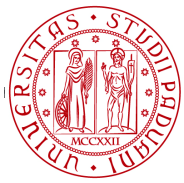


UNIVERSITÀ
DEGLI STUDI
DI PADOVA

TESI DI DOTTORATO

Università degli Studi di Padova

Dipartimento di Pediatria

SCUOLA DI DOTTORATO DI RICERCA IN

Medicina dello Sviluppo e Scienze della Programmazione

INDIRIZZO: Ematooncologia e Immunologia

CICLO XXIII°

Identification of new subgroups and prognostic markers in pediatric B cell precursor acute lymphoblastic leukemia by gene expression profiling

Direttore della Scuola : Ch.mo Prof. Giuseppe Basso

Coordinatore d'indirizzo: Ch.mo Prof. Giuseppe Basso

Supervisore :Ch.ma Prof.ssa Geertruy te Kronnie

Dottoranda : Elena Vendramini

SUMMARY

Treatment of pediatric acute lymphoblastic leukemia (ALL) is increasingly successful, achieving cure rates of over 80%. Early identification of patients with high risk for relapse has led to improved outcome, however, two third of patients encountering relapse were initially stratified into low-intermediate risk groups. The identification of better upfront prognostic factors remains an important challenge in childhood ALL.

In this thesis, gene expression profiling (GEP) was applied to different research approaches aiming to dissect subgroups and to find novel therapeutic targets in B cell precursor ALL (BCP ALL).

Among BCP ALL, the patients lacking major genomic aberrations (B-others) represent the subgroup that is most in need of in depth investigations in order to indentify new prognostic factors and improve of risk stratification. To advance biological knowledge of B-others we performed an integrated study of gene and non coding RNAs expression and genetic aberrancies.

Chapter 1 reports a study on profiling by gene expression arrays of 145 Italian B-others BCP ALL patients and in a representative subcohort of patients microRNAs (miRNAs) expression profiling and genome-wide DNA copy number variation analysis. In this study we found that 25% of Italian B-others patients fits in a group with unique signature and is associated to a favourable outcome. MicroRNAs expression profiling of this group revealed an unique miRNAs signature characterized by over expression of hsa-miR-125b, -125b-2*, -99a, -100, -125a-3p and has-miR-491-5p. Over expression of cluster miR-125b-2 in region 21q21.1 goes along with over expression of genes in the same chromosomal region. Genome-wide analysis excluded copy number alteration of the 21q21.1 region.

The frequent involvement of human chromosome 21 (Hsa21) aberrations in ALL (e.g. hyperdiploidy (HD), t(12;21) or iAmp21) and the involvement of the 21q21.1 region suggest a direct and functional contribution of Hsa21 genes to the malignant transformation of hematopoietic cells. There for there is high interest in studying ALL in children with Down Syndrome (DS), where trisomy 21 is constitutional and where the incidence of ALL is approximately 20-fold higher than in the general population.

In Chapter 2 is presented a study of genomic analysis of a large group of DS ALLs that characterizes molecular abnormalities specific of this ALL group.

Gene expression analysis revealed that DS ALL is a highly heterogeneous disease not definable as a unique ALL subtype with an enrichment of DNA damage and BCL6 responsive genes suggesting B-cell lymphocytic genomic instability. Surprisingly, only a single Hsa21 gene, *SON* was included in the DS ALL signature and it was only slightly upregulated. Furthermore gene expression data suggested that DS ALL and HD ALL are very different leukemias, reflecting the fundamental differences between constitutional and acquired trisomy such as the developmental stage in which the trisomy occurs and the fact that a constitutional trisomy is present both in the leukemia cells and in their microenvironment.

The study further revealed that 62% of the DS ALL samples were characterized by aberrant expression of the type I cytokine receptor *CRLF2*. Two kind of aberrations involving *CRLF2* were identified: a cryptic translocation involving *IGH@* and *CRLF2* in the pseudoautosomal region (PAR1) of the sex chromosomes and a deletion within PAR1. This aberration resulted in the *P2RY8-CRLF2* fusion and leads to overexpression of *CRLF2*. Furthermore a novel activating somatic mutation, F232C, in *CRLF2* was identified. We demonstrated that CRLF2 and mutated JAK2 cooperate in conferring cytokine independent growth to pro-B cells suggesting that the DS ALL children with *CRLF2* aberrant expression may benefit from therapy blocking the CRLF2-JAK2 pathway.

Since *CRLF2* aberrations were found also among non DS patients, we further analyzed the incidence and prognostic impact of this potential new marker in BCP ALL Italian patients enrolled into the AIEOP-BFM ALL2000 study.

Chapter 3, presents the study of a representative cohort of 464 non DS BCP ALL patients that was analyzed for the expression levels of *CRLF2* and for the occurrence of *CRLF2* rearrangements.

In this study we found that the *P2RY8-CRLF2* rearrangements in association with 20 times over expression of *CRLF2* identifies BCP ALL patients with a very poor prognosis and, among them, an important subset of patients currently stratified in the intermediate risk need to be considered for treatment adaptation.

Investigating the pathways highlighted by GEP analysis and testing drug effects require a substantial availability of leukemia samples. Primary ALL samples are difficult to culture *in vitro* and currently available cell lines poorly reflect the heterogeneous nature of the disease. Mouse xenotransplantation models are therefore widely used for *in vivo* testing and to amplify the number of leukemia cells to be used for various analyses.

In Chapter 4 study we assessed the capability of xenografted samples to recapitulate their respective primary leukemia, and we investigated whether the murine microenvironment selects for leukemia initiating cells leading to a bulk tumor markedly different from the diagnostic patient sample. We analysed the gene expression profiles of 7 primary paediatric ALL samples at diagnosis as well as of their respective xenograft leukaemia samples after serial primary, secondary and tertiary passages in the NOD/SCID/huALL transplant model.
In this study we demonstrated that the NOD/SCID/huALL transplant model recapitulates the primary human leukaemia, mimics the features of the primary malignancy and retains these characteristics over serial passages without selection for a subclone of the initial leukaemia.
Chapter 5 reports on a study that investigated engraftment properties of 50 pediatric ALL samples transplanted into NOD/SCID mice. Time to leukemia (TTL) was determined for each patient sample engrafted as weeks from transplant to overt leukemia.
The study shows that short TTL was strongly associated with high risk for early relapse, identifying a new independent prognostic factor. The high risk phenotype is reflected by a gene signature that identified patients with early relapse in an independent patient cohort. Gene expression profiling revealed a set of genes associated with this aggressive phenotype providing a potential strategy to identify these high-risk patients. Most importantly, pathways involving mTOR regulating cell growth were identified, providing targets for alternative therapeutic strategies for these high risk patients.

Concluding, ten years after its introduction in oncohematology, GEP constitutes to be a valuable research tool, efficacious in subtype discovery, biomarkers identification and discoveries of deregulated molecular pathways.

RIASSUNTO

La cura della leucemia linfoblastica acuta (ALL) sta migliorando con successo, raggiungendo un tasso di guarigione che va oltre l'80%. L'identificazione precoce dei pazienti con alto rischio di ricaduta ha portato ad un miglioramento generale dell'*outcome*, tuttavia, due terzi dei pazienti che incorrono nell'evento di ricaduta vengono inizialmente stratificati in gruppi a basso rischio o rischio intermedio. L'identificazione di migliori fattori prognostici rimane un'importante sfida nelle ALL pediatriche.

In questa tesi, lo studio del profilo di espressione genica è stato applicato a diversi approcci di ricerca, con lo scopo di individuare sottogruppi e trovare nuovi target terapeutici nelle ALL a cellule precursori B (BCP ALL)

Tra le BCP ALL, i pazienti privi delle aberrazioni genomiche più riccorrenti (*B-others*) rappresentano il sottogruppo che più necessita di studi approfonditi, tesi ad identificare nuovi fattori prognostici e migliorare la loro stratificazione nelle classi di rischio. Per aumentare le conoscenze biologiche riferite al gruppo dei *B-others*, è stato eseguito uno studio integrato di espressione genica, espressione di *non coding* RNAs e analisi delle aberrazioni genetiche.

Il Capitolo 1 riporta lo studio mediante *microarrays* di espressione genica di 145 pazienti Italiani affetti da BCP ALL e, in una sotto-coorte rappresentativa, lo studio dell'espressione dei microRNAs (miRNAs) e l'analisi di variazione di DNA *copy number* estesa all'intero genoma.

Da questo studio è emerso che il 25% dei pazienti Italiani di tipo *B-others* rientrano in un gruppo con una *signature* specifica e sono associati ad un *outcome* favorevole. Lo studio del profilo di espressione dei miRNAs rivela in questo gruppo una specifica *signature* di miRNAs caratterizzata dalla sovra espressione di hsa-miR-125b, -125b-2*, -99a, -100, -125a-3p e has-miR-491-5p. La sovra espressione del *cluster* miR-125b-2 nella regione 21q21.1 è accompagnata dalla concomitante sovra espressione dei geni nella stessa regione cromosomica. Le analisi sul genoma hanno portato ad escludere la presenza di alterazioni di DNA *copy number* nella regione 21q21.1.

Il frequente coinvolgimento di aberrazioni a carico del cromosoma 21 nelle ALL (come nel caso di iperdiploidia (HD), t(12;21) o iAmp21) e il coinvolgimento della regione 21q21.1, suggeriscono un diretto e funzionale contributo dei geni nel cromosoma 21 alla trasformazione maligna delle cellule ematopoietiche. A questo proposito c'è un grande

interesse nello studio delle ALL nei bambini affetti dalla Sindrome di Down (DS), nei quali la trisomia 21 è costituzionale e per i quali l'incidenza di ALL è approssimativamente 20 volte maggiore che nel resto della popolazione.

Nel Capitolo 2 viene presentato uno studio di analisi genomica di un grande gruppo di DS ALL che mira a caratterizzare le anomalie molecolari specifiche di questo gruppo di ALL.

L'analisi di espressione genica ha rivelato che le DS ALL sono leucemie molto eterogenee, non definibili come un unico sottotipo di ALL, con un arricchimento di geni rispondenti al *signalling* di BCL6 e di risposta al danno al DNA, che suggerisce un'instabilità genomica dei linfociti B. Sorprendentemente, solamente un gene appartenente al cromosoma 21, *SON*, è compreso nella *signature* delle DS ALL e risulta solo debolmente *up*-regolato. Inoltre, i dati di espressione genica suggeriscono che le DS ALL e le HD ALL sono leucemie molto diverse, riflettendo le differenze fondamentali tra trisomia costituzionale e acquisita, quali lo stadio di sviluppo nel quale la leucemia insorge e il fatto che la trisomia costituzionale è presente sia nelle cellule leucemiche che nel microambiente.

Lo studio ha inoltre rilevato che il 62% delle DS ALL sono caratterizzate da un'aberrante espressione del recettore per le citochine di tipo I *CRLF2*. Due tipi di aberrazioni che coinvolgono *CRLF2* sono state identificate: una traslocazione criptica che coinvolge il locus *IGH@* e *CRLF2* nella regione pseudoautosomale PAR1 dei cromosomi sessuali e una delezione in PAR1. Queste aberrazioni danno luogo alla formazione del trascritto di fusione *P2RY8-CRLF2* che determina la sovra espressione di *CRLF2*. Inoltre una nuova mutazione somatica attivante, F232C, in *CRLF2* è stata identificata. E' stato dimostrato che CRLF2 e JAK2 mutato cooperano nel conferire capacità di crescita indipendente da citochine a cellule pro-B suggerendo che i bambini affetti da DS e ALL con un'espressione aberrante di *CRLF2* possono trarre beneficio da terapie mirate a bloccare il *pathway* di CRLF2-JAK.

Dal momento che le aberrazioni a carico di *CRLF2* sono state trovate anche tra i pazienti non affetti dalla Sindrome di Down, è stata analizzata l'incidenza e l'impatto prognostico di questo potenziale nuovo marcatore nei pazienti Italiani con BCP ALL arruolati nello studio AIEOP-BFM ALL2000.

Il Capitolo 3 presenta lo studio di una coorte rappresentativa di 464 pazienti con BCP ALL non affetti da DS che è stata analizzata per l'espressione di *CRLF2* e per la presenza di riarrangiamenti a carico di *CRLF2*.

Da questo studio è emerso che il riarrangiamento *P2RY8-CRLF2* in associazione con la sovra espressione di *CRLF2* (di almeno 20 volte maggiore che nel resto della coorte), identifica

pazienti con una prognosi molto sfavorevole e, tra essi, un inportante sottogruppo di pazienti attualmente stratificati nella classe di rischio intermedia e che necessitano di essere considerati per un adeguamento della terapia.

Per investigare i *pathways* emersi dalle analisi di espressione genica e per testare l'effetto dei farmaci è necessaria una grande disponibilità di cellule leucemiche. Le cellule da leucemia primaria sono difficili da coltivare *in vitro* e le linee cellulari attualmente disponibili non riescono a riflettere la natura eterogenea della malattia. Per questo motivo i modelli di xenotrapianto in topo sono ampiamente usati sia per lo studio *in vivo* che per amplificare il numero di cellule leucemiche da usare nelle varie analisi.

Nello studio riportato nel Capitolo 4 è stata verificata la capacità delle cellule leucemiche ottenute da xenotrapianto di ricapitolare la loro rispettiva leucemia primaria ed è stata valutata la possibilità di una selezione da parte del microambiente murino per particolari cellule "inizianti" la leucemia che portino ad una massa tumorale marcatamente diversa da quella dei pazienti alla diagnosi. E'stato analizzato il profilo di espressione genica di 7 ALL primarie pediatriche alla diagnosi e le rispettive cellule leucemiche ottenute da xenotrapianto dopo un primo, un secondo ed un terzo passaggio seriale nel modello di trapianto di leucemia umana in topo NOD/SCID/huALL.

In questo studio è stato dimostrato che il modello di trapianto NOD/SCID/huALL ricapitola la leucemia primaria umana, mima le caratteristiche del tumore primario e ne trattiene le caratteristiche durante i passaggi seriali senza selezionare per un sottoclone della leucemia primaria iniziale.

Il Capitolo 5 riporta uno studio che ha investigato le proprietà di attecchimento di 50 ALL pediatriche trapiantate in topi NOD/SCID. Il tempo di attecchimento (*Time To Leukemia* – TTL) è stato determinato per ogni campione attecchito in termini di settimane trascorse dal trapianto alla manifestazione della leucemia.

Lo studio ha mostrato che un breve TTL è fortemente associato con un alto rischio di ricaduta precoce, costituendo di fatto un nuovo marcatore prognostico indipendente. Il fenotipo di alto rischio è riflesso in una *signature* in grado di identificare pazienti incorsi precocemente nell'evento di ricaduta in una coorte di pazienti indipendente. Lo studio di espressione genica rivela una serie di geni associati con questo fenotipo aggressivo, mettendo a diposizione una potenziale strategia per identificare i pazienti ad alto rischio. In modo ancora più importante, *pathways* che regolano la crescita cellulare e che coinvolgono mTOR sono stati identificati,

indicando dei *target* per strategie terapeutiche alternative per i pazienti ad alto rischio di riaduta.

Concludendo, dieci anni dopo la sua introduzione in oncoematologia, lo studio del profilo di espressione genica si conferma essere un valido strumento di ricerca, efficace nella scoperta di nuovi sottotipi, nell'individuazione di biomarcatori e nel portare alla luce *pathways* molecolari deregolati.

INDEX

INTRODUCTION

ACUTE LYMPHOBLASTIC LEUKEMIA

Acute lymphoblastic leukaemia (ALL) is a clonal malignant disease of the bone marrow, that originates from lymphoid progenitor cells often carrying specific genetic and epigenetic alterations and characterized by an accumulation of blast cells resulting in the suppression of normal haematopoiesis and infiltration of various extramedullary sites. ALL is the most common pediatric cancer, accounting for a quarter of all malignancies among children aged < 15 years. This potentially catastrophic disease was once fatal in four-fifths of patients, but thanks to the introductiuon of risk-directed treatment and improved supportive care clinical outcome has improved remarkably over the past 50 years. Today the 5-year event-free survival rate for ALL is approximately 80% in children receiving protocol-based therapy in the developed countries[1].

B CELL PRECURSOR ALL

The majority of ALLs are of B-cell lineage in both children (>80%) and adults (>75%)[2, 3]. In B cell precursor ALL (BCP ALL) blast cells are regarded as malignant counterparts of normal B-cell precursors.

Current approaches of risk assessment rely on a number of key clinical and laboratory findings such as the initial leukocyte count, age at diagnosis and early treatment response. Children aged 1–9 years have a better outcome than infants and adolescents. Leukocyte count is a continuous variable, with decreasing counts conferring a better outcome[4].

The most informative prognostic factor is the response to early treatment, as determined by measurements of the rate of clearance of leukemic cells from the blood or bone marrow. This estimate of minimal residual disease accounts for the drug sensitivity or resistance of leukemic cells and the pharmacodynamics of the drugs, which is affected by the pharmacogenetics of the host[5]. Flow cytometric profiling of aberrant immunophenotype markers and polymerase chain reaction amplification of fusion transcripts or antigen-receptor genes, which are at least 100-fold more sensitive than conventional morphological determinants, have allowed minimal residual disease to be detected at very low levels (< 0.01%). This provides a useful means to identify patients at very low or high risk of relapse[6].

Minimal residual disease can be measured by the current techniques in nearly all patients, and has become a key factor for risk stratification in childhood ALL.

In spite of treatment and risk stratification improvement, 20% of children with ALL ultimately relapse, and cure rate after relapse reaches only 25% to 40%. Very high remains the interest to understand the mechanisms of relapse. Genome-wide studies using matched diagnosis and relapse samples from the same patients have shown that both sets of samples are clonally related, and that the relapse clones are often present as minor populations at diagnosis, which suggests that they are selected during treatment. Indeed, many of the genetic alterations that emerge in the dominant clone at relapse involve genes that have been implicated in treatment resistance (e.g. *CDKN2A/B* or *IKZF1*)[7, 8] and gene expression studies have identified a proliferative gene signature that emerges at relapse with consistent up-regulation of genes, such as survivin, that could provide useful targets for novel therapeutic intervention[9].

GENETIC ABNORMALITIES IN BCP ALL

Childhood acute leukemias have long been the best characterized malignancies from a genetic viewpoint. In BCP ALL, individual chromosomal abnormalities are strong independent indicators of outcome, especially to indicate risk of relapse[10].

The most significant impact for risk stratification for treatment are t(9;22)(q34;q11)/*BCR-ABL1* and rearrangements of the *MLL* gene. In particular this applies to t(4;11)(q21;q23)/*MLL-AFF1* (previously known as *MLL-AF4*) The prognosis of the other *MLL* partners may become significant in the future, particularly among infants[11]. The detection of these two abnormalities provides the basic criteria for the classification of high risk groups, which is applied in all American and European protocols.

Other significant structural abnormalities include t(12;21)(p13;q22)/*ETV6-RUNX1* fusion, as well as t(1;19)(q23;p13.3)/*TCF3-PBX1* fusion. However, these are not used in risk stratification on all protocols. The *ETV6-RUNX1* fusion occurs in approximately 25% of younger children with BCP ALL and these patients have an extremely good prognosis. Among patients with *TCF3* rearrangements, those with *TCF3- PBX1* were originally regarded as poor risk on some treatment protocols, but on modern therapy they are classified as standard risk[12, 13]. In contrast the rare variant, t(17;19)(q22;p13)/*HLF-TCF3* fusion, has a dismal outcome on all therapies[14].

Intrachromosomal amplification of chromosome 21 (iAMP21) consists of an abnormal chromosome 21 with highly variable morphology between patients and with a common amplified region that in all cases includes the *RUNX1* gene[15, 16, 17]. This abnormality was originally described as poor risk factor[18, 19, 20, 21], although the outcome has since been shown to be protocol dependent[22].

Numerical cytogenetic abnormalities include: high hyperdiploidy (51–65 chromosomes)[23], near-haploidy (24–29 chromosomes) and low hypodiploidy (31–39 chromosomes)[24, 25]. High hyperdiploidy accounts for approximately 30% of childhood BCP ALL and is characterized by the gain of specific chromosomes. It is associated with a good prognosis in children. Near-haploidy and low hypodiploidy are rare, comprising <1% each of childhood ALL. Their characteristic features are the gain of specific chromosomes onto the haploid chromosome set and, in the majority of patients, the presence of a population of cells with an exact doubling of this chromosome number[26]. Both are linked to a poor outcome and are used to stratify patients as high risk.

Translocations involving *IGH@* at 14q32 are emerging as a significant subgroup in childhood ALL[27, 28, 29, 30]. It is of interest that they occur more frequently in adolescents and, although numbers are small, they appear to have an inferior outcome.

SUBMICROSCOPIC LESION IN BCP ALL

Recurring genomic aberrations that define leukemia subtype are important in leukaemia initiation[31], but alone are insufficient to generate a full leukaemic phenotype, indicating that cooperating oncogenic lesions are present in leukaemia speciments. Although additional mutations have been identified in a subset of cases, the full complement of cooperating lesions and their distribution within the known genetic subtypes of ALL remain to be defined. Recently Mullighan et al. published a comprehensive study of copy number change and loss of heterozigosity on more then 200 ALL pediatric patients[32].

Genomic studies identified a mean of 6.46 somatic copy number alterations per case, with deletions two times more frequent than amplifications. The frequency of genomic deletions and amplifications varied significantly between ALL subtypes. Genomic gains were frequent in B-progenitor ALLs with high hyperdiploidy, while being rare in other B-progenitor ALL subtypes. Deletions were more frequent, ranging from six deletions per case in some genetic subtypes (*ETV6–RUNX1* and hypodiploid ALL) to only a single deletion per case in *MLL*-

rearranged ALL. In this study global genomic instability and high mutational rate were excluded as underlying mechanism in ALL.

The most notable observation was the identification of genomic alterations in genes that regulate B-lymphocyte differentiation in 40% of B-progenitor ALL cases. The most common targets of these genetic alterations were *EBF1*, *PAX5* and *IKZF1* that have central roles in the development of normal B cells[33]. Recurring deletions included 3p14.2 (*FHIT*)[34], 6q16.2-3 (including *CCNC*)[35], 9p21.3 (two regions involving *CDKN2A*[36] and *MLLT3*), 12p13.2 (*ETV6*)[37], 11q23 (including *ATM*)[38], 13q14.2 (*RB1*)[39] and 13q14.2-3 (including mir-16-1 and mir-15a)[40]. New identified deletion included *IKZF3*, *LEF1*, *TCF3* and *BLNK*, encoding regulators of B-cell development. Other deleted loci were *BTG1*, recently associated with glucocorticoids resistance *in vitro*[41], and *ERG*, reported to occur exclusively in a subgroup of BCP ALL associate to a favourable outcome[42].

Recently, a cryptic translocation, t(X;14)(p22;q32) or t(Y;14)(p11;q32), involving *IGH@* and *CRLF2* in the pseudoautosomal region (PAR1) of the sex chromosomes, and a deletion within PAR1, giving rise to the *P2RY8-CRLF2* fusion, have been reported[43, 44, 45, 46]. These genomic alterations lead to overexpression of *CRLF2*, which has been defined as a novel, significant abnormality in BCP ALL. *CRLF2* alterations, including activating mutations of the CRLF2 receptor itself, are associated with activating JAK mutations resulting in constitutive activation of the JAK-STAT signalling pathway[47, 48, 49]. Activation of this pathway has been associated with a worse prognosis in adults and children with BCP ALL[50, 51] and has been highlighted as an important candidate pathway for targeted therapy.

Improvement in the genetic lesion detection has led to the definition of new ALL subtype characterized by new recurrent aberrations, such as patients with iAmp21, *ERG* alteration or *CRLF2* alterations (Figure1). Also these subtypes are currently under investigation, to elucidated their pathobiology and association with outcome. These new recurrent aberrations reduce the group of genetically undefined patients, called B-others, that few year ago counted for 22% of ALL[52] and now are restricted to 7%[53]. Further dissection of B-others group are expected from the applications of new technologies such as next-generation high-throughput sequencing[54].

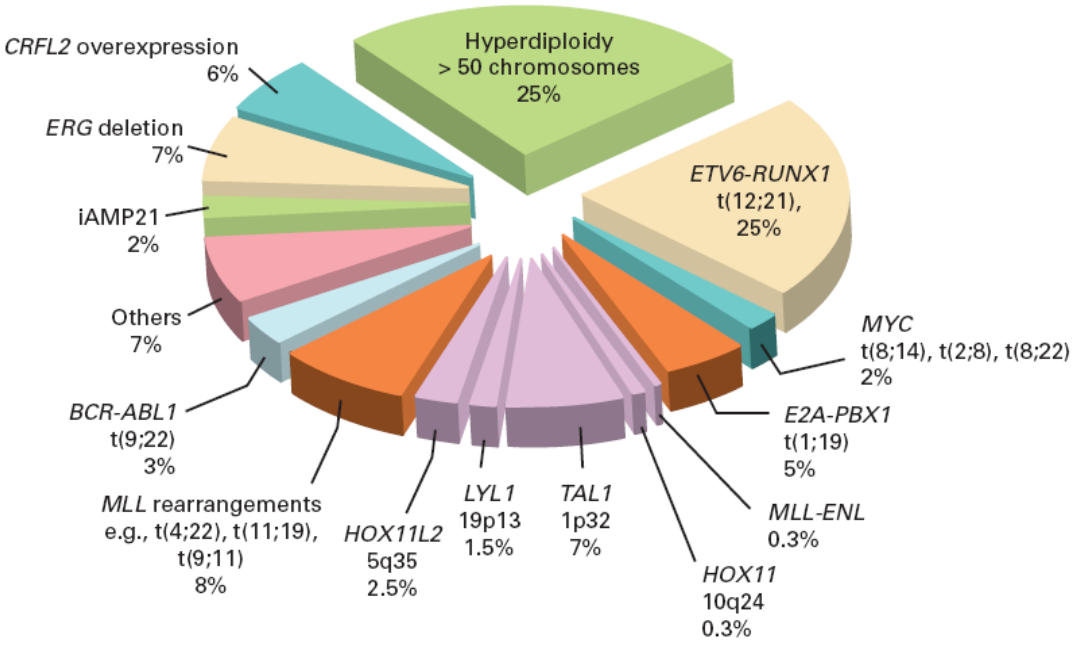


Figure 1. Estimated frequency of genetic abnormalities in ALL. Violet area are referred to T ALL abnormalities, other colours are referred to B ALL abnormalities (From Pui et al., 2011)[55].

THYMIC STROMAL LYMPHOPOIETIN AND CRLF2

Thymic stromal lymphopoietin (TSLP) is an epithelially derived cytokine initially identified in 1994 as a bioactive factor secreted in the supernatants of a murine thymic stromal cell line[56]. TSLP is a four-helix bundle type 1 cytokine, and is closely related to IL-7, with which it shares an overlapping, but distinct, biological profile. TSLP exert its biological activities by binding to a heterodimeric receptor that consists of the IL-7 receptor α-chain (IL-7Rα) and the TSLP receptor chain (TSLPR, also known as cytokine receptor like factor 2, CRLF2), which is closely related to the common receptor-γ chain (γc). CRLF2 has low affinity for TSLP, but in combination with IL-7Rα generates a high affinity binding site for TSLP and triggers signalling[57] (Figure 2).

TSLP mRNA is expressed predominantly by epithelial cells in the thymus, lung, skin, intestine and tonsils as well as stromal cells and mast cells, but is not found in most hematopoietic cell types and endothelial cells[58, 59]. In contrast to the restricted expression of *TSLP*, *CRLF2* is more widely distributed. *CRLF2* mRNA has been detected on many immune cell types, including dendritic cells (DCs), T cells, B cells, mast cells, NKT cells and monocytes as well as in tissues from heart, skeletal muscle, kidney and liver. This suggests that TSLP can function on a broad range of cell types.

TSLP exerts profound influence on the polarization of dendritic cells (DCs) to drive T helper (Th) 2 cytokine production. It also directly promotes T cell proliferation in response to T cell receptor (TCR) activation, and Th2 cytokine production. TSLP also supports B cell expansion and differentiation. TSLP further amplifies Th2 cytokine production by mast cells and NKT cells[60]. TSLP plays essential roles in allergic/inflammatory skin and airway disorders, in helminth infections, and in regulating intestinal immunity[61].

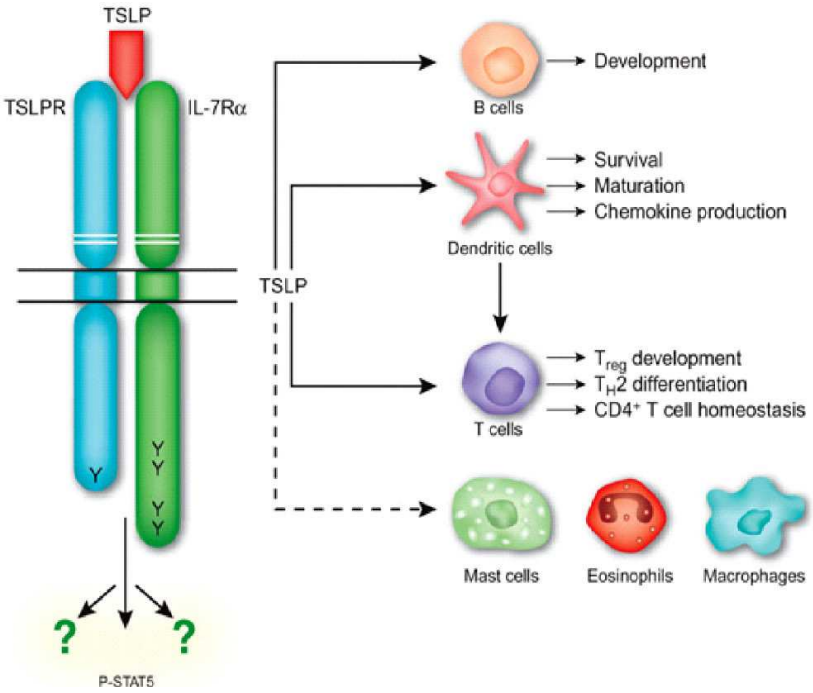


Figure 2. Structure of the heterodimeric TSLP receptor and the its cellular targets immune cells. TSLPR signals via activation and phosphorylation of STAT5 (P-STAT5) and other unidentified pathways (from He and Geha, 2010)[62].

A recent work on primary CD4+ T cells established the critical roles of JAK1 and JAK2 for TSLP-mediated STAT5 activation. The phosphorylation of JAK1 and JAK2 by TSLP correlates with the ability of these JAKs to associate with IL-7Rα and CRLF2, respectively, mediating the docking of STAT5 to IL-7Rα (but not CRLF2) within the receptor complex and its phosphorylation (Figure 3). TSLP, to date, is the only cytokine that uses the combination of JAK1 and JAK2 to principally activate STAT5 proteins[63].

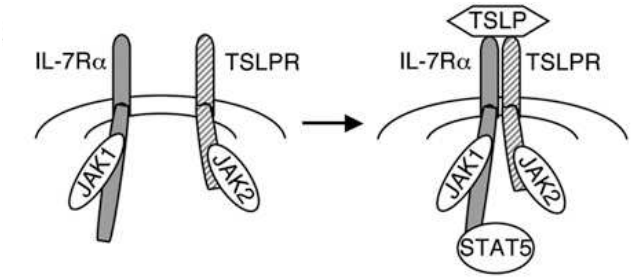


Figure 3. Schematic model for JAK-STAT signaling by TSLP receptors. TSLP induces heterodimerization of CRLF2 with IL-7Rα leading to activation of JAK1, JAK2, and STAT5 (from Rochmana et al., 2010)[64].

STAT5 AND JAK-STAT PATHWAY

Upon Jak-mediated phosphorylation, the STAT (signal transducer and activator of transcription) transcription factors dimerize and translocate to the nucleus where they regulate transcription by binding to target genes. The STAT5 transcription factor consists of two highly related isoforms, STAT5A and STAT5B, which are encoded by separate genes. STAT5A and STAT5B have distinct physiological functions, as inactivation of Stat5a interferes with terminal differentiation of the mammary gland, whereas deletion of Stat5b results in impaired body growth. However within the lymphoid system, the two STAT5 proteins fulfil largely redundant roles, as the individual loss of STAT5A or STAT5B has only minor consequences for lymphocyte function[65].

When both STAT5 proteins are depleted, in the bone marrow of conditional Stat5–/– mice, B cell development is arrested at the uncommitted pre–pro-B cell stage, indicating that STAT5 is an essential mediator in early B cell development[66, 67]. Gain-of-function experiments have implicated STAT5 in activating the expression of the germinal center (GC) B cell regulator Bcl6 and in controlling the differentiation and self-renewal potential of human memory B cells[68]. Late B cell development, memory B cell formation, and plasma cell differentiation appear largely unaffected in the absence of STAT5[69].

STAT5 has also implicated in the pathogenesis of human B cell precursor acute lymphoblastic leukemia (B-ALL). Mutations in either of the two STAT5 genes have, however, not be found in human B-ALL[70]. In contrast, distinct subsets of B-ALL contain activating mutations in upstream signaling components of the JAK-STAT pathway that confer constitutive STAT5 activation.

The BCR–ABL1 kinase (tyrosin kinase constitutively active in t(9;22) ALL) leads to constitutive activation of STAT5 and thus likely contributes to leukemogenesis by signaling

via the JAK-STAT pathway[71, 72]. Indeed, the BCR–ABL1 kinase is unable to induce leukemia in Stat5–/– mice, strongly arguing for an important role of STAT5 in the initiation of B-ALL[73]. The constitutive activation of the JAK-STAT pathway likely activates the antiapoptotic *Mcl-1* gene, which renders B-ALL cells growth-factor-independent by maintaining cell survival in the absence of local cytokines. The STAT5-dependent expression of *Mcl-1* may also explain why STAT5 is continuously required for maintaining the leukemic state, as conditional Stat5 inactivation results in cell cycle arrest and apoptosis of malignant B cells in a BCR–ABL1-dependent mouse leukemia model[74].

The recent discovery of predominantly JAK2 and few JAK1 mutations in 20% of Down syndrome-associated ALL (DS ALL)[75, 76] and 10% of high-risk pediatric B-ALL[77] has demonstrated that direct deregulation of the JAK-STAT pathway can also cause B cell leukemia. Although most mutations are found in the pseudokinase or kinase domain of JAK2, expression of mutant JAK2 alone is insufficient to activate STAT5 and thus to transform cytokine-dependent cell lines to growth-factor independence[78] . Moreover, JAK2 is activated by signaling through common b chain-containing or gp130-containing cytokine receptors, which are not normally associated with lymphocyte development[79]. An explanation was recently provided by the important discovery that the cytokine receptor gene *CRLF2* functions as a proto-oncogene in 15% of adult and pediatric B-ALL. Notably, *BCR-ABL1*-positive and *CRLF2*-overexpressing B-ALL share a common transcription profile with an enrichment of the STAT expression signature, implying that *CRLF2* overexpression supports aberrant JAK-STAT activation in B-ALL. Importantly, all B-ALL cases with JAK2 mutations overexpress *CRLF2* suggesting that the CRLF2 protein functions an essential cytokine receptor scaffold for signaling of the mutant JAK2 proteins. Indeed, overexpression of normal CRLF2 together with mutant JAK2 leads to STAT5 activation and cytokine-independent growth of hematopoietic progenitor cell lines[80] .

CHROMOSOME 21 IN BCP ALL

Chromosome 21 anomalies play an important role in tumor development as acquired somatic mutations[81]. Twenty-three % of all chromosomal abnormalities in ALL involve trisomies or tetrasomies of chromosome 21. Ninety percent of childhood ALL with chromosome numbers >50 show trisomy or tetrasomy 21, but also 30% of B-cell ALLs with a chromosome number of 47–50 present with > 2 chromosome 21 copies[82].

Other aberrations involving chromosome 21 in ALL are the t(12;21) and the iAMP21, an intrachromosomal amplification within the long arm of chromosome 21, both involving the *RUNX1* gene.

The *RUNX1* gene, located in 21q22, is involved in many hematopoietic neoplasms. *RUNX1* is a member of the core binding factor (CBF) family and encodes a subunit of a heterodimeric transcription factor. *RUNX1* is expressed in hematopoietic stem cells, as well as in differentiating myeloid and lymphoid cells[83]. The *RUNX1* gene is essential for normal hematopoiesis as shown by analysis of *RUNX1*-deficient mice[84]. *RUNX1* is often affected by translocations in acute leukemias, with about 40 translocation partners having been reported to date. However, also deletions, amplifications, as mentioned above, and point mutations are known to have a pathogenic effect in acute myeloid and acute lymphoblastic leukemia, in blast crisis of chronic myeloid leukemia, in MDS, and in T-cell lymphomas[85]. In exon 1, affected by the t(12;21) a high density of repetitive Alu sequences mediating recombinations is found[86]. Besides the somatic *RUNX1* mutations, germline mutations have been found in families with thrombocytopenia and a predisposition to leukemia development[87]. Constitutional deletions in 21q affecting the *RUNX1* gene have been demonstrated in patients with thrombocytopenia and also with developmental delay and mental retardation[88].

In Down Syndrome (DS), where trisomy 21 is constitutive, the risk of children to develop ALL is 20-fold increased[89].

In a recent work using the transchromosomic system (mouse embryonic stem cells bearing an extra human chromosome 21 (HSA21)) De Vita et al. demostrated that overdose of more than one HSA21 gene contributes to the disturbance of early haematopoiesis in DS, and that one of the contributors is RUNX1. As the observed T21-driven hyperproduction of multipotential immature precursors precedes the bifurcation to lymphoid and myeloid lineages, authors speculate that this could create conditions of increased chance for acquisition of pre-leukaemogenic rearrangements/mutations in both lymphoid and myeloid lineages during foetal haematopoiesis, contributing to the increased risk of both leukaemia types in DS[90].

GENE EXPRESSION PROFILING

Today approaches in biomedical research often employ high throughput technologies in comprehensive studies named genomics, proteomics, metabolomics (etc.) that give a broad overview of the investigated field. The "Omic" studies disclose molecular correlation networks that are the basis of normal and pathogenic cellular processes.

The power of genomics applying microarray technology lies in the numbers: tens of thousands of genes are analyzed simultaneously and without predetermined bias, novel genes involved in disease processes can be discovered[91]. This knowledge allows investigators to develop not only new diagnostic markers but also prognostic and disease progression markers, ultimately leading to the prospect of patient-tailored therapy.

MICROARRAY TECHNOLOGY

Microarrays consist of numerous regularly spaced DNA probes which are immobilized on a solid surface. The pool of transcripts in a given patient sample is labelled with a fluorescent dye and hybridized to the microarray. The fluorescent signal bound to the probe serves as an indicator of the expression of the corresponding transcript.

Several types of microarray platforms, commercialized by companies such as Affymetrixs, Agilents, Codelinks (GE Healthcare) to name but a few or produced by facilities of nonprofit organizations (e.g. by the Stanford Functional Genomics Facilities, Stanford University, CA, USA). The platforms differ in probe content, design and deposition technology, as well as labelling and hybridizing protocols[92]. Different probe types include cDNA sequences (usually several hundred base pairs in length), short (25–30 mer) DNA oligonucleotides (up to 16 per gene of interest) or longer (60–80 mer) DNA oligonucleotide probes[93]. The probes can be either contact-spotted using pins or deposited using ink-jet devices. Furthermore, oligonucleotides can be synthesized in situ on a quartz wafer (Affymetrix, High Wycombe, UK). Labeling methods vary from in vitro transcribed biotinylated cRNA (stained with streptavidin–phycoerythrin after hybridization) or directly labeled cDNA targets with incorporated fluorescently labeled deoxyribonucleotides. Finally, microarray procedures can be carried out in one color (one labeled target per array, such as used by Affymetrix) or two colors (two separately labeled targets for each array used, for instance, by Agilent Technologies, Palo Alto, CA, USA)[94].

Most laboratories use DNA-oligonucleotide microarrays of 25–60 nucleotide length as this minimizes the risk of cross-hybridization and guarantees a high level of specificity[95], [96]. Currently available commercial arrays have all reached whole transcriptome coverage.

Initially application of microarray technology has aimed at evaluation of expression of coding sequences of genes by assessment of the amount of RNA transcription (gene expression profiling, GEP). Subsequently their use has been extended to investigation of non coding sequences e.g. microRNAs expression (miRNAs profiling). Another widespread application was the quantification of gene dosage on the genomic DNA level[97] allowing to explore single nucleotide polymorphisms (SNPs) and copy number alterations (copy number variations, loss of heterozygosity (LOH), and copy number neutral LOH due to uniparental disomy UPD) in parallel[98]. Currently available arrays incorporate up to ~2.7 million SNP or copy number probes[99]. However, SNP analysis is not capable of directly identifying reciprocal translocations which commonly occur in AL (Acute Leukemia).

DATA ANALYSIS

For interpretation of the large amounts of data being generated by the analysis of the expression of thousands of genes, sophisticated computational methods are needed (Table 1).

Data preprocessing include image analysis, normalization and data transformation. Many image-processing approaches have been developed[100], among which the main differences relate to how spot segmentation — distinguishing foreground from background intensities — is carried out. Another important pre-processing step is normalization, which allows comparisons between microarray experiments and the control of extraneous variation among experiments. Several normalization approaches have been introduced[101], between them robust multi-array average (RMA) and its modification GCRMA, are often consider to perform better than other methods, although there is some controversy about which method is best[102]. RMA corrects arrays for background using a transformation, normalizes them using a formula that is based on a normal distribution, and uses a linear model to estimate expression values on a log scale.

Much appreciated for data visualization is PCA (Principal Component Analysis) that reduces the dimensionality of array data to facilitates visualization and interpretation of large data sets[103].

Classification algorithms are used either to discover new categories within a data set (class discovery; unsupervised classification) or assign cases to a given category (class prediction; supervised classification).

Cluster-analysis algorithms for unsupervised classification group objects on the basis of some sort of similarity metric that is computed for one or more 'features' or variables. Hierarchical cluster analysis graphically presents results in a tree diagram (dendrogram), and is probably the most common unsupervised classification algorithm in microarray analysis. Hierarchical clustering is further used for a reduction of complexity and the visualization of supervised analyses[104]. Non-hierarchical clustering methods divide the cases (samples or genes) into a predetermined number of groups in a manner that maximizes a specific function (for example, the ratio of variability between and within clusters). Cluster-analysis approaches entail making several choices, such as which metric to use to quantify the distance or similarity among pairs of objects, what criteria to optimize in determining the cluster solution, and how many clusters to include in the solution. No consensus or clear guidelines exist to guide these decisions. Cluster analysis always produces clustering, but whether a pattern observed in the sample data characterizes a pattern present in the population remains an open question[105].

Supervised classification (often called 'class assignment', 'prediction' or 'discrimination') entails developing algorithms to assign objects to a priori defined categories. Algorithms are typically developed and evaluated on a 'training' data set and an independent 'test' data set, respectively, in which the categories to which objects belong are known before they are used in practical applications. A common misconception is that the set of the most differentially expressed genes will necessarily give the best predictive accuracy. The gene list that is obtained from hypothesis testing does not necessarily give the best prediction. No one method for constructing prediction algorithms (e.g. support vector machines (SVM)[106], predictive analysis of microarrays (PAM)[107], and k-top scoring pairs (kTSP)[108]) is widely accepted as superior or optimal[109].

Inferential analyses are commonly used to identify differentially regulate genes among two or more groups, parametic and non-parametric tests are used such as t-test, analysis of variance (ANOVA), significance analysis of microarrays (SAM)[110].

A list of differentially expressed genes that constitute a signature profile are accompanied by a false discovery rate (FDR)[111] or a P-value corrected for multiple testing to give an idea of how significant the findings actually are. The robustness of the results should be tested by

means of bootstrapping or leave-one-out cross-validation tests. Concerning the individual genes, further validation experiments are mandatory.

Software	*URL*	*Software type*	*Computer platform*
Affymetrix MAS 5.0	http://www.affymetrix.com/index.affx	Commercial	Windows
Bioconductor (using R)	http://www.bioconductor.org/	Free	Windows, MacOS, Linux/Unix
Cluster	http://rana.lbl.gov/	Free	Windows
D Chip	http://www.dchip.org/	Free	Windows
Genespring	http://www.silicongenetics.com/cgi/SiG.cgi/index.smf	Commercial	Windows
GeneMaths	www.applied-maths.com	Commercial	Windows
Genlab (using Matlab)	http://www.genlab.tudelft.nl/	free	Windows, MacOS, Linux/Unix
GCOS, DMT	www.affymetrix.com/support/	Free	Windows
Ingenuity	www.ingenuity.com	commercial	Java application
LIB SVM	www.csie.ntu.edu.tw/~cjlin/libsvm/	Free	Linux/Unix
PAM	http://www-stat.stanford.edu/~tibs/PAM/	Free	Excel/R under Windows/Unix
Rosetta Resolver	http://www.rosettabio.com	Commercial	Windows
SAM	http://www-stat.stanford.edu/~tibs/SAM/	Free	Excel Add-in
TM4	http://www.tm4.org	Free	Windows
TreeView	http://rana.lbl.gov/	Free	Windows

URL: location in the WWW where more information can be found; software type: commercial or free software; Computer Platform: Operating System necessary.

Table 1. Examples of software used for biostatistical analysis of microarray data (from Staal et al., 2006)[112].

STANDARDIZATION OF MICROARRAY DATA

Completeness of data annotation and specification of data processing and analysis are requirements for repeatability of published microarray studies[113].

With so many variables in place and moreover thousands of raw data parameters to measure and analyze, it has been realized that standardization and quality guidelines are needed[114]. One of the first and very successful attempts has been the requirements put forward in MIAME (Minimum Information About a Microarray Experiment) by the Microarray Gene Expression Data Society[115]. MIAME mainly documents information about the experimental design, RNA extraction and labeling techniques as well as array type used. Many journals (all Nature Press journals including Leukemia) now require submission of microarray data to either of two databases emerging as the main public repositories: GEO (http:// www.ncbi.nlm.nih.gov/geo/) or ArrayExpress (http://www.ebi.ac.uk/arrayexpress). Recently, different reports have been issued about cross-platform comparisons. The current consensus indicates that for well-defined genes such comparisons are certainly possible provided standardized protocols are strictly adhered to[116]. Guidelines are also been published arising from experience of several study groups[117].

CLASSIFICATION IN CANCER AND LEUKEMIA

«*Here, a generic approach to cancer classification based on gene expression monitoring by DNA microarrays is described and applied to human acute leukemias as a test case. A class discovery procedure automatically discovered the distinction between acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL) without previous knowledge of these classes. An automatically derived class predictor was able to determine the class of new leukemia cases.*» [118]

Ten years ago, these phrases were written by Golub et al. to present at the scientific community the feasibility of cancer classification based solely on gene expression profiles independent of previous biological knowledge. This milestone work presented a class discovery method effective in the distinction between AML and ALL, as well as the distinction between B-cell and T-cell ALL. Few years later, Yeoh et al. demonstrated that distinct expression profiles identified each of the prognostically important leukemia subtypes, including T-ALL, *E2A-PBX1*, *BCR-ABL*, *TEL-AML1*, *MLL* rearrangement, and hyperdiploid >50 chromosomes. In addition, another ALL subgroup was identified based on its unique expression profile. Furthermore Yeoh et al reported that gene expression profiling provides a robust and accurate approach for the risk stratification of pediatric ALL patients[119]. Since than, in ALL of the B and T-lineage, several gene expression-based classifiers for the prediction of prognosis have been developed[120, 121, 122].

THE MILE STUDY

The International Microarray Innovations in Leukemia (MILE) [123] Study Group was formed in 2005 around the European LeukemiaNet (ELN) in 11 laboratories across three continents (seven from the ELN, three from the United States, and one in Singapore).

The collaborative MILE study program was designed to assess the clinical accuracy of gene expression profiles (compared with current routine diagnostic work-up) of 16 acute and chronic leukemia subclasses, myelodysplastic syndromes (MDSs), and a so-called “none of the target classes” control group that included nonmalignant disorders and normal bone marrow.

An exploratory retrospective stage I study was designed for biomarker discovery and generated whole-genome expression profiles from 2,143 patients with leukemias and myelodysplastic syndromes (Figure 4). The gene expression profiling–based diagnostic accuracy was further validated in a prospective second study stage of an independent cohort of 1,191 patients. The stage I study achieved 92.2% classification accuracy for all 18 distinct

classes investigated (median specificity of 99.7%) (Figure 5). In stage II, the observed accuracy of the classifier prediction across all 18 classes was 88.1%. The accuracy increased to 91.5% when focused on acute leukemias, representing all 14 distinct classes.

This is thus far the largest gene expression microarray profiling study in hematology and oncology, and clearly underlines the robust performance of this method and demonstrates the possibility of completely standardized laboratory procedures combined with sophisticated data algorithms.

MILE study provided the background for an integration of the microarray technology in routine diagnostics. The coming years thus should see a focus on the definition of clear criteria for the performance of GEP analysis in patients with haematological neoplasms and concentrate on the development of new diagnostic algorithms incorporating GEP and the clinical development of novel and more specific drugs. This will highly contribute to a standardization and optimization of therapeutic decisions in patients with haematological malignancies[124].

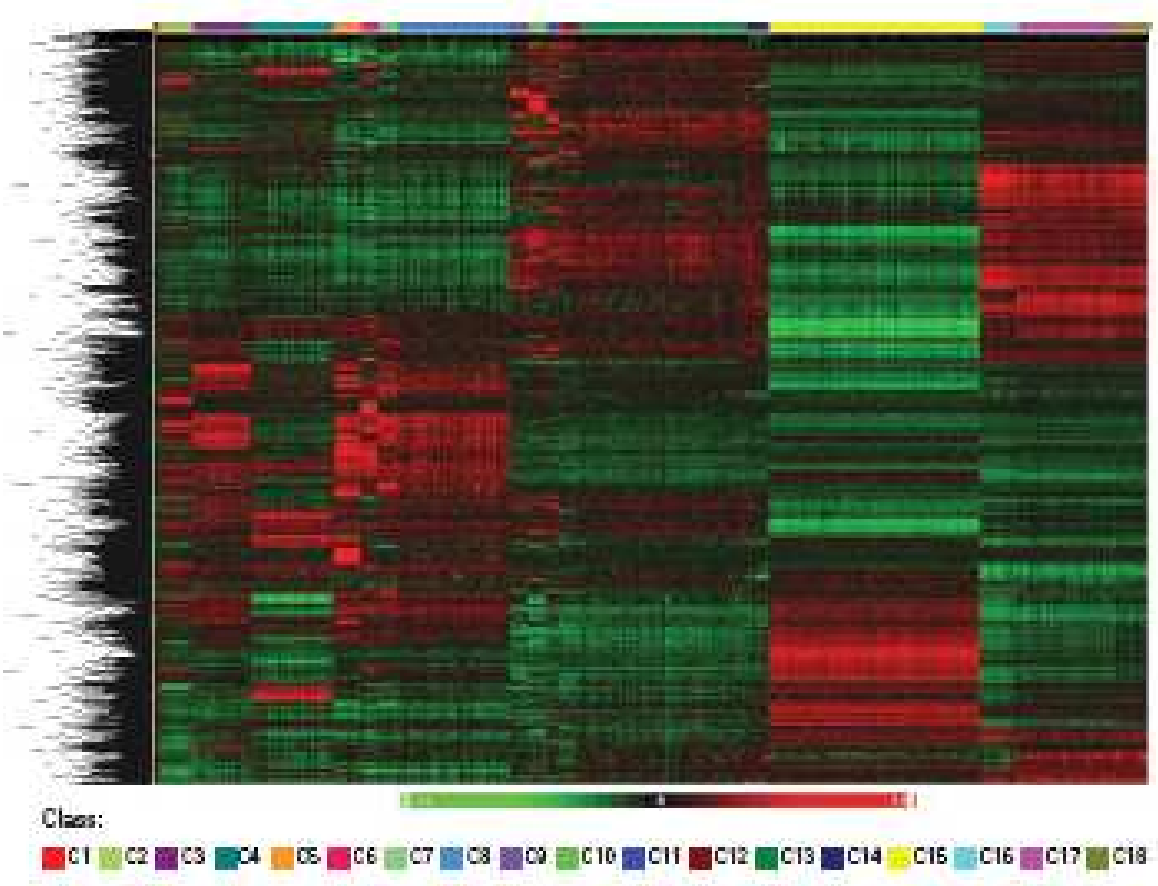


Figure 4. Supervised hierarchical clustering.The exploratory whole-genome clustering analysis was performed for all classes (C1 to C18 in ascending order) including 2,096 samples from stage I. For every class pair, the top

100 differentially expressed probes sets with the largest absolute values of t statistic were selected. The union of these sets contained 3,556 probe sets used in the clustering (from Haferlach et al., 2010)[125].

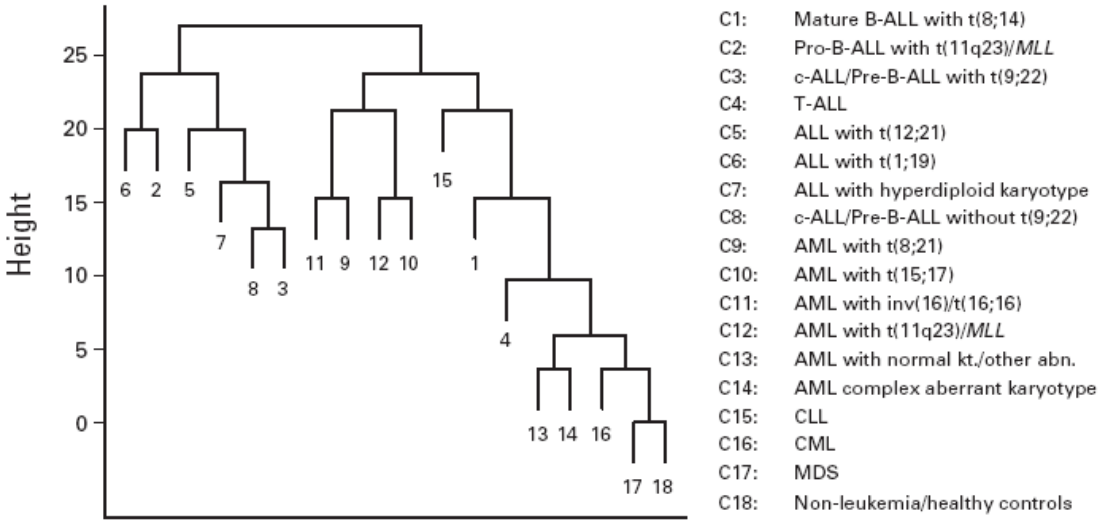


Figure 5. Exploratory margin tree analysis. Margin tree classification is a supervised multiclass support vector machine classification method. The margin tree program was applied to the stage I data set of 2,096 samples, characterized by their 18 class subtype labels (C1 to C18), and was based on 54,630 probe sets (from Haferlach et al., 2010)[126].

B-ALL, B-cell acute lymphoblastic leukemia; MLL, myeloid/lymphoid or mixed-lineage leukemia; pre, precursor; c-ALL, childhood acute lymphoblastic leukemia; T-ALL, T-cell acute lymphoblastic leukemia; ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; kt., karyotype; abn., abnormality; CLL, chronic lymphocytic leukemia; CML, chronic myelogenous leukemia; MDS, myelodysplastic syndrome.

GEP FOR RISK ASSESSMENT AND PREDICTION OFDRUG RESPONSE

A three gene classifier predictive for prognosis in paediatric pre-B-ALL was developed by Hoffmann et al. and were based on the glutamine synthetase (*GLUL*), ornithine decarboxylase antizyme inhibitor (*AZIN*), and immunoglobulin J chain (*IGJ*) genes[127]. Kang et al. investigated 207 children with high-risk B-precursor ALL. A 38-gene expression classifier in combination with flow cytometric measurement of minimal residual disease (MRD) were able to categorize children with high-risk ALL into three groups with relapse-free survival differing between 29% and 87% ($p < 0.001$)[128].

Assays to predict the response to specific ALL treatment approaches were also developed. Hofmann et al. published a 95-gene classifier being able to predict the sensitivity to the tyrosine kinase inhibitor imatinib in *BCR/ABL1* positive ALL[129]. Holleman et al. tested leukaemia cells from 173 children with ALL for in vitro sensitivity to steroids, vincristine, asparaginase, and daunorubicine. The differentially expressed genes were correlated with

treatment outcomes in the original group of patients and an independent cohort of 98 children. Following the assessment of differentially expressed genes being predictive for the single compounds, a combined gene expression score predicting response to the four compounds was developed[130].

In 2009, Den Boer et al. has developed and validated a new gene-expression classifier that identifies the major subtypes of childhood ALL with a high level of accuracy and sensitivity. Importantly, they identified a new high-risk subtype with a gene-expression pattern similar to that of ALL cases containing the *BCR–ABL1* gene fusion. This *BCR–ABL1*-like subtype was characterised by abnormalities in B-cell development genes (*IKZF1*, *PAX5*, *VPREB1*, *TCF3*, *EBF1*) indicating a defective (pre)B-cell receptor signalling pathway[131].
Recently Harvey et al. correlated gene expression data with genome-wide DNA copy number abnormalities (CNAs) in a cohort of high risk BCP ALL children [132]. He identified 8 gene expression-based cluster groups that were also characterized by distinct patterns of genome-wide DNA CNAs, 6 of these groups lacked any previously known cytogenetic lesion and 2 groups were also associated with different preclinical characteristics and treatment outcomes.

In summary, microarray technology can represent a highly useful adjunct to current diagnostic algorithms in haematological malignancies. These novel algorithms might combine GEP with established PCR techniques either for confirmation of microarray results[133] or for the selection of markers being appropriate for minimal residual disease (MRD) monitoring during follow-up[134, 135]. Integration among GEP and microRNAs profiling can improve classification[136] and knowledge of deregulated pathway involved in the diseases[137].
Another interesting option is the integration with next-generation high-throughput sequencing[138]. Finally, the SNP microarray application[139] which allows the most sensitive detection of genetic gains or losses to date, might be combined with GEP which clearly detects the vast majority of reciprocal translocations[140].

