal translocation was balanced, any effort to amplify the reciprocal *CHFR-PAX5* transcript was unsuccessful. Both *CH-001* and *CH-029* transcripts were broadly expressed in non-tumor samples, with the strongest expression of both *CHFR* isoforms in the peripheral blood sample (Supplementary Figure S2E).

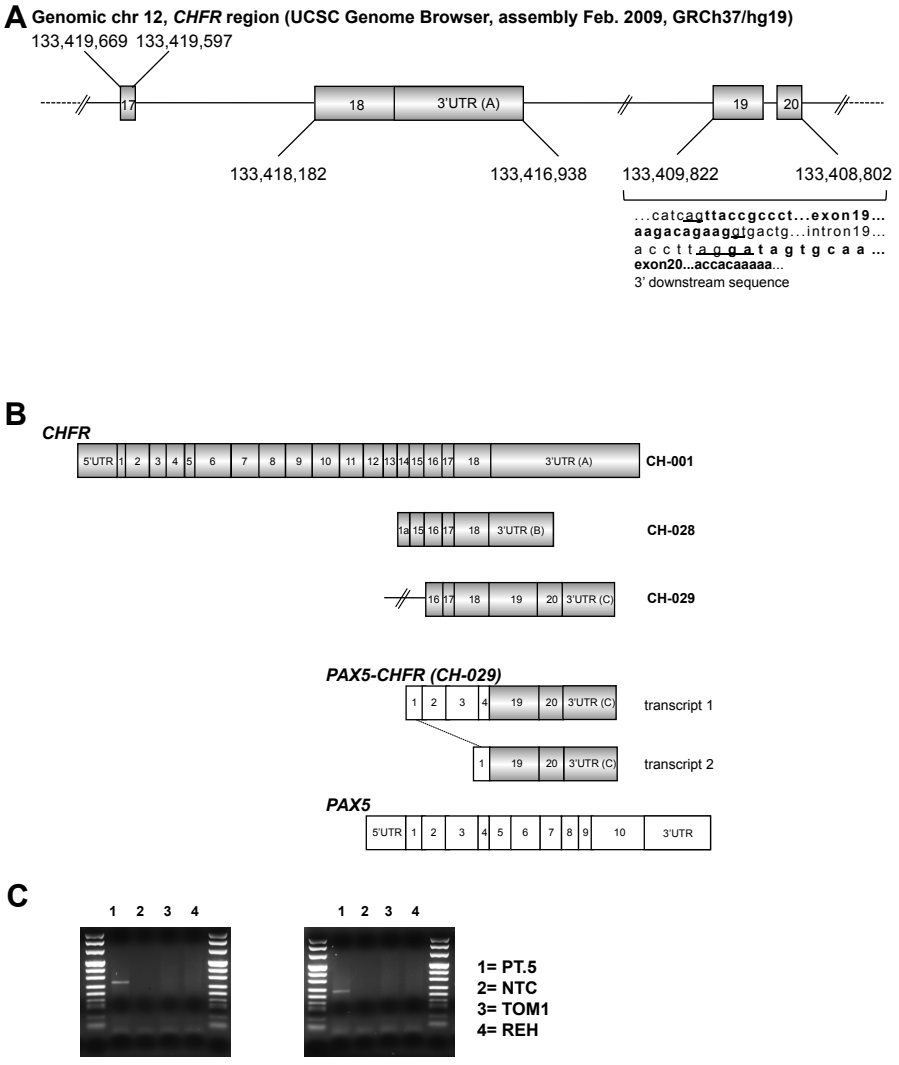


Figure 2. *PAX5/CHFR* fusion transcript and novel *CHFR* RNA isoform. A- Partial genomic organization of *CHFR* gene on chromosome 12, and sequence on newly identified exons. Location of exons is according to UCSC Genome Browser, assembly Feb. 2009, GRCh37/hg19. Partial base sequence of exons 19 and 20 is reported in bold. Splicing donor/acceptor bases are underlined between intron and exon; first/last ten bases are indicated. **B-** Schematic representation of *CHFR* different transcript isoforms (*CH-001* and *CH-028*) and the newly identified alternative transcript (*CH-029*) are represented. In addition, a schematic representation of wild-type *PAX5* and the two alternative *PAX5/CHFR* fusion transcripts are shown. **C-** *PAX5-CH-029* transcript 1 (*left panel*) and transcript 2 (*right panel*) are exclusively expressed in patient 5 (lane 1) and not in control BCP-ALL cell lines (TOM1 –lane 3 and REH - lane 4). Lane 2: no template control.

CNV analysis of patient 5 detected dispersed copy number neutral LOH as well as loss of genomic material, but not in the regions involved in the t(9;12) translocation (Supplementary Table S4 and Figure 3).

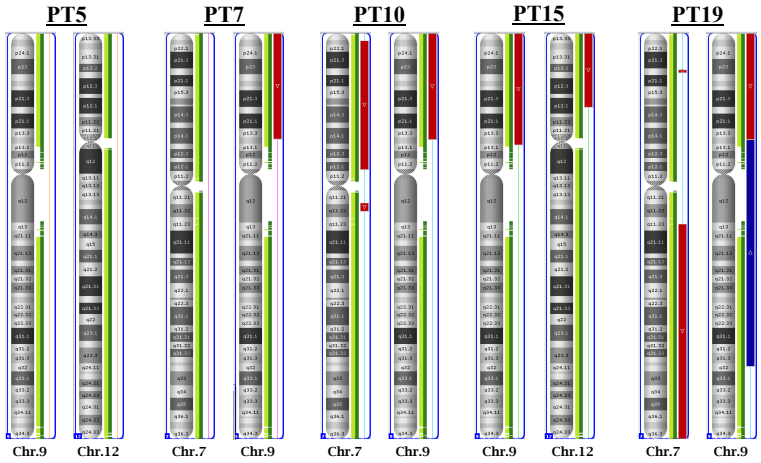


Figure 3. Karyoview of chromosomes involved in the translocation event in each patient. The Chromosome Analysis Suite software was used (Affymetrix®). A minimum number of 20 markers and a minimal length of 100 Kb and 50 Kb respectively were considered for defining gains and losses.

PAX5/SOX5 in a pediatric patient harboring a dic(9;12)(p13;p13) chromosome.

In patient 15, FISH analysis identified the *SRY (Sex Determining Region Y)-Box 5* (*SOX5)* gene on chromosome 12 as fused to *PAX5* on a dic(9;12)(p13;p13) dicentric chromosome. RACE-PCR defined the junction as located between *PAX5* exon 5 and *SOX5* exon 4. The PAX5 protein structure and the predicted PAX5/SOX5 fusion protein are represented in Figure 4A and B, respectively. The breakpoint was further confirmed by RT-PCR (Figure 4B). CNV analysis identified DNA loss at the translocation breakpoints on chromosomes 9 and 12 (Figure 3 and Supplementary Table S4).

POM121C is fused to *PAX5* in a t(7;9)(q11;p13) translocation.

In patient 19, CNV analysis detected deletions of 9p and 7q, focal on *PAX5* gene and *POM121 transmembrane nucleoporin C* (*POM121C)* gene, respectively (Supplementary Table S4 and Figure 3). In the context of a larger study (manuscript in preparation), RNA-sequencing identified reads with *PAX5* exon 6 sequence fused to *POM121C* exon 4 (Supplementary Figure S3), thus confirming that a chromosomal translocation involving

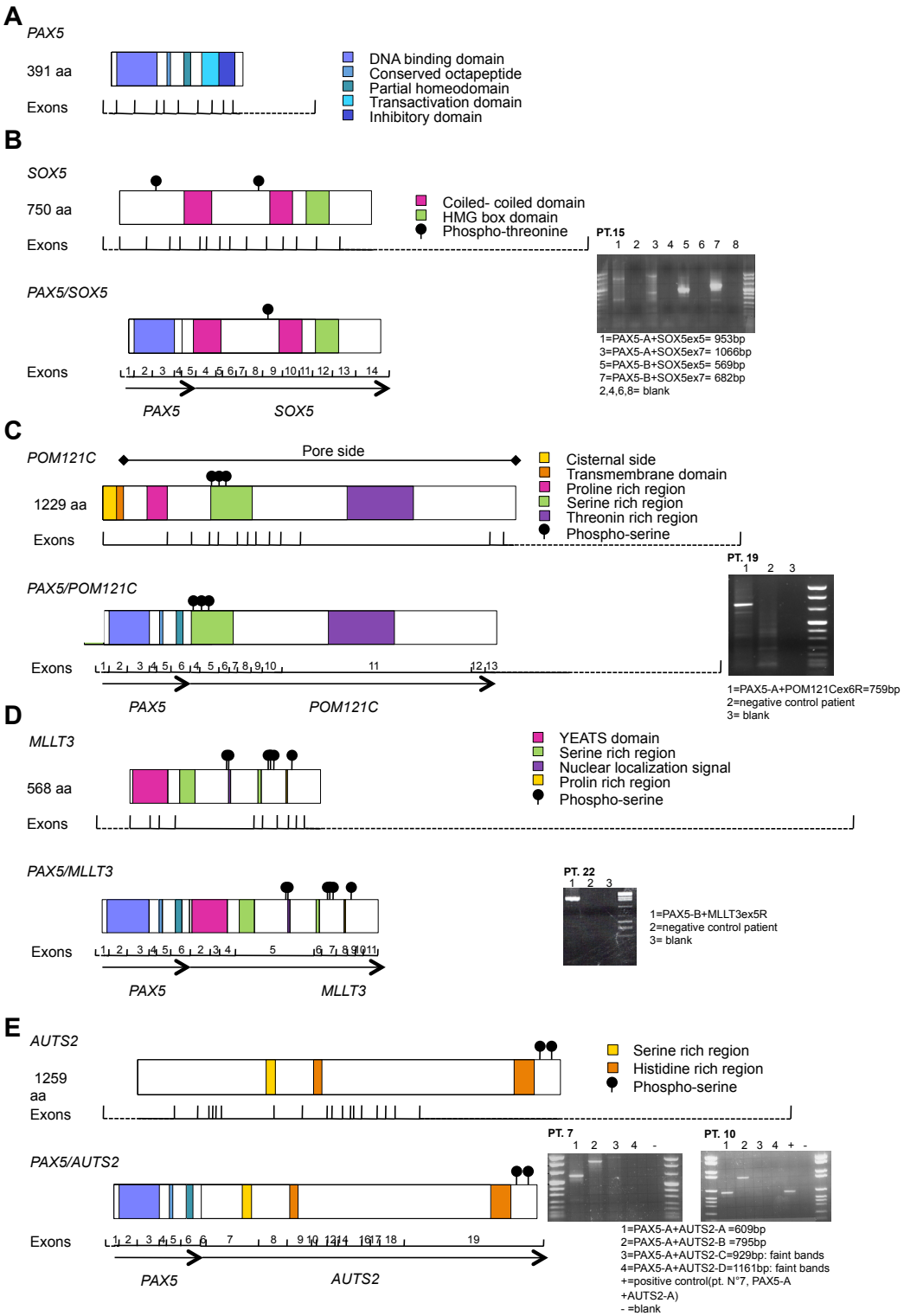


Figure 4. Schematic representation of the novel PAX5 fusion proteins and fusion genes expression analysis. For each novel fusion gene (*panels B-E*), the RT-PCR gel electrophoresis and the scheme of the predicted fusion protein is represented, indicating both the conserved functional domains of the wild type counterparts and the exons breakpoints. **A-**PAX5 protein structure, with indicated functional domains, and exon correspondence; **B-***PAX5/SOX5* in patient 15; **C-***PAX5/POM121C* in patient 19; **D-***PAX5/MLLT3* in patient 22; **E-***PAX5/AUTS2* in patients 7 and 10.

9p and 7q occurred in this case. FISH analysis indeed revealed the presence of a dicentric chromosome dic(7;9)(q11.23;p13.2) and the co-localization of BACs for *PAX5* and *POM121C* (Figure 1). RT-PCR further confirmed a *PAX5*/*POM121C* fusion transcript (Figure 4C).

An intra-chromosomal deletion on 9p originated a *PAX5/MLLT3* fusion gene.

An adult B-ALL patient carrying the t(9;22)(q34;q11) translocation was referred to our attention as having an additional large deletion on chromosome 9p (patient 22, Table 1). CNV analysis detected a loss of 16 Mb between *PAX5* at 9p13 and *MLLT3* at 9p21.3 (Supplementary Figure S4A). RT-PCR revealed a fusion transcript joining *PAX5* exon 6 to *MLLT3* exon 2 (Supplementary Figure S4B). The structure of the chimeric transcript, as well as the putative encoded protein are represented in Figure 4D. Details of CNV results are described in Supplementary Figure S4C.

***PAX5/AUTS2* is a recurrent genetic alteration.**

Patients 7 and 10 had a complex karyotype (Table 1), both characterized by the presence of a t(7;9) translocation. In both patients, FISH analyses showed a deletion of the 3' portion of *PAX5* (Figure 1). The retained segment of RP11-465P6 (5'*PAX5*) on der(9) was found to co-localize with the splitting signals of both BAC probes specific for the 3' portion of *AUTS2* (RP11-666K18, and RP11-485M20) on der(9) (Figure 1). This result indicated the presence of a *PAX5/AUTS2* fusion gene in both cases, further confirmed by RACE- and RT-PCR (Figure 4E). Sequence analysis of the PCR products revealed the joining of the 5' end of *PAX5* (at exon 6) to *AUTS2* gene (at exon 6), as schematically shown in Figure 4E.

In both patient 7 and 10, CNV analysis revealed the loss on the short arm of chromosome 9 with proximal breakpoint within *PAX5*, and in AUTS2 gene locus on chromosome 7, indicating an unbalanced event on chromosomes 7 and 9, compatible with the t(7;9) translocation (Supplementary Table S4 and Figure 3).

Copy Number Variation analysis of translocated cases.

As previously described, the analysis by Cytogenetics Whole Genome 2.7M Array (Affymetrix) detected CNV and/or copy number neutral LOH in different patients.

Moreover, we detected specific IG and/or TCR somatic rearrangements as a marker of the clonal origin of the leukemic cells. Interestingly, besides the translocation events, very few additional genomic lesions were identified in these patients, which were mainly losses (Supplementary Table S4).

In details, 3 patients out of 5 (patients 7, 10 and 15) had only one alteration, and 2 out of the 3 CNV affected a region without known coding genes or miRNAs (UCSC database, human reference assembly GRCh37/hg19). Patient 5 had 4 genetic alterations, which were linked to two main events, such as LOH of chromosome 5 and mosaicism of chromosome 7. Particularly, FISH analysis denoted the presence of a minor leukemic cell clone (about 30%) carrying the chromosome 7 deletion and a minor clone (about 20%) with gain of chromosome 7, in addition to the translocation (data not shown).

Patient 19 had 5 deletions (3 of them on chromosome 2). More interestingly, as a consequence of the translocation, although most of the 9p was lost (from *PAX5* intron 6 to pter), the remaining p arm and most q arm (on 9q32, till *ZFP37* included) was amplified (copy number 3), suggesting that even the fusion gene could be duplicated.

CNV analysis in patient 22, the only adult case, underlined a complex genomic scenario, with several genetic lesions, although most of them located in regions lacking coding genes or sno/miRNAs (Supplementary Figure S4C).

DISCUSSION

We identified four novel PAX5 fusion genes, namely *PAX5/CHFR*, *PAX5/SOX5*, *PAX5/POM121C* and *PAX5/MLLT3* in a cohort of 22 patients, carrying an alteration on the short arm of chromosome 9. Further two patients were positive for the t(7;9) translocation, giving rise to the *PAX5/AUTS2* fusion gene, which has been previously reported as recurrent in BCP-ALL.[11,16,17] In all described PAX5 fusions, including the novel partner described in the present study, the breakpoint on *PAX5* gene was not random, but hot spot of rupture has been observed after exons 4-7, thus retaining the PAX5 DNA binding domain region joined to the partner gene (schematically reported in Figure 5). In addition, most cases previously described and none of the cases here reported had the reciprocal fusion gene, even if carrying a balanced translocation. Moreover, PAX5 partner genes have variable function, making difficult to demonstrate a

common rationale for their contribution, except for the disruption of physiological PAX5 function.[7]

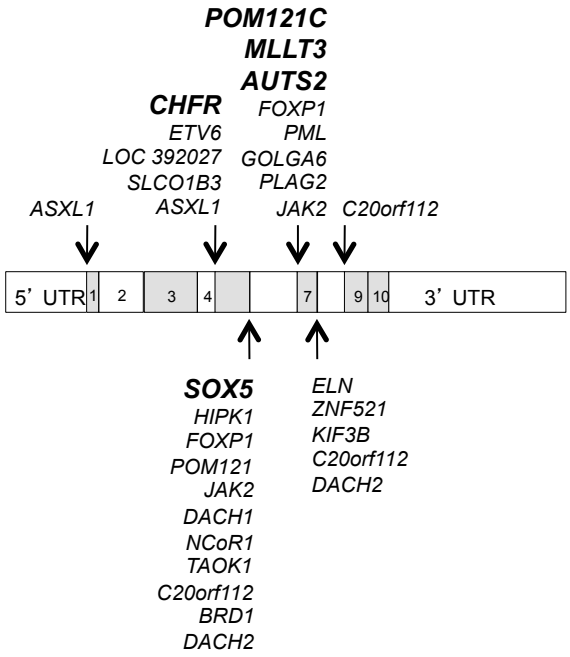


Figure 5. Map of the *PAX5* breakpoints in translocation. Partner genes indicated are derived from a revision of the literature. The names of the *PAX5* partner genes herein described are in bold.

Patient 5 carried the fusion of PAX5 with a new partner gene. Indeed, a previously unknown transcript of CHFR was herewith fused to PAX5. CHFR encodes for a mitotic checkpoint that delays the entry into metaphase in response to mitotic stress. CHFR has been described to be frequently silenced by hypermethylation in human cancers, indicating that it could act as a tumor suppressor.[21]

Patient 15 is the first known example of a fusion between *PAX5* and *SOX5*, a gene encoding for a member of the SOX (SRY-related HMG-box) family of transcription factors involved in the regulation of embryonic development and in the determination of the cell fate. Physiologically, the SOX5 protein acts as a transcriptional regulator after forming a protein complex with other proteins; it plays a role in neuron subtypes development, through the specific repression of target genes in the Central Nervous System (CNS).[22] Notably, this PAX5/SOX5 positive patient experienced a CNS relapse. It

can be hypothesized that the chimeric protein might act as an aberrant transcription factor affecting both B-cell (through PAX5 pathway alteration) and CNS features (by SOX5 pathway alteration). Interestingly, upregulation of the SOX5 has been described in primary splenic follicular lymphoma, by promoter swapping with the *P2RY8* gene,[23] with a mechanism analogous to the one described for the *CRLF2* gene.[24-26]

The third new PAX5 partner gene we identified was *POM121C*, which encodes for an essential component of the nuclear pore complex (NPC), with a repeat-containing domain that may be involved in anchoring components of the pore complex to the pore membrane. In details, *POM121C* gene (UCSC genomic location at chr. 7:75046065-75115568) is a member of the same nucleoporin family of POM121 (UCSC genomic location at chr. 7:72349936-72421979),[27] already identified as a fusion partner of PAX5 in BCP-ALL cases carrying a translocation.[8,28] Although little is known about these genes, haploinsufficiency of both POM121C and POM121 have been associated to developmental diseases such as the Williams and Williams-Beuren syndromes.[29]

In the single adult (patient 22), the novel *PAX5/MLLT3* fusion gene was concomitant to the t(9;22) translocation encoding for *BCR/ABL1.MLLT3* is also known as *AF9*, which is one of the most recurrent partner genes in MLL fusions in ALL and AML.[30] *PAX5/MLLT3* derived from an intra-chromosomal deletion on the short arm of chromosome 9. The patient age as well as the association of the t(9;22) make the genomic profile of this patient rather different compared to other PAX5-translocated cases.

Indeed, besides the translocation events, very few additional genomic lesions were identified in *PAX5* translocated pediatric BCP-ALL patients, similarly to what described for adult *PAX5* translocated cases.[31] We identified a mean of 2.4 alterations/patient (Supplementary Table S4), which is a lower number compared to previously reported childhood BCP-ALL cohorts (with a mean range of 4-8 lesions/patient, as reviewed in Bardini 2010).[32] Although the potential presence of point mutations cannot be ruled out, the low number of cooperative genetic lesions would suggest that the *PAX5* translocation event might have a primary role in leukemia. However, even though an increasing number of PAX5 fusion genes with variable partners have been reported (2-3% incidence)[7,8] and very recent studies proposed a functional role for chimeric proteins,[33,34] several fundamental questions on the role of PAX5 aberrations in leukemogenesis are still unresolved.

Supplementary information is available at Leukemia's website.

Acknowledgments

This work was partially supported by: Associazione Italiana Ricerca sul Cancro (AIRC) and Federazione Italiana Ricerca sul Cancro (FIRC), Fondazione Cariplo, Progetto InterBioMed Regione Lombardia, MIUR, UNIMIB.

Conflict of interest disclosures

The authors declare no competing financial interests or conflict of interest.

REFERENCES

1 Cobaleda, C., Schebesta, A., Delogu, A. & Busslinger, M. Pax5: the guardian of B cell identity and function. *Nature Immunology* **8**, 463-470 (2007).

2 Schebesta, A. *et al.* Transcription Factor Pax5 Activates the Chromatin of Key Genes Involved in B Cell Signaling, Adhesion, Migration, and Immune Function. *Immunity* **27**, 49-63 (2007).

3 Delogu, A. *et al.* Gene repression by Pax5 in B cells is essential for blood cell homeostasis and is reversed in plasma cells. *Immunity* **24**, 269-281 (2006).

4 Mullighan, C. G. *et al.* Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* **446**, 758-764 (2007).

5 Pui, C. H., Relling, M. V. & Downing, J. R. Acute lymphoblastic leukemia. *N Engl J Med* **350**, 1535-1548 (2004).

6 Cazzaniga, G. *et al.* The paired box domain gene PAX5 is fused to ETV6/TEL in an acute lymphoblastic leukemia case. *Cancer Res* **61**, 4666-4670 (2001).

7 Fazio, G., Biondi, A. & Cazzaniga, G. *The role of PAX5 in ALL*. 1st edn, (InTech Open Access Publisher, 2011).

8 Nebral, K. *et al.* Incidence and diversity of PAX5 fusion genes in childhood acute lymphoblastic leukemia. *Leukemia* **23**, 134-143 (2009).

9 Strehl, S., Konig, M., Dworzak, M. N., Kalwak, K. & Haas, O. A. PAX5/ETV6 fusion defines cytogenetic entity dic(9;12)(p13;p13). *Leukemia* **17**, 1121-1123 (2003).

10 An, Q. *et al.* Variable breakpoints target PAX5 in patients with dicentric chromosomes: a model for the basis of unbalanced translocations in cancer. *Proc Natl Acad Sci U S A* **105**, 17050-17054 (2008).

11 Kawamata, N. *et al.* Molecular allelokaryotyping of pediatric acute lymphoblastic leukemias by high-resolution single nucleotide polymorphism oligonucleotide genomic microarray. *Blood* **111**, 776-784 (2008).

12 Zhou, M. H. *et al.* Detection of ETV6 gene rearrangements in adult acute lymphoblastic leukemia. *Ann Hematol* **91**, 1235-1243, doi:10.1007/s00277-012-1431-4 (2012).

13 Nebral, K. *et al.* Identification of PML as novel PAX5 fusion partner in childhood acute lymphoblastic leukaemia. *British Journal of Haematology* **139**, 269-274 (2007).

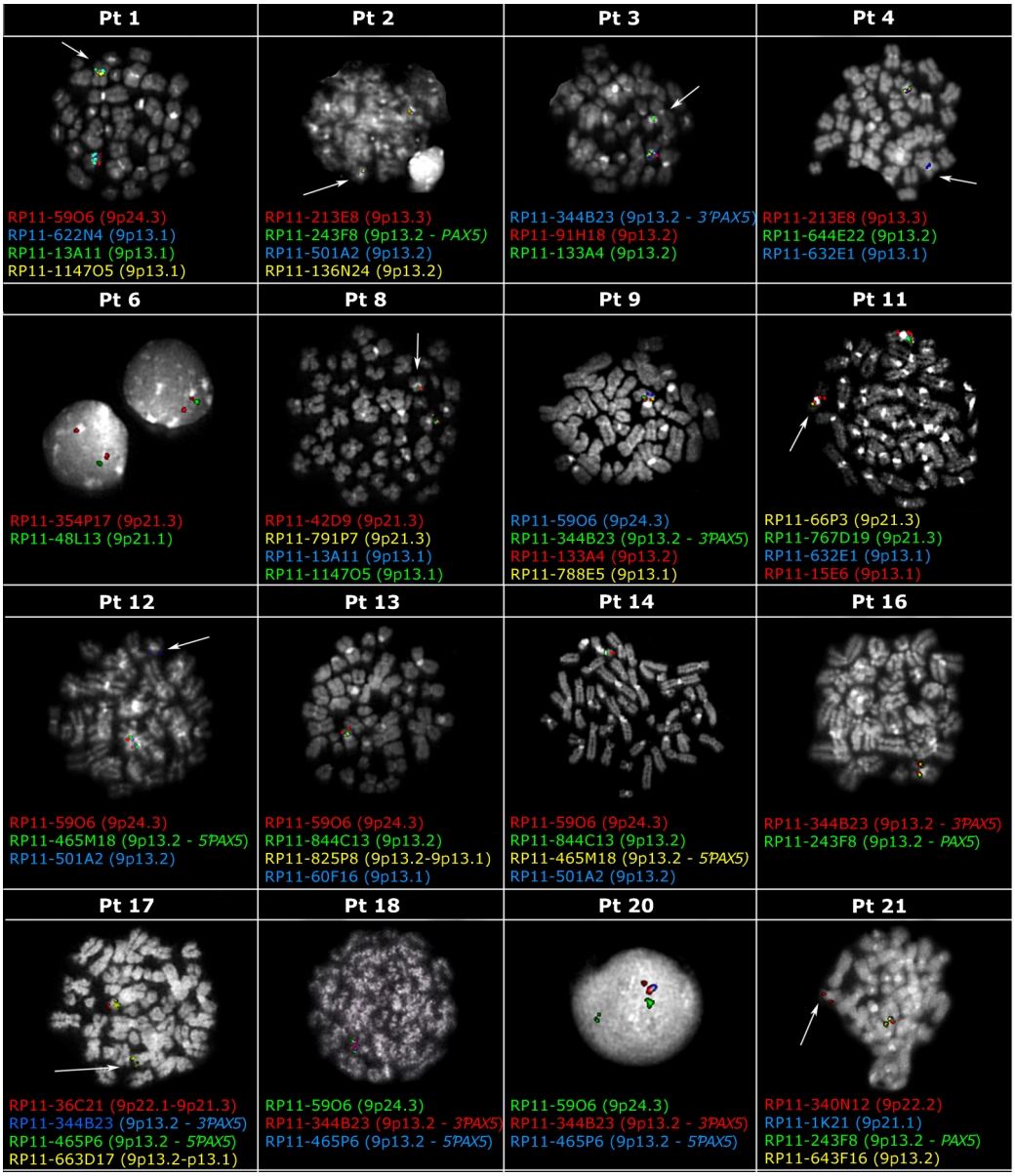
14 Bousquet, M. *et al.* A novel PAX5-ELN fusion protein identified in B-cell acute lymphoblastic leukemia acts as a dominant negative on the wild type PAX5. *Blood* **109**, 3417-3423 (2007).
15 Denk, D., Bradtke, J., Konig, M. & Strehl, S. PAX5 fusion genes in t(7;9)(q11.2;p13) leukemia: a case report and review of the literature. *Mol Cytogenet* **7**, 13, doi:1755-8166-7-13 [pii]10.1186/1755-8166-7-13 (2014).
16 Coyaud, E. *et al.* PAX5-AUTS2 fusion resulting from t(7;9)(q11.2;p13.2) can now be classified as recurrent in B cell acute lymphoblastic leukemia. *Leuk Res* **34**, e323-325, doi:S0145-2126(10)00378-4 [pii]10.1016/j.leukres.2010.07.035 (2010).
17 Denk, D. *et al.* PAX5-AUTS2: a recurrent fusion gene in childhood B-cell precursor acute lymphoblastic leukemia. *Leuk Res* **36**, e178-181, doi:S0145-2126(12)00190-7 [pii]10.1016/j.leukres.2012.04.015 (2012).
18 Storlazzi, C. T., Albano, F., Dencic-Fekete, M., Djordjevic, V. & Rocchi, M. Late-appearing pseudocentric fission event during chronic myeloid leukemia progression. *Cancer Genet Cytogenet* **174**, 61-67, doi:S0165-4608(06)00772-2 [pii]10.1016/j.cancergencyto.2006.11.009 (2007).
19 Chomczynski, P. & Sacchi, N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* **162**, 156-159, doi:10.1006/abio.1987.99990003-2697(87)90021-2 [pii] (1987).
20 Livak, K. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-408 (2001).
21 Kim, J. M. *et al.* CHFR functions as a ubiquitin ligase for HLTF to regulate its stability and functions. *Biochem Biophys Res Commun* **395**, 515-520, doi:S0006-291X(10)00721-7 [pii]10.1016/j.bbrc.2010.04.052 (2010).
22 Lai, T. *et al.* SOX5 controls the sequential generation of distinct corticofugal neuron subtypes. *Neuron* **57**, 232-247, doi:S0896-6273(08)00026-3 [pii]10.1016/j.neuron.2007.12.023 (2008).
23 Storlazzi, C. T. *et al.* Upregulation of the SOX5 by promoter swapping with the P2RY8 gene in primary splenic follicular lymphoma. *Leukemia* **21**, 2221-2225, doi:2404784 [pii]10.1038/sj.leu.2404784 (2007).
24 Mullighan, C. G. *et al.* Rearrangement of CRLF2 in B-progenitor- and Down syndrome-associated acute lymphoblastic leukemia. *Nat Genet* **41**, 1243-1246, doi:ng.469 [pii]10.1038/ng.469 (2009).
25 Russell, L. J. *et al.* Deregulated expression of cytokine receptor gene, CRLF2, is involved in lymphoid transformation in B-cell precursor acute lymphoblastic leukemia. *Blood* **114**, 2688-2698, doi:blood-2009-03-208397 [pii]10.1182/blood-2009-03-208397 (2009).
26 Palmi, C. *et al.* Poor prognosis for P2RY8-CRLF2 fusion but not for CRLF2 over-expression in children with intermediate risk B-cell precursor acute lymphoblastic leukemia. *Leukemia* **26**, 2245-2253, doi:leu2012101 [pii]10.1038/leu.2012.101 (2012).
27 Funakoshi, T. *et al.* Two distinct human POM121 genes: requirement for the formation of nuclear pore complexes. *FEBS Lett* **581**, 4910-4916, doi:S0014-5793(07)01003-4 [pii]10.1016/j.febslet.2007.09.021 (2007).

28 Coyaud, E. *et al.* Wide diversity of PAX5 alterations in B-ALL: a Groupe Francophone de Cytogenetique Hematologique study. *Blood* **115**, 3089-3097, doi:blood-2009-07-234229 [pii]10.1182/blood-2009-07-234229 (2010).
29 Merla, G., Ucla, C., Guipponi, M. & Reymond, A. Identification of additional transcripts in the Williams-Beuren syndrome critical region. *Hum Genet* **110**, 429-438, doi:10.1007/s00439-002-0710-x (2002).
30 Marschalek, R. Mixed lineage leukemia: roles in human malignancies and potential therapy. *FEBS J* **277**, 1822-1831, doi:EJB7608 [pii]10.1111/j.1742-4658.2010.07608.x (2010).
31 Familiades, J. *et al.* PAX5 mutations occur frequently in adult B-cell progenitor acute lymphoblastic leukemia and PAX5 haploinsufficiency is associated with BCR-ABL1 and TCF3-PBX1 fusion genes: a GRAALL study. *Leukemia* **23**, 1989-1998, doi:leu2009135 [pii]10.1038/leu.2009.135 (2009).
32 Bardini, M. *et al.* DNA copy-number abnormalities do not occur in infant ALL with t(4;11)/MLL-AF4. *Leukemia* **24**, 169-176, doi:leu2009203 [pii]10.1038/leu.2009.203 (2010).
33 Fazio, G. *et al.* PAX5/ETV6 alters the gene expression profile of precursor B cells with opposite dominant effect on endogenous PAX5. *Leukemia* **27**, 992-995, doi:leu2012281 [pii]10.1038/leu.2012.281 (2013).
34 Fortschegger, K., Anderl, S., Denk, D. & Strehl, S. Functional Heterogeneity of PAX5 Chimeras Reveals Insight for Leukemia Development. *Mol Cancer Res*, doi:1541-7786.MCR-13-0337 [pii]10.1158/1541-7786.MCR-13-0337 (2014).

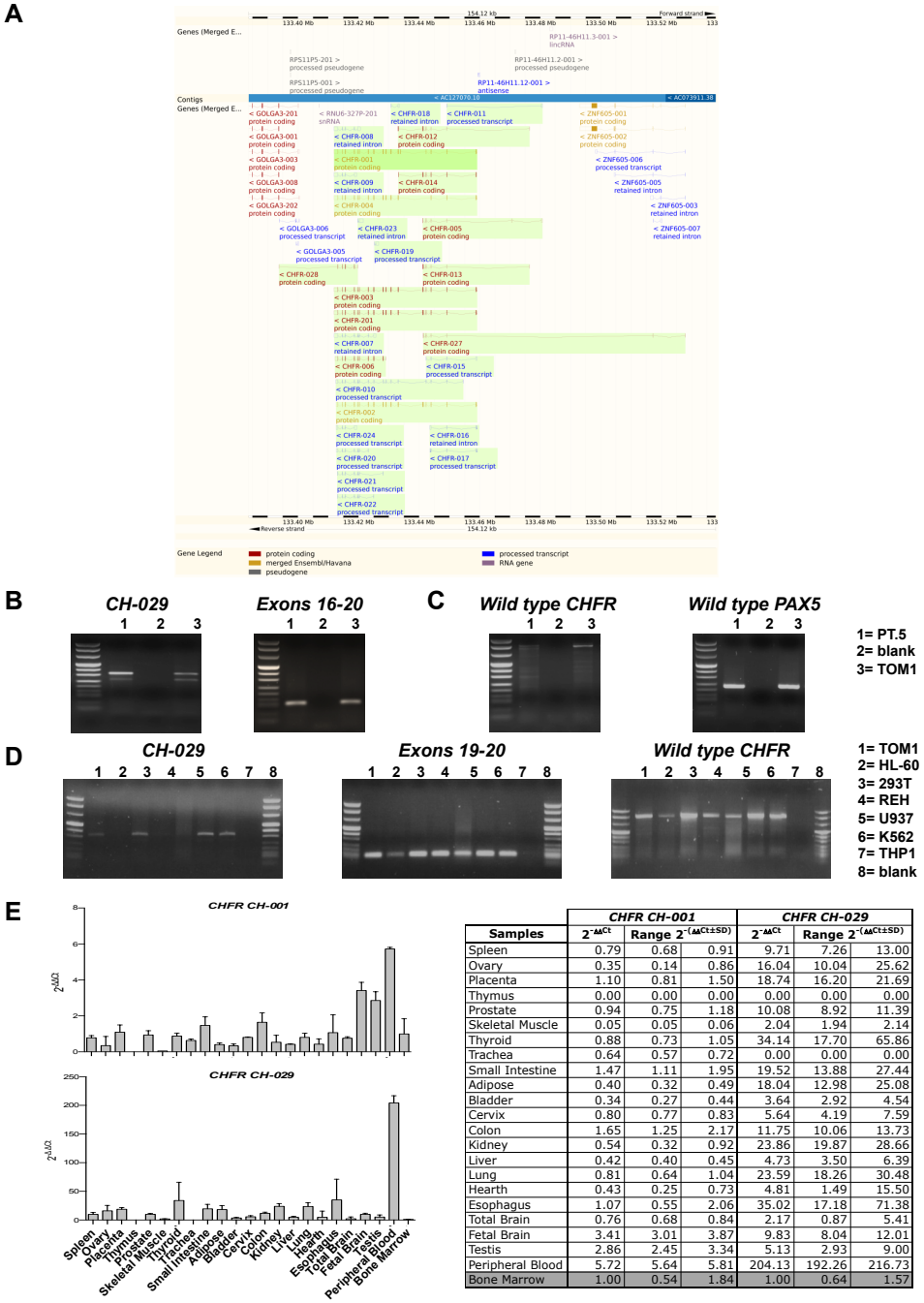
Pt	Gender	Age	WBC count*	Immunophenotype	Karyotype	*PAX5* Status (FISH)
1	M	4	nk	T-ALL	44,XY,der(1)t(1;5)(p32;q11.2),-5,del(6)(q?21),dic(9;12)(p13;p11),der(17)t(5;17)(p13;p11.2)	**deleted**
2	M	5	71,200	cALL	45,XY,del(6)(q?),der(9)t(9;?)(p13;?),-13,der(19)t(19;?)(p or q;?)[6]/46,XY[2]	**deleted**
3	M	5	42,060	cALL	45,XY,dic(9;20)(p13;q11)[15]/46,idem,+20[3]/46,XY[1]	**deleted**
4	M	7	65,490	cALL	46,XY,dic(9;17)(p13;p11),der(12)t(Y;12)(q1?;p1?),+ mar [9]	**deleted**
5	M	4	19,800	cALL	46,XY,t(9;12)(p13;q24.3)[5]	**traslocated-*PAX5/CHFR***
6	M	16	152,200	mat T-ALL	46,XY,der(9)?inv(9)(p13q13)dup(9)(p13p11)[5]/46,XY[2]	**deleted**
7	F	1	370,000	cALL	46,XX,der(7)t(7;11)(q11.2;q12),der(9)t(7;9)(q11.2;p13),-11,+ mar[10]/46,XX[2]	**traslocated-*PAX5/AUTS2***
8	M	15	13,240	pre-pre-B-ALL	47,XY,+8,del(9)(p?)	**deleted**
9	F	9	29,000	cALL	46,XX,del(9)(p12)[5]/46,XX[5]	**deleted**
10	M	3	1,240	pre-B-ALL	45,XY,-7,der(9)t(7;9)(q22;p22)[3]/46,XY[8]	**traslocated-*PAX5/AUTS2***
11	F	8	109,180	cALL – biclonal B	nk	**deleted**
12	M	1	81,370	pre-B ALL	nk	**deleted**
13	M	4	17,690	cALL	46,XY,dic(9;20)(p11;q11),+21[9]/47,idem,+22[4]	**deleted**
14	M	15	50,690	cALL	nk	**deleted**
15	M	10	124,000	cALL	45,XY,dic(9;12)(p13;p13)[6]/46,XY[3]	**traslocated-*PAX5/SOX5***
16	M	4	19,300	pre-pre-B / cALL	46-49,XY,+X,del(9)(p?) or dic(9;20),+21,+21[cp7]/46,XY[2]	**deleted**
17	M	1	123,000	cALL	46,XY,t(9;10)(q33;q23),del(9)(p12p21)[11]/46,XY[1]	**deleted**
18	M	13	147,000	pre-preB/cALL	46,XY,r(5),del(9)(p12),der(10),der(11)t(11;13)(q23;q12)t(13;22)(q14;q1?),der(13)t(11;13),add(17)(p1?),der(22)t(13;22)	**deleted**
19	M	1	8,040	cALL	nk	**traslocated-*PAX5/POM121C***
20	F	1	5,500	cALL	46,XX,t(2;11)(q10;p10),inv(7)(p15q11),del(9)(p22)[4]/46,XX[13]	**deleted**
21	M	4	10,800	cALL	46,XY,del(9)(p?)[2]/46,XY[7]	**deleted**
22	M	57	75,000	ALL	46,XY,t(9;22)(q34;q11)[4]/46,XX[6]	**deleted-*PAX5/MLLT3***

Table 1. Clinical features of patients carrying a *PAX5* alteration.
M, male; F, female; WBC count, number of white blood cells/μl; nk, not known.

SUPPLEMENTARY FIGURES

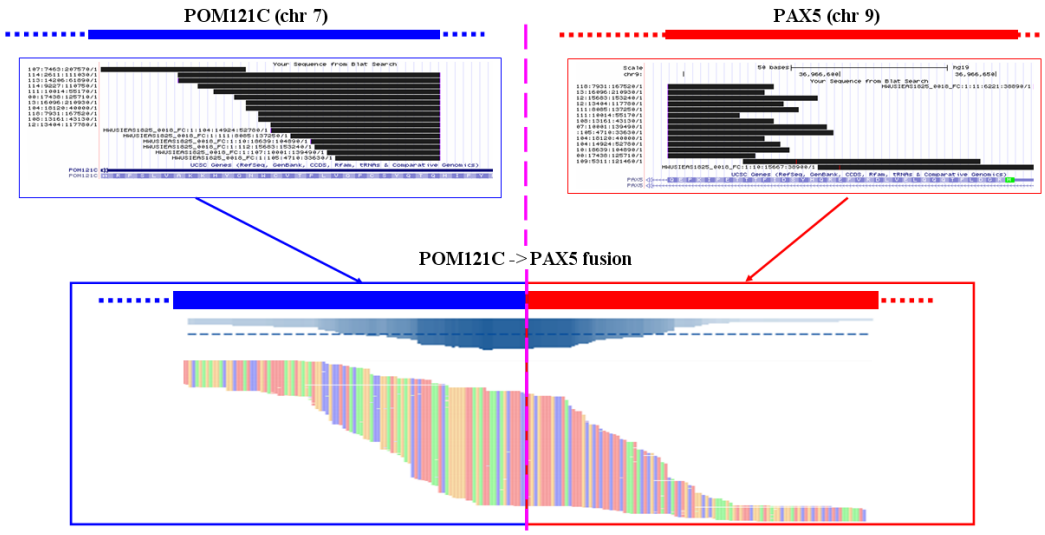


Supplementary Figure S1. FISH results obtained on 16 cases with probes identifying the deleted 9p segment including *PAX5*. BAC probes, shown below each box, are coloured as their corresponding signals in the merged image. The white arrows indicate the deleted chromosome 9 [del(9)].

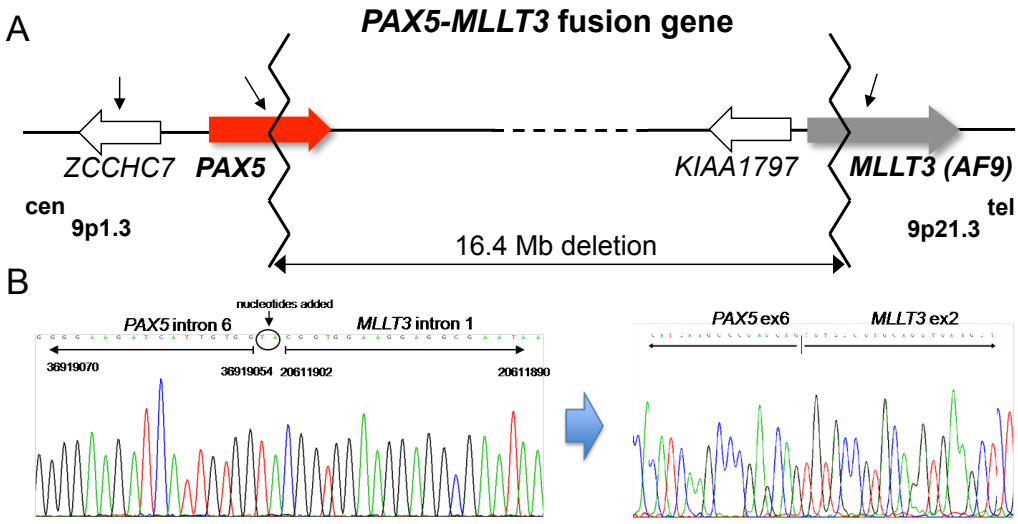


	CHFR CH-001			CHFR CH-029		
Samples	$2^{-\Delta\Delta Ct}$	Range $2^{-(\Delta\Delta Ct \pm SD)}$		$2^{-\Delta\Delta Ct}$	Range $2^{-(\Delta\Delta Ct \pm SD)}$	
Spleen	0.79	0.68	0.91	9.71	7.26	13.00
Ovary	0.35	0.14	0.86	16.04	10.04	25.62
Placenta	1.10	0.81	1.50	18.74	16.20	21.69
Thymus	0.00	0.00	0.00	0.00	0.00	0.00
Prostate	0.94	0.75	1.18	10.08	8.92	11.39
Skeletal Muscle	0.05	0.05	0.06	2.04	1.94	2.14
Thyroid	0.88	0.73	1.05	34.14	17.70	65.86
Trachea	0.64	0.57	0.72	0.00	0.00	0.00
Small Intestine	1.47	1.11	1.95	19.52	13.88	27.44
Adipose	0.40	0.32	0.49	18.04	12.98	25.08
Bladder	0.34	0.27	0.44	3.64	2.92	4.54
Cervix	0.80	0.77	0.83	5.64	4.19	7.59
Colon	1.65	1.25	2.17	11.75	10.06	13.73
Kidney	0.54	0.32	0.92	23.86	19.87	28.66
Liver	0.42	0.40	0.45	4.73	3.50	6.39
Lung	0.81	0.64	1.04	23.59	18.26	30.48
Hearth	0.43	0.25	0.73	4.81	1.49	15.50
Esophagus	1.07	0.55	2.06	35.02	17.18	71.38
Total Brain	0.76	0.68	0.84	2.17	0.87	5.41
Fetal Brain	3.41	3.01	3.87	9.83	8.04	12.01
Testis	2.86	2.45	3.34	5.13	2.93	9.00
Peripheral Blood	5.72	5.64	5.81	204.13	192.26	216.73
Bone Marrow	1.00	0.54	1.84	1.00	0.64	1.57

Supplementary Figure S2. *CHFR* novel RNA isoform *CH-029*, *exons 19-20*and *CH-00 1*in patient 5, cell lines and normal tissues (RT-PCR analysis). A- Alternative transcripts of the *CHFR* gene are reported, named as *CH-001* to *CH-028*, as shown in UCSC genome browser (assembly Feb. 2009, GRCh37/hg19). **B-** The new transcript *CH-029*(*left panel*)as well as the newly identified *exons 19-20*(*right panel*) are expressed both in patient n°5 (lane 1) and in TOM1 BCP-ALL control cell line (lane 3). **C**-*CH-001* (*left panel*)as well as the wild type *PAX5* (*right panel*), are expressed in patient 5 (lane 1) and in the TOM1 cells(lane 3).**D**- RT-PCR screening of *CH-029* expression in 293T (kidney fibroblasts, lane 3) and several haematological cell lines [TOM1, and REH (lanes 1 and 4, respectively; B-ALL); HL-60, and THP1 (lanes 2 and 7, respectively; AML); U937 (lane 5, histiocytic lymphoma); K562 (lane 6, CML). **E**-Tissue expression pattern of *CH-001* (dark gray, *upper panel*) and *CH-029* (light gray, *bottom panel*). Quantitative data are shown in the table.



Supplementary Figure S3.Schematic representation of *POM121C/PAX5* fusion gene by RNA-Seq analysis.Mapping of reads on hg 19 reference sequence (from UCSC genome browser) is shown in the upper part of the plot. Sequencing reads supporting the fusion are shown as mapping across the ad-hoc built reference of the fusion transcript. The blue track displays the depth of coverage across the junction.