MICRORNA

MicroRNAs (miRNAs) are short strands of RNA (21–23 nucleotides) that act as regulatory molecules in eukaryotic cells by binding to a noncoding region within target messenger RNAs (mRNAs), namely the 3'-untranslated region (3'-UTR) and lead to translational repression or mRNA degradation. Through this mechanism miRNAs regulate self-renewal, differentiation, and division of cells via post-transcriptional gene silencing[141]. MiRNAs play important roles in many cellular processes such as development[142], stem cell division[143], apoptosis[144], disease[145], and cancer[146].

BIOGENESIS AND FUNCTION

Canonical pathway for miRNAs biogenesis begins with transcription of the encoded genes by RNA polymerase II. The majority of human miRNAs are expressed from introns, that are non-coding regions of the genome[147]. An intron of about 400 nucleotides is excised from the primary transcript and becomes the primary miRNA (pri-miRNA). The pri-miRNA is then processed by the RNase Drosha into hairpin loops about 70 nucleotides in length, forming the pre-miRNA. The pre-miRNA is exported to the cytoplasm by the nuclear membrane protein Exportin-5. Once in the cytoplasm, the RNase Dicer completes processing, forming mature miRNAs[148] (Figure 6).

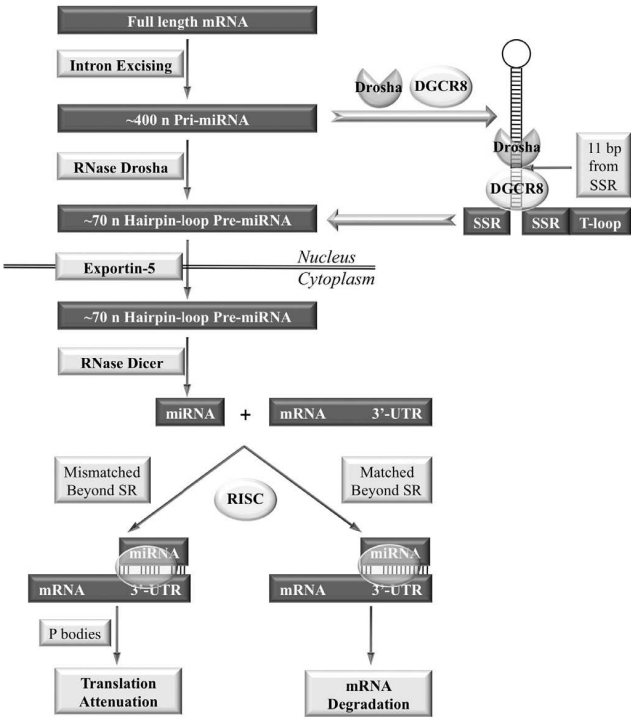


Figure 6. MicroRNAs biogenesis and function (From Zimmerman and Wu, 2011) [149].

MiRNAs regulate gene expression by either inhibiting translation or promoting degradation of specific mRNA transcripts. Post-transcriptional gene silencing begins when a miRNA recruits RISC (RNA-induced silencing complex), which is a complex of proteins that localizes the miRNA to its complementary target mRNA. Nucleotides 2–8 of the miRNA, called the seed region (SR), must bind contiguously to a perfectly complementary sequence on the target mRNA. The binding sites for the miRNA seed region lie in the 3'-UTR of the mRNA, and the complementary sequence usually repeats multiple times within the 3'-UTR. Whether an miRNA promotes degradation or represses translation of its target mRNA likely depends on the degree of complementary binding beyond the seed region[150].

Degradation of the mRNA appears to occur when the sequence beyond the seed region of the miRNA is perfectly complementary to the remaining region in the 3'-UTR of the mRNA. This degradation involves deadenylation, decapping, and exonucleolytic cleavage of the target mRNA transcript[151], but the exact mechanism is still unknown.

Translational repression occours when there is mismatched base pairs between the miRNA and mRNA. MiRNAs repress translation via different mechanisms either at the initiation or elongation step of translation, based on how they modulate the interaction between the 5' cap and the 3' polyadenylated tail of the mRNA[152]. Although the exact mechanism is not well understood, translational repression involves cytoplasmic mRNA-processing bodies (P bodies) that house various proteins that control messenger ribonucleoprotein (mRNP) complexes[153]. These mRNP complexes consist of mRNA and repressor proteins, and they lack translation initiation factors[154]. Thus, the localization of target mRNAs to P bodies, and subsequent formation of the mRNP complexes, inhibits translation.

MICRORNA IN NORMAL HEMATOPOIESIS

In normal hematopoiesis, long-term reconstituting hematopoietic stem cells divide into their short-term counterparts, which then give rise to multipotent progenitors. These multipotent progenitors differentiate into either lymphoid stem cells or myeloid stem cells; the latter differentiation step is inhibited by miR-128a and miR- 181a and activated by miR-223. The differentiation of lymphoid stem cells into common lymphoid progenitors is inhibited by miR-146 and activated by miR-181[155]. Common lymphoid progenitors then differentiate into T cells (activated by miR-150), B cells, and natural killer cells. Differentiation of myeloid stem cells into common myeloid progenitors is inhibited by miR-155, miR-24a, and miR-

17[156]. Common myeloid progenitors differentiation into either granulocyte–macrophage progenitors, inhibited by miR- 16, miR-103, and miR-107, or megakaryotic-erythroid progenitors. Granulocyte–macrophage differentiation into granulocytes is inhibited by miR-223, while differentiation into monocytes is inhibited by miR-17-5p, miR-20a, and miR-106a[157]. Megakaryotic-erythroid progenitors differentiate into either megakaryocyte progenitors, which develop into platelets, or erythroid progenitors, the formation of which is inhibited by miR-24[158]. Erythroid progenitors development into red blood cells is activated from miR-451 and miR-16, or inhibited by miR-150, miR-155, miR-221, and miR-222[159, 160] (Figure 7).

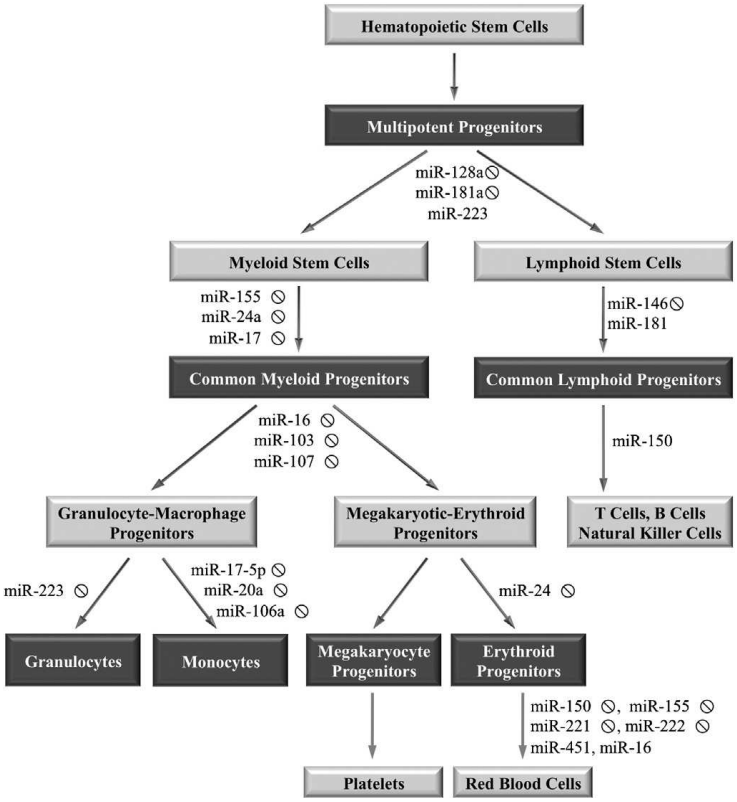


Figure 7. Regulation of hematopoiesis by miRNAs (inhibitory effect shown with Ø) (From Zimmerman and Wu, 2011)[161].

MICRORNA IN LEUKEMIA

Since miRNAs are critical for both stem cell development and cancer pathogenesis, they are being examined for their regulatory roles in self-renewal, proliferation, and differentiation of cancer cells[162]. MiRNAs have been shown to act as both tumor suppressors, which help control growth, and oncogenes, which promote rapid growth[163]. Global inhibition

of miRNAs processing increased tumorigenicity and transformation as reported in a recent study[164], which suggests their important regulatory role. In addition, many cancer-associated regions of the genome contain miRNA genes[165]. Aberrant miRNA levels, specifically an overall downregulation, are present in many cancers, as compared to their normal tissue counterparts. Specific miRNAs expression patterns have been described in several cancers and expression profiling of miRNAs has been shown to be a more accurate method of classifying cancer subtypes than that of mRNAs[166]. Although it is not entirely clear if aberrant miRNAs expression is a cause or an effect of the tumorigenic state, the significance of these regulatory molecules in cancer is apparent.

A screening of deregulated microRNAs in ALL has been published, providing a list of microRNAs involved in leukemogenesis. The five highly expressed microRNAs in ALL were mir-128b, mir-204, mir-218, mir-331 and mir-181b-1. The most commonly represented microRNA in ALL is mir-128b with a fivefold difference compared to normal CD19+ B cells. The second most highly expressed microRNA in ALL is mir-204. On the contrary, the five microRNAs with the lowest expression levels are mir-135b, mir-132, mir-199s, mir-139 and mir-150[167]. The mir-17-92 cluster was found to be up-regulated in ALL and to antagonize the expression of the pro-apoptotic protein BIM favouring the survival of B-cell progenitors[168]. Mi et al. performed a large-scale genome wide microRNAs expression profile assay and identified 27 microRNAs that are differentially expressed between AML and ALL. Among them, mir-128a and mir-128b are significantly over-expressed, whereas let-7b and mir-223 are significantly down-regulated in ALL compared to AML[169].

MicroRNAs potential application as therapeutic agents or target is of great importance. The potential use of anti-microRNAs molecules, termed antagomirs, locked nucleic acid (LNA)-anti-miR oligonucleotides or anti-microRNA oligonucleotides, has been studied by several groups, *in vitro* and *in vivo* and the first clinical trial applying anti-micross agents as drugs has recently been launched[170, 171].

The observation that microRNAs hypermethylation, which results in microRNAs down-regulation, is associated with tumorigenesis raised hope that methylated microRNAs could represent a new target for hypomethylating therapy. In ALL, for example, DNA methylation results in downregulation of several microRNAs while treatment with 5-aza-2'-deoxycytidine results in up-regulation of these microRNAs. In addition, methylation of microRNAs seems to represent an independent prognostic factor associated with poor disease-free and overall survival[172, 173].

Apart from their potential use as therapeutic agents, microRNAs may represent a useful tool in predicting the clinical outcome of a disease or even identifying subgroups of patients at high risk, who need more intensive therapy or closer follow up[174, 175].

OTHER NON CODING RNA: SNORNA AND SCARNA

Several other species of non-coding RNAs (ncRNAs), other than microRNAs, are revealing their involvement in human disease. Between them, small nucleolar RNAs (snoRNAs) and small Cajal body-specific RNAs (scaRNAs) have been implicated in human diseases and cancer[176, 177].

Since the 1990s, a vast collection of snoRNAs in eukaryotic cell have been found to be involved in ribosomal RNA (rRNA) methylation and pseudouridilation[178, 179] - modifications of pre-rRNA that results in the removal of extended spacer regions from the primary transcript. Later in 2001, scaRNAs, was discovered to guide modification in spliceosomal small nuclear RNAs (snRNAs)[180].

Based on distinct sequence motifs and sub-cellular locations, sno/scaRNAs fall into three major groups: box C/D snoRNAs, box H/ACA snoRNAs and scaRNAs[181, 182].

Box C/D snoRNAs (SNORDs) share two short sequence motifs, box C (AUGAUGA) at the 5' ends and box D (CUGA) at the 3' ends, respectively. Two imperfect copies of these boxes, namely box C' and box D', have also been found in some box C/D snoRNAs. Immediately upstream of box D and/or D' is a 10–21 nt antisense element complementary to targeted RNAs[183]. Both the AUGAUGA and CUGA box motifs and the antisense element play essential roles in RNA methylation or processing[184]. Each methylation site exclusively pairs with the fifth nucleotide upstream of box D or box D' in the complementary region between a box C/D snoRNA and targeted RNA[185] .

Box H/ACA snoRNAs (SNORAs) contain two conserved sequence motifs: a box H (ANANNA, where N stands for any nucleotide) and a box ACA (ACANNN), and two stem–loops at 5' and 3' end, respectively. In the internal loop of the one or two stems is an appropriate bipartite guide sequence of 4–10 nts that forms a short snoRNA–rRNA duplex flanking the target site[186]. The pseudouridylation site also obeys a spacing rule and it always appears at 14–16 nts upstream of box H or ACA within the bipartite guide sequence of a box H/ACA snoRNA[187] (Figure 8).

Different from the location of box C/D and box H/ACA snoRNAs in the nucleoli, scaRNAs accumulate within the Cajal bodies, conserved subnuclear organelles that are present in the

nucleoplasm. Moreover, a scaRNA molecule, can possess both box C/D and box H/ACA sequence motifs, guiding both the methylation and pseudouridylation of snRNAs[188].

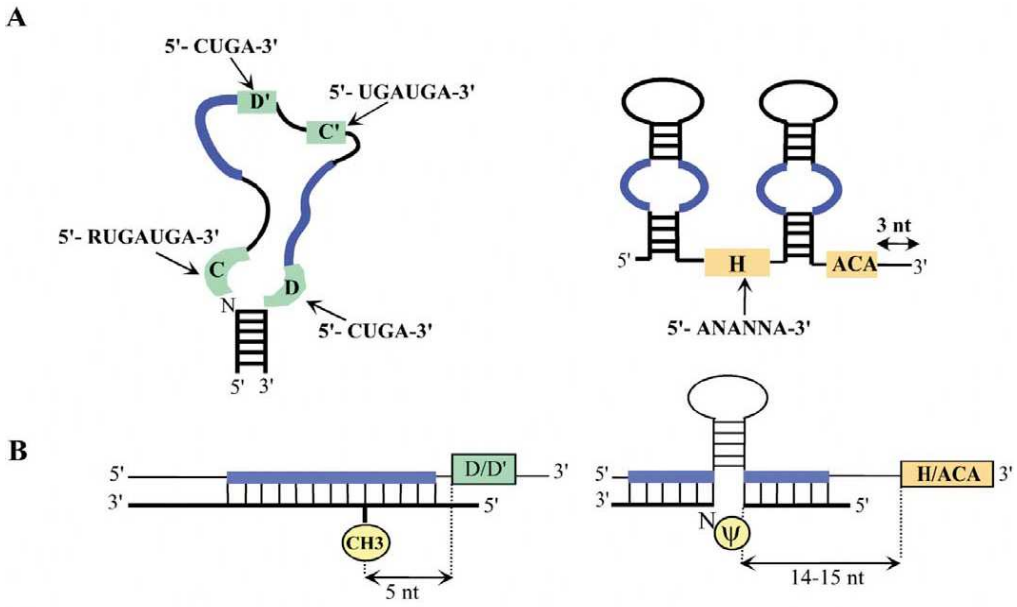


Figure 8. Structural features of the two families of modification guide snoRNAs. (A) Schematic secondary structures of the C/D and H/ACA classes of eukaryotic snoRNAs, with indication of their conserved box motifs (in green and orange boxes, respectively) and sequence tracts complementary to the cognate RNA target, also termed antisense elements (thick blue lines). The nucleotide targeted for modification is denoted by a yellow circle. For C/D snoRNAs (left), the 5'-3' terminal stem allowing the formation of the box C/D structural motif is represented. (B) Canonical structure of each type of guide RNA duplex (From Bachellerie et al., 2002)[189].

A link between snoRNAs and carcinogenesis was first reported by Chang et al, who found that snoRNA h5sn2 was highly expressed in normal brain, but its expression was dramatically reduced in meningioma, suggesting a role for the loss of snoRNA h5sn2 in brain tumorigenesis[190]. Recently, a homozygous 2 bp (TT) deletion in snoRNA U50 was discovered in prostate cancer cell lines and localized prostate tumor tissues, while the heterozygous genotype of the deletion occurred more frequently in women with breast cancer[191, 192]. A recent study has shown that a single microdeletion involving several snoRNA clusters (HBII-85 and HBII-52) results in Prader–Willi syndrome (PWS), suggesting that loss of small RNAs is a causal determinant of the disease[193].

Furhermore, a role of snoRNAs in regulation of alternative splicing of mRNA has been described[194].

Although studies are just emerging, snoRNAs may play malfunction in the development and progression of human malignancy and loss of small RNA loci plays an important role in human diseases.

MODEL SYSTEMS FOR LEUKEMIA STUDY

Understanding the complex molecular pathways leading to disease is critical for the development of effective treatment regimes and novel drug targets. Due to research and resource limitations associated with the use of primary patient material, pre-clinical models are essential to expand our knowledge of cancer biology and for the evaluation of new drugs.

CELL LINES

Cell lines are monoclonal populations of immortalized cells that are able to continually proliferate in culture without undergoing senescence. These continuous cell lines are particularly useful in research settings as they provide a virtually unlimited supply of homogeneous cell material for genetic and biochemical studies. Human cell lines are difficult to develop even from cancerous tissues[195] despite this apparent hindrance more than 1000 human hematopoietic cell lines have been generated since the 1960s by setting up long term cell culture from bone marrow (BM), peripheral blood (PB) or pleural effusions of ALL, AML or lymphoma patients[196, 197]. Whereas normal hematopoietic cells can only survive *in vitro* for days to weeks, these cell lines are able to proliferate continuously in culture while preserving the majority of their characteristic genetic alterations and phenotypic features.

Cell lines carrying many of the key translocations and oncogenes implicated in human leukemogenesis are established over the past years. When leukemic cell lines bearing an oncogene of interest are not readily available, an alternate approach can be followed to introduce this gene into a cell line and to characterize the effects on proliferation, differentiation and intracellular signaling[198].

Cell lines derived from sources other than human hematopoietic malignancies have also been used to study the functional consequences of oncogene expression. For example, Ba/F3 cells are an interleukin-3 (IL-3)-dependent pro-B cell line derived from mouse PB cells[199]. Following the demonstration that BCR– ABL can confer factor independence to these cells, they have been extensively used to examine the potency and downstream signalling of this and other activated tyrosine kinases[200, 201]. Though leukemic cell lines are often similar to the malignant blast population from which they were derived, the process of immortalization and adaptation to continuous growth *in vitro* likely involves the acquisition of a number of genetic/ epigenetic changes[202]. Also, during the course of extended culture in vitro, further differences (phenotypic and/or genetic) could develop[203]. Together, these points emphasize

that cell lines should not be considered as absolute equivalents of their corresponding primary cancers. Also, as cell lines are already fully transformed and immortalized, they cannot be utilized to study the early events of leukemogenic initiation and progression that lead to the generation of leukemia stem cells. Together, these limitations underscore the need to study leukemia-associated oncogenes in a more relevant cellular context such as primary hematopoietic stem and progenitor cell populations.

MURINE LEUKEMIA MODELS

To gain insight into the early events in leukemic transformation, there is a need for prospective *in vivo* models. The mouse has proven to be an exceptionally useful model system that shares gross physiological, anatomical and genomic similarities with humans and is relatively easy to genetically manipulate. The main approaches used for these studies are the development of transgenic mice and retroviral transduction/transplantation.

The earliest transgenic mice were developed by the direct injection of exogenous DNA cassettes encoding a gene of interest and a heterologous promoter into the pronuclei of fertilized zygotes[204, 205]. The emergence of embryonic stem cell technology has led to the development of a second generation of transgenic mice by a 'knock-in' approach[206]. To date, the most elegant means of generating transgenic mice bearing chromosomal translocations has been developed by Rabbitts and colleagues[207]. In 'translocator' mice, chromosomal translocations are generated de novo in a conditional manner by Cre-mediated interchromosomal recombination between loxP sites engineered at the breakpoint regions of both partner genes.

An alternate approach for functionally characterizing leukemic oncogenes is BM transduction/transplantation (BMT). Murine BM progenitors are isolated, transduced *ex vivo* with an oncogene-carrying retrovirus, then characterized for their effects on proliferation, differentiation and self-renewal. The BM progenitors can also be transplanted back into syngeneic recipients following transduction, enabling an assessment of their leukemogenic potential in vivo[208].

XENOTRANSPLANTATION MODELS

Transplantation models are especially important for assaying tumor cell malignancy and migration capabilities, and are critical for identifying putative cancer stem cell populations.

First experiments in which normal and leukemic human hematopoietic cells were transplanted into irradiated, athymic Nude (nu/nu) mice were only partially successful since these mice still had antibody-producing B cells, complement, and natural killer (NK) cells that rejected the human cells[209]. The development of bg/nu/xid (BNX) mice by Andriole et al.[210] and the discovery of severe combined immune-deficient (SCID) mice by Bosma and Carrol[211] had a dramatic effect on human xenotransplantation. The BNX mouse was constructed by combining three recessive mutations, bg (beige), nu (nude), and xid (X-linked immunodeficiency) to generate a more immune-deficient mouse for human cells. SCID mice are homozygous for a mutation in the *scid* gene. Mutations in the *scid* gene result in unsuccessful DNA rearrangement, preventing productive rearrangement of immunoglobulin and T cell receptor genes, resulting in T and B cell deficiency. However, there is a residual immunity in the form of NK cells, complement, and myeloid cells. Mice carrying the *scid* mutation also have a generalized radiation repair defect that renders them twice as sensitive as wild-type mice to ionizing radiation.

The SCID mouse model was greatly improved by the development of non-obese diabetic (NOD)/SCID mice. Shultz et al.[212] created NOD/SCID mice by backcrossing the *scid* mutation onto the NOD/LtSz strain of mice. The new mice have less residual immunity than SCID mice since NOD mice have defects in the complement pathway and macrophage function. NOD/SCID mice do not have autoimmune diabetes since they lack T and B cells due to the SCID mutation. Shultz et al. also report that un-stimulated “clean” NOD/SCID mice have lower levels of NK cells. As a result NOD/SCID mice have higher levels of engraftment both for normal and leukemic human cells than SCID mice and more importantly enable engraftment with lower cell doses, purified normal CD34+ cells, and their subpopulations[213, 214].

Over the last 15 years, immune-compromised mice, have been the model of choice to study morphological and biological characteristics of human cancers in vivo[215].

Following injection into immunodeficient mice, PB and BM cells from primary human ALLs and AMLs are capable of growing *in vivo* and generating grafts which phenotypically resemble the parent leukemia[216, 217]. Interestingly, it has been found that poor prognosis

leukemia tend to robustly engraft immunodeficient mice, suggesting that a correlation exists between xeno-engraftment potential and disease aggressiveness in humans[218, 219].

The kinetics of engraftment reflects the human disease, leading to bone marrow (BM) infiltration, followed by migration to the spleen, peripheral blood and other haematopoietic organs [220, 221].

Xenotransplantation assays have also been extensively used to characterize the cell types responsible for initiating leukemic growth *in vivo*. These cells, operationally defined as SCID-leukemia initiating cells (SL-ICs), provide evidence for the existence of a phenotypically distinct population of leukemia stem cells (LSCs), responsible for initiating and propagating disease *in vivo*[222].

It is important to note that the growth of human leukemic cells in xenotransplant recipients is undoubtedly limited by residual elements of the recipient immune system, the absence of cross-species reactivity of some cytokines, and differences between the murine and human microenvironment. Together, these factors may limit the ability of some human leukemia samples to successfully engraft, and also lead to an underestimation of the absolute frequency of LSC[223].

Optimization of xenotransplantation assays has been developed during the past years and newer strains of mice had been engineered. The introduction of target deletion of the β2-microglobulin gene within a NOD/SCID background have resulted in models with decreased NK cell function better suited for studying the progression of diseases[224]. More recently, reports have demonstrated that a targeted deletion in the γ-common chain in NOD/SCID mice (NSG mice) results in the elimination of residual NK cell activity and provides an improved environment for growth and development of human cells[225, 226].

An alternate approach to render xenotransplant recipients more ‘humanized’ could involve the co-transplantation of human mesenchymal stem cells (hMSCs) and human hematopoietic cells into immunodeficient mice. Following intra-osseous injection, hMSCs can contribute to the hematopoietic microenvironment within the mouse marrow, generating stromal cells, osteoblasts, endothelial cells and other BM cell types[227].

OTHER IN VIVO MODELS

Alternative animal system has been proposed such as canine model. The out-bred dog offers several advantages including dimensions similar to children and similar immune

systems[228], propensity toward the same genetic diseases[229], homology of genes[230] compared to humans.

In recent years, zebrafish have been proposed as a cost-effective alternative to mammals such as rodents and dogs[231]. Transgenic fish provide excellent models of transplantable T-cell[232] and B-cell[233] acute lymphoid leukemia that could be used for assessing drug efficacy[234]. Several studies demonstrate that transplanted leukemia cells can engraft and perpetuate the tumor in zebrafish models of leukemia[235].

PRIMARY HUMAN HEMATOPOIETIC CELLS

In vitro - To asses oncogenes in the appropriate cellular context, experimental protocols have been developed where genetic 'hits' are induced in normal human primitive hematopoietic cells by retroviral-mediated oncogene overexpression, allowing for the early steps of leukemogenesis to be effectively modeled[236]. Human hematopoietic cells are available in the form of adult BM, mobilized PB or umbilical cord blood[237].

In vivo - Leukemic oncogenes studied in the appropriate cellular context (primary human hematopoietic cells genetically modified) are assayed for their ability to initiate disease *in vivo.* Such studies are performed by transplantion of blast cell populations generated during *in vitro* propagation of transduced human hematopoietic cells or transplantation of transduced cells immediately following the infection period[238].

REFERENCES

[1] J Formos Med Assoc. 2010 Nov;109(11):777-87. Recent research advances in childhood acute lymphoblastic leukemia. Pui CH.

[2] Hjalgrim, L. L., Rostgaard, K., Schmiegelow, K., So¨ derha¨ ll, S., Kolmannskog, S., et al. Age- and sex-specific incidence of childhood leukemia by immunophenotype in the Nordic countries. J Natl Cancer Inst 2003. 95: 1539–1544.

[3] Pui, C. H., Relling, M. V. and Downing, J. R., Mechanisms of Disease: acute lymphoblastic leukemia. N Engl J Med 2004. 350: 1535–1548.

[4] Pui CH, Evans WE: Treatment of acute lymphoblastic leukemia. N Engl J Med 2006;354:166–78.

[5] Campana D. Minimal residual disease in acute lymphoblastic leukemia. Semin Hematol 2009;46:100–6.

[6] Stow P, Key L, Cjen X, et al. Clinical significance of low levels of minimal residual disease at the end of remission induction therapy in childhood acute lymphoblastic leukemia. Blood 2010;115:4657–63.

[7] Yang JJ, Bhojwani D, Yang W, et al. Genome-wide copy number profiling reveals molecular evolution from diagnosis to relapse in childhood acute lymphoblastic leukemia. *Blood* 2008;112:4178–83.

[8] Mullighan CG, Phillips LA, Su X, et al. Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. Science 2008;322:1377–80.

[9] Bhojwani D, Kang H, Moskowitz NP, et al. Biologic pathways associated with relapse in childhood acute lymphoblastic leukemia: a Children's Oncology Group study. Blood 2006; 108:711–7.

[10] Moorman, A.V., Ensor, H.M., Richards, S.M., Chilton, L., Schwab, C., Kinsey, S.E., Vora, A., Mitchell, C.D. & Harrison, C.J. (2010) Prognostic effect of chromosomal abnormalities in childhood B-cell precursor acute lymphoblastic leukaemia: results from the UK Medical Research Council ALL97/99 randomised trial. The Lancet Oncology, 11, 429–438.

[11] Pieters, R., Schrappe, M., De, L.P., Hann, I., De, R.G., Felice, M., Hovi, L., Leblanc, T., Szczepanski, T., Ferster, A., Janka, G., Rubnitz, J., Silverman, L., Stary, J., Campbell, M., Li, C.K., Mann, G., Suppiah, R., Biondi, A., Vora, A. & Valsecchi, M.G. (2007) A treatment protocol for infants younger than 1 year with acute lymphoblastic leukaemia (Interfant-99): an observational study and a multicentre randomised trial. Lancet, 370, 240–250.

[12] Pui, C.H., Raimondi, S., Hancock, M.L., Rivera, G.K., Ribeiro, R.C., Mahmoud, H.H., Sabnlund, J.T., Crist, W.M. & Behm, F.G. (1994) Immunologic, cytogenetic, and clinical characterization of childhood acute lymphoblastic leukemia with the t(1;19)(q23;p13) or its derivative. Journal of Clinical Oncology, 12, 2601–2606.

[13] Kager, L., Lion, T., Attarbaschi, A., Koenig, M., Strehl, S., Haas, O.A., Dworzak, M.N., Schrappe, M., Gadner, H. & Mann, G. (2007) Incidence and outcome of TCF3-PBX1-positive acute lymphoblastic leukemia in Austrian children. Haematologica, 92, 1561–1564.

[14] Hunger, S.P. (1996) Chromosomal translocations involving the E2A gene in acute lymphoblastic leukemia: clinical features and molecular pathogenesis. Blood, 87, 1211–1224.

[15] Harewood, L., Robinson, H., Harris, R., Al Obaidi, M.J., Jalali, G.R., Martineau, M., Moorman, A.V., Sumption, N., Richards, S., Mitchell, C. & Harrison, C.J. (2003) Amplification of AML1 on a duplicated chromosome 21 in acute lymphoblastic leukemia: a study of 20 cases. Leukemia, 17, 547–553.

[16] Strefford, J.C., Van Delft, F.W., Robinson, H.M., Worley, H., Yiannikouris, O., Selzer, R., Richmond, T., Hann, I., Bellotti, T., Raghavan, M., Young, B.D., Saha, V. & Harrison, C.J. (2006) Complex genomic

alterations and gene expression in acute lymphoblastic leukemia with intrachromosomal amplification of chromosome 21. Proceedings of the National Academy of Sciences USA, 103, 8167–8172.

[17] Robinson, H.M., Harrison, C.J., Moorman, A.V., Chudoba, I. & Strefford, J.C. (2007) Intrachromosomal amplification of chromosome Chromosome 21 (iAMP21) may arise from a breakage-fusion-bridge cycle. Genes, Chromosomes and Cancer, 46, 318–326.

[18] Robinson, H.M., Broadfield, Z.J., Cheung, K.L., Harewood, L., Harris, R.L., Jalali, G.R., Martineau, M., Moorman, A.V., Taylor, K.E., Richards, S., Mitchell, C. & Harrison, C.J. (2003) Amplification of AML1 in acute lymphoblastic leukemia is associated with a poor outcome. Leukemia, 17, 2249–2250.

[19] Soulier, J., Trakhtenbrot, L., Najfeld, V., Lipton, J.M., Mathew, S., Avet-Loiseau, H., De Braekeleer, M., Salem, S., Baruchel, A., Raimondi, S.C. & Raynaud, S.D. (2003) Amplification of band q22 of chromosome 21, including AML1, in older children with acute lymphoblastic leukemia: an emerging molecular cytogenetic subgroup. Leukemia, 17, 1679–1682.

[20] Soulier, J., Trakhtenbrot, L., Najfeld, V., Lipton, J.M., Mathew, S., Avet-Loiseau, H., De Braekeleer, M., Salem, S., Baruchel, A., Raimondi, S.C. & Raynaud, S.D. (2003) Amplification of band q22 of chromosome 21, including AML1, in older children with acute lymphoblastic leukemia: an emerging molecular cytogenetic subgroup. Leukemia, 17, 1679–1682.

[21] Moorman, A.V., Richards, S.M., Robinson, H.M., Strefford, J.C., Gibson, B.E., Kinsey, S.E., Eden, T.O., Vora, A.J., Mitchell, C.D. & Harrison, C.J. (2007) Prognosis of children with acute lymphoblastic leukemia (ALL) and intrachromosomal amplification of chromosome 21 (iAMP21). Blood, 109, 2327–2330.

[22] Attarbaschi, A., Mann, G., Panzer-Grumayer, R., Rottgers, S., Steiner, M., Konig, M., Csinady, E., Dworzak, M.N., Seidel, M., Janousek, D., Moricke, A., Reichelt, C., Harbott, J., Schrappe, M., Gadner, H. & Haas, O.A. (2008) Minimal residual disease values discriminate between low and high relapse risk in children with B-cell precursor acute lymphoblastic leukemia and an intrachromosomal amplification of chromosome 21: the Austrian and German acute lymphoblastic leukemia Berlin-Frankfurt-Munster (ALL-BFM) trials. Journal of Clinical Oncology, 26, 3046–3050.

[23] Moorman, A.V., Richards, S.M., Martineau, M., Cheung, K.L., Robinson, H.M., Jalali, G.R., Broadfield, Z.J., Harris, R.L., Taylor, K.E., Gibson, B.E., Hann, I.M., Hill, F.G., Kinsey, S.E., Eden, T.O., Mitchell, C.D. & Harrison, C.J. (2003) Outcome heterogeneity in childhood high-hyperdiploid acute lymphoblastic leukemia. Blood, 102, 2756–2762.

[24] Harrison, C.J., Moorman, A.V., Broadfield, Z.J., Cheung, K.L., Harris, R.L., Reza, J.G., Robinson, H.M., Barber, K.E., Richards, S.M., Mitchell, C.D., Eden, T.O., Hann, I.M., Hill, F.G., Kinsey, S.E., Gibson, B.E., Lilleyman, J., Vora, A., Goldstone, A.H., Franklin, I.M., Durrant, J. & Martineau, M. (2004) Three distinct subgroups of hypodiploidy in acute lymphoblastic leukaemia. British Journal of Haematology, 125, 552–559.

[25] Nachman, J.B., Heerema, N.A., Sather, H., Camitta, B., Forestier, E., Harrison, C.J., Dastugue, N., Schrappe, M., Pui, C.H., Basso, G., Silverman, L.B. & Janka-Schaub, G.E. (2007) Outcome of treatment in children with hypodiploid acute lymphoblastic leukemia. Blood, 110, 1112–1115.

[26] Harrison, C.J., Moorman, A.V., Broadfield, Z.J., Cheung, K.L., Harris, R.L., Reza, J.G., Robinson, H.M., Barber, K.E., Richards, S.M., Mitchell, C.D., Eden, T.O., Hann, I.M., Hill, F.G., Kinsey, S.E., Gibson, B.E., Lilleyman, J., Vora, A., Goldstone, A.H., Franklin, I.M., Durrant, J. & Martineau, M. (2004) Three distinct subgroups of hypodiploidy in acute lymphoblastic leukaemia. British Journal of Haematology, 125, 552–559.

[27] Chapiro, E., Russell, L., Radford-Weiss, I., Bastard, C., Lessard, M., Cave, H., Fert-Ferrer, S., Barin, C., Maarek, O., Della-Valle, V., Strefford, J.C., Berger, R., Harrison, C.J., Bernard, O.A. & Nguyen- Khac, F. for the Groupe Francophone de Cytogenetique, Hematologique. (2006) Overexpression of CEBPA resulting from the translocation t(14;19)(q32;q13) of human precursor B acute lymphoblastic leukemia. Blood, 10 3560–3563.

[28] Akasaka, T., Balasas, T., Russell, L.J., Sugimoto, K.J., Majid, A., Walewska, R., Karran, E.L., Brown, D.G., Cain, K., Harder, L., Gesk S., Martin-Subero, J.I., Atherton, M.G., Bruggemann, M., Calasanz, M.J., Davies, T.,

Haas, O.A., Hagemeijer, A., Kempski, H., Lessard, M., Lillington, D.M., Moore, S., Nguyen-Khac, F., Radford-Weiss, I., Schoch, C., Struski, S., Talley, P., Welham, M.J., Worley, H., Strefford, J.C., Harrison, C.J., Siebert, R. & Dyer, M.J. (2007) Five members of the CEBP transcription factor family are targeted by recurrent IGH translocations in B-cell precursor acute lymphoblastic leukemia (BCP ALL). Blood, 109, 3451–3461.

[29] Russell, L.J., Akasaka, T., Majid, A., Sugimoto, K.J., Loraine, K.E., Nagel, I., Harder, L., Claviez, A., Gesk, S., Moorman, A.V., Ross, F., Mazzullo, H., Strefford, J.C., Siebert, R., Dyer, M.J. & Harrison, C.J. (2008) t(6;14)(p22;q32): a new recurrent IGH@ translocation involving ID4 in B-cell precursor acute lymphoblastic leukemia (BCP ALL). Blood, 111, 387–391.

[30] Russell, L.J., Capasso, M., Vater, I., Akasaka, T., Bernard, O.A., Calasanz, M.J., Chandrasekaran, T., Chapiro, E., Gesk, S., Griffiths, M., Guttery, D.S., Haferlach, C., Harder, L., Heidenreich, O., Irving, J., Kearney, L., Nguyen-Khac, F., Machado, L., Minto, L., Majid, A., Moorman, A.V., Morrison, H., Rand, V., Strefford, J.C., Schwab, C., Tonnies, H., Dyer, M.J., Siebert, R. & Harrison, C.J. (2009a) Deregulated expression of cytokine receptor gene, CRLF2, is involved in lymphoid transformation in B-cell precursor acute lymphoblastic leukemia. Blood, 114, 2688–2698.

[31] Greaves, M. F. & Wiemels, J. Origins of chromosome translocations in childhood leukaemia. Nature Rev. Cancer 3, 639–649 (2003).

[32] Mullighan, C.G., Goorha, S., Radtke, I., Miller, C.B., Coustan-Smith, E., Dalton, J.D., Girtman, K., Mathew, S., Ma, J., Pounds, S.B., Su, X., Pui, C.H., Relling, M.V., Evans, W.E., Shurtleff, S.A. & Downing, J.R. (2007) Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature, 446, 758–764.

[33] Busslinger, M. Transcriptional control of early B cell development. Annu. Rev. Immunol. 22, 55–79 (2004).

[34] Peters, U. R. et al. Aberrant FHIT mRNA transcripts are present in malignant and normal haematopoiesis, but absence of FHIT protein is restricted to leukaemia. Oncogene 18, 79–85 (1999).

[35] Jackson, A. et al. Deletion of 6q16-q21 in human lymphoid malignancies: a mapping and deletion analysis. Cancer Res. 60, 2775–2779 (2000).

[36] Okuda, T. et al. Frequent deletion of p16INK4a/MTS1 and p15INK4b/MTS2 in pediatric acute lymphoblastic leukemia. Blood 85, 2321–2330 (1995).

[37] Raynaud, S. et al. The 12;21 translocation involving TEL and deletion of the other TEL allele: two frequently associated alterations found in childhood acute lymphoblastic leukemia. Blood 87, 2891–2899 (1996).

[38] Raimondi, S. C. et al. Acute lymphoblastic leukemias with deletion of 11q23 or a novel inversion (11)(p13q23) lack MLL gene rearrangements and have favorable clinical features. Blood 86, 1881–1886 (1995).

[39] Heerema, N. A. et al. Abnormalities of chromosome bands 13q12 to 13q14 in childhood acute lymphoblastic leukemia. J. Clin. Oncol. 18, 3837–3844 (2000).

[40] Cave, H. et al. Deletion of chromosomal region 13q14.3 in childhood acute lymphoblastic leukemia. Leukemia 15, 371–376 (2001).

[41] BTG1 regulates glucocorticoid receptor autoinduction in acute lymphoblastic leukemia. van Galen JC, Kuiper RP, van Emst L, Levers M, Tijchon E, Scheijen B, Waanders E, van Reijmersdal SV, Gilissen C, van Kessel AG, Hoogerbrugge PM, van Leeuwen FN. Blood. 2010 Jun 10;115(23):4810-9. Epub 2010 Mar 30.

[42] Blood. 2010 Dec 2;116(23):4874-84. Epub 2010 Aug 10. Identification of novel cluster groups in pediatric high-risk B-precursor acute lymphoblastic leukemia with gene expression profiling: correlation with genome-wide DNA copy number alterations, clinical characteristics, and outcome. Harvey RC, Mullighan CG, Wang X, Dobbin KK, Davidson GS, Bedrick EJ, Chen IM, Atlas SR, Kang H, Ar K, Wilson CS, Wharton W, Murphy M, Devidas M, Carroll AJ, Borowitz MJ, Bowman WP, Downing JR, Relling M, Yang J, Bhojwani D, Carroll WL, Camitta B, Reaman GH, Smith M, Hunger SP, Willman CL.

[43] Mullighan, C.G., Collins-Underwood, J.R., Phillips, L.A., Loudin, M.G., Liu, W., Zhang, J., Ma, J., Coustan-Smith, E., Harvey, R.C., Willman, C.L., Mikhail, F.M., Meyer, J., Carroll, A.J., Williams, R.T., Cheng, J., Heerema, N.A., Basso, G., Pession, A., Pui, C.H., Raimondi, S.C., Hunger, S.P., Downing, J.R., Carroll, W.L. & Rabin, K.R. (2009) Rearrangement of CRLF2 in B-progenitor- and Down syndrome-associated acute lymphoblastic leukemia. Nature Genetics, 41, 1243–1246.

[44] Russell, L.J., De Castro, D.G., Griffiths, M., Telford, N., Bernard, O., Panzer-Grumayer, R., Heidenreich, O., Moorman, A.V. & Harrison, C.J. (2009) A novel translocation, t(14;19)(q32;p13), involving IGH@ and the cytokine receptor for erythropoietin. Leukemia, 23, 614–617.

[45] Yoda, A., Yoda, Y., Chiaretti, S., Bar-Natan, M., Mani, K., Rodig, S.J., West, N., Xiao, Y., Brown, J.R., Mitsiades, C., Sattler, M., Kutok, J.L., DeAngelo, D.J., Wadleigh, M., Piciocchi, A., Dal Cin, P., Bradner, J.E., Griffin, J.D., Anderson, K.C., Stone, R.M., Ritz, J., Foa, R., Aster, J.C., Frank, D.A. & Weinstock, D.M. (2009) Functional screening identifies CRLF2 in precursor B-cell acute lymphoblastic leukemia. Proceedings of the National Academy of Sciences of the USA, 107, 252–257.

[46] Hertzberg, L., Vendramini, E., Ganmore, I., Cazzaniga, G., Schmitz, M., Chalker, J., Shiloh, R., Iacobucci, I., Shochat, C., Zeligson, S., Cario, G., Stanulla, M., Strehl, S., Russell, L.J., Harrison, C.J., Bornhauser, B., Yoda, A., Rechavi, G., Bercovich, D., Borkhardt, A., Kempski, H., Kronnie, G.T., Bourquin, J.P., Domany, E. & Izraeli, S. (2010) Down syndrome acute lymphoblastic leukemia: a highly heterogeneous disease in which aberrant expression of CRLF2 is associated with mutated JAK2: a report from the iBFM Study Group. Blood, 115, 1006–1017.

[47] Russell, L.J., De Castro, D.G., Griffiths, M., Telford, N., Bernard, O., Panzer-Grumayer, R., Heidenreich, O., Moorman, A.V. & Harrison, C.J. (2009) A novel translocation, t(14;19)(q32;p13), involving IGH@ and the cytokine receptor for erythropoietin. Leukemia, 23, 614–617.

[48] Yoda, A., Yoda, Y., Chiaretti, S., Bar-Natan, M., Mani, K., Rodig, S.J., West, N., Xiao, Y., Brown, J.R., Mitsiades, C., Sattler, M., Kutok, J.L., DeAngelo, D.J., Wadleigh, M., Piciocchi, A., Dal Cin, P., Bradner, J.E., Griffin, J.D., Anderson, K.C., Stone, R.M., Ritz, J., Foa, R., Aster, J.C., Frank, D.A. & Weinstock, D.M. (2009) Functional screening identifies CRLF2 in precursor B-cell acute lymphoblastic leukemia. Proceedings of the National Academy of Sciences of the USA, 107, 252–257.

[49] Chapiro, E., Russell, L., Lainey, E., Kaltenbach, S., Ragu, C., Della- Valle, V., Hanssens, K., Macintyre, E.A., Radford-Weiss, I., Delabesse, E., Cave, H., Mercher, T., Harrison, C.J., Nguyen-Khac, F., Dubreuil, P. & Bernard, O.A. (2009) Activating mutation in the TSLPR gene in B-cell precursor lymphoblastic leukemia. Leukemia, Leukemia advance online publication 20 May 2010; doi: 10.1038/ leu.2010.93.

[50] Cario, G., Zimmermann, M., Romey, R., Gesk, S., Vater, I., Harbott, J., Schrauder, A., Moericke, A., Izraeli, S., Akasaka, T., Dyer, M.J., Siebert, R., Schrappe, M. & Stanulla, M. (2010) Presence of the *P2RY8-CRLF2* rearrangement is associated with a poor prognosis in non-high-risk precursor B-cell acute lymphoblastic leukemia in children treated according to the ALL-BFM 2000 protocol. Blood, doi:10.1182/blood-2009-11-256131.

[51] Harvey, R.C., Mullighan, C.G., Chen, I.M., Wharton, W., Mikhail, F.M., Carroll, A.J., Kang, H., Liu, W., Dobbin, K.K., Smith, M.A., Carroll, W.L., Devidas, M., Bowman, W.P., Camitta, B.M., Reaman, G.H., Hunger, S.P., Downing, J.R. & Willman, C.L.. (2010) Rearrangement of CRLF2 is associated with mutation of JAK kinases, alteration of IKZF1, Hispanic/Latino ethnicity, and a poor outcome in pediatric B-progenitor acute lymphoblastic leukemia. Blood, doi:10.1182/blood-2009-09-245944.

[52] Pui CH, Relling MV, Downing JR. Acute lymphoblastic leukemia. *N Engl J Med.* 2004;350:1535–48.

[53] J Clin Oncol. 2011 Jan 10. [Epub ahead of print] Biology, Risk Stratification, and Therapy of Pediatric Acute Leukemias: An Update. Pui CH, Carroll WL, Meshinchi S, Arceci RJ.

[54] Maher CA, Kumar-Sinha C, Cao X, et al. Transcriptome sequencing to detect gene fusions in cancer. Nature 2009;458:97–101.

[55] J Clin Oncol. 2011 Jan 10. [Epub ahead of print] Biology, Risk Stratification, and Therapy of Pediatric Acute Leukemias: An Update. Pui CH, Carroll WL, Meshinchi S, Arceci RJ.

[56] Friend SL, Hosier S, Nelson A, et al. Exp Hematol 1994;22(3):321. [PubMed: 8112430]

[57] Soumelis V, Reche PA, Kanzler H, et al. Nature immunology 2002;3(7):673. [PubMed: 12055625]

[58] Quentmeier H, Drexler HG, Fleckenstein D, et al. Leukemia 2001;15(8):1286. [PubMed: 11480573]

[59] Rimoldi M, Chieppa M, Salucci V, et al. Nature immunology 2005;6(5):507. [PubMed: 15821737]

[60] Ann N Y Acad Sci. 2010 Jan;1183:13-24. Thymic stromal lymphopoietin. He R, Geha RS.

[61] Thymic stromal lymphopoietin-mediated STAT5 phosphorylation via kinases JAK1 and JAK2 reveals a key difference from IL-7–induced signalling. Yrina Rochmana, Mohit Kashyapa, Gertraud W. Robinsonb, Kazuhito Sakamotoc, Julio Gomez-Rodriguezd, Kay-Uwe Wagnerc, and Warren J. Leonarda. PNAS,November 9, 2010, vol. 107, no. 45, 19455–19460.

[62] Ann N Y Acad Sci. 2010 Jan;1183:13-24. Thymic stromal lymphopoietin. He R, Geha RS.

[63] Thymic stromal lymphopoietin-mediated STAT5 phosphorylation via kinases JAK1 and JAK2 reveals a key difference from IL-7–induced signalling. Yrina Rochmana, Mohit Kashyapa, Gertraud W. Robinsonb, Kazuhito Sakamotoc, Julio Gomez-Rodriguezd, Kay-Uwe Wagnerc, and Warren J. Leonarda. PNAS,November 9, 2010, vol. 107, no. 45, 19455–19460.

[64] Thymic stromal lymphopoietin-mediated STAT5 phosphorylation via kinases JAK1 and JAK2 reveals a key difference from IL-7–induced signalling. Yrina Rochmana, Mohit Kashyapa, Gertraud W. Robinsonb, Kazuhito Sakamotoc, Julio Gomez-Rodriguezd, Kay-Uwe Wagnerc, and Warren J. Leonarda. PNAS,November 9, 2010, vol. 107, no. 45, 19455–19460.

[65] Hennighausen L, Robinson GW: Interpretation of cytokine signaling through the transcription factors STAT5A and STAT5B. Genes Dev 2008, 22:711-721.

[66] Yao Z, Cui Y, Watford WT, Bream JH, Yamaoka K, Hissong BD, Li D, Durum SK, Jiang Q, Bhandoola A et al.: Stat5a/b are essential for normal lymphoid development and differentiation. Proc Natl Acad Sci U S A 2006, 103:1000-1005.

[67] Hoelbl A, Kovacic B, Kerenyi MA, Simma O, Warsch W, Cui Y, Beug H, Hennighausen L, Moriggl R, Sexl V: Clarifying the role of Stat5 in lymphoid development and Abelson-induced transformation. Blood 2006, 107:4898-4906.

[68] Scheeren FA, Naspetti M, Diehl S, Schotte R, Nagasawa M, Wijnands E, Gimeno R, Vyth-Dreese FA, Blom B, Spits H: STAT5 regulates the self-renewal capacity and differentiation of human memory B cells and controls Bcl-6 expression. Nat Immunol 2005, 6:303-313.

[69] Malin S, McManus S, Cobaleda C, Novatchkova M, Delogu A, Bouillet P, Strasser A, Busslinger M: Role of STAT5 I controlling cell survival and immunoglobin gene recombination during pro-B cell development. Nat Immunol 2010, 11:171-179.

[70] Joliot V, Cormier F, Medyouf H, Alcalde H, Ghysdael J: Constitutive STAT5 activation specifically cooperates with the loss of p53 function in B-cell lymphomagenesis. Oncogene 2006, 25:4573-4584.

[71] Xie S, Wang Y, Liu J, Sun T, Wilson MB, Smithgall TE, Arlinghaus RB: Involvement of Jak2 tyrosine phosphorylation in Bcr–Abl transformation. Oncogene 2001, 20:6188-6195.

[72] Buettner R, Mora LB, Jove R: Activated STAT signalling in human tumors provides novel molecular targets for therapeutic intervention. Clin Cancer Res 2002, 8:945-954.

[73] Hoelbl A, Kovacic B, Kerenyi MA, Simma O, Warsch W, Cui Y, Beug H, Hennighausen L, Moriggl R, Sexl V: Clarifying the role of Stat5 in lymphoid development and Abelson-induced transformation. Blood 2006, 107:4898-4906.

[74] Hoelbl A, Schuster C, Kovacic B, Zhu B, Wickre M, Hoelzl MA, Fajmann S, Grebien F, Warsch W, Stengl G, et al.: Stat5 is indispensable for the maintenance of bcr/abl-positive leukaemia. EMBO Mol Med 2010, 2:in press

[75] Bercovich D, Ganmore I, Scott LM, Wainreb G, Birger Y, Elimelech A, Shochat C, Cazzaniga G, Biondi A, Basso G et al.: Mutations of JAK2 in acute lymphoblastic leukaemias associated with Down's syndrome. Lancet 2008, 372:1484-1492.

[76] Kearney L, Gonzalez De Castro D, Yeung J, Procter J, Horsley SW, Eguchi-Ishimae M, Bateman CM, Anderson K, Chaplin T, Young BD et al.: Specific JAK2 mutation (JAK2R683) and multiple gene deletions in Down syndrome acute lymphoblastic leukemia. Blood 2009, 113:646-648.

[77] Mullighan CG, Zhang J, Harvey RC, Collins-Underwood JR, Schulman BA, Phillips LA, Tasian SK, Loh ML, Su X, Liu W et al.: JAK mutations in high-risk childhood acute lymphoblastic leukemia. Proc Natl Acad Sci U S A 2009, 106:9414-9418.

[78] Yoda A, Yoda Y, Chiaretti S, Bar-Natan M, Mani K, Rodig SJ, West N, Xiao Y, Brown JR, Mitsiades C et al.: Functional screening identifies CRLF2 in precursor B-cell acute lymphoblastic leukemia. Proc Natl Acad Sci U S A 2009, 107:252-257.

[79] Ghoreschi K, Laurence A, O'Shea JJ: Janus kinases in immune cell signaling. Immunol Rev 2009, 228:273-287.

[80] Yoda A, Yoda Y, Chiaretti S, Bar-Natan M, Mani K, Rodig SJ, West N, Xiao Y, Brown JR, Mitsiades C et al.: Functional screening identifies CRLF2 in precursor B-cell acute lymphoblastic leukemia. Proc Natl Acad Sci U S A 2009, 107:252-257.

[81] Mitelman F, Heim S, Mandahl N. 1990. Trisomy 21 in neoplastic cells. Am J Med Genet Suppl 7:262–266.

[82] Karrman K, Forestier E, Andersen MK, Autio K, Borgstrom G, Heim S, Heinonen K, Hovland R, Kerndrup G, Johansson B. 2006. High incidence of the ETV6/RUNX1 fusion gene in paediatric precursor B-cell acute lymphoblastic leukaemias with trisomy 21 as the sole cytogenetic change: A Nordic series of cases diagnosed 1989–2005. Br J Haematol 135:352–354.

[83] Nguyen TT, Ma LN, Slovak ML, Bangs CD, Cherry AM, Arber DA. 2006. Identification of novel Runx1 (AML1) translocation partner genes SH3D19, YTHDf2, and ZNF687 in acute myeloid leukemia. Genes Chromosomes Cancer 45:918–932.

[84] Watanabe-Okochi N, Kitaura J, Ono R, Harada H, Harada Y, Komeno Y, Nakajima H, Nosaka T, Inaba T, Kitamura T. 2008. AML1 mutations induced MDS and MDS/AML in a mouse BMT model. Blood 111:4297–4308.

[85] Roche-Lestienne C, Deluche L, Corm S, Tigaud I, Joha S, Philippe N, Geffroy S, Lai JL, Nicolini FE, Preudhomme C. 2008. RUNX1 DNA-binding mutations and RUNX1-PRDM16 cryptic fusions in BCR-ABLþ leukemias are frequently associated with secondary trisomy 21 and may contribute to clonal evolution and imatinib resistance. Blood 111:3735–3741.

[86] Nguyen TT, Ma LN, Slovak ML, Bangs CD, Cherry AM, Arber DA. 2006. Identification of novel Runx1 (AML1) translocation partner genes SH3D19, YTHDf2, and ZNF687 in acute myeloid leukemia. Genes Chromosomes Cancer 45:918–932.

[87] Song WJ, Sullivan MG, Legare RD, Hutchings S, Tan X, Kufrin D, Ratajczak J, Resende IC, Haworth C, Hock R, Loh M, Felix C, Roy DC, Busque L, Kurnit D, Willman C, Gewirtz AM, Speck NA, Bushweller JH, Li FP, Gardiner K, Poncz M, Maris JM, Gilliland DG. 1999. Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. Nat Genet 23:166–175.

[88] Shinawi M, Erez A, Shardy DL, Lee B, Naeem R, Weissenberger G, Chinault AC, Cheung SW, Plon SE. 2008. Syndromic thrombocytopenia and predisposition to acute myelogenous leukemia caused by constitutional microdeletions on chromosome 21q. Blood 112:1042–1047.

[89] Hasle H. 2001. Pattern of malignant disorders in individuals with Down's syndrome. Lancet Oncol 2:429–436.

[90] Oncogene. 2010 Nov 18;29(46):6102-14. Epub 2010 Aug 9. Trisomic dose of several chromosome 21 genes perturbs haematopoietic stem and progenitor cell differentiation in Down's syndrome. De Vita S, Canzonetta C, Mulligan C, Delom F, Groet J, Baldo C, Vanes L, Dagna-Bricarelli F, Hoischen A, Veltman J, Fisher EM, Tybulewicz VL, Nizetic D.

[91] Yeoh EJ, Ross ME, Shurtleff SA, Williams WK, Patel D, Mahfouz R et al. Classification, subtype discovery, and prediction of out come in pediatric acute lymphoblastic leukemia by gene expression profiling. Cancer Cell 2002; 1: 133–143.

[92] Hardiman G. Microarray platforms – comparisons and contrasts. Pharmacogenomics 2004; 5: 487–502.

[93] Iqbal J, Liu Z, Deffenbacher K, Chan WC. Gene expression profiling in lymphoma diagnosis and management. Best Pract Res Clin Haematol 2009;22:191–210.

[94] Stoughton RB. Applications of DNA microarrays in biology. Annu Rev Biochem 2005; 74: 53–82.

[95] Southern E, Mir K, Shchepinov M. Molecular interactions on microarrays. Nat Genet 1999;21:5–9.

[96] Lipshutz RJ, Fodor SP, Gingeras TR, Lockhart DJ. High density synthetic oligonucleotide arrays. Nat Genet 1999;21:20–4.

[97] Tiu RV, Gondek LP, O'Keefe CL, et al. New lesions detected by single nucleotide polymorphism array-based chromosomal analysis have important clinical impact in acute myeloid leukemia. J Clin Oncol 2009;27:5219–26.

[98] Kawamata N, Ogawa S, Gueller S, et al. Identified hidden genomic changes in mantle cell lymphoma using high-resolution single nucleotide polymorphism genomic array. Exp Hematol 2009;37:937–46.

[99] Mullighan CG, Downing JR. Genome-wide profiling of genetic alterations in acute lymphoblastic leukemia: recent insights and future directions. Leukemia 2009;23:1209–18.