C

Chr	Type	From	To	Size (kbp)	Marker Count	Genes	sno/miRNA
1	Loss	p31.1	p31.1	41	45	-----	-----
2	Loss	p11.2	p11.2	393	91	-----	-----
2	Loss	p22.3	p22.3	36	24	-----	-----
3	Loss	p21.31	p21.31	47	42	-----	-----
4	Gain	p15.1	p15.1	51	44	-----	-----
4	Loss	q13.2	q13.2	114	63	UGT2B17	-----
5	Gain	p15.2	p15.2	72	21	-----	-----
6	Loss	q16.3	q16.3	24	45	-----	-----
6	Loss	q14.1	q14.1	58	102	-----	-----
6	Loss	q12	q12	40	57	-----	-----
7	Loss	p12.2	p12.2	42	29	**IKZF1**	-----
8	Gain	p11.23	p11.22	152	58	ADAM3A, ADAM3A, ADAM3A, ADAM5P	-----
8	Loss	p23.1	p23.1	130	40	FAM86B2, DEFB109P1, FAM66A	-----
8	Loss	p23.1	p23.1	633	144	DEFB106A, DEFB106B, SPAG11B, SPAG11B, DEFB104B, DEFB105A, DEFB103A, SPAG11B, SPAG11B, DEFB107B, ...	-----
9	Loss	p13.2	p21.3	16209	11234	**from PAX5 to MLLT3**	hsa-mir-873, hsa-mir-876, hsa-mir-31, hsa-mir-491,
11	Loss	q11	q11	78	61	OR4C6, OR4P4, OR4S2	-----
14	Loss	q32.33	q32.33	757	323	LOC100133469, KIAA0125, ADAM6	-----
15	Loss	q21.1	q21.1	21	20	-----	-----
15	Loss	q11.2	q11.2	414	80	-----	-----
15	Loss	q11.2	q11.2	958	244	LOC646214, POTEB, OR4M2, LOC650137, OR4N4, CXADRP2, LOC727924, BCL8	hsa-mir-1268
15	Loss	q14	Q14	155	79	GOLGA8B, GOLGA8B, GOLGA8A	hsa-mir-1233
16	Gain	p11.2	p11.1	258	150	LOC283914, LOC283914, LOC146481	-----
18	Loss	q21.1	q21.1	25	34	SMAD2, SMAD2, SMAD2	-----
X	Gain	q21.31	q21.31	250	69	-----	-----
X	Gain	q21.32	q21.32	62	36	-----	-----
X	Gain	q21.31	q21.31	530	111	-----	-----
X	LOH	q11.1	Q28	92253	21562	CT47A5, CUL4B, DACH2, CT47A6, MAGEC3, TSPAN6, CXorf39, CXorf48, GLRA4, SLITRK2, ...	hsa-mir-1321, hsa-mir-514-3, hsa-mir-652, hsa-mir-892b, hsa-mir-325, hsa-mir-513b, hsa-mir-507, ...
X	LOH	p22.33	p11.1	55036	14761	RGN, TRAPPC2, GAGE12E, MAGED4, CLCN4, MAGEH1, UXT, HSD17B10, KDM6A, PIR, ...	hsa-mir-222, hsa-mir-660, hsa-mir-651, ACA12, hsa-mir-502, U107, ...

Supplementary Figure S4. *PAX5/MLLT3* fusion gene. A- Schematic representation of the 9p13 deletion, leading to the genesis of the *PAX5*/*MLLT3* fusion gene**. B-** Partial chromatograms showing the fusion junctions at both genomic (*left*) and RNA (*right*) level between *PAX5* and *MLLT3*.**C-** Copy Number Variation analysis performed in patient 22.

SUPPLEMENTARY TABLES

Primer RT-PCR	Sequence
PAX5-A (ex1)	ATGGATTTAGAGAAAAATTATCCGACT
PAX5-B (ex3)	GTGCCTAGCGTCAGTTCCAT
PAX5-R (ex5)	CTTGCGCTTGTTGGTGTCGG
UAP (RACE kit)	CUACUACUACUAGGCCACGCGTCGACTAGTAC
AUAP (RACE kit)	GGCCACGCGTCGACTAGTAC
AAP (RACE kit)	GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG
AUTS2-A (ex6)	GATGCCTTCGTCTCTCTTGC
AUTS2-B (ex7)	ACCACAGGCTCAAAGATTGG
AUTS2-C (ex7)	TACTCAGAGGTGGGGGTTGA
AUTS2-D (ex8)	TGAGTCTTCGCTGGAGTGC
SOX5ex5 - primer A	TTGTTCTTGTTGCTGCTTGG
SOX5ex7 - primer B	GGAGGGAATACGGGAATCAT
POM121Cex6R	CACCAGGGGCTCAAATGCTGA
POM121Cex13R	GCAGGCAGGGTAAAGGTAAATG
CHFR1829-L (exons 16-17)	CCGAGTTGCCAGTGGCCGTAA
CHFR2720-R (exon 18)	GGGCCGCGGTCACTCACAC
Sequence CH-029 - primer exon19L	GCACGCAGAATAGTGCTGGGCT
Sequence CH-029 - primer exon20R	ACGTTTCCTCGCTGCACGATAC
MLLT3 (ex5)	GTTCCTTGAAGGCCATCTTAGGAAC
RT-PCR product length	
PAX5-A + PAX5-R =594bp	
PAX5-B + PAX5-R =210bp	
PAX5-A + AUTS2-A = 609bp	
PAX5-A + AUTS2-B = 795bp	
PAX5-A + AUTS2-C = 929bp	
PAX5-A + AUTS2-D = 1161bp	
CHFR1829L + CHFR2720R =912bp	
CHFR1829L + primer XR =340bp	
Sequence CH-029: exon19L + exon20R=115bp	
PAX5-A + SOX5ex5 = 953bp	
PAX5-A + SOX5ex7 = 1066bp	
PAX5-B + SOX5ex5 = 569bp	
PAX5-B + SOX5ex7 = 682bp	
PAX5-A + POM121Cex6R = 759bp	
PAX5-B + MLLT3ex5 = 1078 bp	
Primer RQ-PCR	**Sequence**
Wild-Type_transcript_CH-001 CHFR_L	GTTTTCACAGCCCCCTGAG
Wild-Type_transcript_CH-001 CHFR_R	AGAAGAGTCACCCCAGAGCA
Novel_transcript_CH-029 CHFR_L	TCAAAAACTAAGCATCCAGAGG
Novel_transcript_CH-029 CHFR_R	GCGAGCCCAGCACTATTCT
GAPDH left	TGCACCACCAACTGCTTAGC
GAPDH right	GGCATGGACTGTGGTCATGAG

Supplementary Table S1. Primers used in RT-PCR in RQ-PCR analyses.

GENE	PROBE NAME	ACCESSION NUMBER	CHROMOSOME BAND	MAP POSITION (GRCh37/hg19)	FISH RESULTS				
					Pt 5	Pt 7	Pt 10	Pt 15	Pt 19
	RP11-59O6	ends	9p24.3	chr9:198,713-383,816	9, der(12)			9	
	RP11-450B8	ends	9p13.2	chr9:36,509,103-36,709,458				9	
3' PAX5	RP11-344B23	ends	9p13.2	chr9:36,774,163-36,949,832	9, der(12)	9, der(7)*		9	9
PAX5	RP11-243F8	ends	9p13.2	chr9:36,844,935-37,033,204	9, der(9)*, der(12)	9, der(9), der(7)*		9, del(9)	9, dic(7;9), der(9)
5' PAX5	RP11-465P6	ends	9p13.2	chr9:36,909,192-37,075,371	9, der(9), der(12)	9, der(9)		9, del(9)	
5' PAX5	RP11-465M18	ends	9p13.2	chr9:36,975,063-37,209,364	9, der(9), der(12)*				9, dic(7;9), der(9)
	RP11-133A4	ends	9p13.2	chr9:37,221,544-37,396,251			9, der(9)		
	RP11-713A20**	ends	7p22.3	chr7:13,709-203,581					
	RP11-1133D5**	ends	7p22.3	chr7:1,135,418-1,297,910					
	RP11-414M15**	ends	7p22.3	chr7:2,072,278-2,274,856					
	RP11-964D13**	ends	7p22.2	chr7:2,862,624-3,048,184					
	RP11-3F6**	ends	7p22.2	chr7:3,540,892-3,709,123					
	RP11-135H12**	ends	7p22.2	chr7:3,640,957-3,800,970					
	RP11-117I8**	ends	7p22.2	chr7:3,692,526-3,867,016					
	RP11-42B7**	ends	7p22.2	chr7:4,159,942-4,322,618					
	RP11-93G19**	ends	7p22.1	chr7:5,552,812-5,706,061					7, dic(7;9), der(?)
	RP11-269M4**	ends	7p15.3	chr7:20,910,403-21,076,176					
	RP11-465M2**	ends	7p15.3	chr7:21,253,297-21,448,738					
	RP11-1132K14**	ends	7p15.2	chr7:27,146,581-27,303,174					
	RP11-147A4**	ends	7p14.1	chr7:37,337,348-37,517,010					
	RP11-107G20**	ends	7p14.1	chr7:38,539,793-38,694,006					
	RP11-177B1**	ends	7p14.1	chr7:41,277,099-41,446,653					
	RP11-781C22**	ends	7p11.2	chr7:54,987,970-55,164,028					
	RP11-433C12**	ends	7p11.2	chr7:55,082,305-55,259,309					
	RP11-771A19**	ends	7p11.2	chr7:55,153,402-55,342,655					
	pZ7.5								7, dic(7;9)
	RP11-958B20	ends	7q11.23	chr7:74,564,923-74,912,389					7, dic(7;9)
POM121C	RP11-1054F23	ends	7q11.23	chr7:74,924,365-75,079,589					7, dic(7;9), der(9)*
AUTS2	RP11-409E13	ends	7q11.22	chr7:69,113,431-69,267,822		7, der(7)			
AUTS2	RP11-168B15	ends	7q11.22	chr7:69,470,146-69,653,704		7, der(7)			
AUTS2	RP11-666K18	ends	7q11.22	chr7:70,037,637-70,232,635		7, der(7), der(9)*			
3' AUTS2	RP11-485M20	ends	7q11.22	chr7:70,197,568-70,373,244		7, der(7), der(9)			
	RP11-867L5	ends	7q36.3	chr7:158,840,095-159,032,247			7, der(9)		
	RP11-636N1	ends	12p13.33	chr12:1,896,751-2,073,841				12	
	RP11-639O1	ends	12p13.2	chr12:11,737,815-11,916,241				12	
	RP11-418C2	ends	12p13.2	chr12:11,916,329-12,086,277				12	
	RP11-502N13	AC008114.25	12p13.1	chr12:14,632,639-14,757,140				12	
	RP11-1018J8	ends	12p12.3	chr12:15,158,390-15,370,563				12	
	RP11-489N6	ends	12p12.3	chr12:16,193,015-16,279,962				12	
	RP11-69C13	ends	12p12.3	chr12:16,748,391-16,895,630				12	
	RP11-871F6	ends	12p12.3	chr12:17,532,546-17,749,334				12	
	RP11-1144H2	ends	12p12.3	chr12:19,415,511-19,596,445				12	
	RP11-3F5	ends	12p12.3	chr12:19,509,245-19,698,514				12	
	RP11-352D19	ends	12p12.3	chr12:19,596,473-19,774,537				12	
	RP11-145J1	ends	12p12.3	chr12:19,654,367-19,806,117				12	
	RP11-412N19	ends	12p12.3	chr12:19,713,852-19,857,192				12	
	RP11-1006K22	ends	12p12.3	chr12:19,819,709-20,011,494				12	

	RP11-138F13	ends	12p12.2	chr12:20,029,484-20,186,763				12	
	RP11-12D15	ends	12p12.1	chr12:22,319,120-22,478,292				12	
	RP11-136L21	ends	12p12.1	chr12:22,589,365-22,771,823				12	
	RP11-449P1	ends	12p12.1	chr12:22,786,727-22,949,390				12	
	RP11-347F21	ends	12p12.1	chr12:23,052,014-23,215,093				12	
	RP11-1082A20	ends	12p12.1	chr12:23,287,922-23,485,440				12	
	RP11-760E23	ends	12p12.1	chr12:23,497,977-23,680,073				12	
	G248P80439E6	ends	12p12.1	chr12:23,634,673-23,674,964				12	
	G248P88992C6	ends	12p12.1	chr12:23,655,369-23,694,733				12	
3' SOX5	RP11-98J14	ends	12p12.1	chr12:23,671,226-23,822,046				12, dic(9;12)	
SOX5	RP11-1152G11	ends	12p12.1	chr12:23,803,421-23,946,124				12, dic(9;12)	
SOX5	G248P82399C2	ends	12p12.1	chr12:23,929,057-23,970,777				12, dic(9;12)	
SOX5	RP11-672O12	ends	12p12.1	chr12:23,946,131-24,126,509				12	
SOX5	RP11-162F10	ends	12p12.1	chr12:24,012,390-24,167,969				12	
SOX5	RP11-681A17	AC090675.10	12p12.1	chr12:24,132,898-24,304,940				12	
SOX5	G248P84907B3	ends	12p12.1	chr12:24,285,277-24,327,630				12	
SOX5	RP11-628A8	ends	12p12.1	chr12:24,304,948-24,469,455				12, dic(9;12)	
SOX5	RP11-77H23	ends	12p12.1	chr12:24,459,186-24,647,483				12, dic(9;12)	
5' SOX5	RP11-958I6	ends	12p12.1	chr12:24,647,516-24,831,923				12, dic(9;12)	
	RP11-485K18	AC024939.24	12p11.22	chr12:28,518,058-28,576,560				12, dic(9;12)	
	RP11-313F23	AC140847.4	12p11.1	chr12:34,267,708-34,400,580				12, dic(9;12)	
	RP11-450O16	ends	12q24.12-q24.13	chr12:112,119,304-112,311,122	12, der(12)				
	RP11-1128D1	ends	12q24.13	chr12:112,799,929-112,950,786	12, der(12)				
	RP11-136L23	ends	12q24.23	chr12:118,479,911-118,640,443	12, der(12)				
	RP11-80G18	ends	12q24.23	chr12:120,013,477-120,190,144	12, der(12)				
	RP11-1081C6	ends	12q24.31	chr12:121,535,680-121,718,399	12, der(12)				
	RP11-90B6	ends	12q24.31	chr12:123,224,657-123,387,601	12, der(12)				
	RP11-161A18	ends	12q24.31	chr12:125,294,205-125,465,944	12, der(12)				
	RP11-886H4	ends	12q24.32	chr12:128,965,212-129,117,533	12, der(12)				
	RP11-60M19	ends	12q24.32	chr12:129,946,465-130,123,216	12, der(12)				
	RP11-72E15	ends	12q24.33	chr12:130,907,263-131,047,343	12, der(12)				
	RP11-825F9	ends	12q24.33	chr12:131,117,968-131,301,835	12, der(12)				
	RP11-270L16	ends	12q24.33	chr12:132,069,545-132,238,137	12, der(12)				
	RP11-357I22	ends	12q24.33	chr12:132,856,378-133,065,041	12, der(12)				
POLE	RP11-148L11	ends	12q24.33	chr12:133,038,915-133,243,361	12, der(12)				
POLE	G248P80226B8	ends	12q24.33	chr12:133,238,813-133,281,615	12, der(12), der(9)				
	G248P86040D7	ends	12q24.33	chr12:133,270,749-133,308,916	12, der(9)				
	G248P83013G3	ends	12q24.33	chr12:133,332,740-133,372,365	12, der(9)				
	RP11-46H11	AC127070.10	12q24.33	chr12:133,347,391-133,526,403	12, der(9)				
	RP11-637F20	ends	12q24.33	chr12:133,465,282-133,613,019	12, der(9)				

Supplementary Table S2. BAC used for FISH analysis to characterize the *PAX5* deleted cases.

*Signal intensity on the derivative chromosome was significantly fainter than the one on the normal homolog;

** BACs used as a pool probe.

GENE	BAC NAME	ACCESSION NUMBER	CHROMOSOME BAND	MAP POSITION (GRCh37/hg19)	PT 1	2	3	4	6	8	9	10	11	12	13	14	16	17	18	20	21
DOCK8	RP11-59O6	ends	9p24.3	chr9:198,713-383,816	9	9	9	9	9; del(9)	9; del(9)	9	9	9; del(9)	9	9	9	9	9	9	9; del(9)	9; del(9)
	RP11-10G22	ends	9p24.3	chr9:1,533,223-1,709,628														9		9	
PTPRD	RP11-338L20	ends	9p24.1-9p23	chr9:8,981,678-9,142,717					9; del(9)	9; del(9)			9; del(9)								
	RP11-298E19	ends	9p22.3	chr9:14,310,338-14,474,348														9		9	
CNTLN	RP11-340N12	ends	9p22.2	chr9:17,146,369-17,308,494														9		9	9; del(9)
NFIB	RP11-296P7	ends	9p22.1	chr9:19,026,890-19,028,127					9; del(9)	9; del(9)			9; del(9)								
DENND4C	RP11-101L23	ends	9p22.1	chr9:19,142,752-19,339,389														9		9	
ACER2; SLC24A2	RP11-361B3	ends	9p22.1	chr9:19,390,694-19,576,389						9; del(9)											
	RP11-36C21	ends	9p22.1-9p21.3	chr9:19,818,493-19,989,520						9; del(9)								9		9	
	RP11-410F21	ends	9p21.3	chr9:20,143,896-20,316,463																	9; del(9)
MLLT3	RP11-1100C5	ends	9p21.3	chr9:20,422,666-20,591,558						9; del(9)			9; del(9)								
MLLT3; FOCAD	RP11-42D9	ends	9p21.3	chr9:20,570,274-20,746,915						9; del(9)*			9; del(9)								
FOCAD	RP11-791P7	ends	9p21.3	chr9:20,663,654-20,851,416						9											
FOCAD	RP11-66P3	ends	9p21.3	chr9:20,784,480-20,945,583						9			9; del(9)								
FOCAD; PTPLAD2	RP11-767D19	ends	9p21.3	chr9:20,917,109-21,086,861						9			9					9			9
	RP11-354P17	ends	9p21.3	chr9:21,343,280-21,536,271					9; del(9)	9			9								
MTAP; CDKN2A	RP11-615P15	ends	9p21.3	chr9:21,764,403-21,981,385					9												
DMRTA1	RP11-624G3	ends	9p21.3	chr9:22,304,545-22,471,940									9								
	RP11-12N11	ends	9p21.3	chr9:23,436,448-23,597,889																	
ELAVL2	RP11-110I5	ends	9p21.3	chr9:23,688,378-23,854,636																	9
ELAVL2	RP11-31K16	AL161628.9	9p21.3	chr9:23,722,809-23,833,246					9				9								
	RP11-477G9	AL353753.6	9p21.2	chr9:25,920,470-26,099,334					9												
	RP11-48L13	AL354676.10	9p21.1	chr9:29,577,130-29,649,069					9												
TAF1L; TMEM215	RP11-1K21	ends	9p21.1	chr9:32,629,173-32,804,393					9	9											9
	RP11-573M23	AL356494.13	9p13.3	chr9:34,335,596-34,417,345					9	9			9								
	RP11-145E14	ends	9p13.3	chr9:35,574,708-35,746,923					9	9			9								
	RP11-213E8	ends	9p13.3	chr9:35,751,414-35,902,246	9	9	9	9	9	9											
RECK	RP11-626F21	ends	9p13.3	chr9:36,020,455-36,184,946	9	9	9	9	9	9			9								
RNF38	RP11-844C13	ends	9p13.2	chr9:36,331,767-36,515,625								9	9	9	9	9					
	RP11-450B8	ends	9p13.2	chr9:36,509,103-36,709,458	9	9	9	9	9	9		9									
3' PAX5	RP11-344B23	ends	9p13.2	chr9:36,774,163-36,949,832	9	9	9	9	9	9	9	9					9	9	9	9	
PAX5	RP11-243F8	ends	9p13.2	chr9:36,844,935-37,033,204	9	9	9	9	9	9	9	9; del(9)*					9	9; del(9)*		9	9
5' PAX5	RP11-465P6	ends	9p13.2	chr9:36,909,192-37,075,371	9	9	9	9	9	9	9	9; del(9)*						9; del(9)*	9	9	
5' PAX5	RP11-465M18	ends	9p13.2	chr9:36,975,063-37,209,364			9				9	9; del(9)	9	9	9	9		9; del(9)			
5' PAX5	RP11-652D9	ends	9p13.2	chr9:36,986,566-37,189,379	9	9	9	9	9	9	9	9; del(9)		9							
ZCCHC7	RP11-91H18	ends	9p13.2	chr9:37,138,096-37,318,047			9					9; del(9)		9					9		
ZCCHC7	RP11-133A4	ends	9p13.2	chr9:37,221,544-37,396,251			9; del(9)*				9	9; del(9)		9; del(9)*		del					
	RP11-501A2	ends	9p13.2	chr9:37,392,581-37,596,054	9	9	9; del(9)	9	9	9	9	9; del(9)	9	9; del(9)	9	del					
	RP11-643F16	ends	9p13.2	chr9:37,628,199-37,777,278							9					del		9; del(9)			9
	RP11-644E22	ends	9p13.2	chr9:37,777,285-37,950,151	9	9	9; del(9)	9	9	9	9					del					
SHB	RP11-666G2	ends	9p13.2	chr9:37,945,191-38,138,311		9							9		9	9; del(9)					
	RP11-136N24	ends	9p13.2	chr9:38,090,149-38,279,495		9; del(9)		9													
	RP11-663D17	ends	9p13.2-9p13.1	chr9:38,267,084-38,441,277							9							9; del(9)	9		9
	RP11-825P8	ends	9p13.2-9p13.1	chr9:38,291,105-38,491,094		9; del(9)		9					9		9	9; del(9)					
	RP11-1050J3	ends	9p13.1	chr9:38,444,223-38,618,809				9; del(9)*			9		9								
	RP11-632E1	ends	9p13.1	chr9:38,512,027-38,684,880	9	9; del(9)		9; del(9)		9	9		9								
	RP11-788E5	ends	9p13.1	chr9:38,568,002-38,733,846	9	9; del(9)	9; del(9)	9; del(9)	9	9	9										
	RP11-15E6	ends	9p13.1	chr9:38,733,908-38,890,388	9								9; del(9)		9	9; del(9)					
	RP11-622N4	ends	9p13.1	chr9:38,797,151-38,978,531	9		9; del(9)		9	9			9; del(9)								
	RP11-13A11	ends	9p13.1	chr9:38,798,813-38,973,416	9; del(9)*				9	9			9		9	9; del(9)					
	RP11-1147O5	ends	9p13.1	chr9:38,890,456-39,046,992	9; del(9)				9	9; del(9)*			9		9						
	RP11-1150N16	ends	9p13.1	chr9:38,995,367-39,128,143					9	9; del(9)			9		9						
	RP11-60F16	ends	9p13.1	chr9:38,978,534-39,150,161					9	9; del(9)			9		9						

Supplementary Table S3. BAC used for FISH analysis to characterize the *PAX5* translocated cases.
*signal intensity on the deleted chromosome 9 [del(9)] was significantly fainter than the one on the normal homolog

PT5 PAX5/CHFR					
Chr	**Type**	**From**	**To**	**Genes**	**sno/miRNA**
2	Loss	p11.2	p11.2	**IGKC**	-----
2	Loss	p11.2	p11.2	**IGKV1D-27**	-----
5	LOH	q23.3	q31.1	SLC27A6, ISOC1, ADAMTS19, CHSY3, HINT1, LYRM7, CDC42SE2, RAPGEF6, FNIP1, ACSL6	-----
7	Mosaicism/Gain	q11.22	q34	From WBSCR17 to MTRNR2L6	hsa-mir-4284, hsa-mir-590, hsa-mir-1285-1, hsa-mir-653, hsa-mir-489, hsa-mir-591, hsa-mir-25, hsa-mir-93, hsa-mir-106b, hsa-mir-4285, hsa-mir-5480, hsa-mir-592, hsa-mir-593, hsa-mir-129-1, hsa-mir-182, hsa-mir-96, hsa-mir-183, hsa-mir-335, hsa-mir-29a, hsa-mir-29b-1, hsa-mir-490
7	Mosaicism/Loss	q34	qter	PRSS2, PRSS1, TRY6 to telomer (about 16MB deletion)	hsa-mir-548f-4, hsa-mir-1975, hsa-mir-671, hsa-mir-153-2, hsa mir-595
7	Mosaicism/Loss	pter	p14.2	TARP, TARP deletion from TARP included (only TARP in homozygosis) to telomer (about 38MB deletion)	hsa-mir-339, hsa-mir-589, hsa-mir-1302-6, hsa-mir-3146, hsa-mir-1183, HBII-336, hsa-mir-148a, hsa-mir-196b, hsa-mir-550-1, hsa-mir-550-2, hsa-mir-548n
14	Loss	q11.2	q11.2	**TCRA**	

PT7 PAX5/AUTS2					
Chr	**Type**	**From**	**To**	**Genes**	**sno/miRNA**
8	Gain	q24.21	q24.21	-----	-----
9	Loss	pter	p13.2	**From pter to *PAX5*. CDKN2A homozygously deleted.**	hsa-mir-101-2, hsa-mir-3152, U92, hsa-mir-491, hsa-mir-31, hsa-mir-876, hsa-mir-873, SNORD121B, SNORD121A
14	Loss	q11.2	q11.2	**TCRA**	-----

PT10 PAX5/AUTS2					
Chr	**Type**	**from**	**to**	**Genes**	**sno/miRNA**
3	Loss	q26.32	q26.32	TBL1XR1	-----
7	Loss	pter	q11.22	**from PDE1C to AUTS2 (IKZF1 included)**	-----
9	Loss	pter	p13.2	**from DOCK8 to PAX5 (CDKN2A and MLLT3 included)**	-----
14	Loss	q11.2	q11.2	**TCRA**	-----

PT15 PAX5/SOX5					
Chr	**Type**	**From**	**to**	**Genes**	**sno/miRNA**
7	Loss	p14.1	p14.1	TARP	-----
8	Loss	q12.1	q12.1	-----	SNORA51
9	Loss	pter	p13.2	**From DMRT3 to PAX5. Homozygously deleted from MLLT3 to CDKN2B-AS, position 19.932.623-22.046.390. MOBKL2B homozygously deleted.**	hsa-mir-101-2, hsa-mir-3152, U92, hsa-mir-491, hsa-mir-31, hsa-mir-876, hsa-mir-873, SNORD121B, SNORD121A
12	Loss	pter	p12.1	**From FAM138D to SOX5 (TEL deletion included).** Except copy number=2 in region p12.1, from 23.702.993 to 23.957.694)	hsa-mir-613, hsa-mir-614
14	Loss	q11.2	q11.2	TCRA, TCRD,TCRVD2,TRA@... Homozygously deleted	-----

PT19 PAX5/POM121C					
Chr	**Type**	**From**	**To**	**Genes**	**sno/miRNA**
2	loss	q14.3	q14.3	CNTNAP5	-----
2	loss	q21.1	q21.1	from CFC1 to MCM6	-----
2	loss	q34	q36.3	from LANCL1 to TRIP12	hsa-mir-548f-2, hsa-mir-26b, hsa-mir-375, hsa-mir-3131, hsa-mir-153-1, hsa-mir-3132, hsa-mir-4268
6	loss	q22.1	q22.1	-----	-----
7	loss	p21.2	p21.2	DGKB, TMEM195	-----
7	Loss	p14.1	p14.1	TARP	-----
7	Loss	q11.23	qter	**From POM121C to qter**	ACA14a, hsa-mir-1285-1, hsa-mir-653, hsa-mir-489, hsa-mir-591, hsa-mir-25, hsa-mir-93, hsa-mir-106b, hsa-mir-4285, hsa-mir-548o, hsa-mir-592, hsa-mir-593, hsa-mir-129-1, hsa-mir-182, hsa-mir-96, hsa-mir-183, hsa-mir-335, hsa-mir-29a, hsa-mir-29b-1, hsa-mir-490, hsa-mir-548f-4, hsa-mir-1975, hsa-mir 671
9	loss	pter	p13.2	**From PAX5 to pter. Homozygously deleted at CDKN2A, CDKN2B, CDKN2B-AS, DMRTA1, ELAVL2, TUSC1. Homozygously deleted from 26409648 to 28005524, no genes.**	hsa-mir-101-2, hsa-mir-3152, U92, hsa-mir-491, hsa-mir-31, hsa-mir-876, hsa-mir-873, SNORD121B, SNORD121A
9	Gain	p13.2	q32	From PAX5 to ZFP37	hsa-mir-1299, hsa-mir-204, hsa-mir-7-1, hsa-mir-4289, hsa-mir-3153, hsa-mir-4290, SNORA84, hsa-mir-4291, hsa-let-7a-1, hsa-let-7f-1, hsa-let-7d, hsa-mir-2278, hsa-mir-23b, hsa-mir-27b, hsa-mir-3074, hsa-mir-24-1, hsa-mir-1302-8, hsa-mir-32
14	loss	q11.2	q11.2	TCRA	-----

Supplementary Table *S4*. Copy Number Variation analysis of the *PAX5* rearranged cases.

SUPPLEMENTARY METHODS

FISH Analysis. Briefly, chromosome preparations from bone marrow cells were hybridized in situ with 1 μg of each Bacterial Artificial Chromosome (BAC)/fosmid probe labelled by nick translation. Hybridization was performed at 37°C in 2X SSC, 50% (vol/vol) formamide, 10% (wt/vol) dextran sulfate, 5 μg COT1 DNA (Bethesda Research Laboratories, Gaithersburg, MD, USA), and 3 μg sonicated salmon sperm DNA in a volume of 10 μL. Post-hybridization washings were performed at 60°C in 0.1X SSC (three times). In co-hybridization experiments, the probes were directly labeled with Fluorescein, Cy3 and Cy5. Chromosomes were identified by DAPI staining. Digital images were obtained using a Leica DMRXA epifluorescence microscope equipped with a cooled CCD camera (Princeton Instruments, Boston, MA). Cy3 (red; New England Nuclear, Boston, MA, USA), fluorescein (green; Fermentas Life Sciences, Milan, IT), Cy5 (IR; New England Nuclear, Boston, MA, USA) and DAPI (blue) fluorescence signals, which were detected using specific filters, were recorded separately as gray-scale images. Pseudocoloring and merging of images were performed with Adobe Photoshop software.

RNA-Seq Analysis Total RNA was checked for integrity and purity by capillary electrophoresis (Agilent Bioanalyser 2100, Agilent Technologies, Santa Clara, CA, USA), requiring a RNA Integrity Number of 8 or more. 2 μg RNA was used to prepare a double-stranded cDNA library for high-throughput sequencing by Truseq RNA kit (Illumina, San Diego, CA, USA). 6 pM of the library was loaded in one lane of an Illumina flowcell, subjected to cluster bridge amplification in a cBot station and paired-end sequenced on a GAIIx instrument with a 2 × 76 protocol (Illumina pipeline version 1.6). After sequencing run completion, clusters were converted into reads by using Illumina's proprietary software OLB (Offline BaseCaller, v.1.9.0). Raw reads were mapped against the University of California Santa Clara (UCSC) human genome sequence, release 19 (hg19) by using BWA (v. 0.5.9, allowing up to 3 mismatches in the alignment and a trimming of bases having a phred quality below 15, Li H and Durbin R, 2009). Not perfectly mapped or trimmed read pairs (i.e.: those having at least one unmapped mate, more than 7 trimmed bases, or more than 1000 nt between the two mates) were re-aligned against hg19 and RefSeq gene models (http://www.ncbi.nlm.nih.gov/refseq/)

by TopHat (v. 1.3.1, allowing up to 3 mismatches, Trapnell C et al, 2009). Reads having at least one unmapped pair, or characterized by a distance between the two mates above 1000 nucleotides were used as an input to FusionMap software (v. 6.2.1.41, using Human B37 and RefSeq as reference genome and transcriptome model, respectively, Ge H et al, 2011) to search for genomic rearrangements.

Candidate events derived from FusionMap were further filtered as follows:

a) We kept only those events having 3 or more non-clonal supporting reads

b) We excluded events eventually localized in repeated regions of the genome and specifically:

-all the records in FusionMap output sharing one of the breakpoints (even though they had a different mate);

-candidates having the whole junction region (i.e. 60 nt across the junction) completely matched anywhere in the reference genome (in order to compensate for transcripts whose structure was not considered in FusionMap). This was assessed by means of a Blat (Kent WJ, 2002) alignment against the hg19 reference;

-candidates having either of the two halves (i.e. length=30 nt) of the junction matching in multiple locations of the genome according to their Blat alignment.

c) We required coverage of at least 10 reads in the regions flanking the rearrangement, in order to exclude those regions not sufficiently covered by sequencing reads to have a reliable characterization. Sequencing coverage was calculated by using the utility samtools (v. 0.1.17, command mpileup, Li H et al. 2009);

d) We filtered out the events having an incidence of the reads supporting the fusion event below 10% of the total reads on either side.

Copy Number Variation analysis. This array provides dense coverage across the entire genome due to high number of markers, including 400,000 SNPs and 2.3 millions of non-polymorphic copy-number markers. Therefore, it enables to detect small aberrations (gains/losses) and copy number neutral loss of heterozygousity (LOH) regions. Briefly, 100ng of genomic DNA were amplified by an overnight whole-genome amplification reaction and purified by magnetic beads according to the manufacturer's instructions. The samples were then fragmented to generate small (<300 bp) products, which were subsequently biotin-labeled, denatured and loaded into the arrays. After

hybridization, the chips were washed, stained (streptavidin-PE), and scanned using the Gene Chip® Scanner 3000.

Supplementary References

Li H and Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler Transform. Bioinformatics, 25:1754-60

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G. Durbin R and 1000 Genome Project Data Processing Subgroup (2009) The Sequence alignment/map (SAM) format and SAMtools. Bioinformatics, 25, 2078-9

Ge H, Liu K, Juan T, Fang F, Newman M, Hoeck W. FusionMap: detecting fusion genes from next-generation sequencing data at base-pair resolution. Bioinformatics. 2011 Jul 15;27(14):1922-8

Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 2009 May 1;25(9):1105-11

Kent WJ. BLAT--the BLAST-like alignment tool. Genome Res. 2002 Apr;12(4):656-64

Chapter 3

PAX5/ETV6 alters the gene expression profile of precursor B-cells with an opposite dominant effect on endogenous PAX5

PAX5/ETV6 alters the gene expression profile of precursor B-cells with an opposite dominant effect on endogenous PAX5

Leukemia (2013) 27, 992-995; doi: 10.1038/leu.2012.281

Fazio Grazia PhD,[1,2] Cazzaniga Valeria MS,[1,2] Palmi Chiara PhD,[1,2] Galbiati Marta MS,[1] Giordan Marco PhD,[3] te Kronnie Geertruy PhD,[3] Rolink Antonius PhD,[4] Biondi Andrea MD,[1,2] Cazzaniga Giovanni PhD.[1]

[1]Centro Ricerca M. Tettamanti, Clinica Pediatrica, Università di Milano-Bicocca, Ospedale S. Gerardo, Monza, Italy; [2]Dipartimento di Medicina Clinica e Prevenzione, Università di Milano-Bicocca, Monza, Italy; [3]Laboratorio di Oncoematologia, Clinica Pediatrica, Università di Padova, Padova, Italy; [4]University of Basel, Basel, Switzerland.

Biondi Andrea and Cazzaniga Giovanni share the senior authorship of this paper.

Running title: PAX5/ETV6 is a repressor of the PAX5 pathway.

KEYWORDS: PAX5, ETV6, TEL, BCP-ALL

Corresponding author:
Cazzaniga Giovanni
Centro Ricerca M. Tettamanti
Clinica Pediatrica
Università di Milano-Bicocca
Ospedale S. Gerardo
Via Pergolesi 33
20900 – Monza – (MB) – Italy
gianni.cazzaniga@hsgerardo.org
phone: +390392333661; fax: +390392332167

The *PAX5* gene encodes for a transcription factor essential for B cell development,(1,2) which has been recently reported as frequent target of aberrancies, including mutations, deletions and translocations in approximately 30% of pediatric patients with B-cell precursor acute lymphoblastic leukemia (BCP-ALL).(3,4) Among several translocations with different partner genes, the t(9;12) encoding for the *PAX5/ETV6* fusion, is the most recurrent.(2)

The role of PAX5/ETV6 in leukemogenesis is not fully understood, although preliminary evidences suggested that the protein might have a repressor function on transcription.(3,5-6) The PAX5/ETV6 protein retains the ability to interact with the PAX5-consesus sequence contained in the CD19 promoter,(6) and it is able to multimerize and bind the PAX5-consensus sequence with high affinity, determining a *dominant negative* activity on wt PAX5.(7)

Herein, we explored the effect of PAX5/ETV6 on the transcription in precursor B-cells and whether the fusion gene specifically affects the PAX5 and ETV6 physiological pathways. We performed a genome-wide expression analysis in murine wild type cKIT+/B220+/CD19+ pre-BI cells (8) transduced by the retroviral MIGR-PAX5/ETV6 vector, as previously described.(5) We used three different cell populations, called LY5.1FL, FLB6-67 and B6BAFL cells, respectively, isolated from fetal liver of diverse mouse strain donors.

The Gene Expression Profile (GEP) was obtained in two independent experiments on B6BAFL pre-BI cells, while the validation experiments were carried out in all the three cell populations. Here we present the main findings while more detailed information is reported in the Supplementary Methods and Results file.

Statistical analysis selected 340 Differentially Expressed Genes (DEGs) over 12,761 probe-sets, which defined a PAX5/ETV6 specific signature (Figure 1A; Supplementary Table S1). By setting the fold change (FC) as FC≥1.5 or ≤0.7, we observed 245/340 (72%) down-regulated genes, whereas 28% (n=95/340) were up-regulated, thus suggesting that PAX5/ETV6 acts mainly, but not exclusively, as a transcriptional repressor. The expression of 49 transcripts, including 28 down-regulated and 21 up-regulated genes, was successfully validated by RQ-PCR (Supplementary Tables S2-S3). The down-regulation of genes was further confirmed at the protein level. Indeed, by protein arrays, Tumor Necrosis Factor alpha (TNFα) was decreased of 80% (t test,

p=0.0079) in B6BAFL cells (Figure 1A), consistent with reduced TNFα transcript (FC=0.46 in GEP). Interestingly, two additional genes from the TNF family, lymphotoxin A (LTA) and lymphotoxin B (LTB) were strongly repressed by PAX5/ETV6 (Supplementary Table S2). Moreover, we validated by RQ-PCR and FACS analyses the reduction of CD19 and three cell surface PAX5-targets, namely CD22, SLAMF6/LY108 and CD44 antigens (Supplementary Figures S1-S2-S3).

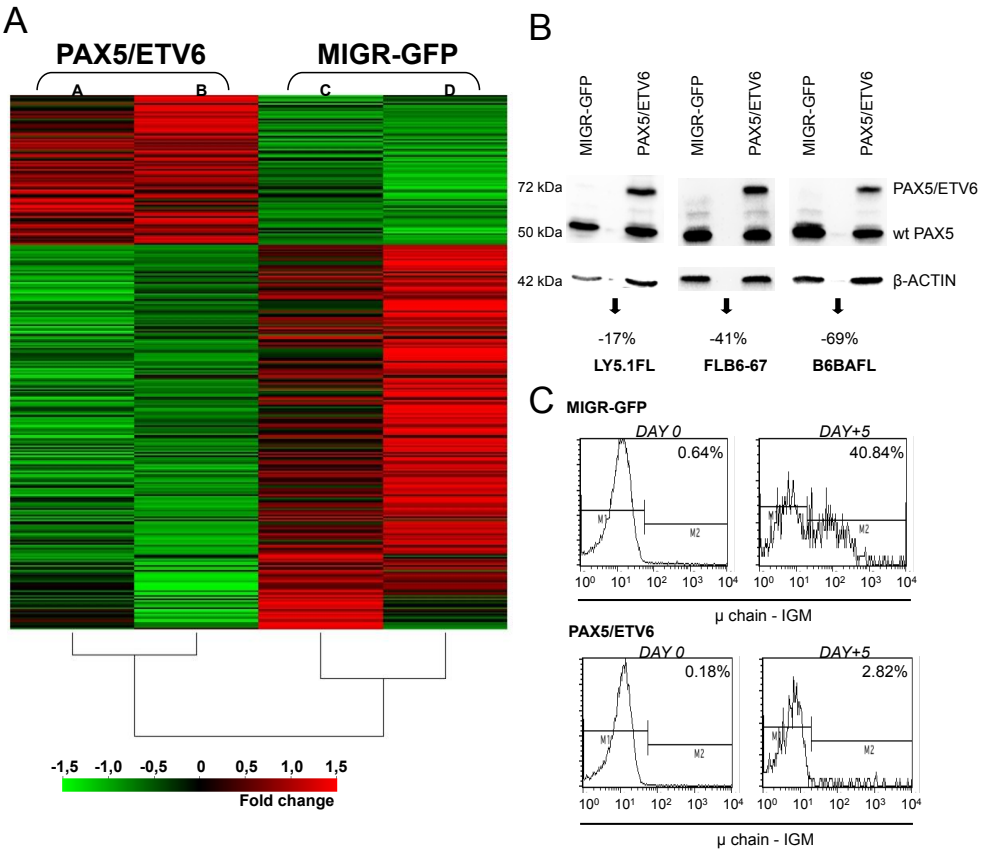


Figure 1. (A) Heatmap of Gene Expression Profile results. Two biological independent transduction-GEP experiments of PAX5/ETV6 (A and B) and MIGR-GFP as a control (C and D) were carried out in B6BAFL pre-BI cells (A vs. C, B vs. D), using the Affymetrix gene chip mouse array 430A 2.0. Differentially expressed genes were defined as *down-regulated* if [log2(signal)<-0.50 = Fold change <0.70] or *up-regulated* if [log2(signal)>0.59 = Fold change >1.5]. DEGs are graphically represented, in fact each colored line corresponded to a single probe/gene. **(B) PAX5/ETV6 and PAX5 protein expression in LY5.1FL, FLB6-67 and B6BAFL pre-BI cells.** Western blot analysis of PAX5/ETV6 (72kDa) and wild type PAX5 (50kDa) was performed using the anti-N-term PAX5(Chemicon). After stripping the same membrane, beta-actin (42kDa), as internal control, was detected. **(C) B-cell receptor signaling is impaired in PAX5/ETV6 pre-BI cells.** Histogram analyses of the cell surface m chain expression on CD19+ control MIGR-GFP and on CD19+ PAX5/ETV6 LY5.1FL pre-BI cells are shown. The cells were cultured in the presence of a sub-optimal concentration of IL-7 (0.5%) and were analyzed at two time points of culture (day 0, +5).

A consistent number of DEGs (n=34/340, 10%) are annotated to be direct transcriptional targets of PAX5 (9-12) (p=1.029E-09 by Fisher's exact test) (Supplementary Table S1 and Supplementary Methods). Indeed, the noteworthy p value confirms that PAX5/ETV6 especially affects the PAX5 transcriptional pathway.

In particular, for the first time we observed that PAX5/ETV6 exerts a dual regulation on wt PAX5 transcription function, by repressing 56% (19/34) and activating 44% (15/34) of PAX5-target genes, respectively (Supplementary Figure S4). In addition, about 63% (12/19) of PAX5-target genes repressed by PAX5/ETV6 are known to be normally activated by wt PAX5 whilst only 37% (7/19) are physiologically repressed by wt PAX5. Similarly, 73% (11/15) of PAX5/ETV6 up-regulated genes, are known to be PAX5-repressed and only 27% (4/15) are activated also by wt PAX5. Comprehensively, PAX5/ETV6 modulated 68% of PAX5-target genes in an opposite direction to wt PAX5, whilst PAX5/ETV6 and PAX activity is concordant only on 32% of targets (Supplementary Figure S4).

A canonical *dominant negative* effect of several PAX5-fusions over wt PAX5 was previously proposed, through the mechanism of DNA-binding competition.(2) Differently, we herewith propose that PAX5/ETV6 has a more complex role on wt PAX5, defined here as an *opposite dominant* function (Supplementary Figure S5).

Furthermore, by setting a less stringent threshold (FC≤0.90), we recognized the presence of 52 additional PAX5-target probes (Supplementary Table S4), giving a total of 71 genes repressed by PAX5/ETV6, the majority of them being activated by wt PAX5 (n=46/71, 65%). Similarly, by setting a less stringent threshold (FC≥1.10) we recognized 46 additional PAX5-target genes (Supplementary Table S4), for a total of 61 physiological PAX5-target genes activated by PAX5/ETV6, being 34 among them (56%) known as to be PAX5-repressed.

Among genes repressed by wt PAX5 but up-regulated by PAX5/ETV6 is the *H19* gene, which is the immature precursor of mir-675-5p and mir-675-3p.(13) By RQ-PCR assay, we confirmed that H19 (Supplementary Table S3), as well as the two mature micro-RNAs, showed a significant up-regulation (mir-675-3p: FC=7.94, t test, p=0.0022; mir-675-5p: FC=5.72, t test, p=0.0117) in PAX5/ETV6 LY5.1FL pre-BI cells (Supplementary Figure S6A-B and Supplementary Figure S6C-D for results obtained in FLB6-67 and B6BAFL cells).

We additionally investigated the expression of PAX5/ETV6-activated but wt PAX5-repressed genes in PAX5$^{-/-}$ pro-B cells. Although in the absence of wt PAX5, PAX5-repressed genes are (obviously) already activated in these cells, PAX5/ETV6 induces a further increase of *RAMP1A*, *TSPAN*, *LCK* and *H19* levels. *H19* has FC=1.67 (p=0.0302 by t test), and accordingly mir-675-3p and -5p increased hundreds time (FC=373.44; p=0.0085 by t test, and FC=24.41; p=0.0001 by t test, respectively) (Supplementary Figure S6E-F).

On the other hands, only a few ETV6 transcriptional targets genes(14) were differentially expressed and validated (*Ldlr* and *Cirbp*, n=2/340, 0.6%), even considering the less stringent threshold (Supplementary Table S1,S2,S5). Therefore, it was not possible to assess whether the fusion protein played a role in the ETV6 pathway.

Certainly, GEP data suggested an opposite dominant interference of the PAX5/ETV6 protein on the wt PAX5 transcription pathway, which could be exerted through a direct competition between PAX5 and PAX5/ETV6, although we cannot exclude alternative mechanisms, which need to be experimentally explored. Indeed, *PAX5* repression was observed at mRNA level (Supplementary Figure S7A). Moreover, by western blot analysis using an anti-N-term PAX5, we observed a significant reduction of the endogenous PAX5 in PAX5/ETV6 cells, even though wt pre-BI cells have two wt *PAX5* alleles (Figure 1B). The PAX5 reduction was confirmed (Supplementary Figure S7B) by using a wt PAX5 specific anti C-terminus antibody, with a -47% mean decrease in LY5.1FL, -34% in FLB6-67 and -35% in B6BAFL (details in Supplementary Results). Although we cannot prove here whether PAX5/ETV6 directly regulates PAX5 protein, its down-regulation may be a potentially important mechanism in transformation: indeed, in human leukemia, one *PAX5* allele is involved in the translocation event, and the second wt PAX5 allele could be partially repressed plus partially hindered by the fusion protein in its function.

To establish the comprehensive functional significance of the PAX5/ETV6 transcription profile, even in the presence of relatively small changes in gene expression, IPA and DAVID pathway analysis software, applied on the DEGs list, consistently identified genes involved in B-cell receptor signaling and adhesion processes (Supplementary Bioinformatics file).

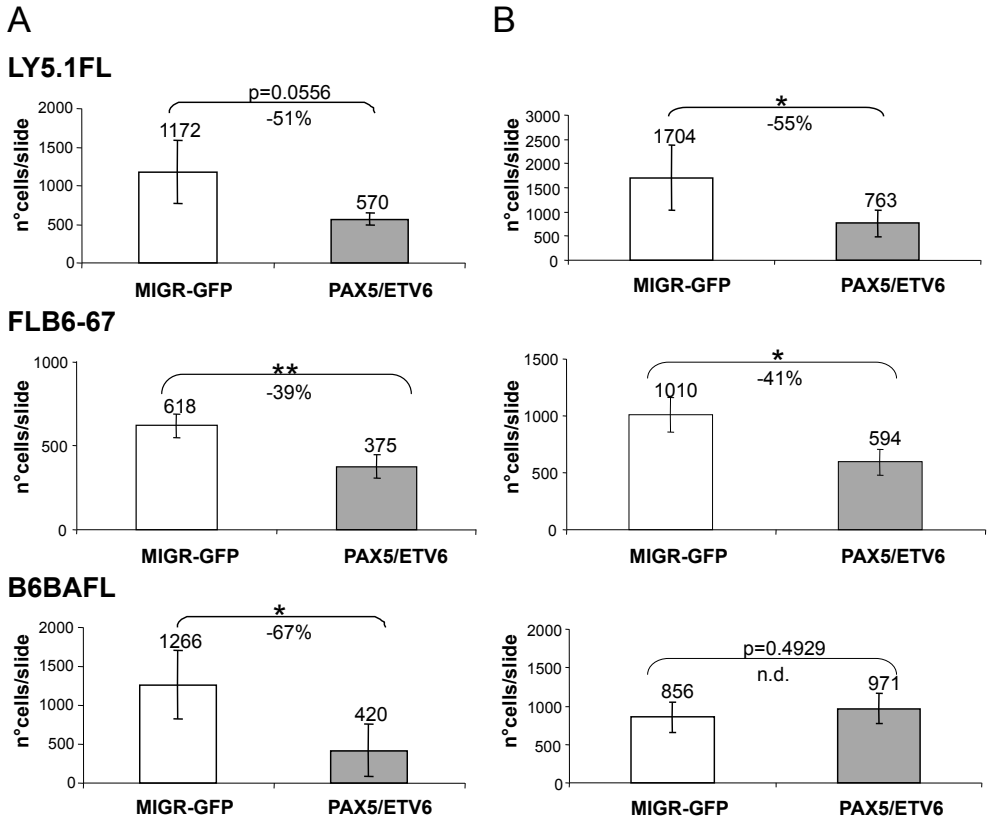


Figure 2. Cell Adhesion is modulated by PAX5/ETV6 protein. (A)*In vitro* assay of VCAM1 adhesion in the absence of the CXCL12 chemokine, in LY5.1FL (p=0.0556, N=2), FLB6-67 (p=0.0027, N=4) and in B6BAFL (p=0.0230, N=4). **(B)***In vitro* assay of VCAM1 adhesion in the presence of CXCL12 chemokine, in LY5.1FL (p=0.0418, N=4), FLB6-67 (p=0.0178, N=4) and in B6BAFL (p=0.4929, N=4). T test, *p<0.05. **p<0.01

The down-regulation was further confirmed for additional PAX5-activated genes known to be involved in pre-BCR signaling but not present in the microarray, as described in Supplementary Results (Supplementary Table S6). To functionally prove whether the assembly and signaling of pre-BCR was impaired, we stimulated IgM rearrangement by culturing LY5.1FL cells in the presence of sub-optimal levels of IL-7 (0.5%). By FACS, we analyzed μ chain expression on cell surface, and representative phenotypes are reported (Figure 1C). At day 0, we detected no protein expression on control or PAX5/ETV6 cells (1.22% and 0.28% μ-positive cells, respectively). At day +5, control

cells expressed remarkable levels of μ chain (40.84% μ-positive cells) as compared to PAX5/ETV6 cells (2.82%), which indeed exhibited a severe defect in pre-BCR assembly.
In parallel, we demonstrated a modulation of genes responsible for extra-cellular binding as well as involved in the intracellular signaling, supporting the role of PAX5/ETV6 in adhesion (Supplementary Results and Supplementary Table S7). In particular, we focused the attention on VCAM1, a VLA4-binding B-cell adhesion molecule in bone marrow,(12) which is composed by integrins CD29 and CD49d. Both were expressed at comparable levels on PAX5/ETV6 and control cells (Supplementary Figure S8). In an *in vitro* adhesion assay to VCAM1, PAX5/ETV6 cells turned out to have a significant reduction of adhesion rate in basal conditions (Figure 2A) (LY5.1FL: -51%, p=0.0556; FLB6-67 cells: -39%, p=0.0027; B6BAFL cells: -67%, p=0.0230). In presence of the CXCL12 chemokine (Figure 2B), which represents a positive signal to VCAM1-mediated adhesion and to migration, PAX5/ETV6 cells efficiently responded to the stimulus by improving their adhesion rate, but the overall response was lower than control cells in LY5.1FL (-55%, p=0.0418;) and in FLB6-67 (-41%, p=0.0178), while B6BAFL cells showed the same response rate. In addition to the previously reported increased migration rate (5), reduced adhesion could lead to tissues infiltration, although in vivo experiments are necessary to prove this.
In conclusion, the *opposite dominance* of PAX5/ETV6 over endogenous PAX5 and the specific transcription profile are responsible for aberrant mechanisms essential in the leukemic transformation (Supplementary Figure S5). First, PAX5/ETV6 determines a PAX5 haplo-insufficiency setting; second, the PAX5/ETV6 fusion protein could be actively responsible for the B-cell development block, mediated by the repression of physiological PAX5-activated genes. Moreover, the fusion protein may induce the acquisition of novel features, due to the aberrant up-regulation of genes which are normally repressed in precursor B-cells. It will be crucial to confirm these phenomena and understand the molecular mechanisms of the opposite dominant function in *in vivo* mouse models and in primary patients' cells carrying PAX5-translocations.

Supplementary information is available at Leukemia's website.

Acknowledgments

Grant support: Associazione Italiana Ricerca sul Cancro (AIRC) and Federazione Italiana Ricerca sul Cancro (FIRC), Fondazione Cariplo, MIUR, UNIMIB, Associazione Gianfranco Lupo.

We thank Dr. Cristina Bugarin (Centro Ricerca Tettamanti, Clinica Pediatrica Università di Milano-Bicocca, Ospedale San Gerardo, Monza, Italy) for the sorting procedures.

Conflict of interest disclosures

The authors declare no competing financial interests or conflict of interest.

Authorship contribution

Fazio Grazia, as post-doc researcher, developed the study, designed the experimental plan, carried out the assays and wrote the manuscript; moreover, she coordinated the work of Cazzaniga Valeria (PhD student), who contributed to the study, especially on bioinformatics analysis and functional assays. Palmi Chiara contributed to the adhesion assay setup and to scientific discussion. Galbiati Marta performed the image acquisition and analyses of adhesion assays. Giordan Marco and te Kronnie Geertruy performed the statistical analysis of microarray data. Prof. Rolink Antonius was involved in the *in vitro* model development and contributed to the scientific discussion of the project. Biondi Andrea and Cazzaniga Giovanni share the senior authorship of this paper, they coordinated the study, with scientific supervision and discussions, helped to interpret the results and contributed to the revision of the manuscript.

REFERENCES

1. Cobaleda C, Schebesta A, Delogu A, Busslinger M. Pax5: the guardian of B cell identity and function. *Nature Immunology* 2007; **8:** 463-470.
2. Fazio G, Biondi A, Cazzaniga G (eds). *The role of PAX5 in ALL. In "Novel aspects in Acute Lymphoblastic Leukemia"*. InTech, 2011.
3. Mullighan CG, Goorha S, Radtke I, Miller CB, Coustan-Smith E, Dalton JD*, et al.* Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* 2007; **446:** 758-764.
4. Nebral K, Denk D, Attarbaschi A, Konig M, Mann G, Haas OA*, et al.* Incidence and diversity of PAX5 fusion genes in childhood acute lymphoblastic leukemia. *Leukemia* 2009; **23:** 134-143.
5. Fazio G, Palmi C, Rolink A, Biondi A, Cazzaniga G. PAX5/TEL acts as a transcriptional repressor causing down-modulation of CD19, enhances migration to CXCL12, and confers survival advantage in pre-BI cells. *Cancer Res* 2008; **68:** 181-189.
6. Kawamata N, Ogawa S, Zimmermann M, Kato M, Sanada M, Hemminki K*, et al.* Molecular allelokaryotyping of pediatric acute lymphoblastic leukemias by high-resolution single nucleotide polymorphism oligonucleotide genomic microarray. *Blood* 2008; **111:** 776-784.
7. Kawamata N, Pennella MA, Woo JL, Berk AJ, Koeffler HP. Dominant-negative mechanism of leukemogenic PAX5 fusions. *Oncogene* 2011.
8. Rolink AG, Schaniel C, Melchers F. Stability and plasticity of wild-type and Pax5-deficient precursor B cells. *Immunol Rev* 2002; **187:** 87-95
9. Delogu A, Schebesta A, Sun Q, Aschenbrenner K, Perlot T, Busslinger M. Gene repression by Pax5 in B cells is essential for blood cell homeostasis and is reversed in plasma cells. *Immunity* 2006; **24:** 269-281.
10. McManus S, Ebert A, Salvagiotto G, Medvedovic J, Sun Q, Tamir I*, et al.* The transcription factor Pax5 regulates its target genes by recruiting chromatin-modifying proteins in committed B cells. *EMBO J* 2011; **30:** 2388-2404.
11. Pridans C, Holmes ML, Polli M, Wettenhall JM, Dakic A, Corcoran LM*, et al.* Identification of Pax5 target genes in early B cell differentiation. *J Immunol* 2008; **180:** 1719-1728.
12. Schebesta A, McManus S, Salvagiotto G, Delogu A, Busslinger GA, M. B. Transcription Factor Pax5 Activates the Chromatin of Key Genes Involved in B Cell Signaling, Adhesion, Migration, and Immune Function. *Immunity* 2007; **27:** 49-63.
13. Cai X, Cullen BR. The imprinted H19 noncoding RNA is a primary microRNA precursor. *Rna* 2007; **13:** 313-316.
14. Boily G, Larose J, Langlois S, Sinnett D. Identification of transcripts modulated by ETV6 expression. *Br J Haematol* 2007; **136:** 48-62.

SUPPLEMENTARY MATERIALS AND METHODS

Cell cultures and retroviral transduction. Three different murine wt pre-BI cell populations were isolated from fetal liver (called LY5.1FL, FLB6-67 and B6BAFL cells) and PAX5-/- pro-B cells were purified from bone marrow of a PAX5-deficient mouse. Both cells were cultured on OP9 bone marrow stroma cells in Iscove's modified Dulbecco's medium supplemented with 2% FBS, 0.03% w/v primatone and 5% IL-7.(1-2) The bicistronic vector pMSCV-IRES- GFP (MIGR) was used for the retroviral transduction which allows the expression of PAX5/ETV6 and GFP.(1) On day +3, sorting of GFP positive cells was performed (FACSAria, BD Biosciences).

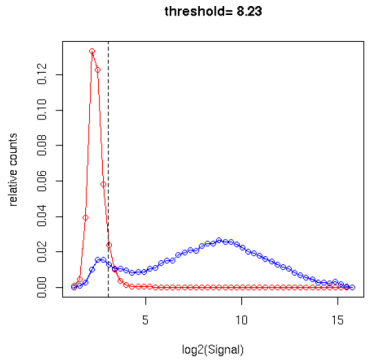
Microarray, statistics and bioinformatics analyses. B6BAFL cells were suspended in Trizol reagent (Invitrogen, Life Technologies) immediately after sorting for GFP. RNA extraction was performed following the manufacturer's protocol and processed by Affymetrix protocols (Genopolis Consortium, University of Milano-Bicocca, Milano, IT). Microarray analysis was carried out on the Affymetrix 430A 2.0 mouse gene chip, and data (as a.CEL file) have been deposited with Array Express repository (accession E-MEXP-3095). The Gene Expression Profile (GEP) by microarray analysis has been executed in two independent biological transduction experiments, in B6BAFL pre-BI cells.

052208_m_A_4_B6_BAFL_PT.CEL
052208_m_B_3_B6_BAFL_PT.CEL
052208_m_C_2_B6_BAFL_V.CEL
052208_m_D_1_B6_BAFL_V.CEL

Expression measures for each probe set were obtained using GCRMA, a multi-array analysis method that estimates probe set signals by considering the physical affinities between probes and targets. Normalization was performed by a quantile method. Two successive steps of data filtering were applied. At first from the dataset were filtered out the probe sets with Absent call in all the samples. Then an empirical threshold was determined on the intensity signal of the 95° percentile of the distribution of the absent call signals. The algorithm selected the probe sets with an intensity signal higher than

the threshold. 12,761 probe sets passed the filtering step, and subsequently probes whose signal was below 40 in all 4 gene chips were further discarded as non efficient. The filtered data were annotated using Bioconductor.

The algorithm selected the probe sets with a signal higher than the threshold (black dotted line). Absent calls: red line, Present calls: blue line.



The log2 (signal) has been calculated by the ratio of the PAX5/ETV6 mean signal over the MIGR-GFP mean signal. Specifically, we considered a differentially expressed gene if log2 (signal) was greater than +0.59 (corresponding to a fold change greater than 1.5) or lesser than -0.50 (corresponding to a fold change lesser than 0.7). For pathway analysis, GEP data were analyzed by on-line software: the Database for Annotation, Visualization and Integrated Discovery (DAVID, http://david.abcc.ncifcrf.gov/home.jsp) and Ingenuity Pathway Analysis (IPA, www.ingenuity.com).

Fisher's exact test. In order to verify whether the enrichment of PAX5-targets was statistically significant, the Fisher's exact test has been performed. In details, we proceeded as follows: i) we identified 581 probe

sets of known PAX5-target genes in any cell environment, derived from the literature (3-6); ii) 502 over 581 probes have been identified on the Affymetrix 430A 2.0 mouse gene chip; iii) we tested our dataset of 12,761 probes and only 389 PAX5-target probes were present in the filtered and normalized dataset, therefore we considered this probe set for Fisher's exact test analysis; iv) 34/389 PAX5-target genes were present among the 340 DEGs, while 355/389 belonged to the group of 12,421 non-DEGs.

Reverse transcription-PCR and real-time quantitative-PCR assays. RNA extraction was performedusing Trizol (Life Technologies). Superscript II enzyme (Life Technologies) was used for cDNA synthesis.Real-time analysis was performed on Light

Cycler 480II with Universal Probe Master system (RocheDiagnostics; F. Hoffmann-La Roche Ltd.). Primers and probes were selected according to the Software Probe Finder (Roche Diagnostics) and are reported in Supplementary Table S8. Data were expressed using the comparative ΔΔCt method, using *Hprt* gene as reference and MIGR-GFP cells as standardization control;(7-8) both t test and SD values refer to triplicates of a single experiment and N=3 biological independent experiments were performed for each gene.

MiRNA analysis. RNA was quantified by Nanodrop2000c (Thermo Fisher Scientific Inc). Reverse transcription with a miRNA-specific primer has been done using the TaqMan® MicroRNA Reverse Transcription Kit (Life Technologies) and followed by real-time PCR with TaqMan® MicroRNA Assays for each single miRNA (mmu-miR-675-5p assay ID n°1940, -3p assay ID n°1941 and assay ID n°001093 for RNU6B snRNA, which has been used as reference miRNA). The comparative ΔΔCt method has been applied.(8)

Antibodies and flow cytometry. Phycoerythrin-conjugated antibodies anti-CD19 (MB19-1), anti-CD22 (2D6), anti-SLAMF6/LY108 (13G3-19D) and allophycocyanin-conjugated antibodies against CD44 (IM7) were used (e-Bioscience Inc). Data were analyzed using CellQuest Software (BD Biosciences).

Protein array and western blot analysis. Protein array was performed on cell supernatant, obtained from B6BAFL cell culture, using RayBio® Cytokine Antibody Arrays - Mouse Array III-IV (Raybiotech Inc), following the manufacturer's protocol. Western blot analysis was performed following either Andrews's extraction method (9-10) or standard RIPA buffer. Anti-beta-actin antibody was used at working dilution 1:1000 (AC-15, Sigma-Aldrich), anti-N-term PAX5 was used at 1:545 (AB4227, Chemicon) and anti-C-term PAX5 at 1:200 (AB15164, Abcam). A StripAblot Stripping Buffer (Euroclone S.p.A.) was used to recover membranes. Densitometry analyses were performed using Alliance instrument and Uviband software (Uvitec Cambridge, UK).

IL-7 starvation assay and IgM/μ heavy chain analysis. Cells were washed out by IL-7 and plated in IMDM+2% FBS+0.5% IL-7 on OP9 stroma layer. At day 0 and day +5, staining for μ heavy chain of IgM was performed using the allophycocyanin-conjugated antibodies against IgM/μ heavy chain (II/41) (e- Bioscience Inc).

VCAM1 adhesion assay. The adhesion of pre-BI cells to VCAM1-coated coverslips was performed as described elsewhere.(6) Briefly, 15mm Ø round coverslips were coated with recombinant mouse VCAM1-Fcprotein (25 µg/ml; R&D Systems) and placed in 12-well dishes, containing 0.5 ml IMDM+2%FBS medium,with $5x10^5$ pre-BI cells, in the presence or lack of 100 ng/ml CXCL12 (Peprotech Inc). The coverslips weremounted on glasses in the presence of DAPI (8µl). Each coverslip was analyzed by a fluorescence microscope, acquiring 30 representative fields, and the single cells were counted.

SUPPLEMENTARY RESULTS

PAX5/ETV6 repressed physiologically activated PAX5-target genes.

In GEP analysis, the value of log2(signal) of CD22 was equal to -1.39 (fold change FC=0.38), and, from RQ- PCR analysis, we obtained fold changes of 0.07 (t test, p=0.000034) in LY5.1FL cells, of 0.29 (t test, p=0.006058) and of 0.42 (t test, p= 0.000034) and in FLB6-67 and B6BAFL pre-BI cells respectively (Supplementary Table 2). More important, phenotypic analysis confirmed previous data, as shown in Supplementary Figures S1B, S2A and S3A, where CD19 overlay is shown as internal control of PAX5/ETV6 activity. In details, considering the CD22 ratio between the Mean Fluorescent Intensity (MFI) of PAX5/ETV6 cells and the MFI of control cells, we obtained a correspondent mean protein repression of - 42% in LY5.1FL (range -19% to -77%, in 9 independent experiments), -44% MFI ratio in FLB6-67 (range - 42% to -45%, in 2 experiments) and of -42% MFI ratio (range -20% to -74%, in 6 experiments) in B6BAFL (Supplementary Figures S1C, S2B and S3B). SLAMF6/LY108 gene was repressed in GEP analysis with a log2(signal)=-0.55 (FC=0.68) confirmed by RQ-PCR analyses (Supplementary Table 2), in LY5.1FL (FC=0.23, t test, p=0.00081), in FLB6-67 (FC=0.59, t test p=0.002125) and in B6BAFL, with a FC=0.57 (p=0.014589). In addition, phenotypic analysis was in agreement with these data, indeed, in LY5.1FL cells (Supplementary Figure S1D), SLAMF6/LY108 was down-regulated of -44% MFI ratio and of -33% MFI ratio in FLB6-67 cells (range -20% to -46%, in 8 experiments, Supplementary Figure S2C) and this result was consistent in B6BAFL, with -54% MFI ratio (Supplementary Figure S3C), range -38% to -71% in 3 experiments. Similarly, in LY5.1FL we observed a 41% decrease in CD44 protein expression on PAX5/ETV6 cells

(MFI ratio range from -19% to -64% in 26 experiments, Supplementary Figure S2E), while the FC was 0.72 by GEP and 0.60 by RQ-PCR (Supplementary Table S2); correspondingly, as shown in Supplementary Figure S2D by FACS analysis, in FLB6-67 cells CD44 was significantly repressed of 28%(in 26 experiments, MFI ratio range -19% to -46%), as well as in B6BAFL (Supplementary Figure S3D), where CD44 was repressed of 43% (in 16 experiments, MFI ratio range -28% to -78%).

PAX5/ETV6 activated physiologically repressed PAX5-target genes.

By RQ-PCR assay, we demonstrated that *H19* and the two micro-RNAs were up-regulated in PAX5/ETV6 LY5.1FL pre-BI cells, which expressed both mir-675-3p and -5p at appreciable levels (Supplementary Figure S6A-B). On the other end, in the FLB6-67, nor mir-675-3p or -5p were amplified with good quality curves, thus we proceeded to a pre-amplification cycle followed by standard RQ-PCR; with this strategy, we technically rescued only mir-675-3p, whereas -5p experiment presented again unacceptable curves of amplification (data not shown). As shown in Supplementary Figure S6C, in PAX5/ETV6 cells, mir-675-3p was significantly up-regulated with a fold increase of 8.75 (t test p=0.014). In addition B6BAFL pre-BI cells, mir-675-3p was present at detectable levels especially in PAX5/ETV6 cells compared to control, leading to a tremendous fold increase equal to 128.27 (t test p=0.005, Supplementary Figure S6D); whereas, mir-675-5p was very low expressed and probably, because of this, we didn't find any statistical difference (data not shown).

PAX5/ETV6 blocked pre-BCR expression and repressed molecules associated to its internal signaling.

We further extended the analysis of genes involved in pre-BCR signaling, but not represented in the microarray. In fact, we showed that PAX5/ETV6 cells repressed PAX5-regulated components of BCR at the surface level, such as CD19 and CD22 (Supplementary Figures S1-S2-S3), CD79a/Igα, CD37, as well as molecules associated to its intracellular signaling, such as *IRF8*, *IRF4*, *IKZF2*/*HELIOS* and *IKZF3*/*AIOLOS* (Supplementary Tables S2-S6) as well as *CD79α* and *BLNK*, as demonstrated in our previous study.(1) Furthermore, many molecules involved in lymphocyte signaling and

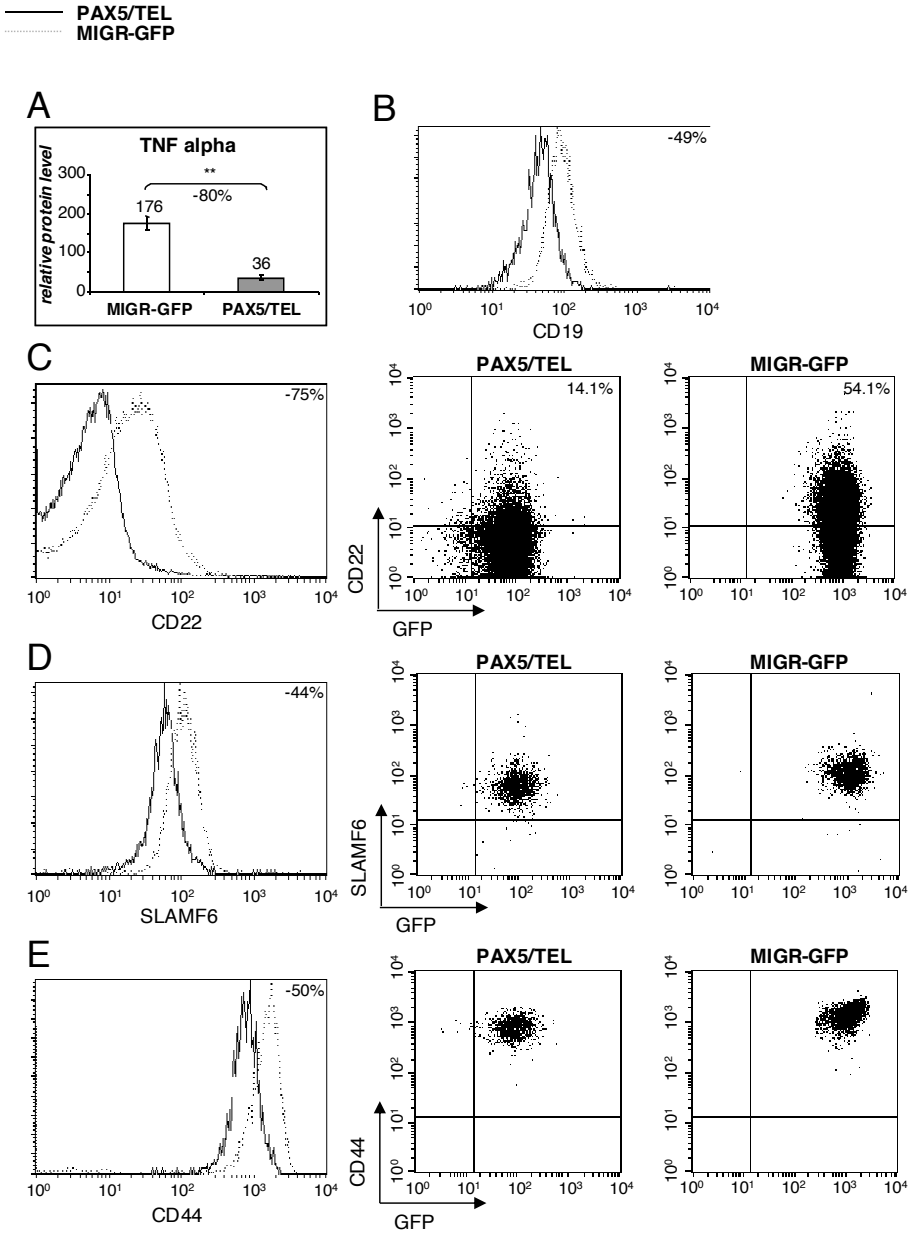
activation, but not regulated by PAX5, are modulated by PAX5/ETV6 (Supplementary Figures S12-S13).

PAX5/ETV6 represses molecules important for B-cell adhesion.

Modulation of genes responsible for extra-cellular binding as well as the molecules involved in the intracellular adhesion signaling supported the role of PAX5/ETV6 in adhesion and in particular respect to VCAM1, a physiological B-cell adhesion molecule expressed in bone marrow niche.(6,11-12) We previously showed the down-regulation of several adhesion molecules, known as PAX5-targets, such as SLAMF6/LY108, CD44, *BCAR3*, *SDC4*, *NEDD9*, *CAPN2*, and *PLEKHA2* genes (Supplementary Table S2;Supplementary Figure S1-S2-S3). Moreover, we demonstrated by RQ-PCR that PAX5/ETV6 caused the down-regulation of *EDG1* and *EPS8*, two PAX5-activated genes having a role in cell adhesion, whose probes were not present in the microarray platform (Supplementary Table S7). Finally, several genes were identified as being involved in cell adhesion, but not regulated by PAX5; examples are: *RHOB* (Supplementary Table S2), TNFα (Supplementary Figure S1A), *THY1*, *CD47*, *CORO1A*, *ITGB1*, *LYN*, *PAFAH1B1*, *THBD*, and *SLAMF1* (Supplementary Figure S14).Altogether, this evidence prompted us to investigate the role of PAX5/ETV6 in B-cell adhesion to VCAM1.

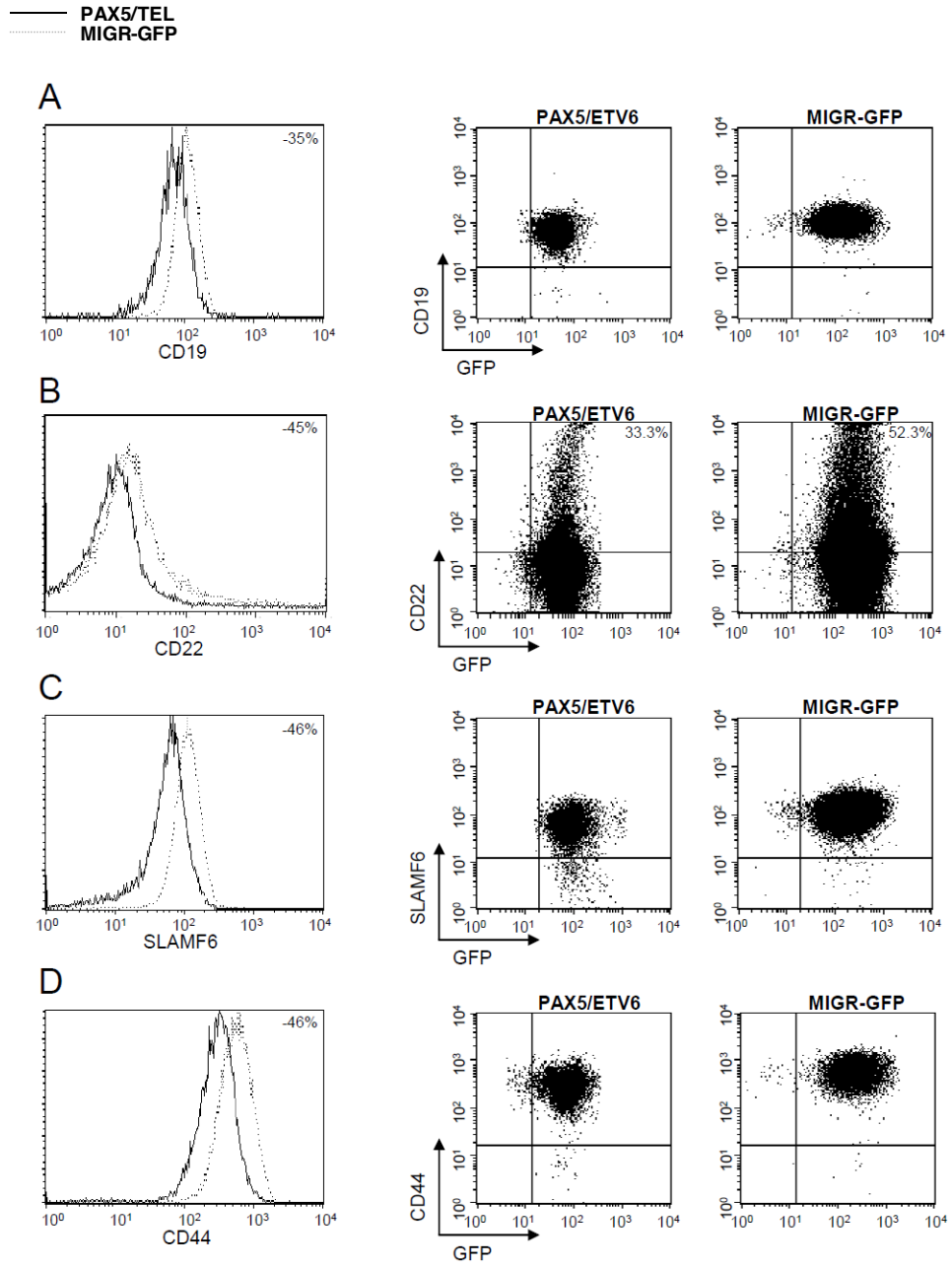
SUPPLEMENTARY FILES

Figure S1



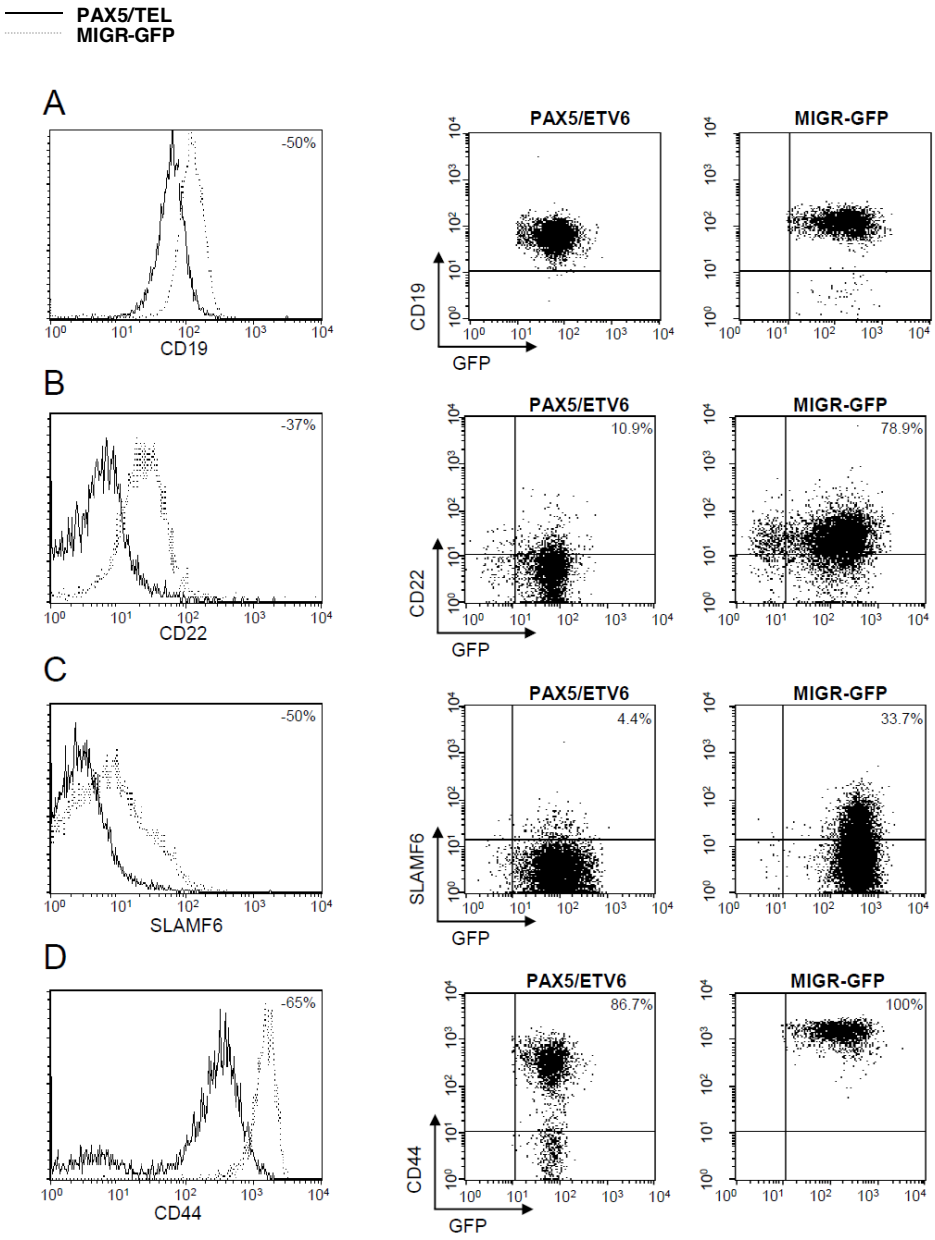
Supplementary Figure S1. Validation experiments: protein analysis of down-regulated molecules inLY5.1FL pre-BI cells. (A) TNF alpha protein levels by protein array analysis of B6BAFL cell supernatant,transduced either by PAX5/ETV6 or by MIGR-GFP vector. From (B) to (E), a representative phenotypicanalysis in LY5.1FL pre-BI cells is shown. The down-regulation of -49% of CD19 in PAX5/ETV6 cellscompared to control MIGR-GFP cells is shown in the histogram overlay analysis in panel (B), as a positivecontrol for fusion protein activity. CD22 expression is -75% (C), SLAMF6 is -44% (D) and CD44 is -50%(E) in PAX5/ETV6 cells compared to control MIGR-GFP cells.

Figure S2



Supplementary Figure S2. Representative phenotype analysis in FLB6-67 cells. (A) CD19 expression is down-regulated of -35%, (B) CD22 expression is -45%, (C) SLAMF6 is -46% and (D) CD44 is -46% in PAX5/ETV6 cells compared to control MIGR-GFP cells.

Figure S3



Supplementary Figure S3. Representative phenotype analysis in B6BAFL cells. (A) CD19 expression is down-regulated of -50%, (B) CD22 expression is -37%, (C) SLAMF6 is -50% and (D) CD44 is -65% in PAX5/ETV6 cells compared to control MIGR-GFP cells.

Figure S4

Affymetrix probe ID	Gene name			
1415936_at	Bcar3	PAX5-activated gene	PAX5/ETV6-repressed gene	
1422476_at	Ifi30	PAX5-activated gene	PAX5/ETV6-repressed gene	
1419768_at	Cd22	PAX5-activated gene	PAX5/ETV6-repressed gene	
1448529_at	Thbd	PAX5-activated gene	PAX5/ETV6-repressed gene	
1415897_a_at	Mgst1	PAX5-activated gene	PAX5/ETV6-repressed gene	
1448890_at	Klf2	PAX5-activated gene	PAX5/ETV6-repressed gene	
1420659_at	Slamf6	PAX5-activated gene	PAX5/ETV6-repressed gene	
1429144_at	Gpcpd1 (Prei4)	PAX5-activated gene	PAX5/ETV6-repressed gene	
1418826_at	Ms4a6b	PAX5-activated gene	PAX5/ETV6-repressed gene	
1436589_x_at	Prkd2	PAX5-activated gene	PAX5/ETV6-repressed gene	
1421998_at	Tor3a	PAX5-activated gene	PAX5/ETV6-repressed gene	
1424375_s_at	Gimap4	PAX5-activated gene	PAX5/ETV6-repressed gene	= 12 opposite
1415812_at	Gsn	PAX5-activated gene	PAX5/ETV6-activated gene	
1436991_x_at	Gsn	PAX5-activated gene	PAX5/ETV6-activated gene	
1417444_at	E2f5	PAX5-activated gene	PAX5/ETV6-activated gene	
1451260_at	Aldh1b1	PAX5-activated gene	PAX5/ETV6-activated gene	= 4 concordant
=16 probes	**=15 genes**			
1460378_a_at	Tes	PAX5-repressed gene	PAX5/ETV6-repressed gene	
1444028_s_at	Dock9 (d14wsu89e)	PAX5-repressed gene	PAX5/ETV6-repressed gene	
1450932_s_at	Dock9 (d14wsu89e)	PAX5-repressed gene	PAX5/ETV6-repressed gene	
1419255_at	Spnb2	PAX5-repressed gene	PAX5/ETV6-repressed gene	
1424246_a_at	Tes	PAX5-repressed gene	PAX5/ETV6-repressed gene	
1435945_a_at	Kcnn4	PAX5-repressed gene	PAX5/ETV6-repressed gene	
1422003_at	Blr1	PAX5-repressed gene	PAX5/ETV6-repressed gene	= 7 concordant
1427139_at	Adamts10	PAX5-repressed gene	PAX5/ETV6-activated gene	
1424039_at	Tmem66	PAX5-repressed gene	PAX5/ETV6-activated gene	
1425396_a_at	Lck	PAX5-repressed gene	PAX5/ETV6-activated gene	
1424567_at	Tspan2	PAX5-repressed gene	PAX5/ETV6-activated gene	
1432417_a_at	Tspan2	PAX5-repressed gene	PAX5/ETV6-activated gene	
1422804_at	Serpinb6b	PAX5-repressed gene	PAX5/ETV6-activated gene	
1448390_a_at	Dhrs3	PAX5-repressed gene	PAX5/ETV6-activated gene	
1448194_a_at	H19	PAX5-repressed gene	PAX5/ETV6-activated gene	
1438511_a_at	1190002H23Rik	PAX5-repressed gene	PAX5/ETV6-activated gene	
1417481_at	Ramp1	PAX5-repressed gene	PAX5/ETV6-activated gene	
1452136_at	Slc5a9	PAX5-repressed gene	PAX5/ETV6-activated gene	= 11 opposite
=18 probes	**=16 genes**			

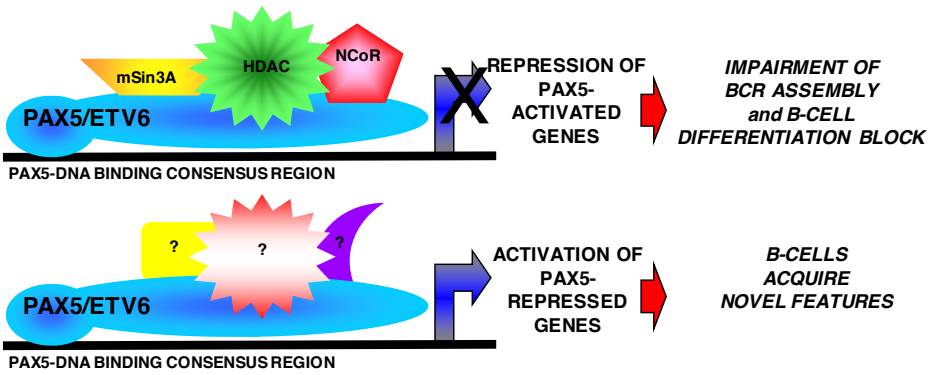
19/34	**56 % PAX5-target REPRESSED by PAX5/ETV6**
15/34	**44 % PAX5-target ACTIVATED by PAX5/ETV6**

23/34	**68 % opposite dominance function**
11/34	**32 % concordant with PAX5 function**

Fisher's exact test $1.029\ e^{-09}$

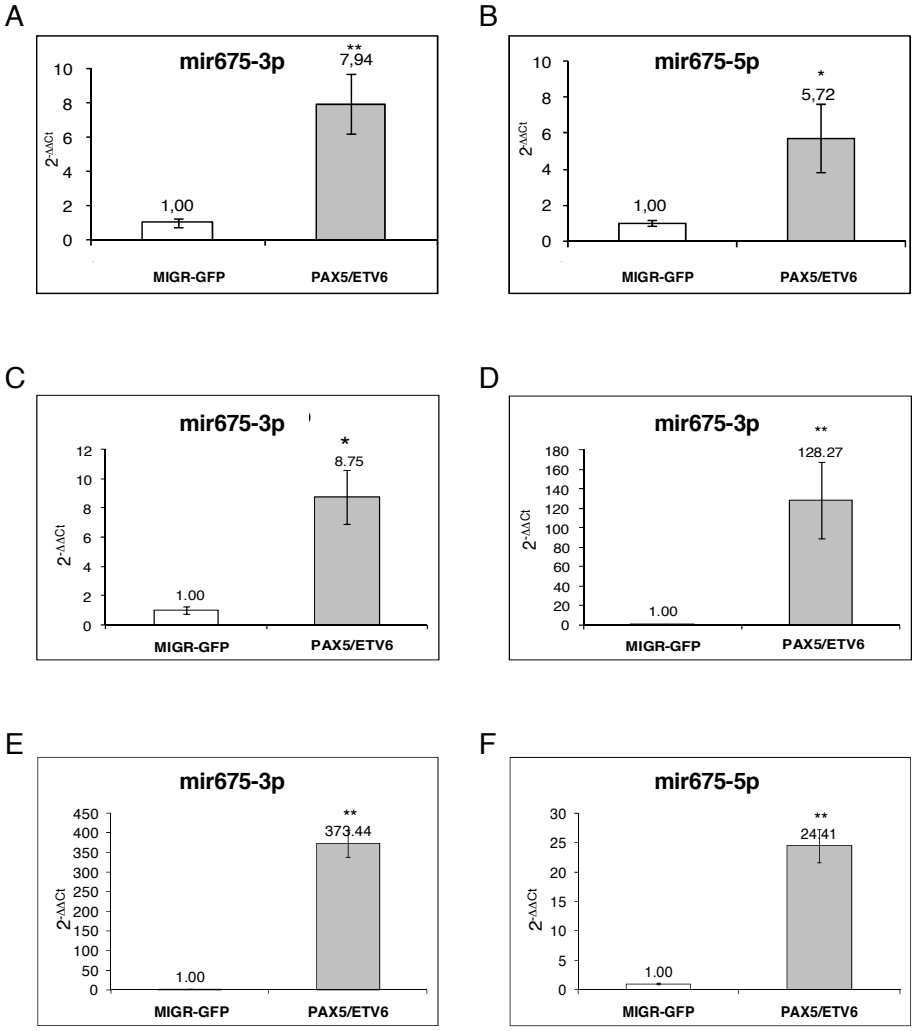
Supplementary Figure S4. Schematic representation of PAX5-target genes among DEGs.

Figure S5



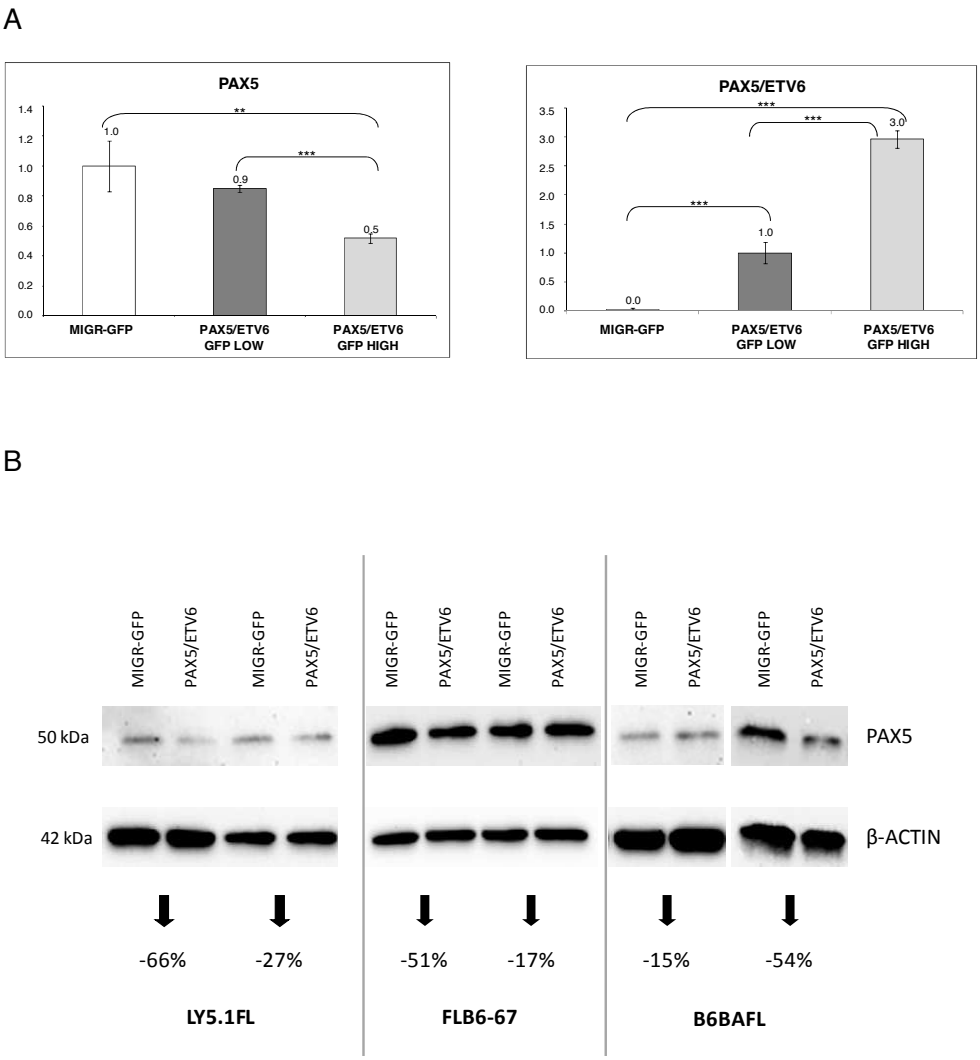
Supplementary Figure S5. Molecular and functional model of PAX5/ETV6 action.