[100] Schadt, E. E., Li, C., Ellis, B. & Wong, W. H. Feature extraction and normalization algorithms for highdensity oligonucleotide gene expression array data. J. Cell Biochem. Suppl. 37, 120–125 (2001).

[101] Quackenbush, J. Microarray data normalization and transformation. Nature Genet. 32, 496–501 (2002).

[102] Bolstad, B. M., Irizarry, R. A., Astrand, M. & Speed, T. P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 19, 185–193 (2003).

[103] Mardia K, Kent J, Bibby J. Multivariate analysis. London: Academic Press; 1979.

[104] Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 1998;95: 14863–8.

[105] Allison DB, Cui X, Page GP, Sabripour M. Microarray data analysis: from disarray to consolidation and consensus. Nat Rev Genet 2006; 7: 55–65.

[106] Brown MP, Grundy WN, Lin D, Cristianini N, Sugnet CW, Furey TS, Ares M, Haussler D. Knowledge-based analysis of microarray gene expression data by using support vector machines. *Proc Natl Acad Sci USA.* 2000;**97**:262-267. doi: 10.1073/pnas.97.1.262.

[107] R. Tibshirani, T. Hastie, B. Narasimhan, G. Chu, Diagnosis of multiple cancer types by shrunken centroids of gene expression, Proc. Natl. Acad. Sci. USA 99 (2002) 6567–6572.

[108] Bioinformatics. 2005 Oct 15;21(20):3896-904. Epub 2005 Aug 16. Simple decision rules for classifying human cancers from gene expression profiles. Tan AC, Naiman DQ, Xu L, Winslow RL, Geman D.

[109] Allison DB, Cui X, Page GP, Sabripour M. Microarray data analysis: from disarray to consolidation and consensus. Nat Rev Genet 2006; 7: 55–65.

[110] Pan W. A comparative review of statistical methods for discovering differentially expressed genes in replicated microarray experiments. Bioinformatics 2002; 18: 546–554.

[111] Storey H. A direct approach to false discovery rates. J R Statist Soc Ser B 2002; 64 (Part 3): 479–498.

[112] Staal FJ, Cario G, Cazzaniga G, et al. Consensus guidelines for microarray gene expression analyses in leukemia from three European leukemia networks. Leukemia 2006;20:1385–92.

[113] Ioannidis JP. Is molecular profiling ready for use in clinical decision making? Oncologist 2007;12:301–11.

[114] Expression profiling – best practices for data generation and interpretation in clinical trials. Nat Rev Genet 2004; 5: 229–237.

[115] Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C et al. Minimum information about a microarray experiment (MIAME) – toward standards for microarray data. Nat Genet 2001; 29: 365–371.

[116] Larkin JE, Frank BC, Gavras H, Sultana R, Quackenbush J. Independence and reproducibility across microarray platforms. Nat Method 2005; 2: 337–344.

[117] Staal FJ, Cario G, Cazzaniga G, et al. Consensus guidelines for microarray gene expression analyses in leukemia from three European leukemia networks. Leukemia 2006;20:1385–92.

[118] Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999;286:531–7.

[119] Cancer Cell. 2002 Mar;1(2):133-43. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. Yeoh EJ, Ross ME, Shurtleff SA, Williams WK, Patel D, Mahfouz R, Behm FG, Raimondi SC, Relling MV, Patel A, Cheng C, Campana D, Wilkins D, Zhou X, Li J, Liu H, Pui CH, Evans WE, Naeve C, Wong L, Downing JR.

[120] Bhojwani D, Kang H, Menezes RX, et al. Gene expression signatures predictive of early response and outcome in high-risk childhood acute lymphoblastic leukemia: A Children's Oncology Group Study. J Clin Oncol 2008;26:4376–84.

[121] Willenbrock H, Juncker AS, Schmiegelow K, Knudsen S, Ryder LP. Prediction of immunophenotype, treatment response, and relapse in childhood acute lymphoblastic leukemia using DNA microarrays. Leukemia 2004;18:1270–7.

[122] Chiaretti S, Li X, Gentleman R, et al. Gene expression profile of adult T-cell acute lymphocytic leukemia identifies distinct subsets of patients with different response to therapy and survival. Blood 2004;103:2771–8.

[123] Haferlach T, Kohlmann A, Wieczorek L, et al. The clinical utility of microarraybased gene expression profiling in the diagnosis and subclassification of leukemia: Report from the International Microarray Innovations in Leukemia Study Group. J Clin Oncol 2010;28:2529–37.

[124] Ulrike Bacher , Alexander Kohlmann , Torsten Haferlach, Gene expression profiling for diagnosis and therapy in acute leukaemia and other haematologic malignancies, Cancer Treatment Reviews 36 (2010) 637–646.

[125] Haferlach T, Kohlmann A, Wieczorek L, et al. The clinical utility of microarraybased gene expression profiling in the diagnosis and subclassification of leukemia: Report from the International Microarray Innovations in Leukemia Study Group. J Clin Oncol 2010;28:2529–37.

[126] Haferlach T, Kohlmann A, Wieczorek L, et al. The clinical utility of microarraybased gene expression profiling in the diagnosis and subclassification of leukemia: Report from the International Microarray Innovations in Leukemia Study Group. J Clin Oncol 2010;28:2529–37.

[127] Hoffmann K, Firth MJ, Beesley AH, et al. Prediction of relapse in paediatric pre-B acute lymphoblastic leukaemia using a three-gene risk index. Br J Haematol 2008;140:656–64.

[128] Kang H, Chen IM, Wilson CS, et al. Gene expression classifiers for relapse-free survival and minimal residual disease improve risk classification and out come prediction in pediatric B-precursor acute lymphoblastic leukemia. Blood 2010;115:1394–405.

[129] Hofmann WK, de Vos S, Elashoff D, et al. Relation between resistance of Philadelphia-chromosome-positive acute lymphoblastic leukaemia to the tyrosine kinase inhibitor STI571 and gene-expression profiles: a geneexpression study. Lancet 2002;359:481–6.

[130] Holleman A, Cheok MH, den Boer ML, et al. Gene-expression patterns in drugresistant acute lymphoblastic leukemia cells and response to treatment. N Engl J Med 2004;351:533–42.

[131] A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide classification study. Den Boer ML, van Slegtenhorst M, De Menezes RX, Cheok MH, Buijs-Gladdines JG, Peters ST, Van Zutven LJ, Beverloo HB, Van der Spek PJ, Escherich G, et al. Lancet Oncol. 2009 Feb; 10(2):125-34. Epub 2009 Jan 8.

[132] Blood. 2010 Dec 2;116(23):4874-84. Epub 2010 Aug 10. Identification of novel cluster groups in pediatric high-risk B-precursor acute lymphoblastic leukemia with gene expression profiling: correlation with genome-wide DNA copy number alterations, clinical characteristics, and outcome. Harvey RC, Mullighan CG, Wang X, Dobbin KK, Davidson GS, Bedrick EJ, Chen IM, Atlas SR, Kang H, Ar K, Wilson CS, Wharton W, Murphy M, Devidas M, Carroll AJ, Borowitz MJ, Bowman WP, Downing JR, Relling M, Yang J, Bhojwani D, Carroll WL, Camitta B, Reaman GH, Smith M, Hunger SP, Willman CL.

[133] Kearney L, Horsley SW. Molecular cytogenetics in haematological malignancy: current technology and future prospects. Chromosoma 2005;114:286–94.

[134] Chen JS, Coustan-Smith E, Suzuki T, et al. Identification of novel markers for monitoring minimal residual disease in acute lymphoblastic leukemia. Blood 2001;97:2115–20.

[135] Steinbach D, Schramm A, Eggert A, et al. Identification of a set of seven genes for the monitoring of minimal residual disease in pediatric acute myeloid leukemia. Clin Cancer Res 2006;12:2434–41.

[136] Nature. 2005 Jun 9;435(7043):834-8. MicroRNA expression profiles classify human cancers.
Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA, Downing JR, Jacks T, Horvitz HR, Golub TR.

[137] BMC Med Genomics. 2011 Jan 17;4(1):8. [Epub ahead of print] Systems biology of interstitial lung diseases: integration of mRNA and microRNA expression changes. Cho JH, Gelinas R, Wang K, Etheridge A, Piper MG, Batte K, Dakhallah D, Price J, Bornman D, Zhang S, Marsh C, Galas D.

[138] Maher CA, Kumar-Sinha C, Cao X, et al. Transcriptome sequencing to detect gene fusions in cancer. Nature 2009;458:97–101.

[139] Walter MJ, Payton JE, Ries RE, et al. Acquired copy number alterations in adult acute myeloid leukemia genomes. Proc Natl Acad Sci USA 2009;106:12950–5.

[140] Haferlach T, Kohlmann A, Schnittger S, et al. Global approach to the diagnosis of leukemia using gene expression profiling. Blood 2005;106:1189–98.

[141] Bartel D: MicroRNAs: Genomics, biogenesis, mechanism, and function. Cell 116:281-297, 2004

[142] B.J. Reinhart, F.J. Slack, M. Basson, et al., The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans, Nature 403 (2000) 901–906.

[143] S.D. Hatfield, H.R. Shcherbata, K.A. Fischer, et al., Stem cell division is regulated by the microRNA pathway, Nature 435 (2005) 974– 978.

[144] A.M. Cheng, M.W. Byrom, J. Shelton, et al., Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis, Nucleic Acids Res. 33 (2005) 1290–1297.

[145] I. Alvarez-Garcia, E.A. Miska, MicroRNA functions in animal development and human disease, Development 132 (2005) 4653– 4662.

[146] P.S. Meltzer, Cancer genomics: small RNAs with big impacts, Nature 435 (2005) 745–746.

[147] A. Rodriguez, S. Griffiths-Jones, J.L. Ashurst, et al., Identification of mammalian microRNA host genes and transcription units, Genome Res. 14 (2004) 1902–1910.

[148] S. Hatfield, H. Ruohola-Baker, MicroRNA and stem cell function, Cell Tissue Res. 331 (2008) 57–66.

[149] Cancer Lett. 2011 Jan 1;300(1):10-9. Epub 2010 Oct 20. MicroRNAs, cancer and cancer stem cells. Zimmerman AL, Wu S.

[150] R.W. Carthew, E.J. Sontheimer, Origins and Mechanisms of miRNAs and siRNAs, Cell 136 (2009) 642–655.

[151] L. Wu, J. Fan, J.G. Belasco, MicroRNAs direct rapid deadenylation of mRNA, Proc. Natl. Acad. Sci. USA 103 (2006) 4034–4039.

[152] W. Filipowicz, S.N. Bhattacharyya, N. Sonenberg, Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?, Nat Rev. Genet. 9 (2008) 102–114.

[153] J.J. Rossi, RNAi and the P-body connection, Nat. Cell Biol. 7 (2005) 643–644.

[154] J. Coller, R. Parker, Eukaryotic mRNA decapping, Annu. Rev. Biochem. 73 (2004) 861–890.

[155] C.Z. Chen, L. Li, H.F. Lodish, et al., MicroRNAs modulate hematopoietic lineage differentiation, Science 303 (2004) 83–86.

[156] H. Bruchova, D. Yoon, A.M. Agarwal, et al., Regulated expression of microRNAs in normal and polycythemia vera erythropoiesis, Exp. Hematol. 35 (2007) 1657–1667.

[157] R.W. Georgantas 3rd, R. Hildreth, S. Morisot, et al., CD34+ hematopoietic stem-progenitor cell microRNA expression andfunction: a circuit diagram of differentiation control, Proc. Natl. Acad. Sci. USA 104 (2007) 2750–2755.

[158] Q. Wang, Z. Huang, H. Xue, et al., MicroRNA miR-24 inhibits erythropoiesis by targeting activin type I receptor ALK4, Blood 111 (2008) 588–595.

[159] L.C. Dore, J.D. Amigo, C.O. Dos Santos, et al., A GATA-1-regulated microRNA locus essential for erythropoiesis, Proc. Natl. Acad. Sci. USA 105 (2008) 3333–3338.

[160] N. Felli, L. Fontana, E. Pelosi, et al., MicroRNAs 221 and 222 inhibit normal erythropoiesis and erythroleukemic cell growth via kit receptor down-modulation, Proc. Natl. Acad. Sci. USA 102 (2005) 18081–18086.

[161] Cancer Lett. 2011 Jan 1;300(1):10-9. Epub 2010 Oct 20. MicroRNAs, cancer and cancer stem cells. Zimmerman AL, Wu S.

[162] J. Lu, G. Getz, E.A. Miska, et al., MicroRNA expression profiles classify human cancers, Nature 435 (2005) 834–838.

[163] S.M. Hammond, MicroRNAs as oncogenes, Curr. Opin. Genet. Dev. 16 (2006) 4–9.

[164] M.S. Kumar, J. Lu, K.L. Mercer, et al., Impaired microRNA processing enhances cellular transformation and tumorigenesis, Nat. Genet. 39 (2007) 673–677.

[165] G.A. Calin, C. Sevignani, C.D. Dumitru, et al., Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers, Proc. Natl. Acad. Sci. USA 101 (2004) 2999–3004.

[166] Lu J, Getz G, Miska E A, et al. MicroRNA expression profiles classify human cancers. Nature, 2005, 435(7043): 834—838

[167] Zanette DL, Rivadavia F, Molfetta GA, Barbuzano FG, Proto-Siqueira R, Silva- WA Jr, Falcao RP, Zago MA. miRNA expression profiles in chronic lymphocytic and acute lymphocytic leukemia. Braz J Med Biol Res 2007; 40:1435–40.

[168] Ventura A, Young AG, Winslow MM, et al. Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters. Cell 2008;132:875–86.

[169] Mi S, Lu J, Sun M, et al. MicroRNA expression signatures accurately discriminate acute lymphoblastic leukemia from acute myeloid leukemia. Proc Natl Acad Sci USA 2007;104:19971–6.

[170] Krutzfeldt J, Rajewsky N, Braich R, Rajeev KG, Tuschl T, Manoharan M, Stoffel M. Silencing of microRNAs in vivo with 'antagomirs'. Nature 2005;438:685–9.

[171] Jopling CL, Yi M, Lancaster AM, Lemon SM, Sarnow P. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. Science 2005;309:1577–81.

[172] Roman-Gomez J, Agirre X, Jimenez-Velasco A, et al. Epigenetic regulation of microRNAs in acute lymphoblastic leukemia. J Clin Oncol 2009;27:1316–22.

[173] Agirre X, Vilas-Zornoza A, Jimenez-Velasco A, et al. Epigenetic silencing of the tumor suppressor microRNA Hsa-miR-124a regulates CDK6 expression and confers a poor prognosis in acute lymphoblastic leukemia. Cancer Res 2009;69:4443–53.

[174] Stamatopoulos B, Meuleman N, Haibe-Kains B, Saussoy P, Van Den Neste E, Michaux L, Heimann P, Martiat P, Bron D, Lagneaux L. microRNA-29c and microRNA- 223 down-regulation has in vivo significance in chronic lymphocytic leukemia and improves disease risk stratification. Blood 2009;113:5237–45.

[175] MicroRNA patterns associated with clinical prognostic parameters and CNS relapse prediction in pediatric acute leukemia. Zhang H, Luo XQ, Zhang P, Huang LB, Zheng YS, Wu J, Zhou H, Qu LH, Xu L, Chen YQ. PLoS One. 2009 Nov 13;4(11):e7826.

[176] Dong,X.Y., Guo,P., Boyd,J., Sun,X., Li,Q., Zhou,W. and Dong,J.T. (2009) Implication of snoRNA U50 in human breast cancer. J. Genet. Genomics, 36, 447–454.

[177] Galasso M, Elena M Sana, Volinia S. Non-coding RNAs: a key to future personalized molecular therapy? Genome Med. 2010;2:12. doi: 10.1186/gm133.

[178] Bachellerie,J.P., Cavaille,J. and Huttenhofer,A. (2002) The expanding snoRNA world. Biochimie, 84, 775–790.

[179] Kiss,T. (2002) Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. Cell, 109, 145–148.

[180] Jady,B.E. and Kiss,T. (2001) A small nucleolar guide RNA functions both in 20-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. EMBO J., 20, 541–551.

[181] Balakin,A.G., Smith,L. and Fournier,M.J. (1996) The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions. Cell, 86, 823–834.

[182] Darzacq,X., Jady,B.E., Verheggen,C., Kiss,A.M., Bertrand,E. and Kiss,T. (2002) Cajal body-specific small nuclear RNAs: a novel class of 20-O-methylation and pseudouridylation guide RNAs. EMBO J., 21, 2746–2756.

[183] Bachellerie,J.P., Caffarelli,J. and Qu,L.H. (2000) Nucleotide modifications of eukaryotic rRNAs: the world of small nucleolar RNA guides revisited. The Ribosome: Structure, Function, Antibiotics and Cellular Interactions. ASM Press, Washington, DC, pp. 191–203.

[184] Weinstein,L.B. and Steitz,J.A. (1999) Guided tours: from precursor snoRNA to functional snoRNP. Curr. Opin. Cell Biol., 11, 378–384.

[185] Kiss-Laszlo,Z., Henry,Y., Bachellerie,J.P., Caizergues-Ferrer,M. and Kiss,T. (1996) Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. Cell, 85, 1077–1088.

[186] Bortolin,M.L., Ganot,P. and Kiss,T. (1999) Elements essential for accumulation and function of small nucleolar RNAs directing site-specific pseudouridylation of ribosomal RNAs. EMBO J., 18, 457–469.

[187] Ganot,P., Bortolin,M.L. and Kiss,T. (1997) Site-specific pseudo uridine formation in preribosomal RNA is guided by small nucleolar RNAs. Cell, 89, 799–809.

[188] Jady,B.E. and Kiss,T. (2001) A small nucleolar guide RNA functions both in 20-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. EMBO J., 20, 541–551.

[189] Bachellerie JP, Cavaille´ J, Hu¨ttenhofer A. 2002. The expanding snoRNA world. Biochimie. 84:775–790.

[190] Biochem Biophys Res Commun. 2002 Nov 29;299(2):196-200. Differential expression of human 5S snoRNA genes. Chang LS, Lin SY, Lieu AS, Wu TL.

[191] SnoRNA U50 is a candidate tumor-suppressor gene at 6q14.3 with a mutation associated with clinically significant prostate cancer. Dong XY, Rodriguez C, Guo P, Sun X, Talbot JT, Zhou W, Petros J, Li Q, Vessella RL, Kibel AS, Stevens VL, Calle EE, Dong JT. Hum Mol Genet. 2008 Apr 1;17(7):1031-42. Epub 2008 Jan 17.

[192] Dong,X.Y., Guo,P., Boyd,J., Sun,X., Li,Q., Zhou,W. and Dong,J.T. (2009) Implication of snoRNA U50 in human breast cancer. J. Genet. Genomics, 36, 447–454.

[193] Sahoo T, Del Gaudio D, German J, Shinawi M, Peters S, Person R, et al. Prader–Willi phenotype caused by paternal deficiency for the HBII-85 C/D box small nucleolar RNA cluster. Nat Genet 2008;40:719–721.

[194] Kishore S, Stamm S. The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. Science 2006;311:230–232.

[195] Shay JW, Wright WE, Werbin H. Defining the molecular mechanisms of human cell immortalization. Biochim Biophys Acta 1991; 1072: 1–7.

[196] Pulvertaft JV. Cytology of Burkitt's tumour (African lymphoma). Lancet 1964; 1: 238–240.

[197] Drexler HG, Matsuo AY, MacLeod RA. Continuous hematopoietic cell lines as model systems for leukemia-lymphoma research. Leuk Res 2000; 24: 881–911.

[198] Kawabe T, Muslin AJ, Korsmeyer SJ. HOX11 interacts with protein phosphatases PP2A and PP1 and disrupts a G2/M cellcycle checkpoint. Nature 1997; 385: 454–458.

[199] Palacios R, Steinmetz M. Il-3-dependent mouse clones that express B-220 surface antigen, contain Ig genes in germ-line configuration, and generate B lymphocytes in vivo. Cell 1985; 41: 727–734.

[200] Daley GQ, Baltimore D. Transformation of an interleukin 3-dependent hematopoietic cell line by the chronic myelogenous leukemia-specific P210bcr/abl protein. Proc Natl Acad Sci USA 1988; 85: 9312–9316.

[201] Warmuth M, Kim S, Gu XJ, Xia G, Adrian F. Ba/F3 cells and their use in kinase drug discovery. Curr Opin Oncol 2007; 19: 55–60.

[202] Drexler HG, MacLeod RA. Leukemia-lymphoma cell lines as model systems for hematopoietic research. Ann Med 2003; 35: 404–412.

[203] Elefanty AG, Cory S. bcr-abl-Induced cell lines can switch from mast cell to erythroid or myeloid differentiation in vitro. Blood 1992; 79: 1271–1281.

[204] Heisterkamp N, Jenster G, ten Hoeve J, Zovich D, Pattengale PK, Groffen J. Acute leukaemia in bcr/abl transgenic mice. Nature 1990; 344: 251–253.

[205] Adams JM, Harris AW, Pinkert CA, Corcoran LM, Alexander WS, Cory S et al. The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. Nature 1985; 318: 533–538.

[206] Forster A, Pannell R, Drynan LF, Codrington R, Daser A, Metzler M et al. The invertor knock-in conditional chromosomal translocation mimic. Nat Methods 2005; 2: 27–30.

[207] Daser A, Rabbitts TH. The versatile mixed lineage leukaemia gene MLL and its many associations in leukaemogenesis. Semin Cancer Biol 2005; 15: 175–188.

[208] Lavau C, Szilvassy SJ, Slany R, Cleary ML. Immortalization and leukemic transformation of a myelomonocytic precursor by retrovirally transduced HRX-ENL. EMBO J 1997; 16: 4226–4237.

[209] Pantelouris E (1971) Observations on the immunobiology of nude mice. Immunol 20:247

[210] Andreiole G, Mule J, Hansen C, Linehan W, Rosenberg S (1985) Evidence that lymphokine-activated killer cells and natural killer cells are distinct based on analysis of congenitally immunodeficient mice. J Immunol 135:2911

[211] Bosma M, Carroll M (1991) The SCID mouse mutant: definitions, characterization, and potential uses. Annu Rev Immunol 9:323

[212] Shultz L, Schweltzar P, Christianson S, Goll B, Schweltzer I, Tennent B, McKenna S, Mobraaton L, Rajan T, Greiner D, Leiter E (1995) Multiple defects in innate and adaptive immunological function in NOD/LtSz-scid mice. J Immunol 154:180–191

[213] Lowary P, Schultz L, Quesenberry P (1994) Multiple immune defects in the NOD/SCID mouse facilitate human hematopoietic engraftment. Blood 84 [Supp 1]:346°

[214] Dick J (1996) Normal and leukemic human stem cells assale in SCID mice. Semin Immunol 3:367–378

[215] Lapidot T, Fajerman Y, Kollet O. Immune-deficient SCID and NOD/SCID mice models as functional assays for studying normal and malignant human hematopoiesis. J Mol Med. 1997;75(9):664–673. doi: 10.1007/s001090050150.

[216] Kamel-Reid S, Letarte M, Sirard C, Doedens M, Grunberger T, Fulop G et al. A model of human acute lymphoblastic leukemia in immune-deficient SCID mice. Science 1989; 246: 1597–1600.

[217] Lapidot T, Sirard C, Vormoor J, Murdoch B, Hoang T, Caceres- Cortes J et al. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. Nature 1994; 367: 645–648.

[218] Uckun FM, Waurzyniak BJ, Sather HN, Sensel MG, Chelstrom L, Nachman J et al. Prognostic significance of T-lineage leukemic cell growth in SCID mice: a Children's Cancer Group study. Leuk Lymphoma 1999; 32: 475–487.

[219] Pearce DJ, Taussig D, Zibara K, Smith LL, Ridler CM, Preudhomme C et al. AML engraftment in the NOD/SCID assay reflects the outcome of AML: implications for our under standing of the heterogeneity of AML. Blood 2006; 107: 1166–1173.

[220] Kamel-Reid S, Letarte M, Sirard C, Doedens M, Grunberger T, Fulop G, Freedman MH, Phillips RA, Dick JE: A model of human acute lymphoblastic leukemia in immune-deficient SCID mice. Science 1989, 246(4937):1597-1600.

[221] Nijmeijer BA, Mollevanger P, van Zelderen-Bhola SL, Kluin-Nelemans HC, Willemze R, Falkenburg JH: Monitoring of engraftment and progression of acute lymphoblastic leukemia in individual NOD/SCID mice. Exp Hematol 2001, 29(3):322-329.

[222] Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nat Med 1997; 3: 730–737.

[223] Kelly PN, Dakic A, Adams JM, Nutt SL, Strasser A. Tumor growth need not be driven by rare cancer stem cells. Science 2007; 317: 337.

[224] Feuring-Buske M, Gerhard B, Cashman J, Humphries RK, Eaves CJ, Hogge DE. Improved engraftment of human acute myeloid leukemia progenitor cells in beta 2-microglobulin-deficient NOD/SCID mice and in NOD/SCID mice transgenic for human growth factors. Leukemia. 2003;17(4):760–763. doi: 10.1038/sj.leu.2402882.

[225] Shultz LD, et al. Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R gamma null mice engrafted with mobilized human hemopoietic stem cells. *J Immunol.* 2005;**174**(10):6477–6489.

[226] Ishikawa F, et al. Development of functional human blood and immune systems in NOD/SCID/IL2 receptor {gamma} chain(null) mice. Blood. 2005;106(5):1565–1573. doi: 10.1182/blood-2005-02-0516.

[227] Muguruma Y, Yahata T, Miyatake H, Sato T, Uno T, Itoh J et al. Reconstitution of the functional human hematopoietic microenvironment derived from human mesenchymal stem cells in the murine bone marrow compartment. Blood 2006; 107:
1878–1887.

[228] Felsburg PJ. Overview of immune system development in the dog: comparison with humans (Review). *Human & Experimental Toxicology.* 2002;21:487–492.

[229] Tsai KL, Clark LA, Murphy KE. Understanding hereditary diseases using the dog and human as companion model systems (Review). *Mamm Genome.* 2007;18:444–451

[230] Joy F, Basak S, Gupta SK, Das PJ, Ghosh SK, Ghosh TC. Compositional correlations in canine genome reflects similarity with human genes. *Journal of Biochemistry and Molecular Biology.* 2006;39:240–246

[231] Zon LI, Peterson RT. In vivo drug discovery in the zebrafish. Nat Rev Drug Discov. 2005;4(1):35-44.

[232] Smith AC, Raimondi AR, Salthouse CD, Ignatius MS, Blackburn JS, Mizgirev IV, et al. High-throughput cell transplantation establishes that tumor-initiating cells are abundant in zebrafish T-cell acute lymphoblastic leukemia. Blood. 2010;115(16):3296-303.

[233] Sabaawy HE, Azuma M, Embree LJ, Tsai HJ, Starost MF, Hickstein DD. TEL-AML1 transgenic zebrafish model of precursor B cell acute lymphoblastic leukemia. Proc Natl Acad Sci U S A 2006;103:15166–15171. [PubMed: 17015828] 35. Sambrook JG, Figueroa F, Beck S. A genome-wide

[234] Mizgirev IV, Revskoy S. A new zebrafish model for experimental leukemia therapy. Cancer Biol Ther. 2010;9(11):895-902.

[235] Frazer JK, Meeker ND, Rudner L, Bradley DF, Smith AC, Demarest B, et al. Heritable T-cell malignancy models established in a zebrafish phenotypic screen. Leukemia. 2009

[236] Darley RL, Hoy TG, Baines P, Padua RA, Burnett AK. Mutant N-RAS induces erythroid lineage dysplasia in human CD34+ cells. J Exp Med 1997; 185: 1337–1347.

[237] Wang JC, Doedens M, Dick JE. Primitive human hematopoietic cells are enriched in cord blood compared with adult bone marrow or mobilized peripheral blood as measured by the quantitative in vivo SCID-repopulating cell assay. Blood 1997; 89: 3919–3924.

[238] Levine RL, Gilliland DG. JAK-2 mutations and their relevance to myeloproliferative disease. Curr Opin Hematol 2007; 14: 43–47.

SCOPE OF THIS THESIS

Among the first genome-wide technologies to be developed, microarray-based gene transcript profiling is extensively applied in bio-medical research. Since ten years, its application in oncohematology aids researchers to deeply investigate the pathobiology of hematologic malignancies.

Overall survival in pediatric B cell precursor acute lymphoblastic leukemia (BCP ALL) exceeds 80% on contemporary treatment regimens, nevertheless, the identification of better prognostic factors and new therapeutic targets remains an important challenge.

This thesis presents the application of gene expression profiling (GEP) for the study of BCP ALL.

To improve biological knowledge of B-others, a group of BCP ALL poor in parameters for risk assessment and lacking gross genetic aberration, an integrated study of gene and non coding RNAs expression and genetic aberrancies was performed. In this study I searched for new subgroups characterization, evidence of deregulated pathways and potentially new prognostic factors.

One of the focal points of my study is the role of chromosome 21 (Hsa21) in the development of BCP ALL. GEP was performed on a large group of ALL in patients with Down Syndrome (DS ALL). The Study aimed to improve knowledge of the DS ALL subgroup, and to compare its signature with that of non DS ALL, particularly hyperdiploid ALL with acquired trisomy 21.

Given the increasing application of mouse xenotransplantation models for *in vivo* testing and amplification of leukemia cells, GEP was employed to asses whether xenograft samples recapitulate their respective primary leukemia. Finally, GEP of xenograft samples was used to investigate biologic features of BCP ALL associated to engraftment properties and clinical parameters of the patients.

The thesis will discuss the findings of this research approaches and their impact on BCP ALL management.

CHAPTER 1

MiR-125b-2 cluster is over expressed in a subgroup of pediatric BCP ALL associated with a unique gene expression signature and favourable outcome

Elena Vendramini, Marco Giordan, Marta Galbiati, Antonella Lettieri, Silvia Bungaro, Daniela Silvestri, Luca Trentin, Emanuela Giarin, Barbara Michielotto, Anna Leszl, Maria Grazia Valsecchi, Gianni Cazzaniga, Giuseppe Basso, Geertruy te Kronnie.

UNDER SUBMISSION

ABSTRACT

In the pediatric B cell precursor acute lymphoblastic leukemia (BCP ALL) scenery B-others represent the subgroup needy of deeper investigations, to find new prognostic factors and improvement of risk stratification. Gene expression profile analysis identified among B-others a group with a unique signature, recently associated to favourable outcome and characterized by the frequent occurrence of *ERG* intragenic deletions.

In this study we profiled by gene expression arrays 145 Italian B-others BCP ALL pediatric patients and found a unique signature in 24.8% of patients. This group shared a CRI of 12,3% compared to 25,6% for the rest of the BCP ALL cohort. A GEP classifier was developed to identify the patients with favourable outcome and reached a sensitivity of 0.98 and a specificity of 1. MicroRNAs expression profiling of this group revealed an unique miRNAs signature characterized by over expression of hsa-miR-125b, -125b-2*, -99a, -100, -125a-3p and has-miR-491-5p. Over expression of cluster miR-125b-2 in region 21q21.1 goes along with over expression of genes in the same chromosomal region, suggesting a direct involvement of this region in the disease pathobiology. Genome-wide analysis excluded copy number alteration of the 21q21.1 region.

Our study adds new insight in the subgroup of B-others with favourable outcome on deregulated genes and microRNAs and point to a particular role for 21q21.1 in leukemia.

INTRODUCTION

Twenty-five percent of B cell precursor acute lymphoblastic leukemia (BCP ALL), defined B-others, lack genetic determinants of outcome[1].

Gene expression profiling proved to be efficacious in subtype discovery[2] and several subgroups were identified among B-others and associated to different outcome, suggesting intrinsic biological heterogeneity of B-others related to clinical behaviour[3, 4, 5].

Fortunately, amelioration of prognostic tools independent from genetic features, as the detection of minimal residual diseases at day 15, improved risk stratification of BCP ALL overall[6].

In line with the somatic mutation theory of cancer, genome-wide analysis of genetic alterations in ALL recently lead to finding of submicroscopic abnormalities, also in B-others, that disrupt genes involved in B-cell development and play an important role in leukemogenesis[7]. Frequently affected were genes involved in the B-cell developmental pathway such as *PAX5, TCF3, EBF1, LEF1, IKZF1* and *IKZF3* or genes controlling cell cycle progression such as *CDKN2A*, *CDKN1B* and *RB1*[8]. Recently reported is also the *P2RY8-CRLF2* fusion, caused by a cryptic translocation involving *IGH@* and *CRLF2* in the pseudoautosomal region (PAR1) of the sex chromosomes or by a deletion within PAR1, and leading to overexpression of *CRLF2*[9, 10, 11, 12].

Genetic lesion are also reported to involve non coding RNAs (ncRNAs) loci leading to disruption of their regulative functions. MicroRNA (miRNA) genes are frequently located at fragile sites and regions of copy number alteration (CNA) associated with cancer in general[13], and a recent work estimated that ~70% of miRNAs are located in regions of leukemia-associated cytogenetic changes[14]. Therefore, there is increasing oncology research to study miRNAs and other classes of ncRNAs such as small nucleolar RNAs (snoRNAs) and the close related small Cajal body-specific RNAs (scaRNAs).

MiRNAs, single stranded RNAs of 18-24 nucleotides, are known to regulate gene expression of targets mRNA[15]. MiRNAs are reported to play fundamental roles in hematopoiesis, hematopoietic malignancies acting as oncogenes or tumor suppressor genes[16] and to be involved in epigenetic regulation in ALL[17]. MiRNAs are increasing their importance in cancer management given they potential as molecular markers for cancer classification, prognostic stratification and drug-response prediction[18].

SnoRNAs and scaRNAs, 60-300 nucleotide long RNAs[19], are known to guide nucleotide modifications respectively in ribosomal RNAs and small nuclear RNAs and snoRNA-related machinery has already been implicated in human diseases and cancer [20].

To improve biological knowledge of B-others we performed an integrated study of gene and ncRNAs expression and genetic aberrancies. One hundred forty-five patients with diagnosis of BCP ALL lacking known genetic aberrations were profiled by mRNA and in a representative subcohort of patients we performed miRNAs expression profiling and genome-wide DNA copy number variation (CNV) analysis. We focused on a group of patients with a unique signature previously described[21], and reported to be associated to frequent *ERG* intragenic deletions[22] and favorable outcome[23].

PATIENTS AND METHODS

Cohort of patients

A cohort of 145 paediatric patients new diagnosed of acute lymphoblastic leukemia between 2002 and 2006 were included in the study. Patients were routinely tested for recurrent genomic aberrations and DNA index of blast cells and treated according with AIEOP ALL 2000 therapeutic protocol. The local ethics committees approved the study and informed consent was obtained for all patients.

RNA and DNA preparation

DNA and RNA were isolated from bone marrow or peripheral blood mononuclear cells separated by Ficoll-Hypaque technique (Pharmacia, Uppsala, Sweden), extraction were performed either from fresh cells at time of diagnosis or from stored frozen diagnostic material. DNA were isolated by Puregene Cell and Tissue Kit (Qiagen Inc., Valencia, USA). Total RNA were isolated by TRIZOL following manufacturer instruction (Invitrogen, Paisley, UK), quality control was done with the 2100 Bioanalyzer using “Eukaryote total RNA Nano Assay” (Agilent Technologies). All material were stored at -80°C.

Genes and miRNAs expression Arrays

Gene expression profile were performed on HG-U133 Plus 2.0 GeneChip® (Affymetrix, Santa Clara, CA, USA). First group of patients (103) were processed for MILE study as previously

described[24]. Second group of patients (42) were processed starting from 100 ng of total RNA using GeneChip® 3'IVT express kit and protocol (Affymetrix, Santa Clara, CA, USA).
MiRNAs expression profile were performed on Mirna array 1.0 GeneChip® (Affymetrix, Santa Clara, California). This array provides a total of 7815 probe sets representing miRNAs of 71 organisms (847 human miRNAs) annotated in Mirbase v.11 and 922 snoRNAs and scaRNAs sequences. Total RNA (1µg) were labelled using FlashTag™ kit (Genisphere, Hatfield, PA) that use 3DNA Dendrimer technology.
For both platforms hybridization, staining and washing were performed using protocols as recommended by the manufacturer, stained chip were scanned on GeneChip Scanner 3000 7G (Affymetrix, Santa Clara, California). Expression files (Affymetrix .CEL files) were generated using the GCOS 1.4 or 1.2 and Affymetrix® GeneChip® Command Console® Software (Affymetrix).

Affymetrix® Cytogenetics Whole Genome 2.7M Array Analysis

Diagnostic samples and their corresponding remission samples were genotyped with Affymetrix® Cytogenetics Whole Genome 2.7M Array (Affymetrix®, Santa Clara, CA). This array provides dense coverage, across the entire genome, of the annotated genes, cancer genes, miRNA regions, and haploinsufficiency genes thanks 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 heterozygosity (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-labelled, denatured and loaded into the arrays. After hybridization the chips were than washed, stained (streptavidin-PE) and scanned using the Gene Chip® Scanner 3000.
CEL files were generated using Affymetrix® GeneChip® Command Console™ (AGCC) v3.1 and analyzed by Chromosome Analysis Suite (Affymetrix®, Santa Clara, CA) software.

Quantitative assay of mature microRNA

MiRNAs expression data were validated measuring mature form amount of five interesting miRNAs by TaqMan® MicroRNA assays (Applied Biosystems, Foster City, CA, USA).

MiRNAs tested were hsa-miR-125b (Assay ID 000449), hsa-miR-125b-2* (Assay ID 002158), hsa-miR-99a (Assay ID 000435), hsa-miR-let-7c (Assay ID 000379), hsa-miR-155 (Assay ID 002623). Control miRNA Assay U6 (RUN6B) was assessed as endogenous control (Assay ID 001093). cDNA was generated by TaqMan® MicroRNA Reverse Transcription (RT) Kit (Applied Biosystems, Foster City, CA, USA) starting from 10 to 50 ng of total RNA. PCR was carried out with the 7900HT Fast Real Time PCR system (Applied Biosystems, USA). Relative expression was calculated using the comparative CT method[25] and using RNA sample from CD19+ cells (sorted from healthy donor bone marrow) as calibrator. Graphpad Prism version 4 software program was used for analyses (GraphPad Software, La Jolla, CA, USA).

Data analysis

Arrays have been normalized using robust multiple-array average (RMA)[26]. When data belonged to different protocols batch effects were removed using Combat[27]. Unsupervised analyses were based on hierarchical clustering (with Euclidean distance and Ward's method). Heatmaps were used to highlight the associations between the clusterings and the expression levels of the genes. The shrinkage approach[28] was used to assess differences in gene expression levels between two groups of interest, using local false discovery rate as method to control false positives. When this approach had been believed inaccurate[29, 30] we used a permutation approach on filtered probe sets (filtering out probe sets with small variance across samples; 90% of the probe sets removed) with tests based on standardized rank sum Wilcoxon statistics and we control false positives with the method of Benjamini and Hochberg[31]. Results from these two approaches were considered significant if they had local false discovery rate <0.05 or adjusted p-value < 0.05, respectively. Fold change of probe sets was calculated as the ration between the mean expression in Group 1and the mean expression in Group 2.

To understand if Group 1 patients can be distinguished from B-others patients on the basis of their gene expression profiles a classifier was adapted on the whole cohort. We used LASSO[32] as prediction method. The method is implemented in the package CMA[33] which is available through bioconductor (www.bioconductor.org). Prediction accuracy was assessed using 5 fold cross validation (10 iterations) with stratified sampling. Hyperparameter tuning was carried out using an inner loop of 3 fold cross validation.

To identify putative miRNA-mRNA functional pairs, we integrated miRNA and mRNA expression data using Partek® Genomics Suite™ software, version 6.5 Copyright © 2010 (Partek Inc., St.Louis, MO, USA). The "combine microRNA with they mRNA targets" function were applied to identify the predicted target of most deregulated miRNAs between the most deregulated mRNA. TargetScan (Release 5.1)[34] and MicroCosm were choosen as algorithms for miRNA target prediction. For genes that were up-regulated in Group 1 patients, we searched for targeting miRNAs that were down-regulated, and *vice versa*. mRNA target predicted by both algorithms were further confirmed in a third prediction algorithm Pictar[35].

Representation of Gene Ontology (GO) terms was examined using the Onto-Express tool[36], functional gene interactions were investigated using Pathway-Express tool (http://vortex.cs.wayne.edu/projects.htm) and Gene Set Enrichment Analysis[37] (http://www.broadinstitute.org/gsea/index.jsp).

Presence of putative CpG islands at 5' of interesting genes were investigated by use of CpG Plot software (http://www.ebi.ac.uk/Tools/emboss/cpgplot/index.html) using following setting: 300bp for scan window, 50 for %GC threshold , 0.6 for Obs/Exp ratio.

Statistical analysis

Cumulative Relapse Incidence (CRI) were estimated by the Kaplan-Meier method; the starting point of the observation time was the date of diagnosis, first relapse were considered as event. Graphpad Prism version 4 software program was used for analyses (GraphPad Software, La Jolla, CA, USA). To assess associations between patients features, Chi-Square test was applied and corrected for multiplicity by Bonferroni-Holm method.

RESULTS

Study cohort

We have studied 145 specimens at diagnosis of BCP ALL patients that lack known genomic aberrations or hyperdiploid karyotype (DNA index< 1,16) named B-others[38]. Patients in the cohort (53,8% male, 46.2% female) were prevalently younger than ten years (74.5%) and presenting low WBC count (74.2%). Most patients were classified as good prednison responders (82.4%), had intermediate risk group classification for MRD at day 15 and day 78 (44.7% and 61% respectively) and were assigned to intermediate risk protocol strata (52.4%)

(Tab.1). Thirty-one patients (21.4%) experienced one or more relapse events. Most of relapse events occurred in patients with intermediated risk classification to indicate the need of new prognostic factors to improve risk assessment in B-others BCP ALL patients.

Identification of unique signature among B-others

To improve molecular biological knowledge of B-others BCP ALL we performed gene expression profiling of an initial cohort of 103 patients.

In a class discovery approach, unsupervised hierarchical cluster analysis identified a subgroup of patients (24.3%, Group 1) that clustered separately from the rest of the cohort (Group 2). To explore the distinction between these two groups both groups were compared to genetically defined BCP ALL subtypes using respective gene expression signatures (Table S-IV, Den Boer et al., 2009). In Group 2 patients high heterogeneity was observed, identifying patients with *BCR-ABL1*-like, *TEL-AML1*-like or hyperdiploid-like signatures (Suppl. Fig 1). Contrary, Group 1 patients didn't share signatures of genetically defined subtypes suggesting distinct biological features in this group.

To confirm the recurrence of patients with Group 1 features among B-others an additional set of 42 patients was profiled. Again, unsupervised hierarchical cluster analysis divided patients in a smaller cluster of 26.2% (11) patients (Group 1'), and a larger cluster with remaining patients (Group 2'). Unsupervised hierarchical cluster analysis on the merge of the two cohorts (145 patients) resulted in two clusters where all Group 1 and Group 1' patients cluster together (Group 1''), revealing to belong to the same biological subtype, distinct from the other patients (Group 2'').

Class comparison analysis, using shrinkage T statistic, between groups in the first cohort, in the second cohort and in the merged cohort (Group1, 1' or 1''vs. Group 2, 2' or 2'' respectively), generated three gene lists that shared the majority of probe sets 254. (49 most deregulated probe sets are listed in Suppl. Tab.1). This uniformity of gene expression levels of Group 1 and Group 1' patients, allowed to consider them as an unique subtype (referred as Group 1 in the further analysis), distinct from the other patients (referred as Group 2 in the further analysis), with an overall incidence of 24.8% (36 out 145) among the Italian BCP ALL B-others cohort (Fig.1).

The Group 1 signature shared several probe sets with a novel ALL subgroup (named "cluster 6") recently described by Harvey and colleagues in a cohort of high-risk pediatric BCP ALL[39]. Fifty-two out of 100 most deregulate probe sets in cluster 6 signature (Table S14,

Harvey et al, 2010) [40] are also found in our Group 1 signature (254 probe sets). Only a fifth (22) were found between the 49 most deregulated (fold change >1.5 or <1/1.5) probe sets in the Group 1 signature.

Development of a classifier for Group 1 patients

All 145 patients were used to construct a classifier to identify Group 1 patients. Applying the LASSO on all the available probe sets we identified a classifier of 79 probe sets (Tab.2). The discriminative ability of the classifier was assessed using the misclassification rate as performance measure. Considering the outer loop cross validation we obtained a misclassification rate of 0.0062 (sensitivity 0.98, and specificity 1). All patients were correctly identified using this classifier.

Gene ontology analysis

Gene ontology of Group 1 up-regulated genes revealed enrichment in structural constituents of ribosomes, RNA and protein binding and insulin receptor binding among *Molecular Functions* categories; traslational elongation, rRNA processing and insulin-like growth factor receptor signalling pathway among *Biological Process* categories; Ribosome and Nucleus among *Cellular Component* categories. Group 1 down-regulated genes revealed enrichment in purinergic nucleotide receptor activity G-protein coupled, SH3 domain binding and GTPase activator activity among *Molecular Functions* categories; puscle development and positive regulation of I-kappaB kinase/NF-kappaB cascade among *Biological Process* categories; Plasma membrane, Integral to plasma membrane, Cytoskeleton, Extracellular region and Endoplasmic reticulum membrane among *Cellular Component* categories.

Group 1 patients features and outcome in Italian cohort

Compared to Group 2, Group 1 had a prevalent female composition (63,9% vs. 40.4%), prevalence of low WBC count (88.8% vs. 69.9%), more intermediate risk stratified patients for MRD at day 15 (56.6% vs. 40.8%) and day 78 (71.4% vs. 57.4%) and more patients assigned to intermediate risk protocol strata (63.9% vs. 48.6%)(Tab.1b). However, Chi-Square test failed to found significant association between Group 1 and other clinical parameters (WBC, age, gender, prednison response, MRD stratification and risk stratification).

Cumulative relapse incidence (CRI) of Group 1 patients is lower then in rest of the cohort (12.3%±5.9 vs. 25.6%±4.2, 4/36 vs.27/109 relapses) showing an important trend but not statistically significant (Fig.2a). Comparing Group 1 patients with the three risk strata of Group 2 the CRI differences are significant (Fig.2b). Apparently Group 1 patients show a clinical behaviour similar to standard risk patients even if they were originally majorly stratified as intermediate risk.

Group 1 patients share a microRNA signature with overexpression of the miR-125b-2 cluster

MiRNAs expression profile analysis was performed on 24 BCP ALL patients of the study cohort, 8 Group 1 patients and 16 Group 2 patients. The miRNAs expression data refer to both mature and immature forms of miRNAs in the samples. Unsupervised cluster analysis on 847 human microRNAs probe sets clustered Group 1 patients separately from Group 2 patients (Fig.3a). Class comparison analysis (Wilcoxon) resulted in a list of 18 deregulated miRNAs (adjusted p-value <0.05) (Tab.3a, Fig.3c). All miRNAs belonging to miR-125b-2 cluster (has-miR-125b, -125b-2*, -99a, let-7c) were over-expressed in Group 1 patients, all but one (miR-let-7c) had a fold change >1.5. Quantitative RT-PCR confirmed the overexpression of the mature forms of miRNAs belong to the miR-125b-2 cluster in Group 1 patients (Suppl. Fig 2a.).

Other top ranked miRNAs were hsa-miR-100, hsa-miR-125a-3p and has-miR-491-5p up-regulated in Group 1, with a fold change greater then 1.5, and has-miR-132* down-regulated in Group 1 with a fold change smaller then 1/1.5 (not assayed by qRT-PCR).

Integration of miRNA and mRNA data using three distinct prediction algorithms separately, allowed to find several targets of deregulated miRNAs among deregulated mRNAs. The merge of results from the 3 algorithms agreed on one prediction: *UGP2* (UDP-glucose pyrophosphorylase 2), down-regulated in Group 1 was predicted to be target of miR-92a (up-regulated in Group1).

Interestingly, unsupervised cluster analysis on 922 human snoRNAs and scaRNAs probe sets showed tight clustering of Group 1 patients (Fig.3b). A class comparison analysis detected a list of 103 differentially expressed probe sets (adjusted p-value < 0.05, 20 probe sets with fold change >1.5) most of thse probe sets are located in the 15q11.2 region (Tab.3b, Fig.3d). Although the function of these non-coding RNA molecules is largely unknown, this finding suggests a considerable role in the biology of leukemic cells.

21q21.1 region with a specific regulation in Group 1 patients

We focused our attention on chromosome 21 where the miR-125b-2 cluster host gene C21ORF34 is mapped. Parallel GEP analyses on patients investigated for miRNAs expression using class comparison analysis between Group 1 and Group 2 on 672 probe sets located on chromosome 21 identified several differentially expressed genes (18 probe sets, adjusted p-value <0.05, 3 probe sets with fold change >1.5). Among up-regulated genes in Group 1 patients were identified: *C21ORF34*, *PRSS7* and *C21ORF91* mapping close to each other into a 2.5 Mb region in 21q21.1. This finding suggests a specific involvement of 21q21.1 region in the pathobiology of Group 1 leukemias (Fig.4, Tab.4).

Genomic region upstream *C21ORF34* annotated sequence (ENSG00000215386) was investigated for the presence of putative CpG islands using CpGPlot software (http://www.ebi.ac.uk/Tools/emboss/cpgplot/index.html). Only one putative CpG site 216 bp long and localized 50kb up-stream of *C21ORF34* was found. Size and position of this CpG site discourage the hypothesis of an epigenetic regulation via DNA methylation of the *C21ORF34* locus, however further investigations might be considered to completely exclude such regulation.

MIR155HG (MIR155 host gene) resulted down-regulated in Group 1 patients compared to group 2 but qRT-PCR analysis revealed no difference in expression of the mature form of miR-155 between the two groups (suppl. Fig.2b).

Genome-wide analysis identified ERG deletions among Group 1 patients

Genome-wide copy number analysis was performed on 8 Group 1 patients to investigate the genomic features specifically related to the genes and miRNAs expression profiles.

No aberrations in *C21ORF34* locus in 21q21.1 region were found. Deletions involving *IKZF1* (2/8) and *CDKN2A/B* (2/8) locus were found (Suppl. Tab.2).

In 4 out 8 patients (50%) an intragenic deletion on *ERG* (v-ets erythroblastosis virus E26 oncogene homolog) was found. Deletions, always found in heterozygosis, extended for about 50 Kb, included exons 6-11 (transcript variant 2, NM_004449) of *ERG* (Suppl. Fig.3) This kind of *ERG* intragenic deletion was earlier described by Mullighan and colleagues in their study using Affymetrix GeneChip Human Mapping 50K Hind 240, 50K Xba 240 and 250K Sty arrays[41, 42]. Our finding is in accordance with Harvey et al. reporting 8 *ERG* deleted

patients out of 21 (38%) on a similar group[43]. No differences in *ERG* expression were observed between Group 1 and Group 2 patients.

ERG intragenic deletion is associated to distinct gene expression profile

In a class comparison analysis comparing Group 1 patients sharing *ERG* deletion and Group 1 patients without *ERG* deletion we found a sub-signature (440 probe sets, lfdr<0.05, 56 probe sets with fold change >1.5 or <1/1.5) (Suppl. Tab.3).

Gene set enrichment analysis demonstrated down-regulation, in patients with *ERG* intragenic deletion, of genes involved in immune systems, signal transduction and regulation of apoptosis. Furthermore, down-regulated genes were enriched for gene sets down-regulated in hematopoietic and leukemic stem cells and gene sets down-regulated in *RUNX1-RUNX1T1* fusion gene (t(8;21)(q22;q22)). Since *RUNX1-RUNX1T1* fusion gene promotes an increase in stem/early progenitors by favouring self-renewal over differentiation[44] the genes enrichment suggest a more immature features of blast cells in patients with the *ERG* intragenic deletion.

DISCUSSION

The occurrence of relapse events in BCP ALL B-others patients initially stratified into low-intermediate risk groups underscore the need to better understand the biology of the disease and the urgency of new prognostic factors to improve the risk assessment.

In this study gene expression profiling permitted to identify among B-others a group of patients (Group 1) with a unique signature distinct from those of genetically defined BCP ALL subtypes.

Signature was similar to that of a BCP ALL group recently associated to frequent *ERG* intragenic deletions and to a favorable outcome, studied among a cohort of high risk stratified BCP ALL patients[45]. In our study, Group 1 patients had heterogeneous risk classification and most of them belong to the intermediate risk class. Given the absence of an association with a specific risk strata, we developed a more appropriate approach to investigate Group 1 features in the contest of an unselected B-others cohort. Indeed, the COG and AIEOP study show only 20% overlap among the most deregulated probe sets of the two signature likely to be due to major differences in cohort composition.

In two independent Italian cohorts, Group 1 counted for 25% of B-others (5% of B ALL) and was associated to a favorable outcome (CRI 12.3%±5.9). A classifier of 79 probe sets,

identify Group1 patients with a sensitivity of 0.98, a specificity of 1 and a misclassification rate of 0.0062.

The Group 1 GEP signature, found through shrinkage T statistics, included among up-regulated probe sets genes known to be associated to tumor suppression (*PCDH17*, *CABLES1*) or to be anti-angiogenic factors such as *ANGPT2*, whereas genes known to be involved in cell migration (*NT5E*, *ECM1*) or to be pro-angiogenic factors such as *BMP2* and *CTGF*, were found among up-regulated probe sets. This signature suggests that bone marrow cells interactions with their microenvironment might contribute to the favorable outcome in Group 1.

Newly found in this study, Group 1 patients shared a unique microRNAs expression signature of which the top ranked up-regulated miRNAs (hsa-miR-125b, hsa-miR-125b-2*, hsa-miR-99a, hsa-miR-100, hsa-miR-125a-3p and has-miR-491-5p) are known to be involved in hematopoietic malignancies and other cancers.

MiR-125b, largely studied in hematopoietic malignancies, is reported to be highly expressed in pediatric AML particularly in patients with the translocation *PML-RARA*[46], is involved in the translocation t(2;11)(p21;q23) in AML and myelodysplasia[47] and in the translocation t(11;14)(q24;q32) in B-ALL[48].

In vitro experiments demonstrated that miR-125b may block the process of differentiation of primary human CD34+ cells and leukemic cell lines (HL60 and NB4) [49]. *In vivo* experiments on mice, over-expressed miR-125b promoted malignant transformation and suggested a role of miR-125b in the differentiation process of lymphoid and myeloid lineages[50].

Regarding cellular regulations, it has been reported that miR-125b negatively regulates p53 and many proteins in the p53 pathway resulting in inhibition of apoptosis[51, 52], furthermore, a recent study showed that miR-125b overexpression confers p53 independent survival advantage *in vitro* in *TEL-AML1* ALL cell lines[53].

How overexpression of miR-125b , with pro-tumorigenic and pro-proliferative behaviour well describe in leukemia, could be associated to favourable outcome found in this study needs to be addressed in future studies.

Below we discus some recent finding involving deregulated miRNAs in solid cancer even if a recent work suggested that solid cancers and leukemia may have different oncogenic miRNAs networks[54].

Mir-125b and miR-100 are reported to be down-regulated in oral scuamous cell carcinoma tumors and to decrease cells proliferation in vitro[55]. MiR-99a and miR-100 are reported to be

potential tumor suppressor genes in prostate cancer[56]. Furthermore, genomic region that hosting these miRNAs (21q21.1 and 11q) are frequently involved in copy number losses and proposed as tumor suppressor locus in various solid cancer type[57, 58, 59].

Two separate studies, on adrenocortical tumors cell lines and clear cell ovarian cancer, showed that miR-100 and miR-99a overexpression inhibits mTOR signalling at mRNA and protein level, increasing the sensitivity of this cells to mTOR inhibitor RAD001 (everolimus)[60, 61].

The hypothesis of mTOR pathway inhibition in Group 1 patients is supported by down-regulation of three genes (*BMP2*, *CTGF*) reported to be related by mTOR signalling: inhibition of mTOR pathway are reported to impair the *BMP2* expression in osteoblasts[62]; *CTGF* is regulated by the mTOR pathway and expression decreases after treatment with mTOR inhibitors in fibroblast[63].

We reported that two miRNAs minor strand forms (also called "passenger") had deregulated expression and, in the case of miR-125b-2*, the distinct levels of the mature miR-125b-2* were also confirmed by qRT-PCR, among the two patients groups. Only few information is available on minor strand forms but is has been suggested that miRNAs guide strand and passenger strand expressions regulation might be independent[64].

Integration of miRNAs and mRNAs data failed to identify considerable correlations. Only one miRNA–mRNA target association were identified and the miRNA/mRNA involved were only slightly deregulated. This finding could suggest that deregulated miRNAs act mainly at a post-trascriptional level, without impairing transcript amount.

Notably, Group 1 patients shared also an unique signature of snoRNAs. The most of differentially expressed probe sets are up-regulated in Group 1, and between those with greater fold change there are many snoRNAs belonging to HBII-85 cluster in the 15q11.2 region. Although studies are just emerging, involvement of snoRNAs in cancer has been observed[65, 66, 67] and loss of snoRNA clusters HBII-85 and HBII-52 in 15q11.2 region is reported to results in Prader–Willi syndrome (PWS) [68], suggesting that deregulation of small RNA loci plays an important role in human disease.

Focusing on miR-125b-2 cluster (miR-125b-2, miR-99a and miR-let-7c) hosted in *C21ORF34* on chromosome 21, we directed our attention to chromosome 21 genes expression and found three genes up-regulated in Group 1 patients compared to Group 2 (*C21ORF34*, *PRSS7* and *C21ORF91*) that map close to each other into a 2.5 Mb region in 21q21.1. This

finding suggested a specific involvement of the 21q21.1 locus in the pathobiology of Group 1 leukemia.

Genome-wide DNA copy number analysis failed to detect aberrations in the 21q21.1 region in Group 1 and this region are not reported to be deleted in general[69]. Two out of eight patients had *CDKN2A/B* deletions and other two had *IKZF1* deletions, recently associated with a very poor outcome in B-cell-progenitor ALL[70]. In view of the overall good outcome of Group 1 the latter observation suggests that in the context of the Group 1 gene expression signature the deleterious effect of the *IKZF1* mutation may be counter balanced.

ERG intragenic deletions, already described as exclusive of this group[71], were identified in 4/8 Group 1 patients. Group 1 patients carrying *ERG* intragenic deletion had a specific signature respect the remaining Group 1 patients and seem to reveal a more immature phenotype of leukemic blasts.

Taken together these finding extend the insight in molecular features of Group 1 patients calling for further research on role of deregulated genes and small ncRNAs, and add magnitude to the contribution of chromosome 21 in leukemia pathobiology.

TABLES

FEATURES	N	%
All patients	145	100
Age		
1-5 Yrs	66	45.5
6-9 Yrs	42	29
≥10 Yrs	37	25.5
Gender		
Male	78	53.8
Female	67	46.2
WBC (x1000/µL)		
<20	89	74.2
20-100	16	13.3
≥100	15	12.5
Unknown	25	-
MRD D15		
HR	25	26.6
MR	42	44.7
SR	27	28.7
Unknown	51	-
MRD D78		
HR	17	12.5
MR	83	61
SR	36	26.5
Unknown	9	-
Prednison Response		
Good	108	82.4
Poor	23	17.6
Unknown	14	-
Final Protocol Strata		
HR	34	23.8
MR	75	52.4
SR	34	23.8
Unknown	2	-
Relapse events	31	21.4

Table 1A. Clinical data of 145 B-others patients in the study cohort.

FEATURES	GROUP 1		GROUP 2	
	N	%	N	%
All patients	36	100	109	100
Age				
1-5 Yrs	12	33.3	54	49.5
6-9 Yrs	15	41.7	27	24.8
≥10 Yrs	9	25	28	25.7
Gender				
Male	13	36.1	65	59.6
Female	23	63.9	44	40.4
WBC (x1000/µL)				
<20	24	88.9	65	69.9
20-100	2	7.4	14	15.05
≥100	1	3.7	14	15.05
Unknown	9	-	16	-
MRD D15				
HR	2	8.7	23	32.4
MR	13	56.5	29	40.8
SR	8	34.8	19	26.8
Unknown	13	-	38	-
MRD D78				
HR	5	14.3	12	11.9
MR	25	71.4	58	57.4
SR	5	14.3	31	30.7
Unknown	1	-	8	-
Prednison Response				
Good	27	81.8	81	82.7
Poor	6	18.2	17	17.3
Unknown	3	-	11	-
Final Protocol Strata				
HR	8	22.2	26	24.3
MR	23	63.9	52	48.6
SR	5	13.9	29	27.1
Unknown	0	-	2	-
Relapse	4	11.1	27	24.8

Table 1B. Clinical data of Group 1 and Group 2 patients.

Probeset ID	Gene Symbol		Probeset ID	Gene Symbol
227289_at	PCDH17		201200_at	CREG1
205656_at	PCDH17		224374_s_at	EMILIN2
228863_at	PCDH17		1569362_at	ALCAM
1554343_a_at	STAP1		239826_at	---
209602_s_at	GATA3		219789_at	NPR3
240758_at	---		226043_at	GPSM1
219489_s_at	NXN		38918_at	SOX13
203329_at	PTPRM		211148_s_at	ANGPT2
206756_at	CHST7		227358_at	ZBTB46
1552398_a_at	CLEC12A		1563113_at	UBR4
202409_at	IGF2 /// INS-IGF2		205423_at	AP1B1
239956_at	---		213272_s_at	TMEM159
225613_at	MAST4		236501_at	SALL4
235146_at	---		244509_at	GPR155
202747_s_at	ITM2A		226499_at	NRARP
233038_at	---		1553863_at	C10orf64
209603_at	GATA3		205054_at	NEB
233225_at	---		203448_s_at	TERF1
234196_at	---		229677_at	SLC39A3
206067_s_at	WT1		200753_x_at	SFRS2
215146_s_at	TTC28		219790_s_at	NPR3
225611_at	MAST4		241545_x_at	---
235968_at	CENTG2			
235343_at	---		236760_at	AMMECR1
209959_at	NR4A3		214716_at	BMP2K
232227_at	---		203110_at	PTK2B
231600_at	CLEC12B		201034_at	ADD3
213058_at	TTC28		201753_s_at	ADD3
207978_s_at	NR4A3		201752_s_at	ADD3
205572_at	ANGPT2		37170_at	BMP2K
225369_at	ESAM		201029_s_at	CD99
242664_at	---		221558_s_at	LEF1
229091_s_at	CCNJ		230069_at	SFXN1
223627_at	MEX3B		228555_at	CAMK2D
229661_at	SALL4		220952_s_at	PLEKHA5
219470_x_at	CCNJ		203066_at	GALNAC4S-6ST
1569401_at	CLEC12A		209781_s_at	KHDRBS3
227329_at	ZBTB46		227486_at	NT5E
219227_at	CCNJL		205289_at	BMP2
231357_at	CLEC12B		205290_s_at	BMP2

Table 2. All 79 probe sets in the classifier for Group 1 patients (misclassification rate 0.0062, sensitivity 0.9764286, specificity 1). Group 1 up-regulated probe sets in white cells, Group 1 down-regulated probe sets in grey cells.

Probeset ID	Locus	PvalueBH	Mean Group 1	Mean Group 2	Fold Change
hsa-miR-125b-2-star_st	21q21.1	0,002	5,126	0,943	5,437137879
hsa-miR-99a_st	21q21.1	0,002	8,819	2,156	4,091059018
hsa-miR-100_st	11q24.1	0,017	4,440	1,387	3,201542161
hsa-miR-125b_st	21q21.1; 11q24.1	0,015	9,926	3,845	2,581418873
hsa-miR-125a-3p_st	19q13.41	0,016	3,506	1,500	2,338077591
hsa-miR-491-5p_st	9p21.3	0,031	3,589	1,592	2,254831889
hsa-miR-126_st	9q34.3	0,002	10,025	6,981	1,436036194
hsa-miR-1275_st	6p21.31	0,035	7,449	6,046	1,232153294
hsa-miR-181d_st	19p13.13	0,022	7,056	5,975	1,180886865
hsa-let-7c_st	21q21.1	0,029	11,243	9,729	1,155527216
hsa-miR-181a_st	1q32.1; 9q33.3	0,035	12,940	12,454	1,039073506
hsa-miR-92a_st	13q31.3; Xq26.2	0,015	12,466	12,063	1,033381857
hsa-miR-24_st	9q22.33; 19p13.13	0,035	9,442	10,234	0,922549142
hsa-miR-15b_st	3q25.33	0,002	8,183	9,006	0,908572506
hsa-miR-155_st	21q21.3	0,003	8,695	10,239	0,849219966
hsa-miR-27a_st	19p13.13	0,035	5,789	7,061	0,819789015
hsa-miR-625_st	14q23.3	0,022	5,550	7,016	0,790995143
hsa-miR-132-star_st	17p13.3	0,035	3,720	5,710	0,651573502

Table 3A. MicroRNAs probe sets differentially expressed between Group 1 and Group 2 patients (18 probe sets, p-value BH<0.05). In white cells are highlighted probe sets whit fold change >1.5 and <1/1.5 (7 probe sets).

Probeset ID	Locus	Gene name	snoRNA type	Pvalue BH	Mean Group 1	Mean Group 2	Fold Change
HBII-85-16_x_st	15q11.2	SNORD116-16	C/D box	0,002	1,343	0,422	3,184689209
HBII-85-22_x_st	15q11.2	SNORD116-22	C/D box	0,004	1,667	0,603	2,762666922
HBII-85-17_x_st	15q11.2	SNORD116-17	C/D box	0,002	1,883	0,753	2,499964778
HBII-85-11_st	15q11.2	SNORD116-11	C/D box	0,004	3,253	1,353	2,404718552
HBII-85-15_x_st	15q11.2	SNORD116-15	C/D box	0,003	1,995	0,837	2,383175437
HBII-85-24_x_st	15q11.2	SNORD116-24	C/D box	0,010	1,700	0,786	2,164138837
HBII-85-14_x_st	15q11.2	SNORD116-14	C/D box	0,004	1,498	0,709	2,11231288
HBII-85-23_x_st	15q11.2	SNORD116-23	C/D box	0,003	4,491	2,243	2,002199542
HBII-438A_s_st	15q11.2	SNORD109A	C/D box	0,004	3,887	1,961	1,981941655
HBII-85-21_x_st	15q11.2	SNORD116-21	C/D box	0,012	1,418	0,717	1,977340673
HBII-13_st	15q11.2	SNORD64	C/D box	0,016	2,955	1,626	1,817114415
HBII-13_x_st	15q11.2	SNORD64	C/D box	0,013	2,673	1,500	1,781780056
ENSG00000212326_st	2p32.2	ENSG00000212326		0,004	1,557	0,875	1,778873714
U47_st	1q25.1	SNORD47	C/D box	0,030	1,224	0,707	1,732639669
mgU6-77_st	17p13.1	SNOR10	C/D box	0,004	1,552	0,909	1,707226125
HBII-296B_st	17p13.3	SNORD91B	C/D box	0,017	1,083	0,640	1,69255785
ACA17_st	9q34.3	SNORA17	H/ACA box	0,003	2,762	1,699	1,625896033
ENSG00000201199_s_st	11q21	SCARNA9		0,035	1,386	0,869	1,595609039
U53_st	2p23.2	SNORA53	C/D box	0,002	4,007	2,573	1,557373998
U35B_st	19q13.33	SNORD35B	C/D box	0,002	2,382	1,556	1,530867912

Table 3B. Small Nucleolar RNAs probe sets differentially expressed between Group 1 and Group 2 patients (20 probe sets whit fold change >1.5, p-value BH<0.05).

Probeset ID	Gene Symbol	Locus	PvalueBH	Mean Group 2	Mean Group 1	Fold change
207638_at	PRSS7	21q21.1	0,00784	3,210666667	7,0561875	2,197732818
239999_at	C21orf34	21q21.1	0,00336	3,08116	5,8805875	1,908562846
1561777_at			0,00336	3,164166667	5,1216875	1,618652884
217269_s_at	PRSS7	21q21.1	0,027328	5,67718	7,7922	1,372547638
211018_at	LSS	21q22.3	0,01176	4,766053333	6,4779125	1,35917751
202245_at	LSS	21q22.3	0,0096	7,528153333	9,72	1,291153297
230631_s_at	LOC100288432	21q22.3	0,01269333	4,508073333	5,772075	1,280386226
220941_s_at	C21orf91	21q21.1	0,00448	6,514433333	8,2515625	1,266658522
1559901_s_at	C21orf34	21q21.1	0,00504	3,176813333	3,6618625	1,152684189
205663_at	PCBP3	21q22.3	0,03957333	5,33284	5,89105	1,104674057
237853_x_at	KRTAP10-12	21q22.3	0,03957333	5,224766667	5,65915	1,083139279
1555265_at	ABCC13	21q11.2	0,02429538	3,379313333	3,2197875	0,952793418
206409_at	TIAM1	21q22.11	0,020832	5,141586667	4,506825	0,876543622
244822_at	GART	21q22.11	0,027328	6,3064	5,216	0,827096283
243091_at			0,02429538	6,811766667	5,5249625	0,81109098
229437_at	MIR155HG	21q21.3	0,02138182	8,777066667	6,6219375	0,754459064
213135_at	TIAM1	21q22.11	0,00784	7,660826667	5,66915	0,740018049
207267_s_at	DSCR6	21q22.13	0,03957333	4,95624	3,4152625	0,689083358

Table 4. All 18 chromosome 21 probe sets differentially expressed among Group 1 and Group 2 in 23 miRNAs profiled patients. (18 probe sets, p-value BH<0.05). In white cells are are highlighted probe sets whit fold change >1.5 (3 probe sets).

FIGURES

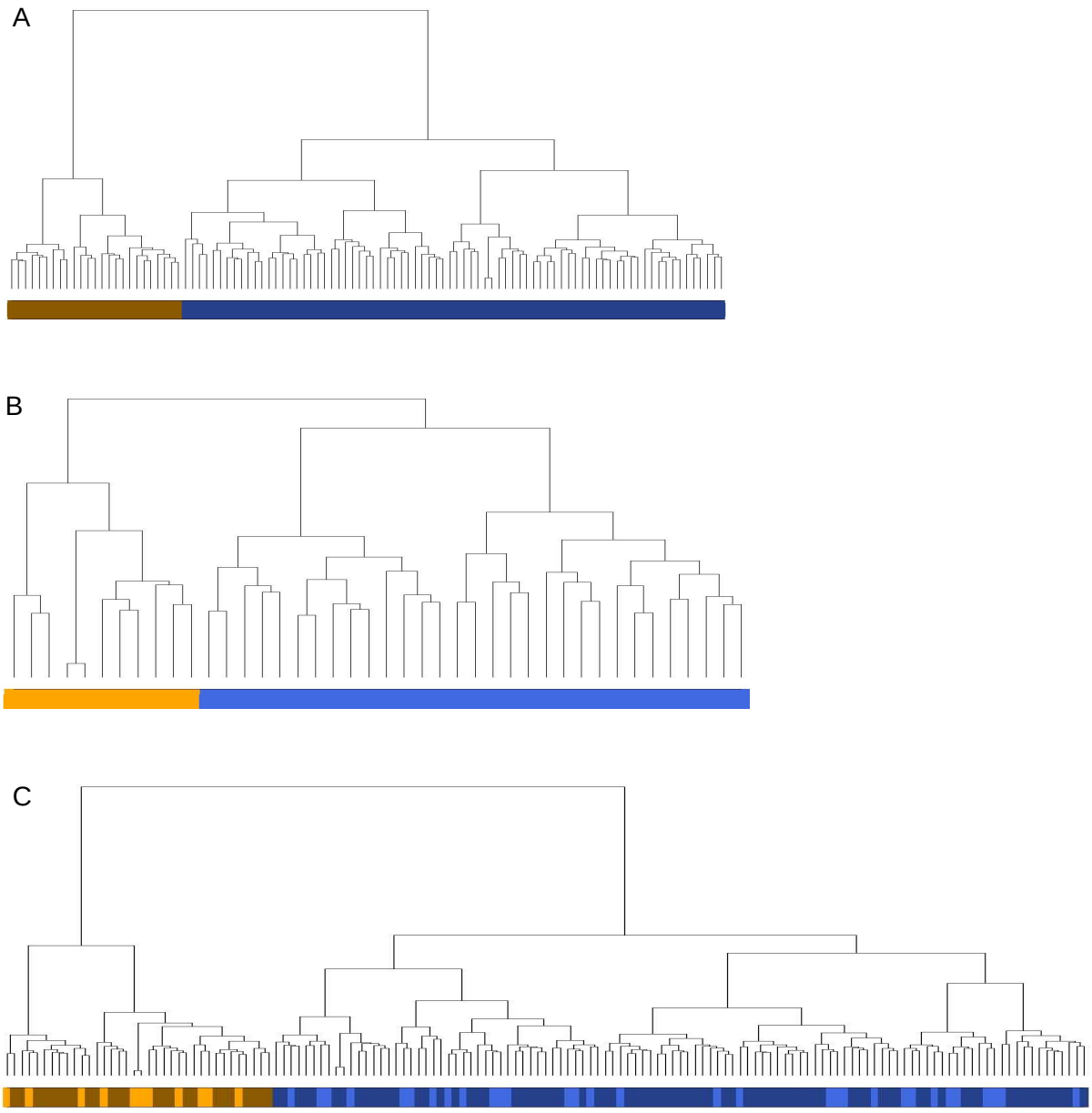
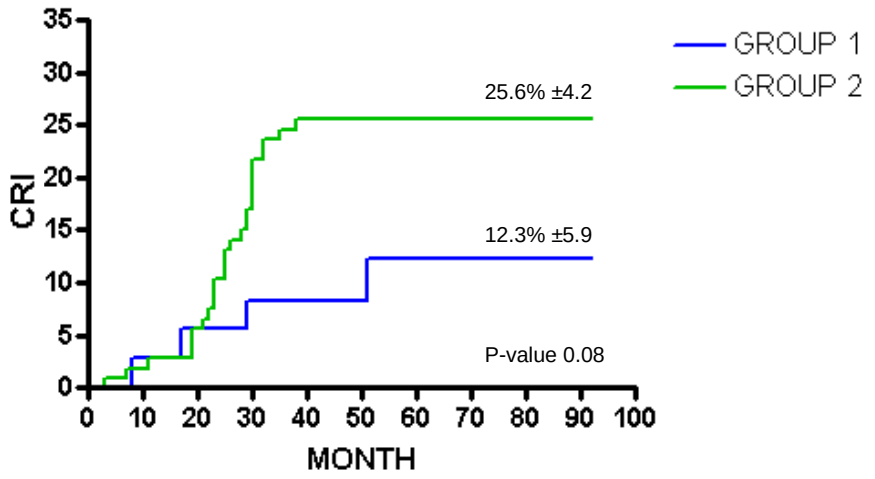


Figure 1. Unsupervised hierarchical cluster of B-others patients on 5463 most variable probe sets. First cohort (103 patients) Group 1 patients (brown) cluster separately from the rest of the cohort, Group 2 (blue) (A). Second cohort (42 patients) Group 1' patients (orange) cluster separately from the Group 2' patients (light blue) (B). In the merge of the two cohorts (145 patients) 36 patients belonging to Group 1 and 1' grouped toghether (Group 1'') separately from the major group (Group 2'').

A



B

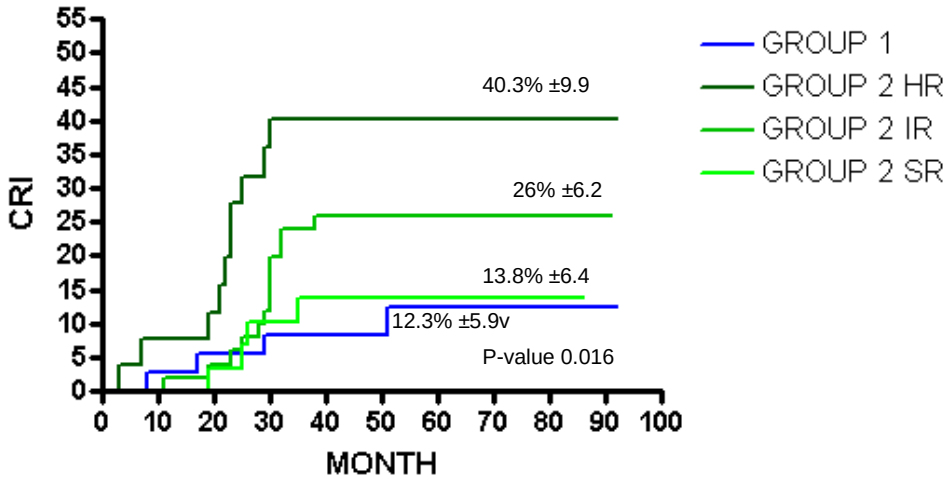


Figure 2. Cumulative relapse incidence in B-others cohort. All 145 patients grouped by Group 1 (4/36 relapsed) and Group 2 (27/109 relapsed) (A). Group 1 patients compared to the three risk strata of Group 2 patients - high risk (HR), intermediate risk (IR) or standard risk (SR). Two patients missing risk classification were excluded.(B).

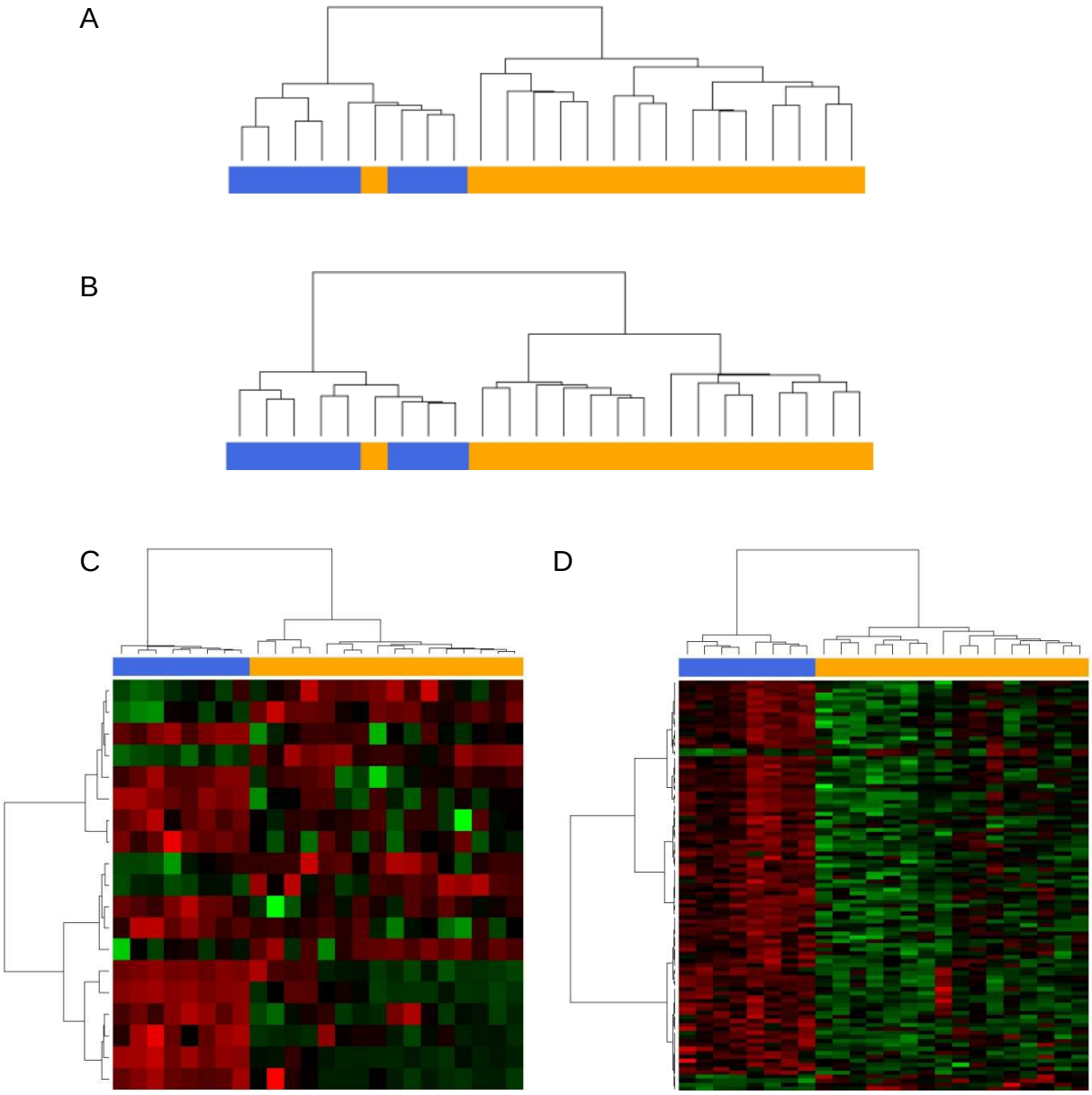


Figure 3A-B. Dendrograms representing unsupervised hierarchical clustering of 24 patients on 847 human miRNAs probe sets (A) and 922 non coding RNAs probe sets (B). Of note: outlayer of Group 2 in cluster A is not the same sample as outlayer of Group 2 in cluster B. Blue: 8 Group 1 patients. Yellow :16 Group 2 patients.

Figure 3C-D. Heatmaps of 24 patients clustered on 18 differentially expressed miRNAs probe sets (C) and on 103 differentially expressed non coding RNAs probe sets (D). Blue: 8 Group 1 patients. Yellow :16 Group 2 patients.

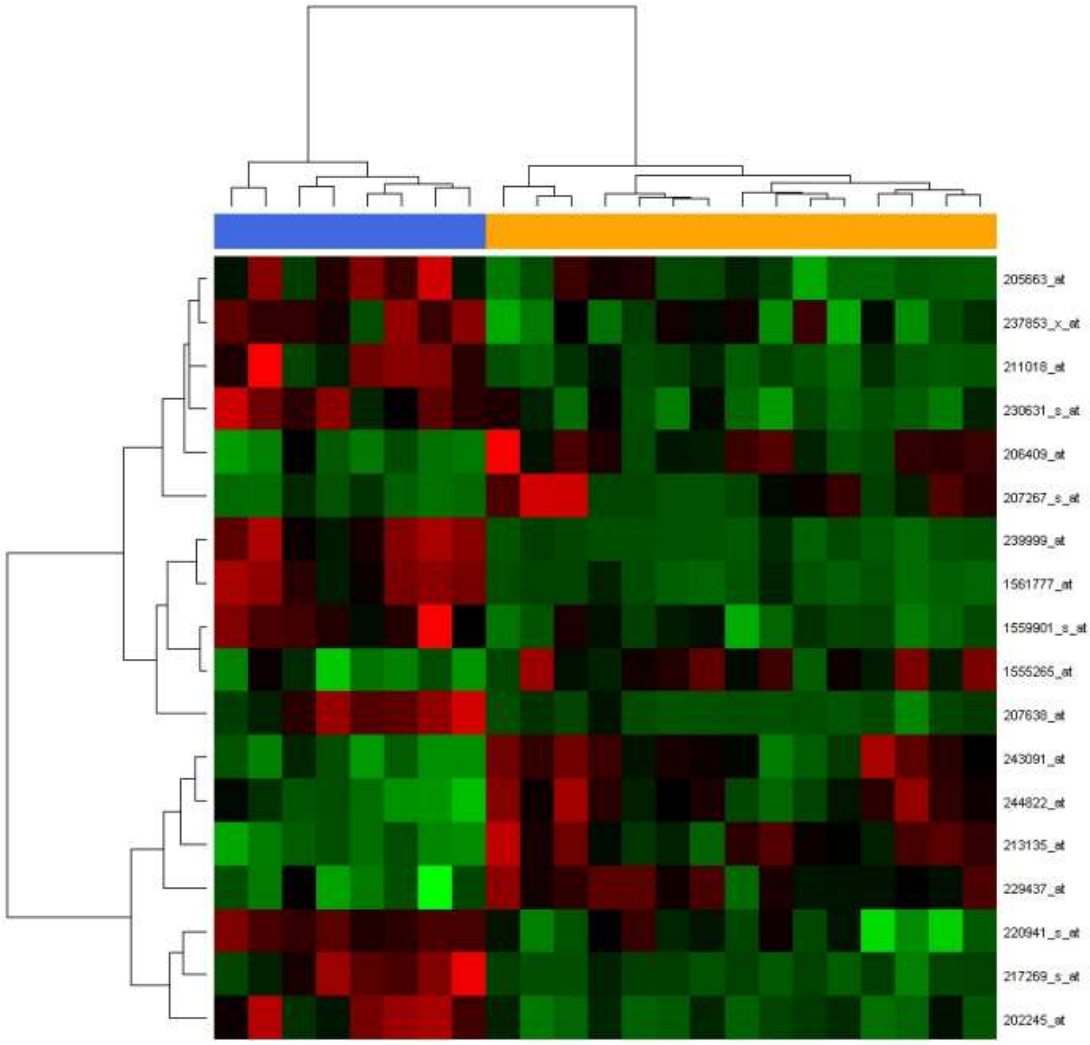


Figure 4. Heatmap of 23 patients clustered on 18 differentially expressed chromosome 21 probe sets. Blue: 8 Group 1 patients. Yellow: 15 Group 2 patients.

SUPPLEMENTARY TABLES

Probeset ID	Gene Symbol	Locus	Mean Group 1	Mean Group 2	Fold change
228863_at	PCDH17	13q21.1	10,19965951	3,034441389	3,361297256
227289_at	PCDH17	13q21.1	9,989171147	3,093574606	3,229006059
231600_at	CLEC12B	12p13.2	6,670410917	3,060136294	2,179775761
205656_at	PCDH17	13q21.1	10,07968937	4,778343703	2,109452564
233038_at	NA	---	7,264057469	3,535118746	2,054827006
1556592_at	NA	---	6,768821494	3,33735406	2,028199997
242664_at	NA	---	6,148296379	3,036717919	2,024651793
1555579_s_at	PTPRM	18p11.2	10,09590578	5,058236456	1,995933933
240758_at	NA	---	8,091554566	4,189554188	1,931364103
230537_at	NA	---	8,211870135	4,28168936	1,91790423
235146_at	TMCC3	---	8,154497698	4,253015635	1,917344867
215146_s_at	TTC28	22q12.1	8,412849067	4,401768495	1,911242964
202409_at	NA	11p15.5	7,723317693	4,05166622	1,906207785
233225_at	NA	---	7,526831417	3,972423707	1,894770541
239956_at	NA	---	8,585001222	4,576454203	1,875906726
204066_s_at	AGAP1	2p24.3-p24.1	10,22457201	5,486468151	1,863598171
234196_at	NA	---	8,40479604	4,613411514	1,821817979
203329_at	PTPRM	18p11.2	8,722099209	4,980196246	1,751356529
213058_at	TTC28	22q12.1	8,68828589	4,995905055	1,739081467
231357_at	CLEC12B	12p13.2	5,666240203	3,278354916	1,728379125
205572_at	ANGPT2	8p23.1	7,304180567	4,241827072	1,721942088
1552398_a_at	NA	12p13.2	9,24158362	5,397668042	1,712143753
237261_at	ANGPT2	---	5,939539081	3,489584358	1,702076371
228698_at	SOX7	---	9,918084242	5,835775813	1,699531401
243478_at	NA	---	5,550755796	3,2853678	1,689538625
1569401_at	CLEC12A	12p13.2	5,845850418	3,467963338	1,685672496
209602_s_at	GATA3	10p15	9,256065613	5,507246145	1,680706722
235968_at	AGAP1	2p24.3-p24.1	6,940851333	4,198187024	1,653297315
206067_s_at	WT1	11p13	9,020141544	5,473515719	1,647961202
1556593_s_at	NA	---	7,74699879	4,835014862	1,602269902
206756_at	CHST7	Xp11.23	9,180497493	5,774717867	1,589774203
236034_at	ANGPT2	---	6,156016379	3,903360212	1,577106914
1554343_a_at	STAP1	4q13.2	11,19465613	7,225371272	1,549353757
1558621_at	CABLES1	18q11.2	6,653683395	4,300257637	1,547275526
209603_at	GATA3	10p15	8,074768681	5,237582794	1,541697573
225613_at	MAST4	5q12.3	8,752815864	5,67821749	1,541472457
226489_at	TMCC3	12q22	8,034650873	5,241863208	1,532785301
215177_s_at	ITGA6	2q31.1	10,67878494	7,114637029	1,500959908
220389_at	CCDC81	11q14.2	4,101945612	6,740111928	0,608587165
224994_at	CAMK2D	4q26	4,311148722	7,148539093	0,603081086
225688_s_at	PHLDB2	3q13.2	4,379150809	7,435351177	0,588963548
209101_at	CTGF	6q23.1	5,68791341	9,797561935	0,580543756
213385_at	CHN2	7p15.3	3,955488668	6,829099899	0,579210837
231042_s_at	NA	4q26	2,97550664	5,142268304	0,578636987
227486_at	NT5E	6q14-q21	4,910139039	8,601811826	0,57082614
202242_at	TSPAN7	Xp11.4	4,034017412	7,266420323	0,555158831
205290_s_at	BMP2	20p12	5,458800893	10,19819154	0,535271462
205289_at	BMP2	20p12	4,867660875	9,3384622	0,521248656
203939_at	NT5E	6q14-q21	4,488371638	8,727601751	0,514273195

ST 1. Forty-nine most deregulated probe sets in Group 1 (lfdr <0,05; fold change >1.5 and <1/1.5). White cells are up-regulated probe sets in Group 1 and grey cells down-regulated probe sets in Group 1.

Patients code	Copy Number State			*ERG* intragenic deletion
	IKZF1	***CDKN2A/CDKN2B***	***C21ORF34***	
Pts.1	2	2	2	Yes
Pts.2	1	2	2	Yes
Pts.3	2	1	2	No
Pts.4	2	0	2	No
Pts.5	1	2	2	No
Pts.6	2	2	2	Yes
Pts.7	2	2	2	Yes
Pts.8	2	2	2	No

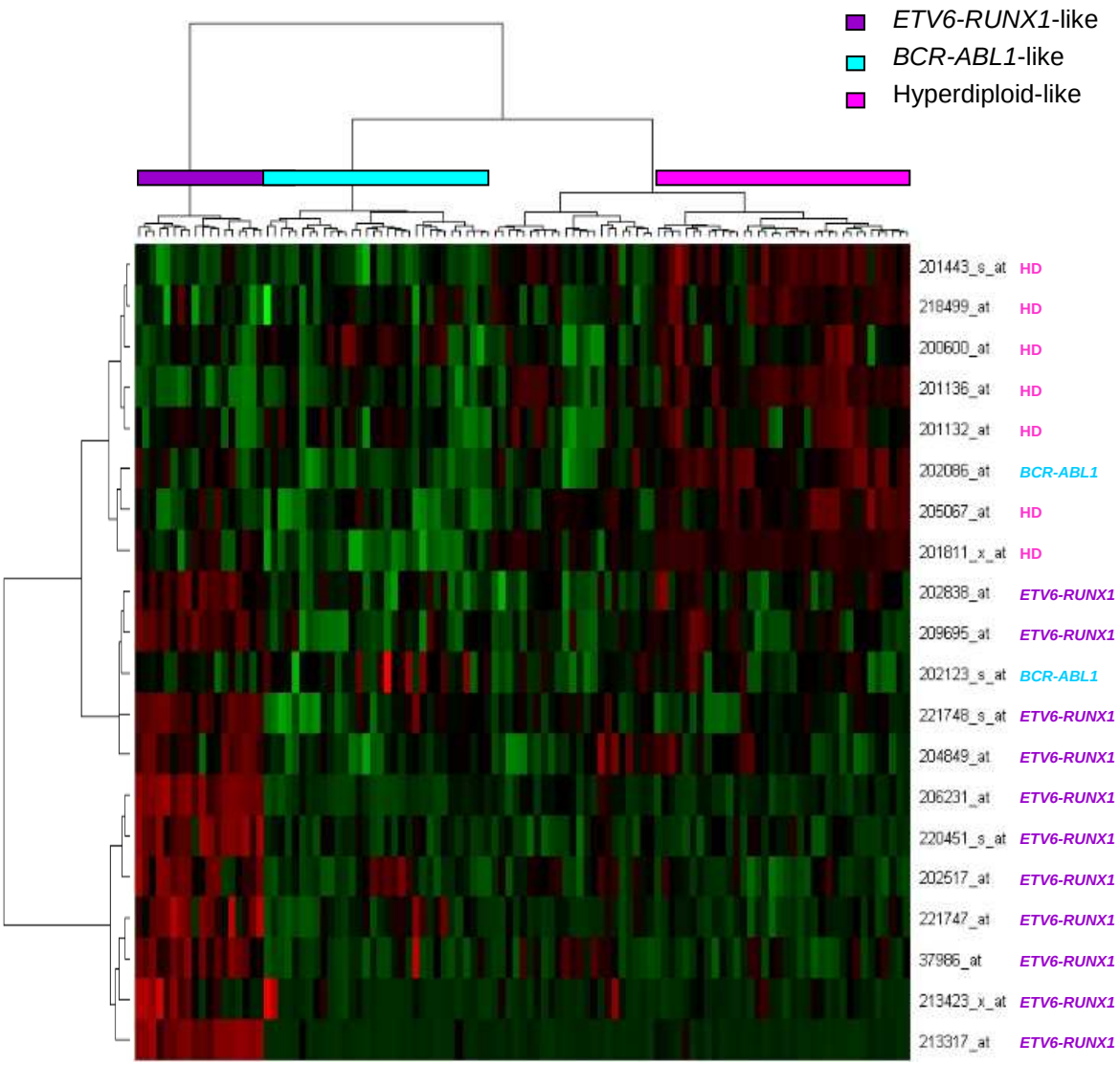
ST 2. Copy Number Variations data of 8 Group 1 patients. CNS 2=wild type, CNS 1=heterozygous loss, CNS 0=homozygous loss.

Probeset ID	Gene Symbol	Locus	Mean *ERG* deleted	Mean Non *ERG* deleted	Fold change
229459_at	FAM19A5	chr22q13.32	79,107	33,9795	2,328080166
229057_at	SCN2A	chr2q23-q24	85,02725	38,5485	2,205721364
237094_at	FAM19A5	chr22q13.32	79,4605	39,71275	2,000881329
224520_s_at	BEST3	chr12q14.2-q15	75,4525	37,8105	1,995543566
228740_at	---	---	65,18625	33,18675	1,96422518
203440_at	CDH2	chr18q11.2	79,551	41,0885	1,936089173
213994_s_at	SPON1	chr11p15.2	68,4875	37,056	1,848216213
205347_s_at	TMSL8	chrXq21.33-q22.3	87,85875	48,07775	1,827430568
213993_at	SPON1	chr11p15.2	58,889	33,35925	1,765297481
229233_at	NRG3	chr10q22-q23	80,4225	46,03125	1,74712831
219855_at	NUDT11	chrXp11.22	91,1535	52,29175	1,743171724
229085_at	LRRC3B	chr3p24	85,435	49,964	1,70993115
228108_at	---	---	67,98125	40,95475	1,659911244
216623_x_at	TOX3	chr16q12.1	100,9225	60,98225	1,654948776
214774_x_at	TOX3	chr16q12.1	102,70725	62,46975	1,64411175
226974_at	NEDD4L	---	66,1775	40,31375	1,641561502
230928_at	---	---	94,214	57,7995	1,6300141
1558388_a_at	---	---	79,82625	49,234	1,621364301
215108_x_at	TOX3	chr16q12.1	100,058	61,84175	1,617968444
203131_at	PDGFRA	chr4q11-q13	113,469	70,78175	1,603082716
228776_at	GJC1	chr17q21.31	74,793	48,283	1,549054533
225627_s_at	CACHD1	chr1p31.3	76,09725	49,405	1,540274264
229655_at	FAM19A5	chr22q13.32	89,3415	58,10525	1,537580511
217901_at	DSG2	chr18q12.1	63,9495	41,8515	1,528009749
205114_s_at	CCL3	chr17q11-q21	63,45525	95,13775	0,666982875
202912_at	ADM	chr11p15.4	65,16625	98,21	0,663539864

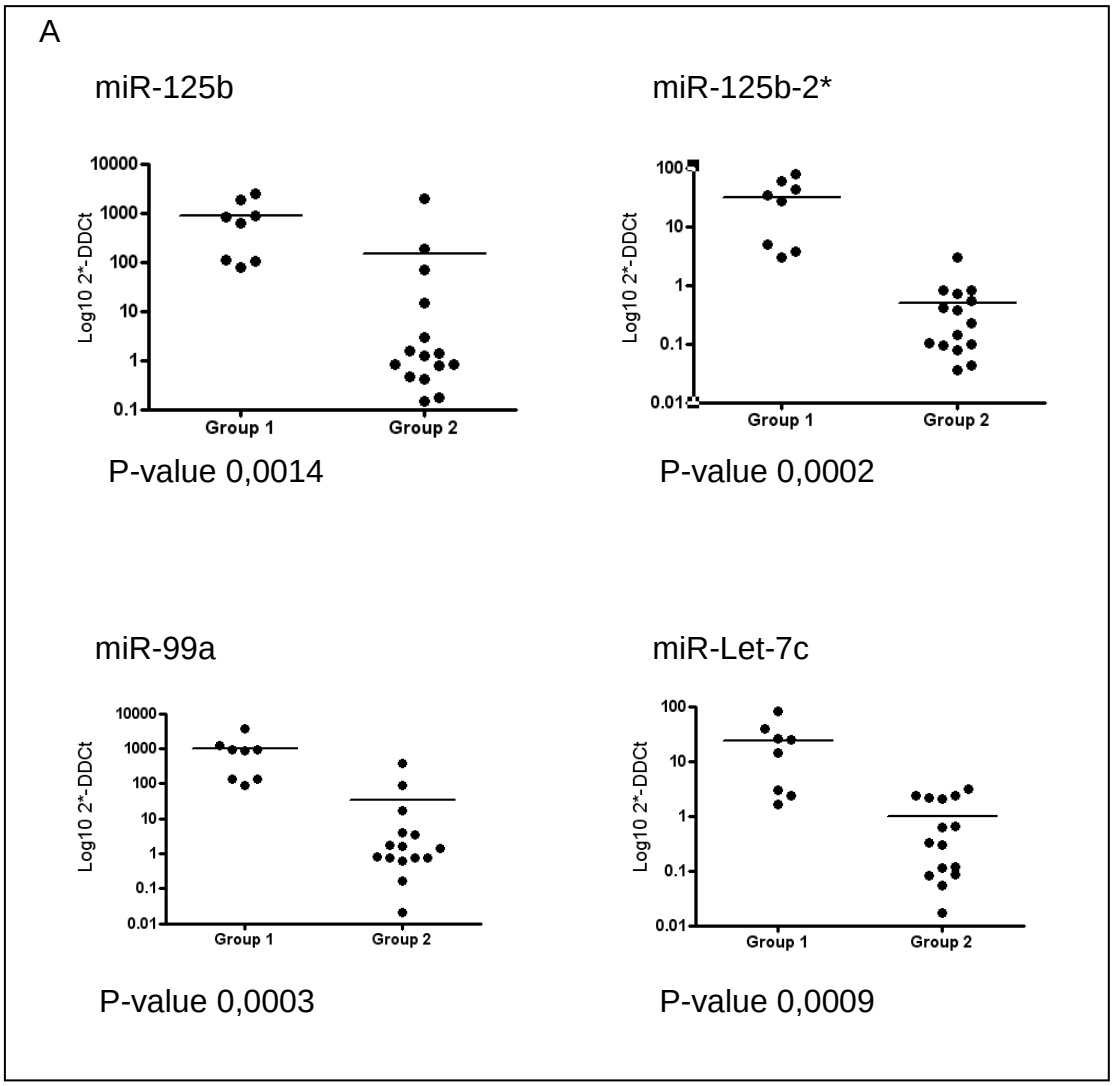
205767_at	EREG	chr4q13.3	38,86525	58,599	0,66324084
206115_at	EGR3	chr8p23-p21	55,9575	84,72375	0,660470057
201888_s_at	IL13RA1	chrXq24	45,703	69,28475	0,659640108
215087_at	C15orf39	chr15q24.2	47,92575	72,923	0,657210345
212769_at	TLE3	chr15q22	46,26025	70,6465	0,654813048
201893_x_at	DCN	chr12q21.33	52,3525	80,35575	0,651509071
1557905_s_at	CD44	chr11p13	67,10925	103,27625	0,649803319
203949_at	MPO	chr17q23.1	63,95475	98,57825	0,648771407
210916_s_at	CD44	chr11p13	65,12725	100,51125	0,647959805
205997_at	ADAM28	chr8p21.2	36,87475	56,987	0,647073017
207341_at	PRTN3	chr19p13.3	46,0915	71,72025	0,642656711
230127_at	---	---	46,65	73,05975	0,638518473
242907_at	---	---	56,60125	89,0195	0,63582979
238865_at	LOC132430	chr4q28.3	55,56225	89,06925	0,623809564
226099_at	ELL2	chr5q15	56,43475	90,6215	0,622752327
1565868_at	CD44	chr11p13	50,028	81,09975	0,616869966
229221_at	CD44	chr11p13	53,9435	88,079	0,61244451
204490_s_at	CD44	chr11p13	60,5885	99,35025	0,609847484
211612_s_at	IL13RA1	chrXq24	40,751	67,1025	0,60729481
210895_s_at	CD86	chr3q21	41,2795	68,07825	0,606353718
209606_at	PSCDBP	chr2q11.2	68,3435	112,75225	0,60613868
209835_x_at	CD44	chr11p13	59,27525	98,43175	0,602196446
203948_s_at	MPO	chr17q23.1	51,631	85,8015	0,601749387
212014_x_at	CD44	chr11p13	59,26675	98,752	0,600157465
225612_s_at	B3GNT5	chr3q28	48,4095	81,0375	0,597371587
227697_at	SOCS3	chr17q25.3	49,26725	84,7175	0,581547496
214146_s_at	PPBP	chr4q12-q13	44,39175	84,23875	0,526975412
227140_at	---	---	37,98875	74,814	0,507775951
231911_at	KIAA1189	chr2q24.1	40,497	85,34	0,474537146
205899_at	CCNA1	chr13q12.3-q13	36,93525	89,704	0,411745853

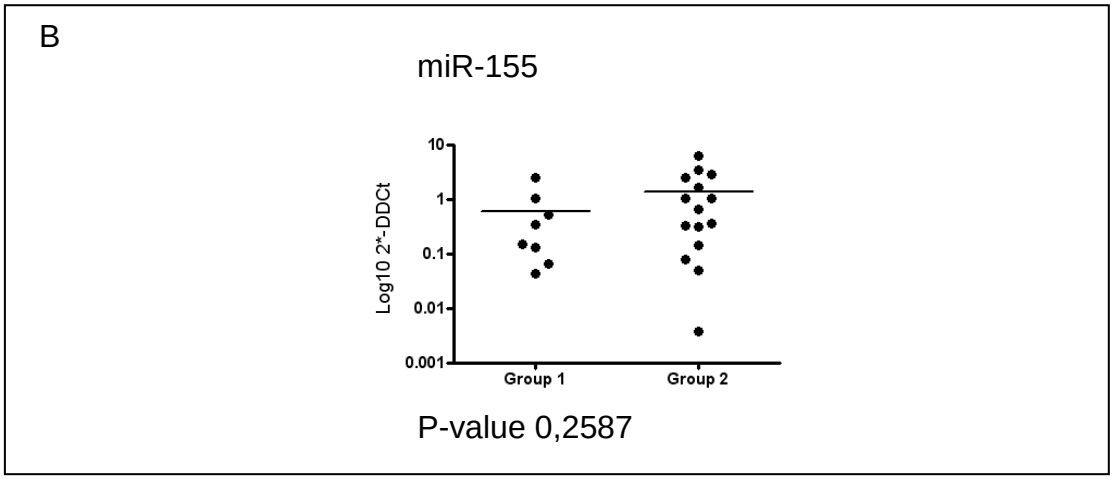
ST 3. Differentially expressed probe sets among Group 1 patients with and without *ERG* intragenic deletions (56 probe sets, lfdr<0.05). Twenty-four probe sets in white cells are up-regulated (fold change >1.5) and thirty-two probe sets in grey cells are down-regulated (fold change <1/1.5) in *ERG* deleted patients.

SUPPLEMENTARY FIGURES

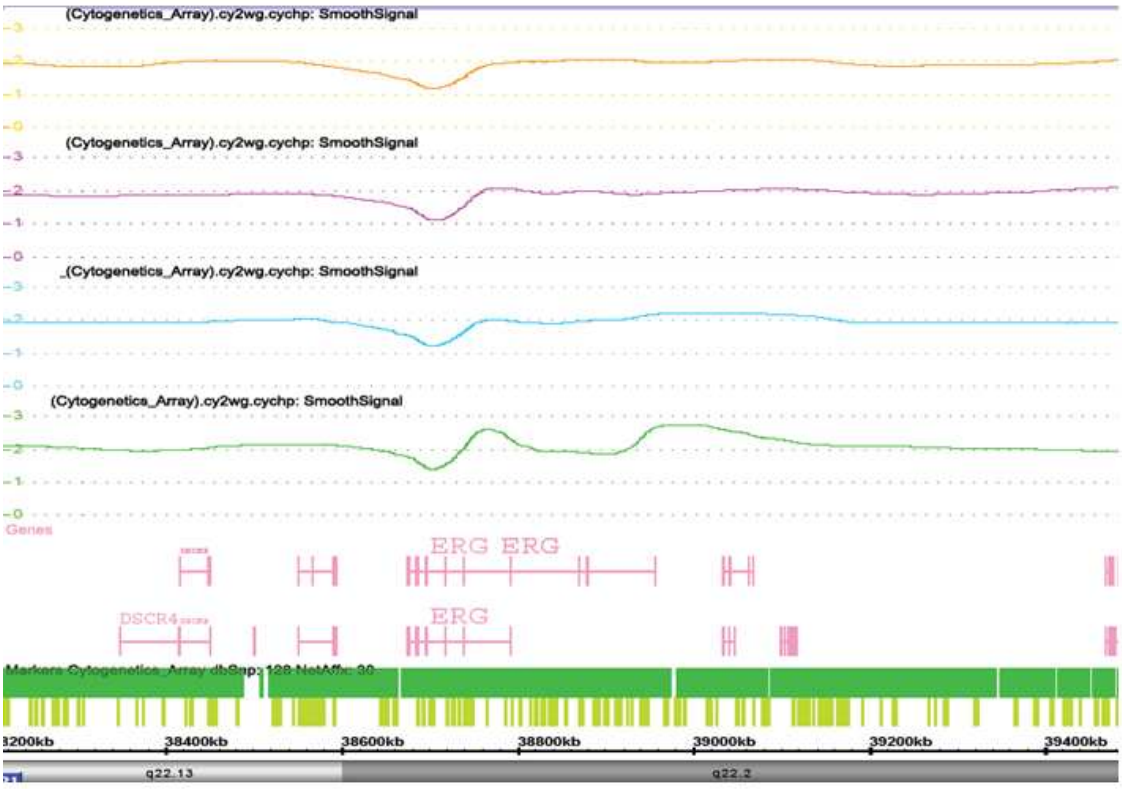


SF 1. Group 2 patients clustered on 20 probe sets from genetic defined ALL subgroups (Table S-IV, Den Boer et al., 2009). Pink label for Hyperdiploid-like samples, light blue label for *BCR-ABL1*-like samples, violet label for *ETV6-RUNX1*-like samples.





SF.2. Results from qRT-PCR analysis of expression of mature miRNAs. MiRNAs in miR-125b-2 cluster (miR-125b, miR-125b-2*, miR-99a and miR-let-7c) are over-expressed in Group 1 patients (A). MiR-155 is not differentially expressed among two groups (B).



SF 3. *Chromosome Analysis Suite* view. Intragenic *ERG* deletion in 21q22.3 region shared by 4 out of 8 investigated Group 1 patients. Deletions, always found in heterozygosis, extended for about 50 Kb, included exons 6-11 (transcript variant 2, NM_004449) of *ERG*. Sample reppresented by green line show two additional amplifications: one gain between exons 4-5 and one gain at 5' of *ERG* (including exon 1). Copy number state (CNS) of 21q22.13-q22.2 region are shown: CNS 2=wild type, CNS 1=heterozigous loss, CNS 0=homozigous loss, CNS >2= gain.

REFERENCES

[1] Pui CH, Relling MV, Downing JR. Acute lymphoblastic leukemia. *N Engl J Med.* 2004;350:1535–48.

[2] Yeoh, E.J., Ross, M.E., Shurtleff, S.A., Williams, W.K., Patel, D., Mahfouz, R., Behm, F.G., Raimondi, S.C., Relling, M.V., Patel, A., Cheng, C., Campana, D., Wilkins, D., Zhou, X., Li, J., Liu, H., Pui, C.H., Evans, W.E., Naeve, C., Wong, L. & Downing, J.R. (2002) Classification, subtype discovery, and prediction of outcome in paediatric acute lymphoblastic leukaemia by gene expression profiling. Cancer Cell, 1, 133–143.

[3] A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide classification study. Den Boer ML, van Slegtenhorst M, De Menezes RX, Cheok MH, Buijs-Gladdines JG, Peters ST, Van Zutven LJ, Beverloo HB, Van der Spek PJ, Escherich G, et al. Lancet Oncol. 2009 Feb; 10(2):125-34. Epub 2009 Jan 8.

[4] Blood. 2010 Dec 2;116(23):4874-84. Epub 2010 Aug 10. Identification of novel cluster groups in pediatric high-risk B-precursor acute lymphoblastic leukemia with gene expression profiling: correlation with genome-wide DNA copy number alterations, clinical characteristics, and outcome. Harvey RC, Mullighan CG, Wang X, Dobbin KK, Davidson GS, Bedrick EJ, Chen IM, Atlas SR, Kang H, Ar K, Wilson CS, Wharton W, Murphy M, Devidas M, Carroll AJ, Borowitz MJ, Bowman WP, Downing JR, Relling M, Yang J, Bhojwani D, Carroll WL, Camitta B, Reaman GH, Smith M, Hunger SP, Willman CL.

[5] Genes Chromosomes Cancer. 2009 Jan;48(1):22-38. Integration of genomic and gene expression data of childhood ALL without known aberrations identifies subgroups with specific genetic hallmarks. Bungaro S, Dell'Orto MC, Zangrando A, Basso D, Gorletta T, Lo Nigro L, Leszl A, Young BD, Basso G, Bicciato S, Biondi A, te Kronnie G, Cazzaniga G.

[6] Risk of relapse of childhood acute lymphoblastic leukemia is predicted by flow cytometric measurement of residual disease on day 15 bone marrow. Basso G, Veltroni M, Valsecchi MG, Dworzak MN, Ratei R, Silvestri D, Benetello A, Buldini B, Maglia O, Masera G, Conter V, Arico M, Biondi A, Gaipa G. J Clin Oncol. 2009 Nov 1;27(31):5168-74. Epub 2009 Oct 5.

[7] Mullighan, C.G., Goorha, S., Radtke, I., Miller, C.B., Coustan-Smith, E., Dalton, J.D., Girtman, K., Mathew, S., Ma, J., Pounds, S.B., Su, X., Pui, C.H., Relling, M.V., Evans, W.E., Shurtleff, S.A. & Downing, J.R. (2007) Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature, 446, 758–764

[8] Kuiper, R.P., Schoenmakers, E.F., van Reijmersdal, S.V., Hehir-Kwa, J.Y., van Kessel, A.G., van Leeuwen, F.N. & Hoogerbrugge, P.M. (2007) High-resolution genomic profiling of childhood ALL reveals novel recurrent genetic lesions affecting pathways involved in lymphocyte differentiation and cell cycle progression. Leukemia, 21, 1258–1266.

[9] Mullighan, C.G., Collins-Underwood, J.R., Phillips, L.A., Loudin, M.G., Liu, W., Zhang, J., Ma, J., Coustan-Smith, E., Harvey, R.C.,Willman, C.L., Mikhail, F.M., Meyer, J., Carroll, A.J., Williams, R.T.,Cheng, J., Heerema, N.A., Basso, G., Pession, A., Pui, C.H., Raimondi,S.C., Hunger, S.P., Downing, J.R., Carroll, W.L. & Rabin,K.R. (2009) Rearrangement of CRLF2 in B-progenitor- and Downsyndrome-associated acute lymphoblastic leukemia. Nature Genetics,41, 1243–1246.

[10] Russell, L.J., De Castro, D.G., Griffiths, M., Telford, N., Bernard, O., Panzer-Grumayer, R., Heidenreich, O., Moorman, A.V. & Harrison, C.J. (2009) A novel translocation, t(14;19)(q32;p13), involving IGH@ and the cytokine receptor for erythropoietin. Leukemia, 23,614–617.

[11] Yoda, A., Yoda, Y., Chiaretti, S., Bar-Natan, M., Mani, K., Rodig, S.J., West, N., Xiao, Y., Brown, J.R., Mitsiades, C., Sattler, M., Kutok, J.L., DeAngelo, D.J., Wadleigh, M., Piciocchi, A., Dal Cin, P., Bradner, J.E., Griffin, J.D., Anderson, K.C., Stone, R.M., Ritz, J., Foa, R., Aster, J.C., Frank, D.A. & Weinstock, D.M. (2009) Functional screening identifies CRLF2 in precursor B-cell acute lymphoblastic leukemia. Proceedings of the National Academy of Sciences of the USA, 107, 252–257.

[12] Hertzberg, L., Vendramini, E., Ganmore, I., Cazzaniga, G., Schmitz, M., Chalker, J., Shiloh, R., Iacobucci, I., Shochat, C., Zeligson, S., Cario, G., Stanulla, M., Strehl, S., Russell, L.J., Harrison, C.J., Bornhauser, B., Yoda,

A., Rechavi, G., Bercovich, D., Borkhardt, A., Kempski, H., Kronnie, G.T., Bourquin, J.P., Domany, E. & Izraeli, S. (2010) Down syndrome acute lymphoblastic leukemia: a highly heterogeneous disease in which aberrant expression of CRLF2 is associated with mutated JAK2: a report from the iBFM Study Group. Blood, 115, 1006–1017.

[13] Calin GA, Sevignani C, Dumitru CD, et al. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci U S A. 2004;101(9):2999-3004.

[14] Genome-wide identification of human microRNAs located in leukemia-associated genomic alterations. Starczynowski DT, Morin R, McPherson A, Lam J, Chari R, Wegrzyn J, Kuchenbauer F, Hirst M, Tohyama K, Humphries RK, Lam WL, Marra M, Karsan A. Blood. 2010 Oct 20

[15] Bartel D: MicroRNAs: Genomics, biogenesis, mechanism, and function. Cell 116:281-297, 2004

[16] Eur J Haematol. 2010 Jan 1;84(1):1-16. Epub 2009 Sep 10. The role of microRNAs in normal and malignant hematopoiesis. Vasilatou D, Papageorgiou S, Pappa V, Papageorgiou E, Dervenoulas J.

[17] J Clin Oncol. 2009 Mar 10;27(8):1316-22. Epub 2009 Jan 21. Epigenetic regulation of microRNAs in acute lymphoblastic leukemia. Roman-Gomez J, Agirre X, Jiménez-Velasco A, Arqueros V, Vilas-Zornoza A, Rodriguez-Otero P, Martin-Subero I, Garate L, Cordeu L, San José-Eneriz E, Martin V, Castillejo JA, Bandrés E, Calasanz MJ, Siebert R, Heiniger A, Torres A, Prosper F.