Figure S6



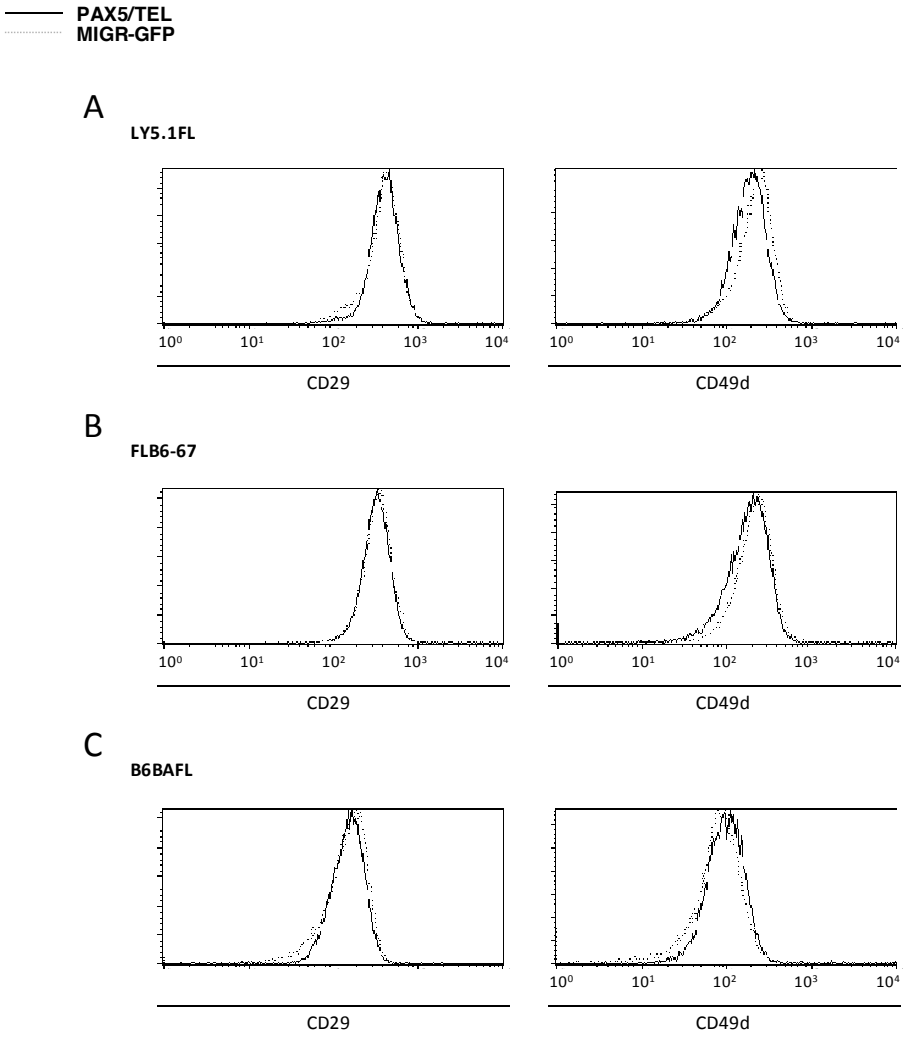
Supplementary Figure S6. (A-B) Mir675-3p and mir675-5p expression in LY5.1FL pre-BI cells. (C) Mir675-3p expression in FLB6-67 and (D) in B6BAFL cells. (E-F) Mir675-3p and mir675-5p expression in PAX5-/- pro-B cells. T test, *p<0.05. **p<0.01

Figure S7



Supplementary Figure S7. Down-regulation of endogenous PAX5 expression. (A) *PAX5/ETV6* and endogenous *PAX5* mRNA levels, in two sub-populations of PAX5/ETV6 B6BAFL cells, sorted for low and high levels of GFP-PAX5/ETV6. Probes and primers at 3'-term of wt *PAX5* and at fusion mRNA breakpoint were designed, demonstrating an inverse correlation between *PAX5/ETV6* and *PAX5* mRNA expression, confirming the down regulation of wt *PAX5* in presence of increasing levels of the fusion gene (RQ-PCR analysis). (B) Protein analysis by western blot using anti-C-term PAX5 antibody and beta-actin (as internal control) were performed on the same membrane, for each cell population. T test, *p<0.05. **p<0.01 ***p<0.001.

Figure S8



Supplementary Figure S8. A representative phenotypic analysis of CD29 and CD49d expression in (A) LY5.1FL, (B) FLB6-67 and (C) B6BAFL pre-BI cells. The histogram overlay analysis shows PAX5/ETV6 cells and control MIGR-GFP cells.

Table S1

PAX5/ETV6-REPRESSED GENES [log2(signal)<-0.50 = Fold change <0.70]

	Affymetrix probe ID	gene name	log2(signal)	FC	gene description
1	**1449394_at**	**Slco1b2**	-4.08	0.06	**solute carrier organic anion transporter family, member 1b2**
2	**1452494_s_at**	**Slco1b2**	-3.73	0.08	**solute carrier organic anion transporter family, member 1b2**
3	**1420353_at**	**Lta**	-3.48	0.09	**lymphotoxin A**
4	1419703_at	Col5a3	-3.29	0.10	collagen, type V, alpha 3
5	1425917_at	H28	-2.88	0.14	histocompatibility 28
6	1421836_at	Mtap7	-1.91	0.27	microtubule-associated protein 7
7	1423319_at	Hhex	-1.91	0.27	hematopoietically expressed homeobox
8	1449991_at	Cd244	-1.83	0.28	CD244 natural killer cell receptor 2B4
9	**1422476_at**	**Ifi30**	-1.68	0.31	**interferon gamma inducible protein 30**
10	1452614_at	Gm566	-1.63	0.32	BCLI2-like 15
11	AFFX-18SRNAMur/ X00686_3_at	NA	-1.58	0.33	NA
12	**1415936_at**	**Bcar3**	-1.55	0.34	**breast cancer anti-estrogen resistance 3**
13	1448078_at	C76533	-1.52	0.35	expressed sequence C76533
14	1450917_at	Myom2	-1.49	0.36	myomesin 2
15	1436798_at	Rpl9	-1.49	0.36	ribosomal protein L9
16	1452352_at	Ctla2b	-1.48	0.36	cytotoxic T lymphocyte-associated protein 2 beta
17	1425923_at	Mycn	-1.44	0.37	v-myc myelocytomatosis viral related oncogene, neuroblastoma derived (avian)
18	1418386_at	2510005D08Rik	-1.44	0.37	N-6 adenine-specific DNA methyltransferase 2 (putative)
19	**1419768_at**	**Cd22**	-1.39	0.38	**CD22 antigen**
20	1449105_at	Sh2d2a	-1.35	0.39	SH2 domain protein 2A
21	1425570_at	Slamf1	-1.34	0.39	signaling lymphocytic activation molecule family member 1
22	1450932_s_at	Dock9	-1.31	0.40	dedicator of cytokinesis 9
23	1444028_s_at	Dock9	-1.29	0.41	dedicator of cytokinesis 9
24	1449072_a_at	Hbb-b2	-1.29	0.41	N-6 adenine-specific DNA methyltransferase 2 (putative)
25	1449937_at	Pp11r	-1.28	0.41	placental protein 11 related
26	1434484_at	1100001G20Rik	-1.28	0.41	RIKEN cDNA 1100001G20 gene
27	1460378_a_at	Tes	-1.27	0.42	testis derived transcript
28	1424447_at	1700030K09Rik	-1.27	0.42	RIKEN cDNA 1700030K09 gene
29	1421598_at	NA	-1.25	0.42	NA
30	AFFX-18SRNAMur/ X00686_5_at	NA	-1.24	0.42	NA
31	1422003_at	Blr1	-1.22	0.43	chemochine (C-X-C motif) receptor 5
32	1448529_at	Thbd	-1.21	0.43	thrombomodulin
33	1454231_a_at	Rpgrip1	-1.18	0.44	retinitis pigmentosa GTPase regulator interacting protein 1
34	1419255_at	Spnb2	-1.17	0.44	spectrin beta 2
35	1421504_at	Sp4	-1.12	0.46	trans-acting transcription factor 4
36	1419607_at	Tnf	-1.12	0.46	tumor necrosis factor
37	1448501_at	Tspan6	-1.10	0.47	tetraspanin 6
38	**1449110_at**	**Rhob**	-1.10	0.47	**ras homolog gene family, member B**
39	1426568_at	Slc2a9	-1.10	0.47	solute carrier family 2 (facilitated glucose transporter), member 9
40	1450988_at	Lgr5	-1.10	0.47	leucine rich repeat containing G protein coupled receptor 5
41	1416154_at	Srp54	-1.07	0.48	signal recognition particle 54a
42	1436677_at	1810032O08Rik	-1.05	0.48	RIKEN cDNA 1810032O08 gene
43	1452353_at	2510005D08Rik	-1.04	0.49	G protein-coupled receptor 155
44	1423465_at	Frrs1	-1.03	0.49	ferric-chelate reductase 1
45	1425519_a_at	Cd74	-1.03	0.49	CD74 antigen (invariant polypeptide of major histocompatibility complex, class II antigen-associated)
46	**1415897_a_at**	**Mgst1**	-1.01	0.50	**microsomal glutathione S-transferase 1**
47	1419521_at	Zfp94	-1.00	0.50	zinc finger protein 94
48	**1423868_at**	**Txnrd3**	-0.99	0.50	**thioredoxin reductase 3**
49	1448325_at	Myd116	-0.99	0.50	myeloid differentiation primary response gene 116
50	**1419135_at**	**Ltb**	-0.99	0.50	**lymphotoxin B**
51	1459908_at	NA	-0.98	0.51	NA
52	1416190_a_at	Sec61a1	-0.98	0.51	Sec61 alpha 1 subunit (S. cerevisiae)
53	1449818_at	Abcb4	-0.98	0.51	ATP-binding cassette, sub-family B (MDR/TAP), member 4
54	1447631_at	Myst2	-0.97	0.51	MYST histone acetyltransferase 2
55	1454970_at	Ptcd1	-0.96	0.51	BUD31 homolog (yeast)
56	1433863_at	Btf3	-0.95	0.52	basic transcription factor 3
57	1438527_at	Rpl3	-0.95	0.52	ribosomal protein L3
58	1441023_at	Eif2s2	-0.93	0.52	eukaryotic translation initiation factor 2, subunit 2 (beta)
59	1424142_at	Ikbkap	-0.93	0.52	inhibitor of kappa light polypeptide enhancer in B-cells, kinase complex-associated protein
60	1452008_at	9130422G05Rik	-0.93	0.53	tetratricopeptide repeat domain 39B
61	1455678_at	Sema4b	-0.92	0.53	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4B
62	1432151_at	2810409H07Rik	-0.91	0.53	Obg-like ATPase 1
63	1452148_at	Gpr155	-0.91	0.53	low density lipoprotein receptor-related protein associated protein 1
64	1423162_s_at	Spred1	-0.91	0.53	sprouty protein with EVH-1 domain 1, related sequence
65	**1423869_s_at**	**Txnrd3**	-0.91	0.53	**thioredoxin reductase 3**
66	1425556_at	Crkrs	-0.91	0.53	CDC2-related kinase, arginine/serine-rich
67	1422134_at	Fosb	-0.90	0.53	FBJ osteosarcoma oncogene B
68	1426526_s_at	Ovgp1	-0.89	0.54	oviductal glycoprotein 1
69	1420994_at	B3gnt5	-0.89	0.54	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 5

70	1451893_s_at	Magi1	-0.88	0.54	membrane associated guanylate kinase, WW and PDZ domain containing 1
71	**1425736_at**	**Cd37**	-0.88	0.54	**CD37 antigen**
72	1449439_at	Klf7	-0.88	0.54	Kruppel-like factor 7
73	1455930_at	LOC673503	-0.87	0.55	NA
74	1460352_s_at	Pik3r4	-0.87	0.55	phosphatidylinositol 3 kinase, regulatory subunit, polypeptide 4, p150
75	1416795_at	Cryl1	-0.87	0.55	crystallin, lambda 1
76	1451505_at	Chchd5	-0.86	0.55	coiled-coil-helix-coiled-coil-helix domain containing 5
77	1448748_at	Plek	-0.86	0.55	pleckstrin
78	1417755_at	Topors	-0.85	0.56	topoisomerase I binding, arginine/serine-rich
79	1424246_a_at	Tes	-0.85	0.56	testis derived transcript
80	1438501_at	Rps17	-0.84	0.56	ribosomal protein S17
81	1419356_at	Klf7	-0.84	0.56	Kruppel-like factor 7
82	**1431947_at**	**Ldlr**	-0.84	0.56	**low density lipoprotein receptor**
83	1453753_at	Dtl	-0.83	0.56	denticleless homolog (Drosophila)
84	1448773_at	Rpl11	-0.83	0.56	ribosomal protein L11
85	1422901_at	Mgea5	-0.83	0.56	meningioma expressed antigen 5 (hyaluronidase)
86	1425470_at	NA	-0.82	0.56	NA
87	1436781_at	Man2b1	-0.82	0.57	mannosidase 2, alpha B1
88	1433678_at	Pld4	-0.82	0.57	phospholipase D family, member 4
89	1424683_at	1810015C04Rik	-0.82	0.57	family with sequence similarity 134, member B
90	1436384_at	Rps10	-0.81	0.57	ribosomal protein S10
91	1452361_at	Rnf20	-0.80	0.57	ring finger protein 20
92	1433598_at	9430010O03Rik	-0.80	0.57	arginine and glutamate rich 1
93	1434403_at	Spred2	-0.80	0.58	sprouty-related, EVH1 domain containing 2
94	**1438397_a_at**	**Rbm39**	-0.79	0.58	**RNA binding motif protein 39**
95	1452987_at	Josd3	-0.77	0.59	TATA box binding protein (Tbp)-associated factor, RNA polymerase I, D
96	1438741_at	Rbm13	-0.77	0.59	cDNA sequence BC019943
97	1423596_at	Nek6	-0.77	0.59	NIMA (never in mitosis gene a)-related expressed kinase 6
98	**1449623_at**	**Txnrd3**	-0.77	0.59	**thioredoxin reductase 3**
99	1425673_at	Lpp	-0.76	0.59	LIM domain containing preferred translocation partner in lipoma
100	1434469_at	Otud4	-0.75	0.59	OTU domain containing 4
101	1425855_a_at	Crk	-0.75	0.60	v-crk sarcoma virus CT10 oncogene homolog (avian)
102	1450391_a_at	Mgll	-0.74	0.60	monoglyceride lipase
103	1426697_a_at	Lrpap1	-0.73	0.60	low density lipoprotein receptor-related protein associated protein 1
104	1416832_at	Slc39a8	-0.73	0.60	solute carrier family 39 (metal ion transporter), member 8
105	1422875_at	Cd84	-0.73	0.60	CD84 antigen
106	1421852_at	Kcnk5	-0.73	0.60	potassium channel, subfamily K, member 5
107	**1416332_at**	**Cirbp**	-0.72	0.61	**cold inducible RNA binding protein**
108	1453470_a_at	Gna13	-0.72	0.61	guanine nucleotide binding protein, alpha 13
109	1451353_at	Tm6sf1	-0.71	0.61	transmembrane 6 superfamily member 1
110	1438502_x_at	Rps17	-0.71	0.61	predicted gene, 100040123
111	1452429_s_at	Abcf1	-0.71	0.61	ATP-binding cassette, sub-family F (GCN20), member 1
112	1435572_at	2310014L17Rik	-0.71	0.61	RIKEN cDNA 2310014L17 gene
113	1416333_at	Dok2	-0.70	0.62	docking protein 2
114	1430533_a_at	Ctnnb1	-0.70	0.62	catenin (cadherin associated protein), beta 1
115	1450379_at	Msn	-0.69	0.62	moesin
116	1421912_at	Slc23a1	-0.69	0.62	solute carrier family 23 (nucleobase transporters), member 1
117	1420386_at	Seh1l	-0.68	0.62	SEH1-like (S. cerevisiae)
118	**1438398_at**	**Rbm39**	-0.68	0.62	**RNA binding motif protein 39**
119	**1450971_at**	**Gadd45b**	-0.68	0.63	**growth arrest and DNA-damage-inducible 45 beta**
120	1433589_at	D6Wsu116e	-0.68	0.63	DNA segment, Chr 6, Wayne State University 116, expressed
121	1420974_at	1810034K20Rik	-0.67	0.63	RIKEN cDNA 1810034K20 gene
122	1448381_at	Gfm1	-0.67	0.63	G elongation factor, mitochondrial 1
123	1418826_at	Ms4a6b	-0.67	0.63	membrane-spanning 4-domains, subfamily A, member 6B
124	1438076_at	Rpl30	-0.66	0.63	ribosomal protein L30
125	1420513_at	Efcab2	-0.66	0.63	EF-hand calcium binding domain 2
126	1449025_at	Ifit3	-0.66	0.63	interferon-induced protein with tetratricopeptide repeats 3
127	1450229_at	Crsp2	-0.66	0.63	mediator complex subunit 14
128	1418911_s_at	Acsl4	-0.65	0.64	acyl-CoA synthetase long-chain family member 4
129	1436547_at	Dgke	-0.65	0.64	diacylglycerol kinase, epsilon
130	1459880_at	Hnrpab	-0.65	0.64	alanine-glyoxylate aminotransferase 2-like 2
131	1450105_at	Adam10	-0.65	0.64	a disintegrin and metallopeptidase domain 10
132	1425098_at	Zfp106	-0.65	0.64	zinc finger protein 106
133	1455195_at	Rps24	-0.65	0.64	ribosomal protein S24
134	1418635_at	Etv3	-0.65	0.64	ets variant gene 3
135	1417781_at	Lass4	-0.65	0.64	LAG1 homolog, ceramide synthase 4
136	1425471_x_at	NA	-0.65	0.64	NA
137	AFFX-18SRNAMur/ X00686_M_at	NA	-0.64	0.64	NA
138	1435270_x_at	2510005D08Rik	-0.64	0.64	N-6 adenine-specific DNA methyltransferase 2 (putative)
139	1438674_a_at	Sfrs8	-0.64	0.64	splicing factor, arginine/serine-rich 8
140	1449360_at	Csf2rb2	-0.64	0.64	colony stimulating factor 2 receptor, beta 2, low-affinity (granulocyte-macrophage)
141	1424936_a_at	Dnahc8	-0.64	0.64	dynein, axonemal, heavy chain 8
142	1436512_at	Arl4c	-0.63	0.64	ADP-ribosylation factor-like 4C

143	1417884_at	Slc16a6	-0.63	0.65	solute carrier family 16 (monocarboxylic acid transporters), member 6
144	1420928_at	St6gal1	-0.62	0.65	beta galactoside alpha 2,6 sialyltransferase 1
145	1427763_a_at	Camk2d	-0.62	0.65	calcium/calmodulin-dependent protein kinase II, delta
146	1424927_at	Glipr1	-0.62	0.65	GLI pathogenesis-related 1 (glioma)
147	1438675_at	Sfrs8	-0.62	0.65	splicing factor, arginine/serine-rich 8
148	1419911_at	Coro1c	-0.62	0.65	coronin, actin binding protein 1C
149	1460239_at	Tspan13	-0.62	0.65	tetraspanin 13
150	1451004_at	Acvr2a	-0.62	0.65	activin receptor IIA
151	1419236_at	Helb	-0.62	0.65	helicase (DNA) B
152	1418482_at	Cyb561d2	-0.62	0.65	cytochrome b-561 domain containing 2
153	1420708_at	Traip	-0.62	0.65	TRAF-interacting protein
154	1433530_at	Rpl41	-0.61	0.65	ribosomal protein L41
155	1426603_at	Rnasel	-0.61	0.65	ribonuclease L (2', 5'-oligoisoadenylate synthetase-dependent)
156	1430523_s_at	Igl-V1	-0.61	0.66	immunoglobulin lambda chain, variable 1
157	1437101_at	Lats2	-0.61	0.66	large tumor suppressor 2
158	1427540_at	Zwint	-0.61	0.66	ZW10 interactor
159	1418660_at	Clock	-0.60	0.66	circadian locomoter output cycles kaput
160	1457455_at	Suhw4	-0.60	0.66	suppressor of hairy wing homolog 4 (Drosophila)
161	1416191_at	Sec61a1	-0.60	0.66	Sec61 alpha 1 subunit (S. cerevisiae)
162	1424687_at	2700008B19Rik	-0.60	0.66	HEAT repeat containing 6
163	1423161_s_at	Spred1	-0.60	0.66	sprouty protein with EVH-1 domain 1, related sequence
164	1433758_at	Nisch	-0.60	0.66	nischarin
165	1427145_at	Iqsec1	-0.60	0.66	IQ motif and Sec7 domain 1
166	1422646_at	Mga	-0.60	0.66	MAX gene associated
167	1424414_at	Ogfrl1	-0.60	0.66	opioid growth factor receptor-like 1
168	1420161_at	AA409749	-0.60	0.66	expressed sequence AA409749
169	1420117_at	NA	-0.59	0.66	NA
170	1425598_a_at	Lyn	-0.59	0.66	Yamaguchi sarcoma viral (v-yes-1) oncogene homolog
171	1428141_at	Gga2	-0.59	0.66	golgi associated, gamma adaptin ear containing, ARF binding protein 2
172	1427629_at	Ptprj	-0.59	0.66	protein tyrosine phosphatase, receptor type, J
173	1438064_at	Ybx1	-0.59	0.66	Y box protein 1
174	1420022_s_at	Suz12	-0.58	0.67	NA
175	1429144_at	Prei4	-0.58	0.67	preimplantation protein 4
176	1417084_at	Eif4ebp2	-0.58	0.67	eukaryotic translation initiation factor 4E binding protein 2
177	1451127_at	AW146242	-0.58	0.67	expressed sequence AW146242
178	1437729_at	Rpl27a	-0.58	0.67	ribosomal protein L27a
179	1418598_at	Ubox5	-0.58	0.67	U box domain containing 5
180	1438427_at	4932442K08Rik	-0.58	0.67	family with sequence similarity 120, member B
181	1438860_a_at	Slc44a2	-0.58	0.67	solute carrier family 44, member 2
182	1456447_at	Rpl18	-0.58	0.67	ribosomal protein L18
183	1417483_at	Nfkbiz	-0.58	0.67	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta
184	1450045_at	Srrm1	-0.57	0.67	serine/arginine repetitive matrix 1
185	1435288_at	Coro1a	-0.57	0.67	coronin, actin binding protein 1A
186	1427146_at	AI790298	-0.57	0.67	ATP-grasp domain containing 1
187	1437566_at	Gnl2	-0.57	0.67	guanine nucleotide binding protein-like 2 (nucleolar)
188	1425458_a_at	Grb10	-0.57	0.67	growth factor receptor bound protein 10
189	1427703_at	Pafah1b1	-0.57	0.67	platelet-activating factor acetylhydrolase, isoform 1b, beta1 subunit
190	1422198_a_at	Shmt1	-0.57	0.67	serine hydroxymethyltransferase 1 (soluble)
191	1450951_at	Smc3	-0.57	0.67	structural maintenace of chromosomes 3
192	1421998_at	Tor3a	-0.57	0.68	torsin family 3, member A
193	1427797_s_at	Ctse	-0.57	0.68	NA
194	1418495_at	Zc3h8	-0.57	0.68	zinc finger CCCH type containing 8
195	1426555_at	Scpep1	-0.57	0.68	serine carboxypeptidase 1
196	1446147_at	C79248	-0.57	0.68	RNA binding motif protein 39
197	1451421_a_at	Rogdi	-0.56	0.68	rogdi homolog (Drosophila)
198	1449098_a_at	Poli	-0.56	0.68	polymerase (DNA directed), iota
199	1450276_a_at	Scin	-0.56	0.68	scinderin
200	1424303_at	Depdc7	-0.56	0.68	DEP domain containing 7
201	1439548_at	Rap2b	-0.56	0.68	RAP2B, member of RAS oncogene family
202	1452688_at	Prpf39	-0.56	0.68	PRP39 pre-mRNA processing factor 39 homolog (yeast)
203	1417782_at	Lass4	-0.56	0.68	LAG1 homolog, ceramide synthase 4
204	1451496_at	Mtss1	-0.56	0.68	metastasis suppressor 1
205	1429400_at	Clcn5	-0.56	0.68	chloride channel 5
206	1452941_at	0610038F07Rik	-0.56	0.68	RIKEN cDNA 0610038F07 gene
207	1449870_a_at	Atp6v0a2	-0.56	0.68	ATPase, H+ transporting, lysosomal V0 subunit A2
208	**1456386_at**	**Rbm39**	-0.55	0.68	**RNA binding motif protein 39**
209	**1448890_at**	**Klf2**	-0.55	0.68	**Kruppel-like factor 2 (lung)**
210	1416029_at	Klf10	-0.55	0.68	Kruppel-like factor 10
211	1436898_at	Sfpq	-0.55	0.68	splicing factor proline/glutamine rich (polypyrimidine tract binding protein associated)
212	1417434_at	Gpd2	-0.55	0.68	glycerol phosphate dehydrogenase 2, mitochondrial
213	1430986_at	Farslb	-0.55	0.68	phenylalanyl-tRNA synthetase, beta subunit
214	1417203_at	Ethe1	-0.55	0.68	ethylmalonic encephalopathy 1
215	**1420659_at**	**Slamf6**	-0.55	0.68	**SLAM family member 6**
216	1415913_at	Rps13	-0.55	0.68	ribosomal protein S13

217	1452545_a_at	Itgb1	-0.54	0.69	integrin beta 1 (fibronectin receptor beta)
218	1452402_at	Uchl3	-0.54	0.69	NA
219	1451629_at	Lbh	-0.54	0.69	limb-bud and heart
220	1448471_a_at	Ctla2a	-0.54	0.69	cytotoxic T lymphocyte-associated protein 2 alpha
221	1455517_at	Rbm4	-0.54	0.69	RNA binding motif protein 4
222	1424375_s_at	Gimap4	-0.54	0.69	GTPase, IMAP family member 4
223	**1449773_s_at**	**Gadd45b**	-0.54	0.69	**growth arrest and DNA-damage-inducible 45 beta**
224	1452050_at	Camk1d	-0.54	0.69	calcium/calmodulin-dependent protein kinase ID
225	1437658_a_at	Snord22	-0.54	0.69	small nucleolar RNA, C/D box 22
226	1455987_at	Sec61a1	-0.54	0.69	Sec61 alpha 1 subunit (S. cerevisiae)
227	1421127_at	Tmem42	-0.54	0.69	transmembrane protein 42
228	1423555_a_at	Ifi44	-0.54	0.69	interferon-induced protein 44
229	**1436589_x_at**	**Prkd2**	-0.53	0.69	**protein kinase D2**
230	1450075_at	Polh	-0.53	0.69	polymerase (DNA directed), eta (RAD 30 related)
231	1423812_s_at	AW146242	-0.53	0.69	expressed sequence AW146242
232	1419372_at	Gosr2	-0.53	0.69	golgi SNAP receptor complex member 2
233	1419967_at	Seh1l	-0.53	0.69	SEH1-like (S. cerevisiae
234	1421910_at	Tcf20	-0.53	0.69	transcription factor 20
235	1449621_s_at	Thsd1	-0.53	0.69	thrombospondin, type I, domain 1
236	1427097_at	Wwp1	-0.52	0.70	WW domain containing E3 ubiquitin protein ligase 1
237	1451469_at	D530005L17Rik	-0.52	0.70	centlein, centrosomal protein
238	1450648_s_at	H2-Ab1	-0.52	0.70	response to metastatic cancers 5
239	1426287_at	Atxn7	-0.52	0.70	ataxin 7
240	1419644_at	Cstf2	-0.52	0.70	cleavage stimulation factor, 3' pre-RNA subunit 2
241	1423170_at	Taf7	-0.52	0.70	TAF7 RNA polymerase II, TATA box binding protein (TBP)-associated factor
242	1425617_at	Dhx9	-0.52	0.70	DEAH (Asp-Glu-Ala-His) box polypeptide 9
243	1442484_at	Syncrip	-0.52	0.70	DNA segment, Chr 9, ERATO Doi 306, expressed
244	1452629_at	Safb2	-0.52	0.70	scaffold attachment factor B2
245	**1435945_a_at**	**Kcnn4**	-0.52	0.70	**potassium intermediate/small conductance calcium-activated channel, subfamily N, member 4**

PAX5/ETV6-ACTIVATED GENES [log2(signal)>0.59 = Fold change >1.5]

	Affymetrix probe ID	gene name	log2(signal)	FC	gene description
246	1426812_a_at	9130404D14Rik	0.59	1.51	family with sequence similarity 129, member B
247	1424189_at	Pigc	0.60	1.51	phosphatidylinositol glycan anchor biosynthesis, class C
248	1451975_at	2810453I06Rik	0.60	1.52	RIKEN cDNA 2810453I06 gene
249	1425534_at	Stau2	0.61	1.52	staufen (RNA binding protein) homolog 2 (Drosophila)
250	1448185_at	Herpud1	0.61	1.52	homocysteine-inducible, endoplasmic reticulum stress-inducible, ubiquitin-like domain member 1
251	1419239_at	Zfp54	0.61	1.53	zinc finger protein 54
252	1435176_a_at	Id2	0.61	1.53	inhibitor of DNA binding 2
253	1455938_x_at	Rad21	0.62	1.53	RAD21 homolog (S. pombe)
254	1449370_at	Sox4	0.62	1.54	SRY-box containing gene 4
255	1422444_at	Itga6	0.62	1.54	integrin alpha 6
256	**1415812_at**	**Gsn**	0.62	1.54	**gelsolin**
257	1418108_at	Plekhk1	0.62	1.54	rhotekin 2
258	1421594_a_at	Sytl2	0.63	1.55	synaptotagmin-like 2
259	**1419466_at**	**Nkd2**	0.63	1.55	**naked cuticle 2 homolog (Drosophila)**
260	1415972_at	Marcks	0.63	1.55	myristoylated alanine rich protein kinase C substrate
261	1431609_a_at	Acp5	0.63	1.55	acid phosphatase 5, tartrate resistant
262	1428794_at	Specc1	0.63	1.55	cytospin B
263	1419003_at	Bves	0.64	1.56	blood vessel epicardial substance
264	1426563_at	Zfp553	0.65	1.57	zinc finger protein 553
265	1424418_at	BC010801	0.66	1.58	solute carrier family 25, member 38
266	1424633_at	Camk1g	0.66	1.58	calcium/calmodulin-dependent protein kinase I gamma
267	1438865_at	H13	0.66	1.58	histocompatibility 13
268	1418536_at	H2-Q8	0.66	1.58	histocompatibility 2, Q region locus 7
269	1449310_at	Ptger2	0.66	1.58	prostaglandin E receptor 2 (subtype EP2)
270	1424706_at	Zfp51	0.66	1.58	zinc finger protein 51
271	1452679_at	Tubb2b	0.66	1.58	tubulin, beta 2B
272	1420398_at	Rgs18	0.66	1.58	regulator of G-protein signaling 18
273	1460651_at	Lat	0.67	1.59	linker for activation of T cells
274	**1436991_x_at**	**Gsn**	0.68	1.60	**gelsolin**
275	1418380_at	Terf1	0.69	1.61	telomeric repeat binding factor 1
276	1419810_x_at	Arhgap9	0.69	1.61	Rho GTPase activating protein 9
277	**1452462_a_at**	**Banp**	0.69	1.62	**BTG3 associated nuclear protein**
278	1420772_a_at	Tsc22d3	0.70	1.62	TSC22 domain family, member 3
279	**1425396_a_at**	**Lck**	0.71	1.63	**lymphocyte protein tyrosine kinase**
280	1455531_at	A930031D07Rik	0.71	1.63	major facilitator superfamily domain containing 4
281	1455551_at	Uevld	0.71	1.64	UEV and lactate/malate dehyrogenase domains
282	1424831_at	Cpne2	0.72	1.65	copine II
283	1435941_at	Rhbdl3	0.72	1.65	rhomboid, veinlet-like 3 (Drosophila)
284	1451765_a_at	Entpd5	0.73	1.66	ectonucleoside triphosphate diphosphohydrolase 5
285	**1450680_at**	**Rag1**	0.73	1.66	**recombination activating gene 1**
286	1449514_at	Gprk5	0.73	1.66	G protein-coupled receptor kinase 5

287	1419358_at	Sorcs2	0.74	1.68	sortilin-related VPS10 domain containing receptor 2
288	1422804_at	Serpinb6b	0.76	1.69	serine (or cysteine) peptidase inhibitor, clade B, member 6b
289	**1424039_at**	**Tmem66**	0.76	1.69	**transmembrane protein 66**
290	1449383_at	Adssl1	0.76	1.70	adenylosuccinate synthetase like 1
291	1424249_a_at	Arhgap9	0.77	1.70	Rho GTPase activating protein 9
292	1421840_at	Abca1	0.77	1.70	ATP-binding cassette, sub-family A (ABC1), member 1
293	1451972_at	Glcci1	0.78	1.72	similar to glucocorticoid induced transcript 1
294	1455869_at	Camk2b	0.80	1.74	NA
295	1426968_a_at	Rdh10	0.81	1.75	retinol dehydrogenase 10 (all-trans)
296	1421592_at	Ncam2	0.81	1.76	neural cell adhesion molecule 2
297	1425128_at	B3gnt8	0.82	1.77	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 8
298	1417444_at	E2f5	0.84	1.78	E2F transcription factor 5
299	**1424567_at**	**Tspan2**	0.84	1.79	**tetraspanin 2**
300	1416268_at	Ets2	0.84	1.79	E26 avian leukemia oncogene 2, 3' domain
301	1448754_at	Rbp1	0.86	1.82	retinol binding protein 1, cellular
302	1451644_a_at	H2-Q1	0.86	1.82	MHC class I like protein GS10
303	**1422701_at**	**Zap70**	0.87	1.82	**zeta-chain (TCR) associated protein kinase**
304	**1432417_a_at**	**Tspan2**	0.87	1.83	**tetraspanin 2**
305	1427139_at	Adamts10	0.87	1.83	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 10
306	1436021_at	A930031D07Rik	0.88	1.84	major facilitator superfamily domain containing 4
307	1451475_at	Plxnd1	0.89	1.85	plexin D1
308	1452343_at	D18Ertd653e	0.89	1.86	DNA segment, Chr 18, ERATO Doi 653, expressed
309	**1448383_at**	**Mmp14**	0.89	1.86	**matrix metallopeptidase 14 (membrane-inserted)**
310	1428122_s_at	2610528K11Rik	0.92	1.89	family with sequence similarity 125, member B
311	**1438511_a_at**	**1190002H23Rik / RGC-32**	0.93	1.90	**RIKEN cDNA 1190002H23 gene**
312	1427216_at	Ifnz	0.93	1.90	interferon zeta
313	**1427144_at**	**Hnrpll**	0.97	1.96	**heterogeneous nuclear ribonucleoprotein L-like**
314	1451181_at	Tmem121	0.97	1.96	transmembrane protein 121
315	1423893_x_at	Apbb1	0.97	1.96	amyloid beta (A4) precursor protein-binding, family B, member 1
316	1423135_at	Thy1	0.99	1.98	thymus cell antigen 1, theta
317	**1425255_s_at**	**Hnrpll**	1.02	2.02	**heterogeneous nuclear ribonucleoprotein L-like**
318	1451260_at	Aldh1b1	1.02	2.03	aldehyde dehydrogenase 1 family, member B1
319	1422811_at	Slc27a1	1.04	2.06	solute carrier family 27 (fatty acid transporter), member 1
320	1421965_s_at	Notch3	1.07	2.10	Notch gene homolog 3 (Drosophila)
321	1424394_at	Selm	1.09	2.13	selenoprotein M
322	**1448507_at**	**Efhd1**	1.15	2.22	**EF hand domain containing 1**
323	**1448194_a_at**	**H19**	1.34	2.53	**H19 fetal liver mRNA**
324	1420399_at	Gfi1b	1.35	2.55	growth factor independent 1B
325	1417694_at	Gab1	1.37	2.59	growth factor receptor bound protein 2-associated protein 1
326	**1448390_a_at**	**Dhrs3**	1.46	2.75	**dehydrogenase/reductase (SDR family) member 3**
327	**1427029_at**	**Htra3**	1.46	2.75	**HtrA serine peptidase 3**
328	1418726_a_at	Tnnt2	1.48	2.78	troponin T2, cardiac
329	1424967_x_at	Tnnt2	1.49	2.80	troponin T2, cardiac
330	**1417943_at**	**Gng4**	1.57	2.97	**guanine nucleotide binding protein (G protein), gamma 4**
331	1424404_at	0610040J01Rik	1.57	2.97	RIKEN cDNA 0610040J01 gene
332	**1417481_at**	**Ramp1**	1.85	3.59	**receptor (calcitonin) activity modifying protein 1**
333	1416221_at	Fstl1	1.87	3.64	follistatin-like 1
334	1421425_a_at	Dscr1l1	1.87	3.66	regulator of calcineurin 2
335	**1419292_at**	**Htra3**	1.91	3.75	**HtrA serine peptidase 3**
336	1448259_at	Fstl1	1.98	3.94	follistatin-like 1
337	1416882_at	Rgs10	2.17	4.50	regulator of G-protein signalling 10
338	1417917_at	Cnn1	2.58	5.97	calponin 1
339	1427387_a_at	Itgb4	3.27	9.66	integrin beta 4
340	**1452136_at**	**Slc5a9**	3.35	10.20	**solute carrier family 5 (sodium/glucose cotransporter), member 9**

Supplementary Table S1. PAX5/ETV6 repressed and activated genes. Probes in bold have been successfully validated. Grey, PAX5-activated target genes; yellow, PAX5-repressed target genes; green, ETV6-target genes. FC, fold change.

Table S2

NEITHER PAX5- NOR ETV6- TARGET GENES

	GEP		LY5.1FL		FLB6-67		B6BAFL	
gene name	log2(signal)	FC	FC	p value	FC	p value	FC	p value
SLCO1B2[1]	-4.08	0.06	0.06	0.004977	0.07	0.007706	0.08	0.000023
LTA	-3.48	0.09	0.17	0.000135	0.15	0.000037	0.17	0.000002
RHOB	-1.10	0.47	0.18	0.000018	0.36	0.000744	0.48	0.000847
LTB	-0.99	0.50	0.30	0.000069	0.14	0.000321	0.32	0.000120
TXNRD3[1]	-0.99	0.50	0.48	0.000196	0.44	0.000759	0.57	0.001253
CD37	-0.88	0.54	0.67	0.000359	0.47	0.002114	0.48	0.007250
RBM39[1]	-0.79	0.58	0.62	0.001205	0.69	0.004662	0.71	0.018462
GADD45b[1]	-0.54	0.69	0.58	0.000093	0.39	0.0000001	0.72	0.001440
PAX5-ACTIVATED TARGET								
IFI30	-1.68	0.31	0.29	0.002660	0.37	0.000051	0.37	0.000070
BCAR3	-1.55	0.34	0.32	0.000003	0.24	0.000395	0.32	0.000833
CD22	-1.39	0.38	0.07	0.000034	0.29	0.006058	0.42	0.003456
MGST1	-1.01	0.50	0.17	0.000202	0.45	0.001099	0.53	0.000885
SLAMF6	-0.55	0.68	0.23	0.000812	0.59	0.002125	0.57	0.014589
KLF2	-0.55	0.68	0.30	0.000002	0.38	0.000096	0.64	0.000027
PRKD2	-0.53	0.69	0.43	0.000001	0.70	0.000042	0.71	0.002661
CD44[2]	-0.48	0.72	0.60	0.004632	0.60	0.000429	0.68	0.005192
BCAT1[1-2]	-0.45	0.73	0.65	0.000166	0.71	0.000081	0.78	0.011586
CAPN2[2]	-0.43	0.74	0.41	0.000115	0.51	0.000023	0.81	0.019466
IRF8[1-2]	-0.42	0.75	0.65	0.000965	0.34	0.000050	0.63	0.004453
SDC4[2]	-0.39	0.76	0.61	0.012851	0.40	0.000450	0.59	0.003626
PLEKHA2[1-2]	-0.37	0.77	0.46	0.000025	0.51	0.000025	0.79	0.000577
LCP2[2]	-0.33	0.80	0.33	0.002490	0.55	0.000388	0.67	0.022835
NEDD9[1-2]	-0.27	0.83	0.64	0.003254	0.34	0.000001	0.54	0.004976
PAX5-REPRESSED TARGET								
KCNN4	-0.52	0.70	0.52	0.000033	0.34	0.000328	0.72	0.000738
FXYD5[2]	-0.47	0.72	0.82	0.039443	0.61	0.002052	0.77	0.000038
MYLC2PL[2]	-0.44	0.74	0.04	0.000001	0.18	0.000250	0.60	0.010091
ETV6-TARGET								
CIRBP	-0.72	0.61	0.31	0.040072	0.59	0.039270	0.76	0.002304
LDLR	-0.84	0.56	0.51	0.014012	0.64	0.000169	0.57	0.007251

Supplementary Table S2. Validation results of PAX5/ETV6-repressed genes (RQ-PCR).[1]genes with more than one probe; [2] genes listed in supplementary table S4.

Table S3

NEITHER PAX5- NOR ETV6- TARGET GENES

	GEP		LY5.1FL		FLB6-67		B6BAFL	
gene name	log2(signal)	FC	FC	p value	FC	p value	FC	p value
HTRA3[1]	1.91	3.75	3.95	0.001640	2.12	0.002921	2.47	0.001428
GNG4	1.57	2.97	6.23	0.001687	3.66	0.000002	2.07	0.000155
EFHD1	1.15	2.22	1265.88	0.000197	1.57	0.009830	1.38	0.000062
HNRPLL[1]	1.02	2.02	1.55	0.152580	2.14	0.014574	2.01	0.006791
MMP14	0.89	1.86	1.90	0.048041	2.15	0.004762	1.45	0.001582
ZAP70	0.87	1.82	6.87	0.000001	2.57	0.000028	2.50	0.000004
RAG1	0.73	1.66	20.25	0.000012	1.32	0.011905	1.68	0.000193
BANP	0.69	1.62	2.96	0.000001	2.01	0.001173	1.43	0.011338
NKD2	0.63	1.55	2.02	0.028243	1.26	0.008172	1.39	0.009511
PAX5-REPRESSED TARGET								
SLC5A9	3.35	10.2	7.91	0.006501	10.89	0.000793	8.47	0.029866
RAMP1A	1.85	3.59	5.36	0.000190	2.63	0.001441	4.31	0.002162
RAMP1B	1.85	3.59	4.93	0.000068	1.69	0.003354	3.12	0.005552
DHRS3	1.46	2.75	4.14	0.000002	1.27	0.013871	2.01	0.000270
H19	1.34	2.53	7.02	0.000018	1.42	0.000028	2.15	0.001410
RGC32[1]	0.93	1.90	2.47	0.000004	1.43	0.005911	1.70	0.000703
TSPAN2[1]	0.84	1.79	7.36	0.000001	1.60	0.000160	1.79	0.004058
TMEM66	0.76	1.69	8.79	0.000242	1.40	0.044344	1.35	0.010556
LCK	0.71	1.63	2.04	0.000070	1.78	0.000050	1.55	0.000630
EMP1[2]	0.46	1.38	6.71	0.000228	1.61	0.033579	1.69	0.000773
PAX5-ACTIVATED TARGET								
GSN[1]	0.68	1.60	3.51	0.000125	1.82	0.000275	1.76	0.008839
IFI27[2]	0.34	1.27	3.63	0.000001	1.63	0.026759	1.91	0.000180

Supplementary Table S3. Validation results of PAX5/ETV6-activated genes (RQ-PCR).[1]genes with more than one probe; [2] genes listed in supplementary table S4.