[18] Expert Rev Mol Diagn. 2010 Apr;10(3):297-308. Micromarkers: miRNAs in cancer diagnosis and prognosis. Ferracin M, Veronese A, Negrini M.

[19] Cell. 2002 Apr 19;109(2):145-8. Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. Kiss T.

[20] Dong,X.Y., Guo,P., Boyd,J., Sun,X., Li,Q., Zhou,W. and Dong,J.T. (2009) Implication of snoRNA U50 in human breast cancer. J. Genet. Genomics, 36, 447–454.

[21] Yeoh, E.J., Ross, M.E., Shurtleff, S.A., Williams, W.K., Patel, D., Mahfouz, R., Behm, F.G., Raimondi, S.C., Relling, M.V., Patel, A., Cheng, C., Campana, D., Wilkins, D., Zhou, X., Li, J., Liu, H., Pui, C.H., Evans, W.E., Naeve, C., Wong, L. & Downing, J.R. (2002) Classification, subtype discovery, and prediction of outcome in paediatric acute lymphoblastic leukaemia by gene expression profiling. Cancer Cell, 1, 133–143.

[22] Mullighan CG, Miller CB, Su X, et al. *ERG* deletions define a novel subtype of B-progenitor acute lymphoblastic leukemia [abstract]. Blood. 2007; 110(11):212-213.

[23] Blood. 2010 Dec 2;116(23):4874-84. Epub 2010 Aug 10. Identification of novel cluster groups in pediatric high-risk B-precursor acute lymphoblastic leukemia with gene expression profiling: correlation with genome-wide DNA copy number alterations, clinical characteristics, and outcome. Harvey RC, Mullighan CG, Wang X, Dobbin KK, Davidson GS, Bedrick EJ, Chen IM, Atlas SR, Kang H, Ar K, Wilson CS, Wharton W, Murphy M, Devidas M, Carroll AJ, Borowitz MJ, Bowman WP, Downing JR, Relling M, Yang J, Bhojwani D, Carroll WL, Camitta B, Reaman GH, Smith M, Hunger SP, Willman CL.

[24] Campo Dell'Orto, M., Zangrando, A., Trentin, L., Li, R., Liu, W.M., te Kronnie, G., Basso, G., and Kohlmann, A. (2007). New data on robustness of gene expression signatures in leukemia: comparison of three distinct total RNA preparation procedures. BMC Genomics 8, 188.

[25] Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2-DDCT method. Methods 25:402–408

[26] Irizarry, RA, Hobbs, B, Collin, F, Beazer-Barclay, YD, Antonellis, KJ, Scherf, U, Speed, TP (2003) Exploration, Normalization, and Summaries of High Density Oligonucleotide Array Probe Level Data. Biostatistics .Vol. 4, Number 2: 249-264

[27] Adjusting batch effects in microarray expression data using empirical Bayes methods. Johnson WE, Li C, Rabinovic A. Biostatistics. 2007 Jan;8(1):118-27. Epub 2006 Apr 21.

[28] Opgen-Rhein, R., and K. Strimmer. 2007. Accurate ranking of differentially expressed genes by a distribution-free shrinkage approach. Statist. Appl. Genet. Mol. Biol. 6:9.

[29] Bradley Efron (2004) Large-Scale Simultaneous Hypothesis Testing: The Choice of a Null Hypothesis. Journal of the American Statistical Association, Vol. 99, No. 465 (Mar., 2004), pp. 96- 104

[30] Strimmer, K. (2008).A unified approach to false discovery rate estimation. BMC Bioinformatics 9: 303.

[31] Benjamini Y HY. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J Roy Stat Soc. 1995;57(Ser B):289-300.

[32] Regression Shrinkage and Selection via the Lasso. Robert Tibshirani Journal of the Royal Statistical Society. Series B (Methodological), Vol. 58, No. 1. (1996), pp. 267-288.

[33] Martin Slawski, Anne-Laure Boulesteix and Christoph Bernau. (2009). CMA: Synthesis of microarray-based classification. R package version 1.6.0.

[34] Most Mammalian mRNAs Are Conserved Targets of MicroRNAs Robin C Friedman, Kyle Kai-How Farh, Christopher B Burge, David P Bartel. Genome Research, 19:92-105 (2009).

[35] Nature Genetics 37, 495 - 500 (2005) Combinatorial microRNA target predictions. Azra Krek, Dominic Grün, Matthew N Poy, Rachel Wolf, Lauren Rosenberg, Eric J Epstein, Philip MacMenamin, Isabelle da Piedade, Kristin C Gunsalus, Markus Stoffel, Nikolaus Rajewsky

[36] Khatri P, Voichita C, Kattan K, Ansari N, Khatri A, Georgescu C, Tarca AL, Draghici S (2007) Onto-Tools: new additions and improvements in 2006. Nucleic Acids Res 35: W206–W211

[37] Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA 102, 15545–15550 (2005).

[38] Pui CH, Relling MV, Downing JR. Acute lymphoblastic leukemia. *N Engl J Med.* 2004;350:1535–48.

[39] Blood. 2010 Dec 2;116(23):4874-84. Epub 2010 Aug 10. Identification of novel cluster groups in pediatric high-risk B-precursor acute lymphoblastic leukemia with gene expression profiling: correlation with genome-wide DNA copy number alterations, clinical characteristics, and outcome. Harvey RC, Mullighan CG, Wang X, Dobbin KK, Davidson GS, Bedrick EJ, Chen IM, Atlas SR, Kang H, Ar K, Wilson CS, Wharton W, Murphy M, Devidas M, Carroll AJ, Borowitz MJ, Bowman WP, Downing JR, Relling M, Yang J, Bhojwani D, Carroll WL, Camitta B, Reaman GH, Smith M, Hunger SP, Willman CL.

[40] Blood. 2010 Dec 2;116(23):4874-84. Epub 2010 Aug 10. Identification of novel cluster groups in pediatric high-risk B-precursor acute lymphoblastic leukemia with gene expression profiling: correlation with genome-wide DNA copy number alterations, clinical characteristics, and outcome. Harvey RC, Mullighan CG, Wang X, Dobbin KK, Davidson GS, Bedrick EJ, Chen IM, Atlas SR, Kang H, Ar K, Wilson CS, Wharton W, Murphy M, Devidas M, Carroll AJ, Borowitz MJ, Bowman WP, Downing JR, Relling M, Yang J, Bhojwani D, Carroll WL, Camitta B, Reaman GH, Smith M, Hunger SP, Willman CL.

[41] Mullighan CG, Miller CB, Su X, et al. *ERG* deletions define a novel subtype of B-progenitor acute lymphoblastic leukemia [abstract]. Blood. 2007; 110(11):212-213.

[42] Mullighan, C.G., Goorha, S., Radtke, I., Miller, C.B., Coustan-Smith, E., Dalton, J.D., Girtman, K., Mathew, S., Ma, J., Pounds, S.B., Su, X., Pui, C.H., Relling, M.V., Evans, W.E., Shurtleff, S.A. & Downing, J.R. (2007) Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature, 446, 758–764.

[43] Blood. 2010 Dec 2;116(23):4874-84. Epub 2010 Aug 10. Identification of novel cluster groups in pediatric high-risk B-precursor acute lymphoblastic leukemia with gene expression profiling: correlation with genome-

wide DNA copy number alterations, clinical characteristics, and outcome. Harvey RC, Mullighan CG, Wang X, Dobbin KK, Davidson GS, Bedrick EJ, Chen IM, Atlas SR, Kang H, Ar K, Wilson CS, Wharton W, Murphy M, Devidas M, Carroll AJ, Borowitz MJ, Bowman WP, Downing JR, Relling M, Yang J, Bhojwani D, Carroll WL, Camitta B, Reaman GH, Smith M, Hunger SP, Willman CL.

[44] Effects of the leukemia-associated AML1-ETO protein on hematopoietic stem and progenitor cells. Nimer SD, Moore MA. Oncogene. 2004 May 24;23(24):4249-54. Review.

[45] Blood. 2010 Dec 2;116(23):4874-84. Epub 2010 Aug 10. Identification of novel cluster groups in pediatric high-risk B-precursor acute lymphoblastic leukemia with gene expression profiling: correlation with genome-wide DNA copy number alterations, clinical characteristics, and outcome. Harvey RC, Mullighan CG, Wang X, Dobbin KK, Davidson GS, Bedrick EJ, Chen IM, Atlas SR, Kang H, Ar K, Wilson CS, Wharton W, Murphy M, Devidas M, Carroll AJ, Borowitz MJ, Bowman WP, Downing JR, Relling M, Yang J, Bhojwani D, Carroll WL, Camitta B, Reaman GH, Smith M, Hunger SP, Willman CL.

[46] Zhang H, et al. (2009) MicroRNA patterns associated with clinical prognostic parameters and CNS relapse prediction in pediatric acute leukemia. PLoS ONE 4: e782

[47] Bousquet M, et al. (2008) Myeloid cell differentiation arrest by miR-125b-1 in myelodysplastic syndrome and acute myeloid leukemia with the t(2;11)(p21;q23) translocation. J Exp Med 205:2499–2506.

[48] Chapiro E, et al. (2010) A new recurrent translocation t(11;14)(q24;q32) involving IGH@ and miR-125b-1 in B-cell progenitor acute lymphoblastic leukemia. Leukemia 24:1362–1364

[49] Bousquet M, et al. (2008) Myeloid cell differentiation arrest by miR-125b-1 in myelodysplastic syndrome and acute myeloid leukemia with the t(2;11)(p21;q23) translocation. J Exp Med 205:2499–2506.

[50] Proc Natl Acad Sci U S A. 2010 Nov 30. [Epub ahead of print] MicroRNA miR-125b causes leukemia. Bousquet M, Harris MH, Zhou B, Lodish HF.

[51] Genes Dev. 2009 Apr 1;23(7):862-76. Epub 2009 Mar 17. MicroRNA-125b is a novel negative regulator of p53. Le MT, Teh C, Shyh-Chang N, Xie H, Zhou B, Korzh V, Lodish HF, Lim B.

[52] J Biol Chem. 2010 Jul 9;285(28):21496-507. Epub 2010 May 11. MicroRNA-125b confers the resistance of breast cancer cells to paclitaxel through suppression of pro-apoptotic Bcl-2 antagonist killer 1 (Bak1) expression. Zhou M, Liu Z, Zhao Y, Ding Y, Liu H, Xi Y, Xiong W, Li G, Lu J, Fodstad O, Riker AI, Tan M.

[53] Gefen N, et al. (2010) Hsa-mir-125b-2 is highly expressed in childhood ETV6/RUNX1 (TEL/AML1) leukemias and confers survival advantage to growth inhibitory signals independent of p53. Leukemia 24(1):89–96.

[54] Reprogramming of miRNA networks in cancer and leukemia. Volinia S, Galasso M, Costinean S, Tagliavini L, Gamberoni G, Drusco A, Marchesini J, Mascellani N, Sana ME, Abu Jarour R, Desponts C, Teitell M, Baffa R, Aqeilan R, Iorio MV, Taccioli C, Garzon R, Di Leva G, Fabbri M, Catozzi M, Previati M, Ambs S, Palumbo T, Garofalo M, Veronese A, Bottoni A, Gasparini P, Harris CC, Visone R, Pekarsky Y, de la Chapelle A, Bloomston M, Dillhoff M, Rassenti LZ, Kipps TJ, Huebner K, Pichiorri F, Lenze D, Cairo S, Buendia MA, Pineau P, Dejean A, Zanesi N, Rossi S, Calin GA, Liu CG, Palatini J, Negrini M, Vecchione A, Rosenberg A, Croce CM. Genome Res. 2010 May;20(5):589-99.

[55] Decreased expression of miR-125b and miR-100 in oral cancer cells contributes to malignancy. Henson BJ, Bhattacharjee S, O'Dee DM, Feingold E, Gollin SM. Genes Chromosomes Cancer. 2009 Jul;48(7):569-82.

[56] Cancer Res. 2011 Jan 6. [Epub ahead of print] miR-99 family of microRNAs suppresses the expression of prostate specific antigen and prostate cancer cell proliferation. Sun D, Lee YS, Malhotra A, Kim HK, Matecic M, Evans C, Jensen RV, Moskaluk CA, Dutta A.

[57] Genes Chromosomes Cancer. 2008 Sep;47(9):810-8. Detailed characterization of a homozygously deleted region corresponding to a candidate tumor suppressor locus at 21q11-21 in human lung cancer. Yamada H, Yanagisawa K, Tokumaru S, Taguchi A, Nimura Y, Osada H, Nagino M, Takahashi T.

[58] Am J Hum Genet. 2008 Aug;83(2):243-53. ATM gene mutations result in both recessive and dominant expression phenotypes of genes and microRNAs. Smirnov DA, Cheung VG.

[59] Genes Chromosomes Cancer. 2009 Jul;48(7):569-82. Decreased expression of miR-125b and miR-100 in oral cancer cells contributes to malignancy. Henson BJ, Bhattacharjee S, O'Dee DM, Feingold E, Gollin SM.

[60] Mol Endocrinol. 2010 Feb;24(2):447-63. Epub 2010 Jan 15. A link between mir-100 and FRAP1/mTOR in clear cell ovarian cancer. Nagaraja AK, Creighton CJ, Yu Z, Zhu H, Gunaratne PH, Reid JG, Olokpa E, Itamochi H, Ueno NT, Hawkins SM, Anderson ML, Matzuk MM.

[61] Cancer Res. 2010 Jun 1;70(11):4666-75. Epub 2010 May 18. Regulation of insulin-like growth factor-mammalian target of rapamycin signaling by microRNA in childhood adrenocortical tumors. Doghman M, El Wakil A, Cardinaud B, Thomas E, Wang J, Zhao W, Peralta-Del Valle MH, Figueiredo BC, Zambetti GP, Lalli E.

[62] Hypoxia induces BMP-2 expression via ILK, Akt, mTOR, and HIF-1 pathways in osteoblasts. Tseng WP, Yang SN, Lai CH, Tang CH. J Cell Physiol. 2010 Jun;223(3):810-8

[63] J Cell Physiol. 2006 Aug;208(2):336-43. Upregulation of secretory connective tissue growth factor (CTGF) in keratinocyte-fibroblast coculture contributes to keloid pathogenesis. Khoo YT, Ong CT, Mukhopadhyay A, Han HC, Do DV, Lim IJ, Phan TT.

[64] Int J Cancer. 2010 Dec 2. [Epub ahead of print] Human tumor MicroRNA signatures derived from large-scale oligonucleotide microarray datasets. Wang W, Peng B, Wang D, Ma X, Jiang D, Zhao J, Yu L.

[65] Biochem Biophys Res Commun. 2002 Nov 29;299(2):196-200. Differential expression of human 5S snoRNA genes. Chang LS, Lin SY, Lieu AS, Wu TL.

[66] SnoRNA U50 is a candidate tumor-suppressor gene at 6q14.3 with a mutation associated with clinically significant prostate cancer. Dong XY, Rodriguez C, Guo P, Sun X, Talbot JT, Zhou W, Petros J, Li Q, Vessella RL, Kibel AS, Stevens VL, Calle EE, Dong JT. Hum Mol Genet. 2008 Apr 1;17(7):1031-42. Epub 2008 Jan 17.

[67] Dong,X.Y., Guo,P., Boyd,J., Sun,X., Li,Q., Zhou,W. and Dong,J.T. (2009) Implication of snoRNA U50 in human breast cancer. J. Genet. Genomics, 36, 447–454.

[68] Sahoo T, Del Gaudio D, German J, Shinawi M, Peters S, Person R, et al. Prader–Willi phenotype caused by paternal deficiency for the HBII-85 C/D box small nucleolar RNA cluster. Nat Genet 2008;40:719–721.

[69] Mullighan, C.G., Goorha, S., Radtke, I., Miller, C.B., Coustan-Smith, E., Dalton, J.D., Girtman, K., Mathew, S., Ma, J., Pounds, S.B., Su, X., Pui, C.H., Relling, M.V., Evans, W.E., Shurtleff, S.A. & Downing, J.R. (2007) Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature, 446, 758–764.

[70] N Engl J Med. 2009 Jan 29;360(5):470-80. Epub 2009 Jan 7. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. Mullighan CG, Su X, Zhang J, Radtke I, Phillips LA, Miller CB, Ma J, Liu W, Cheng C, Schulman BA, Harvey RC, Chen IM, Clifford RJ, Carroll WL, Reaman G, Bowman WP, Devidas M, Gerhard DS, Yang W, Relling MV, Shurtleff SA, Campana D, Borowitz MJ, Pui CH, Smith M, Hunger SP, Willman CL, Downing JR; Children's Oncology Group.

[71] Mullighan CG, Miller CB, Su X, et al. *ERG* deletions define a novel subtype of B-progenitor acute lymphoblastic leukemia [abstract]. Blood. 2007; 110(11):212-213.

CHAPTER 2

Down Syndrome acute lymphoblastic leukemia: A highly heterogeneous disease in which aberrant expression of CRLF2 is associated with mutated JAK2 – a report from the iBFM-Study Group

Libi Hertzberg, Elena Vendramini, Ithamar Ganmore, Gianni Cazzaniga, Maike Schmitz, Jane Chalker, Ruth Shiloh, Ilaria Iacobucci, Chen Shochat, Sharon Zeligson, Gunar Cario, Martin Stanulla, Sabine Strehl, Lisa J Russell, Christine J Harrison, Beat Bornhauser, Akinori Yoda, Gideon Rechavi, Dani Bercovich, Arndt Borkardt, Helena Kempski, Geertruy te Kronnie, Jean Pierre Bourquin, Eytan Domany, Shai Izraeli.

PUBBLISHED IN BLOOD, Vol. 115 (1006–1017)

ABSTRACT

We report gene expression and other analyses to elucidate the molecular characteristics of acute lymphoblastic leukemia (ALL) in children with Down Syndrome (DS). We find that by gene expression DS ALL is a highly heterogeneous disease not definable as a unique entity. Nevertheless, 62% (33/53) of the DS ALL samples analyzed were characterized by high expression of the type I cytokine receptor *CRLF2* caused by either *IgH@* translocations or by interstitial deletions creating chimeric transcripts *P2RY8-CRLF2*. In 3 of these 33 patients a novel activating somatic mutation, F232C in *CRLF2* was identified. Consistent with our previous research, mutations in R683 of *JAK2* were identified in 10 specimens (19% of the patients) and interestingly all 10 had high *CRLF2* expression. CRLF2 and mutated Jak2 cooperated in conferring cytokine independent growth to BaF3 pro-Bcells. Intriguingly the gene expression signature of DS ALL is enriched with DNA damage and BCL6 responsive genes, suggesting the possibility of B-cell lymphocytic genomic instability. Thus DS confers increased risk for genetically highly diverse ALLs with frequent overexpression of *CRLF2*, associated with activating mutations in the receptor itself or in *JAK2*. Our data also suggest that the majority of DS children with ALL may benefit from therapy blocking the CRLF2-JAK2 pathways.

INTRODUCTION

Children with Down Syndrome (DS) have a higher rate of acute lymphoblastic leukemia (DSALL). DS ALLs are mostly of B-cell precursor (BCP) origin and similar in the age of diagnosis and immunophenotype to high hyperdiploid (HD) or *TEL-AML*1 ALLs,[1] the two most common genetic subtypes of childhood ALL. Given that these cytogenetic abnormalities are less frequent in DS ALL,[2] the existence of unique collaborating somatic genetic events in DS ALL, similar to the *GATA1* mutation in DS-AMKL[3] has been postulated. We and others reported the presence of somatic activating mutations in *JAK2* in approximately 20% of DS ALL.[4-6] Similar mutations are present in about 10% of high-risk ALL in non-DS children corresponding to approximately 3% of unselected childhood ALLs.[7] We hypothesized that the mutated JAK2 may cooperate with a type I cytokine receptor that is aberrantly expressed in DS ALL.4 To characterize additional molecular abnormalities in DS ALL, we performed genomic analysis of a large group of DS ALLs. This analysis reveals, next to a striking heterogeneity of these leukemias, an aberrant expression of the cytokine receptor *CRLF2* in 62% of the patients, associated with somatic activating mutations in *JAK2* or in the receptor itself.

MATERIAL AND METHODS

Patient Samples.

RNA and DNA were derived from diagnostic bone marrow samples of children with DS and BCP ALL enrolled on treatment protocols with an informed consent and approval of local and national ethic committees. Samples were anonymized for the study. Patients' clinical data is described in supplementary tables 1S and 2S. 76 of these patients were included in our previous publication describing the *JAK2* mutations in DS ALL.4 The study was approved by the Israeli Health Ministry Ethic committee, approval # 920070771.

Genomic Studies.

RNA processing and hybridization to Affymetrix arrays was performed according to manufacturer instructions and as previously published.[8,9] Only specimens containing more than 70% blasts were included. There were four datasets obtained by different teams as summarized in Table 1. AIEOP (Associazione Italiana Ematologia Oncologia Pediatrica) is

the main dataset used for gene expression analysis, comprising 97 diagnostic ALL samples: 25 DS ALL, and 72 non DS ALL samples, described in Table 1. The additional three datasets were utilized for validations. The primary gene expression data files have been deposited in NCBIs Gene Expression omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) GEO Series accession number GSE17459. Genomic DNA from 42 diagnostic bone marrow ALLs and 34 paired remission samples were genotyped with Affymetrix GeneChip Human Mapping 100K set (Affymetrix, Santa Clara, CA) according to the manufacturer's directions. See supplementary files for details.

Mutation analysis was performed as we previously described.[4,11] *CRLF2* (NM_022148.2) primers are described in Table 3S. Quantitative real-time-PCR (qRT-PCR) was performed using Applied-Biosystems TaqMan® Gene-Expression Assays (*CRLF2* Hs_00845692, *GAPDH* Hs_99999905) according to the manufacturer's instructions. Each sample was run in triplicate. The endogenous control gene was *GAPDH*.

Fluorescence In-Situ Hybridization (FISH) for detection of *IGH@-CRLF2* translocation or the presence of a microdeletion upstream to *CRLF2* was performed as described[10].

Flow Cytometry analysis (Becton-Dickinson Canto-II, FlowJo software) was performed on primary cryopreserved ALL cells after the first xenotransplantation in Nod/LtSzScid IL2γnull mice. Antibodies used were anti-CRLF2 (Abcam clone 1D3, ab48482), goat anti-mouse Alexa Fluor 488 (Invitrogen), anti-IL7RA Alexa 647 (CD127, BD clone HIL-7R-M21), anti-CD19 PE (AbD-Serotec, clone LT19, MCA1940) and 7-AAD (AbD-Serotec). All samples were gated on the viable (7-AAD negative) and leukemic (CD19 positive) population before analysis of CRLF2 and IL7RA. For the calculation of delta MFI (mean fluorescence intensity), background non-specific staining was evaluated in populations gated by CD19, comparing tubes with or without anti-CRLF2 antibodies. This background MFI was similar to the MFI of CRLF2 negative populations in normal human blood.

Plasmid construction The FLAG-mJAK2 wild-type and R683S were cloned into the pHRSINCSGW lentivirus11 which carries SFFV promoter and an emerald-GFP reporter. pMX-Puro-hCRLF2 was used as a template for the generation of *CRLF2* mutations by site-directed mutagenesis (QuikChangeTM-II-XL, Stratagene).

Cell lines

BaF3 cells were cultured in RPMI-1640 containing 10% fetal calf serum and 10% WHEI-3B conditioned media as a source of interleukin 3. Parental BaF3 cells were transduced with

pMX-Puro-hCRLF212 and *hCRLF2* expressing cells were selected with puromycin (2μg/ml). Parental BaF3 and BaF3-CRLF2 cells were transduced with the appropriate *Jak2* expressing vector and GFP-positive cells were sorted by flow cytometry 3-4 days later.

BaF3 proliferation assays and Western blotting were performed as descrive before.[4] Antibodies used were anti-JAK2 (C-20, Santa Cruz), anti-STAT5, antiphospho-JAK2 Tyr-1007 (Cell Signaling), anti-phospho-STAT5 Tyr-694 (Epitomics), anti-hTSLPR (AF981)(R&D Systems), anti-FLAG-M2 and anti-α-tubulin (Sigma- Aldrich).

Pharmacological inhibition of JAK2 BaF3 cells expressing Jak2 R683S and BaF3/CRLF2 cells expressing wt or R683S Jak2 were cultured without cytokines in different concentrations of JAK inhibitor I (Calbiochem, La Jolla, CA, USA). Controls were BaF3/EpoR cells expressing BCR-ABL. Viable cells were counted after 72 h. Data from three independent experiments were combined for analysis. We calculated the normalized viability by dividing the cell number at each inhibitor concentration by the cell number with vehicle alone.

Bioinformatics

Gene expression preprocessing is described in Supplementary Methods.

Combining probesets of the same gene. For those genes that were represented by more than one probe-set, we used, when needed, a combination procedure to create a single representation of a gene's expression (Supplementary Methods).

Gene Set Enrichment Analysis (GSEA). The first ingredient of GSEA[13] is a list of genes *L*, ranked by some attribute *A*, ordered from low to high values of *A*. The second ingredient is a set of genes *S* that are a subset of *L*. GSEA aims at answering if the members of *S* are randomly distributed along the ranked list *L*, or if they are skewed towards one of the sides For details see Supplementary Methods and[13].

Refining DS ALL profile genes. The preliminary DS ALL profile gene list, that was constructed using the AIEOP data set was narrowed down using GSEA[13] to select genes that show consistent expression pattern in at least two of the other three datasets (Table 1). We used the up-regulated members of the preliminary DS ALL genes as our set *S* (see above) and the genes of one of the other three experiments constituted *L*. The genes of *L* were ordered according to their differential expression in DS versus the rest of the samples (see Refining DS ALL profile genes in Supplementary Methods). The process was repeated for each of the three datasets and for the down-regulated genes, yielding for each case genes that were

identified as consistently up (or down) regulated in the AIEOP dataset and the other dataset tested (Figure 2A).

RESULTS

Marked heterogeneity of DS ALLs revealed by unsupervised gene expression analysis.

We first explored the extent of similarity between DS ALLs and other defined BCP ALL genetic subtypes using in the analysis the 1500 probe-sets with the highest standard deviation among the AIEOP samples. Two unsupervised analysis algorithms were used: SPIN (Sorting Points Into Neighborhoods),[14] that places samples with similar expression profiles near each other (Figure 1A), and PCA (Principal Component Analysis, MATLAB 7.4 software) (Figure 1B). Both gave similar results. In agreement with previous studies,[15] unsupervised analysis of gene expression tends to group the pediatric ALL samples according to their genetic subtypes. As can be seen in the Euclidian distance matrix in Figure 1A, the ALL subgroups that are the most homogenous (exhibiting high similarity of samples of the same subtype) are *E2A-PBX1* positive ALL and *MLL-AF4* positive ALL, followed by *TEL-AML1* positive ALL. HD-ALL samples are also grouped together, but are relatively more distant from each other than the above mentioned samples. Although *BCR-ABL* ALLs are clustered together, they are less homogeneous, consistent with previous reports.[15] In contrast, DS ALLs are very heterogeneous (Figures 1A, 1B). About half are grouped together, relatively close to both *BCR-ABL* and HD-ALL. However even here, individual DS samples are more separated from each other than a typical pair of samples within the other ALL subtypes (Figures 1A,1B). The other half are grouped with other ALL subtypes: six with *TEL-AML1*, six with HD, two with *BCR-ABL* and one with *E2A-PBX1*. Of these 15 DS ALLs only three carried the chromosomal translocation of the subtype of ALL to which they are most similar (one *E2A-PBX1*, two *TEL-AML1*) and only one DS ALL sample was found to be also HD. Even the five DS ALL samples with somatic mutations in *JAK2* (blue boxes below Figure 1A, black circles at Figure 1B) are not clustered together. This unsupervised gene expression analysis reveals that DS ALLs are markedly less homogenous than the other ALL genetic subtypes. It suggests that DS is a *predisposing condition* to several genetic subtypes of B-cell precursor ALLs, and that unlike the myeloid leukemia of DS should not be considered as a unique molecular entity.

Genomic analysis of DS ALL

We performed 100K SNP-array analysis of 34 paired diagnosis and remission samples (15 DS, 9 HD- and 10 *TEL-AML* ALLs, Supplementary Methods). Copy number and loss of heterozigosity (LOH) analyses (Supplementary Figure 3S and Table 10S) generally confirm previous reports[5,16] that deletions are more common in DS and *TEL-AML1* compared with HD ALLs. The frequency of deletions in genes regulating normal B-lymphoid development (supplementary table 9S) in DS ALL was 53%, slightly higher than the 40% reported for of BCP ALL.[17] Recently, deletions in the *IKZF1* gene were reported in the majority of patients with *BCR/ABL* and "*BCR/ABL* like" ALL.[18,19] Since most of these deletions involve only a subset of exons (most commonly exons 4-7), the 100K SNP platform is inadequate to detect these abnormalities. Therefore 38 additional diagnostic DS ALL specimens were screened for *IKZF1* deletions by PCR analysis as previously reported.[20] Monoallelic *IKZF1* deletions were identified in nine patients (24%) (Supplementary Table 8S and Figure 3S). Thus the frequency of deletions in B-cell differentiation genes, including *IKZF1*, in DS ALL is similar to other non *BCR-ABL* subtypes of BCP ALL.

DS ALLs gene expression profile.

We hypothesized that despite their heterogeneity, DS ALLs share a common gene expression signature. We reasoned that by comparing the gene expression in DS ALL to the relatively similar groups, HD and *TEL-AML1*, we could potentially isolate the “DS ALL” characteristics from the other ALL characteristics that might be similar between these groups. In addition, the analysis was done in a way that only genes which differentiate DS from *TEL-AML1 and* from HD ALLs are depicted. The fact that *TEL-AML1* and HD ALLs have dissimilar expression profiles (denoted by dark red entries in Figure 1A) helps to identify genes that characterize DS ALL, and not one of the groups to which it is compared. We first identified probe-sets that had significant differential expression in DS ALL samples, compared to both HD and *TEL-AML1* ALLs in the AIEOP dataset (Supplementary Methods). This "preliminary DS ALL profile" consisted of 792 genes upregulated and 535 genes downregulated in DS ALL. To check consistency with each of the other three gene expression datasets ("BFM", "ICH" and "IL"), we performed GSEA[21] separately on each of the three. A representative enrichment analysis is shown in Figure 2A; here the genes are ordered according to their DS ALL differential expression in the ICH dataset (used as the list *L*, see Methods), and the 535

genes downregulated in DS ALL are used as the set *S*, tested for enrichment. Six such analyses showed significant enrichment of the preliminary gene lists (both up and down-regulated) obtained from AIEOP, in the other three datasets (Table 2). We refined our AIEOP-based lists by including only genes that showed consistent expression patterns in at least two of the three other datasets. The "refined DS ALL profile genes" (Figure 2B) consists of 152 up and 199 downregulated genes (Tables 4S and 5S).

Pathway analysis and BCL6 signature

To identify molecular pathways that showed differential expression in the refined DSALL profile we interrogated the DAVID database of Gene Ontology functional categories[22]. Constituent genes of 8 pathways were significantly (False Discovery Rate, FDR < 10%) over-represented in the DS ALL expression profile (Table 3 and Supplementary Table 6S). The most enriched pathway (p=5×10-4) is "Response to DNA damage stimulus": Ten of the 341 genes assigned by DAVID to this pathway are downregulated and 6 are upregulated in DS ALL. One of the upregulated genes is *BCL6*, with a mean fold change of 1.46 in DS ALL compared with non DS ALL (Table 4S). BCL6 is a transcription factor primarily expressed in mature B-cells at the germinal centers, where it facilitates Ig affinity maturation by repressing the DNA damage response. It is also a known oncogene in diffuse large B-cell lymphomas (DLBCLs).[23,24]

To search for evidence for BCL6 activity in the DS ALL gene expression profile, we used the Oncomine (http://www.oncomine.org)[25] database, in which cancer gene expression signatures derived from different expression analyses are stored as Molecular Concept Maps (MCMs). These are lists of differentially expressed genes between two logical groupings of normal or malignant human tissue or cell lines. We tested BCL6 direct targets and each of the 24 Oncomine MCMs that involve BCL6 (Table 7S) for enrichment in DS ALL up and down-regulated genes, and eight of these 25 gene groups passed at False Discovery Rate (FDR)[26] of 15% (Table 4). These include the target genes of BCL6,[27] genes modified by ectopic expression of BCL6 in lymphoblastoid B-cells[28] and the gene expression signature of B-cell lymphomas with oncogenic activation of BCL6.[29] Hence the targets and pathways downstream to BCL6 in lymphomas and mature B-cells are modified in the DS ALL expression profile.

Aberrant expression of the cytokine receptor CRLF2 in DS ALLs

We have previously hypothesized that a cytokine receptor may be aberrantly expressed in DS ALL and cooperate with JAK2 carrying the "lymphoid" mutation in R683.[4] Examination of the DS ALL expression signature (Table 4S) reveals that the third most differentially expressed gene is *CRLF2* (cytokine receptor-like factor 2, *TSLPR*) located at the pseudoautosomal region of the sex chromosomes. As depicted in Figure 3A, increased expression of *CRLF2* was noted in 23 (62.1%) DS ALLs out of 37 samples that were hybridized to U133 family of arrays (*CRLF2* is not represented on the Exon Arrays used in the "IL" dataset), compared with other ALL subtypes. *CRLF2* expression along DS ALL versus all other ALL subtypes yielded ttest P values of 3.7×10^{-11}, 2.2×10^{-7}, 5.6×10^{-5} for AIEOP, BFM and ICH datasets, respectively. CRLF2 is known to dimerize with IL7RA to form the heterodimeric receptor for TSLP (thymic stromal lymphopoietin).[30] While *CRLF2* is aberrantly expressed in DSALLs (Figure 3A), expression of *IL7RA* is similar in the different ALL subtypes (Figure 3B).

To validate the findings of the expression arrays and to analyze additional DS ALL samples we measured the expression of *CRLF2* by qRT-PCR in 32 patients (Figure 3C). Microarray data was available for 16 of these cases. The qRT-PCR confirms the *CRLF2* expression levels seen in the arrays (Pearson correlation=0.85, $p<0.001$). In two patients *CRLF2* expression was analyzed in RNA derived from diagnostic and remission bone marrows (BM) and was seen only in the diagnostic sample. In one patient similar *CRLF2* expression levels were seen in BM samples from diagnosis and relapse (Supplementary Figure 2S). Altogether, 33 out of 53 (62.3%) DS ALL patients analyzed by either qRT-PCR or Microarrays overexpressed *CRLF2*. The surface expression of the CRLF2 protein was also verified on 4 samples by flow cytometry (Figure 3D). IL7RA is also expressed on the leukemic blasts independent of CRLF2 expression.

Recently Russell et al[10] reported aberrant expression of *CRLF2* caused by either chromosomal translocations to the *IGH@* locus or interstitial deletions upstream to *CRLF2* juxtaposing *CRLF2* with the *P2RY8* regulatory elements in about 5% of childhood ALLs. To examine if the increased *CRLF2* expression in our specimens was caused by the same genomic aberrations, 12 available diagnostic DS ALL samples overexpressing *CRLF2* were analyzed by FISH (Figure 4A). *IGH@* translocations were seen in 4 specimens and interstitial deletions in 7. In the remaining sample (#DS-32 Table 2S), in which the CRLF2 expression level was just above the threshold, the FISH pattern of CRLF2 appeared normal. Further evidence

supporting the presence of the deletions is provided by a statistically significant inverse correlation between *CRLF2* and *P2RY8* expression (p=0.02, Figure 4B).

To test if the deletion caused a fusion between the *P2RY8* and *CRLF2* we performed RT-PCR with primers derived from both genes (table 3S). A transcript fusing the first non-coding exon of *P2RY8* and the first exon of *CRLF2* prior to the ATG was detected in the two DS patients with the deletion detected by FISH but not in the patient with the IgH@ translocation (Figure 4C). A similar chimeric transcript was described in a single patient with splenic lymphoma, fusing *P2RY8* to *SOX5* resulting in overexpression of SOX5[31]. We extended the analysis and identified the chimeric transcript in seven of 10 patients with overexpression of *CRLF2* and in none of 8 samples with no expression of *CRLF2* (Table 2S). Thus, consistent with the FISH findings[10], the interstitial deletion is more common than the IgH@ translocation.

To explore the effect of CRLF2 on gene expression we compared the 30% of DSALLs with the highest *CRLF2* expression and the 30% of DS ALL with the lowest *CRLF2* expression in the AIEOP database. Only 5 probe-sets passed FDR of 30%, with *CRLF2* being one of the 5 (Table 5). This is consistent with the finding that samples that over-express *CRLF2* (red marks under Figure 1A) do not cluster separately from DS ALLs that do not express *CRLF2*. Interestingly, the *IGJ* gene which differentiate these two groups (Fold change 36.4, Table 5) is also the most differentiating gene between DS ALL and non DS ALL in our datasets (Table 4S).

Clinical significance of CRLF2 expression in DS ALL

Clinically, children with high/medium expression of *CRLF2* were diagnosed younger (table 6) than children with no/low expression of *CRLF2* (5.56 vs. 9.87 yrs, p=0.004). No significant differences between the two groups regarding their sex or WBC count at diagnosis were found. Patients expressing *CRLF2* tended to have a lower probability for event free survival (Supplementary Figure 4S, p=0.12 log-rank test).

Cooperation between JAK2 R683 mutations and CRLF2 aberrant expression

Among the 53 DS ALL samples for which *CRLF2* expression was available, 10 had somatic mutations in *JAK2* R683. We identified chimeric *P2RY8-CRLF2* transcripts in three additional patients with *JAK2* R683 mutations (Table 2S). Thus all mutations occurred in

specimens with aberrant expression of *CRLF2,* supporting our initial hypothesis that CRLF2 may act as type I cytokine receptor for mutated JAK2.

To examine if CRLF2 and mutated JAK2 cooperate, we generated BaF3 cells that express hCRLF2 (BaF3-CRLF2) and transduced both BaF3 and BaF3-CRLF2 cells with wild-type m*Jak2-FLAG*, R683S m*Jak2-FLAG*, and empty vector. As depicted in Figure 5A there was synergism between CRLF2 and both wt Jak2 and R683S mutated Jak2, with the best cytokine independent growth observed in cells expressing CRLF2 and the mutated Jak2. These functional effects on cell growth are reflected in protein analysis of the JAK-STAT pathway (Figure 5B). Interestingly, despite identical levels of CRLF2 at the time of transduction, the levels of CRLF2 were consistently higher in cells transduced with Jak2 compared to empty vector or parental cells. Examination of STAT5 and Jak2 phosphorylation five hours after cytokine withdrawal reveals that when CRLF2 was expressed, phosphorylation levels in cells transduced with wt Jak2 were increased, while no change was observed in the already high phosphorylation levels in cells expressing the mutated Jak2. The marked advantage in cytokine independent growth rate of cells co-expressing CRLF2 and R683S Jak2 despite similar STAT5 phosphorylation may indicate the involvement of additional signalling pathways.

To test if the cells expressing CRLF2 and/or either wt or R683 mutated Jak2 depend on activated JAK-signaling, we incubated BaF3 cells transduced with the different vectors cultured without IL3 in the presence of different concentrations of JAK inhibitor 1 (Figure 5C). Although BaF3 cells transduced with CRLF2/Jak were more sensitive to the inhibitor compared with the control cells expressing BCR-ABL ($p=0.04$, ANOVA), the cells expressing CRLF2 and mutated Jak2 were the least sensitive.

Activating mutations of CRLF2 in DS ALL

To identify additional events leading to *CRLF2* activation we screened 87 diagnostic DS ALL samples for mutations in *CRLF2* (Figure 6). In addition to polymorphisms V136M and V244M that were present also in remission samples and in healthy controls, we identified in three patients a somatic mutation replacing phenylalanine 232, located at the juxtamembraneous domain, with cysteine (F232C). Genomic data was available for one of the patients (#DS-97) who displayed the *P2YR8-CRLF2* transcript. Although F232C induced constitutive STAT5 phosphorylation in cytokine deprived BaF3 cells (Figure 6E) it did not provide a consistent survival advantage; While during the first few days after cytokine

withdrawal more cells expressing F232C CRLF2 were alive compared with cells expressing wt CRLF2, at day 7 almost all BaF3 cells were dead (not shown). To examine the collaboration with wt Jak2, BaF3 cells stably expressing wt Jak2 were transduced with retroviral vectors expressing either wt CRLF or F232C CRLF2 (Figure 6D, E). In the presence of exogenous wt Jak2, there was about fifteen fold increase in the growth rate of cells expressing the mutant CRLF2 compared with those expressing wt CRLF2 ($p=0.02$, paired t-test). Together these observations demonstrate that the F232C CRLF2 activates JAK-STAT signaling and cooperates with JAK2 to provide significant growth advantage in a cytokine deprived environment.

DISCUSSION

Here we report the results of a genome wide study of DS ALL based on a dataset of unprecedented size. Unexpectedly, the molecular phenotype obtained by gene expression profiling is strikingly less homogeneous in DS ALL than any of the common genetic subtypes of childhood BCP ALLs. However, despite this heterogeneity, we describe a major feature that is shared by up to two-thirds of the patients - the aberrant expression of the wt or mutated cytokine receptor *CRLF2* and its association with mutations in *JAK2*.

That DS ALL is less uniform than the specific DS-associated myeloid leukemia has been suggested by a large cytogenetic study performed by the iBFM-SG.[2] However, neither that study, nor the genomic analysis reported here or previously[5,16] explain the level of inhomogeneity in gene expression. Even those DS ALLs that clustered together were not similar to each other. Such heterogeneity suggests that unlike the common aberrations of childhood ALL (*TEL-AML1*, Hyperdiploidy, *E2A-PBX1* etc), constitutional trisomy 21 is not a typical initiating event. Rather, DS is a predisposing condition to multiple genetic subtypes of BCP ALLs.

To identify genes and pathways common to DS ALLs we have generated a DS ALL gene expression signature, exploiting the advantage of having several datasets. *CRLF2* is one of the three genes most differentiating between DS and non-DS ALLs. Confirming the expression of *CRLF2* RNA and protein in DS ALLs and extending these observations to patients for whom array data was not available, we observed increased expression of *CRLF2* in 62% of 53 patients with DS ALL. These data are corroborated by the recent report describing IGH@ translocations or interstitial deletions upstream to *CRLF2* in 5% of non selected childhood ALL, and in 35 of 68 DS ALL (52%) consecutively enrolled in UK treatment protocols[10].

We report that the interstitial deletion results in fusion transcript in which the first non-coding exon of *P2YR8* fuses to the coding region of *CRLF2*, thereby driving *CRLF2* expression by the *P2YR8* promoter. A similar mechanism was reported in a single patient with splenic lymphoma and *P2YR8-SOX5* fusion[31] and is reminiscent of the common *SIL-SCL (STIL-TAL1)* rearrangement in T-ALL[32]. Cloning of the genomic breakpoints is required to determine if, like *SIL-SCL* the deletion is caused by aberrant V(D)J activity.

Although we do not have genomic data for all *CRLF2* expressing samples, the FISH and RT-PCR results of 17 out of 33 specimens overexpressing *CRLF2*, the inverse correlation between *CRLF2* and *P2YRY8* expression and the similar frequency of *CRLF2* overexpression in our and Russell's et al[10] two independent cohorts, suggest that most, if not all, aberrant *CRLF2* expression is caused by genomic rearrangements.

CRLF2 dimerizes with IL7RA to form the receptor to thymic stromal-derived lymphopoietin (TSLP), an epithelial derived cytokine that plays a role in inflammation and lymphoid development.[12, 33-35] The expression of *IL7RA* on the leukemic blasts suggest that some of the aberrantly expressed CRLF2 may interact with IL7RA and form a TSLP receptor on the leukemic cells. However we also demonstrate that CRLF2 cooperates with Jak2 to transform BaF3 cells lacking expression of IL7RA (Figure 5A and Supplementary Figure 5S). This suggests that CRLF2 may act independently of IL7RA, possibly through homo-dimerization similar to other type I cytokine receptors.

We report an unusual cooperation between CRLF2 and ectopically expressed wt Jak2 in BaF3 cells, a phenomenon not observed with other type I cytokine receptors such as EPOR or TPOR. CRLF2 is an atypical type I cytokine receptor that contains only one of the two "boxes" that mediate binding of JAK enzymes and only one tyrosine in its C-terminal domain. Hence it is a weak activator of JAK236. This may explain the requirements for higher levels of Jak2 for activation of the Jak-Stat pathway. Interestingly the levels of CRLF2 were higher in the presence of ectopically expressed wt or mutated Jak2. Positive regulation of the expression of a type I cytokine receptor by JAK2 and Tyk2 was previously reported.[37-39] Thus one mechanism by which Jak2 may cooperate with CRLF2 is by increasing the expression of the latter.

We observed two acquired events associated with the increased expression of *CRLF2*. The most common event is activating "lymphoid" somatic mutation in *JAK2*. All DS ALLs specimens with *JAK2* mutations in our series and in the cohort reported by Russell et al[10] had aberrant expression of *CRLF2*, strongly implying that CRLF2 is the cytokine receptor

cooperating with R683 mutated JAK2. Indeed, in BaF3 cytokine weaning assays, only the combination of CRLF2 and mutated Jak2 lead to a robust cytokine independent growth, demonstrating for the first time that these two proteins cooperate in providing growth and survival advantage.

The second less common event is an activating mutation in *CRLF2*. Weinstock et al[40] reported an E40G activating somatic mutation in *CRLF2* in a single patient with adult BCP ALL. We now found that 3 of the 33 patients with DS ALL overexpressing *CRLF2* have a somatic mutation replacing phenylalanine in the juxtamembrane position 232 by cysteine. This mutation caused constitutive phosphorylation of STAT5 associated with robust cytokine independent growth of BaF3 cells ectopically transduced with wt Jak2. Introduction of cysteines in this region in the erythropoietin receptor, another type I cytokine receptor signaling through JAK2, caused its constitutive activation by enhancing ligand independent dimerization.[41]

Although several scenarios may be possible, a reasonable model (Figure 7) is that the overexpression of *CRLF2* is the first event occurring in about 60% of DS ALL patients. The expanded pre-leukemic clone then acquires additional genetic aberrations, among them an activating mutation in *JAK2* or *CRLF2* or thus far unidentified events that may involve the JAK-STAT pathway. This model explains three key observations: (a) All samples with mutated *JAK2* and the only evaluable patient with mutated *CRLF2* also had aberrant CRLF2 expression (b) Many *CRLF2* overexpressing samples do not have mutation in *JAK2* (c) In one reported patient 10 an aberrant *CRLF2* genomic rearrangement was present at diagnosis while mutant JAK2 was present only in the relapse sample.

The most intriguing question is why there is a dramatic 10-fold increase in genomic lesions causing *CRLF2* overexpression in DS (60% in DS ALL compared with 5% in sporadic ALL) and how this relates to trisomy 21? Only a single Hsa21 gene, *SON* was included in the DS ALL signature and it was only slightly (1.3) upregulated (see Table 4S). Indeed we found no major difference in the gene expression from the trisomic chromosome 21 between DS ALL and HD-ALL (data not shown). Yet our data suggests that DS ALL and HD-ALL are to a great extent different leukemias. There are obvious fundamental differences between constitutional and acquired trisomy[42] such as the developmental stage in which the trisomy occurs and the fact that a constitutional trisomy is present both in the leukemia cells and in their microenvironment.

Regardless the role of the constitutional trisomy, our data generates an intriguing hypothesis. We observe a significant enrichment in DNA damage and repair genes in DS ALL and identify increased expression and clear "footprint" of BCL6 in these leukemias. BCL6 regulates the germinal center B cell maturation, through its effects on the DNA damage response. Recent studies by Muschen et al[43] suggest for the first time a role for BCL6 in BCP ALL. We speculate that DS may predispose to ALL through B cell lymphocytic specific genomic instability involving *BCL6*. The signatures of BCL6 and the DNA damage response pathway may be related to previous reports on impaired cellular response to DNA damage in DS[44] and to the increased prevalence of *IgH@* chromosomal translocation in DS ALL[10,45]. At present, however, it is impossible to determine if the *BCL6* signature precedes or follows the *CRLF2* rearrangements. As high expression of *CRLF2* blocks B cell differentiation[10], one cannot exclude the possibility that it causes a developmental arrest of the preleukemic cell in a stage in which *BCL6* is active. Distinguishing between these two hypotheses will require the identification and study of pre-leukemic cells in children with DS.

Finally our data imply that therapeutics targeting JAK-STAT signaling may be of potential benefit to the majority of DS ALLs not limited only to those with mutated JAK2. Although we demonstrate that BaF3 cells co-expressing CRLF2 and mutated Jak2 are more susceptible to JAK inhibitor 1 than cells transformed with BCR-ABL, they were relatively resistant in comparison with cells transformed only with mutated Jak2. This preliminary observation requires further testing in primary leukemic cells. It may indicate that targeting other pathways activated by CRLF2 or the utilization of anti CRLF2 specific antibodies will synergize with JAK2 inhibitors in treatment of DS ALL and non-DS ALL with aberrant *CRLF2* expression.

TABLES

Table 1. Description of the gene expression datasets analyzed

Dataset Symbol	No. samples	DS ALL	HD	*TEL-AML1*	Other	platform
AIEOP	97	25	26	29	4 *E2A-PBX1* 6 *MLL* 7 *BCR-ABL*	Affymetrix HG- U133 Plus 2.0
ICH	15	6	5	4		Affymetrix HG- U133 Plus 2.0
BFM[1] + St Jude[15]	7 29	7	 12	 17		 Affymetrix HG-U133A
IL	27	11	10	6		Affymetrix Exon 1.0 ST

***** IL is partially overlapping with AIEOP (5 DS ALLs) and BFM (4 different DS ALLs). Abbreviations: AIEOP Associazione Italiana Ematologia Oncologia Pediatrica; ICH – Institute Child Health; BFM – Berlin Frankfurt Munster; IL- Israel

Table 2. GSEA for genes of the DS ALL expression signature, identified from the AIEOP dataset, in three other datasets

DS ALL up-regulated genes	ES (enrichment score)	Nominal p-value	FDR q-value
ICH dataset	0.437993	0.01002	0.008979
IL dataset	0.361621	0.035124	0.022851
BFM dataset	0.344479	0.001996	0.009101
DS ALL down-regulated genes			
ICH dataset	-0.50601554	0.005617978	0.012959
IL dataset	-0.5008136	0.01192843	0.006397
BFM dataset	-0.55240697	0.001976285	0.001131

Table 3. Gene ontology pathways over-represented in the differential DS ALL signature. For details see table 6S

Gene Ontlology Group	Size	P Value	FDR
GO:0006974~ response to DNA damage stimulus	16	0.000503	0.77%
GO:0006397~ mRNA processing	13	0.001857	2.80%
GO:0015031~ protein transport	24	0.00216	3.30%
GO:0008104~ protein localization	26	0.002438	3.70%
GO:0046907~ intracellular transport	24	0.002913	4.40%
GO:0065003~ macromolecular complex assembly	20	0.003825	5.80%
GO:0051649~ establishment of cellular localization	27	0.005078	7.60%
GO:0043067~ regulation of programmed cell death	19	0.006188	9.10%

Table 4. Enrichment of BCL6 related gene expression signatures and direct targets in DS ALL profile genes (FDR < 15%).

Oncomine "Molecular Concepts" enriched in DS ALL up-regulated genes	Size	P-value	Q-value
EREB Lymphoblastoid CellLine BCL6 transfection top 10% over-expressed [28]	17	0.003733	0.06533
Ramos Burkitt Lymphoma CellLine BCL6 Pest mutant top 5% under-expressed in Anti IgM [28]	11	0.01138	0.09962
Lymphoma BCL6 break top 5% over-expressed[29]	11	0.01866	0.1291
Lymphoma BCL6 break top 10% over-expressed[29]	17	0.02213	0.1291
BCL6 direct targets[27]	11	0.02951	0.1291
Oncomine "Molecular Concepts" enriched in DS ALL down-regulated genes			
Lymphoma BCL6 break top 10% under-expressed[29]	36	7.89×10^{-5}	0.00209
EREB Lymphoblastoid CellLine BCL6 transfection top 10% under-expressed [28]	26	0.00011	0.00209
Lymphoma BCL6 break top 5% under-expressed[29]	20	0.001007	0.009569

Table 5. CRLF2 differentiating genes. List of statistically significant differentiating genes (FDR 30%) between 8 DS ALL samples with highest *CRLF2* expression and 8 DS ALL samples with lowest *CRLF2* expression in AIEOP dataset. The fold change value is between the mean expression in the two groups

Probe set Id	Gene Symbol	Description	Band	Fold change
212592_at	*IGJ*	immunoglobulin J chain	4q13.3	36.4
208303_s_at	*CRLF2*	cytokine receptor-like factor 2 isoform 1	Xp22.33	8.47
244871_s_at	*USP32*	ubiquitin specific protease 32	17q23.2	2.77
221523_s_at	*RRAGD*	Ras-related GTP binding D	6q15	0.551
208765_s_at	*HNRNPR*		1p36.12	0.716

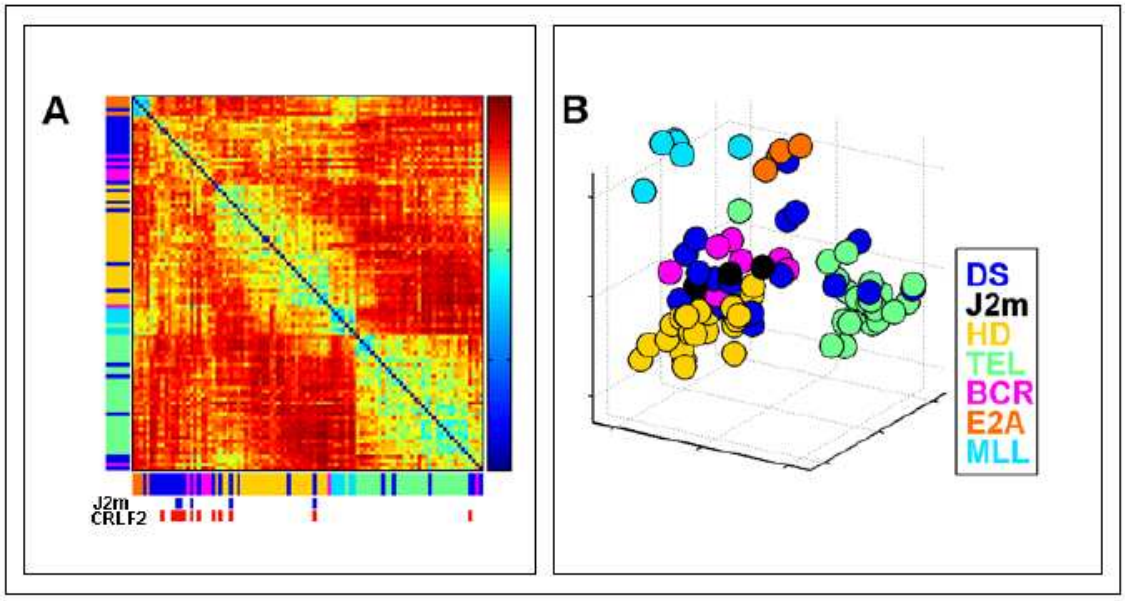
Table 6. Clinical and diagnostic characteristics of patients with DS ALL with high/medium expression of *CRLF2* vs. low/no expression of *CRLF2*

		No/low *CRLF2* expression (n=20)	High/medium *CRLF2* expression (n=33)	p-value
Sex	Male	11	17	1[a]
	Female	9	16	
Age (years)	Mean (SE)	9.87 (1.21)	5.56 (0.67)	0.004[b]
	Median (range)	11.62 (1.74-18.7)	4.06 (1.99-20.22)	
White blood cell count at diagnosis (cell/μl)	Mean (SE)	41613 (8733)	39792 (9738)	0.392 [b]
	Median (range)	23900 (2400-130000)	18930 (1500-259000)	
JAK2 R683 mutations	Yes	0	10	0.008 [a]
	No	20	22	

[a] According to Fisher's exact test
[b] According to Mann-Whitney U test

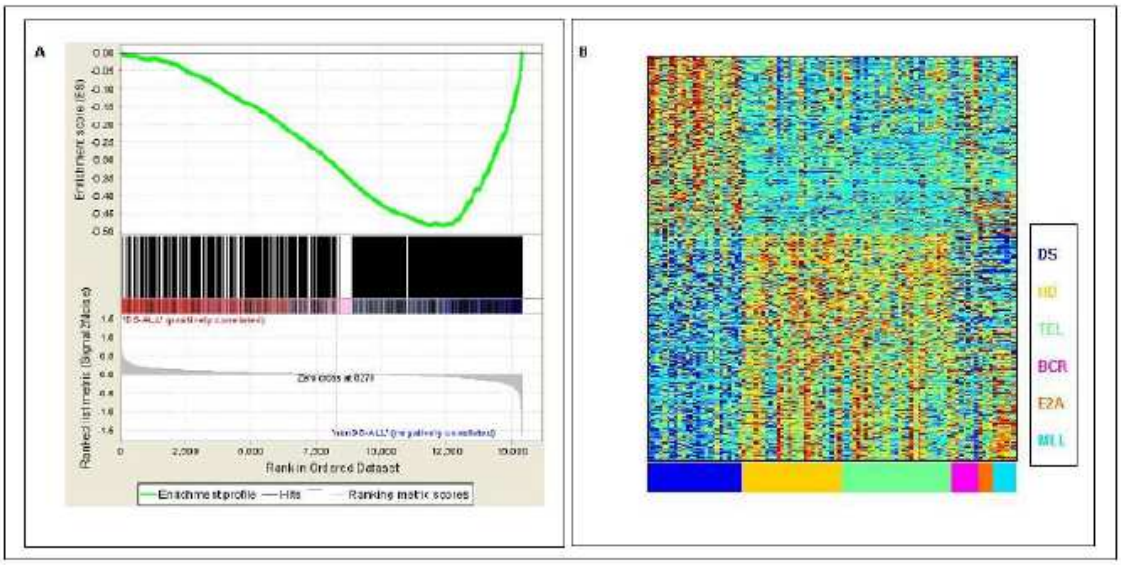
FIGURES

Figure 1. Unsupervised analysis of the AIEOP dataset.