Table S4

PAX5/ETV6-REPRESSED GENES [log2(signal)<-0.15 = Fold change <0.90]

	Affymetrix probe ID	gene name	log2(signal)	FC	regulation by PAX5	gene description
1	1416008_at	Satb1	-0.49	0.71	*PAX5-repressed target*	special AT-rich sequence binding protein 1
2	**1423760_at**	**Cd44**	-0.48	0.72	***PAX5-activated target***	**CD44 antigen**
3	**1418296_at**	**Fxyd5**	-0.47	0.72	***PAX5-repressed target***	**FXYD domain-containing ion transport regulator 5**
4	**1430111_a_at**	**Bcat1**	-0.45	0.73	***PAX5-activated target***	**branched chain aminotransferase 1, cytosolic**
5	**1416257_at**	**Capn2**	-0.43	0.74	***PAX5-activated target***	**calpain 2**
6	1437667_a_at	Bach2	-0.43	0.74	*PAX5-activated target*	BTB and CNC homology 2
7	**1420805_at**	**Mylc2pl**	-0.44	0.74	***PAX5-repressed target***	**myosin light chain 2, precursor lymphocyte-specific**
8	**1448452_at**	**Irf8**	-0.42	0.75	***PAX5-activated target***	**interferon regulatory factor 8**
9	1421027_a_at	Mef2c	-0.41	0.75	*PAX5-repressed target*	myocyte enhancer factor 2C
10	**1448793_a_at**	**Sdc4**	-0.39	0.76	***PAX5-activated target***	**syndecan 4**
11	**1417289_at**	**Plekha2**	-0.37	0.77	***PAX5-activated target***	**pleckstrin homology domain-containing, family A (phosphoinositide binding specific) member 2**
12	1436212_at	Tmem71	-0.38	0.77	*PAX5-activated target*	transmembrane protein 71
13	**1450871_a_at**	**Bcat1**	-0.36	0.78	***PAX5-activated target***	**branched chain aminotransferase 1, cytosolic**
14	1429265_a_at	Rnf130	-0.36	0.78	*PAX5-repressed target*	ring finger protein 130
15	1460711_at	4930461P20Rik	-0.33	0.80	*PAX5-activated target*	RIKEN cDNA 4930461P20 gene
16	**1418641_at**	**Lcp2**	-0.33	0.80	***PAX5-activated target***	**lymphocyte cytosolic protein 2**
17	1436819_at	Sept6	-0.32	0.80	*PAX5-repressed target*	septin 6
18	1421733_a_at	Tpst1	-0.30	0.81	*PAX5-activated target*	protein-tyrosine sulfotransferase 1
19	1424374_at	Gimap4	-0.30	0.81	*PAX5-activated target*	GTPase, IMAP family member 4
20	1416007_at	Satb1	-0.31	0.81	*PAX5-repressed target*	special AT-rich sequence binding protein 1
21	1451507_at	Mef2c	-0.29	0.82	*PAX5-repressed target*	myocyte enhancer factor 2C
22	1452097_a_at	Dusp7	-0.27	0.83	*PAX5-activated target*	dual specificity phosphatase 7
23	**1422818_at**	**Nedd9**	-0.27	0.83	***PAX5-activated target***	**neural precursor cell expressed, developmentally down-regulated gene 9**
24	1423890_x_at	Atp1b1	-0.27	0.83	*PAX5-activated target*	ATPase, Na+/K+ transporting, beta 1 polypeptide
25	1416049_at	Gldc	-0.26	0.83	*PAX5-activated target*	glycine decarboxylase
26	1436890_at	Uap1l1	-0.26	0.83	*PAX5-activated target*	UDP-N-acteylglucosamine pyrophosphorylase 1-like 1
27	1415903_at	Slc38a1	-0.28	0.83	*PAX5-repressed target*	solute carrier family 38, member 1
28	**1417288_at**	**Plekha2**	-0.26	0.84	***PAX5-activated target***	**pleckstrin homology domain-containing, family A (phosphoinositide binding specific) member 2**
29	1429428_at	Tcf7l2	-0.25	0.84	*PAX5-activated target*	transcription factor 7-like 2, T-cell specific, HMG-box
30	1423431_a_at	Mybbp1a	-0.26	0.84	*PAX5-repressed target*	MYB binding protein (P160) 1a
31	1438968_x_at	Spint2	-0.23	0.85	*PAX5-repressed target*	serine protease inhibitor, Kunitz type 2
32	1448160_at	Lcp1	-0.24	0.85	*PAX5-repressed target*	lymphocyte cytosolic protein 1
33	**1437132_x_at**	**Nedd9**	-0.22	0.86	***PAX5-activated target***	**neural precursor cell expressed, developmentally down-regulated gene 9**
34	1453472_a_at	Slamf7	-0.21	0.86	*PAX5-activated target*	SLAM family member 7
35	**1416714_at**	**Irf8**	-0.21	0.86	***PAX5-activated target***	**interferon regulatory factor 8**
36	1426083_a_at	Btg1	-0.21	0.86	*PAX5-activated target*	B-cell translocation gene 1, anti-proliferative
37	1451152_a_at	Atp1b1	-0.21	0.86	*PAX5-activated target*	ATPase, Na+/K+ transporting, beta 1 polypeptide
38	1416162_at	Rad21	-0.21	0.86	*PAX5-activated target*	RAD21 homolog (S. pombe)
39	1418465_at	Ncf4	-0.21	0.86	*PAX5-repressed target*	neutrophil cytosolic factor 4
40	1450655_at	Pten	-0.22	0.86	*PAX5-repressed target*	phosphatase and tensin homolog
41	1423430_at	Mybbp1a	-0.22	0.86	*PAX5-repressed target*	MYB binding protein (P160) 1a
42	1421028_a_at	Mef2c	-0.21	0.87	*PAX5-repressed target*	myocyte enhancer factor 2C
43	1418453_a_at	Atp1b1	-0.18	0.88	*PAX5-activated target*	ATPase, Na+/K+ transporting, beta 1 polypeptide
44	1436421_s_at	Arpc5l	-0.19	0.88	*PAX5-activated target*	actin related protein 2/3 complex, subunit 5-like
45	**1450570_a_at**	**Cd19**	-0.17	0.89	***PAX5-activated target***	**CD19 antigen**
46	1460716_a_at	Cbfb	-0.16	0.89	*PAX5-activated target*	core binding factor beta
47	1452708_a_at	Pabpc1	-0.17	0.89	*PAX5-repressed target*	poly A binding protein, cytoplasmic 1
48	1420196_s_at	Tbc1d14	-0.15	0.9	*PAX5-activated target*	TBC1 domain family, member 14
49	1415827_a_at	D3Ucla1	-0.15	0.9	*PAX5-activated target*	DNA segment, Chr 3, University of California at Los Angeles 1
50	1416630_at	Id3	-0.15	0.90	*PAX5-activated target*	inhibitor of DNA binding 3
51	1416723_at	Tcf4	-0.16	0.90	*PAX5-activated target*	transcription factor 4
52	1454722_at	Pten	-0.16	0.90	*PAX5-repressed target*	phosphatase and tensin homolog

PAX5/ETV6-ACTIVATED GENES [log2(signal)>0.14 = Fold change >1.10]

	Affymetrix probe ID	gene name	log2(signal)	FC	regulation by PAX5	gene description
1	1453569_s_at	Mark4	0.14	1.10	*PAX5-repressed target*	MAP/microtubule affinity-regulating kinase 4
2	1425514_at	Pik3r1	0.14	1.10	*PAX5-repressed target*	phosphatidylinositol 3-kinase, regulatory subunit, polypeptide 1 (p85 alpha)
3	1460407_at	Spib	0.14	1.10	*PAX5-activated target*	Spi-B transcription factor (Spi-1/PU.1 related)
4	1448118_a_at	Ctsd	0.15	1.11	*PAX5-repressed target*	cathepsin D
5	1419083_at	Tnfsf11	0.15	1.11	*PAX5-repressed target*	tumor necrosis factor (ligand) superfamily, member 11
6	1448361_at	Ttc3	0.15	1.11	*PAX5-repressed target*	tetratricopeptide repeat domain 3
7	1448182_a_at	Cd24a	0.16	1.12	*PAX5-activated target*	CD24a antigen
8	1417621_at	Nfatc1	0.16	1.12	*PAX5-activated target*	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 1
9	1424753_at	Nudt14	0.16	1.12	*PAX5-repressed target*	nudix (nucleoside diphosphate linked moiety X)-type motif 14
10	1417786_a_at	Rgs19	0.16	1.11	*PAX5-repressed target*	regulator of G-protein signaling 19
11	1415818_at	Anxa6	0.17	1.13	*PAX5-repressed target*	annexin A6
12	1450168_at	Ankrd12	0.18	1.13	*PAX5-activated target*	ankyrin repeat domain 12
13	1418634_at	Notch1	0.18	1.13	*PAX5-repressed target*	Notch gene homolog 1 (Drosophila)
14	1450093_s_at	Zbtb7a	0.18	1.13	*PAX5-activated target*	zinc finger and BTB domain containing 7a
15	1428116_a_at	Dynlt1	0.19	1.14	*PAX5-repressed target*	dynein light chain Tctex-type 1
16	1426236_a_at	Glul	0.20	1.15	*PAX5-activated target*	glutamate-ammonia ligase (glutamine synthetase)
17	1426236_a_at	Glul	0.20	1.15	*PAX5-activated target*	glutamate-ammonia ligase (glutamine synthetase)
18	1420936_s_at	Cpsf2	0.21	1.16	*PAX5-activated target*	cleavage and polyadenylation specific factor 2
19	1452061_s_at	Strbp	0.21	1.16	*PAX5-activated target*	spermatid perinuclear RNA binding protein
20	1417894_at	Gpr97	0.22	1.17	*PAX5-repressed target*	G protein-coupled receptor 97
21	1416484_at	Ttc3	0.22	1.16	*PAX5-repressed target*	tetratricopeptide repeat domain 3
22	1416483_at	Ttc3	0.23	1.18	*PAX5-repressed target*	tetratricopeptide repeat domain 3
23	1421303_at	Zfpn1a1	0.23	1.17	*PAX5-repressed target*	zinc finger protein, subfamily 1A, 1 (Ikaros)
24	1418483_a_at	Ggta1	0.24	1.18	*PAX5-repressed target*	glycoprotein galactosyltransferase alpha 1, 3
25	1448188_at	Ucp2	0.24	1.18	*PAX5-activated target*	uncoupling protein 2 (mitochondrial, proton carrier)
26	1416161_at	Rad21	0.26	1.20	*PAX5-activated target*	RAD21 homolog (S. pombe)
27	1426411_a_at	Strbp	0.26	1.20	*PAX5-activated target*	spermatid perinuclear RNA binding protein
28	1451150_at	Zfp410	0.26	1.2	*PAX5-activated target*	zinc finger protein 410
29	1416424_at	M6prbp1	0.27	1.21	*PAX5-repressed target*	mannose-6-phosphate receptor binding protein 1
30	1425212_a_at	Tnfrsf19	0.27	1.21	*PAX5-activated target*	tumor necrosis factor receptor superfamily, member 19
31	**1456312_x_at**	**Gsn**	0.30	1.23	***PAX5-activated target***	**gelsolin**
32	1418512_at	Stk3	0.31	1.24	*PAX5-activated target*	serine/threonine kinase 3 (Ste20, yeast homolog)
33	**1426278_at**	**Ifi27**	0.34	1.27	***PAX5-activated target***	**interferon, alpha-inducible protein 27**
34	**1437171_x_at**	**Gsn**	0.36	1.28	***PAX5-activated target***	**gelsolin**

35	1448548_at	Tulp4	0.36	1.29	*PAX5-repressed target*	tubby like protein 4
36	1434940_x_at	Rgs19	0.37	1.30	*PAX5-repressed target*	regulator of G-protein signaling 19
37	1417575_at	Otub2	0.39	1.31	*PAX5-activated target*	OTU domain, ubiquitin aldehyde binding 2
38	1448147_at	Tnfrsf19	0.40	1.32	*PAX5-activated target*	tumor necrosis factor receptor superfamily, member 19
39	1428706_at	Prr6	0.41	1.33	*PAX5-repressed target*	proline-rich polypeptide 6
40	1429246_a_at	Anxa6	0.42	1.34	*PAX5-repressed target*	annexin A6
41	**1418003_at**	**1190002H23Rik/ RGC-32**	0.44	1.36	***PAX5-repressed target***	**RIKEN cDNA 1190002H23 gene**
42	1425099_a_at	Arntl	0.44	1.36	*PAX5-activated target*	aryl hydrocarbon receptor nuclear translocator-like
43	**1416529_at**	**Emp1**	0.46	1.38	***PAX5-activated target***	**epithelial membrane protein 1**
44	1423754_at	Ifitm3	0.50	1.41	*PAX5-repressed target*	interferon induced transmembrane protein 3
45	1427600_at	Tnfrsf19	0.50	1.41	*PAX5-activated target*	tumor necrosis factor receptor superfamily, member 19
46	1416930_at	Ly6d	0.57	1.48	*PAX5-repressed target*	lymphocyte antigen 6 complex, locus D

Supplementary Table S4. List of additional PAX5-target genes. Probes in bold have been successfully validated. FC, fold change.

Table S5

PAX5/ETV6-REPRESSED GENES [log2(signal)<-0.15 = Fold change <0.90]

	Affymetrix probe ID	gene name	log2(signal)	FC	regulation by ETV6	gene description
1	1417316_at	Them2	-0.18	0.88	*ETV6-repressed target*	thioesterase superfamily member 2
2	1417922_at	Kbtbd4	-0.18	0.89	*ETV6-repressed target*	kelch repeat and BTB (POZ) domain containing 4

PAX5/ETV6-ACTIVATED GENES [log2(signal)>0.14 = Fold change >1.10]

	Affymetrix probe ID	gene name	log2(signal)	FC	regulation by ETV6	gene description
1	1416419_s_at	Gabarapl1	0.30	1.23	*ETV6-repressed target*	gamma-aminobutyric acid (GABA(A)) receptor-associated protein-like 1
2	1416418_at	Gabarapl1	0.31	1.24	*ETV6-repressed target*	gamma-aminobutyric acid (GABA(A)) receptor-associated protein-like 1
3	1451959_a_at	Vegfa	0.33	1.26	*ETV6-activated target*	vascular endothelial growth factor A
4	1418835_at	Phlda1	0.41	1.33	*ETV6-activated target*	pleckstrin homology-like domain, family A, member 1
5	1423804_a_at	Idi1	0.56	1.47	*ETV6-repressed target*	isopentenyl-diphosphate delta isomerase

Supplementary Table S5. List of additional ETV6-target genes.FC, fold change.

Table S6

	LY5.1FL		FLB6-67		B6BAFL	
gene name	FC	p value	FC	p value	FC	p value
IRF4	0.46	0.000010	0.17	0.000009	0.74	0.015881
IKZF2	0.34	0.011486	0.44	0.000062	0.32	0.000628
IKZF3	0.53	0.000029	0.63	0.004848	0.25	0.001794

Supplementary Table S6. Additional PAX5-target genes involved in BCR (RQ-PCR).

Tabke S7

	LY5.1FL		FLB6-67		B6BAFL	
gene name	FC	p value	FC	p value	FC	p value
EPS8	0.35	0.000080	0.22	0.002869	0.6	0.000001
EDG1	0.12	0.000002	0.22	0.000569	0.53	0.000001

Supplementary Table S7. Additional PAX5-target genes involved in adhesion (RQ-PCR).

Table S8

gene name	UPL #	left primer	right primer
BANP	20	tcaccctgaactccgaagag	ccgcatctcagggttattct
BCAR3	80	gctttctctgccatcatgaaa	agagccgtccatgttttctc
BCAT1	1	caggagtgaccaggcagag	ggtgaggtgtctctcacatacctt
CAPN2	82	tggcttcggcatctatgag	aagtttttgccgaggtggat
CD22	58	tgaccgcagggactgaac	aaggttggaacggtttctcc
CD37	19	gggcatcctcatttccact	cgtcctcaacaccaattcct
CD44	29	gcatcgcggtcaatagtagg	caccgttgatcaccagctt
CIRBP	64	tgatggctatgaatgggaagt	ggaccggttgtcagaagact
DHRS3	9	ccaaagctgtccgagagaa	ctttttccatggaccacagc
EDG1	70	caggatctactccttggtcagg	gagatgttcttgcggaaggt
EFHD1	11	tctggagaagatgttcaaaacg	tttcagctccatcaggtcaa
EMP1	66	caccatgctggtgtgctg	tccaggtcaggatgaaacaa
EPS8	47	agagcaagccagaccttcac	tcgatatcttcactaattaggtttgc
FXYD5	26	ctgcgactactcgtgacaatg	gatgagaggggtggtctgg
GADD45b	10	ctgcctcctggtcacgaa	ttgcctctgctctcttcaca
GNG4	79	ctgcatggacagggtgaa	gcacgtgggcttcacagta
GSN	22	caaagtcgggtgtctgagg	cttccctgccttcaggaat
H19	67	gtctcgaagagctcggactg	actggcaggcacatccac
HNRPLL	72	tgcttgtatggaaatattgaaaagg	tcacccatttctacgagtgct
HPRT	22	ctgcagaagggtgcctgt	gaacttgtaccgcggactg
HTRA3	3	tagccacactccaatcagca	agagcaaggctccaacagc
IFI27	25	gcgttcctaaccatcgtctg	gtgacacctcaggagcatacac
IFI30	56	gacctcacctcaagcacacc	tccagctttacagaatttgtgc
IKZF2	47	cataagcgaagccataccg	gtaaaagatgccccgcact
IKZF3	3	acagcaccttatggctctctg	atggggtggcatcatgtagt
IRF4	17	gtgtggccgctcagagat	ggctcctcttggtcataccc
IRF8	26	tctgcacgctgagatgttgt	accaggagcaggtacagcac
KCNN4	48	ctaaaggcgcatctgcgta	tagtggcgggtaagctcgt
KLF2	21	cgtgtgtgaaaactgccact	gagatccctcataggtgaccag
LCK	83	ggagacagaaagccacgaag	ttgggactttcatagtcatcttca
LCP2	64	gatggctatacctacccctcaa	tgctcatgccacatcgtc
LDLR	62	tccctcagaagcacttgacc	gagttctgcttgctggggta
LTA	76	cctggtgaccctgttgttg	tgctcctgagccaatgatct
LTB	27	gcccttctccctggattc	ggccatcaacacctcattgt
MGST1	42	aacttcgtgttgcctgatga	tttgtgggtgaccctgactt
MMP14	63	aaaggctttgtcaaggctga	gggaaagctgcaaacatctg
MYLC2PL	63	gcggagttggatggatgat	ctccttctgctgtcgctca
NEDD9	17	tgtggaacatcgctcacg	ggcccctccttagggtct
NKD2	33	aaccaagccagcaagatcac	aacttctgtggccaagggta
PAX5	83	gacgctgacagggatggt	ggggaacctccaagaatcat
PAX5/ETV6	33	cctagcgtcagttccatcaac	gcaagcctatgctgtgactg
PLEKHA2	10	gatatccaggagggtgacctg	tctggaaatcctcgaaggtg
PRKD2	46	aggcctgtggagcaaggta	gctcagggtagacggcaag

RAG1	17	ggaccctgactatgggactct	cgtgcttggtgcagtaagtg
RAMP1A	17	gggaccctgactatgggact	agtgcgcatgcagatcct
RAMP1B	33	cgatctaagggatatggatttattaca	aacgtgaccaactttcattgg
RBM39	55	gagaagctgaattctccaacca	cgagctctttagtgtcacctaattt
RGC32	20	cagactgcctgacatctgct	gtgcccacgctaattctcag
RHOB	64	cccttccctgaagtgattga	agttccttgggctctgagg
SDC4	29	tggtctggctctttccactt	tcattagctgggggtctgag
SLAMF6	72	agtcgagggaccgttggt	ccacattgctggacattagaga
SLC5A9	89	cactgttcaatcatggaccaaa	agggctgccaaaaatatcct
SLCO1B2	26	ctcctgctgtcactcctgct	tcccgcaagagtattctgtca
TMEM66	62	tgcggtgcatcaagtatctg	ataacggctgatccggcta
TSPAN2	53	gggaatccagggagataaaga	aagaagcgcctacgactaatgt
TXNRD3	12	cagaccgacggcaagttc	ccatagaccagggacagtgc
ZAP70	69	ggagcggtagcacctcct	ctggttcatcatcgctaatcac

Supplementary Table S8. UPL probe number and primers (RQ-PCR validation assays). Universal Probe Master system (Roche Diagnostics; F. Hoffmann-La Roche Ltd.)

SUPPLEMENTARY BIOINFORMATICS

PAX5/ETV6 affects fundamental cellular processes.

The Ingenuity Pathway Analysis (IPA) software identified 25 biological networks (Supplementary Figure S9), associated to specific bio-functions, such as: cellular development, cellular growth and proliferation, cell morphology, cellular assembly and organization, hematological system development and function, and gene expression. Supplementary Figures S10 and S11 are examples of the graphical views of the top 2 networks, in which relationships between different molecules are shown. The most significant Network 1 (Supplementary Figure S10) is related to antigen presentation, and it is centered on the TNF pathway, including some of the molecules validated in previous experiments, such as TNFα itself, LTA and LTB. The second network in order of significance is centered on B-cell receptor signaling (Supplementary Figure S11), including cell surface as well as intracellular signaling molecules, many of them corresponding to PAX5-target genes, such as CD19 and CD22.

Single networks were linked to each other on the basis of recurrence of molecules and biological bio-functions (Supplementary Figure S12); the main correlation involved networks (such as 1, 2, 3, 7 and 9) whose common biological functions are cell assembly and organization, cell to cell signaling, and cellular development. All the top bio-functions are summarized in Supplementary Figure S13. Among the most relevant functions, basic hematopoietic cell processes were identified, such as cell movement and adhesion, cell development and signaling the list of genes, function and statistical analyses are reported in Supplementary Figures S14-S15). Notably several genes, not regulated by PAX5 but involved in the adhesion process, were found.

Similarly to IPA, The Database for Annotation, Visualization and Integrated Discovery (DAVID) allowed us to identify groups of molecules with similar functions with two different approaches: gene functional classification and functional annotation analysis. Here we reported the highest stringency classification that defined 5 gene enrichment clusters (reported in Supplementary Figure S16), which are mainly related to the following functions: transcription and protein synthesis (gene groups 1 and 3), cell signaling and proliferation (gene group 2), and B-cell receptor and adhesion (gene

groups 4 and 5). The graphical view of this analysis is shown in Supplementary Figure S17.

IPA analysis
340 DEGs list
version 8.0
content version 2602

Analysis settings

View
Reference set: Mouse Genome 430A 2.0 Array
Relationship to include: Direct and Indirect
Includes Endogenous Chemicals
Optional Analyses: My Pathways My List

Filter Summary:
Consider all molecules and/or relationships

Top Networks

ID	Associated Network Functions	Score
1	Antigen Presentation, Cellular Development, Cellular Function and Maintenance	47
2	Cellular Growth and Proliferation, Hematological System Development and Function, Cell Morphology	40
3	Cellular Assembly and Organization, Cellular Function and Maintenance, Hematological System Development and Function	33
4	Cell Cycle, Hair and Skin Development and Function, Protein Synthesis	27
5	Connective Tissue Disorders, Hematological Disease, Organismal Injury and Abnormalities	26

Top Bio Functions

Diseases and Disorders

Name	p-value	# Molecules
Inflammatory Response	6,58E-06 - 1,55E-02	47
Immunological Disease	4,75E-05 - 1,71E-02	20
Infectious Disease	4,75E-05 - 1,09E-02	45
Neurological Disease	4,75E-05 - 1,71E-02	21
Infection Mechanism	7,30E-05 - 5,04E-03	7

Molecular and Cellular Functions

Name	p-value	# Molecules
Cell Morphology	3,64E-08 - 1,35E-02	31
Cellular Development	3,64E-08 - 1,71E-02	66
Cellular Growth and Proliferation	1,65E-07 - 1,44E-02	89
Cellular Function and Maintenance	3,98E-06 - 1,57E-02	48
Cell-To-Cell Signaling and Interaction	4,34E-06 - 1,71E-02	54

Physiological System Development and Function

Name	p-value	# Molecules
Hematological System Development and Function	1,65E-07 - 1,71E-02	65
Cell-mediated Immune Response	3,98E-06 - 1,57E-02	26
Hematopoiesis	3,98E-06 - 1,57E-02	33
Immune Cell Trafficking	3,98E-06 - 1,71E-02	34
Lymphoid Tissue Structure and Development	5,74E-06 - 1,02E-02	14

Top Canonical Pathways

Name	p-value	Ratio
Reelin Signaling in Neurons	2,95E-03	6/77 (0,078)
Crosstalk between Dendritic Cells and Natural Killer Cells	4,35E-03	6/98 (0,061)
Primary Immunodeficiency Signaling	1E-02	4/53 (0,075)
Phospholipase C Signaling	1,35E-02	10/253 (0,04)
Role of NFAT in Regulation of the Immune Response	1,83E-02	8/195 (0,041)

Top Molecules

Log Ratio up-regulated

Molecules	Exp. Value	Exp. Chart
ITGB4	↑3,271	
CNN1	↑2,577	
RGS10	↑2,171	
FSTL1*	↑1,979	
RCAN2	↑1,870	
RAMP1	↑1,846	
C4ORF19	↑1,572	
GNG4	↑1,571	
TNNT2*	↑1,486	
DHRS3	↑1,458	

Log Ratio down-regulated

Molecules	Exp. Value	Exp. Chart
SLCO1B3*	↓-4,077	
LTA	↓-3,482	
COL5A3	↓-3,294	
IFI44L	↓-2,876	
MAP7	↓-1,913	
HHEX	↓-1,906	
CD244	↓-1,832	
IFI30	↓-1,684	
BCL2L15 (includes EG:229672)	↓-1,634	
BCAR3	↓-1,551	

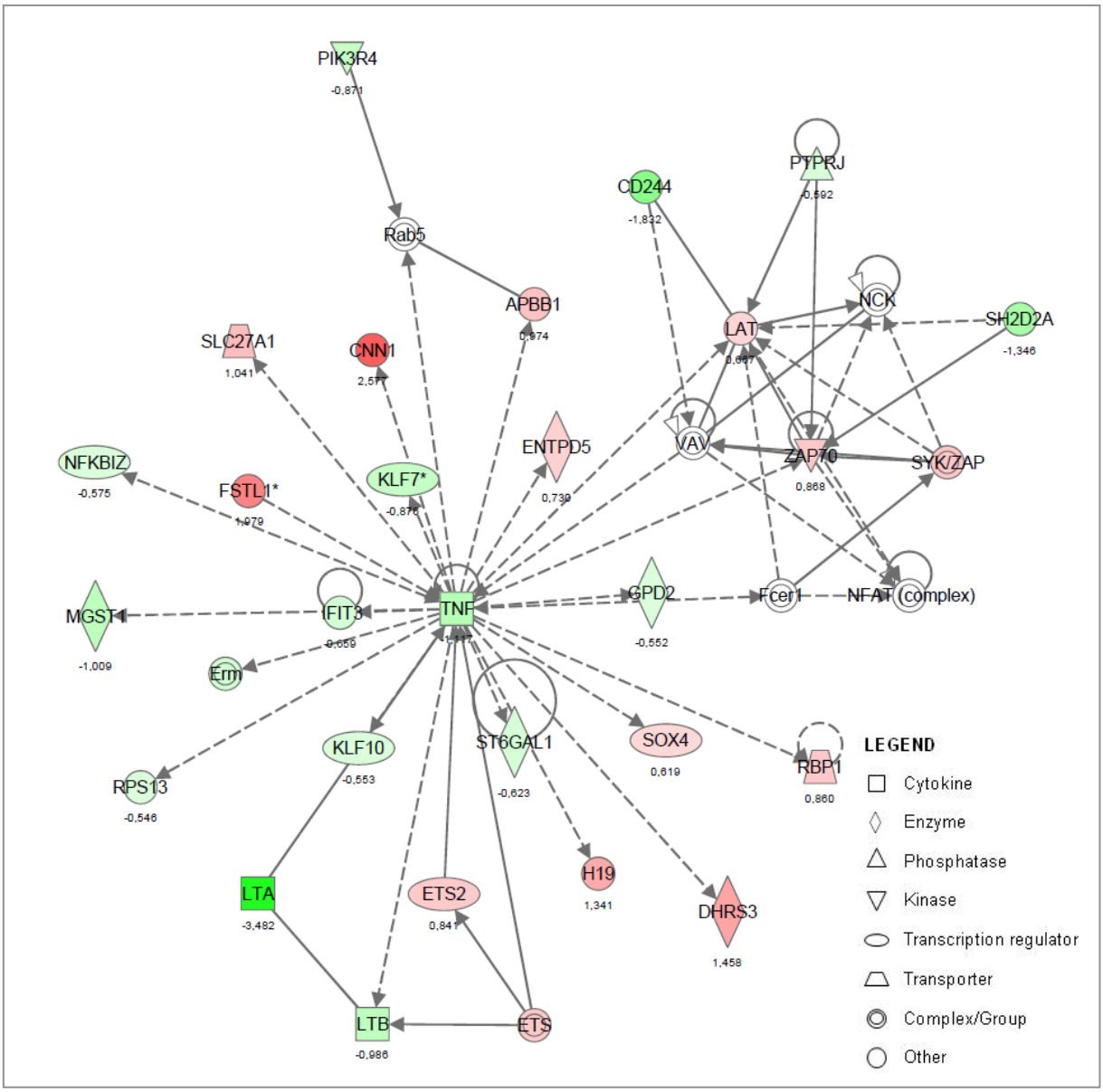
FIGURE S9

Ingenuity Pathway Analysis 8.0

ID	Molecules in Network	Score	Focus Molecules	Top Functions
1	↑**AP8B1,**↓**CD244,**↑**CNN1,**↑**DHRS3,**↑**ENTPD5,**Erm, ETS, ↑**ETS2,** Fcer1, ↑**FSTL1*,**↓**GPD2,**↑**H19,**↓**IFIT3,** ↓**KLF7*,**↓**KLF10,**↑**LAT,**↓**LTA,**↓**LTB,**↓**MGST1,** NCK, NFAT (complex),↓**NFKBIZ,**↓**PIK3R4,**↓**PTPRJ,** Rab5, ↑**RBP1,**↓**RPS13,**↓**SH2D2A,**↑**SLC27A1,**↑**SOX4,**↓**ST6GAL1,** SYK/ZAP, ↓**TNF,** VAV, ↑**ZAP70**	47	27	Antigen Presentation, Cellular Development, Cellular Function and Maintenance
2	↑**ALDH1B1,**↓**B3GNT5,**BCR, CD8, ↓**CD19,**↓**CD22,**↓**CD37,**↓**CD84,**↓**CLOCK,**↓**CORO1A,**↓**DOK2,**↓**GADD45B*,** ↑**ID2,** Igm, ↓**KCNN4,**↓**KLF2,**↑**LCK,** LRP, ↓**LYN,**↓**MGEA5,** MHC CLASS I (family),↓**MTSS1,** NFkB (complex), Notch, ↑**NOTCH3,**↓**PRKD2,** Proteasome, Sapk, Secretase gamma, ↓**SLAMF1,**↓**SP4,**↓**THBD,**↓**TRAIP,**↓**VOPP1*,**↓**WWP1**	40	25	Cellular Growth and Proliferation, Hematological System Development and Function, Cell Morphology
3	↓**BTF3,** Caspase, Ck2, ↓**CRKRS,**↓**CSTF2 (includes EG:1478),**Cyclin E, ↓**FXYD5,**↓**HELB,**↑**HERPUD1,** Histone h3, Histone h4, ↑**HLA-B,** Hsp70, Hsp90, ↓**IKBKAP,**↓**MED14,**↓**MYST2,**↓**NEK6,**P38 MAPK, ↓**PAFAH1B1,** PDGF BB, Pka, Pka catalytic subunit, PLC gamma,↑**RAD21,** Rb, ↓**RBM39*,**↑**RGS10,** RNA polymerase II, ↓**RPS10,**↓**SDHAF2,** ↓**SFPQ,**↓**SMC3,**↑**TSC22D3,**↓**YBX1**	33	21	Cellular Assembly and Organization, Cellular Function and Maintenance, Hematological System Development and Function
4	↓**ACSL4,** CDKN2A, DDX11, DLEU1, E2F3, ↑**E2F5,**↓**ETV3,** GRB2, IPO4, ↓**MGA,** MIR103-1 (includes EG:406895),MYC, MYL12A, ↓**NISCH,**↓**RPL3,** RPL7, ↓**RPL11,** RPL13, ↓**RPL18,** RPL19, RPL27, ↓**RPL30,**↓**RPL9 (includes EG:20005),** ↓**RPS13,** RPS16, RPS20, RPS26, ↓**SAFB2,**↓**SCPEP1,**↓**SEC61A1*,**↓**SEH1L*,**↓**SHMT1,**↓**SLC23A1,** TFDP3, ↓**ZFP106**	27	18	Cell Cycle, Hair and Skin Development and Function, Protein Synthesis
5	Actin, Alpha Actinin, Alpha catenin, ↓**ATXN7,**Calmodulin, Camk, CaMK I, ↓**CAMK1D,**↑**CAMK1G,**↓**CAMK2D,**↓**CD44,** ↓**COL5A3,** Collagen type I, Collagen(s), ↓**DHX9,** ERK, F Actin, Focal adhesion kinase,↑**GSN*,** Ige, Integrin alpha 2 beta 1, Laminin,↓**LPP,**↓**LRPAP1*,**↑**MARCKS (includes EG:4082),**↓**MSN,**↓**OVGP1,** Pak, Rock, ↓**SCIN,**↓**SPRED1*,**↓**SPRED2,**↓**SPTBN1,**↓**TES*,** Tropomyosin	26	18	Connective Tissue Disorders, Hematological Disease, Organismal Injury and Abnormalities
6	14-3-3, Akt, Calpain, CD3, ↓**CD47,** Fibrinogen, ↑**GAB1,**↓**GRB10,** Integrin, Integrina, ↑**ITGA6,**↓**ITGB1,**↑**ITGB4,** ↓**MAGI1,** MAP2K1/2, Mek, ↓**MYCN,**PP2A, Rap1, Ras, ↓**RBM4,**↓**RPL3,**↓**RPL11,**↓**RPL18,**↓**RPL30,**↓**RPL27A,** Shc, Sos, SRC, TCR, ↓**TM6SF1,** TSPAN, ↑**TSPAN2*,**↓**TSPAN6,** Vla-4	25	17	Cell-To-Cell Signaling and Interaction, Hair and Skin Development and Function, Cell Morphology
7	8-oxo-7-hydrodeoxyguanosine, ABCC1, ↓**CD244,**CD33 (includes EG:945),↓**CIRBP,**CTSC, DECR1, FCGR1B, ↓**GFM1,** GH1, glutathione, ↓**HEATR6,**↓**IFI44,**↓**IFI44L,**↓**IFIT3,**IFNG, IGHG1,↓**IGL-V1,**IL25, ↓**MGST1,** PAPPA, PDCD1, PI3, ↓**PLEK,**PTPN11, ↓**RAP2B,**↑**RTKN2,**SH2D1A, SH2D1B, SIGLEC9, ↓**SLAMF1,**↓**SLAMF6,**↓**SLCO1B3*,** ↓**TXNRD3*,**ursodeoxycholic acid	23	16	Lipid Metabolism, Molecular Transport, Small Molecule Biochemistry
8	↑**ACP5,**↓**ACVR2A,**↓**ADAM10,** Ap1, ↓**ARL4C,**↓**BCAR3,** CaMKII, ↓**CRK,**↓**DTL,**↓**EIF2S2,**↓**EIF4EBP2,**FSH, G alpha, G-protein beta, ↓**GNA13,**↑**GRK5,**hCG, Insulin, Jnk, ↓**LBH,** Lh, Mapk, Nfat (family), PI3K, Pkc(s), Pld, ↓**PLD4,** Rac, ↑**RAG1,** Ras homolog,↓**RHOB,**↓**SDC4*,**↑**THY1,**↑**TNNT2*,**Vegf	23	18	Cellular Assembly and Organization, Cell-To-Cell Signaling and Interaction, Neurological Disease
9	AMBP, ↓**C19ORF44,**↑**C4ORF19,**↓**C8ORF41,**↓**CHCHD5,**↓**CYB561D2,**EMG1, ↑**FAM125B,** GALM, HNF4A, IMMT, LRSAM1, ↓**MAP7,**↓**MYST2,** NEDD4L, ONECUT1, ORC1L, POLB, PPFIBP2, PSMB7, RFC3, ROCK1, ↓**SFRS8*,** ↑**SLC25A38,** SRP19, ↓**SRP54,**↓**TAF1D,**↓**TRAIP,** TRIM23, TSG101, ↑**UEVLD,**VPS28, VPS37B, ↓**YBX1,** ZSCAN16	21	15	Cellular Assembly and Organization, Digestive System Development and Function, Hepatic System Development and Function
10	ALDOA, ATP5O, BLMH, dihydrotestosterone, DLG4, ↓**DNAH8,**↑**EFHD1,**FOS, ↓**GGA2,**↑**GLCCI1,**↓**GPD2,**↓**HHEX,** HLA-A, ↓**IQSEC1,**↓**LASS4*,** MIR133A-1, NR3C1, NUCB1, ↓**OLA1,** palmitic acid, PHGDH, ↓**RPL3,** RPL8, RPL19, RPL27, ↓**RPL30,** RPLP0 (includes EG:6175),RPS23, ↓**RPS17 (includes EG:20068)*,**↓**RPS24 (includes EG:20088),** ↓**SEMA4B,** SGK1, SLC2A4, TUBA1B, ↑**TUBB2B**	21	15	Infection Mechanism, Gene Expression, Lipid Metabolism
11	↓**ABCF1,** BMP1, CIDEC, CLDN4, COL5A1, ↓**CORO1C,**CPT2, ERBB2, ESPL1, ↓**FAM134B,**↓**FRRS1,**↑**FSTL1*,** ↓**GOSR2,**↑**HM13,** IGFBP6, LIF, ↓**LTA,** LTBP2, MXD3, ↑**NKD2,**NNMT, PPARG, RASA3, ↓**RPS-1022P6.2,**↑**SELM,** ↓**SLC39A8,** SRP19, TAX1BP3, TGFA, TGFB1, ↓**THSD1,** UCK1, XBP1, ↑**ZFP54,**↓**ZWINT (includes EG:11130**	21	15	Cellular Development, Respiratory System Development and Function, Embryonic Development
12	ARPC1B, BMP1, BTG3, ↓**DOCK9*,**↓**ETHE1,**↑**FAM129B,**↓**GLIPR1,** HDC, HTT, KIAA1377, ↓**LATS2,**↓**LTB,** MCAM, MIR124, PERP, PMF1, POLB, ↓**POLH,**↓**POLI,** POLM, PPP1R13L, ↑**RDH10,** REV1, RFC3, ↓**RNF20,**RNF40, ↓**ROGDI,** SAT1, SP1, ↑**STAU2,**↓**TAF7,**↑**TMEM66,**↓**TOPORS,** TP53, UCHL1	20	15	Cell Morphology, DNA Replication, Recombination, and Repair, Molecular Transport
13	↓**ABCB4,** AMPD3, APP, beta-carotene, ↑**C13ORF15,**CD5L, ↓**CLCN5,**CLEC7A, Cofilin, CRCP, ↓**CRYL1,**CTSF, ENPP2, GATA3, ↑**GFI1B,**↑**H19,**↑**HNRPLL*,**IL2, IL1B, IRG1, ↓**LGR5,**↓**LTA,** phosphatidylinositol 4,5-diphosphate, PLA2G7, progesterone, ↑**RAMP1,**↓**SLC2A9,** SNCA, ↑**SYTL2,** TENC1, TNFRSF21, TYMP, ↓**UBOX5,**VCP, ↓**ZC3H8**	18	14	Carbohydrate Metabolism, Lipid Metabolism, Molecular Transport
14	↑**ADAMTS10 (includes EG:81794),**AKAP12, ↓**ATP6V0A2,** BDKRB1, CDH13, ↓**CSF2RB,**↓**CTNNB1,** CTSZ (includes EG:1522),↑**CYTSB,**EPAS1, EPO, FZD9, IL10, JUN, L-glutamic acid, MAFF, ↓**MGLL,** MIR24-1 (includes EG:407012),↑**MMP14,**↓**MYOM2,** NFYB, NOSIP, ↓**OTUD4,** peptidase, ↑**PIGC,** RBP3, SIM2, SLC1A3, TAX1BP3, ↓**TCF20,** Tcf/lef, Timp,↓**TOR3A,**↓**TTC39B,**↓**ZNF45**	18	14	Cellular Development, Cellular Growth and Proliferation, Cancer
15	↑**ABCA1,** ALP, Cbp/p300, ↓**CD74,**CHEMOKINE, Creb, ↓**CTLA2A,** ERK1/2, ↓**FOSB,** HDL, ↓**HLA-DQB1,** HLA-DR, ↓**IFI30,** Ifn, IFN Beta, ↑**IFNZ,**IgG, IL1, IL12 (complex), Il12 (family), Interferon alpha, LDL, ↓**LDLR,** MHC Class II, Mmp, NADPH oxidase, Pdgf, ↓**PPP1R15A (includes EG:23645),**↑**PTGER2,**↑**RGS18,**↓**RNASEL,**↓**SLC16A6,** STAT5a/b, Tgf beta, Tlr	17	13	Inflammatory Response, Cardiovascular Disease, Genetic Disorder
16	ALOX5AP, ↑**BANP,**↓**BCL2L15 (includes EG:229672)**BMP1, CASP3, CASP12 (includes EG:12364), CCL6, CD274, ↓**CTLA2B,**↓**CXCR5,**↓**FAM120B,** GNB5, ↑**GNG4,**↓**GPR155,** HSPB2, IL4, KITLG, LTC4S, MADD, ↑**NCAM2,** NRP1, ↑**PLXND1,**PRNP, RAF1, Sema3, SEMA3E, ↓**SLC44A2,** STK3, ↑**TERF1,** TIFA, TNFRSF4, TNFSF9, TRAF2, ↓**USP53,** ZFP36	15	12	Lipid Metabolism, Small Molecule Biochemistry, Molecular Transport
17	ACTB, ↑**ADSSL1,**↑**ARHGAP9*,** ARRB1, CBY1, ↑**CPNE2,**CTPS, ↓**DGKE,**EIF3C, ↓**FARSB,** L-triiodothyronine, LIMA1, LMO7, ↓**MAN2B1,**MYCBP2, NCKIPSD, PFKP, PLEC1, ↑**RCAN2,**↓**RPL3,** RPL6, RPL22, RPLP0 (includes EG:6175),RPS7, RPS8, SFRS4, SMAD9, ↓**SRRM1,** SRRM2, TMOD3, TPI1, TRIP12, UCP3, YWHAG, YWHAZ	10	9	Carbohydrate Metabolism, Amino Acid Metabolism, Cellular Assembly and Organization
18	TGFB3, ↓**TSPAN13**	2	1	Embryonic Development, Tissue Morphology, Cellular Development
19	MIR188 (includes EG:406964),↑**TMEM121**	2	1	
20	MIR96 (includes EG:407053),↓**ZNF280C**	2	1	
21	↑**C7ORF70,**MIRN352	2	1	
22	↓**MDP1,**SYT17	2	1	Cancer, Reproductive System Disease
23	↑**BVES,**GEFT	2	1	Cell Death, Cellular Development, Nervous System Development and Function
24	↓**ATPGD1,**TRIP12	2	1	
25	↓**AGXT2L2,** TRANSAMINASE	2	1	

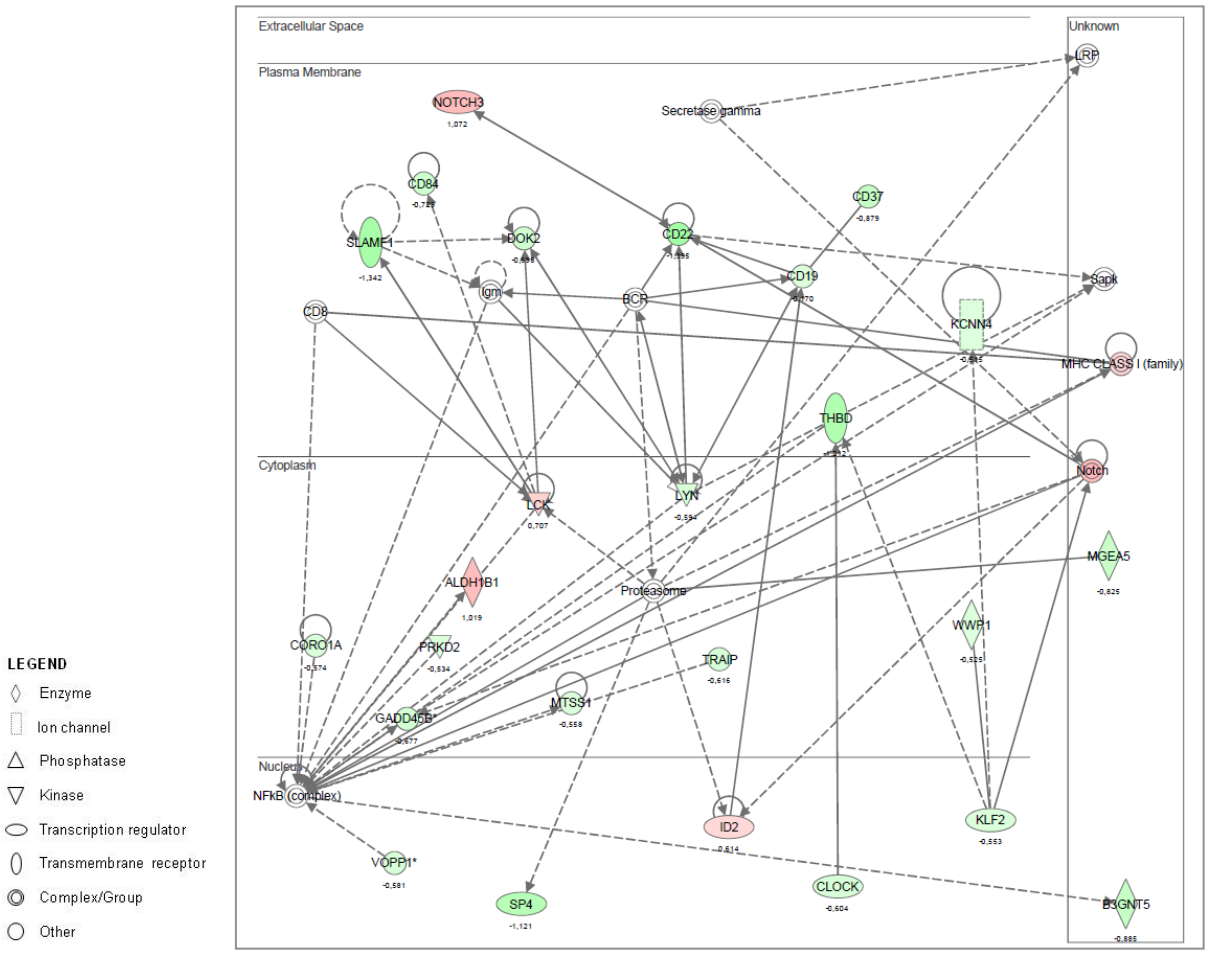
Supplementary Figure S9. IPA network list.

FIGURE S10



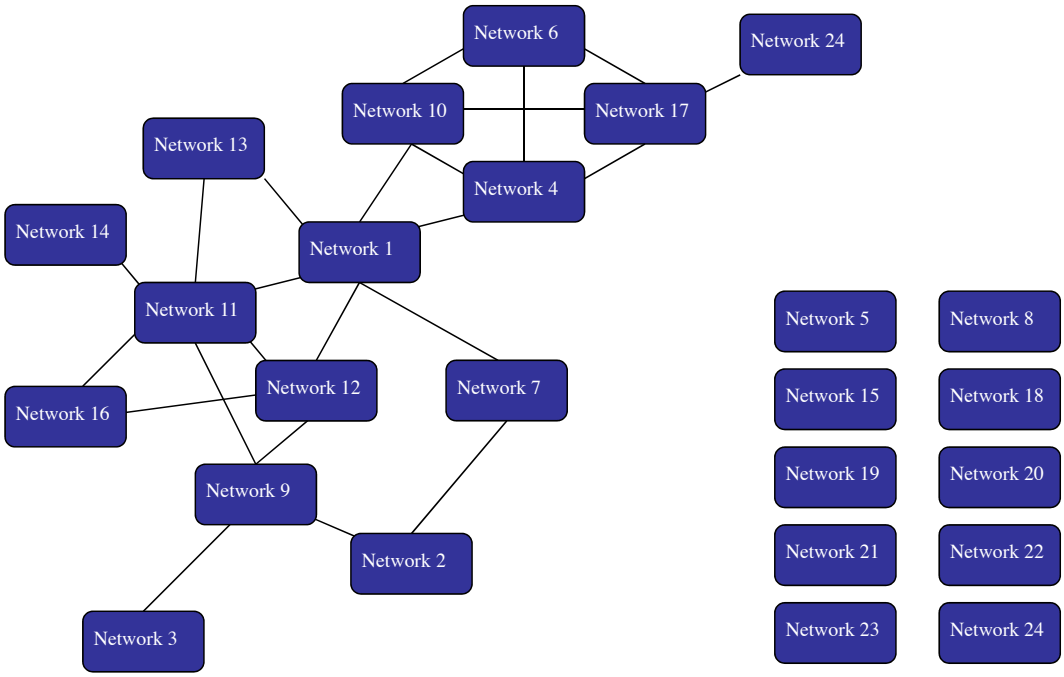
Supplementary Figure S10. Graphycal view of IPA network n°1. For each gene, the correspondent GEP value of log2(signal) is indicated. Green molecule, down-regulated gene; red molecule, up-regulated gene. The intensity of the color is directly related to the GEP value.

FIGURE S11



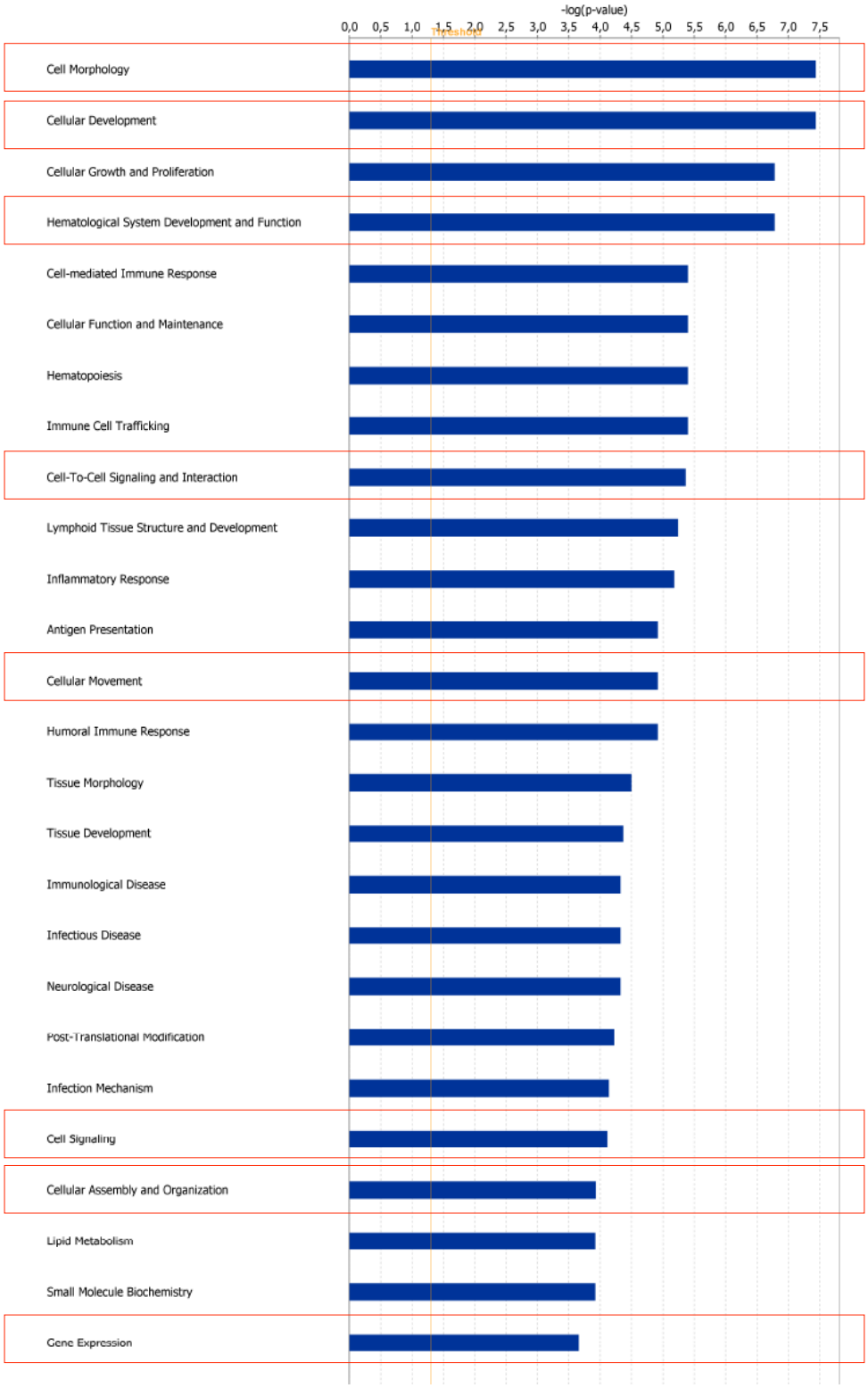
Supplementary Figure S11. Graphycal view of IPA network n°2. For each gene, the correspondent GEP value of log2(signal) is indicated. Green molecule, down-regulated gene; red molecule, up-regulated gene. The intensity of the color is directly related to the GEP value.

FIGURE S12



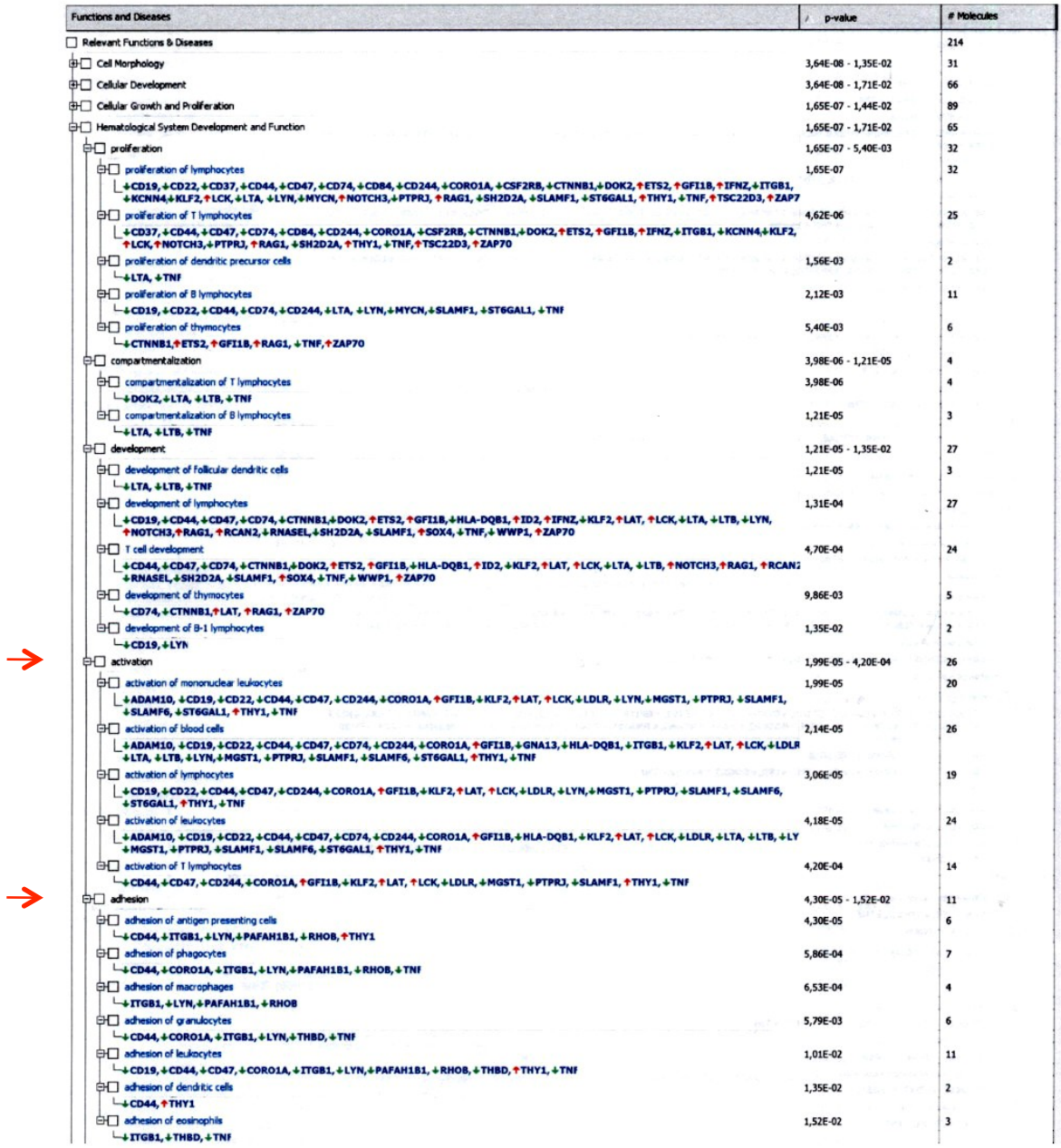
Supplementary Figure S12. IPA networks correlation.

FIGURE S13



Supplementary Figure S13. IPA top Bio-functions. IPA organized the top bio-functions in a histogram view, giving a priority score to each function, on the basis of the number of molecules present among DEGs list, which belong to the specific function.

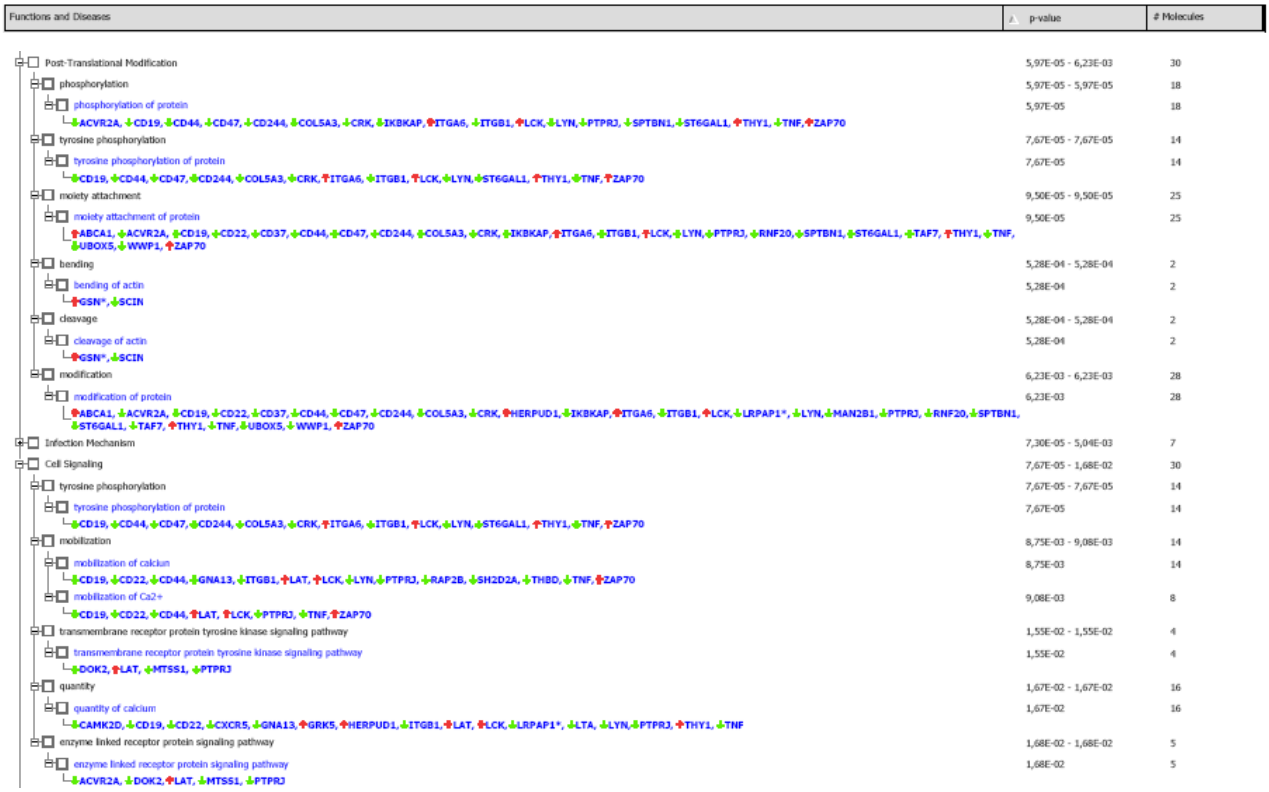
FIGURE S14



Functions and Diseases	p-value	# Molecules
Relevant Functions & Diseases		214
Cell Morphology	3,64E-08 - 1,35E-02	31
Cellular Development	3,64E-08 - 1,71E-02	66
Cellular Growth and Proliferation	1,65E-07 - 1,44E-02	89
Hematological System Development and Function	1,65E-07 - 1,71E-02	65
proliferation	1,65E-07 - 5,40E-03	32
proliferation of lymphocytes	1,65E-07	32
↓CD19,↓CD22,↓CD37,↓CD44,↓CD47,↓CD74,↓CD84,↓CD244,↓CORO1A,↓CSF2RB,↓CTNNB1,↓DOK2,↑ETS2,↑GFI1B,↑IFNZ,↓ITGB1, ↓KCNN4,↓KLF2,↑LCK,↓LTA,↓LYN,↓MYCN,↑NOTCH3,↓PTPRJ,↑RAG1,↓SH2D2A,↓SLAMF1,↓ST6GAL1,↑THY1,↓TNF,↑TSC22D3,↑ZAP7		
proliferation of T lymphocytes	4,62E-06	25
↓CD37,↓CD44,↓CD47,↓CD74,↓CD84,↓CD244,↓CORO1A,↓CSF2RB,↓CTNNB1,↓DOK2,↑ETS2,↑GFI1B,↑IFNZ,↓ITGB1,↓KCNN4,↓KLF2, ↑LCK,↑NOTCH3,↓PTPRJ,↑RAG1,↓SH2D2A,↑THY1,↓TNF,↑TSC22D3,↑ZAP70		
proliferation of dendritic precursor cells	1,56E-03	2
↓LTA,↓TNF		
proliferation of B lymphocytes	2,12E-03	11
↓CD19,↓CD22,↓CD44,↓CD74,↓CD244,↓LTA,↓LYN,↓MYCN,↓SLAMF1,↓ST6GAL1,↓TNF		
proliferation of thymocytes	5,40E-03	6
↓CTNNB1,↑ETS2,↑GFI1B,↑RAG1,↓TNF,↑ZAP70		
compartmentalization	3,98E-06 - 1,21E-05	4
compartmentalization of T lymphocytes	3,98E-06	4
↓DOK2,↓LTA,↓LTB,↓TNF		
compartmentalization of B lymphocytes	1,21E-05	3
↓LTA,↓LTB,↓TNF		
development	1,21E-05 - 1,35E-02	27
development of follicular dendritic cells	1,21E-05	3
↓LTA,↓LTB,↓TNF		
development of lymphocytes	1,31E-04	27
↓CD19,↓CD44,↓CD47,↓CD74,↓CTNNB1,↓DOK2,↑ETS2,↑GFI1B,↓HLA-DQB1,↑ID2,↑IFNZ,↓KLF2,↑LAT,↑LCK,↓LTA,↓LTB,↓LYN, ↑NOTCH3,↑RAG1,↑RCAN2,↓RNASEL,↓SH2D2A,↓SLAMF1,↑SOX4,↓TNF,↓WWP1,↑ZAP70		
T cell development	4,70E-04	24
↓CD44,↓CD47,↓CD74,↓CTNNB1,↓DOK2,↑ETS2,↑GFI1B,↓HLA-DQB1,↑ID2,↓KLF2,↑LAT,↑LCK,↓LTA,↓LTB,↑NOTCH3,↑RAG1,↑RCAN2 ↓RNASEL,↓SH2D2A,↓SLAMF1,↑SOX4,↓TNF,↓WWP1,↑ZAP70		
development of thymocytes	9,86E-03	5
↓CD74,↓CTNNB1,↑LAT,↑RAG1,↑ZAP70		
development of B-1 lymphocytes	1,35E-02	2
↓CD19,↓LYN		
activation	1,99E-05 - 4,20E-04	26
activation of mononuclear leukocytes	1,99E-05	20
↓ADAM10,↓CD19,↓CD22,↓CD44,↓CD47,↓CD244,↓CORO1A,↑GFI1B,↓KLF2,↑LAT,↑LCK,↓LDLR,↓LYN,↓MGST1,↓PTPRJ,↓SLAMF1, ↓SLAMF6,↓ST6GAL1,↑THY1,↓TNF		
activation of blood cells	2,14E-05	26
↓ADAM10,↓CD19,↓CD22,↓CD44,↓CD47,↓CD74,↓CD244,↓CORO1A,↑GFI1B,↓GNA13,↓HLA-DQB1,↓ITGB1,↓KLF2,↑LAT,↑LCK,↓LDLR ↓LTA,↓LTB,↓LYN,↓MGST1,↓PTPRJ,↓SLAMF1,↓SLAMF6,↓ST6GAL1,↑THY1,↓TNF		
activation of lymphocytes	3,06E-05	19
↓CD19,↓CD22,↓CD44,↓CD47,↓CD244,↓CORO1A,↑GFI1B,↓KLF2,↑LAT,↑LCK,↓LDLR,↓LYN,↓MGST1,↓PTPRJ,↓SLAMF1,↓SLAMF6, ↓ST6GAL1,↑THY1,↓TNF		
activation of leukocytes	4,18E-05	24
↓ADAM10,↓CD19,↓CD22,↓CD44,↓CD47,↓CD74,↓CD244,↓CORO1A,↑GFI1B,↓HLA-DQB1,↓KLF2,↑LAT,↑LCK,↓LDLR,↓LTA,↓LTB,↓LY ↓MGST1,↓PTPRJ,↓SLAMF1,↓SLAMF6,↓ST6GAL1,↑THY1,↓TNF		
activation of T lymphocytes	4,20E-04	14
↓CD44,↓CD47,↓CD244,↓CORO1A,↑GFI1B,↓KLF2,↑LAT,↑LCK,↓LDLR,↓MGST1,↓PTPRJ,↓SLAMF1,↑THY1,↓TNF		
adhesion	4,30E-05 - 1,52E-02	11
adhesion of antigen presenting cells	4,30E-05	6
↓CD44,↓ITGB1,↓LYN,↓PAFAH1B1,↓RHOB,↑THY1		
adhesion of phagocytes	5,86E-04	7
↓CD44,↓CORO1A,↓ITGB1,↓LYN,↓PAFAH1B1,↓RHOB,↓TNF		
adhesion of macrophages	6,53E-04	4
↓ITGB1,↓LYN,↓PAFAH1B1,↓RHOB		
adhesion of granulocytes	5,79E-03	6
↓CD44,↓CORO1A,↓ITGB1,↓LYN,↓THBD,↓TNF		
adhesion of leukocytes	1,01E-02	11
↓CD19,↓CD44,↓CD47,↓CORO1A,↓ITGB1,↓LYN,↓PAFAH1B1,↓RHOB,↓THBD,↑THY1,↓TNF		
adhesion of dendritic cells	1,35E-02	2
↓CD44,↑THY1		
adhesion of eosinophils	1,52E-02	3
↓ITGB1,↓THBD,↓TNF		

Supplementary Figure S14. Details of "Hematological System Development and Function".

FIGURE S15



Functions and Diseases	p-value	# Molecules
Post-Translational Modification	5,97E-05 - 6,23E-03	30
phosphorylation	5,97E-05 - 5,97E-05	18
phosphorylation of protein	5,97E-05	18
ACVR2A, CD19, CD44, CD47, CD244, COL5A3, CRK, IKBKAP, ITGA6, ITGB1, LCK, LYN, PTPRJ, SPTBN1, ST6GAL1, THY1, TNF, ZAP70		
tyrosine phosphorylation	7,67E-05 - 7,67E-05	14
tyrosine phosphorylation of protein	7,67E-05	14
CD19, CD44, CD47, CD244, COL5A3, CRK, ITGA6, ITGB1, LCK, LYN, ST6GAL1, THY1, TNF, ZAP70		
moiety attachment	9,50E-05 - 9,50E-05	25
moiety attachment of protein	9,50E-05	25
ABCA1, ACVR2A, CD19, CD22, CD37, CD44, CD47, CD244, COL5A3, CRK, IKBKAP, ITGA6, ITGB1, LCK, LYN, PTPRJ, RNF20, SPTBN1, ST6GAL1, TAF7, THY1, TNF, UBOX5, WWP1, ZAP70		
bending	5,28E-04 - 5,28E-04	2
bending of actin	5,28E-04	2
GSN*, SCIN		
cleavage	5,28E-04 - 5,28E-04	2
cleavage of actin	5,28E-04	2
GSN*, SCIN		
modification	6,23E-03 - 6,23E-03	28
modification of protein	6,23E-03	28
ABCA1, ACVR2A, CD19, CD22, CD37, CD44, CD47, CD244, COL5A3, CRK, HERPUD1, IKBKAP, ITGA6, ITGB1, LCK, LRPAP1*, LYN, MAN2B1, PTPRJ, RNF20, SPTBN1, ST6GAL1, TAF7, THY1, TNF, UBOX5, WWP1, ZAP70		
Infection Mechanism	7,30E-05 - 5,04E-03	7
Cell Signaling	7,67E-05 - 1,68E-02	30
tyrosine phosphorylation	7,67E-05 - 7,67E-05	14
tyrosine phosphorylation of protein	7,67E-05	14
CD19, CD44, CD47, CD244, COL5A3, CRK, ITGA6, ITGB1, LCK, LYN, ST6GAL1, THY1, TNF, ZAP70		
mobilization	8,75E-03 - 9,08E-03	14
mobilization of calcium	8,75E-03	14
CD19, CD22, CD44, GNA13, ITGB1, LAT, LCK, LYN, PTPRJ, RAP2B, SH2D2A, THBD, TNF, ZAP70		
mobilization of Ca2+	9,08E-03	8
CD19, CD22, CD44, LAT, LCK, PTPRJ, TNF, ZAP70		
transmembrane receptor protein tyrosine kinase signaling pathway	1,55E-02 - 1,55E-02	4
transmembrane receptor protein tyrosine kinase signaling pathway	1,55E-02	4
DOK2, LAT, MTSS1, PTPRJ		
quantity	1,67E-02 - 1,67E-02	16
quantity of calcium	1,67E-02	16
CAMK2D, CD19, CD22, CXCR5, GNA13, GRK5, HERPUD1, ITGB1, LAT, LCK, LRPAP1*, LTA, LYN, PTPRJ, THY1, TNF		
enzyme linked receptor protein signaling pathway	1,68E-02 - 1,68E-02	5
enzyme linked receptor protein signaling pathway	1,68E-02	5
ACVR2A, DOK2, LAT, MTSS1, PTPRJ		

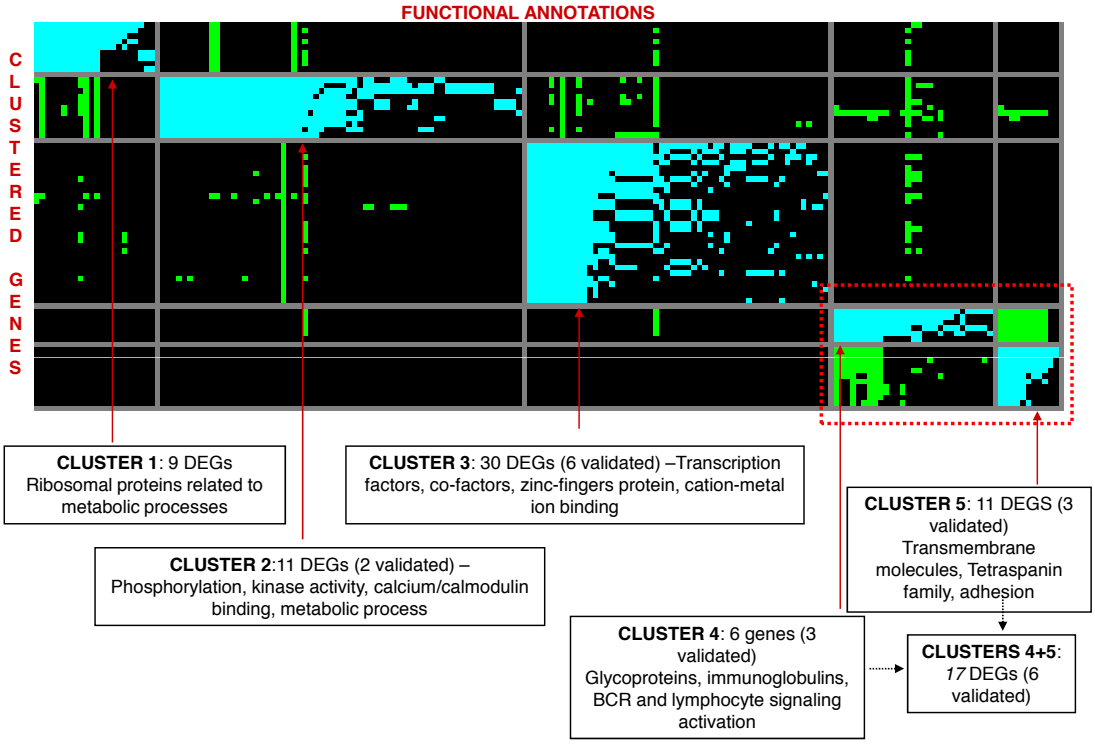
Supplementary Figure S15. Details of "Cell Signaling".

FIGURE S16

Gene group 1	Enrichment score: 3.08		PAX5/ETV6 reg
1437729_at	Rpl27a	ribosomal protein L27a	↓
1438502_x_at, 1438501_at	Rps17	ribosomal protein S17	↓
1456447_at	Rpl18	ribosomal protein L18	↓
1438527_at	Rpl3	ribosomal protein L3	↓
1436798_at	Rpl9	ribosomal protein L9	↓
Gene group 2	**Enrichment score: 1.96**		
1455869_at	Camk2b	NA	↑
1449514_at	Gprk5	G protein-coupled receptor kinase 5	↑
1460352_s_at	Pik3r4	phosphatidylinositol 3 kinase, regulatory subunit, polypeptide 4, p150	↓
1452050_at	Camk1d	calcium/calmodulin-dependent protein kinase ID	↓
1436589_x_at	**Prkd2**	**protein kinase D2**	↓
1460300_a_at	Ltk		↓
1425556_at	Crkrs	CDC2-related kinase, arginine/serine-rich	↓
1423596_at	Nek6	NIMA (never in mitosis gene a)-related expressed kinase 6	↓
1437101_at	Lats2	large tumor suppressor 2	↓
1427763_a_at	Camk2d	calcium/calmodulin-dependent protein kinase II, delta	↓
Gene group 3	**Enrichment score: 1.79**		
1438397_a_at, 1456386_at, 1438398_at	Rbm39	RNA binding motif protein 39	↓
1420399_at, 1427144_at	Gfi1b	growth factor independent 1B	↑
1447631_at	Myst2	MYST histone acetyltransferase 2	↓
1417654_at, 1448793_a_at	**Sdc4**		↓
1421504_at	Sp4	trans-acting transcription factor 4	↓
1451629_at	Lbh	limb-bud and heart	↓
1452361_at	Rnf20	ring finger protein 20	↓
1417444_at	E2f5	E2F transcription factor 5	↓
1421322_a_at	**Isgf3g**		↑
1460231_at	**Irf5**		↓
1422646_at	Mga	MAX gene associated	↓
1417483_at	Nfkbiz	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta	↓
1449439_at, 1419356_at	Klf7	Kruppel-like factor 7 (ubiquitous)	↓
1419521_at	Zfp94	zinc finger protein 94	↓
1424142_at	Ikbkap	inhibitor of kappa light polypeptide enhancer in B-cells, kinase complex-associated protein	↓
1433863_at	Btf3	basic transcription factor 3	↓
1421910_at	Tcf20	transcription factor 20	↓
1422134_at	Fosb	FBJ osteosarcoma oncogene B	↓
1423170_at	Taf7	TAF7 RNA polymerase II, TATA box binding protein (TBP)-associated factor	↓
1448890_at	Klf2	Kruppel-like factor 2 (lung)	↓
1452462_a_at	**Banp**	**BTG3 associated nuclear protein**	↑
1424706_at	Zfp51	zinc finger protein 51	↑
Gene group 4	**Enrichment score: 1.25**		
1419768_at	**Cd22**	**CD22 antigen**	↓
1449991_at	Cd244	CD244 natural killer cell receptor 2B4	↓
1450570_a_at	**Cd19**		↓
1420659_at	**Slamf6**	**SLAM family member 6**	↓
Gene group 5	**Enrichment score: 0.72**		
1424927_at	Glipr1	GLI pathogenesis-related 1 (glioma)	↓
1432417_a_at, 1424567_at	**Tspan2**	**tetraspanin 2**	↑
1451353_at	Tm6sf1	transmembrane 6 superfamily member 1	↓
1425736_at	Cd37	CD37 antigen	↓
1424039_at	**Tmem66**	**transmembrane protein 66**	↑
1460239_at	Tspan13	tetraspanin 13	↓

Supplementary Figure S16. DAVID analysis. From Gene Functional Classification, performed with High Stringency criteria, we defined 5 gene groups.

FIGURE S17



Supplementary Figure S17. Adapted graphical view of DAVID analysis, from Gene Functional Classification performed with High Stringency criteria.

Chapter 4

PAX5 fusion genes drive *LCK* over-expression and enable oncogenic signaling in pre-BI cells

PAX5 fusion genes drive LCK over-expression and enable oncogenic signaling in pre-BI cells

Manuscript

Valeria Cazzaniga,[1] Cristina Bugarin,[1] Marco Giordan,[2] Geertruy te Kronnie,[2] Giuseppe Basso,[2] Andrea Biondi,[1] Grazia Fazio[1] and Giovanni Cazzaniga[1]

[1]Centro Ricerca M. Tettamanti, Clinica pediatrica, Università di Milano-Bicocca, Ospedale San Gerardo/Fondazione MBBM, Monza, Italy; [2]Department of Women's and Children's Health, Univeristy of Padova, Padova.

KEYWORDS: PAX5, ETV6, LCK, STAT5

Corresponding author:
Cazzaniga Giovanni
Centro Ricerca M. Tettamanti
Clinica Pediatrica
Università di Milano-Bicocca
Ospedale S. Gerardo
Via Pergolesi 33
20900 – Monza – (MB) – Italy
gianni.cazzaniga@hsgerardo.org
phone: +390392333661; fax: +390392332167

INTRODUCTION

PAX5 is a transcription factor that belongs to the *PAX* family and is essential for B lymphoid cell commitment.[1] It functions both as a transcriptional activator and a repressor of different target genes involved in lineage development. Furthermore, *PAX5* has been recently reported to be a target of aberrancies, including mutations, deletions and translocations, in around 30% of pediatric patients affected by BCP-ALL,[2] the most frequent leukemia subset in children.[3] Translocations occur in approximately 2-3% of patients,[4] with a variety of partner genes,[2,4,5] and the translocation t(9;12) is the most recurrent. It encodes for the *PAX5/ETV6* fusion gene,[2,5-8] thus juxtaposing two transcription factors with a fundamental role in hematopoiesis and especially in B-cell development.[9]

Mechanistically, the PAX5/ETV6 fusion protein retains the DNA-binding domain of PAX5 while it substitutes its regulatory domains with functually critical ETV6 domains, responsible for DNA-binding, dimerization and transcription regulation.[7]

We previously reported that PAX5/ETV6 is an aberrant transcription factor that localizes in the nucleus.[10,11] It is able to significantly alter the transcription profile of pre-BI cells, mainly disregulating genes involved in pre-BCR assembly and signaling, as evidently shown by the failure of PAX5/ETV6 positive cells to successfully rearrange the μ heavy chain.[11] Moreover, PAX5/ETV6 pre-BI cells showed reduced adhesion to VCAM1, while their migration towards CXCL12, an essential pro-survival chemokine, is enhanced.[10]

While these findings suggest an important role in physiological pre-BI cells, the function of *PAX5* translocations in leukemogenesis is poorly understood. Only very recent studies described the impact of PAX5 on gene expression,[10-12] while functional implications on signaling pathway responsible for cellular proliferation and survival remain to be elucidated.

Indeed, recent studies described the ability of PAX5 fusion proteins in impairing the rearrangement of the u heavy chain and down-regulating many genes involved in the signaling pathway downstream the pre-BCR. Since the pre-BCR signaling pathway is responsible not only for pre-BI cells differentiation but also for their survival. Herein, we employed primary cultures of wild type (wt) pre-BI cells stably transduced with the

retroviral PAX5/ETV6 vector and the correspondent empty control vector (MIGR-GFP), aiming to identify the pro-survival signaling pathway activated by PAX5 fusion proteins.

MATHERIAL AND METHODS

Cell cultures and retroviral transduction. Wt murine pre-BI cell population have been co-cultured on OP9 stroma and transduced as previously described.[10,11]

Reverse transcription-PCR and real-time quantitative-PCR assays have been performed as previously described. Primers and probes have been selected according to the Software Probe Finder (Roche Diagnostics) and are reported in Supplementary Table S1. Data were expressed using the comparative ΔΔCt method, using HPRT gene as reference and MIGR-GFP cells as standardization control; both t test and SD values refer to triplicates of a single experiment and N=3 biological independent experiments were performed for each gene.

Flow cytometry and Phosphoflow analysis. For phosphoflow analysis, cells have been stimulated with 5% IL7 after synchronization by overnight IL7 withdrawal. Afterwards, at different time points the cells have been harvested, fixed in 1.5% PFA and permeabilized with 90% Methanol according to Nolan's Lab protocol (http://www.cytobank.org/nolanlab/experiment_protocols/general_protocol.html). Samples have been stored in 1 ml 90% methanol at -20°C. After recovering they have been stained by specific antibodies (Supplementary Table S2).

Western blot analysis. Pre-BI cells have been lysed in Tris-HCl pH 7.4 20mM, NaCl 20mM, EDTA pH8 2mM, Na3VO4 0.2mM, Triton 1%, NaF 25mM, β-glycerolphosphate 25mM, PMSF 1mM, Protease inhibitor cocktail 1x. StripAblot Stripping Buffer (Euroclone S.p.A.) was used to recover membranes. Densitometry analyses were performed using Alliance instrument and Uviband software (Uvitec Cambridge, UK). Antibodies used and correspondent dilutions are listed in Supplementary Table S3.

BIBF1120 treatment. Pre-BI cells have been treated with either 50nM BIBF1120 or vehicle (DMSO) after synchronization by overnight IL7 withdrawal.

Proliferation assay. Cells have been stained 15 minutes with 1μM CellTrace™ Far Red DDAO-SE (Life Technologies), washed twice with PBS-20%FBS and then analyzed with BD FacsCanto 2 (Becton Dickinson Biosciences).

Cell cycle analysis. Pre-BI cells have been fixed in PFA 1%, permeabilized with ethanol 70% and stained with PI solution (40μg/ml PI and RNase). Samples have been acquired with BD FacsCanto2 and analyzed with FlowJo software.

RESULTS AND DISCUSSION

We previously demonstrated that PAX5/ETV6 is an aberrant transcription factor and it is able to significantly alter the gene expression profile in pre-BI cells. Indeed, gene expression analysis (GEP) revealed a significant number of genes to be differentially expressed in PAX5/ETV6 transduced pre-BI cells compared to the empty vector control (MIGR-GFP). Interestingly, among them, a significant number has been described to be transcriptional direct targets of *Pax5*. One of the top-ranking up-regulated genes was the *Pax5* repressed target gene *Lck* (FC=1.63, $p<0.05$).[11]

While Lck has been implicated to be a fundamental Src kinase in T cells and especially in TCR triggering,[13] very recent studies suggested a potential role in a wide range of hematological malignancies. Indeed, LCK ha been shown to be essential for the proliferation and survival in a subset of T-ALL characterized by NUP214/ABL1 activity (6% of T-ALL cases).[14] Moreover, high expression of *LCK* correlates with poor prognosis in chronic lymphoid leukemia (CLL), where it is involved in BCR signaling and mediates glucocorticoid resistance.[15,16] A similar involvement in response to treatment has been described in MLL-rearranged ALL.[17] In addition, LCK down-modulated activity has been correlated to BCP-ALL prednisone good responders patients.[18] Taken together, these studies suggest an emerging role of LCK in the progression of many subsets of leukemia.

In order to avoid the lack of material from primary patients, we developed a murine model to mimic PAX5 fusion genes and in particular PAX5/ETV6, using three murine pre-BI cell populations, namely LY5.1FL, B6BAFL and FLB6-67, which were retrovirally transduced either with PAX5/ETV6 or the empty vector control (MIGR-GFP).

Quantitative PCR confirmed *Lck* over-expression both at basal level (FC=2.04; 1.55; 1.78; in LY5.1FL, B6BAFL and FLB6-67 respectively; $p<0.001$) (Figure 1A, Supplementary Figure S1A and S2A), and at all time points (0, 24, 48 and 72 hours) after synchronization by overnight withdrawal of IL7. Interestingly, increased *Lck* expression is not altered by overnight withdrawal of IL7, (FC= 7.96 $p<0.001$, FC=15.93 $p<0.01$, FC=3,83 $p<0,001$, FC=4,32 $p<0,001$ respectively) (Figure 1B).

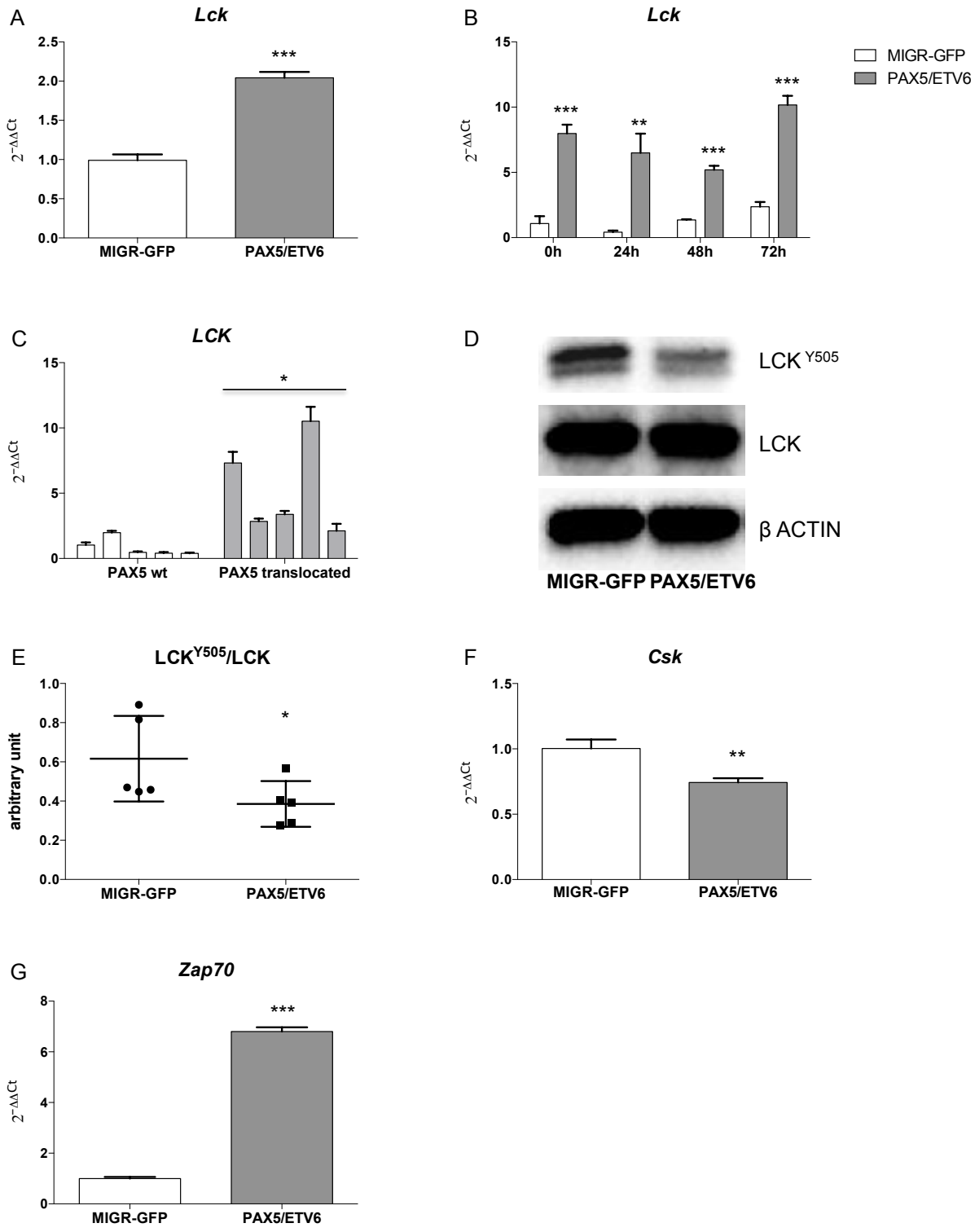


Figure 1. *PAX5* fusion genes disregulate Lck expression. *Lck* is up-regulated in PAX5/ETV6 transduced LY5.1FL pre-BI cells both A) in basal conditions and B) in time course experiments in synchronized cells after overnight withdrawal of IL7. C) PAX5-fusion genes induce *LCK* over-expression in primary patients samples, compared to wt PAX5 BCP-ALL patients. D) LCK protein expression in LY5.1FL pre-BI cells and E) schematic representation of LCK^{Y505}/LCK ratio, summarizing multiple western blot experiments. F) *Csk* and G) *Zap70* mRNA expression levels.

Next, we studied if *Lck* up-regulation was a common feature of *PAX5*-fusion genes or was exclusively found in cases carrying *PAX5/ETV6*. We had the chance to access to the RNA from newly diagnosed childhood BCP-ALL patients enrolled in the AIEOP-IBFM-ALL2000 protocol. We analyze *LCK* mRNA expression levels in patients carrying translocations juxtaposing *PAX5* to several partner genes (our unpublished data, reported in Chapter 2). Quantitative PCR confirmed that *LCK* was up-regulated among *PAX5* translocated patients (n=5) compared to wt *PAX5* BCP-ALL patients (n=5), (mean FC=6.27, $p<0.05$) (Figure 1C). To verify if *LCK* up-regulation was only due to *PAX5* haploinsufficiency, we tested *LCK* transcript levels in PAX5-deleted but not translocated BCP-ALL patients (n=4), where only a not significant up-regulation was found in *PAX5* deleted patients, thus suggesting that *LCK* over-expression is mainly due to the activity of PAX5 fusion proteins and not only to *PAX5* haploinsufficiency (data not shown).

While the increased mRNA levels of *Lck* suggested an important role in PAX5/ETV6 positive pre-BI cells, on the protein level an important function of Lck is mediated through phosphorylation of its residue Tyr505. Despite no striking total Lck overexpression has been observed, PAX5/ETV6 induced dephosphorylation of the Lck inhibitory domain (LCK^{Y505}), causing Lck over-activation, (Figure 1D-E, Supplementary Figure S1B-C, Supplementary Figure S2B-C) as similarly described in Poor Responders BCP-ALL patients.[18]

Mechanistically, Lck phosphorylation in the residue Tyr505 is regulated by the C-terminal SRC kinase Csk, which is responsible for inducing a close conformation to the Lck protein, thus preventing the accessibility to its catalytic domain.[13] Therefore, we wondered whether the dephosphorylation in the inhibitory domain could be due to the down-regulation of Csk itself. Indeed, when we measured *Csk* expression levels we identified a significant down-regulation in all the analyzed populations, demonstrating that the dephosphorylation at the Lck inhibitory domain is caused by the down-regulation of the mRNA of the regulatory protein *Csk* (Figure 1F, Supplementary Figure S1D andS2D).

Thereby, in PAX5/ETV6 cells Lck could be over-activated through two main mechanisms: a) the effect of *PAX5* fusion gene, that up-regulates *Lck* transcription; b) the down-regulation of Csk, thus resulting in the dephosphorylation of its inhibitory domain.

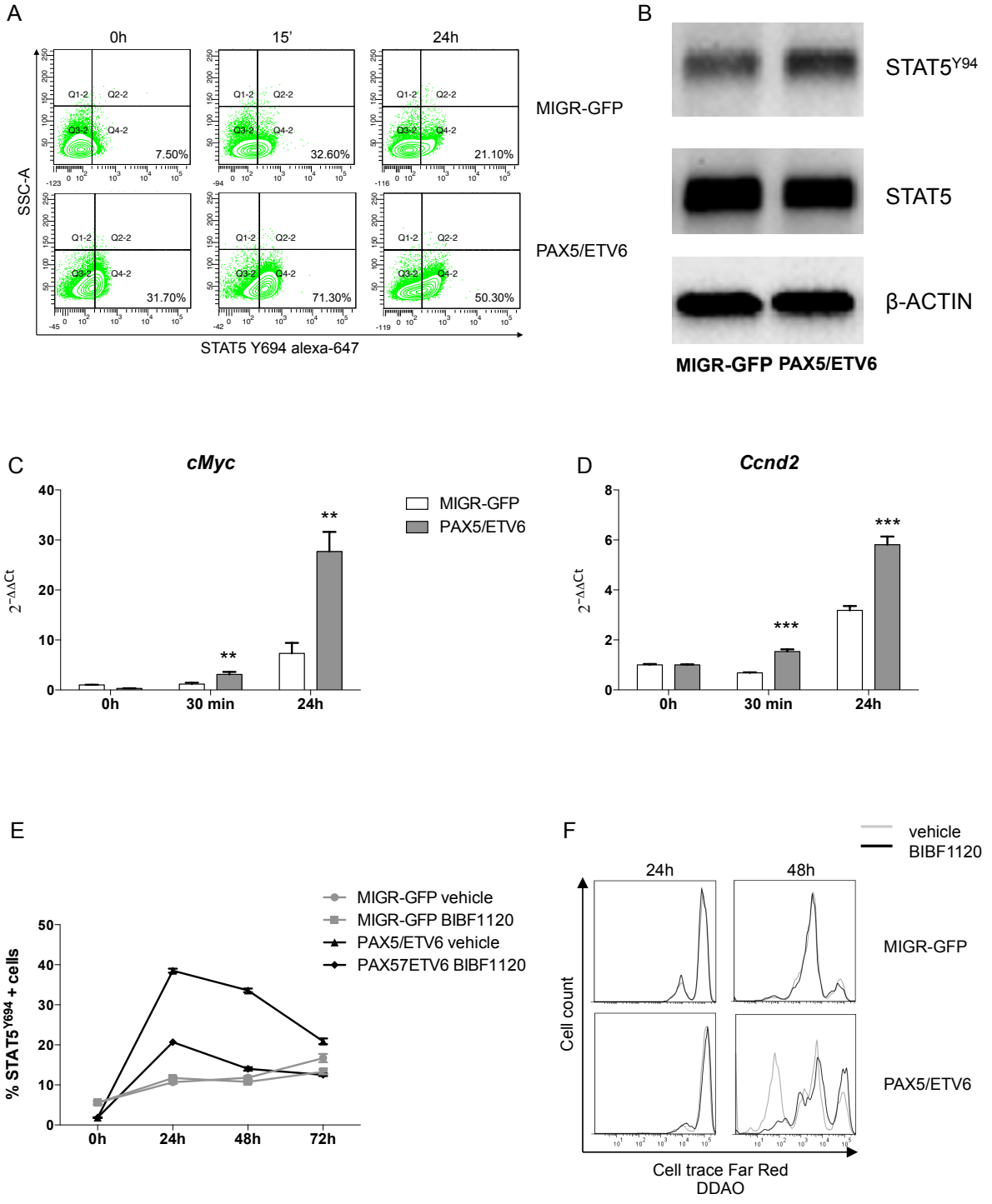


Figure 2. STAT5 hyper-activation is abolished by the administration of the LCK inhibitor BIBF1120 in LY5.1FL pre-BI cells. A) Representative dot plots of STAT5^{Y694} phenotypes. B) Confirmation of STAT5 hyper-phosphorylation by western blot. C) RQ-PCR of *cMyc* and D) *Ccnd2* at early and late time points after IL7 administration. E) Schematic representation of the percentage of STAT5^{Y694} positive cells after administration of either BIBF1120 or vehicle. F) Increased number of replication cycles in PAX5/ETV6 transduced pre-BI cells is abrogated by BIBF1120 administration.

Moreover, *Zap70*, another element of the TCR triggering machinery, has been found up-regulated in PAX5/ETV6 pre-BI cells, describing once again a similar scenario to the one already reported in CLL, where it has been associated to worse prognosis in terms of progression and survival (Figure 1G, Supplementary Figure S1E andS2E).[19]

However, despite the up-regulation of *Lck* and *Zap70* as previously described in CLL, a similar mechanism of coupling to the pre-BCR is not feasible because PAX5/ETV6 cells fail to rearrange the μ heavy chain and then displaying pre-BCR on their cellular surface. Furthermore, the pre-BCR signaling is fundamental in pre-BI cells, not only for their differentiation, but also for their survival and proliferation together with the IL7R pathway,[20] which leads to the activation of STAT5, thus over-activating inducer of cell cycle progression.

Since PAX5/ETV6 provided a survival advantage upon IL7 withdrawal,[10] we wondered if it could be due to an over-activation of the IL7R-STAT5 pathway independently from the ligand.

In accordance to this hypothesis, *Lck* over-expression had been previously described to induce Stat5 hyper-phosphorylation in Ba/F3 pro-B cells.[21] Therefore we investigated if a similar phenomena occured in PAX5/ETV6 transduced cells as well. Phosphoflow analysis revealed increased Stat5 phosphorylation after IL7 stimulation in time course experiments (Figure 2A). Furthermore, Stat5 phosphorylation was observed at basal level, both by phosphoflow analysis and by western blot (Figure 2B and Supplementary Figure S3A). Since the IL7Rα (CD127) was expressed at the same level both in PAX5/ETV6 and MIGR-GFP cells, this suggested that increased Stat5 phosphorylation was independent from classical IL7R signaling (Supplementary Figure S3B and S4B).

Downstream in the pathway, we observed the up-regulation of *cMyc* and *Ccnd2*, at early and late time points after IL7 stimulation (30 minutes and 24 hours, respectively), thus indicating that Stat5 hyper-phosphorylation led to the activation of the transcription of its target genes (Figure 2C-D, Supplementary Figure S3C-D, Supplementary Figure S4C-D).