(A) Samples' Euclidian distance matrix. The color in each entry (i,j) represents the Euclidian distance between the expression profiles of samples i and j. It was measured after centering and normalization of each sample's expression, using 1500 probe sets with highest standard deviation. The samples are ordered by SPIN along both the xaxis and y-axis. The color bars next to both axes represent the different ALL subtypes, listed on the right of the Figure. The blue marks at bottom specify DS ALL samples with mutant JAK2 (J2m), and the red marks specify samples with high *CRLF2* expression levels (CRLF2) (see results section on CRLF2).
(B) Projection of all samples onto the first three principle components of the expression. Abbreviations: DS: Down Syndrome ALL; J2m: Down Syndrome ALL with mutated JAK2 R683; HD: High Hyperdiploid; TEL: TEL-AML1; BCR: BCRABL; E2A: E2A-PBX1; MLL: MLL-AF4;

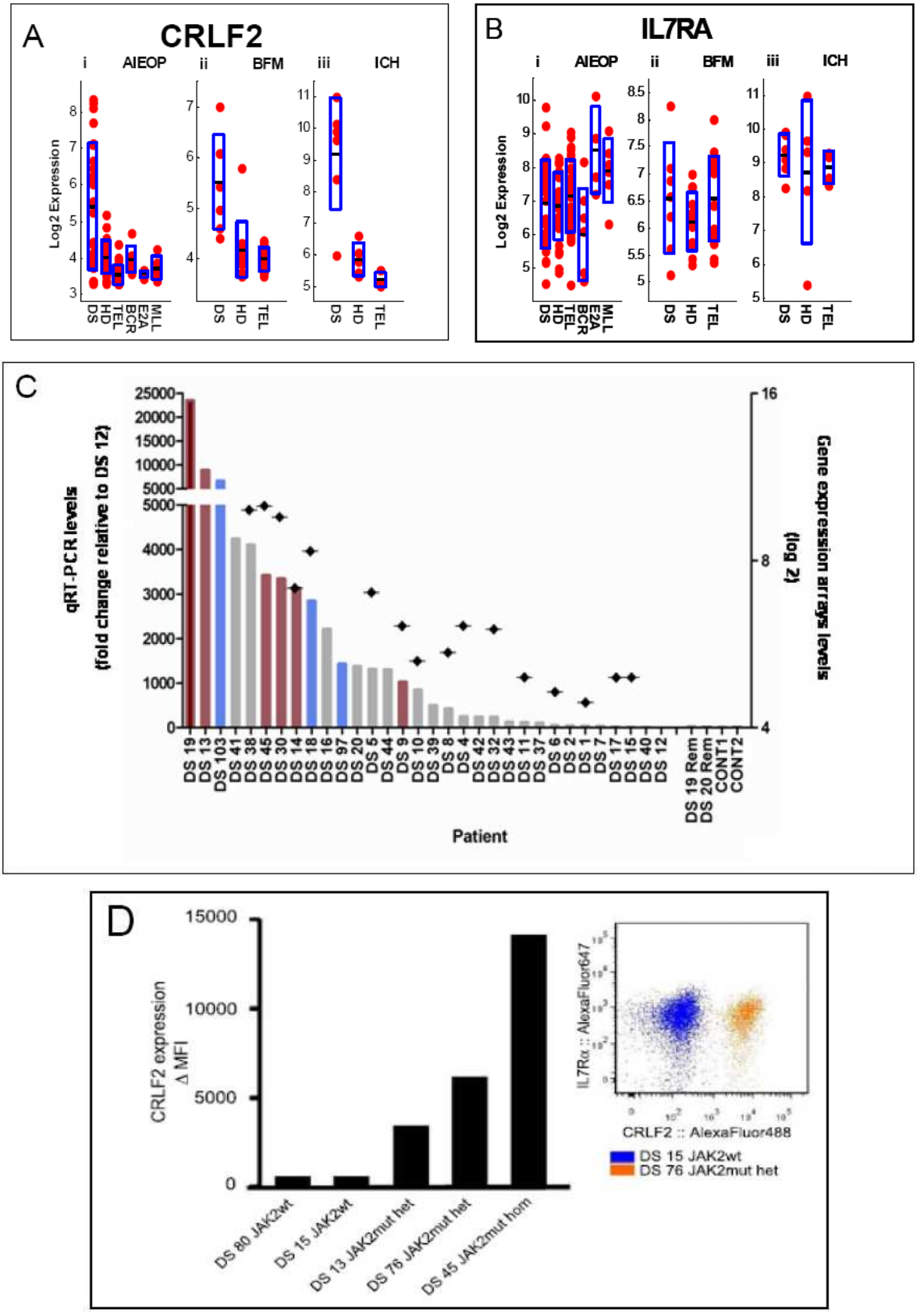
Figure 2. DS ALL gene expression profile.



(A) GSEA analysis on ICH data set. Genes are ranked (bottom of panel, gray) according to their expression in DS ALL samples vs. the rest of the samples, by GSEA, using the default parameters. The members of a gene set S (here - the set of 535 genes down-regulated in DS ALL, AIEOP data) are tested: are they randomly distributed in the ranked gene list, or primarily found at the top or bottom. Occurrences of members of the gene set S in the ranked gene list are shown as vertical black lines above the ranked signature. The green curve and upper y-axis represent the enrichment score (ES) as a function of the number of ranked genes tested for enrichment of gene set S. See Supplementary Methods for full details. 36

(B) Expression levels of the genes from the refined DS ALL lists, measured on the AIEOP dataset. 423 Probe sets that belong to the refined DS ALL profile gene lists are centered and normalized. Values for each individual case are represented by a color, with red representing deviation above the mean and blue representing deviation below the mean. The colors along the x-axis represent the different ALL subtypes, listed on the right of the plot. Abbreviations: DS: Down Syndrome ALL; HD: High Hyperdiploid; TEL: TEL-AML1; BCR: BCR-ABL; E2A: E2A-PBX1; MLL: MLLAF4;

Figure 3. *CRLF2* expression in DS ALL.



(A) *CRLF2* expression in the AIEOP (i), BFM (ii), ICH (iii) datasets. The y-axis represents *CRLF2* Log basis 2 expression. The x-axis represents the different ALL subtypes. Each point corresponds to a sample. The black line

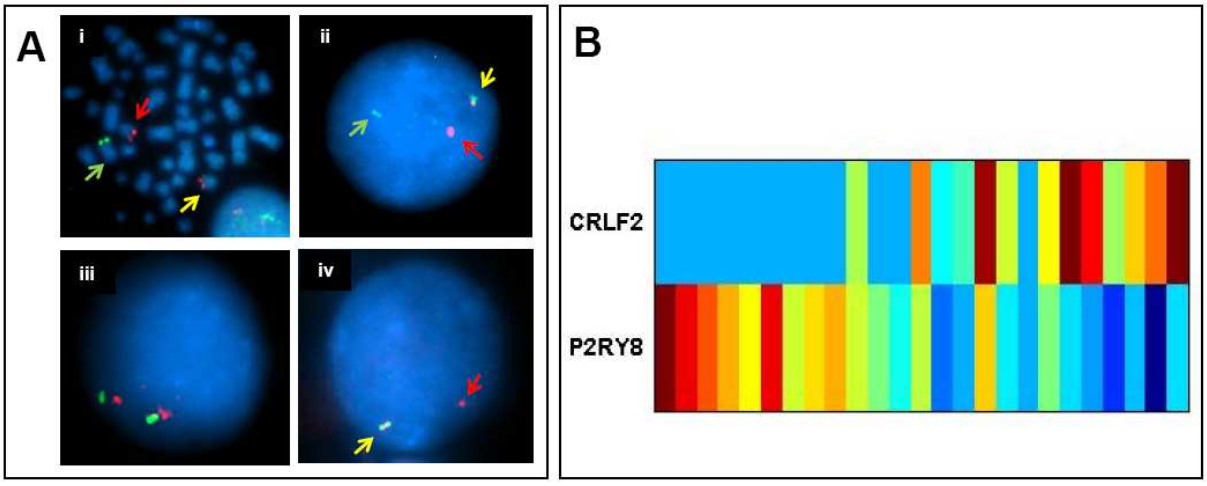
in each ALL subtype is the *CRLF2* mean (Log basis 2) expression in this subtype. The height of the blue rectangle in each ALL subtype is the measured standard deviation of *CRLF2* (Log basis 2) expression. DS ALL versus all other ALL yielded t-test P values of 3.7×10-11, 2.2 ×10-7, 5.6 ×10-5 for AIEOP (i), BFM (ii) and ICH (iii), respectively. Abbreviations: DS: Down Syndrome ALL; HD: High Hyperdiploid; TEL: TELAML1; BCR: BCR-ABL; E2A: E2A-PBX1; MLL: MLL-AF4;

(B) IL7RA expression in the AIEOP (i), BFM (ii), ICH (iii) datasets. There are no statistically significant differences between DS and non DS ALLs.

(C) Verification of *CRLF2* expression levels by qRT-PCR. Bars represent qRT-PCR *CRLF2* expression levels (left Y-axis- fold change relative to patient #DS-12 - the lowest *CRLF2* expresser). Rhombuses represent gene expression arrays *CRLF2* expression levels (right Y-axis- Log basis 2). Red barspatients with JAK2 R683 mutation; Blue bars- patients with CRLF2 F232C mutation see Figure 6; "Rem"- *CRLF2* levels of available remission samples (patients #DS- 19 and #DS-20); "CONT"- Control- *CRLF2* expression levels in peripheral white blood cells of healthy donors.

(D) CRLF2 and IL7RA protein expression on the surface of DS ALL leukemic blasts. Left panel – Delta mean fluorescence intensity of the signal detected by flow cytometry using specific anti-CRLF2 antibodies compared to background unspecific staining (see Methods), indicating an apparent association between the JAK2 mutational status and the level of expression of CRLF2 on DS ALL blasts.; right panel – Dot plot of two representative CRLF2 and IL7RA co-stainings. IL7RA is highly expressed on leukemic blasts independent of JAK2 mutational status and level of CRLF2 expression in all cases examined. wt: wild-type, mut: mutant, het: heterozygous, hom: homozygous.

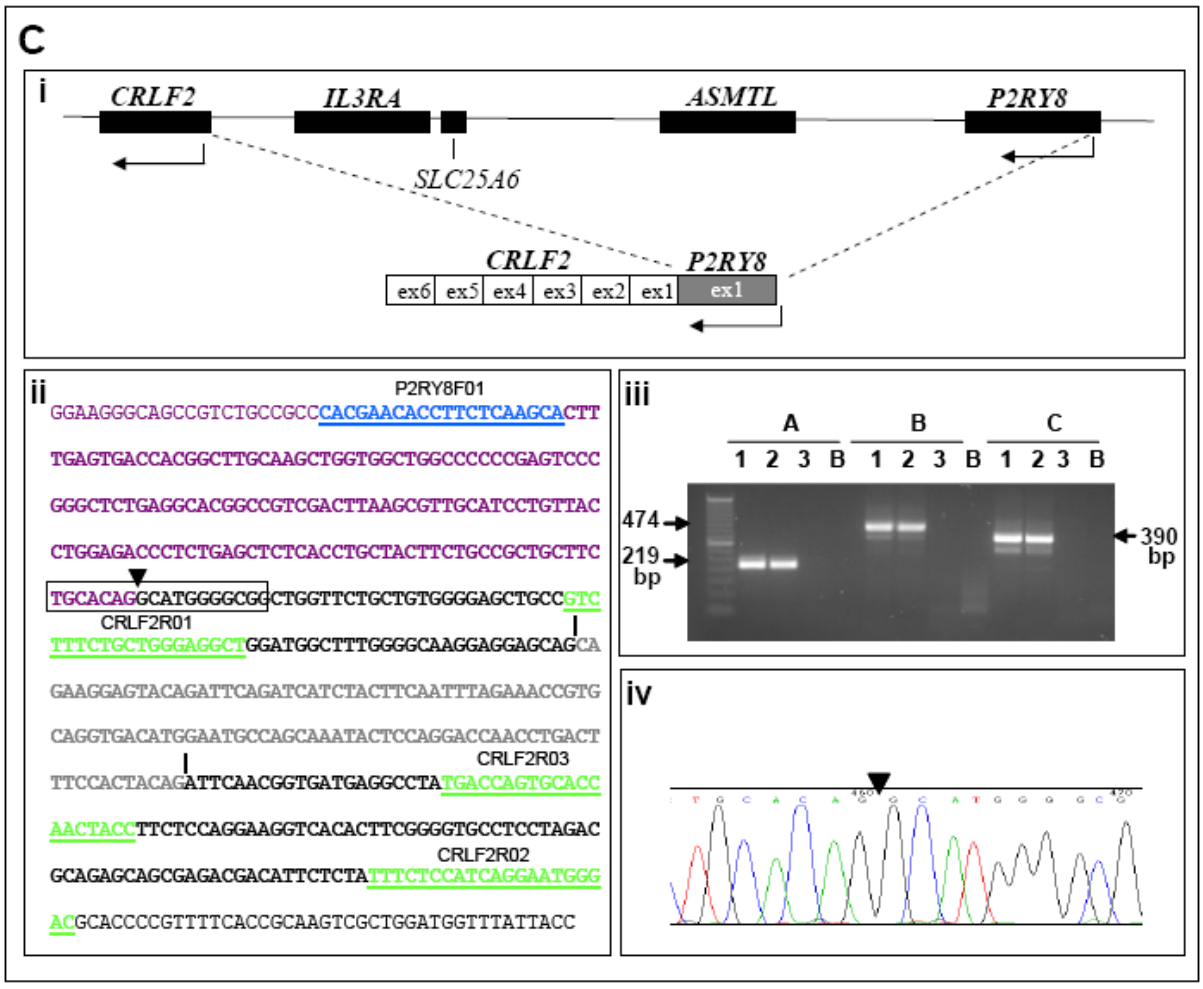
Figure 4. Genomic analysis of CRLF2 aberrations.



(A) FISH analysis of DS ALL expressing *CRLF2* (i-ii) *IGH@-CRLF2* translocation, patient # DS-85: (i) Metaphase showing a positive result with the LSI IGH@ break-apart rearrangement probe (Abbot Molecular): normal chromosome 14 (yellow arrow) derived chromosome 14 (red arrow), derived X chromosome (green arrow). (ii) Interphase nucleus from the same patient hybridized with the homegrown *CRLF2* probe showing a split signal pattern, 1R1G1F confirming its involvement in the translocation [1 fusion signal (yellow arrow), 1

red signal (red arrow) and 1 green signal (green arrow)]. (iii-iv) *CRLF2* microdeletion, patient # DS-82 (iii) Interphase nucleus hybridized with the IGH@ probe showing the normal 0R0G2F signal pattern 38 confirming the presence of two normal copies of IGH@. (iv) Interphase nucleus from the same patient hybridized with the homegrown *CRLF2* probe showing the deletion of the green portion of the probe [1 red signal (red arrow) and 1 fusion signal (yellow arrow)] denoting the presence of a centromeric interstitial deletion.

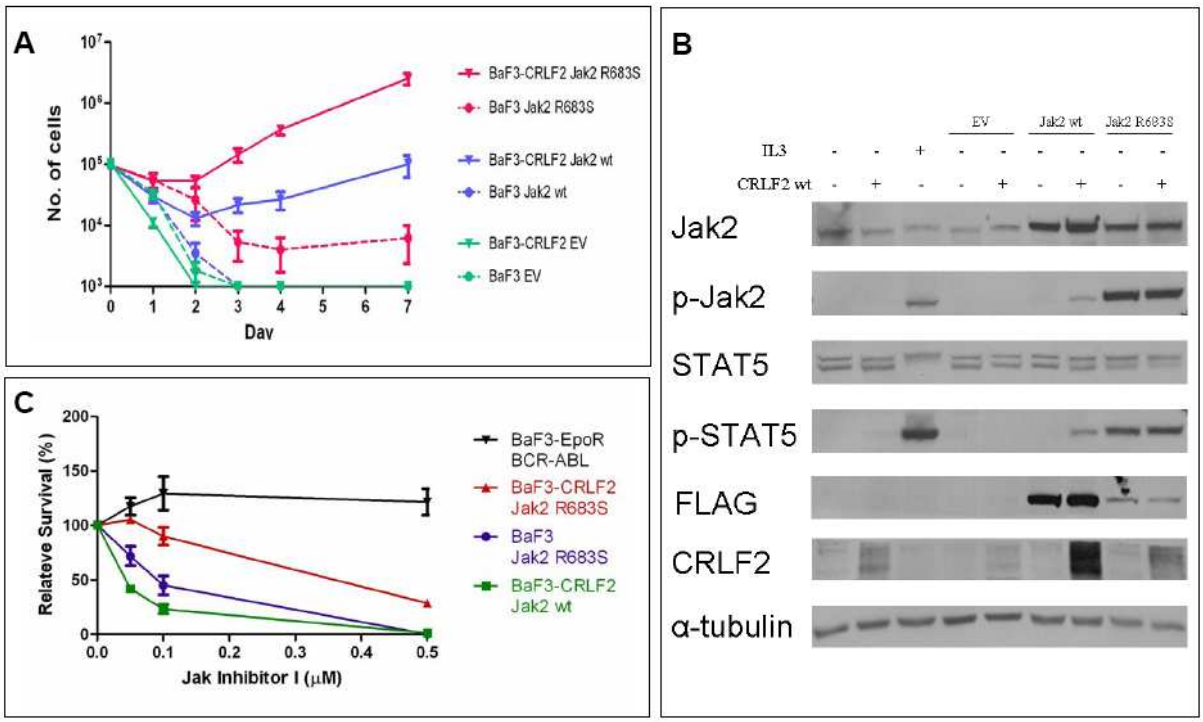
(B) *CRLF2* and *P2RY8* expression in DS ALL samples. Centered and normalized log basis 2 expression of *CRLF2* and *P2RY8* along DS ALL samples in AIEOP data set. Values for each individual case are represented by a color, with red representing deviation above the mean and blue representing deviation below the mean. The samples are sorted using SPIN. Pearson correlation between *CRLF2* and *P2RY8*: -0.45 (P = 0.02),



(C) Detection of the *P2RY8-CRLF2* fusion transcript. (i) schematic representation of the deletion break point region at the telomeric end of chromosome X/Y with gene locations. The dashed lines represent the genomic deletion leading to the fusion of the first non coding exon of *P2RY8* and to the first (coding) exon of *CRLF2*. (iii)RT-PCR experiments on cDNA of DS patients. Lanes 1-3 DS diagnostic ALL samples (#DS93, #DS82 with FISH determined deletion and #DS92 with FISH determined IgH@ translocation, respectively) lane 4: Blank. The three patient samples were positive for ABL amplification (not shown). Primer sets used are: A) *P2RY8* F01/ *CRLF2* R01; B) *P2RY8* F01/ *CRLF2* R02; C) *P2RY8* F01/*CRLF2* R03 shown on the sequence on the left

(ii). Chimeric transcripts are present in the first two lanes of each set. (ii)Nucleotide sequencing of the largest PCR fragment confirming the fusion transcript; vertical lines indicate exon boundries, the arrowhead indicates the *P2RY8-CRLF2* transcript junction. As seen the fusion is just upstream to the ATG of *CRLF2*. Reference sequences are respectively *P2RY8*-001 (ENST00000381297) 39 and *CRLF2*-001 (ENST00000400841). Boxed sequence around the transcript junction is represented in the electropherogram on the right lower side (iv).

Figure 5. Functional significance of CRLF2 expression.

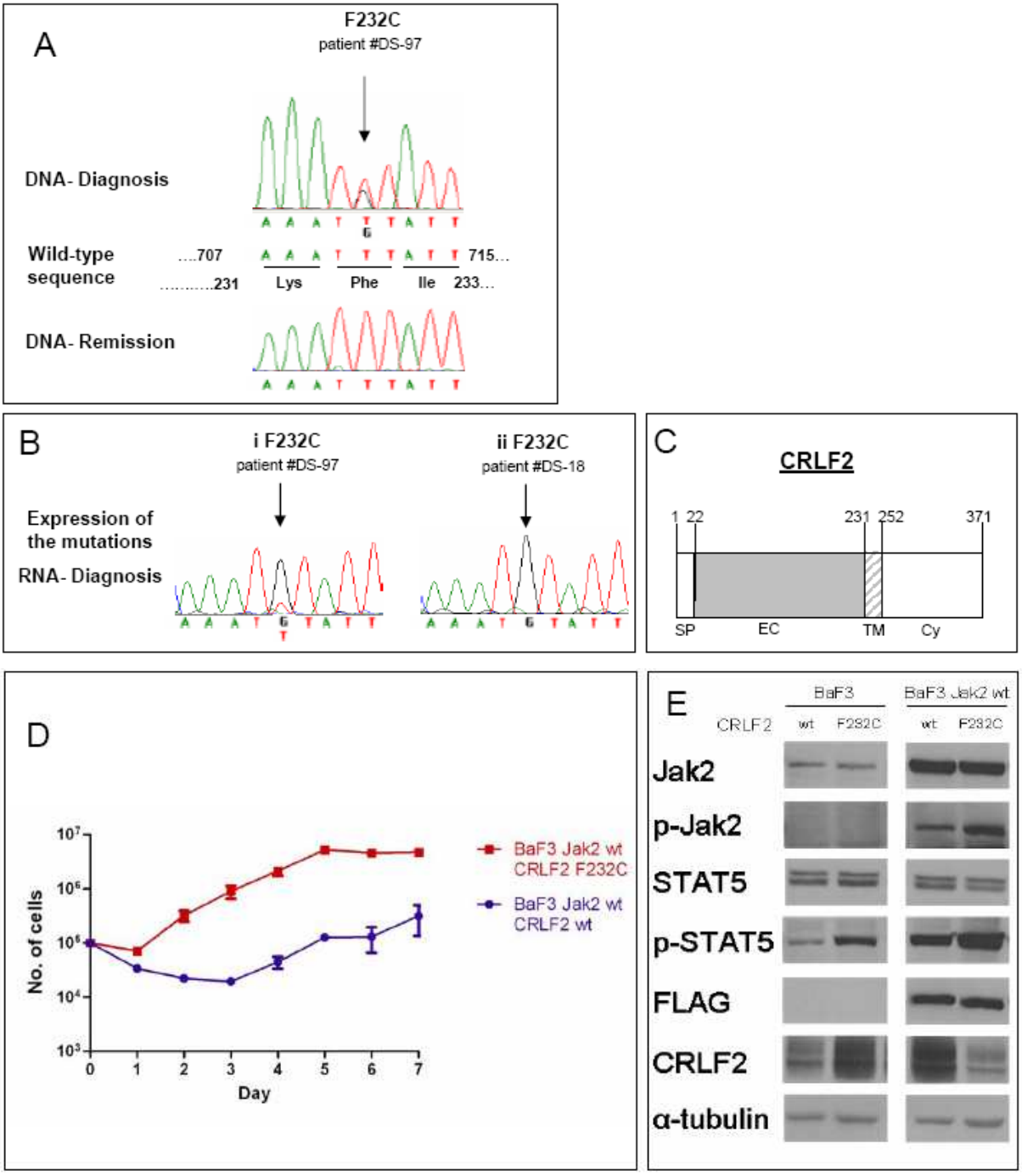


(A) Cytokine withdrawal assay of BaF3 and BaF3-CRLF2 cells infected with either empty vector (EV), mouse FLAG-Jak2 wild-type (wt) or mouse FLAG-Jak2 R683S. Error bars represent SE.

(B) Constitutive activation of the JAK-STAT5 pathway in BaF3 and BaF3-CRLF2 cells expressing mouse FLAG-Jak2 wild-type (wt) or R683S, after 5 hours of cytokines deprivation. "IL3+" are cells harvested after 5 hours of interleukin 3 deprivation followed by 15 minutes of interleukin 3 stimulation.

(C) Effect of JAK inhibitor I on growth of BaF3 cells expressing Jak2 R683S and BaF3-CRLF2 cells expressing either wild-type Jak2 or Jak2 R683S.

Figure 6. Mutations of CRLF2 in patients with Down syndrome-associated acute lymphoblastic leukemia.



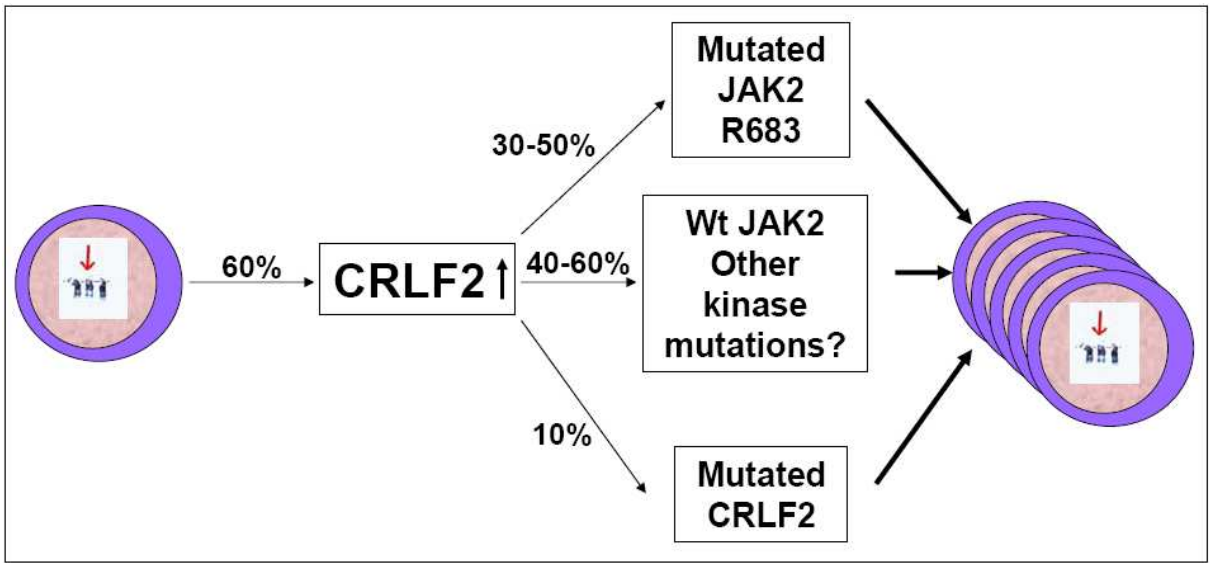
(A) Example of sequences depicting the F232C in CRLF2. The F232C (arrowed) is present at diagnosis but not in remission. The wild-type sequence denotes positions of both nucleotides and amino-acids.

(B) Expression of CRLF2 F232C mutation. Examples of two patients- in one (i) both alleles, wild-type and mutated, are expressed, while in the other (ii) only the mutated allele is expressed. 40

(C) Schematic presentation of CRLF2. “SP”- Signal peptide, “EC”- Extra-cellular region, “TM”- trans-membrane region, “Cy”- cytoplasmic region. Numbers indicate amino-acids position.

(D) Cytokine withdrawal assay of BaF3 cells stably expressing wild-type mouse FLAG-Jak2, that were transduced with either wild-type human CRLF2 or human CRLF2 F232C. Error bars represent SE.
(E) Constitutive activation of the JAK-STAT5 pathway in BaF3 cells expressing wild-type mouse FLAG-Jak2 and either wild-type human CRLF2 (wt) or human CRLF2 F232C (F232C), after 5 hours of cytokines deprivation. IL3 + are cells harvested after 5 hours of interleukin 3 deprivation followed by 15 minutes of interleukin 3 stimulation.

Figure 7. CRLF2 in DS ALL – a model.



Increased CRLF2 expression caused by genomic aberration is followed by progression event consisting of activating mutations in CRLF2, in JAK2 or other alterations in yet unidentified kinases. The percentages in the figure are approximations based on combination of the data in our manuscript and in Russell et al.10

For all supplementary materials see:

http://bloodjournal.hematologylibrary.org/cgi/content/full/blood-2009-08-235408/DC1

REFERENCES

1. Malinge S, Izraeli S, Crispino JD. Insights into the manifestations, outcomes, and mechanisms of leukemogenesis in Down syndrome. Blood. 2009;113:2619-2628.

2. Forestier E, Izraeli S, Beverloo B, et al. Cytogenetic features of acute lymphoblastic and myeloid leukemias in pediatric patients with Down syndrome: an iBFM-SG study. Blood. 2008;111:1575-1583.

3. Rainis L, Bercovich D, Strehl S, et al. Mutations in exon 2 of GATA1 are early events in megakaryocytic malignancies associated with trisomy 21. Blood. 2003;102:981-986.

4. Bercovich D, Ganmore I, Scott LM, et al. Mutations of JAK2 in acute lymphoblastic leukaemias associated with Down's syndrome. Lancet. 2008;372:1484- 1492.

5. Kearney L, Gonzalez De Castro D, Yeung J, et al. Specific JAK2 mutation (JAK2R683) and multiple gene deletions in Down syndrome acute lymphoblastic leukemia. Blood. 2009;113:646-648.

6. Gaikwad A, Rye CL, Devidas M, et al. Prevalence and clinical correlates of JAK2 mutations in Down syndrome acute lymphoblastic leukaemia. Br J Haematol. 2009;144:930-932.

7. Mullighan CG, Zhang J, Harvey RC, et al. JAK mutations in high-risk childhood acute lymphoblastic leukemia. Proc Natl Acad Sci U S A. 2009;106:9414- 9418.

8. Bourquin JP, Subramanian A, Langebrake C, et al. Identification of distinct molecular phenotypes in acute megakaryoblastic leukemia by gene expression profiling. Proc Natl Acad Sci U S A. 2006;103:3339-3344.

9. Bungaro S, Dell'Orto MC, Zangrando A, et al. Integration of genomic and gene expression data of childhood ALL without known aberrations identifies subgroups with specific genetic hallmarks. Genes Chromosomes Cancer. 2009;48:22- 38.

10. Russell LJ, Capasso M, Vater I, et al. Deregulated expression of cytokine receptor gene, CRLF2, is involved in lymphoid transformation in B-cell precursor acute lymphoblastic leukemia. Blood. 2009;114:2688-2698.

11. Hong D, Gupta R, Ancliff P, et al. Initiating and cancer-propagating cells in TEL-AML1-associated childhood leukemia. Science. 2008;319:336-339.

12. Fujio K, Nosaka T, Kojima T, et al. Molecular cloning of a novel type 1 cytokine receptor similar to the common gamma chain. Blood. 2000;95:2204-2210.

13. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102:15545-15550.

14. Tsafrir D, Tsafrir I, Ein-Dor L, Zuk O, Notterman DA, Domany E. Sorting points into neighborhoods (SPIN): data analysis and visualization by ordering distance matrices. Bioinformatics. 2005;21:2301-2308.

15. Ross ME, Zhou X, Song G, et al. Classification of pediatric acute lymphoblastic leukemia by gene expression profiling. Blood. 2003;102:2951-2959.

16. Lo KC, Chalker J, Strehl S, et al. Array comparative genome hybridization analysis of acute lymphoblastic leukaemia and acute megakaryoblastic leukaemia in patients with Down syndrome. Br J Haematol. 2008;142:934-945.

17. Mullighan CG, Goorha S, Radtke I, et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature. 2007;446:758-764.

18. Den Boer ML, van Slegtenhorst M, De Menezes RX, et al. A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genomewide classification study. Lancet Oncol. 2009;10:125-134.

19. Mullighan CG, Miller CB, Radtke I, et al. BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. Nature. 2008;453:110-114.

20. Iacobucci I, Storlazzi CT, Cilloni D, et al. Identification and molecular characterization of recurrent genomic deletions on 7p12 in the IKZF1 gene in a large cohort of BCR-ABL1-positive acute lymphoblastic leukemia patients: on behalf of Gruppo Italiano Malattie Ematologiche dell'Adulto Acute Leukemia Working Party (GIMEMA AL WP). Blood. 2009;114:2159-2167.

21. Subramanian A, Tamayo P, Mootha VK, et al. From the Cover: Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles 10.1073/pnas.0506580102. PNAS. 2005;102:15545-15550.

22. Dennis G, Jr., Sherman BT, Hosack DA, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. 2003;4:P3.

23. Lossos IS, Czerwinski DK, Alizadeh AA, et al. Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. N Engl J Med. 2004;350:1828-1837.

24. Parekh S, Polo JM, Shaknovich R, et al. BCL6 programs lymphoma cells for survival and differentiation through distinct biochemical mechanisms. Blood. 2007;110:2067-2074.

25. Rhodes DR, Yu J, Shanker K, et al. ONCOMINE: a cancer microarray database and integrated data-mining platform. Neoplasia. 2004;6:1-6.

26. Benjamini Y HY. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J Roy Stat Soc. 1995;57(Ser B):289-300.

27. Polo JM, Juszczynski P, Monti S, et al. Transcriptional signature with differential expression of BCL6 target genes accurately identifies BCL6-dependent diffuse large B cell lymphomas. Proc Natl Acad Sci U S A. 2007;104:3207-3212.

28. Basso K, Margolin AA, Stolovitzky G, Klein U, Dalla-Favera R, Califano A. Reverse engineering of regulatory networks in human B cells. Nat Genet. 2005;37:382-390.

29. Hummel M, Bentink S, Berger H, et al. A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. N Engl J Med. 2006;354:2419- 2430.

30. Liu YJ, Soumelis V, Watanabe N, et al. TSLP: an epithelial cell cytokine that regulates T cell differentiation by conditioning dendritic cell maturation. Annu Rev Immunol. 2007;25:193-219.

31. Storlazzi CT, Albano F, Lo Cunsolo C, et al. Upregulation of the SOX5 by promoter swapping with the P2RY8 gene in primary splenic follicular lymphoma. Leukemia. 2007;21:2221-2225.

32. Aplan PD, Lombardi DP, Ginsberg AM, Cossman J, Bertness VL, Kirsch IR. Disruption of the human SCL locus by "illegitimate" V-(D)-J recombinase activity. Science. 1990;250:1426-1429.

33. Rochman Y, Leonard WJ. The role of thymic stromal lymphopoietin in CD8+ T cell homeostasis. J Immunol. 2008;181:7699-7705.

34. Ziegler SF, Liu YJ. Thymic stromal lymphopoietin in normal and pathogenic T cell development and function. Nat Immunol. 2006;7:709-714.

35. Pandey A, Ozaki K, Baumann H, et al. Cloning of a receptor subunit required for signaling by thymic stromal lymphopoietin. Nat Immunol. 2000;1:59-64.

36. Carpino N, Thierfelder WE, Chang MS, et al. Absence of an essential role for thymic stromal lymphopoietin receptor in murine B-cell development. Mol Cell Biol. 2004;24:2584-2592.

37. Ragimbeau J, Dondi E, Alcover A, Eid P, Uze G, Pellegrini S. The tyrosine kinase Tyk2 controls IFNAR1 cell surface expression. Embo J. 2003;22:537-547.

38. Huang LJ, Constantinescu SN, Lodish HF. The N-terminal domain of Janus kinase 2 is required for Golgi processing and cell surface expression of erythropoietin receptor. Mol Cell. 2001;8:1327-1338.

39. Royer Y, Staerk J, Costuleanu M, Courtoy PJ, Constantinescu SN. Janus kinases affect thrombopoietin receptor cell surface localization and stability. J Biol Chem. 2005;280:27251-27261.

40. Yoda A, Bar-Natan M, Sattler M, Ritz J, Frank AD, Weinstock DM. CRLF2 is a proto-oncogene in pre-B acute lymphoblastic leukemia with normal cytogenetics. Proceedings of the 100th Annual Meeting of the American Association for Cancer Research. 2009:Abstract Nr 5610.

41. Lu X, Gross AW, Lodish HF. Active conformation of the erythropoietin receptor: random and cysteine-scanning mutagenesis of the extracellular juxtamembrane and transmembrane domains. J Biol Chem. 2006;281:7002-7011.

42. Ganmore I, Smooha G, Izraeli S. Constitutional aneuploidy and cancer predisposition. Hum Mol Genet. 2009;18:R84-93.

43. Duy C, Yu j, J, Cerchietti L, et al. BCL6-Mediated Survival Signaling Promotes Drug-Resistance in BCR-ABL1-Driven Acute Lymphoblastic Leukemia. Proceedings of the 50th Annual Meeting of the American Society of Hematology. 2008:Abstract Nr 295.

44. Morawiec Z, Janik K, Kowalski M, et al. DNA damage and repair in children with Down's syndrome. Mutat Res. 2008;637:118-123.

45. Lundin C, Heldrup J, Ahlgren T, Olofsson T, Johansson B. B-cell precursor t(8;14)(q11;q32)-positive acute lymphoblastic leukemia in children is strongly associated with Down syndrome or with a concomitant Philadelphia chromosome. Eur J Haematol. 2009;82:46-53.

CHAPTER 3

Combined *P2RY8-CRLF2* Fusion and *CRLF2* Overexpression As a New Marker of Poor Prognosis for Children with intermediate risk BCP ALL

Chiara Palmi, Elena Vendramini, Daniela Silvestri, Giulia Longinotti, AnnaMaria Di Meglio, Daniela Frison, Emanuela Giarin, V Rossi, Anna Lezl, Grazia Fazio, Silvia Bungaro, Shai Israeli, Giuseppe Basso, Andrea Biondi, Maria Grazia Valsecchi, Valentino Conter, Geertruy te Kronnie, Gianni Cazzaniga.

UNDER SUBMISSION

ABSTRACT

Purpose

Treatment of pediatric BCP ALL is increasingly successful, achieving cure rates of over 80%. However, relapse is still the most frequent adverse event, occurring mostly in patients initially stratified as at intermediate risk emphasizing the need for upfront identification of these patients. Recently, overexpression and genetic aberrations related to *CRLF2* have been proposed as new markers for the identification of patients with poor prognosis. We analyzed the incidence and prognostic impact of these potential new markers in BCP ALL Italian patients enrolled into the AIEOP-BFM ALL2000 study.

Patients and Methods

A representative study cohort of 464 non DS and non Ph+ BCP ALL patients from the AIEOP ALL-2000 cohort was analyzed for the expression levels of *CRLF2* and for the occurrence of *CRLF2* rearrangements. Positive patients were analyzed for *JAK2* and *CRLF2* mutations and for recurrent *DNA copy number variations*.

Results

In 22 out of 464 samples (4,7%) *CRLF2* expression levels were 20 times higher than the overall median, and these patients had a cumulative incidence of relapse of 37,1% compared to 15,2% in the rest of the cohort. The *CRLF2* related genetic aberrancies were more frequent in *CRLF2* overexpressed patients (50% vs 3.5%) and the presence of *P2RY8-CRLF2* fusion gene strongly contributes to the adverse outcome (CI 42,8%). Among Intermediate risk patients with high levels of *CRLF2*, the *P2RY8-CRLF2* rearrangements occurred with a

frequency of 53% and this association significantly contributed to the reduced EFS for patients with hi*CRLF2* and *P2RY8-CRLF2* rearrangements (37.5%).

Conclusion

The *P2RY8-CRLF2* rearrangement in association with 20 times over expression of *CRLF2* identifies BCP ALL patients with a very poor prognosis and identifies an important subset of patients currently stratified as intermediate risk patients that need to be considered for treatment adaptation.

INTRODUCTION

Despite cure rates of more than 80%, the prognosis of children with acute lymphoblastic leukaemia (ALL) who relapse after treatment is only around 30%[1]. Therefore, the identification of better upfront prognostic factors remains a formidable challenge in childhood ALL.

Genetic alterations have been identified as key factors in the pathogenesis of ALL[2] and provide essential diagnostic and prognostic hallmarks that have been incorporated in risk stratification of patient for treatment[3]. Moreover, genetic alterations are suggested to offering the potential for the development of targeted therapeutic approaches.

Recently, in a subset of BCP ALL patients without known chromosomal aberrations new genomic abnormalities have been reported that involve the pseudoautosomal region (PAR1) of the sex chromosomes, leading to deregulated *CRLF2* signaling. These abnormalities include small deletions within PAR1 as well as translocations of this region with the IGH@ locus on chromosome 14[4]. Several studies provided evidence for the mechanistic basis of the overexpression of *CRLF2* as a consequence of these chromosomal abnormalities. The PAR1 deletion juxtaposes the first non-coding exon of *P2RY8* to the first exon of *CRLF2*, leading to *CRLF2* expression driven by the promoter of *P2RY8*; whereas translocation of the PAR1 region to the IGH@ locus brings *CRLF2* under the control of *IGH@* enhancer elements[5, 6]. Elevated *CRLF2* expression was further associated with *JAK2* and *CRLF2* point mutations[7, 8]. Interestingly, *CRLF2* overexpression correlated with poor outcome in high-risk BCP ALL patients[9].

Here we present data on the incidence and prognostic impact of *CRLF2* pathway aberrations and *CRLF2* overexpression at diagnosis and relapse in 464 BCP ALL Italian children treated on the AIEOP-BFM ALL2000 protocol.

CRLF2 aberration seems to represent a new prognostic marker in childhood BCP ALL, and inhibition of CRLF2-JAK2 signaling has the potential to become a therapeutic approach for this subgroup of patients.

PATIENTS AND METHODS

Patients

B Cell Precursor Acute Lymphoblastic Leukaemia (BCP ALL) patients enrolled in the AIEOP (Associazione Italiana Ematologia Oncologia Pediatrica Group) ALL 2000 study

from February 2003 to July 2005 were included in the study cohort. Patients with Down Syndrome and patients with t(9;22) were excluded. Diagnosis was made according to standard cytomorphology, cytochemistry and immunophenotypic criteria[10]. Diagnostic bone marrow specimens were available for 464 BCP ALL patients, and relapse specimens were available for 34 patients of the study cohort. DNA and RNA were isolated from mononuclear cells and cDNA was synthesized according to standard method (Biomed). Data on routinely tested recurrent genomic aberrations and DNA index of blast cells was available for most patients. The local ethics committees approved the study and informed consent was obtained for all patients.

The 464 patients cohort has a balanced gender composition, prevalence of patients below 5 years of age and a prevalence of low WBC count (<20,000cell/µl). The majority of patients were good prednisone responders (93.6%), more than half of the patients were intermediate risk by PCR-MRD (60.2%) and included on intermediate risk final protocol strata (64,7%). Clinical characteristics of the study cohort (Table 1) and cumulative incidence of relapse (Supplementary Figure.1) were representative for the entire AIEOP-ALL 2000 study cohort.

Protocol stratification

Definition of patients risk groups was as it follows. The high-risk group included patients with any of the following criteria: t(4;11) or MLL/AF4; prednisone poor response (≥1,000 blasts/µL on day 8 peripheral blood after 7 days of prednisone and one dose of intrathecal methotrexate on day 1); inability to achieve clinical remission after Induction; high burden ($\geq 10^{-3}$) of PCR-MRD at day 78. The standard-risk group included patients who lacked high-risk criteria and tested negative by PCR-MRD performed by using two sensitive markers ($\geq 1\times10^{-4}$) at both day 33 and day 78. The intermediate-risk group included the remaining patients, and those not evaluated by PCR-MRD. PCR-MRD was detected by real-time quantitative polymerase chain reaction (RQ-PCR) of receptor gene rearrangements, in bone marrow samples collected at the end of induction phases TP1 (day 33), and TP2 (day 78); data were interpreted according to published guidelines[11].

Quantitative expression of CRLF2

CRLF2 transcript levels on diagnostic and relapse samples were analyzed using the TaqMan Gene Expression Assay Hs00913509_s1 (Applied Biosystems, Foster City, CA, US); the

housekeeping *GUS* gene transcript was tested as an internal control by using the Universal Probe Library (UPL) system (Roche Diagnostics, Basel, Switzerland), following the manufacturers' instructions. Each cDNA sample (20ng RNA equivalent) was tested in duplicate. Relative gene expression (indicated as *fold change*) was quantified by the 2^{-DDCt} method[12]. The DDCt were calculated subtracting the median of the DCt of all (n=464) tested patients at diagnosis to the DCt of each sample.

CRLF2 rearrangements

The presence of the fusion transcript *P2RY8-CRLF2* as resulting from the PAR1 deletion, was analyzed by RT-PCR using primers designed in the first exon of *P2RY8* (5'-GGACAGATGGAACTGGAAGG-3') and the third exon of *CRLF2* (5'-GTCCCATTCCTGATGGAGAA-3'). PCR product was approximately 511bp.
Fluorescence in situ hybridization (FISH) on interphase nuclei was performed on 24 diagnostic samples, using the Vysis LSI *IGH* Dual Color Break-Apart Rearrangement Probe (Abbott Molecular) and the CEP® Chromosomes Enumeration Probes DNA FISH Probe to detect chromosome X (Abbott Molecular). The nuclei were counterstained with DAPI (4', 6-diamino-2-phenylindol). Results were recorded using a fluorescence Leica DMRB miscroscope fitted with a 100x/1.30 oil objective, CCD camera and digital imaging software from Metasystem (ISIS, FISH imaging system).

Other genetic aberrations

High Resolution Melting (HRM) analysis was performed to identify JAK2 mutations in exon 16 on 160 samples at diagnosis and 60 at relapse using the High Resolution Melting Master and the LightCycler® 480 (Roche Diagnostics), following the manufacturer's instructions. DNA from a pool of peripheral blood buffy coat of healty donors was used as wild type reference, while DNA from the MUTZ5 human cell line (DSMZ), carrying R683G Jak2 mutation was used as a positive control. All the fragments with abnormal melting profile were sequenced. The PCR products were cloned by the Zero Blunt® PCR Cloning Kit (Invitrogen) following the manufacturer's instructions and then sequenced using the *ABI-3130 Genetic Analyzer* instrument (Applied Biosystems).
Patients positive for *CRLF2* overexpression and/or *CRLF2* rearrangements at diagnosis were analysed for *DNA copy number variations* of the gene *IKZF1*, frequently deleted in BCP ALL

patients, by Multiplex Ligation-dependent Probe Amplification (MLPA) technique, using 100ng of DNA and the Salsa MLPA kit P335-A3 ALL-IKZF1 kit (MRC-Holland, Amsterdam, the Netherlands), according to the manufacturer's instructions. Samples of ALL pediatric patients in complete remission were used as wild type controls. The fragments were separated with the *ABI-3130 Genetic Analyzer* instrument (Applied Biosystems) and data analysed using the Coffalyser software (http://old.mlpa.com/coffalyser).

Statistical analysis

To analyze patients' clinical outcome in relation to the various parameters analyzed in this study, a statistical method was used that calculates the development of relapse over a period of time. The Relapse Cumulative Incidence (CRI) curves were estimated by the Kaplan-Meier method; the starting point of the observation time was the date of diagnosis.

RESULTS

CRLF2 expression and genomic aberrations at diagnosis

CRLF2 expression evaluated by RQ-PCR on 464 pediatric B-ALL patients at diagnosis ranged from 0.01 to 800 fold change compared to the overall median value. In particular, 111 cases (24.0%) were 3 times higher than the median, 68 cases (14.6%) 5 times higher, 39 cases (8.5%) 10 times higher and 22 (4.7%) cases 20 times higher than the median value. The comparison of Event Free Survival (EFS) and Cumulative Incidence of Relapse (CIR) of each subgroup (see below and supplementary figure 2) relative to the 'rest' showed a significant difference only for the *CRLF2* ≥20 times expression group which was here after named *"hiCRLF2"*. *HiCRLF2* patients had a mean age of 5.9 years (median 4,9) and mean WBC of 14,144 cell/µl (median 8,245 cell/µl) at time of diagnosis.

The *P2RY8-CRLF2* rearrangement was detected in 22/365 cases overall (6.0%), and in 10 out of 20 cases within the *hiCRLF2* group (50.0%). Patients carrying the fusion gene had a mean age of 5.9 years (median 4.9) and mean WBC of 21,766 cell/µl (median 11,150 cell/µl) at time of diagnosis.

The IGH@ translocations were identified by FISH in two samples out of 24 tested; samples tested had variable CRLF2 expression; within the *hiCRLF2* group one out of eight (negative for *P2RY8-CRLF2)* was *positive* for the translocation (Fig. 1A). Fourteen patients showed additional chromosomes X and/or 14, and three patients had a deletion of the IGH@ locus.

All *hiCRLF2* patients were negative for the recurring chromosomal translocations t(4;11), t(9;22) and t(12;21), while 3/12 *CRLF2*-low expressing, *P2RY8-CRLF2* positive patients (25.0%) carried the t(12;21) translocation. All PAR1 deleted patients had DNA index <1.16, while 5/22 *hiCRLF2* patients were classified as high hyperdiploid.

In 5/157 cases (3.2%) we observed JAK2 mutations (two R683G, one R683S and two insertion mutations). In particular 3/5 mutations were identified in *hiCRLF2* patients *P2RY8-CRLF2* positive, 1/5 in a *hiCRLF2* patient *P2RY8-CRLF2* negative and 1/5 in a *CRLF2*-low expressing patient *P2RY8-CRLF2* positive. The *CRLF2* F232C mutation was detected in two *hiCRLF2* patients, positive for *P2RY8-CRLF2* and for *IKZF1* deletions (Fig 1B and 1C).

No significative differences in the frequency of deletions in *IKZF1* gene was found in patients positive or negative for *CRLF2* overexpression and *P2RY8-CRLF2* rearrangement (Fig 1C).

Prognostic impact of the CRLF2 gene overexpression at diagnosis.

Eight out of 22 *hiCRLF2* patients (36,4%) experienced relapse. Kaplan-Meyer survival analysis at 5 years revealed an inferior EFS (62.9%±10.5 vs. 82.6% ±1.8, P=0.05) and increased CIR (37.1%±10.5 vs. 15.2% ±1.7, P=0.02) for *hiCRLF2* patients (Figure 2 A-B).

Prognostic impact of the P2RY8-CRLF2 fusion gene at diagnosis.

Nine out of 22 patients (40.9%) carrying the fusion gene experienced relapse. Kaplan-Meyer survival analysis at 5 years revealed an inferior EFS (57.2%±10.9 vs. 83.5% ±2, P=0.004) and increased CIR (42.8%±10.9 vs. 14.5% ±1.9, P=0.001) for patients carrying the *P2RY8-CRLF2* fusion gene (Figure 2 C-D).

Outcome and risk group

Considering the unfavourable outcome associated to *CRLF2*-deregulation in our and other studies, we analyzed the prognostic value of *CRLF2* over expression and PAR1 deletion within Intermediate risk (IR) patients, the subgroup of AIEOP ALL2000 patients with higher absolute number of relapses and lack of independent prognostic markers.

Seventeen out 300 (5.6%) patients stratified as IR had *CRLF2* expression 20 times higher than the median (hiCRLF2). Thirteen out of 219 (5.95%) were defined as IR by PCR-MRD stratification. These patients had a considerably higher CIR (42.5%±12.3 vs. 18.3% ±2.3, P=0.04) with respect to IR patients with CRLF2 lower expression (<20 fold) (Figure 3A).

Incidence of PAR1 deletion on IR patients was 15 out 229 (6.5%) (12/219, 5.5% IR by PCR-MRD stratification), and was associated to increased CIR (61.1%±12.9 vs. 17.6% ±2.6 , P=0.0001) (Figure 3B).

Prognostic impact of the combined P2RY8-CRLF2 fusion gene and CRLF2 overexpression at diagnosis.

These finding suggested to further investigate which of the studied features had the strongest impact on outcome on the entire patients cohort. Patients presenting the fusion gene and concomitant over expression of *CRLF2* had a mean age of 6,1 years (median 5,5) and mean WBC of 9,010 cell/µl (median 6,415 cell/µl) at time of diagnosis. Patients carrying the fusion gene without *CRLF2* over expression had a mean age of 5 years (median 3,8) and mean WBC of 32,395 cell/µl (median 28,545 cell/µl) at time of diagnosis.

In Figure 3 we report the CIR of PAR1 deleted patients stratified by *CRLF2* expression (37.5%±16.1 in hi*CRLF2* (≥20 fold change) vs. 75% ±12.5 in *CRLF2*<20 fold change, P=0.24, Figure 3C) and CIR of *CRLF2* over expressed patients stratified by presence of PAR1 deletion (37.5%±16.1 in PAR1 deleted vs. 90% ±9.5 in PAR1 non-deleted, P=0.03, Figure 3D). Taken together, these data indicate that the combined presence of both *CRLF2* over expression and PAR 1 deletion influence patient outcome and should both be taken into account when determining prognosis in BCP ALL patients.

CRLF2 expression and genomic aberrations at relapse

Thirty-four out 62 consecutive patients with medullary relapse were analyzed for *CRLF2* expression and *P2RY8-CRLF2* rearrangement, 25 for *CRLF2* mutations and 41 for *JAK2* mutations. Samples at relapse showed a median value of *CRLF2* expression two times higher than the respective samples at diagnosis (2.82 vs. 1.46). Two out 33 relapse cases (6%) are placed in the *hiCRLF2* group, and one of these was positive for *CRLF2* overexpression already at diagnosis (Fig. 4A).

Interestingly, the relapse sample with the highest *CRLF2* expression level was the only sample positive for both *P2RY8-CRLF2* rearrangement and *JAK2* mutation (L681-I682insLR). Two *P2RY8-CRLF2* positive patients at diagnosis lost the rearrangement at relapse and no *de novo* fusion genes were detected in the other 30 relapse cases. *CRLF2*

mutations remained stable at relapses, while JAK2 mutations were lost in 1/36 patients at relapse (2.8%) and acquired in 2/36 cases (5.6%). (Fig. 4C).

DISCUSSION

The improvement of cure rate among BCP ALL in recent years was been strongly related to the amelioration of risk assessment and establishment of effective prognostic factor. The recent finding of *CRLF2* aberrations suggests an association with unfavourable outcome and an additional value as new prognostic marker. To gain information on incidence, outcome associations and signalling deregulation in relation to *CRLF2* aberrations, more study on other patient cohorts and in distinct therapeutic contests are need.

In this study, *CRLF2* expression has been investigated in a representative cohort of BCP ALL Italian patients. We found a great variability of *CRLF2* expression, ranging from 0.01 to 800 fold change compared to the overall median value. Only the group with higher *CRLF2* expression, named hi*CRLF2* (≥20 fold change, 5% of patients) was associated to an inferior outcome compared to the rest of the cohort, in line with previous finding[13].

The *P2RY8-CRLF2* rearrangement, reported to lead to *CRLF2* overexpression, was detected in 6% of patients (concordant with data reported on German patients treated on the same BFM protocol[14], or in a english cohort treated on MRC ALL97 protocol [15]), and this group had an inferior outcome compared to the non rearranged patients. Interestingly more than 50% of *P2RY8-CRLF2* were not associated to *CRLF2* overexpression.

Comparing the prognostic impact of *CRLF2* overexpression and PAR 1deletion separately or in combination, we assessed that combined presence of both *CRLF2* over expression and PAR 1 deletion indentified patient with a very inferior outcome.

Moreover, the majority of patients belonging to the hi*CRLF2* group or PAR1 deleted group were stratified as IR and were associated to an inferior outcome compared to the rest of IR stratified patients. This finding support the utility of *CRLF2* aberrations as prognostic marker to improve risk assessment among IR stratified patients.

Further investigation are required to enlighten the mechanisms leading to *CRLF2* overexpression in patients without PAR 1 deletions (translocation involving *IGH@* locus seems to explain only the minority of the cases). Also the absence of *CRLF2* overespression in PAR1 deleted patients in 50% of patients and the particular contribution to outcome requires further investigations.

Concluding, the *CRLF2-P2RY8* rearrangement in association with 20 times over expression of *CRLF2* identifies BCP ALL patients with a very poor prognosis and identifies an important subset of patients currently stratified as intermediate risk patients that need to be considered for treatment adaptation.

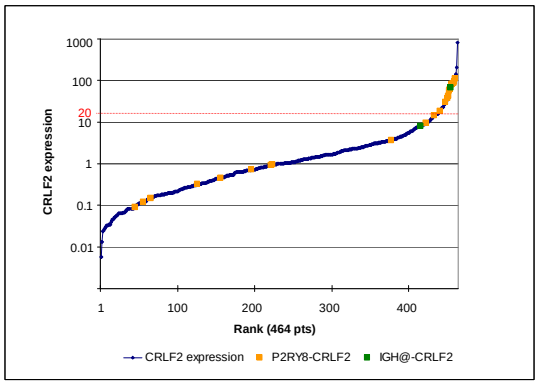
TABLE

FEATURES	Analyzed		Not Analyzed		P-value
	N	%	N	%	
All patients	464	100	1245	100	
Age					
1-5 Yrs	300	64,66	793	63,69	0,31
6-9 Yrs	95	20,47	231	18.55	
≥10 Yrs	69	14,87	221	17,75	
Gender					
Male	239	51,51	661	53,09	0,56
Female	225	48,49	584	46,91	
WBC (x1000/µL)					
<20	327	70,47	865	69,48	0,23
20-100	105	22,63	316	25,38	
≥100	32	6,9	64	5,14	
Traslocations					
MLL					
pos	7	1,51	122	9,9	0,86
neg	464	98,49	1108	90.1	
Unknown	0		15	-	
TEL/AML1					
pos	92	20,81	242	22,34	0,51
neg	350	79,19	841	77,66	
Unknown	22	-	162	-	
DNA Index					
<1.16	336	77,60	862	76,35	0,60
≥1.16	97	22,40	267	23,65	
Unknown	31	-	116	-	
MRD D78					
HR	7	1,93	77	8.09	
MR	218	60,22	508	53,36	0,0001
SR	137	37,85	367	38,55	
Unknown	102	-	293	-	
Prednison Response					
Good	434	93,94	1143	91,95	0,17
Poor	28	6,06	100	8,05	
Unknown	2	-	2	-	
Final Protocol Strata					
HR	35	7,54	170	13,66	0,0015
MR	300	64,66	722	57,99	
SR	129	27,8	353	28,35	

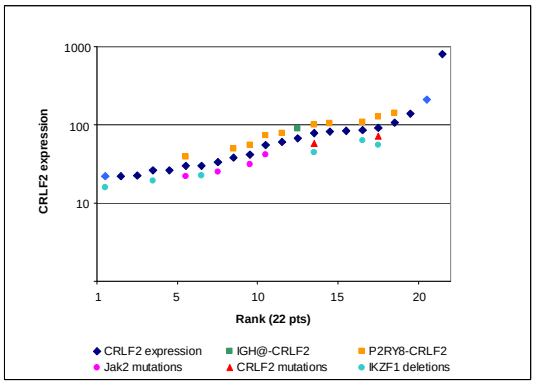
Table.1 Clinical and biological features of patients enrolled in AIEOP ALL 2000 study from February 2003 to July 2005 grouped by analyzed and non analyzed in the present study. Patients with Down Syndrome and patient with t(9;22) were excluded.

FIGURES

A



B



C

DIAGNOSIS	Total	P2RY8-CRLF2	IgH@-CRLF2	CRLF2 mutations	JAK2 mutations	IKZF1 deletions
CRLF2 expression (f.c.) ≥ 20	22/464 *(4.7%)*	10/20 *(50%)*	1/9 *(11.1%)*	2/22 *(9.1%)*	4/22 (18.2%)	5/22 (22.7%)
CRLF2 expression (f.c.) < 20	449/464 *(95.3%)*	12/345 *(3.5%)*	1/17 *(5.9%)*	0/27 *(0%)*	1/135 *(0.7%)*	1/12 (8.3%)
P2RY8-CRLF2 positive	22/365 *(6.0%)*	-	0/1 *(0%)*	2/14 (14.3%)	4/21 *(19.05%)*	4/21 (19.0%)
P2RY8-CRLF2 negative	343/365 *(94.0%)*	-	2/24 *(8.3%)*	0/31 *(0%)*	1/121 (0.8%)	1/12 (8.3%)
CRLF2 expression (f.c.) ≥ 20 & P2RY8-CRLF2 positive	10/365 *(2.7%)*	-	0/1 *(0%)*	2/10 *(20%)*	3/10 *(30%)*	3/10 *(30%)*

Figure 1. *CRLF2* expression reported as fold change on median expression value of the cohort. Expression ranged from 0.01 to 800 fold change (A). Association between *CRLF2* expression ≥20 fold change and investigated genomic aberrations. In bright blue if no tested for *P2RY8-CRLF2* (B). Table of genomic aberrations investigated in the study cohort and results stratified by *CRLF2* expression and PAR1 deletion (C).

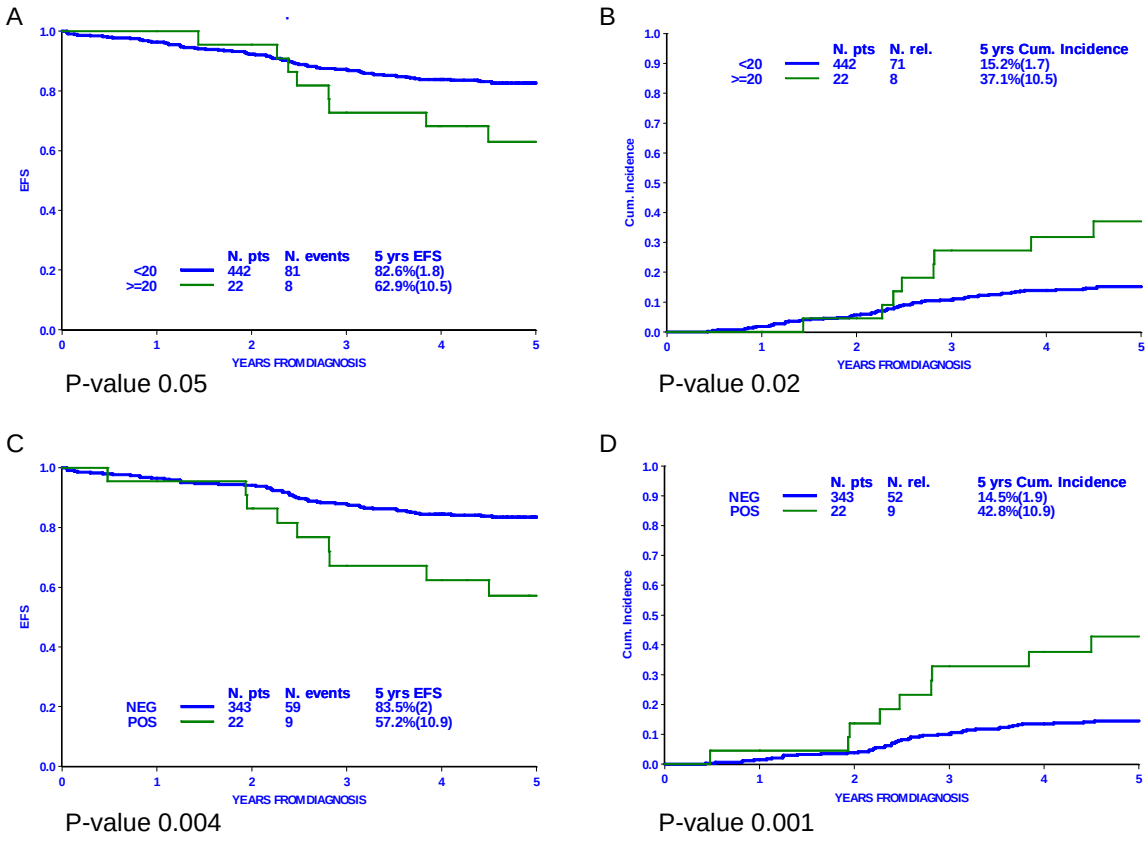


Figure 2. Event free survival at 5 years from diagnosis of study cohort grouped by *CRLF2* expression range: <20 fold change and ≥20 fold change (A). Cumulative incidence of relapse at 5 years from diagnosis of study cohort grouped by *CRLF2* expression range: <20 fold change and ≥20 fold change (B). Event free survival at 5 years from diagnosis of PAR1 deletion investigate patients grouped by presence or absence of fusion gene *P2RY8-CRLF2* (C). Cumulative incidence of relapse at 5 years from diagnosis of PAR1 deletion investigate patients grouped by presence or absence of fusion gene *P2RY8-CRLF2* (D).

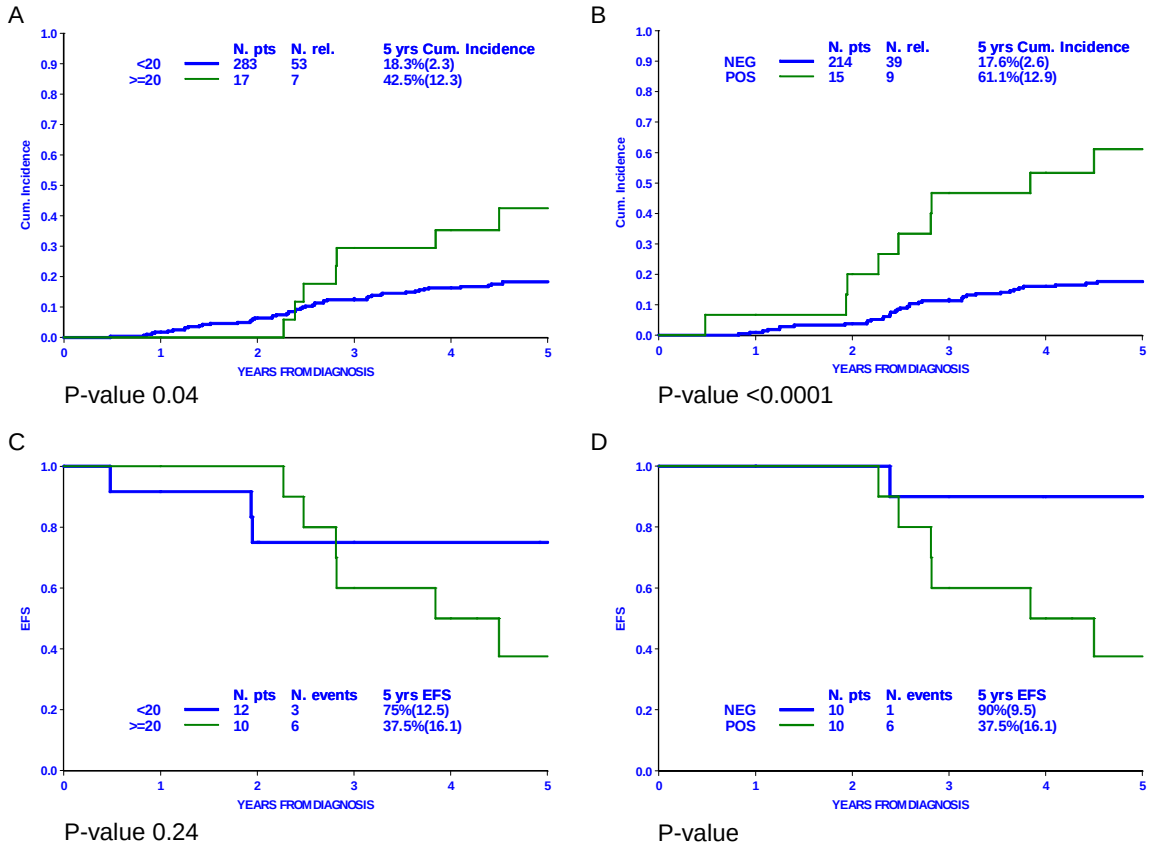
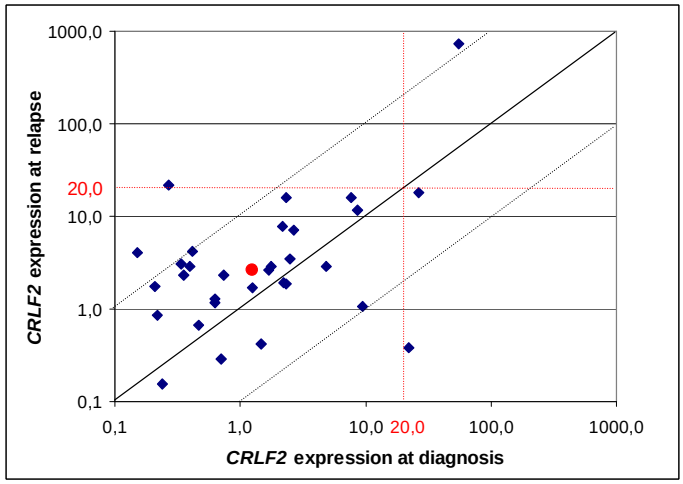


Figure 3. Cumulative incidence of relapse at 5 years from diagnosis of intermediated risk patients grouped by *CRLF2* expression range: <20 fold change and ≥20 fold change (A). Cumulative incidence of relapse at 5 years from diagnosis of PAR1 deletion investigate intermediated risk patients grouped by presence or absence of fusion gene *P2RY8-CRLF2* (B). Cumulative incidence of relapse at 5 years from diagnosis of PAR1 deleted patients grouped by *CRLF2* expression range: <20 fold change and ≥20 fold change (C). Cumulative incidence of relapse at 5 years from diagnosis of PAR1 deletion investigate patients with *CRLF2* expression ≥20 fold change grouped by presence or absence of fusion gene *P2RY8-CRLF2* (D).

A



B

DIAGNOSES vs RELAPSES

DX \ REL	NEG	POS
NEG	21 *(87.5%)*	-
POS	2 *(8.3%)*	1 *(4.2%)*

P2RY8-CRLF2 (N=24)

DX \ REL	NEG	POS
NEG	4 *(66.7%)*	-
POS	-	2 *(33.3%)*

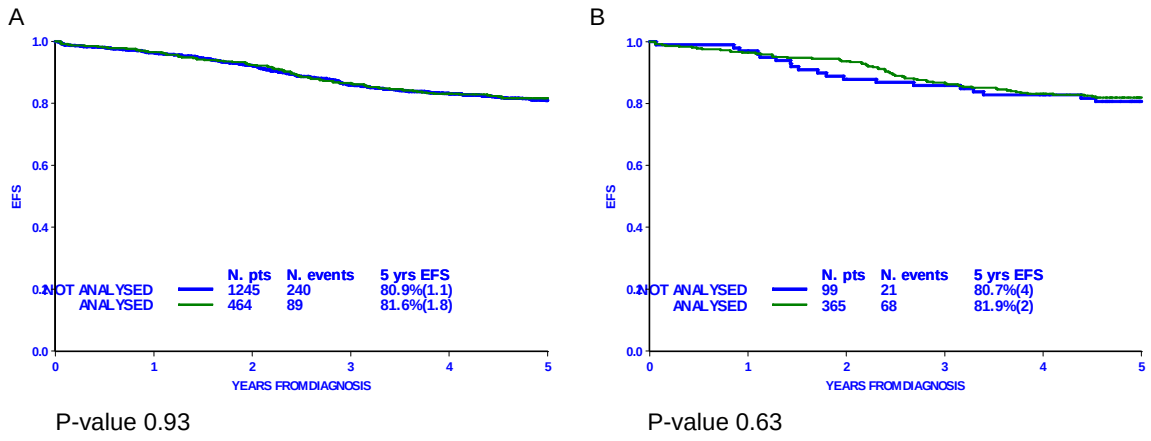
CRLF2 MUTATIONS (N=6)

DX \ REL	NEG	POS
NEG	30 *(88.2%)*	2 *(5.9%)*
POS	1 *(2.9%)*	1 *(2.9%)*

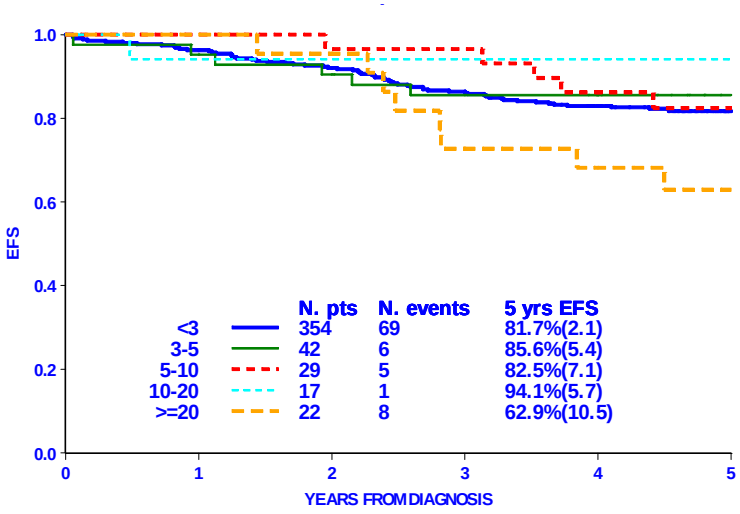
JAK2 MUTATIONS (N=34)

Figure 4. Log-log plot of *CRLF2* expression value for 34 paired diagnosis and relapsed specimens. *CRLF2* expression at relapse were double on median than paired diagnostic samples (median fold change = 2.70 vs 1.35) (A). Paired diagnosis-relapse data of genomic aberrations investigated (B).

SUPPLEMENTARY FIGURES



SF 1. Event free survival at 5 years from diagnosis among AIEOP-ALL 2000 patients (from February 2003 to July 2005, Down Syndrome and Ph+ patients excluded) grouped by included and non-included in the study cohort (A). Event free survival at 5 years from diagnosis among study cohort grouped by analysed and non-analysed for PAR1 deletion (B).



SF 2. Event free survival at 5 years from diagnosis of study cohort grouped by *CRLF2* expression range: <3 fold change, 3-5 fold change, 5-10 fold change, 10-20 fold change and ≥20 fold change .

REFERENCES

[1] Pui CH, Relling MV, Downing JR. Acute lymphoblastic leukemia. N Engl J Med. 2004;350:1535–48.

[2] Greaves, M. F. & Wiemels, J. Origins of chromosome translocations in childhood leukaemia. Nature Rev. Cancer 3, 639–649 (2003).

[3] Moorman, A.V., Ensor, H.M., Richards, S.M., Chilton, L., Schwab, C., Kinsey, S.E., Vora, A., Mitchell, C.D. & Harrison, C.J. (2010) Prognostic effect of chromosomal abnormalities in childhood B-cell precursor acute lymphoblastic leukaemia: results from the UK Medical Research Council ALL97/99 randomised trial. The Lancet Oncology, 11, 429–438.

[4] Blood. 2009 Sep 24;114(13):2688-98. Epub 2009 Jul 29. Deregulated expression of cytokine receptor gene, CRLF2, is involved in lymphoid transformation in B-cell precursor acute lymphoblastic leukemia. Russell LJ, Capasso M, Vater I, Akasaka T, Bernard OA, Calasanz MJ, Chandrasekaran T, Chapiro E, Gesk S, Griffiths M, Guttery DS, Haferlach C, Harder L, Heidenreich O, Irving J, Kearney L, Nguyen-Khac F, Machado L, Minto L, Majid A, Moorman AV, Morrison H, Rand V, Strefford JC, Schwab C, Tönnies H, Dyer MJ, Siebert R, Harrison CJ.

[5] Blood. 2009 Sep 24;114(13):2688-98. Epub 2009 Jul 29. Deregulated expression of cytokine receptor gene, CRLF2, is involved in lymphoid transformation in B-cell precursor acute lymphoblastic leukemia. Russell LJ, Capasso M, Vater I, Akasaka T, Bernard OA, Calasanz MJ, Chandrasekaran T, Chapiro E, Gesk S, Griffiths M, Guttery DS, Haferlach C, Harder L, Heidenreich O, Irving J, Kearney L, Nguyen-Khac F, Machado L, Minto L, Majid A, Moorman AV, Morrison H, Rand V, Strefford JC, Schwab C, Tönnies H, Dyer MJ, Siebert R, Harrison CJ.

[6] Mullighan, C.G., Collins-Underwood, J.R., Phillips, L.A., Loudin, M.G., Liu, W., Zhang, J., Ma, J., Coustan-Smith, E., Harvey, R.C.,Willman, C.L., Mikhail, F.M., Meyer, J., Carroll, A.J., Williams, R.T.,Cheng, J., Heerema, N.A., Basso, G., Pession, A., Pui, C.H., Raimondi,S.C., Hunger, S.P., Downing, J.R., Carroll, W.L. & Rabin,K.R. (2009) Rearrangement of CRLF2 in B-progenitor- and Downsyndrome-associated acute lymphoblastic leukemia. Nature Genetics,41, 1243–1246.

[7] Blood. 2009 Sep 24;114(13):2688-98. Epub 2009 Jul 29. Deregulated expression of cytokine receptor gene, CRLF2, is involved in lymphoid transformation in B-cell precursor acute lymphoblastic leukemia. Russell LJ, Capasso M, Vater I, Akasaka T, Bernard OA, Calasanz MJ, Chandrasekaran T, Chapiro E, Gesk S, Griffiths M, Guttery DS, Haferlach C, Harder L, Heidenreich O, Irving J, Kearney L, Nguyen-Khac F, Machado L, Minto L, Majid A, Moorman AV, Morrison H, Rand V, Strefford JC, Schwab C, Tönnies H, Dyer MJ, Siebert R, Harrison CJ.

[8] Hertzberg, L., Vendramini, E., Ganmore, I., Cazzaniga, G., Schmitz, M., Chalker, J., Shiloh, R., Iacobucci, I., Shochat, C., Zeligson, S., Cario, G., Stanulla, M., Strehl, S., Russell, L.J., Harrison, C.J., Bornhauser, B., Yoda, A., Rechavi, G., Bercovich, D., Borkhardt, A., Kempski, H., Kronnie, G.T., Bourquin, J.P., Domany, E. & Izraeli, S. (2010) Down syndrome acute lymphoblastic leukemia: a highly heterogeneous disease in which aberrant expression of CRLF2 is associated with mutated JAK2: a report from the iBFM Study Group. Blood, 115, 1006–1017.

[9] Rearrangement of CRLF2 is associated with mutation of JAK kinases, alteration of IKZF1, Hispanic/Latino ethnicity, and a poor outcome in pediatric B-progenitor acute lymphoblastic leukemia. Harvey RC, Mullighan CG, Chen IM, Wharton W, Mikhail FM, Carroll AJ, Kang H, Liu W, Dobbin KK, Smith MA, Carroll WL, Devidas M, Bowman WP, Camitta BM, Reaman GH, Hunger SP, Downing JR, Willman CL. Blood. 2010 Jul 1;115(26):5312-21. Epub 2010 Feb 4.

[10] Basso G, Buldini B, De Zen L, Orfao A (2001) New methodologic approaches for immunophenotyping acute leukemias. Heamatologica 86: 675–692.

[11] Molecular response to treatment redefines all prognostic factors in children and adolescents with B-cell precursor acute lymphoblastic leukemia: results in 3184 patients of the AIEOP-BFMALL 2000 study. Valentino Conter, Claus R. Bartram, Maria Grazia Valsecchi, Andre´ Schrauder, Renate Panzer-Gru¨mayer, Anja Mo¨

ricke, Maurizio Arico` , Martin Zimmermann, Georg Mann, Giulio De Rossi, Martin Stanulla,5Franco Locatelli, Giuseppe Basso, Felix Niggli, Elena Barisone, Gu¨ nter Henze, Wolf-Dieter Ludwig, Oskar A. Haas, Giovanni Cazzaniga, Rolf Koehler, Daniela Silvestri, Jutta Bradtke, Rosanna Parasole, Rita Beier, Jacques J. M. van Dongen,1 Andrea Biondi, and Martin Schrappe. Blood. 2010 115: 3206-3214

[12] Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2-DDCT method. Methods 25:402–408

[13] Mullighan, C.G., Collins-Underwood, J.R., Phillips, L.A., Loudin, M.G., Liu, W., Zhang, J., Ma, J., Coustan-Smith, E., Harvey, R.C., Willman, C.L., Mikhail, F.M., Meyer, J., Carroll, A.J., Williams, R.T., Cheng, J., Heerema, N.A., Basso, G., Pession, A., Pui, C.H., Raimondi, S.C., Hunger, S.P., Downing, J.R., Carroll, W.L. & Rabin, K.R. (2009) Rearrangement of CRLF2 in B-progenitor- and Down syndrome-associated acute lymphoblastic leukemia. Nature Genetics, 41, 1243–1246.

[14] Presence of the *P2RY8-CRLF2* rearrangement is associated with a poor prognosis in non-high-risk precursor B-cell acute lymphoblastic leukemia in children treated according to the ALL-BFM 2000 protocol. Cario G, Zimmermann M, Romey R, Gesk S, Vater I, Harbott J, Schrauder A, Moericke A, Izraeli S, Akasaka T, Dyer MJ, Siebert R, Schrappe M, Stanulla M. Blood. 2010 Jul 1;115(26):5393-7. Epub 2010 Apr 8.

[15] Blood. 2010 Nov 29. [Epub ahead of print] Demographic, clinical and outcome features of children with acute lymphoblastic leukemia and CRLF2 deregulation: results from the MRC ALL97 clinical trial. Ensor HM, Schwab C, Russell LJ, Richards SM, Morrison H, Masic D, Jones L, Kinsey SE, Vora AJ, Mitchell CD, Harrison CJ, Moorman AV.

CHAPTER 4

High fidelity of gene expression profiles in serial pediatric NOD/SCID/huALL

Manon Queudeville, Elena Vendramini, Marco Giordan, Sarah M. Eckhoff, Giuseppe Basso, Klaus-Michael Debatin, Gertruuy te Kronnie, and Lüder H. Meyer.

UNDER SUBMISSION (Brief Report)

ABSTRACT

Acute leukemia is supposed to arise from particular subclones within the bulk of leukemia cells. We analyzed the gene expression pattern of primary pediatric acute lymphoblastic leukemia (ALL) samples at diagnosis and compared them to their respective xenograft leukemias after serial transplantation into NOD/SCID mice. After first passage in the NOD/SCID mouse, immunophenotype and gene expression profiles were highly similar to the diagnostic leukemia, with minor differences arising from the presence of normal human hematopoietic cells in primary patient material. Moreover, both expression profile and phenotype remained stable after secondary and tertiary passages. We conclude that the NOD/SCID/huALL model recapitulates the primary human leukemia in the mouse and does not select for a subclone of leukemia initiating cells leading to a bulk tumor with a gene expression distinct from the profile at diagnosis.

INTRODUCTION

Primary childhood ALL samples are difficult to culture *in vitro* and currently available cell lines poorly reflect the heterogeneous nature of the disease. Mouse xenotransplantation models are therefore widely used for *in vivo* testing and to amplify the number of leukemia cells to be used for various analyses.[1] To date, it remains unclear as to what extent the xenografted samples recapitulate their respective primary leukemia. Although a consistent immunophenotype of leukemia cells grown in immunocompromised mice has been shown in the past[2-4], this does not exclude the possibility that the murine microenvironment selects for leukemia initiating cells with a specific gene expression profile which consequently could lead to a bulk tumor markedly different from diagnosis.

STUDY DESIGN

Patients leukemia cell samples were obtained after informed consent was given. The study was approved by the institutional ethical review board and carried out in accordance of the Declaration of Helsinki. Animal studies were conducted according to national animal welfare law and approved by the appropriate authority. Diagnostic leukemia samples (n= 7) were obtained from pediatric ALL patients and were transplanted intravenously into unconditioned NOD/SCID mice (Charles River Laboratories). At manifestation of disease, cell suspensions from spleen and bone marrow were isolated and further analyzed. Leukemia cells were stained with monoclonal antibodies and appropriate isotype controls and analyzed on an LSRII (BD Biosciences). RNA was isolated from leukemia samples using a standardized procedure (RNeasy kit, Qiagen). Gene expression analysis was performed using the Affymetrix HG-U133 Plus 2.0 oligonucleotide microarrays and Affymetrix protocols.
The quality parameters included percentage of present called probe sets (mean 25,909, 47.4%, std 1,045) and the ratio of intensities of 3' probes to 5' probes for the housekeeping gene *GAPDH* (mean 1.01, std 0.003). Statistical analyses were performed in R (http://www.R-project.org).

RESULTS AND DISCUSSION

Patient characteristics are summarized in table 1A. All leukemia samples transplanted resulted in overt leukemia of the recipient NOD/SCID-mice, with high percentages of human leukemia cells both in spleen and bone marrow (BM) (table 1B). However, mice transplanted with

samples #1, #3 and #7 developed high percentages of leukemia cells in peripheral blood (PB), while mice transplanted with #2, #4, #5 and #6 presented leukemia-related symptoms with less than 15% human cells in PB. These clinical features were stable over secondary and tertiary transplantation. Interestingly, the primary patients of #1, #3 and #7 also displayed higher total blast counts in PB at diagnosis than the other patients. Additionally, the immunophenotypes from diagnosis remained stable after serial passages (Figure 1A). Gene expression profiles (GEPs) of diagnostic material and xenograft samples were analyzed after primary, secondary and tertiary transplantation. In an unsupervised clustering analysis of all samples, the matching xenograft samples all clustered tightly together, while the diagnostic samples of all the B-cell precursor ALL samples clustered closer to each other than to their derived xenograft samples (Figure 1B). To clarify differences between diagnosis and xenograft passage we performed a class comparison analysis of the 7 diagnostic versus the 7 primary xenograft samples, resulting in 231 genes upregulated at diagnosis and 23 genes upregulated in the xenografts. The high number of genes upregulated at diagnosis is due to presence of normal PB or BM cells in primary patient samples as (i) 64% of the genes can be attributed to blood cells of various lineages, and (ii) immunophenotyping shows disappearance of CD45-positive non-leukemic cells after the first xenograft passage. Consistently, CD14- and glycophorin A- expression on the leukemia cells (flowcytometry: n=7, mean 2.5% and 8.3%) was also detected by upregulation of the corresponding genes *CD14* and *GYPA,* indicating presence of erythroid and myeloid cells in the diagnostic samples while absent in xenografts. However, when comparing primary to secondary or tertiary passages, no significant differences in GEP were found.

When performing unsupervised clustering of diagnosis and first xenograft passages only, we found that all diagnostic samples cluster together with their respective xenograft samples (Figure 1C). Clonal identity was furthermore confirmed by consistent classification according to MILE[5], SNP profiles or gene fusions in selected samples (supp. Table S1).

Systematic differences in expression levels of primary tumors and xenografts appear largely to be due to the presence of normal hematopoietic cells in the diagnostic sample. Similar results were seen by others when comparing the GEP of diagnostic liver tumors to the profiles of the respective xenografts.[6] Another study comparing primary neuroblastoma tumors and derivative cell lines found that stromal contamination of primary tumors is a major contributing factor in underestimation of loss of heterozygosity and copy numbers.[7] However, stromal cells in the xenotransplanted samples would be of murine origin and presumably not

be detected by the array which is insensitive to murine RNA due to differences in the 3' sequence of homolog murine and human genes.

Furthermore, genes differentially regulated between D and xenograft passages separated an independent cohort of 30 ALL diagnostic samples into 2 clusters: one cluster grouping samples with low leukemia infiltration in the BM (< 70%) in contrast to the second cluster comprising patients with a high proportion of blast cells (>96%). Interestingly, samples with a lower percentage of leukemia cells show expression profiles similar to those at diagnosis (D) (Figure 1D).

The 23 genes upregulated in xenograft samples compared to diagnosis are mainly involved in cell cycle regulation and mitosis and have been described in connection with cancer subtypes.[8-15] Their upregulation might therefore be an indicator of a high proliferative state in general and argue towards a more aggressive potential of the engrafted leukemia cells but alternatively could also simply be due to the fact that the xenograft samples are pure leukemic blasts and do not contain up to 15% of non-cycling healthy bone marrow cells as do the diagnostic samples.

A study comparing a large panel of pediatric cancers to the tumor type of origin concluded that characteristic expression patterns of primary tumors were maintained in the corresponding xenografts in the majority of samples and that xenografts more closely resemble primary tumors than cell lines. But the study did not compare xenografts to the matching diagnostic sample, as the xenografts in the study were not related to the diagnostic samples analyzed.[16] Similar results were seen analyzing small cell lung cancer xenografts, cell lines derived from these as well as secondary xenografts from the cell lines showing that xenografts were much more similar to primary tumors on gene expression level than cell lines or secondary xenografts.[17]

It is widely accepted that transcriptional similarity reflects the overall biology of the cancer, as seen in gene expression profiling studies on primary pediatric ALL.[18,19] Our study show that the NOD/SCID-model recapitulates the primary human leukemia, reflects the inherent variability of different ALL and closely mimics the features of the primary malignancy and, most importantly, retains these over serial passages.

Our data moreover have implications for the current issue of the nature of leukemia stem cells. The results seem to exclude the possibility of selection of a specific subclone of leukemia initiating cells in the NOD/SCID-model, leading to a distinct GEP from its diagnostic counterpart. The data rather support recent findings that SCID-repopulating cells

reconstitute the original bulk tumor.[20] We conclude that monitoring 7 leukemias different in origin and leukemia characteristics there is no systematic influence of xenotransplantation on the GEP between diagnosis and xenopassages in the mouse.

TABLES

Table 1. Clinical characteristics of primary patient samples and xenograft samples

sample	immuno-phenotype	cyto-genetics	gender	age [yrs]	WBC [1,000/µl]	leukemia cells PB [%]	total leukemia cell count PB [1,000/µl]	leukemia cells BM [%]
# 1	cortical T-ALL	t(10;14)	m	12	36,5	45.5	16,6	92.5
# 2	common ALL	neg.	f	8.8	17,1	57.5	9,8	97.5
# 3	pro-B ALL	t(4;11)	f	0.7	164,4	81	133,2	90
# 4	pre-B ALL	t(9;14)	m	3.6	10,0	47	4,7	91.5
# 5	pre-B ALL	t(9;22)	f	5.6	14,9	71	10,6	98
# 6	common ALL	neg.	m	16	9,3	86.5	8,0	92.5
# 7	pro-B-ALL	t(11;19)	m	0.8	31,5	91	28,7	n. a.

(A) Clinical characteristics of the 7 patients at diagnosis, the leukemia subtypes representing the common immunophenotypic subtypes of childhood ALL (pro-, common- and pre-B ALL as well as one T- cell precursor ALL). Three of the patients displayed high white blood cell counts (WBC) at diagnosis (#1, #3 and #7), while the others had normal (#4 and #6) or only mildy elevated (#2 and #5) leukocyte counts (PB indicates peripheral blood; BM, bone marrow; m, male; f, female).

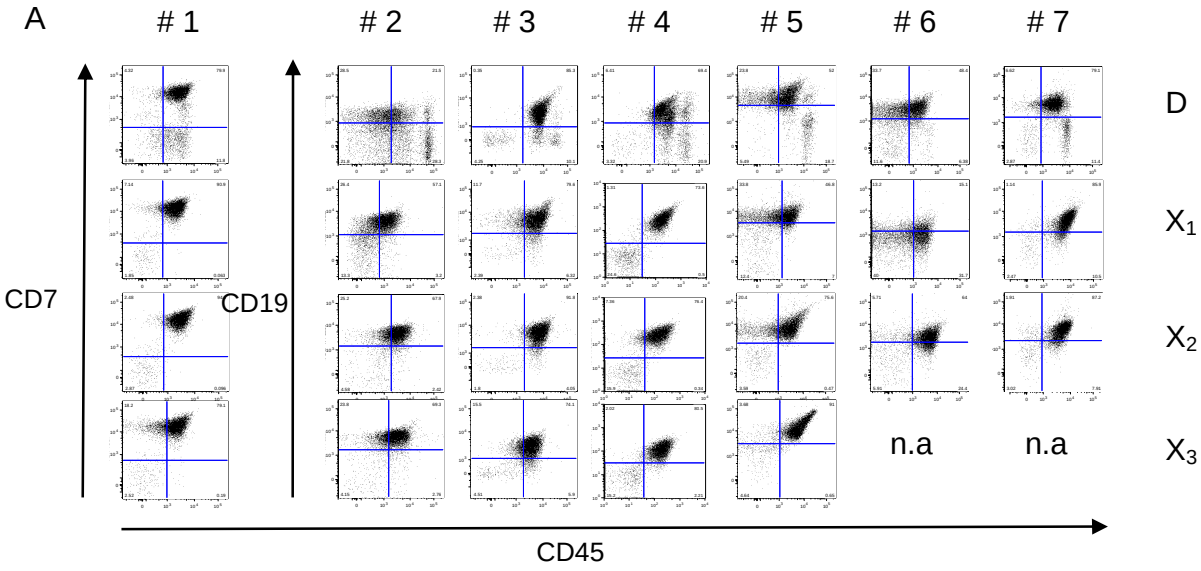
sample	xenograft passage	human leukemia cells BM [%]	total cell count spleen [$*10^6$ cells]	human leukemia cells spleen [%]	human leukemia cells PB [%]
# 1	X_1	82	200	92	55
	X_2	84	300	89	40
	X_3	86	920	79	22
# 2	X_1	87	180	89	5
	X_2	86	270	89	2
	X_3	91	200	94	7
# 3	X_1	92	100	94	44
	X_2	85	342	93	45
	X_3	87	340	92	48
# 4	X_1	70	19	87	1
	X_2	81	26	82	2
	X_3	85	13	74	3
# 5	X_1	83	340	88	5
	X_2	80	815	92	12
	X_3	86	330	96	12
# 6	X_1	74	200	88	14
	X_2	84	360	77	12
	X_3	83	320	86	12
# 7	X_1	87	260	89	55
	X_2	84	260	88	60
	X_3	86	300	87	58

(B) Features of the leukemia- bearing mice at autopsy after primary (X_1), secondary (X_2) and tertiary (X_3) transplantation. The values represent mean values of two mice. The total cell count from spleen was assessed

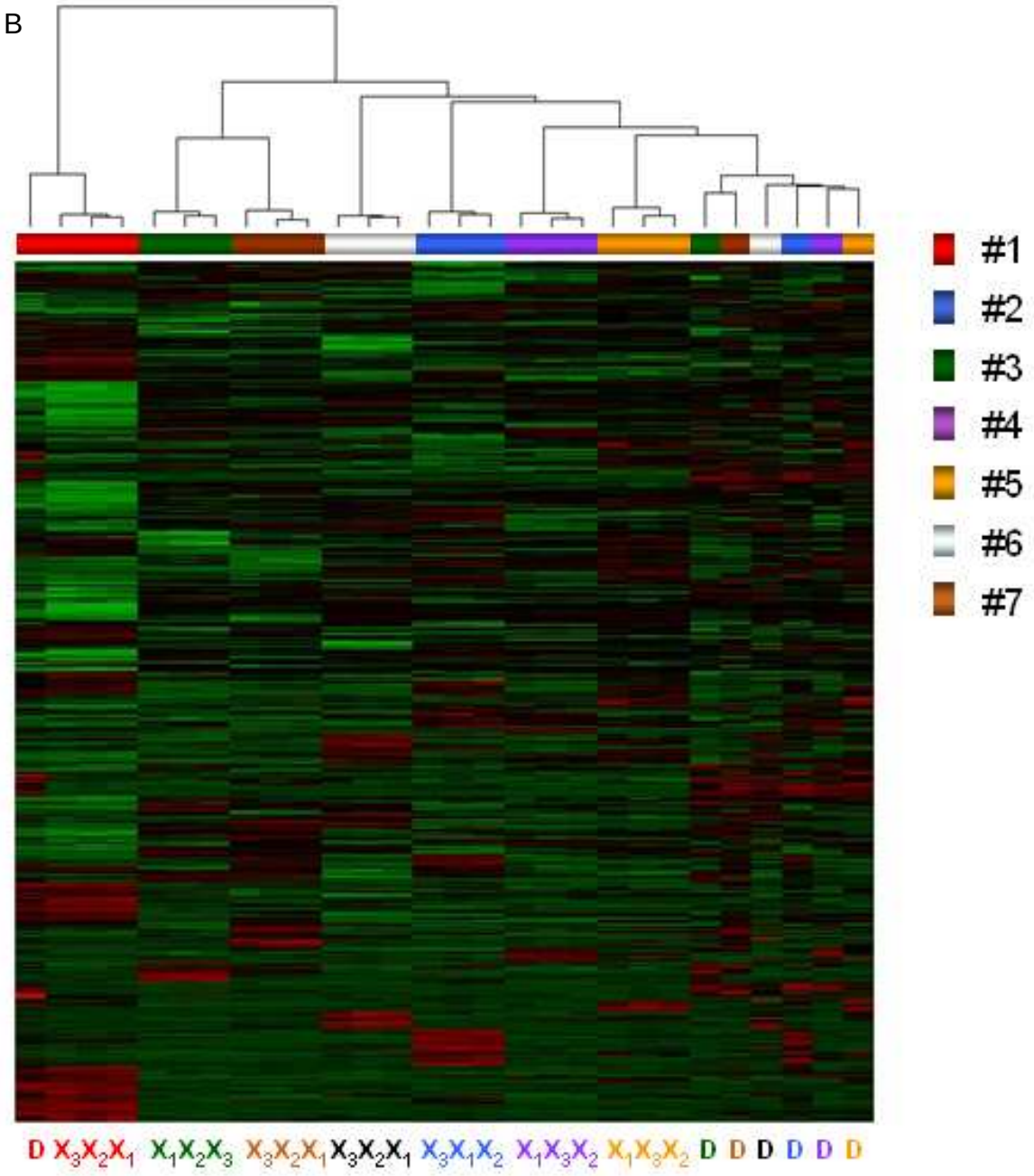
after cell isolation by density gradient centrifugation. The percentages of human leukemia cells were assessed by flow cytometry, by staining human CD45, human CD19, human CD7 and human CD4 and murine CD45 (Ly-5).

FIGURES

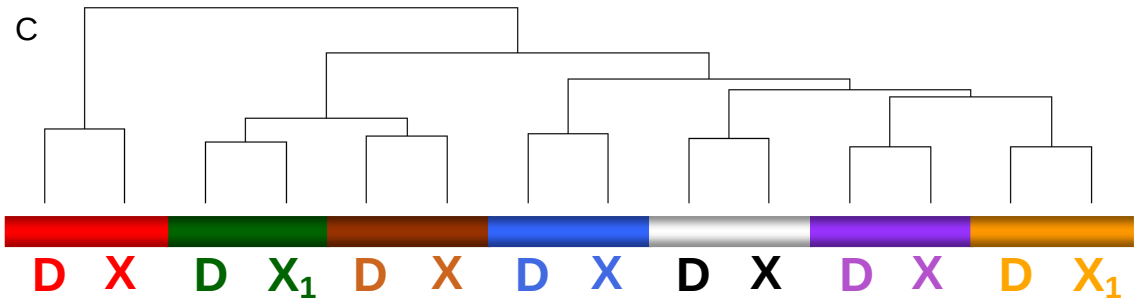
Figure 1. Immunophenotype and gene expression profile remain stable after serial transplantation



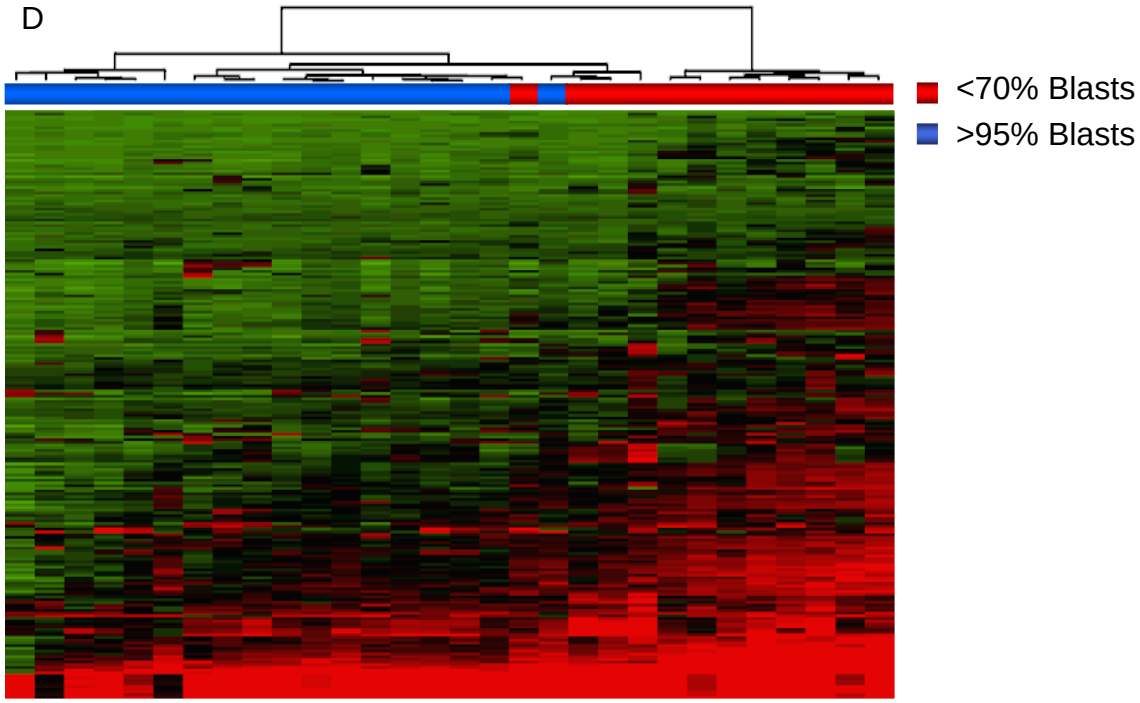
(A) Immunophenotypic analysis of leukemia cells at diagnosis (D) and after primary (X_1), secondary (X_2) and tertiary (X_3) transplantation. The expression of the typical leukemia surface markers analyzed (shown here: CD7 for T- cell precursor ALL and CD19 for B- cell precursor ALL, both against CD45) remained constant over several passages in the NOD/SCID mice, while the CD45-positive non-leukemic cells present at diagnosis disappeared after the first xenograft passage.



(B) Unsupervised hierarchical clustering of all samples using a filter on variance 90% (Ward's method, Euclidean distance). The cortical T- ALL and the derived xenografts cluster separately (#1, red), while all the diagnostic B- cell precursor ALL samples cluster together (D), although their derived xenograft samples all cluster correctly with each other (X_1-X_3).



(C) Unsupervised hierarchical clustering of diagnostic (D) and primary xenograft (X_1) samples using a filter on variance 90% (Ward's method, Euclidean distance). In this analysis the diagnostic sample always clusters together with it's derived primary xenograft sample.



(D) Clustering of independent cohort of ALL diagnostic samples selected by very high (>95%) or low (<70%) bone marrow blasts invasion. Samples were clustered using 254 differentially expressed genes among D and X1. Supervised analysis separates high blasts invasion samples from others. Low blasts invasion samples (<70%) showed expression profiles similar to those at diagnosis (D), indicating that different gene expression between diagnostic samples and their respective xenograft are likely to be due to the presence of a higher number of residual healthy bone marrow cells in diagnostic samples.

SUPPLEMENTARI TABLE

ST 1. Confirmation of clonal identity in serial xenograft passages

sample	type	MILE classifier	Identical SNPs [%]	fusion gene PCR
# 1	D	C4	D – X_1: 99.8	NA
	X_1	C4	D – X_3: 97.8	NA
	X_2	C4	X_1- X_3: 99.8	NA
	X_3	C4		NA
# 2	D	C8	D – X_1: 99.9	NA
	X_1	C8	D – X_3: 99.9	NA
	X_2	C8	X_1- X_3: 99.9	NA
	X_3	C8		NA
# 3	D	C2		*MLL/AF4*
	X_1	C2	ND	*MLL/AF4*
	X_2	C2		*MLL/AF4*
	X_3	C2		*MLL/AF4*
# 4	D	C8	D – X_1: 99.8	NA
	X_1	C8	D – X_3: 99.9	NA
	X_2	C8	X_1- X_3: 99.9	NA
	X_3	C8		NA
# 5	D	C3		ND
	X_1	C3	ND	ND
	X_2	C3		ND
	X_3	C3		ND
# 6	D	C8		NA
	X_1	C3	ND	NA
	X_2	C3		NA
	X_3	C3		NA
# 7	D	C2		ND
	X_1	C2	ND	ND
	X_2	C2		ND
	X_3	C2		ND

MILE-classification, percentage of similarities in SNP-analysis (Affymetrix Genome-Wide Human SNP 6.0 Array) between diagnostic and xenograft samples and monitoring for the presence of the diagnostic fusion transcript. MILE classes: C1= mature B-ALL with t(8;14) C2= MLL/pro-B ALL, C3= B-ALL with BCR/ABL, C4= T-cell precursor ALL, C5= B-ALL with TEL/AML1, C6= B-ALL with t(1;19), C7= B-ALL with hyperdiploid karyotype, C8= all other B-ALL.

REFERENCES

1. Liem NL, Papa RA, Milross CG, et al. Characterization of childhood acute lymphoblastic leukemia xenograft models for the preclinical evaluation of new therapies. Blood. 2004;103:3905-3914.
2. Borgmann A, Baldy C, von Stackelberg A, et al. Childhood ALL blasts retain phenotypic and genotypic characteristics upon long-term serial passage in NOD/SCID mice. Pediatr Hematol Oncol. 2000;17:635-650.
3. Dialynas DP, Shao L, Billman GF, Yu J. Engraftment of human T-cell acute lymphoblastic leukemia in immunodeficient NOD/SCID mice which have been preconditioned by injection of human cord blood. Stem Cells. 2001;19:443-452.
4. Lock RB, Liem N, Farnsworth ML, et al. The nonobese diabetic/severe combined immunodeficient (NOD/SCID) mouse model of childhood acute lymphoblastic leukemia reveals intrinsic differences in biologic characteristics at diagnosis and relapse. Blood. 2002;99:4100-4108.
5. Haferlach T, Kohlmann A, Wieczorek L, et al. The Clinical Utility of Microarray-Based Gene Expression Profiling in the Diagnosis
and Subclassification of Leukemia: Report on 3334 Cases from the International
MILE Study Group. J Clin Onc in press
. 2009.
6. Mischek D, Steinborn R, Petznek H, et al. Molecularly characterised xenograft tumour mouse models: valuable tools for evaluation of new therapeutic strategies for secondary liver cancers. J Biomed Biotechnol. 2009;2009:437284.
7. Volchenboum SL, Li C, Li S, et al. Comparison of primary neuroblastoma tumors and derivative early-passage cell lines using genome-wide single nucleotide polymorphism array analysis. Cancer Res. 2009;69:4143-4149.
8. Zhou X, Temam S, Oh M, et al. Global expression-based classification of lymph node metastasis and extracapsular spread of oral tongue squamous cell carcinoma. Neoplasia. 2006;8:925-932.
9. Kanehira M, Harada Y, Takata R, et al. Involvement of upregulation of DEPDC1 (DEP domain containing 1) in bladder carcinogenesis. Oncogene. 2007;26:6448-6455.
10. Eun SY, Woo IS, Jang HS, et al. Identification of cytochrome c oxidase subunit 6A1 as a suppressor of Bax-induced cell death by yeast-based functional screening. Biochem Biophys Res Commun. 2008;373:58-63.
11. Hsieh WJ, Hsieh SC, Chen CC, Wang FF. Human DDA3 is an oncoprotein down-regulated by p53 and DNA damage. Biochem Biophys Res Commun. 2008;369:567-572.
12. Troncone G, Guerriero E, Pallante P, et al. UbcH10 expression in human lymphomas. Histopathology. 2009;54:731-740.
13. Corson TW, Huang A, Tsao MS, Gallie BL. KIF14 is a candidate oncogene in the 1q minimal region of genomic gain in multiple cancers. Oncogene. 2005;24:4741-4753.
14. Lin SY, Pan HW, Liu SH, et al. ASPM is a novel marker for vascular invasion, early recurrence, and poor prognosis of hepatocellular carcinoma. Clin Cancer Res. 2008;14:4814-4820.
15. Wang YH, Takanashi M, Tsuji K, et al. Level of DNA topoisomerase IIalpha mRNA predicts the treatment response of relapsed acute leukemic patients. Leuk Res. 2009;33:902-907.
16. Whiteford CC, Bilke S, Greer BT, et al. Credentialing preclinical pediatric xenograft models using gene expression and tissue microarray analysis. Cancer Res. 2007;67:32-40.
17. Daniel VC, Marchionni L, Hierman JS, et al. A primary xenograft model of small-cell lung cancer reveals irreversible changes in gene expression imposed by culture in vitro. Cancer Res. 2009;69:3364-3373.
18. Yeoh EJ, Ross ME, Shurtleff SA, et al. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. Cancer Cell. 2002;1:133-143.
19. Holleman A, Cheok MH, den Boer ML, et al. Gene-expression patterns in drug-resistant acute lymphoblastic leukemia cells and response to treatment. N Engl J Med. 2004;351:533-542.
20. Le Viseur C, Hotfilder M, Bomken S, et al. In childhood acute lymphoblastic leukemia, blasts at different stages of immunophenotypic maturation have stem cell properties. Cancer Cell. 2008;14:47-58.

CHAPTER 5

Early Relapse in Pediatric ALL is identified by Time To Leukemia in NOD/SCID mice and is characterized by a gene signature involving survival pathways

Lüder Hinrich Meyer, Sarah Mirjam Eckhoff, Manon Queudeville, Johann Michael Kraus, Marco Giordan, Jana Stursberg, Andrea Zangrando, Elena Vendramini, Anja Möricke, Martin Zimmermann, Andre Schrauder, Georgia Lahr, Karlheinz Holzmann, Martin Schrappe, Giuseppe Basso, Karsten Stahnke, Hans Armin Kestler , Geertruy te Kronnie, Klaus-Michael Debatin.

IN PRESS CANCER CELL, 18th February 2011

ABSTRACT

Summary

We investigated engraftment properties and impact on patient outcome of 50 pediatric acute lymphoblastic leukemia (ALL) samples transplanted into NOD/SCID mice. Time to leukemia (TTL) was determined for each patient sample engrafted as weeks from transplant to overt leukemia. Short TTL was strongly associated with high risk for early relapse, identifying a new independent prognostic factor. This high risk phenotype is reflected by a gene signature which upon validation in an independent patient cohort (N=197) identified a high risk cluster of patients with early relapse. Furthermore, the signature points to independent pathways, including mTOR, involved in cell growth and apoptosis. The pathways identified can directly be targeted thereby offering additional treatment approaches for these high risk patients.

Significance

Treatment of pediatric ALL is increasingly successful, achieving cure rates of over 80%. Relapse is associated with poor outcome, however the majority of relapsed patients is initially stratified into low risk groups based on current stratification markers. This emphasizes the need for additional factors for upfront identification of high risk patients. In this study we describe a strong association of the engraftment phenotype found in a series of transplanted primary pediatric ALL cells in a mouse in vivo model and early patient relapse. Gene expression profiling revealed a set of genes associated with this aggressive phenotype

providing a potential strategy to identify these high-risk patients. Most importantly, pathways regulating cell growth are identified, providing targets for alternative therapeutic strategies.

INTRODUCTION

Acute lymphoblastic leukemia (ALL) is the most frequent malignant disease in children and adolescents. Although intensification of multiagent chemotherapy regimens and advances in supportive care have lead to improvement of remission induction and long term survival over the last decades presently achieving cure rates of over 80%, about 20% of the patients suffer from relapse associated with an inferior prognosis [1-3]. Particularly, in 10% of all patients relapse occurs at early time points associated with a substantially reduced survival of 5% or less [4-6]. Stratification of patients into therapy regimens of different intensity is based on the individual risk for relapse by using prognostic factors including cytogenetics and response to therapy to define different risk groups (high-risk, HR and non-high risk, non-HR) [1, 7, 8]. Leukemia cell clearance in response to steroid treatment is evaluated and has been used as a prognostic marker for almost 20 years demonstrating inferior survival for prednisone poor responding (PPR) patients [1, 2, 9-11]. In addition, detection of residual leukemia cells at submicroscopic levels after remission induction therapy (minimal residual disease, MRD) qualifies for HR-treatment [12-16]. Early identification of patients with high risk for relapse has led to improved outcome. However, two thirds of patients encountering relapse were initially stratified into non-HR groups[2]. This limitation of prognostic factors emphasizes the need to identify novel parameters which ideally also reflect the functional biology of the disease for the development of molecular based risk stratification and identification of therapeutic targets.

While leukemia cells cannot sufficiently be cultured *in vitro*, xenograft-models can overcome these limitations. Transplantation of primary leukemia cells into NOD/SCID mice results in recipients exhibiting leukemia recapitulating the human disease [17-20]. We adopted the NOD/SCID/huALL model and report for the first time that rapid development of leukemia in NOD/SCID mice engrafted with primary ALL cells is characteristic for patients with early relapse. Furthermore, we identified a gene signature for relapse pointing to pathways involved in regulation of cell growth and survival. The signature was applied on an independent patient cohort and identified a high risk cluster of patients with early relapse.

RESULTS

NOD/SCID/huALL

Leukemia cells isolated from diagnostic samples of 50 pediatric de novo B-cell precursor (BCP) ALL patients were transplanted onto NOD/SCID mice ($1x10^7$ cells per recipient). Recipient mice were investigated regularly for the presence of leukemia and sacrificed at manifestation of disease. Leukemia was confirmed by detection of a high percentage of human leukemia cells in the peripheral blood of the mice and infiltration of bone marrow and spleen by leukemia cells using flowcytometry. Of 50 samples, 29 led to manifestation of overt leukemia during the observation time of 20 weeks while 21 transplants did not manifest disease within this time. Engraftment of leukemia was also monitored in different organ compartments at different time points. As exemplified for two samples, leukemia was consistently first detected in bone marrow and spleen followed by the occurrence of ALL cells in peripheral blood (Fig. 1A).

TIME TO LEUKEMIA (TTL)

When we analyzed leukemia engraftment of different samples we observed a prominent difference in time elapsing from transplantation to manifestation of leukemia. Rapid onset of leukemia related morbidity was observed in some samples (4, 5, 6, 8 and 9 weeks after transplant) in contrast to other samples giving rise to leukemia after a substantially longer period of time (12 to 26 weeks). No onset of leukemia was observed between 9 and 12 weeks after transplant suggesting distinct engraftment phenotypes of different leukemia cell samples. The time from transplant to leukemia was defined as 'Time To Leukemia' (TTL) and quantified for each leukemia sample. A threshold of 10 weeks was used to distinguish the rapid from the slow engraftment phenotype, a TTL of less than 10 weeks was designated TTL^{short} and longer time periods were classified as TTL^{long} (Fig. 1B).

Additional cutoff points representing intervals at 16 and 21 weeks (Supp. Fig. SF 1) and leukemia characteristics (ALL-BFM high risk stratification, TEL/AML1 fusion, prednisone poor response, age, and white blood cells at diagnosis) were tested to divide the cohort. Age (unfavorable: 0-1 and above 9 years; favorable: 1-9 years) and hyperleukocytosis at diagnosis (above or below 50 white blood cells [1000/µl]) were classified according to St. Judes Total Therapy Study XIIIB protocol criteria [21]. Time to overt leukemia (in the recipient mice) and relapse free survival (of the patients) were compared between the subgroups divided

according to the different cutoff points and pretreatment and clinical response features by Kaplan Meier analysis and log-rank test. Resulting P values were adjusted for multiplicity (Bonferroni's method) with the most distinct grouping according to the lowest adjusted significant P value. The 10 weeks cutoff point was shown to be most discriminatory with respect to engraftment and relapse free survival (Supp. Tab. ST 1A and B).

Of the 50 samples transplanted, 21 samples did not expand *in vivo* to substantial cell numbers within the observation time. Leukemia manifestation after week 20 was not analyzed systematically in our study. However, two leukemias were followed up for a prolonged time and showed disease manifestation with a TTL of 36 and 44 weeks, suggesting that eventually all of the transplanted leukemia cell samples lead to overt leukemia if the observation time is long enough. The samples showing no overt leukemia at week 20 were included into the TTL^{long} group.