Since Lck function positively affected cell survival, we decided to probe its role using the Lck inhibitor BIBF1120.[15] Interestingly, treatment with BIBF1120 caused a significant reduction of $Stat5^{Y694}$ phosphorylation in PAX5/ETV6 cells, as suggested by the decreased number of positive cells and the Mean Fluorescence Intensity (MFI), while it

had only minor effect on MIGR-GFP control, where Lck is expressed at lower levels (Figure 2E and Supplementary Figure S5). Moreover, cell cycle analysis revealed that BIBF1120 abrogated the increase in the replicative S phase in PAX5/ETV6 cells, reducing it to the level of MIGR-GFP control cells (Supplementary Figure S6-S7). Furthermore, tracking of the replication cycles showed that PAX5/ETV6 cells have a faster proliferation rate and strikingly its replicative advantage is abrogated by the administration of BIBF1120, which reduces the replicative rate, as show in Figure 2F.

In conclusion, this data suggests that LCK over-expression driven by PAX5/ETV6, and PAX5 translocations in general, can lead to the over-activation of STAT5 signaling pathway, therefore enabling leukemia cells survival (Supplementary Figure S8). Lck inhibition can revert this phenomenon, down-tuning STAT5 signaling and reducing leukemia cells replicative potential, thus suggesting new potential targeting.

REFERENCES

1 Cobaleda, C., Schebesta, A., Delogu, A. & Busslinger, M. Pax5: the guardian of B cell identity and function. *Nature Immunology* **8**, 463-470 (2007).

2 Mullighan, C. G. *et al.* Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* **446**, 758-764 (2007).

3 Pui, C. H., Relling, M. V. & Downing, J. R. Acute lymphoblastic leukemia. *N Engl J Med* **350**, 1535-1548 (2004).

4 Nebral, K. *et al.* Incidence and diversity of PAX5 fusion genes in childhood acute lymphoblastic leukemia. *Leukemia* **23**, 134-143 (2009).

5 Coyaud, E. *et al.* Wide diversity of PAX5 alterations in B-ALL: a Groupe Francophone de Cytogenetique Hematologique study. *Blood* **115**, 3089-3097, doi:blood-2009-07-234229 [pii]10.1182/blood-2009-07-234229 (2010).

6 Strehl, S., Konig, M., Dworzak, M. N., Kalwak, K. & Haas, O. A. PAX5/ETV6 fusion defines cytogenetic entity dic(9;12)(p13;p13). *Leukemia* **17**, 1121-1123 (2003).

7 Cazzaniga, G. *et al.* The paired box domain gene PAX5 is fused to ETV6/TEL in an acute lymphoblastic leukemia case. *Cancer Res* **61**, 4666-4670 (2001).

8 An, Q. *et al.* Variable breakpoints target PAX5 in patients with dicentric chromosomes: a model for the basis of unbalanced translocations in cancer. *Proc Natl Acad Sci U S A* **105**, 17050-17054 (2008).

9 Bohlander, S. K. ETV6: a versatile player in leukemogenesis. *Semin Cancer Biol* **15**, 162-174 (2005).

10 Fazio, G., Palmi, C., Rolink, A., Biondi, A. & Cazzaniga, G. PAX5/TEL acts as a transcriptional repressor causing down-modulation of CD19, enhances migration to CXCL12, and confers survival advantage in pre-BI cells. *Cancer Res* **68**, 181-189 (2008).

11 Fazio, G. *et al.* PAX5/ETV6 alters the gene expression profile of precursor B cells with opposite dominant effect on endogenous PAX5. *Leukemia* **27**, 992-995, doi:leu2012281 [pii]10.1038/leu.2012.281 (2013).

12 Fortschegger, K., Anderl, S., Denk, D. & Strehl, S. Functional Heterogeneity of PAX5 Chimeras Reveals Insight for Leukemia Development. *Mol Cancer Res*, doi:1541-7786.MCR-13-0337 [pii]10.1158/1541-7786.MCR-13-0337 (2014).

13 Filipp, D., Ballek, O. & Manning, J. Lck, membrane microdomains, and TCR triggering machinery: defining new rules of engagement. *Frontiers in Immunology* **3**, 1-14 (2012).

14 De Keersmaecker, K. *et al.* NUP214-ABL1-mediated cell proliferation in T-cell acute lymphoblastic leukemia is dependent on the LCK kinase and various interacting proteins. *Haematologica* **99**, 85-93, doi:haematol.2013.088674 [pii]10.3324/haematol.2013.088674 (2014).

15 Harr, M. W. *et al.* Inhibition of Lck enhances glucocorticoid sensitivity and apoptosis in lymphoid cell lines and in chronic lymphocytic leukemia. *Cell Death Differ* **17**, 1381-1391, doi:cdd201025 [pii]10.1038/cdd.2010.25 (2010).

16 Talab, F., Allen, J. C., Thompson, V., Lin, K. & Slupsky, J. R. LCK is an important mediator of B-cell receptor signaling in chronic lymphocytic leukemia cells. *Mol*

Cancer Res **11**, 541-554, doi:1541-7786.MCR-12-0415-T [pii]10.1158/1541-7786.MCR-12-0415-T (2013).

17 Spijkers-Hagelstein, J. A., Mimoso Pinhancos, S., Schneider, P., Pieters, R. & Stam, R. W. Src kinase-induced phosphorylation of annexin A2 mediates glucocorticoid resistance in MLL-rearranged infant acute lymphoblastic leukemia. *Leukemia* **27**, 1063-1071, doi:leu2012372 [pii]10.1038/leu.2012.372 (2013).

18 Accordi, B. *et al.* Functional protein network activation mapping reveals new potential molecular drug targets for poor prognosis pediatric BCP-ALL. *PLoS One* **5**, e13552, doi:10.1371/journal.pone.0013552 (2010).

19 Crespo, M. *et al.* ZAP-70 expression in normal pro/pre B cells, mature B cells, and in B-cell acute lymphoblastic leukemia. *Clin Cancer Res* **12**, 726-734, doi:12/3/726 [pii]10.1158/1078-0432.CCR-05-1531 (2006).

20 Herzog, S., Reth, M. & Jumaa, H. Regulation of B-cell proliferation and differentiation by pre-B-cell receptor signalling. *Nat Rev Immunol* **9**, 195-205 (2009).

21 Shi, M., Cooper, J. C. & Yu, C. L. A constitutively active Lck kinase promotes cell proliferation and resistance to apoptosis through signal transducer and activator of transcription 5b activation. *Mol Cancer Res* **4**, 39-45 (2006).

SUPPLEMENTARY TABLES

Gene name	Primer sequence	Probe number
h_*ABL* L	aggaatccagtatctcagacgaa	57
h_*ABL* R	ggaggtcctcgtcttggtg	
h_*LCK* L	agtcagatgtgtggtcttttgg	18
h_*LCK* R	cctccgggttggtcatc	
m_*Ccnd2* L	ctgtgcatttacaccgacaac	45
m_*Ccnd2* R	cactaccagttcccactccag	
m_*Hprt* L	ggagcggtagcacctcct	69
m_*Hprt* R	ctggttcatcatcgctaatcac	
m_*cMyc* L	cctagtgctgcatgaggaga	77
m_*cMyc* R	tccacagacaccacatcaattt	
m_*Csk* L	gcctgaagccttgagagaga	31
m_*Csk* R	ttcaggggaattcttgggta	
m_*Lck* L	cgtgtgtgaaaactgccact	21
m_*Lck* R	gagatccctcataggtgaccag	
m_*Zap70* L	cagaccgacggcaagttc	12
m_*Zap70* R	ccatagaccagggacagtgc	

Supplementary Table S1. List of primers and UPL probe numbers used.

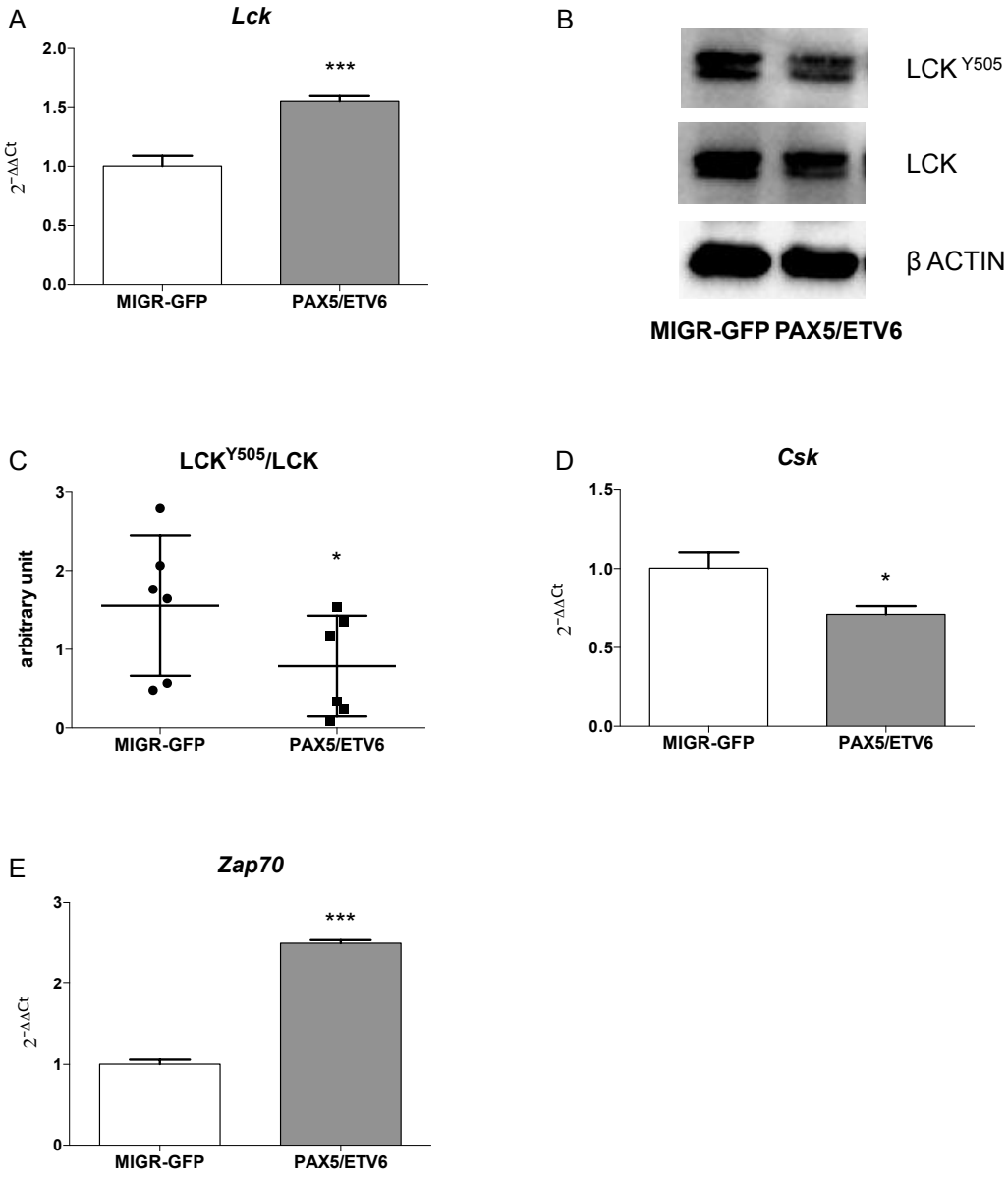
Antigen	Fluorochrome	Source
anti $STAT5^{Y694}$ mouse	Alexa-647	BD
IgG1 κ isotype control	Alexa-647	BD
CD127 mouse	APC	eBioscience

Supplementary Table S2. Antibodies used for FACS analysis.

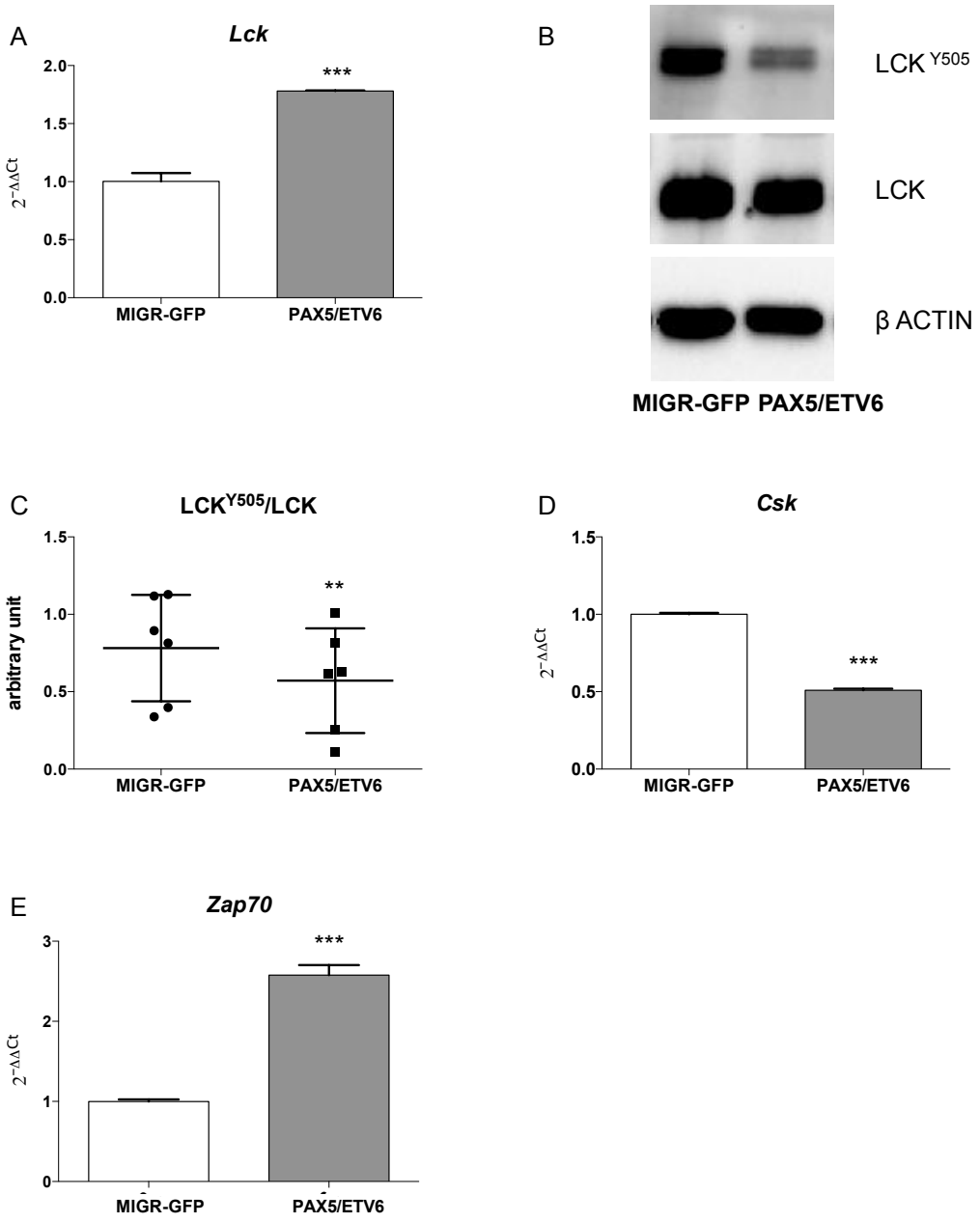
Antigen	Source
LCK^{Y505}	Cell Signaling Technology
LCK	Cell Signaling Technology
$STAT5^{Y694/Y699}$	Merck-Millipore
STAT5 (3H7)	Cell Signaling Technology
ACTB	Sigma

Supplementary Table S2. List of antibodies used for western blotting.

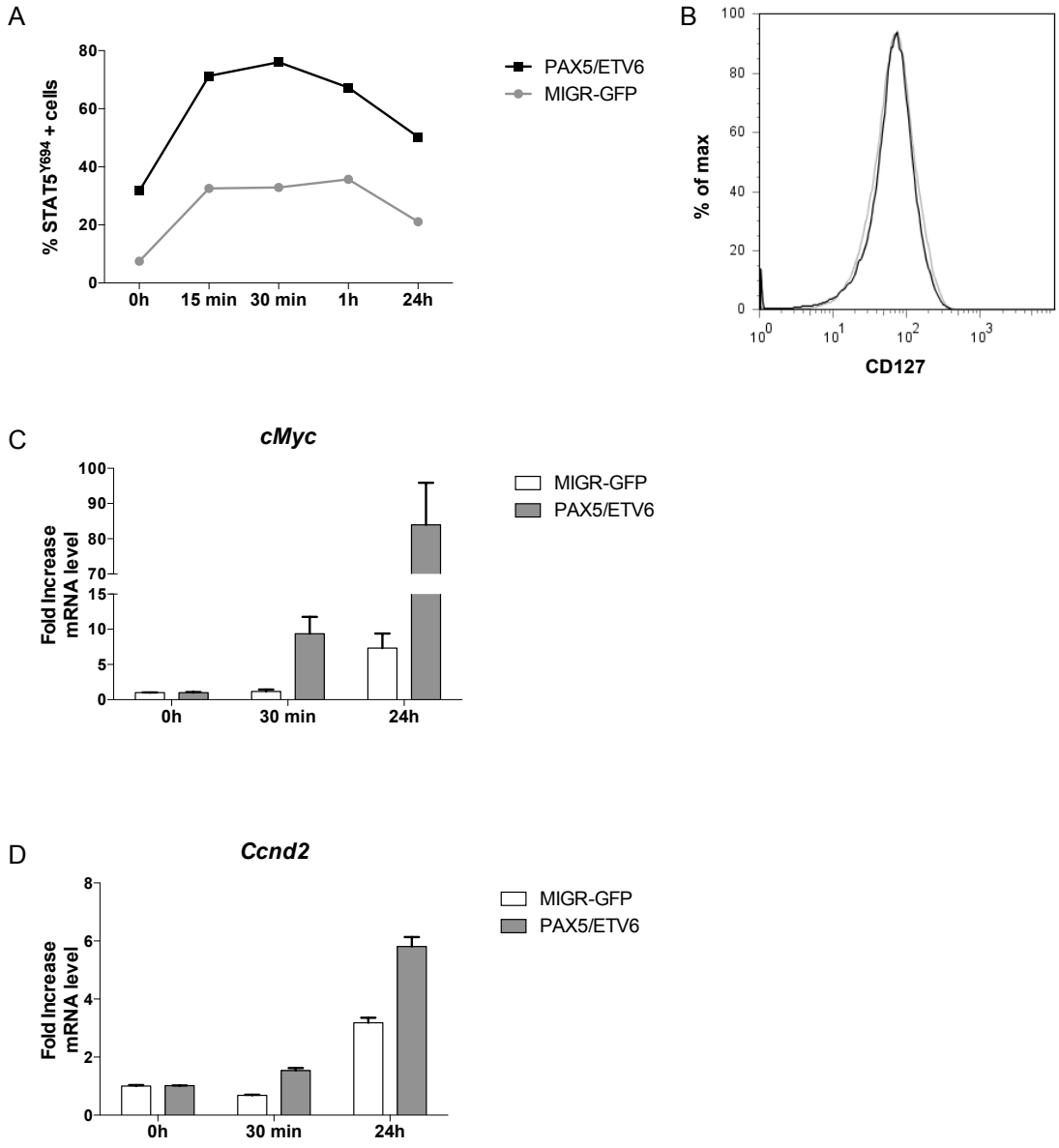
SUPPLEMENTARY FIGURES



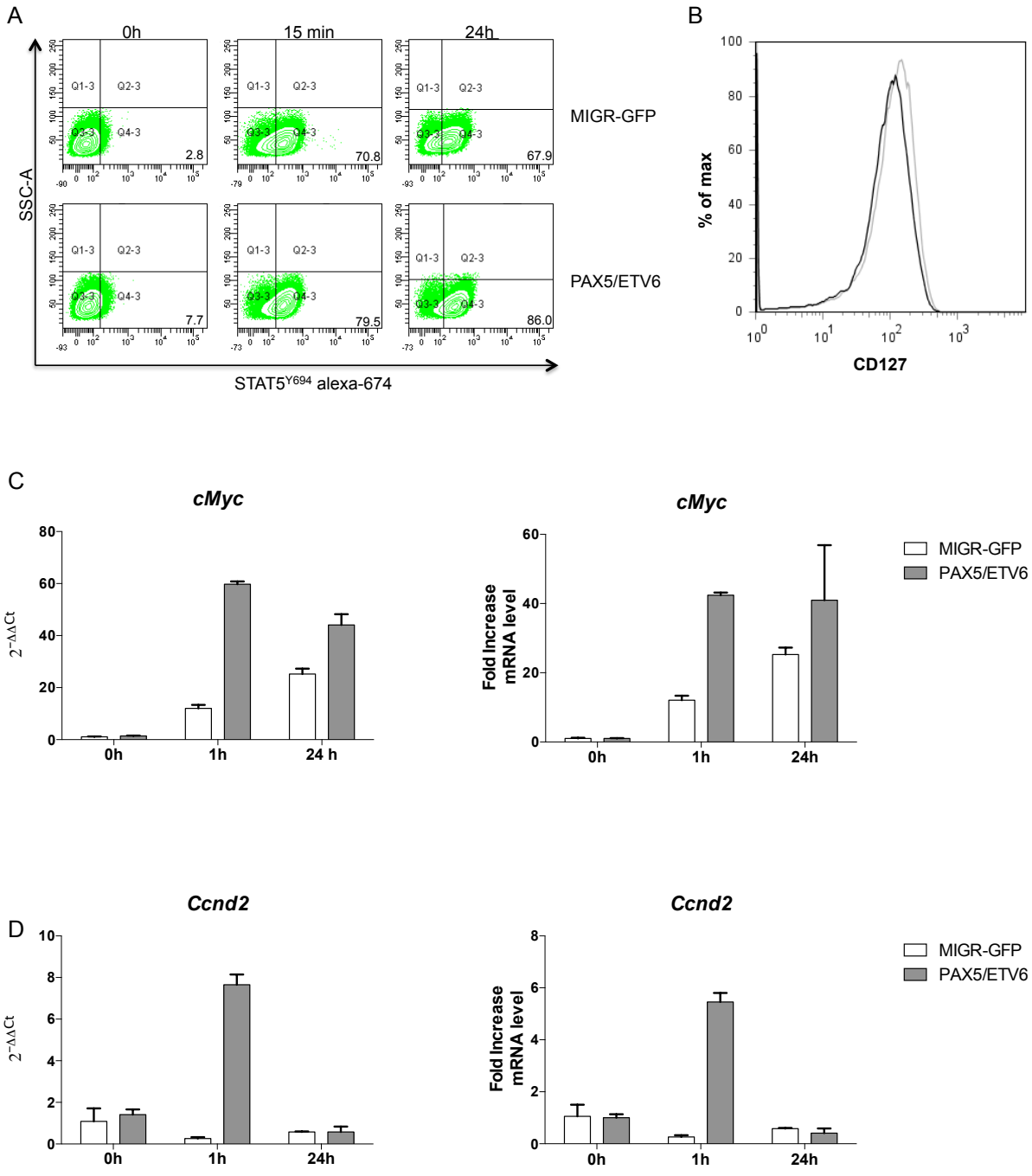
Supplementary Figure S1. A) *Lck* mRNA expression levels in PAX5/ETV6 transduced B6BAFL pre-BI cells. B) LCK protein expression and C) schematic representation of LCK^{Y505}/LCK ratio in B6BAFL pre-BI cells, summarizing multiple western blot experiments. D) *Csk* and E) *Zap70* mRNA expression levels determined by RQ-PCR.



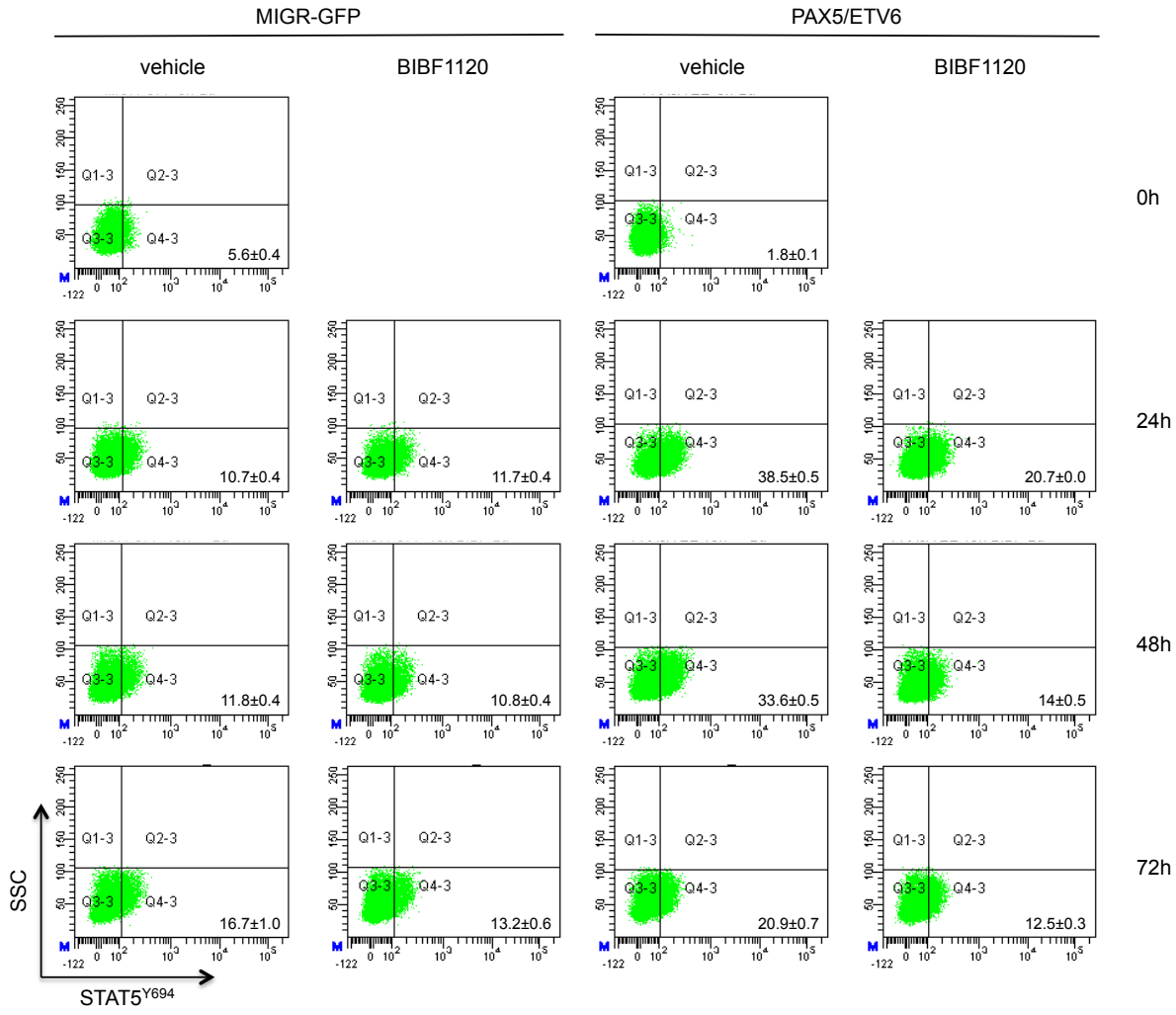
Supplementary Figure S2. A) *Lck* mRNA expression levels in PAX5/ETV6 transduced FLB6-67 pre-BI cells. B) LCK protein expression and C) schematic representation of LCK^{Y505}/LCK ratio in FLB6-67 pre-BI cells. D) *Csk* and E) *Zap70* mRNA expression levels determined by RQ-PCR.



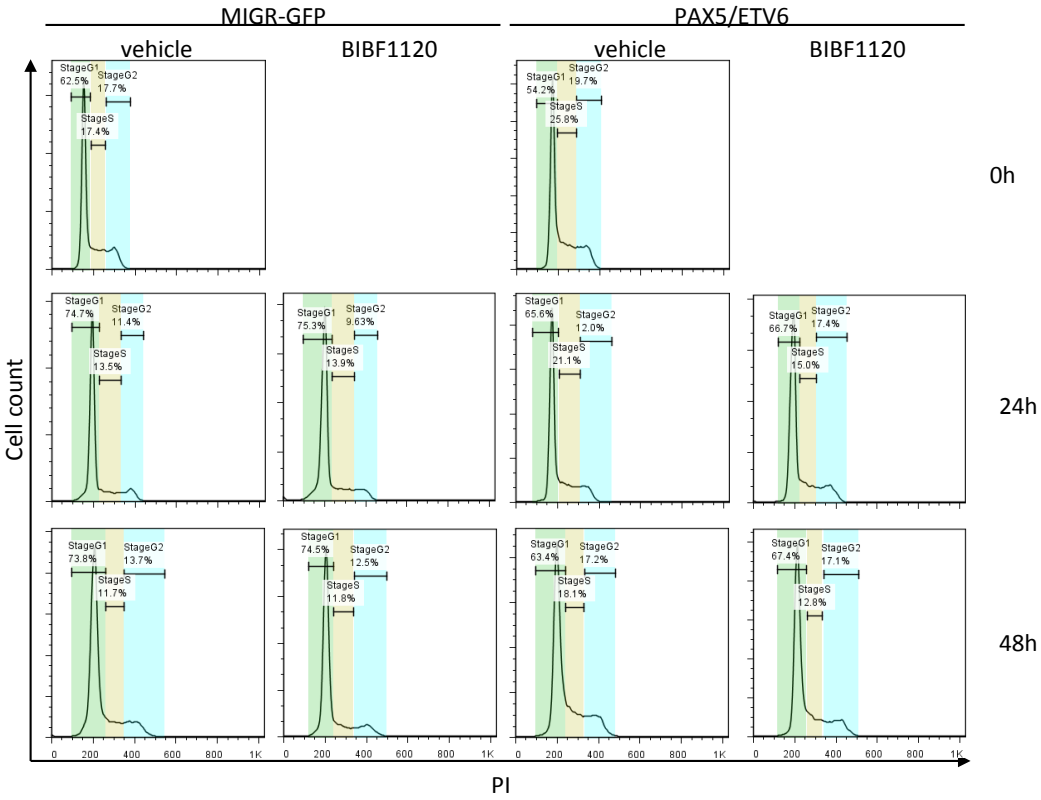
Supplementary Figure S3. A) Schematic representation of the percentage of STAT5^{Y694} positive cells in time course experiment in LY5.1FL pre-BI cells. B) CD127 protein expression is similar in PAX5/ETV6 (black line) versus MIGR-GFP transduced cells (grey line). C) Fold increase calculated on time 0h of *cMyc* and *Ccnd2* mRNA expression levels.



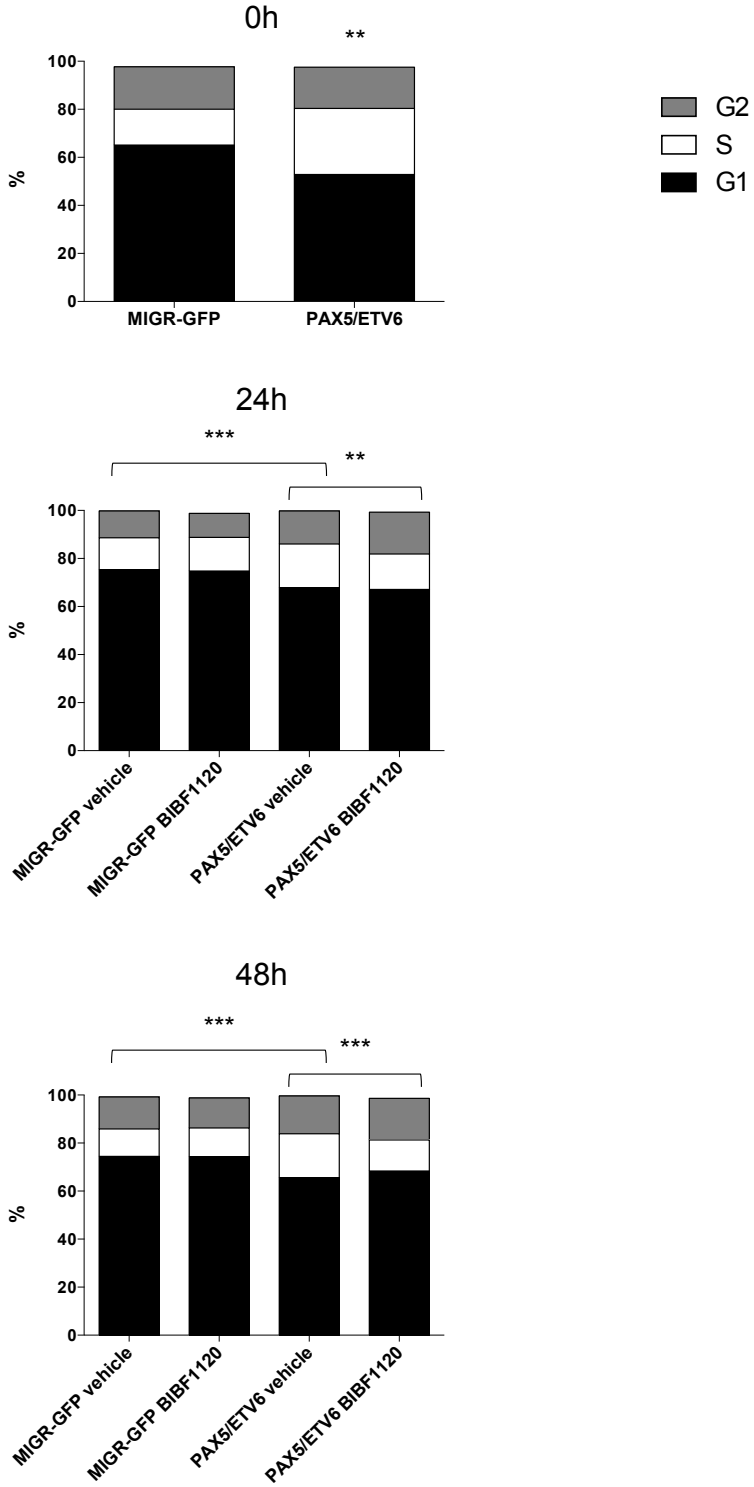
Supplementary Figure S4. A) Representative dot plots of STAT5^{Y694} phenotypes in B6BAFL cells. B) CD127 protein expression is similar in PAX5/ETV6 (black line) versus MIGR-GFP transduced cells (grey line). C) RQ- PCR of *cMYC* and D)*Ccnd2* at early and late time points after IL7 administration and correspondent fold increase calculated on time 0h of *cMyc* and *Ccnd2* mRNA expression levels.



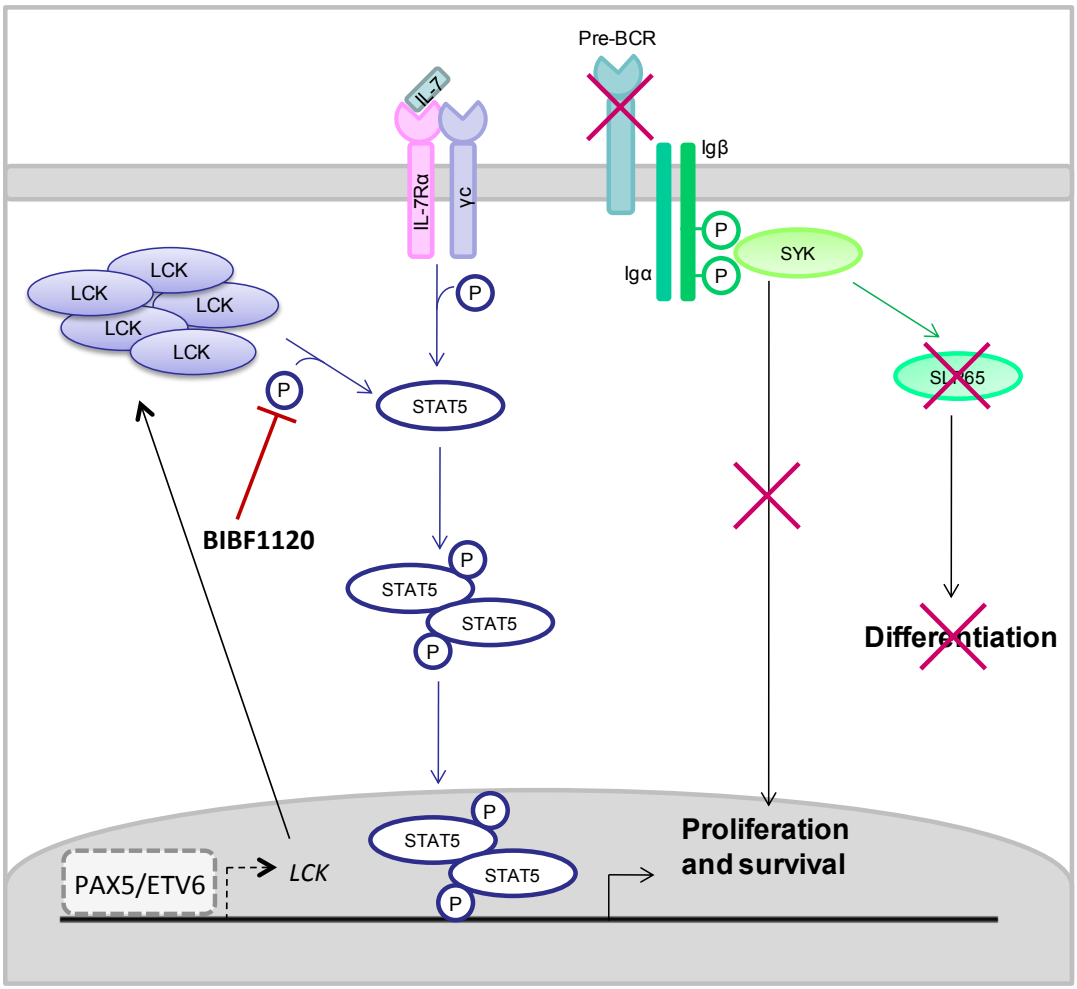
Supplementary Figure S5. Representative dot plots of STAT5^{Y694} phenotypes after BIBF1120 administration in LY5.1FL cells.



Supplementary Figure S6. Example of cell cycle analysis in LY5.1FL cells. G1 phase, green; S phase, yellow; G2 phase, blue.



Supplementary Figure S7. Example evaluation of cell cycle analysis in LY5.1FL cells.



Supplementary Figure S8. Schematic representation of the effect of PAX5/ETV6 in pre-BI cells.

Chapter 5

Additional results:
Mechanism of PAX5 tumor suppression in ALL

RESULTS

Patients cohort

In order to study the effect of *PAX5* in disease progression in ALL, we selected 2 main cohorts: Ph+ and Ph- patients.

Ph+ cases are characterized by the translocation t(9;22), that fuses *BCR* to the *ABL1* proto-oncogene. The 2 major forms of BCR/ABL proteins, P190 and P210, differ only in the amount of BCR amino acid residues included.[1] The Ph- subgroup is formed by cases negative for the translocation t(4;11) (MLL-r), t(12;21) (ETV6/RUNX1), t(9;22) (BCR-ABL), t(1;19) (E2A-PBX1).

We screened the diagnosis of BCP-ALL pediatric patients in the 2 subgroups by MLPA and/or SNP in order to identify cases carrying *PAX5* alterations.

We selected 14 pediatric patients with wt or deleted *PAX5* status both in Ph- and Ph+ cohort. In detail, we chose 5 PAX5-deleted Ph+ cases (MXP1-5) and 2 Ph+ with wt PAX5 status as correspondent control (MXP11-12), 5 PAX5-deleted Ph- samples (MXP6-10) and 2 wt PAX5 Ph- as controls (MXP13-14). The patients' characteristics and their identified lesions are reported in Table 1.

		PAX5	IKZF1	EBF1	CDKN2A	CDKN2B	ETV6	BTG1	RB1	P2RY8-CRLF2
MXP1	Ph+	1	1	0	1	1	0	0	0	0
MXP2	Ph+	1	1	0	1	1	0	0	0	0
MXP3	Ph+	1	1	0	1	1	0	0	0	0
MXP4	Ph+	1	1	0	1	1	1	0	0	0
MXP5	Ph+	1	1	1	0	0	0	0	0	0
MXP6	Ph-	1	0	0	0	0	0	0	0	0
MXP7	Ph-	1	0	0	1	2	0	2	2	0
MXP8	Ph-	1	1	0	2	2	0	0	0	0
MXP9	Ph-	1	0	0	0	0	0	0	2	0
MXP10	Ph-	1	0	0	0	0	1	0	0	0
MXP11	Ph+	0	1	0	0	0	0	0	0	0
MXP12	Ph+	0	1	0	0	0	0	0	0	0
MXP13	Ph-	0	0	0	1	1	0	0	0	1
MXP14	Ph-	0	1	0	1	0	0	0	0	0

Table 1. Patients cohort and lesions identified by MLPA/SNPs analysis.

Establishment of ALL xenografts, as a representative model of human disease

In order to unravel the role of PAX5 in ALL, we have established xenografts using BM cells collected at the diagnosis of selected patients (Table 2).

	UPN	PAX5	Lesions	Age	Material
MXP1	1027306	Deleted	t(9;22) P190	5	BM
MXP2	1028641	Deleted	t(9;22) P190	6	BM
MXP3	1021100	Deleted	t(9;22) P190/P210	13	BM
MXP4	1025787	Deleted	t(9;22) P190	14	BM
MXP5	1029905	Deleted ex(2-6)	t(9;22) P190	6	BM
MXP6	1021326	Deleted ex(2-6)	t(4;11), t(12;21), t(1;19) t(9;22) neg	5	BM
MXP7	1022289	Deleted ex(2-6)	t(4;11), t(12;21), t(1;19) t(9;22) neg	16	BM
MXP8	1023217	Deleted ex(2-6)	t(4;11), t(12;21), t(1;19) t(9;22) neg	12	BM
MXP9	1024420	Deleted ex(2-6)	t(4;11), t(12;21), t(1;19) t(9;22) neg	2	BM
MXP10	1024891	Deleted ex(2-6)	t(4;11), t(12;21), t(1;19) t(9;22) neg	3	BM
MXP11	1036075	wt	t(9;22) P190	8	BM
MXP12	1032706	wt	t(9;22) P210	10	BM
MXP13	1025341	wt	t(4;11), t(12;21), t(1;19) t(9;22) neg	5	BM
MXP14	1037380	wt	t(4;11), t(12;21), t(1;19) t(9;22) neg	17	BM

Table 2. Clinical feauters of the patients cohort. Wt, wild type; BM, bone marrow.

In particular, 1 million cells have been injected into the femur of sub-lethally irradiated NSG mice. We periodically checked the percentage of the engraftment in NSG mice, in order to monitor the leukemia progression. Two months after transplantation, cells from peripheral blood have been collected and stained for hCD45 and mCD45, in order to estimate the engraftment of human cells. A representative analysis is reported in Figure 1.

Once the mice manifested overt signs of leukemia, they were sacrificed and BM and spleen were collected. A schematic representation of the system is reported in Figure 2.

This model offers the following advantages:

a) leukemia cells harvested from the mice are available in abundance (more than 200 million cells can be collected from each mouse spleen);
b) xenografts cells are highly representative of the human disease, showing phenotypically and genotypically stable characteristics of the primary patients;
c) splenocytes collected from secondary mice are able to proliferate on OP9 stroma, representing a physiological model to study their interactions with the BM niche.

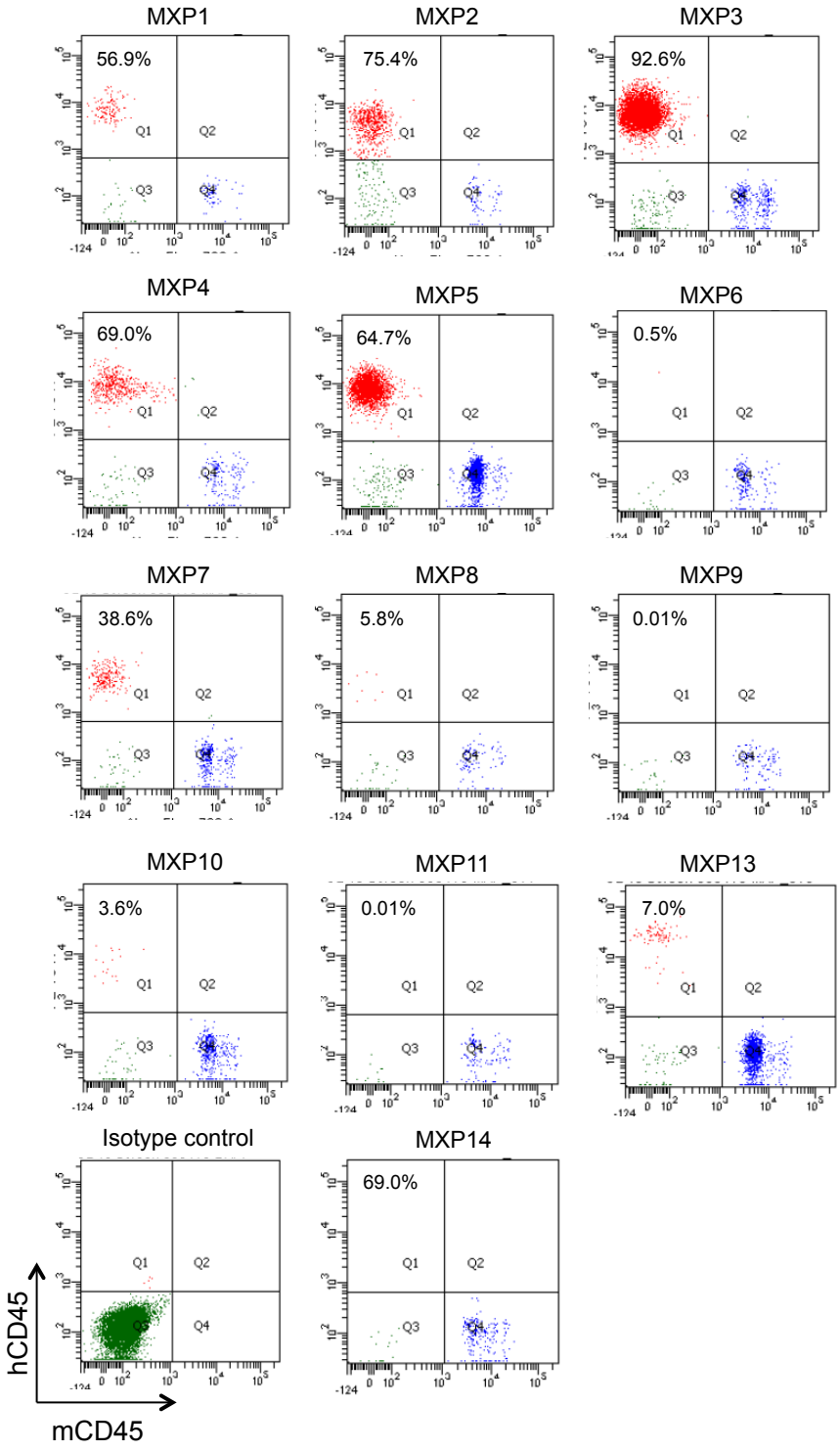


Figure 1. Representative FACS analysis on the peripheral blood of primary xenografts 2 months after the transplantation. In order to estimate the engraftment of human samples, we used specific antibodies against human and and murine CD45 marker.

We completed the injection and collection of material from primary mice of all the transplanted patients (n=14). Secondary xenografts are still ongoing in NSG mice.

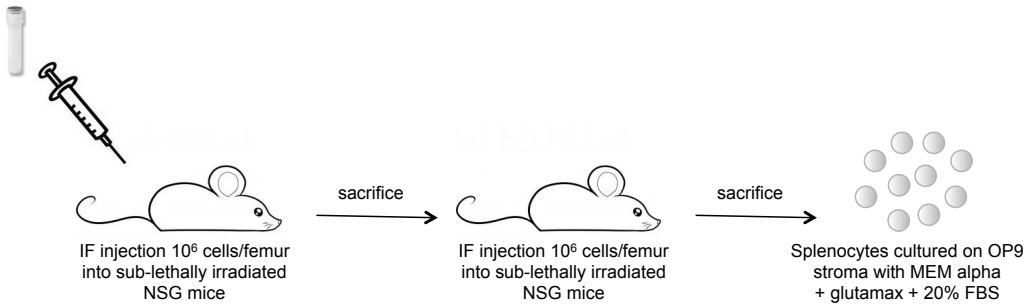


Figure 2. Schematic representation showing the establishment of xenografts from the patients BM to in vitro co-culture on OP9 stroma.

Identification of PAX5 targets by ChIP sequencing

Identifying *PAX5* direct target genes have been the focus of many studies in the last years. It is widely known that PAX5 acts both as a transcriptional activator and a repressor on different target genes, which are involved in lineage development.[2] In particular PAX5 activates genes involved in B cell development and signaling, such as CD19, BLNK, CD79a and BACH2.[3] On the contrary, it represses genes involved in the commitment to other lineages, such as NOTCH1, FLT3 and MCSFR.[4]

However, little is still known about PAX5 alterations and especially about their ability of binding the DNA.[5,6] Therefore, aim of the present task has been to perform ChIP sequencing analysis in order to identify the promoters bound by PAX5 in leukemia setting and especially in presence of *PAX5* alterations.

The optimal conditions to carry ChIP sequencing have been identified performing single locus qChIP using Ph+ ALL xenografts cells. We immunoprecipitated the chromatin using either anti-PAX5 antibody, that recognizes the N-terminal region (PAX5 N19, Santa Cruz Biotechnology), or the correspondent IgG control (anti Goat IgG). We referred to the promoter of *CD19* gene, a known PAX5 target, as positive control, and to the *GAPDH* promoter as negative control. The workflow and an example of single locus qChIP performed are reported in Figure 3. The analysis of the ChIP sequencing, that will identify PAX5 targets in ALL, is still ongoing.

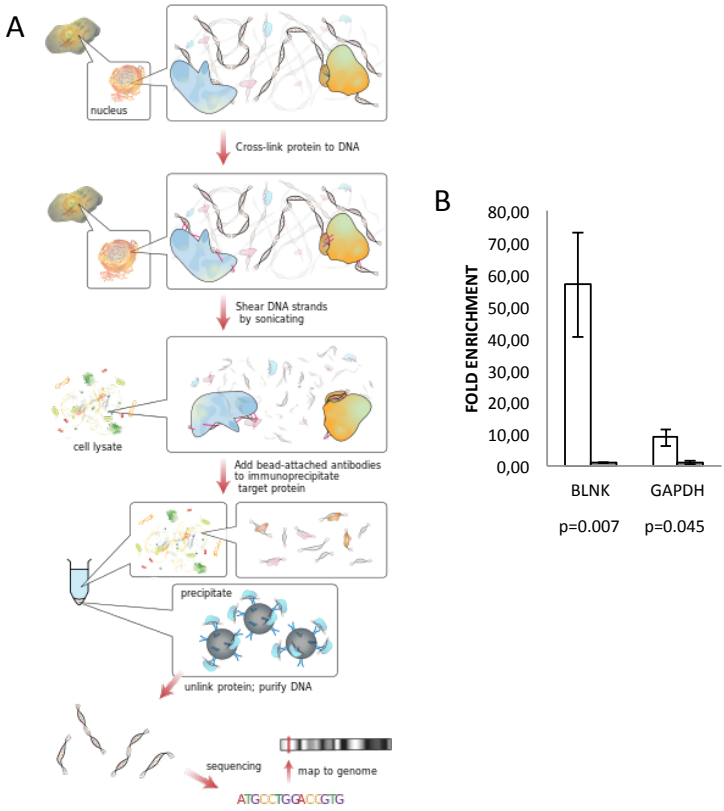


Figure 3. PAX5 ChIP sequencing. A) Workflow of ChIP sequencing procedure. B) Single locus qChIP using the ChIP grade antibody PAX5 N19 (Santa Cruz Biotechnolgies), showing enrichment on the CD19 promoter (positive control) and low binding activity on the GAPDH promoter (negative control).

Analysis of the role of PAX5 in phoshorylation status

Moreover we looked at the perturbations induced by *PAX5* alterations in key players of different signaling pathways with a phospho-protein array (Human Phospho-Kinase Array Kit, Proteome Profiler Array, R&D). In detail, 10 million cells collected from the spleen of primary xenografts have been lysed and incubated on the phospho-protein array. We normalized the obtained results on the phopsphorylation profile of the wt *PAX5* sample. The results obtained in Ph+ ALL xenografts are reported in Figure 4.

The phosphorylation profile in Ph+ xenografts carrying *PAX5* deletions showed perturbations in many proteins. In particular, we identified over-activation both of mediators of pro-survival signaling pathways (e.g. STAT5) and pro-apoptotic molecules (e.g P53), therefore suggesting that PAX5 could be responsible of maintaining the balance between pro-apoptotic and pro-survival signals in Ph+ ALL.

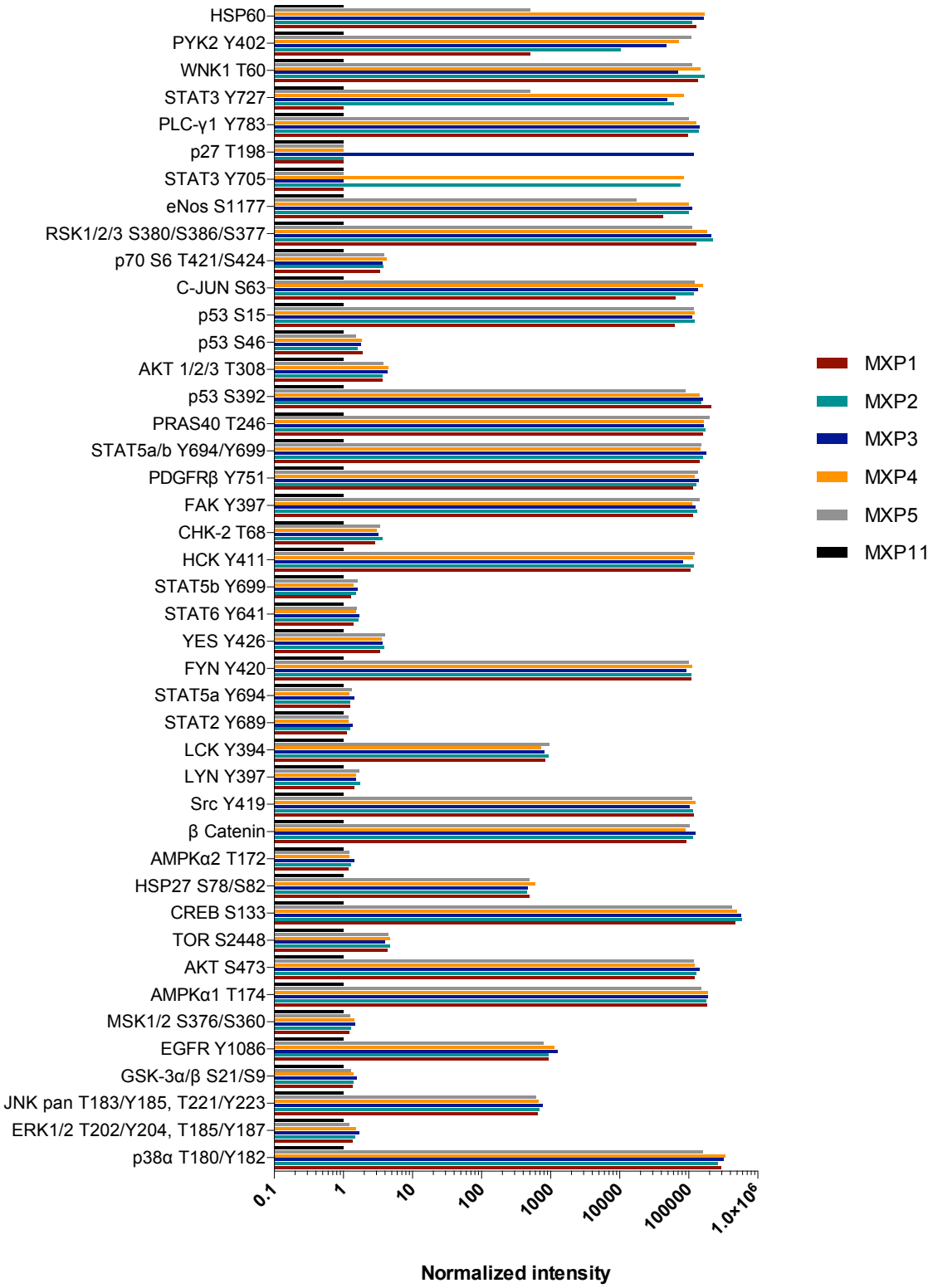


Figure 4. Analysis of the phopsho-protein profiler array in Ph+ xenografts cells. Intensity values of PAX5 deleted cases have been normalized on the intensity of the wt PAX5 sample.

Generation of an *in vitro* model to unravel PAX5 role in Ph+ ALL

Since phospho-profiling of Ph+ ALL xenografts showed a potential involvement of PAX5 in balancing signaling pathway, we decided to establish an *in vitro* model to further study *PAX5* role in Ph+ ALL.

We generated inducible vectors encoding either wt *PAX5* or the *PAX5* dominant negative *PAX5/ETV6*. Specifically, we cloned either *PAX5* or *PAX5/ETV6* sequences into the lentiviral ER^{T2}-GFP vector, that carries the GFP as a selection marker and the selected gene fused to the ER^{T2} tag. The HSP90 retains the fusion protein into the cytoplasm through the binding to the ER^{T2} tag. When the tamoxifene (4-OHT) is administred, it binds the ER^{T2} tag, discarding the HSP90 and enabling the translocation of the protein to the nucleus and the activation of its target genes (Figure 5).

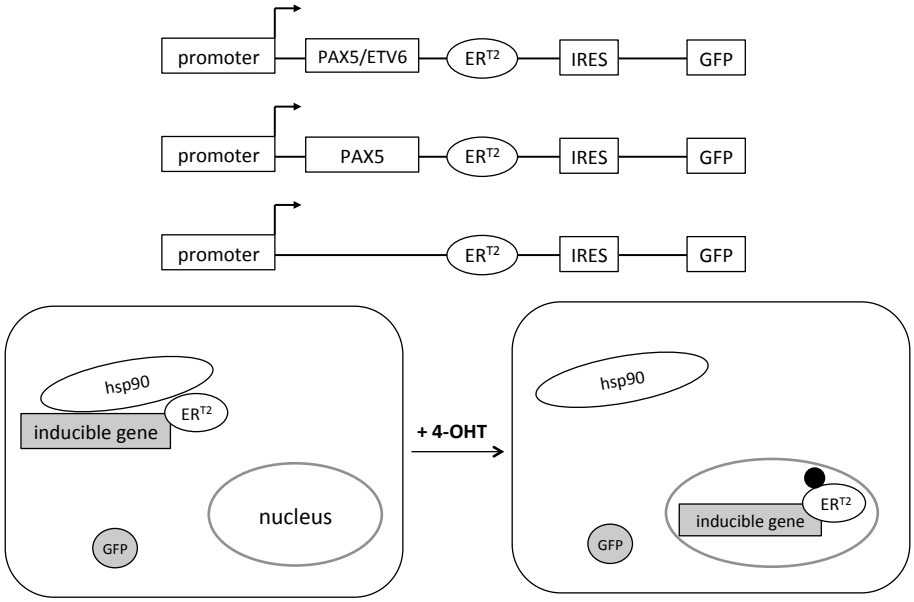


Figure 5. Description of the ER^{T2} lentiviral vectors generated.

We first transduced a previously established xenograft carrying MLL-rearrangement and homozigously deleted in *PAX5*, in order to verify the efficiency of the transduction and the vector. Indeed, the cells showed a good transduction efficiency (>70%) and western blot analysis demonstrated that the protein of interesting translocates to the nucleus after tamoxifene administration (3h), thus validating the system (Figure 6).

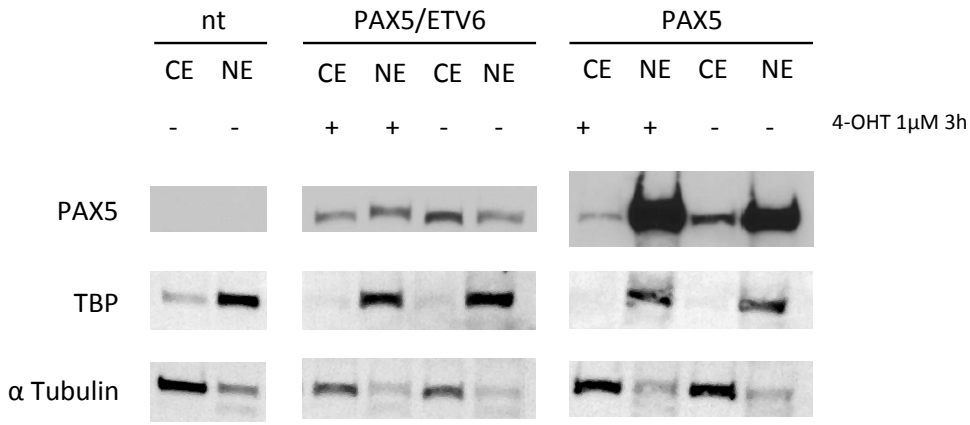


Figure 6. Western blot analysis in lentivirally transduced MLL-rearranged xenograft, carrying homozygous *PAX5* deletion showed the translocation from the cytoplasm (CE) to the nucleous (NE) of PAX5 and PAX5/ETV6 after 4-OHT administration. TBP (TATA binding protein) and α tubulin represent the nuclear and cytoplasmic internal control, respectively.

Then, we transduced cells from previously established Ph+ xenografts [7] carrying wt *PAX5* status either with *PAX5* over-expression, *PAX5/ETV6* or the correspondent empty vector control (EV).

Fourteen days after the induction FACS analysis showed CD19 down-regulation in PAX5/ETV6 cells, as previously described (Figure 7).[8,9]

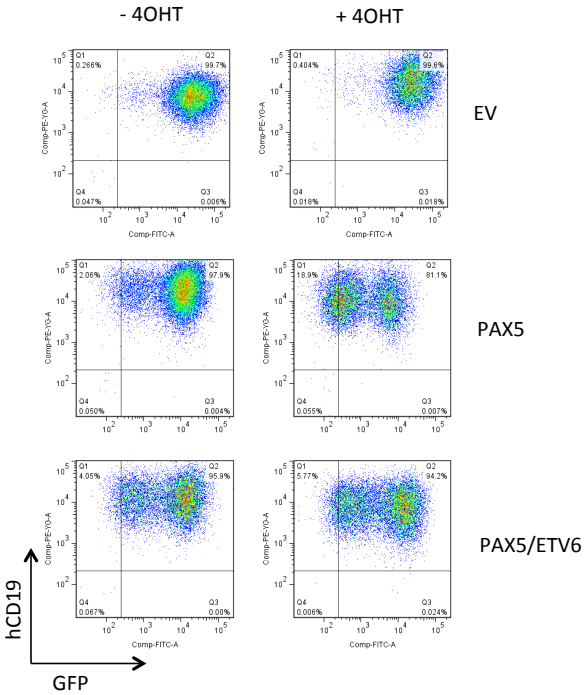


Figure 7. FACS analysis showing the downregulation of CD19 in PAX5/ETV6 positive cells.

Growth competition assays, calculated as the fold increase of GFP+ versus GFP- cells after induction, showed that cells transduced with PAX5 over-expression vector have a growth disadvantage, indicating that PAX5 can act as tumor suppressor in Ph+ ALL. Moreover, the induction of PAX5/ETV6 transduced cells displayed a long term decrease in the percentage of GFP positive cells indicating that the PAX5/ETV6 oncogene can induced hyper-signaling, and thus leading to senescence in Ph+ ALL xenograft cells (Figure 8). [10]

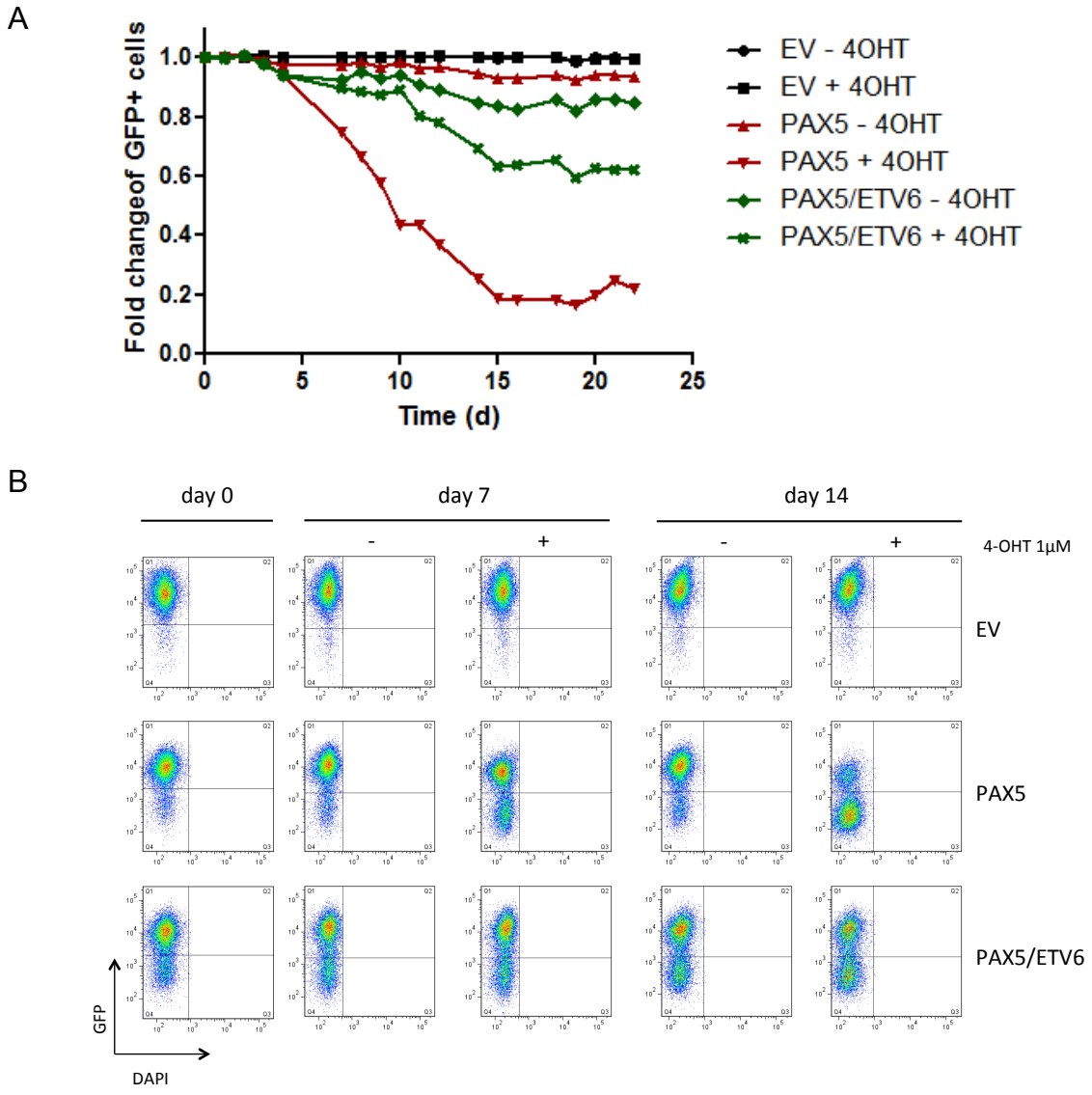


Figure 8. A) Growth competition assay performed in xenografts cells derived from the BM of a Ph+ patient. B) Key time points of FACS analyses performed on Ph+ xenografts transduced cells, using GFP as marker of transduction and DAPI- as a marker of viability.

Moreover, western blot analysis of GFP positive cells 3 hours post tamoxifene induction, indicated that PAX5 overexpression causes the de-phosphorylation of STAT5, while PAX5/ETV6 induces its over-activation, thus partially explaining the different growth ability (Figure 9).

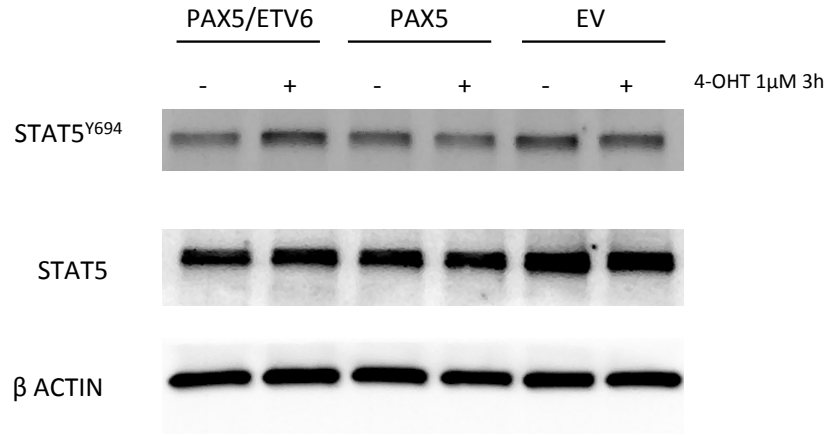


Figure 9. Western blot analysis of transduced Ph+ xenografts cells, showing that PAX5/ETV6 induces STAT5 activation, whereas PAX5 over-expression leads to STAT5 dephosphorylation.

CONCLUSIONS AND FUTURE PERSPECTIVE

In order to study the role of *PAX5* in ALL we screened and selected a small cohort of patients, characterized by Ph+ and Ph- cases with or without *PAX5* wt status. Then, we established a xenograft system that offers the advantage of expanding cells from primary patients and the opportunity of studying them *in vitro* on OP9 stroma, thus mimicking the bone marrow micro-environment.

Since PAX5 has been described as a frequent target of alteration in ALL, we wondered if its alteration could affect its ability of binding the promoters of target genes described in physiological conditions. In particular, our aim was to comprehend if PAX5 deleted cases showed different target genes and thus to understand if the gene dosage is important for the activation of the transcription of target genes.[11] The ChIP sequencing analysis on PAX5 wt and PAX5 deleted patients will elucidate this hypothesis.

Moreover, phospho-array profiling showed that PAX5 deletion is able to induce perturbations in the phosphorylative profile of Ph+ ALL xenografts compared to wt PAX5 cases. These alterations include both over-activation of pro-survival and pro-

apoptotic mediators of signaling, thus suggesting that PAX5 may be responsible of the balance between survival and apoptosis in Ph+ ALL.

The generation of lentiviral inducible vectors has been fundamental to study the role of *PAX5* in ALL, in particular creating the condition of studying the effect of the over-expression of *PAX5* and the dominant negative *PAX5/ETV6*, as a model of induced *PAX5* haplo-insufficiency.

The transduction of human patient-derived xenograft cells with the described vectors demonstrated that PAX5 acts as a tumor suppressor in Ph+ ALL, whereas long term PAX5/ETV6 cells show the characteristics of senescence induced by hyper-signaling.[10] The exact mechanism has still to be unraveled, and in particular we aim to determine if the growth disadvantage is due to either increased apoptosis or cell cycle block in PAX5 over-expressed human cells.

REFERENCES

1 Voncken, J. W. *et al.* BCR/ABL P210 and P190 cause distinct leukemia in transgenic mice. *Blood* **86**, 4603-4611 (1995).

2 Cobaleda, C., Schebesta, A., Delogu, A. & Busslinger, M. Pax5: the guardian of B cell identity and function. *Nature Immunology* **8**, 463-470 (2007).

3 Schebesta, A. *et al.* Transcription Factor Pax5 Activates the Chromatin of Key Genes Involved in B Cell Signaling, Adhesion, Migration, and Immune Function. *Immunity* **27**, 49-63 (2007).

4 Delogu, A. *et al.* Gene repression by Pax5 in B cells is essential for blood cell homeostasis and is reversed in plasma cells. *Immunity* **24**, 269-281 (2006).

5 Kawamata, N., Pennella, M. A., Woo, J. L., Berk, A. J. & Koeffler, H. P. Dominant-negative mechanism of leukemogenic PAX5 fusions. *Oncogene*, doi:onc2011291 [pii]10.1038/onc.2011.291 (2011).

6 Fortschegger, K., Anderl, S., Denk, D. & Strehl, S. Functional Heterogeneity of PAX5 Chimeras Reveals Insight for Leukemia Development. *Mol Cancer Res*, doi:1541-7786.MCR-13-0337 [pii]10.1158/1541-7786.MCR-13-0337 (2014).

7 Trageser, D. *et al.* Pre-B cell receptor-mediated cell cycle arrest in Philadelphia chromosome-positive acute lymphoblastic leukemia requires IKAROS function. *J Exp Med* **206**, 1739-1753 (2009).

8 Fazio, G., Palmi, C., Rolink, A., Biondi, A. & Cazzaniga, G. PAX5/TEL acts as a transcriptional repressor causing down-modulation of CD19, enhances migration to CXCL12, and confers survival advantage in pre-BI cells. *Cancer Res* **68**, 181-189 (2008).

9 Fazio, G. *et al.* PAX5/ETV6 alters the gene expression profile of precursor B cells with opposite dominant effect on endogenous PAX5. *Leukemia* **27**, 992-995, doi:leu2012281 [pii]10.1038/leu.2012.281 (2013).

10 Braig, M. & Schmitt, C. A. Oncogene-induced senescence: putting the brakes on tumor development. *Cancer Res* **66**, 2881-2884, doi:66/6/2881 [pii]10.1158/0008-5472.CAN-05-4006 (2006).

11 Sigvardsson, M. Transcription factor dose links development to disease. *Blood* **120**, 3630-3631, doi:120/18/3630 [pii]10.1182/blood-2012-09-455113 (2012).

Chapter 6

Discussion and future perspectives

In the present study, we aimed to analyze the role of *PAX5* alterations in BCP-ALL. Therefore, we analyzed the samples of newly diagnosed patients carrying abnormalities in the 9p chromosome to identify PAX5 alterations. Then, we used an established *in vitro* model of wt pre-BI cells transduced with PAX5/ETV6, the most recurrent PAX5 fusion gene, in order to identify changes in the gene expression profile[1] and in signaling pathways.

We described *PAX5* fusions to four new partner genes (*CHFR*, *SOX5*, *POM121C* and *MLLT3*) and we reported 2 new cases of the *PAX5/AUTS2* fusion gene, confirming it as a frequent partner of *PAX5* in translocations.[2]

Overall, *PAX5* fusion genes can be classified in several subgroups on the basis of the molecular function of the partner gene. As shown in Figure 1, the partner genes can be transcription factors or molecules involved in the transcription machinery, structural proteins, kinases, transporters, and others.[3]

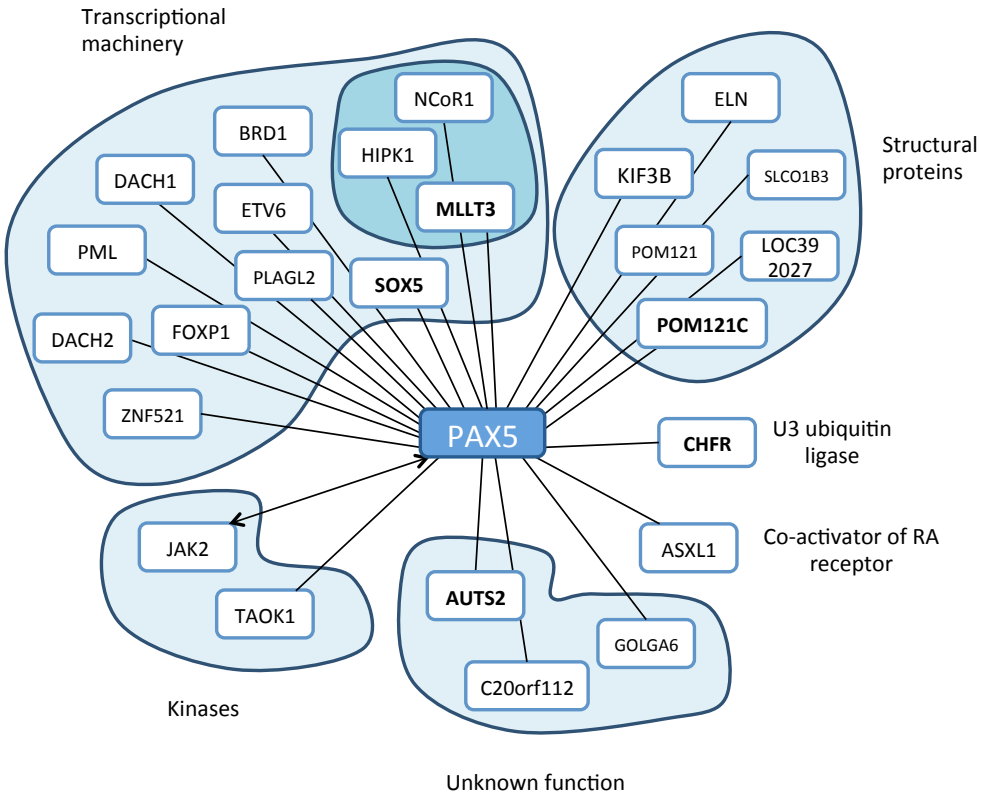


Figure 1. Schematic representation of *PAX5* partner genes and their molecular functions.

The main characteristic of *PAX5* fusion genes is the retention of the PAX5 DNA binding domain. It has been described that the chimeric proteins show the tendency to multimerize. The multimerization of the DNA binding domains leads to increased affinity of binding to the chromatin compared to wt PAX5. Higher stability on the chromatin results in dominant negative activity by competition with the endogenous transcription factor for binding to PAX5-binding sites.[4]

In order to overcome the lack of material from primary patients, we established an *in vitro* model in which murine wt pre-BI cells have been retrovirally transduced with *PAX5/ETV6*, the most recurrent *PAX5* fusion gene, as a model to recapitulate *PAX5* translocations.

Indeed, in the present study we showed that PAX5/ETV6 is an aberrant transcription factor. It disregulates the gene expression profile of pre-BI cells, mainly acting on PAX5 target genes. In detail, PAX5/ETV6 exhibits not only the characteristics of a dominant negative, thus down-regulating *PAX5* activated genes and up-regulating *PAX5* repressed targets, but it has a more complex role herein defined as opposite dominant function.[1,5]

Moreover, PAX5/ETV6 induces alterations in the expression levels of many molecules involved in adhesion process, thereby causing decreased adhesive capacity toward VCAM1, a molecule expressed in the BM niche.[1] On the contrary, PAX5/ETV6 cells showed increased migratory capacity towards CXCL12, therefore depicting the scenario of a typical tumor cell characterized by metastatic potential.[6]

Our group previously demonstrated that PAX5/ETV6 cells are immature and show a differentiation block.[6,7] Indeed, they cannot complete the IgM rearrangement and they directly down-regulate many of the pre-BCR components, such as the μ chain itself, Igα/MB-1, and BLNK/SLP-65, as previously demonstrated in the contest of BCR/ABL1 fusione gene.[8]

Herein, we showed that PAX5/ETV6 induces changes not only in the gene expression profile but also in the signaling pathway. In particular, we showed the *PAX5* fusion gene driven up-regulation of *LCK*, thus leading to over-activation of the signaling mediator STAT5 and eventually causing a survival advantage, as previously described in the presence of TGFβ1 and after withdrawal of IL7.[6]

Taken together, this data suggests that *PAX5* fusion genes are able to induce perturbations at multiple levels, thus giving to the cells the typical features of cancer:

a) altered gene expression profile;[1]
b) increased pro-survival signaling;[6]
c) survival advantage in presence of pro-apoptotic signals (TGFβ1) and in absence of physiological conditions (IL7 withdrawal);[6]
d) block in the differentiation;[6]
e) decreased adhesion and increased migration.[1,6]

Since PAX5 fusion proteins are able to induce leukemic phenotype, we wondered about the role of *PAX5* as a tumor suppressor in ALL. Therefore, we established patient-derived xenografts from a small group of *PAX5* deleted childhood BCP-ALL patients. We aimed to identify the PAX5 target genes in a leukemic subset and evaluate whether the gene dosage could be important for the transcription of target genes identified in physiological situations.[9,10]

Moreover, we identified PAX5 as a potential key player in keeping the balance between pro-survival and pro-apoptotic signals by phospho-protein profile in xenografts cells. The establishment of the *in vitro* model and the generation of inducible vectors encoding for PAX5 and PAX5/ETV6 will be a fundamental to contribution to comprehend the mechanism of PAX5 tumor suppression in ALL.

Very recently, gene expression profiling (GEP) in childhood BCP-ALL cases has identified a genetic subgroup of patients, not genetically classified, with poor outcome, that clusters close to the subgroup characterized by the *BCR/ABL1* fusion gene, thus called 'BCR/ABL1-like'. This group of patients carries mutations in genes that are fundamental in B cell development, such as *IKZF1*, *PAX5*, *VPREB1*, *TCF3*, *EBF1* and pre-BCR.[11,12] In this subgroup, patients carrying *PAX5* deletions are more than 60% (our unpublished data).

PAX5 abnormalities have not been associated to unfavorable prognosis, but it's important to underline that the deletions in the BCR/ABL1-like group not only affect the *PAX5* gene locus but also often include larger regions on chromosome 9p and other genes involved in B cell development. This suggests that the BCR/ABL1-like group does

not exactly overlap with the subgroup of BCP-ALL patients carrying exclusively *PAX5* abnormalities.[13]

A genome-wide DNA copy number analysis on matched diagnosis and relapse samples from pediatric patients affected by ALL detected an increasing number of additional regions of deletion at relapse, including *PAX5* and *IKZF1*. This suggests that genomic abnormalities contributing to ALL relapse, including *PAX5* alterations, might be selected during treatment, paving the way for designing a new therapeutic intervention to target these abnormalities.[14]

One question still remains open: are *PAX5* alterations driver or passengers in ALL?

Multiple studies are aiming to address this question, trying to define whether *PAX5* aberrations are part of a complex scenario of cooperating events or whether they are a unique genetic aberration event that drives leukemogenesis.

Patients carrying *PAX5* deletion have a more complex karyotype than patients with translocation or dicentric chromosomes. Frequently, classical deletion events or dicentric chromosomes giving rise to prematurely truncated *PAX5* transcripts have been described as co-existing in leukemic blast cells together with other *PAX5* unrelated genetic lesions, such as *ETV6/AML1*,[7] *BCR/ABL1* and *TCF3/PBX1* fusion genes.[11,15-17] This observation led us to hypothesize that there is a cooperative role for *PAX5* deletions over the main genetic lesion. Similarly to micro-deletions on fundamental genes in hematopoiesis (such as *IKZF1*, *EBF1*, *TCF3*, *LEF1*),[7,15] *PAX5* aberrancies can contribute to block cell development in B cell precursors, as reflected in the mouse model of homozygous *PAX5* deletion, which is characterized by complete blockage at an early stage of differentiation.[9]

By contrast, patients carrying *PAX5* fusion genes (as a consequence of translocations) were mainly reported as negative for the most common genetic aberrations found in childhood.[18,19] High resolution Copy Number Abnormality (CNA) analysis on diagnostic samples with *PAX5* fusion genes revealed that in both childhood and adult ALL, *PAX5* translocated cases have a simple karyotype.[19] Indeed, in childhood patients carrying *PAX5* translocations, we reported a mean of 2.4 lesions in addition to the translocation event itself (our unpublished data), depicting a simpler scenario compared to the whole cohort of BCP-ALL cases (>4.2 lesions/patient).[20]

In conclusion, it becomes evident that *PAX5* translocations giving rise to fusion genes have a completely different impact on the biology and development of leukemia compared to *PAX5* deletions. It can be argued that gene fusions occur early, whereas deletions can be regarded as a late/secondary event.[19] Therefore, we can presume that *PAX5* fusion genes are driver genetic lesions, whereas *PAX5* deletions can be cooperative aberrancies, which require further alteration to determine the disease.

The final word on the real impact of *PAX5* alterations will come in the near future from robust functional studies, including the development of an appropriate *in vivo* leukemogenesis model.

REFERENCES

1 Fazio, G. *et al.* PAX5/ETV6 alters the gene expression profile of precursor B cells with opposite dominant effect on endogenous PAX5. *Leukemia* **27**, 992-995, doi:leu2012281 [pii]10.1038/leu.2012.281 (2013).

2 Coyaud, E. *et al.* PAX5-AUTS2 fusion resulting from t(7;9)(q11.2;p13.2) can now be classified as recurrent in B cell acute lymphoblastic leukemia. *Leuk Res* **34**, e323-325, doi:S0145-2126(10)00378-4 [pii]10.1016/j.leukres.2010.07.035 (2010).

3 Fazio, G., Biondi, A. & Cazzaniga, G. *The role of PAX5 in ALL*. 1st edn, (InTech Open Access Publisher, 2011).

4 Kawamata, N., Pennella, M. A., Woo, J. L., Berk, A. J. & Koeffler, H. P. Dominant-negative mechanism of leukemogenic PAX5 fusions. *Oncogene*, doi:onc2011291 [pii]10.1038/onc.2011.291 (2011).

5 Fortschegger, K., Anderl, S., Denk, D. & Strehl, S. Functional Heterogeneity of PAX5 Chimeras Reveals Insight for Leukemia Development. *Mol Cancer Res*, doi:1541-7786.MCR-13-0337 [pii]10.1158/1541-7786.MCR-13-0337 (2014).

6 Fazio, G., Palmi, C., Rolink, A., Biondi, A. & Cazzaniga, G. PAX5/TEL acts as a transcriptional repressor causing down-modulation of CD19, enhances migration to CXCL12, and confers survival advantage in pre-BI cells. *Cancer Res* **68**, 181-189 (2008).

7 Mullighan, C. G. *et al.* Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* **446**, 758-764 (2007).

8 Trageser, D. *et al.* Pre-B cell receptor-mediated cell cycle arrest in Philadelphia chromosome-positive acute lymphoblastic leukemia requires IKAROS function. *J Exp Med* **206**, 1739-1753 (2009).

9 Nutt, S. L., Heavey, B., Rolink, A. G. & Busslinger, M. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature* **401**, 556-562 (1999).

10 Sigvardsson, M. Transcription factor dose links development to disease. *Blood* **120**, 3630-3631, doi:120/18/3630 [pii]10.1182/blood-2012-09-455113 (2012).

11 Den Boer, M. L. *et al.* A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide classification study. *Lancet Oncol* **10**, 125-134, doi:S1470-2045(08)70339-5 [pii]10.1016/S1470-2045(08)70339-5 (2009).

12 Mullighan, C. G. *et al.* Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. *N Engl J Med* **360**, 470-480, doi:NEJMoa0808253 [pii]10.1056/NEJMoa0808253 (2009).

13 Mullighan, C. G. *et al.* Rearrangement of CRLF2 in B-progenitor- and Down syndrome-associated acute lymphoblastic leukemia. *Nat Genet* **41**, 1243-1246, doi:ng.469 [pii]10.1038/ng.469 (2009).

14 Mullighan, C. G. *et al.* Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science* **322**, 1377-1380, doi:322/5906/1377 [pii]10.1126/science.1164266 (2008).

15 Paulsson, K. *et al.* Microdeletions are a general feature of adult and adolescent acute lymphoblastic leukemia: Unexpected similarities with pediatric disease. *Proc Natl Acad Sci U S A* **105**, 6708-6713, doi:0800408105 [pii]10.1073/pnas.0800408105 (2008).
16 Familiades, J. *et al.* PAX5 mutations occur frequently in adult B-cell progenitor acute lymphoblastic leukemia and PAX5 haploinsufficiency is associated with BCR-ABL1 and TCF3-PBX1 fusion genes: a GRAALL study. *Leukemia* **23**, 1989-1998, doi:leu2009135 [pii]10.1038/leu.2009.135 (2009).
17 Iacobucci, I. *et al.* The PAX5 gene is frequently rearranged in BCR-ABL1-positive acute lymphoblastic leukemia but is not associated with outcome. A report on behalf of the GIMEMA Acute Leukemia Working Party. *Haematologica* **95**, 1683-1690, doi:haematol.2009.020792 [pii]10.3324/haematol.2009.020792 (2010).
18 Nebral, K. *et al.* Incidence and diversity of PAX5 fusion genes in childhood acute lymphoblastic leukemia. *Leukemia* **23**, 134-143 (2009).
19 Coyaud, E. *et al.* Wide diversity of PAX5 alterations in B-ALL: a Groupe Francophone de Cytogenetique Hematologique study. *Blood* **115**, 3089-3097, doi:blood-2009-07-234229 [pii]10.1182/blood-2009-07-234229 (2010).
20 Bardini, M. *et al.* DNA copy-number abnormalities do not occur in infant ALL with t(4;11)/MLL-AF4. *Leukemia* **24**, 169-176, doi:leu2009203 [pii]10.1038/leu.2009.203 (2010).

Appendix

Manuscripts not included in the thesis

Plasma cell-specific unfolded protein response molecules enable oncogenic signaling in pre-B acute lymphoblastic leukemia

Manuscript under revision

Behzad Kharabi Masouleh[1], Huimin Geng[1], Christian Hurtz[1], Misook Chang[3], Chuanxin Huang[4], Lai N. Chan[1], Srividya Swaminathan[1], Haibo Sun[5], Valeria Cazzaniga[1,6], H. Phillip Koeffler[5], Giovanni Cazzaniga[6], Ari Melnick[4], Elisabeth Paietta[7], Aaron C Logan[2], Laurie H. Glimcher[4], Markus Müschen[1,2]

[1]Department of Laboratory Medicine and [2]Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco CA 94143; [3]Children's Hospital Los Angeles, Los Angeles CA 90027; [4]Department of Medicine and Pharmacology, Weill Cornell Medical College, New York, NY 10065; [5]Cedars Sinai Medical Center, Los Angeles CA; [6]Centro Ricerca Tettamanti, Clinica Pediatrica, Università di Milano-Bicocca, Ospedale S. Gerardo, Monza, Italy; [7]Albert Einstein College of Medicine, Bronx, NY

Running title: Ectopic activation of UPR-molecules reveals novel vulnerability in pre-B ALL

Abbreviations: ALL, acute lymphoblastic leukemia; ChIP, chromatin immunoprecipitation; EV, empty vector; Ig, immunoglobulin; IL, interleukin; IM, Imatinib; n, denotes the number of independent experiments

For correspondence:
Markus Müschen,
Department of Laboratory Medicine
University of California San Francisco
521 Parnassus Ave
San Francisco CA 94143
E-mail: markus.muschen@ucsf.edu

Abstract

The unfolded protein response (UPR) pathway, a stress-induced signaling cascade emanating from the endoplasmic reticulum (ER), regulates the expression and activity of molecules including BiP (HSPA5), IRE1 (ERN1), (Blimp-1) PRDM1 and XBP1. These molecules are required for terminal differentiation of B cells into plasma cells and expressed at high levels in plasma cell-derived multiple myeloma. While these molecules have no known role at early stages of B-cell development, we here we show that their expression transiently peaks at the pre-B cell receptor checkpoint. Inducible, Cre-mediated deletion of *Hspa5*, *Prdm1* and *Xbp1* consistently induces cellular stress and cell death both in normal pre-B cells as well as in pre-B cell acute lymphoblastic leukemia (ALL) driven by BCR-ABL1- and NRASG12D oncogenes. Mechanistically, expression and activity of the UPR downstream effector XBP1 was positively regulated by STAT5 and negatively by the B cell-specific transcriptional repressors BACH2 and BCL6. In two clinical trials for children and adults with ALL, high *XBP1* mRNA levels at the time of diagnosis predicted poor outcome. A new small molecule inhibitor of ERN1-mediated XBP1-activation induced selective cell death of patient-derived pre-B ALL cells in vitro and significantly prolonged survival of transplant recipient mice in vivo. Collectively, these studies reveal that pre-B ALL cells are uniquely vulnerable to ER-stress and identify the UPR pathway and its downstream effector XBP1 as novel therapeutic targets to overcome drug-resistance in pre-B ALL.

ETV6-RUNX1 de-regulates the cytoskeleton and migration properties of B cell progenitor cells

Submitted

Chiara Palmi[1], Grazia Fazio[1], Angela M. Savino[1], Julia Procter[2], Louise Howell[2], Giulia Longinotti[1], Ilaria Brunati[1], Valentina Andrè[1], Valeria Cazzaniga[1], Pamela Della Mina[3], Antonello Villa[3], Mel Greaves[2], Andrea Biondi[1], Giovanna D'Amico[1], Anthony Ford[2] and Giovanni Cazzaniga[1].

[1]Centro Ricerca Tettamanti, Clinica Pediatrica, Università di Milano-Bicocca, Monza, Italy.
[2]Haemato-Oncology Research Unit, Division of Molecular Pathology, The Institute of Cancer Research, Sutton, Surrey SM2 5NG, UK.
[3]Microscopy and Image Analysis Consortium, Università di Milano-Bicocca, Monza, Italy.

Running title: ETV6-RUNX1 inhibits CXCL12 driven cell migration.

Keywords: ETV6-RUNX1, pediatric BCP-ALL, pre-leukemia, CXCL12, CXCR4

Corresponding author:
Andrea Biondi,
Clinica Pediatrica,
Università di Milano Bicocca,
Ospedale San Gerardo,
Via Pergolesi 33,
20900 Monza (MB), Italy.
E-mail: abiondi.unimib@gmail.com
Tel: +39 (0)39 233.3661
Fax: +39 (0)39 233.2167

Abstract

Although ETV6-RUNX1 fusion is a frequent initiating event in childhood leukemia**,** its role in leukemogenesis is still unknown. The main impact of the fusion itself is to generate and sustain a clone of clinically silent pre-leukemic B cell progenitor cells. Second hits, occurring even several years later, are required for overt disease. The understanding of the features and interactions of ETV6-RUNX1 positive cells during this "latency" period might help explaining how they can persist, and whether they could be prone to additional genetic changes. In this study, by using two murine models, we investigated whether ETV6-RUNX1 alters the cellular adhesion and migration properties of B cell progenitor cells. ETV6-RUNX1 expressing cells showed a significant defect in the chemotactic response to CXCL12, caused by a block in CXCR4 signaling, as demonstrated by inhibition of CXCL12-associated calcium flux and lack of ERK kinase phosphorylation. Moreover, the induction of ETV6-RUNX1 caused changes in the expression of cell-surface adhesion molecules. The expression of genes regulating the cytoskeleton was also affected, in particular those involved in the CDC42 pathway. The abnormalities described here could alter the interaction of ETV6-RUNX1 pre-leukemic B cell progenitor cells with the microenvironment and contribute to the pathogenesis of the disease. **Implications:** the cytoskeleton and migration abnormalities described here represent early events in the evolution of the disease, from the pre-leukemic phase to the clinical onset, with potential implications to develop strategies for effective eradication of leukemia.