STABILITY, CONSISTENCY AND INDEPENDENCY OF THE TTL PHENOTYPE

Time to leukemia (TTL) was in all cases assessed after transplantation of $1x10^7$ of primary patient leukemia cells. Sets of primary leukemia samples obtained at diagnosis were transplanted in parallel ($1x10^7$ cells) and TTL was assessed for both recipients of each leukemia sample. Similar TTL values were observed within the replicates of 21 primary leukemia samples (Supp. Tab. ST 2A). Furthermore, two leukemia samples obtained at relapse were transplanted as replicates and compared to the respective matching diagnosis samples. Similar TTL values were also observed for the replicates of the relapse samples. Moreover, similar time to leukemia was seen comparing diagnosis-relapse pairs (Supp. Tab. ST 2B), indicating stability of the TTL phenotype even at leukemia relapse.

We also analyzed the consistency of TTL in subsequent recipients. Leukemia cells of diagnostic patient samples and the consecutive xenograft passages were transplanted in parallel onto secondary and tertiary recipients using the identical dose of $1x10^7$ cells for each recipient. TTL was assessed for both recipients of each passage and similar TTL values were observed constantly within the replicates. Thus, TTL phenotypes did not change and remained TTL^{short} or TTL^{long} between diagnosis and consecutive passages in the NOD/SCID/huALL model (Supp. Tab. ST 2C).

Also later passages were investigated. Representative xenograft samples for TTL^{long} and TTL^{short}, as characterized upon initial transplantation of the diagnostic sample, were retransplanted onto primary, tertiary or 6^{th} recipients. The dose of $1x10^7$ cells per recipient as

used throughout the study recapitulated the TTL phenotype of the respective leukemia remaining stable also after up to 5 passages (Fig. 1C and Suppl. Tab. ST 2D).

We also addressed whether the TTL phenotype might be influenced by varying numbers of cells transplanted. Higher numbers of TTL^{long} leukemia cells ($3x10^7$ cells per animal) did not accelerate engraftment, nor did a decrease of cell numbers ($1x10^6$, $1x10^5$ cells per mouse) prolong TTL^{short} demonstrating that the TTL phenotypes are largely independent from cell numbers transplanted and represent an intrinsic feature of the individual leukemia (Fig. 1C and Suppl. Tab. ST 2D).

TTL AND PATIENT OUTCOME

Six patients exhibited TTL^{short} and 44 TTL^{long}. Five of these 6 TTL^{short} patients encountered medullar relapse within 24 months after diagnosis. This corresponds to the expected proportion of about 10% of very high risk/early relapsing patients.

We analyzed relapse free survival of the patients whose leukemia cells were used for transplant. Patients whose transplanted cells showed a TTL^{long} phenotype revealed a marked superior relapse free survival compared to TTL^{short} patients. Most interestingly, all relapses of the patients characterized by TTL^{short} occurred early within the first two years after diagnosis (Fig. 2A). Thus, the TTL^{short} phenotype is characteristic for patients at high risk for early relapse.

Since the majority of relapse patients originates from the low risk groups we analyzed whether time to leukemia (TTL) identifies patients who would encounter an early relapse despite stratification into low risk groups. Therefore, the analysis was focused on the non-high risk group patients (N=43; TTL^{short} n=4, TTL^{long} n=39). A clear inferior relapse free survival of TTL^{short} patients was observed, all 4 non-HR patients characterized by TTL^{short} developed early medullar relapse within 24 months after diagnosis (Fig. 2B). This demonstrates that TTL is associated with patient outcome independent of risk stratification.

21 leukemia samples did not lead to leukemia in the recipients within the observation time and were therefore included into the TTL^{long} group. We also analyzed whether this non-engraftment subgroup would be associated with a distinct patient outcome. Most importantly, the group of patients whose leukemia cells upon transplant did not lead to leukemia within 20 weeks observation exhibited further increased superior relapse free survival, even achieving 100% survival in the non-HR group (Fig. 2C and D).

A superior relapse free survival was also observed if only patient samples leading to overt leukemia (all risk groups, N=29; non- high risk groups, N=25) were analyzed and patients whose leukemia cells did not engraft were excluded (Fig. 2E and F). Furthermore, a significant correlation of TTL (mouse model) and remission duration (patients) was observed analyzing engrafted samples of relapsed patients (N=9, all risk groups; N=7, non- HR groups) (Fig. 2G and H).

In summary, the TTL phenotype is highly associated to patient survival, i.e. TTL^{short} identifies early relapse patients and TTL^{long} is characteristic for superior relapse free survival of up to 100%, irrespective of stratification into risk groups.

PROGNOSTIC IMPACT OF TTL

Different prognostic factors have been implemented in different treatment protocols to stratify patients based on their individual risk. We compared the prognostic impact of TTL to established prognostic factors: poor leukemia cell clearance in response to steroid treatment (prednisone poor response, PPR) , hyperleukocytosis (highly elevated white blood cell counts > 50 [1000/µl] in the peripheral blood at diagnosis) and unfavorable age (age at diagnosis below one and above nine years) were included [21]. By multivariate Cox's regression, TTL^{short} patients exhibited an 45-fold increased risk for relapse in contrast to no significant elevated risk for patients with prednisone poor response, high white blood cell count at diagnosis or unfavorable age in this cohort (Tab. 1).

Also in comparison to additional thresholds to determine TTL (16 and 21 weeks), TTL^{short} defined by the 10 weeks cutoff was the most significant prognostic factor (Supp. Tab. ST 3).

Genetic alterations are common in pediatric ALL and might be associated with favorable or unfavorable outcome and are therefore used for treatment stratification. The presence of BCR/ABL or MLL/AF4 fusion genes qualify for HR stratification [2, 14]. In the group studied, one patient was positive for MLL/AF4 and displayed a TTL^{short} phenotype. No BCR/ABL positive patients were included. TEL/AML1-fusion is associated with favorable outcome [7]. Most interestingly, all 19 TEL/AML1 positive patients showed TTL^{long}. In addition, only 1 of the 50 patients included in this study met criteria for MRD based high risk stratification. However, this patient was high risk stratified due to prednisone poor response and showed TTL^{long}. Pre- treatment characteristics and response features such as age, gender, peripheral leukocyte count at diagnosis and achievement of remission at the end of induction therapy (day 33) are not associated with TTL^{long} or TTL^{short} (Tab. 2).

TTL IS CONFIRMED IN AN INDEPENDENT LEUKEMIA SUBSET

To evaluate the significance of the TTL phenotype, an independent subset of cell bank BCP ALL samples (4 pairs matched for age, gender and low risk group classification either showing early relapse or not) were transplanted. Early relapse patients always revealed a shorter TTL than patients with late or no relapse. However, transplant of one frozen cell bank sample derived from an early relapse patient showed a 'borderline' TTL of 13 weeks (Suppl. Fig. SF 2A). TTL was significantly associated with the time from diagnosis to relapse (remission duration) (Suppl. Fig. SF 2B) confirming the prognostic value of TTL.

GENE EXPRESSION ANALYSIS OF XENOGRAFT LEUKEMIA

To gain insight into molecular mechanisms responsible for distinct *in vivo* leukemia proliferation gene expression profiles were analyzed using a human whole genome array approach. ALL cells isolated from leukemia bearing recipients (TTL^{long} or TTL^{short}) were investigated. Cytogenetic abnormalities leading to fusion genes and involving transcription factors are recognized by specific transcription profiles [22] and might overcast other differences in gene regulation. Therefore, the analysis has been focused on xenografts derived from patients without known gene fusions and also sufficient follow up time (N=12, TTL^{short} n=5, TTL^{long} n=7) (Suppl. Tab. ST 4A and B). The expression profiles were analyzed employing a model-free shrinkage estimate of the variance across genes ranking the data set ("shrinkage t" statistic [23]). 73 genes (88 probe sets) were identified to be differentially regulated between TTL^{short} and TTL^{long} (FDR < 5%), 51 genes up-regulated and 22 down-regulated in TTL^{short} (Fig. 3A and Tab. 3). Among the probe sets with the highest fold changes two genes involved in regulation of mTOR signaling were identified: *DDIT4L* and *RHEB*. *DDIT4L* codes for DNA-damage-inducible-transcript-4-like, a molecule negatively regulating mTOR. Consistent with this function (i.e. inhibiting mTOR, a central molecule regulating cell proliferation and survival) *DDIT4L* is up-regulated in TTL^{long} (inhibition of mTOR) and down-regulated in TTL^{short} (lack of mTOR inhibition). *RHEB*, coding for the positive mTOR regulator Ras homolog enriched in brain, was identified to be up-regulated in TTL^{short} leukemia by two probe sets. In line with its mTOR activating function, *RHEB* was found to be up-regulated in TTL^{short}/early relapse leukemia.

In addition, two genes coding for molecules involved in regulation apoptosis were identified displaying high fold changes: *PDE4A* and *DAPK1* (Tab. 3). The type 4 cyclic AMP phosphodiesterase PDE4A down-regulates cAMP levels leading to impaired apoptosis sensitivity. In line, *PDE4A* is up-regulated in TTL^{short}/early relapse leukemia. *DAPK1* codes for the pro-apoptotic death associated protein kinase and is, congruent with its apoptosis sensitizing effect, highly expressed in TTL^{long}/good prognosis leukemia.

DIFFERENTIAL REGULATION OF IDENTIFIED GENES IN XENOGRAFT ALL

The differential regulation of transcripts identified in gene array analysis was analyzed by quantitative RT-PCR. ALL xenograft samples (N=24; n=12 used for array analysis and n=12 without analyzed gene profile; no known gene fusions, n=9 TTL^{short}, n=15 TTL^{long}) were analyzed. A significant differential regulation of *PDE4A*, *DDIT4L* and *RHEB* in TTL^{short} versus TTL^{long} leukemia was confirmed (Fig. 3B, C and D). Additionally, also *FRAP1* coding for mTOR was analyzed and itself found to be significantly up-regulated in TTL^{short} (Fig. 3E). Moreover, transcript levels of the mTOR regulators *DDIT4L*, *RHEB* and *FRAP1* itself were also highly associated with TTL (Suppl. Tab. ST 6).

APPLICATION OF THE XENOGRAFT SIGNATURE ON AN INDEPENDENT PA-TIENT COHORT

To corroborate the relevance of the specific expression profile identified in xenograft leukemia samples the TTL signature was applied onto profiles of an independent cohort of pediatric BCP ALL patients (N=197, samples isolated at diagnosis; Supp. Tab. ST 5) who have also been treated according to a BFM-based protocol (AIEOP-LLA-2000). This independent set of patient profiles was analyzed with respect to the best distinction into subgroups according to the signature by a cluster number estimation procedure for k-means via repeated clusterings on re-sampled versions of the data including a correction for random partitions [24]. Based on the 88 probe sets, two robust clusters were identified within the 197 patients (Supp. Fig. SF 3A).

Most importantly, patients grouped into the TTL^{short} cluster displayed a significantly inferior relapse free survival in contrast to patients clustering with the TTL^{long} profile (log-rank test, P = 2.5 e-4) (Fig. 3F).

A CLASSIFIER BASED ON THE TTL SIGNATURE IDENTIFIES EARLY RELAPSE PATIENTS

The TTL^{short} phenotype and corresponding expression signature are strongly associated with early relapse leukemia. For this reason we also analyzed, whether based on the TTL profile a classifier could be obtained to identify early relapse patients. A classifier was generated by conjunction of threshold decisions on different genes using a set covering machine (SCM) with data dependent rays [25, 26]. Starting from the whole signature of 88 probes sets, the SCM reduced the number of genes employed in the decision to a set showing best early relapse prediction. Utilizing 20 probe sets (including *RHEB* and *DAPK1*; Suppl. Fig. SF 3B) all early relapse patients were identified in the patients. The set was refined to further improve the robustness yielding a subset of 5 probe sets (Suppl. Fig. SF 3C). Applying this rule to the 197 patients all 26 early relapse cases are detected (100% sensitivity) with an overall classification accuracy of 75%. Cross-validating (leave-one-out) this classifier achieved a sensitivity of 77% and an accuracy of 72%.

CELL DEATH IS PREDOMINANTLY INDUCED IN TTL^{short} BUT NOT TTL^{long} LEUKEMIA CELLS

An activated (*DDIT4L*^{low}/*RHEB*^{high}/*FRAP1*^{high}) mTOR pathway was found to be characteristic for the TTL^{short} phenotype. Additionally, *PDE4A*^{high} was also associated with TTL^{short} leukemia. Inhibition of mTOR or PDE4A might therefore successfully target TTL^{short}/early relapse leukemia (Fig. 4A and 5A). Xenograft ALL samples (TTL^{short} or TTL^{long}) isolated from recipient mice with overt leukemia were treated with either the mTOR inhibitor rapamycin or the phosphodiesterase inhibitor rolipram and cell death was analyzed. Cell death was found to be increased upon *ex vivo* rapamycin treatment in most TTL^{short} (activated mTOR pathway) ALL cells but not in TTL^{long} (*DDIT4L*^{high}) leukemia samples (Fig. 4B). Treatment with rolipram also showed a stronger effect on the TTL^{short} (*PDE4A*^{high}) samples in contrast to TTL^{long} (*PDE4A*^{low}) leukemia cells (Fig. 5B).

These first data on primary pediatric ALL showing an effect of both inhibitors predominantly on TTL^{short} leukemia samples point to a potential use as targets for directed therapy in patients at high risk for early relapse.

DISCUSSION

We identified *in vivo* proliferation of de novo BCP ALL cells transplanted into NOD/SCID mice to be of prognostic and predictive value for patient outcome. While slow engraftment (TTL^{long}) is associated with favorable outcome, TTL^{short} is found in patients with poor prognosis. This biological phenotype is associated with a distinct cellular phenotype characterized by a specific gene signature identifying early relapse patients and pointing to pathways regulating cell proliferation and apoptosis.

Consistent with previous reports, leukemia cells engrafted in our model of NOD/SCID/huALL retained the immunophenotypic and clinical characteristics of the primary leukemia [17-19, 27] and the proportion of leukemia cells in peripheral blood correlated with infiltration in different organs [17, 20]. The observed engraftment phenotype remained stable not only upon transplant of identical cell numbers of the same leukemia sample but was also consistent during consecutive passages.

Intriguingly, we found a most significant correlation of *in vivo* leukemia growth and overall patient outcome (ALL-BFM therapy). Patients whose leukemia cells showed a short time to overt leukemia in the recipient animals (TTL^{short}) relapsed early within 24 months while still on therapy, a feature repeatedly shown to be associated with poor outcome [4, 5]. Most relapses occur in non-high risk patients initially responding to treatment [2] indicating that these patients are not identified by the current stratification strategies. When we focused our analysis on this low risk group only, we also observed a clear cut inferior survival of TTL^{short} patients in this non-HR patient group. A putative prognostic impact of *in vivo* leukemia growth on patient outcome was investigated in a number of studies, so far with conflicting results. This might be explained by the small number of patients studied, the use of less permissive recipients or because leukemia cells from patients at relapse were used [17, 28-30].

Pretreatment characteristics such as age, white blood cell count at diagnosis but also initial treatment response were not significantly associated with increased relapse risk. Most importantly, patients with short TTL had a 45-fold increased relapse risk demonstrating the power of TTL as independent prognostic factor. TEL/AML1- fusion is associated with favorable outcome [7], concordantly with this all TEL/AML1 positive patients (N=19) exhibited TTL^{long} in our xenotransplant model. One TTL^{long} patient included in our study was high risk stratified based on MRD, all 5 TTL^{short} patients with available information were not detected by MRD as high risk. Impaired reduction of leukemia cells after treatment as

assessed by MRD is likely to be due to resistance. Despite relapse, TTL^{short} patients are not characterized by poor treatment response. By transplanting different numbers of TTL^{short} and TTL^{long} leukemia cells we found that engraftment properties are indeed intrinsic to the leukemia cell itself and are retained independently of the cell number transplanted. Also, NOD/SCID/huALL established from cryopreserved leukemia cells in the test set showed an identical correlation of TTL and time to relapse as NOD/SCID/huALL established from fresh samples in our study. These results are in line with observations showing the engraftment ability to be an inherent property also of AML cells transplanted [31, 32].

Whole genome expression analysis was used to characterize specific profiles and/or functionally relevant pathways associated with the TTL phenotype. We identified a signature of 73 genes (FDR < 5%) differentially regulated in TTL^{long} and TTL^{short} xenografts and applied this signature onto gene expression arrays from an independent cohort of 197 ALL patients analyzed with the same platform and treated by a similar risk adapted BFM type protocol. According to the TTL signature the patients clustered robustly into two subgroups with an inferior relapse free survival of the TTL^{short} signature group.

Interestingly, based on the TTL profile a classifier was obtained precisely identifying all early relapse patients of the independent cohort. Although signatures associated with treatment response and also early relapse in pediatric ALL have been previously reported [22, 33-37], this is to our knowledge the first report on a decision tool robustly identifying this very high risk group of early relapse patients. Importantly, these findings also point to activated pathways associated with poor outcome indicating additional therapeutic targets.

Mammalian target of rapamycin (mTOR) is a central downstream switch integrating diverse pathways involved in cell growth and survival [38]. DDIT4L negatively regulates the mTOR pathway inhibiting cell growth and facilitating cell death [39] and is highly expressed in TTL^{long}, suggesting a more efficient control of mTOR in this subgroup. Vice versa, leukemias showing down-regulated *DDIT4L* might lack repression of mTOR thereby exhibiting TTL^{short} and impaired outcome.

Most interestingly, RHEB, directly binding to the mTOR complex and thereby enabling mTOR activation [40] was identified in the TTL signature by two probe sets. *RHEB* is up-regulated in TTL^{short} accounting for increased activation of the mTOR pathway. Furthermore, also up-regulation of *FRAP1* (mTOR) itself, was strongly associated with a short TTL underlining the relevance of this pathway positively regulating cell proliferation for early relapse leukemia.

Taken together, the TTL^{short} phenotype is characterized by *DDIT4L*^{low}/*RHEB*^{high}/*FRAP1*^{high} indicating an activated mTOR pathway in the TTL^{short}/early relapse leukemias. These results have important therapeutic implications since the activated mTOR pathway can be directly targeted inhibiting growth of TTL^{short} leukemia. Rapamycin and related mTOR inhibitors have been shown to inhibit leukemia growth and induce apoptosis in BCP- ALL *in vitro* and *in vivo* alone or in combination with established chemotherapeutic drugs [41-44] and similarities of drug associated signatures identified rapamycin to induce glucocorticoid sensitivity in ALL cells [45]. In fact, TTL^{short} leukemias were more sensitive to *ex vivo* rapamycin treatment than TTL^{long} samples. The individual xenograft leukemias showed heterogeneous sensitivities suggesting additional mechanisms responsible for the TTL phenotype. Indeed, additional genes involved in cell proliferation and apoptosis were identified in the TTL signature.

PDE4A down-regulates cAMP levels leading to impaired apoptosis sensitivity [46] and consistently with this function was found to be up-regulated in TTL^{short} leukemia. Inhibition of PDE4A activity using compounds like rolipram resulted in increased glucocorticoid sensitivity and apoptosis in leukemia cells [47-49]. Treatment of xenograft leukemia samples with rolipram in fact induced cell death predominantly in TTL^{short} leukemia cells.

DAPK1, a tumor suppressor gene whose expression is lost by epigenetic silencing in a number of different tumors including hematopoietic malignancies [50-54] codes for a serine/threonine kinase which positively regulates apoptosis and suppresses tumor progression *in vivo* [55-57]. Consistent with these reports, *DAPK1* was identified to be down-regulated in TTL^{short} suggesting that TTL^{short} cells evade from cell death by loss of this pro-apoptotic regulator. Along this line, we previously reported that deficient apoptosis signaling via the intrinsic pathway is a feature of poor prognosis in primary pediatric ALL and AML [58, 59]. In contrast to its pro-apoptotic and tumor suppressing function, up-regulated *DAPK1* expression was reported in pediatric ALL samples resistant to *in vitro* prednisolone treatment [60]. However, TTL is independent of glucocorticoid resistance with 5 of 6 TTL^{short} patients being prednisone good responders indicating that the engraftment phenotype of primary ALL cells reflects constitutional hallmarks of the leukemia.

Taken together, the proliferative capacity of primary ALL cells in NOD/SCID mice appears to reflect intrinsic properties of the leukemia cells and determines relapse in pediatric ALL. TTL^{short} is characteristic for early relapse patients and associated with a specific gene profile. The biological impact of our observations in the xenotransplant model could be demonstrated

in an independent group of patients and led to identification of distinct pathways involved in regulation of cell growth and apoptosis which can be targeted by well known drugs such as rapamycin. Thus, our NOD/SCID/huALL model provides a powerful tool to identify prognostic factors in acute leukemia at a cellular and molecular level leading to rational targets for therapeutic intervention strategies.

EXPERIMENTAL PROCEDURES

Detailed information with respect to methods and materials used is available in the supplementary appendix.

PATIENTS AND LEUKEMIA SAMPLES

Leukemia samples were obtained at diagnosis from pediatric de novo BCP ALL patients enrolled into the ALL-BFM 2000 study protocol [2] or the AIEOP LLA 2000 study protocol [61] after informed consent was given. Both studies are performed in accordance to the Declaration of Helsinki and registered at http://clinicaltrials.gov (ALL BFM 2000: NCT00430118 and AIEOP LLA 2000: NCT00613457).

NOD/SCID/huALL AND TIME TO LEUKEMIA

Animal studies were approved by the appropriate authority. Patient samples were isolated from diagnostic specimens and consisted of more than 90% leukemia cells. Unconditioned NOD/LtSz-scid/scid mice were transplanted by intravenous injection of $1x10^7$ cells per recipient throughout the study. As an exception, different numbers were injected if the impact of the cell dose was analyzed. TTL was determined for each patient sample transplanted as weeks from transplant to clinical leukemia manifestation. ALL cells were detected and leukemia confirmed in peripheral blood, bone marrow or spleen by flowcytometry [17]. Mice without evidence for disease at week 20 after transplant were killed and absent leukemia infiltration of spleen or bone marrow confirmed no leukemia.

GENE EXPRESSION AND STATISTICAL ANALYSIS

Gene expression analysis was carried out on xenograft leukemia samples with sufficient follow up time and negative for known cytogenetic abnormalities (N=12; TTL^{short} n=5 and TTL^{long} n=7) isolated from recipients with overt leukemia. Expression profiles of the

validation cohort (N=197) were obtained from diagnostic specimens from patients treated according to the AIEOP LLA 2000 protocol. Expression was analyzed using the Human Genome U133 Plus 2.0 Array platform. Gene expression microarray files (Affymetrix .CEL files) were generated using the GCOS 1.4 or 1.2 software (Affymetrix). Arrays have been normalized using robust multiple-array average (RMA) [62]. Expression data were analyzed using Bioconductor package for R (v2.11.1). Data have been deposited in NCBI's Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo) and are accessible through GEO Series accession number GSE13576.

Differentially expressed genes have been calculated by the shrinkage T-statistic [23]. Multiple comparison results were controlled by maintaining a false discovery rate of less than .05 [63]. The signature of 88 differentially expressed probe sets was applied to profiles of 197 independent BCP ALL patients. A robust clustering method based on the K-means algorithm has been applied to estimate the correct number of clusters in this data set [24]. Kaplan-Meier analysis has been performed of relapse free survival in the two clusters obtained from the cluster analysis. Classification into TTL^{short} and TTL^{long} classes was computed with a set covering machine algorithm [25, 26] and a leave-one-out method has been used to validate the classifier.

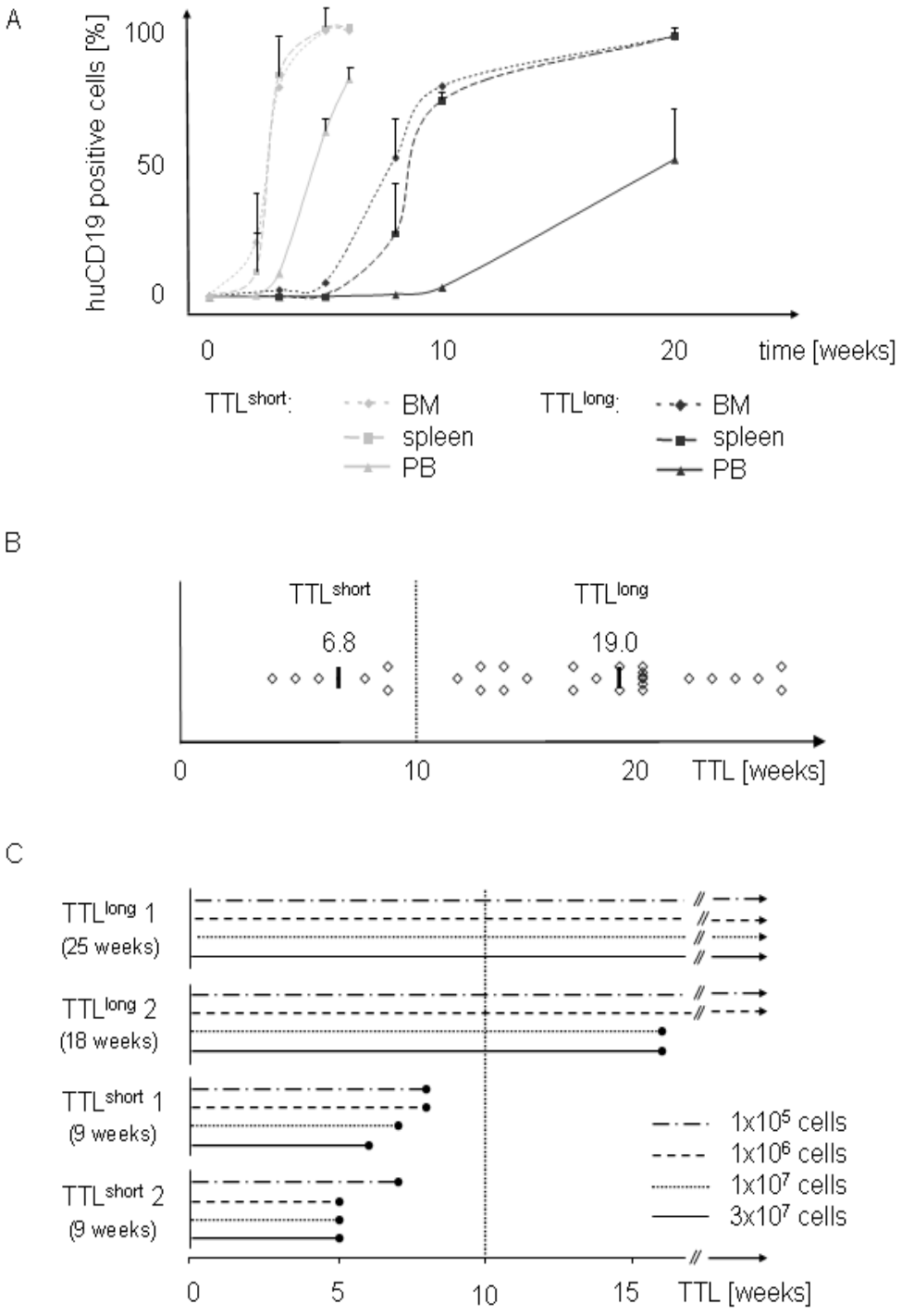
All other statistical analyses were carried out using SPSS 11.0 software (SPSS, Munich, Germany) or 'R' (v. 2.11.1) [64]. P- values of $< .05$ were considered significant in all tests carried out in this study.

EX VIVO TREATMENT OF XENOGRAFT LEUKEMIA

Leukemia cells were isolated from recipient animals as described and incubated with either rapamycin (10 and 100 nM) or rolipram (10 and 100 µM) diluted in dimethyl sulphoxide (DMSO) or DMSO alone and cell viability was assessed by flowcytometry (forward/side scatter criteria). Specific cell death was calculated as described previously [65].

FIGURES

Figure1. Distinct engraftment properties of NOD/SCID/huALL



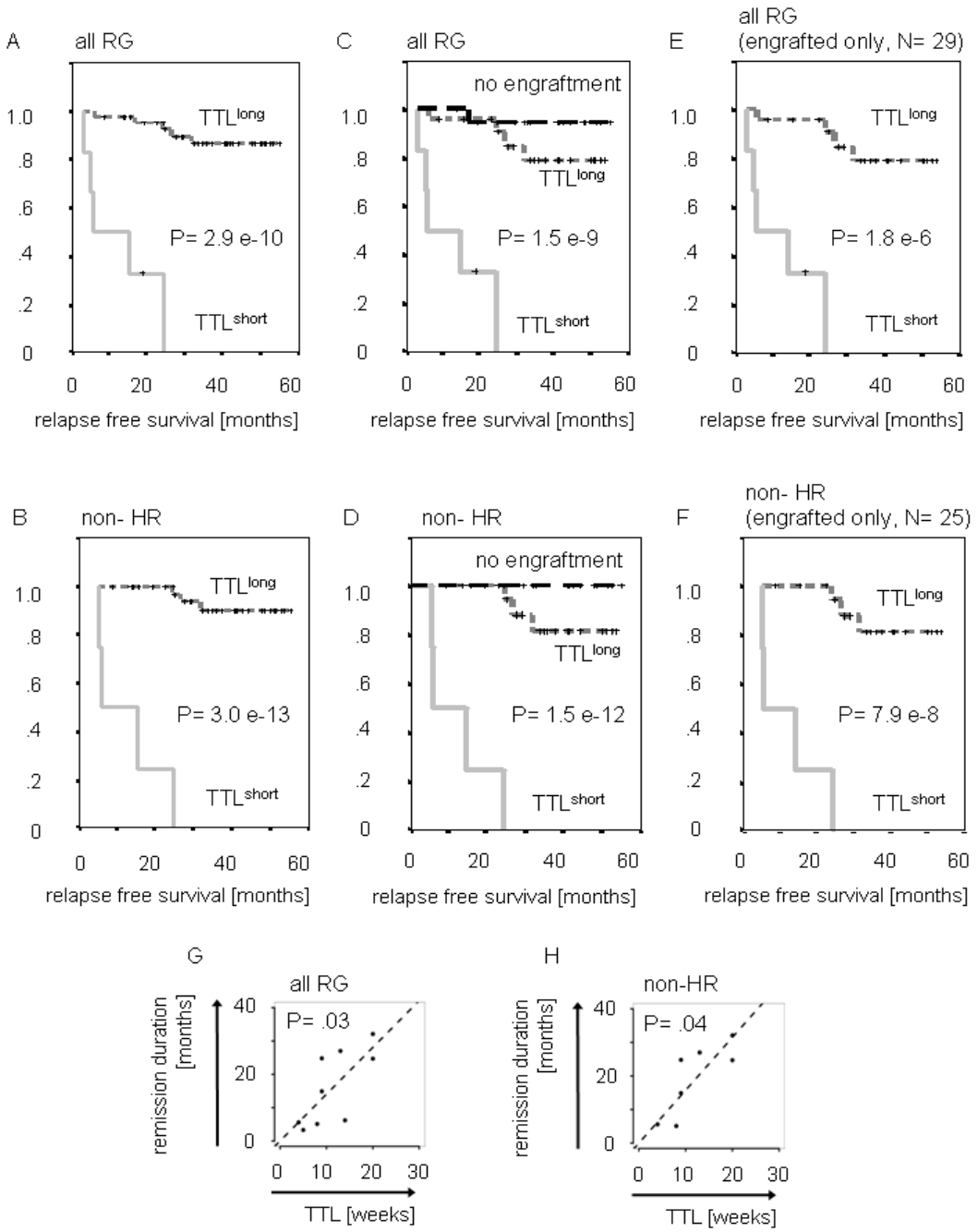
A *In vivo* expansion of leukemia cells in different organ compartments of recipients over time. Two NOD/SCID/huALL samples (TTL^{short}, light grey; TTL^{long}, dark grey) were transplanted into recipient mice (n=3 per group) and the proportions of huCD19 positive ALL cells were detected in bone marrow (rhombs, dotted

lines), spleen (squares, dashed lines) or peripheral blood (triangles, solid lines) by flowcytometry. Data points represent mean values of the proportions of huCD19 positive cells (the standard error of the mean is indicated if > 1%). Infiltration of leukemia was first observed in bone marrow and spleen followed by appearance in peripheral blood.

B Distinct differences in TTL of all transplants engrafted (N=29) are observed. Each data point represents the mean TTL in weeks for every individual leukemia transplanted (n=2 recipients). Some samples lead to disease within a short period of time (TTL^{short}, n=6, mean 6.8 weeks) in contrast to samples showing leukemia manifestation after a longer time period (TTL^{long}, n=23, mean 19.0 weeks; T- test P< .001). (See also: Supp. Fig. SF 1 and Supp. Tab. ST 1).

C. TTL is independent of cell numbers transplanted. NOD/SCID/huALL (N=4) with either a TTL^{long} (25 or 18 weeks) or TTL^{short} (9 weeks) phenotype were re-transplanted onto subsequent recipients injecting $1x10^5$, $1x10^6$, $1x10^7$ or $3x10^7$ cells per animal (n=2 per group) showing a stable TTL phenotype independent of the amount of cells transplanted. (See also Supp. Tab. ST 2).

Figure 2. Superior relapse free survival of TTL^{long} patients



A. Relapse free survival (RFS) of patients with TTL^{long} or TTL^{short} phenotypes (for A-F P values by log-rank test). Superior RFS of TTL^{long} patients (all RG; N=50; TTL^{long} n=44; TTL^{short} n=6; P= 2.9 e-10);

B. but also in the group of non-high risk (non-HR) stratified patients only (N=43; TTL^{long} n=39, TTL^{short} n=4; P= 3.0 e-13).

C. Additionally, the non-engraftment subgroup (no leukemia manifestation within 20 weeks post-transplant) was considered separately and compared to TTL^{long} and TTL^{short} resulting in an even more marked RFS of patients with non-engraftment in the whole risk group (N=50; non-engrafter n=21; TTL^{long} n=23; TTL^{short} n=6; P= 1.5 e-9).

D. In the non-HR patient group even 100% RFS was observed (N=43; non-engrafter n=18; TTL^{long} n=21; TTL^{short} n=4; P= 1.5 e-12).

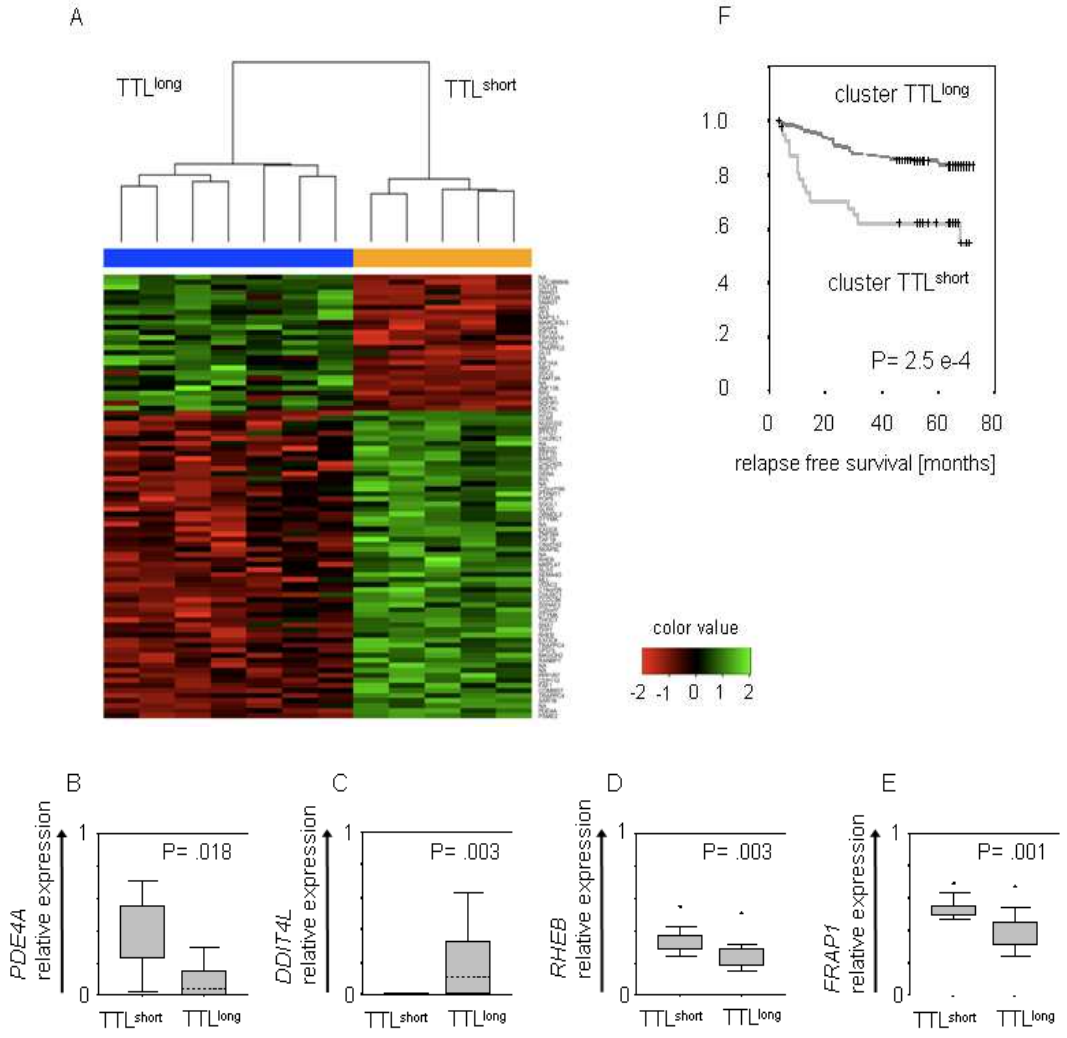
E. Analysis of the group of patients whose samples engrafted. Superior relapse free survival for TTL^{long} patients (N=29; TTL^{long}, n=23; TTL^{short}, n=6; P= 1.8 e-6) and

F. non-HR patients (N=25; TTL^{long}, n=21; TTL^{short}, n=4; P= 7.9 e-8).

G. TTL correlates to remission duration in ALL patients. Spearman correlation of TTL (mouse) to time from diagnosis to relapse (remission duration) of relapsed patients; all risk groups (N=9, rho= .731, P= .03) and

H. non-HR patients (N=7, rho=782, P= .04). (See also Supp. Fig. SF 2).

Figure 3. The TTL^{short} signature is associated with poor outcome

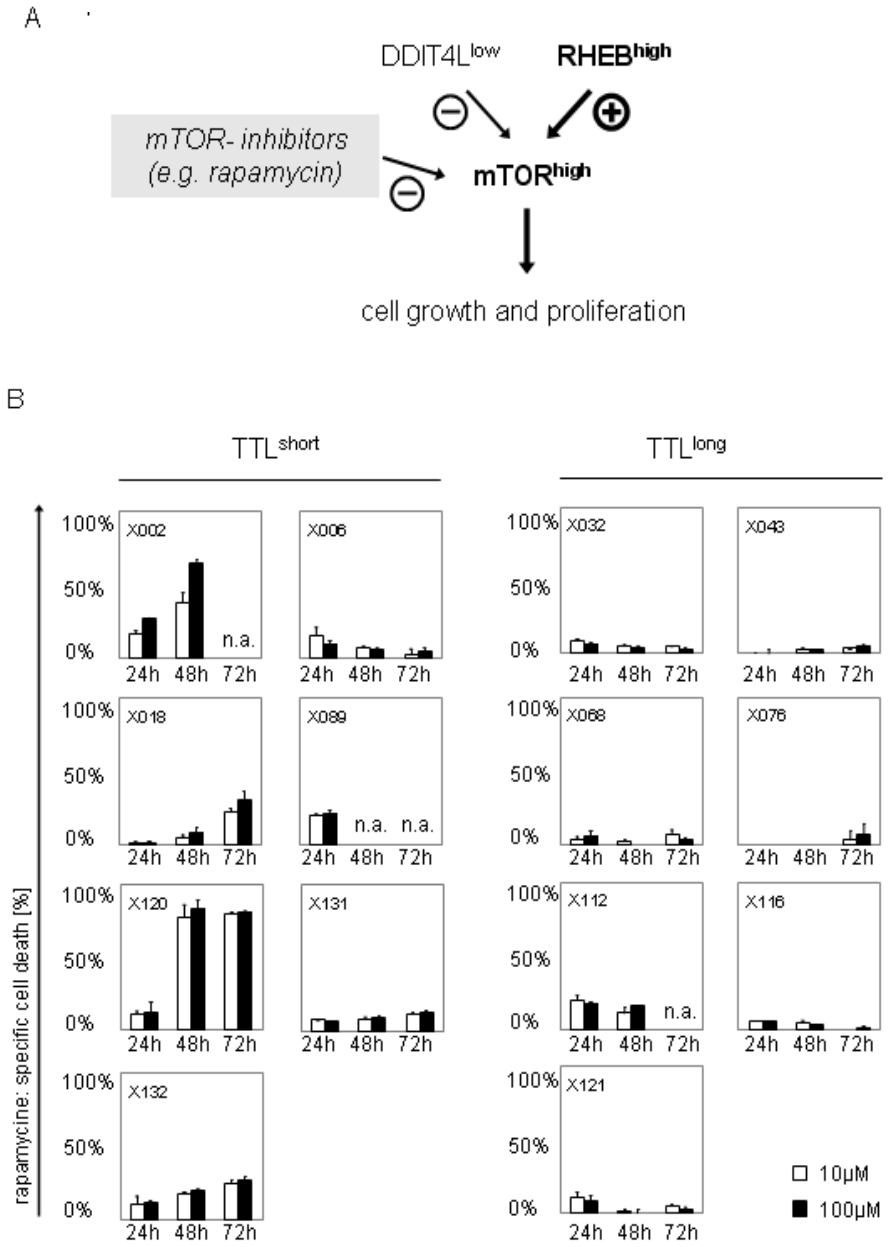


A. Signature of 88 differentially regulated probe sets (73 genes; “shrinkage t” statistic, FDR < 5%) in xenograft leukemia. Unsupervised cluster analysis of xenograft samples according to the TTL signature showing two clusters of TTL^{long} and TTL^{short} xenograft leukemia samples.

B to E: Differential expression of genes identified in the TTL signature in xenograft ALL samples (N=24, TTL^{short} n=9, TTL^{long} n=15) and additionally of *FRAP1* (coding for mTOR). Wilcoxon rank sum test, B: *PDE4A*: P= .018; C: *DDIT4L*: P= .003; D: *RHEB*: P= .003.; and E: *FRAP1*: P= .001.

F. Significant inferior relapse free survival of patients of an independent BCP ALL patient cohort (N= 197) who cluster with the TTL^{short} profile. Clustering into two groups according to the TTL signature as determined by cluster number estimation analysis (See also Supp. Fig. SF 3 and Supp. Tab. ST 4, 5 and 6).

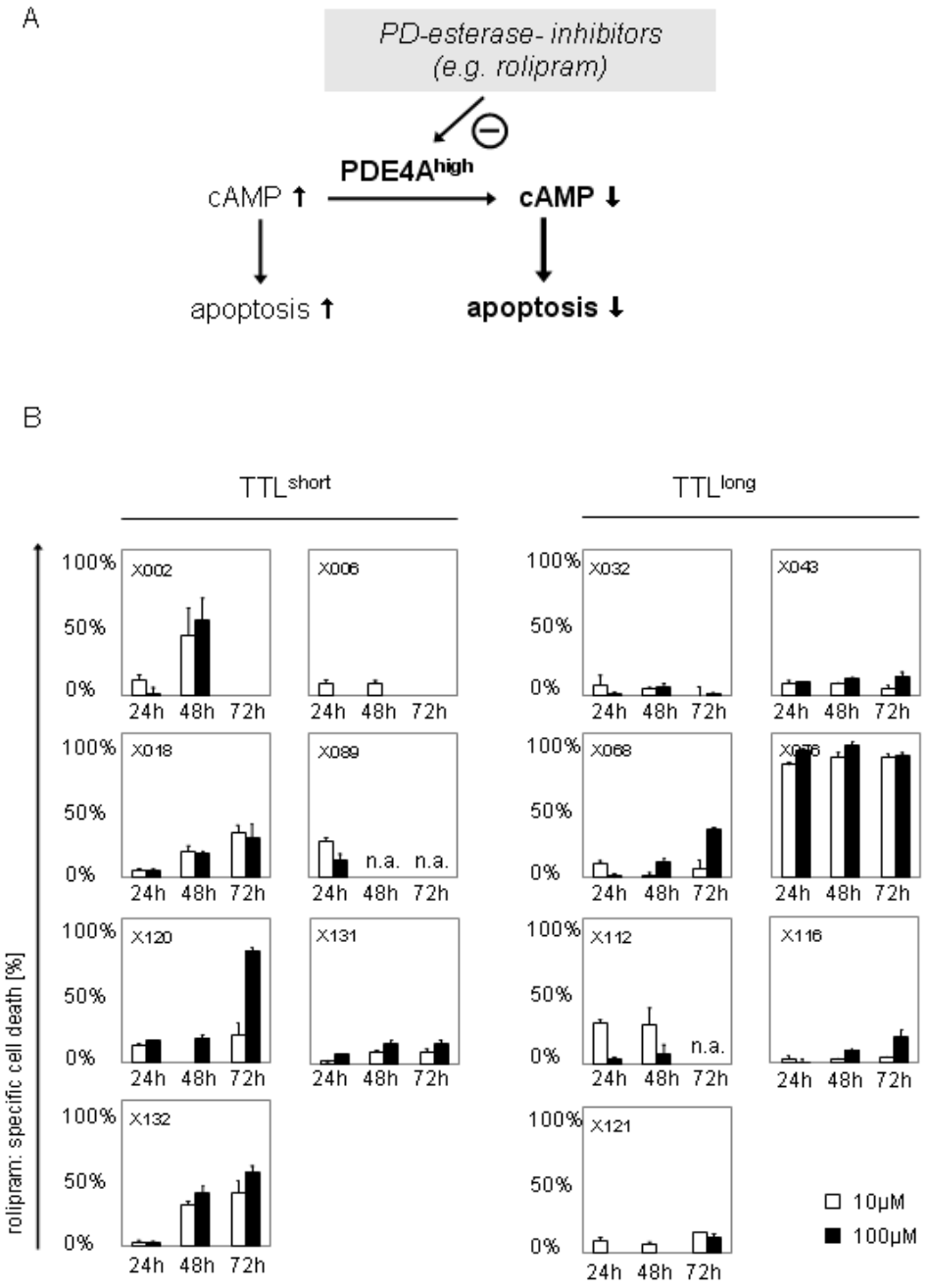
Figure 4. mTOR as potential therapeutic target for TTL^{short}/early relapse leukemia



A. TTL^{short} is characterized by an activated mTOR pathway (*DDIT4L*^{low}/*RHEB*^{high}/*FRAP1*^{high}) serving as potential target for directed therapy of TTL^{short}/early relapse leukemia.

B. mTOR inhibition increases cell death predominantly in TTL^{short} leukemia. Specific cell death of xenograft ALL upon *ex vivo* treatment with rapamycin (10 nM, white; 100 nM, black columns) for 24, 48 or 72 hours. Time points with >90% spontaneous cell death were not analyzed (n.a.). Data are given as mean of triplicates with standard deviation.

Figure 5. Inhibition of PDE4A induces cell death predominantly in TTL^{short}/early relapse leukemia



A. Predominant effect of phosphodiesterase inhibitors on TTL^{short}/early relapse leukemia associated with high PDE4A.

B. PDE4A inhibition increases cell death predominantly in TTL^{short} leukemia. Specific cell death of xenograft leukemia samples *ex vivo* incubated with rolipram (10 µM, white; 100 µM, black columns) for 24, 48 or 72 hours. Time points with >90% spontaneous cell death were not analyzed (n.a.). Data are given as mean of triplicates with standard deviation.

TABLES

Table 1. Prognostic relevance of TTL

N=50	n	P	risk ratio (relapse)	CI
TTL^{short}	6	6.25 e-5	45.08	6.98 – 290.94
PPR	6	.05	5.61	.99 – 31.70
unfavorable age	18	.11	3.34	.76 – 14.73
hyperleukocytosis	16	.51	1.68	.36 – 7.86

Multivariate analysis on relapse free survival (Cox regression, N=50) including different risk factors: TTL^{short} phenotype, prednisone poor response (PPR), hyperleukocytosis (leukocyte count at diagnosis higher than 50 [1000/µl]), and unfavorable age (0-1 and > 9 years); CI: confidence interval. (See also Supp. Tab. ST 3).

Table 2. Patient characteristics of de novo BCP ALL samples transplanted.

Table 2 A. categorical variables

		total		TTL^{long}		TTL^{short}	
		N	*%*	N	*%*	N	*%*
total		50	*100*	44	*100*	6	*100*
gender	female	26	*52*	23	*52*	3	*50*
	male	24	*48*	21	*48*	3	*50*
prednisone response (PR)	good	44	*88*	39	*89*	5	*83*
	poor	6	*12*	5	*11*	1	*17*
immunophenotype	pro-B ALL	1	*2*	0	-	1	*17*
	c-ALL	36	*72*	33	*75*	3	*50*
	pre-B ALL	13	*26*	11	*25*	2	*33*
fusion gene	TEL/AML1	19	*38*	19	*44*	0	-
	BCR/ABL	0	-	0	-	0	-
	MLL/AF4	1	*2*	0	-	1	*17*
	MLL/ENL	2	*4*	1	*2*	1	*17*

	no	26	*52*	22	*50*	4	*66*
	not done	2	*4*	2	*4*	0	-
minimal residual disease (MRD)	MRD-HR	1	*2*	1	*2*	0	-
	MRD-non-HR	41	*82*	36	*82*	5	*83*
	no MRD available	8	*16*	7	*16*	1	*17*
final risk groups	non-HR	43	*86*	39	*89*	4	*67*
	HR	7	*14*	5	*11*	2	*33*

Table 2 B. continuous variables

	total			TTL^{long}			TTL^{short}			T-test
	N	range	mean	N	range	mean	N	range	mean	P
age	50	0-19.3	6.7	44	0.3-19.3	6.8	6	0-10.7	6.3	.815
WBC at diagnosis [1000/µl]	50	1.7-272	47.0	44	1.7-272	40.5	6	7.7-264	95.0	.224
PB blast cells at diagnosis [%]	47	0-96	53.3	42	0-96	53.1	5	0-94	55.3	.908
WBC at day 8 [1000/µl]	50	0.5-31.7	37.2	44	0.5-31.7	38.5	6	1-7.5	28.3	.427
PB blast cells at day 8 [%]	49	0-59.5	8.0	43	0-59.5	8.3	6	0.5-24	6.0	.600
BM blast cells at day 33 [%]	47	-	0	43	-	0	4	-	0	-

Table 3. Gene signature associated with the TTL phenotype

up-regulated in TTL^{short}			down-regulated in TTL^{short}		
probe set	gene symbol	FC	probe set	gene symbol	FC
204735_at	**PDE4A**	**3,480**	227798_at	SMAD1	-5,526
226736_at	CHURC1	2,072	210993_s_at	SMAD1	-5,441
229814_at	NA	1,821	**203139_at**	**DAPK1**	**-4,497**
236443_at	NA	1,638	**228057_at**	**DDIT4L**	**-3,809**
222556_at	ALG5	1,345	217800_s_at	NDFIP1	-3,409
226643_s_at	NUDCD2	1,301	202973_x_at	FAM13A	-2,754
213409_s_at	**RHEB**	**1,247**	227376_at	GLI3	-2,427
206508_at	CD70	1,194	217047_s_at	FAM13A	-2,216
208844_at	VDAC3	1,188	206142_at	ZNF135	-1,798
244110_at	MLL	1,110	1559005_s_at	CNTLN	-1,781
223479_s_at	CHCHD5	1,070	212158_at	SDC2	-1,605
221995_s_at	MRP63	1,050	200998_s_at	CKAP4	-1,361
214395_x_at	EEF1D	1,046	213955_at	MYOZ3	-1,313
229525_at	THOC7	1,023	224151_s_at	AK3	-1,208
217426_at	NA	1,017	200644_at	MARCKSL1	-1,128
233924_s_at	EXOC6	1,009	201016_at	EIF1AX	-1,062
227364_at	NA	0,995	205191_at	RP2	-1,039
1553690_at	SGOL1	0,993	201019_s_at	EIF1AP1	-1,038
223210_at	CHURC1	0,975	201017_at	EIF1AX	-0,986
235878_at	TAF1B	0,878	219351_at	TRAPPC2	-0,938
204084_s_at	CLN5	0,854	216929_x_at	ABO	-0,804
227352_at	C19orf39	0,851	208655_at	NA	-0,786
216554_s_at	NA	0,846	214678_x_at	ZFX	-0,783
226259_at	EXOC6	0,843	221002_s_at	TSPAN14	-0,753
209276_s_at	GLRX	0,824	213864_s_at	NAP1L1	-0,681
203270_at	DTYMK	0,823	1564129_a_at	NA	-0,650
203554_x_at	PTTG1	0,821	1556102_x_at	LOC389906	-0,598
233380_s_at	RUFY1	0,821	-	-	-
235063_at	C20orf196	0,819	-	-	-
218064_s_at	AKAP8L	0,795	-	-	-
221598_s_at	MED27	0,785	-	-	-
219324_at	NOL12	0,777	-	-	-
223481_s_at	MRPL47	0,759	-	-	-
218556_at	ORMDL2	0,756	-	-	-
225901_at	PTPMT1	0,746	-	-	-
226070_at	C9orf142	0,726	-	-	-
204839_at	POP5	0,723	-	-	-
205345_at	BARD1	0,708	-	-	-
218102_at	DERA	0,702	-	-	-
224815_at	COMMD7	0,673	-	-	-
207877_s_at	NVL	0,662	-	-	-
203528_at	SEMA4D	0,659	-	-	-
205642_at	CEP110	0,653	-	-	-
1554482_a_at	SAR1B	0,652	-	-	-
217958_at	TRAPPC4	0,644	-	-	-
1553984_s_at	DTYMK	0,639	-	-	-
222894_x_at	C20orf7	0,638	-	-	-
212491_s_at	DNAJC8	0,635	-	-	-
218996_at	TFPT	0,627	-	-	-
1553957_at	ZNF564	0,627	-	-	-
223048_at	SDHAF2	0,620	-	-	-
217959_s_at	TRAPPC4	0,620	-	-	-

201214_s_at	PPP1R7	0,618	-	-	-
201452_at	**RHEB**	**0,599**	-	-	-
209103_s_at	UFD1L	0,597	-	-	-
204335_at	CCDC94	0,589	-	-	-
217692_at	MAGOH2	0,579	-	-	-
201716_at	SNX1	0,570	-	-	-
218080_x_at	FAF1	0,555	-	-	-
201762_s_at	PSME2	0,520	-	-	-
221915_s_at	RANBP1	0,519	-	-	-

Differentially regulated genes comparing TTL^{short} and TTL^{long} xenograft expression profiles (88 probe sets, 73 genes) identified by "shrinkage t" statistic, FDR < 5%. Fold change (FC) is given as logarithm (base 2), probe sets up-regulated in TTL^{short} (n=61, 51 genes) display a positive, down-regulated probe sets (n=27, 22 genes) a negative FC value.

REFERENCES

1. Schrappe, M., Reiter, A., Zimmermann, M., Harbott, J., Ludwig, W., Henze, G., Gadner, H., Odenwald, E., and Riehm, H. (2000). Long-term results of four consecutive trials in childhood ALL performed by the ALL-BFM study group from 1981 to 1995. Berlin-Frankfurt-Münster. Leukemia *14*, 2205-2222.
2. Schrappe, M., Reiter, A., Ludwig, W.D., Harbott, J., Zimmermann, M., Hiddemann, W., Niemeyer, C., Henze, G., Feldges, A., Zintl, F., Kornhuber, B., Ritter, J., Welte, K., Gadner, H., and Riehm, H. (2000). Improved outcome in childhood acute lymphoblastic leukemia despite reduced use of anthracyclines and cranial radiotherapy: results of trial ALL-BFM 90. German-Austrian-Swiss ALL-BFM Study Group. Blood *95*, 3310-3322.
3. Pui, C.-H., and Evans, W.-E. (2006). Treatment of Acute Lymphoblastic Leukemia. N Engl J Med *354*, 166-178.
4. Gaynon, P., Qu, R.P., Chappell, R.J., Willoughby, M.L., Tubergen, D.G., Steinherz, P.G., and Trigg, M.E. (1998). Survival after relapse in childhood acute lymphoblastic leukemia: impact of site and time to first relapse--the Children's Cancer Group Experience. Cancer *82*, 1387-1395.
5. Henze, G., Fengler, R., Hartmann, R., Kornhuber, B., Janka-Schaub, G., Niethammer, D., and Riehm, H. (1991). Six-year experience with a comprehensive approach to the treatment of recurrent childhood acute lymphoblastic leukemia. Blood *78*, 1166-1172.
6. Lawson, S.E., Harrison, G., Richards, S., Oakhill, A., Stevens, R., Eden, O.B., and Darbyshire, P.J. (2000). The UK experience in treating relapsed childhood acute lymphoblastic leukaemia: a report on the medical research council UKALLR1 study. Br J Haematol *108*, 531-543.
7. Pui, C.H., Relling, M.V., and Downing, J.R. (2004). Acute lymphoblastic leukemia. N Engl J Med *350*, 1535-1548.
8. Silverman, L.B., and Sallan, S.E. (2003). Newly diagnosed childhood acute lymphoblastic leukemia: update on prognostic factors and treatment. Curr Opin Hematol. *10*, 290-296.
9. Riehm, H., Reiter, A., Schrappe, M., Berthold, F., Dopfer, R., Gerein, V., Ludwig, R., Ritter, J., Stollmann, B., and Henze, G. (1987). Corticosteroid-dependent reduction of leukocyte count in blood as a prognostic factor in acute lymphoblastic leukemia in childhood (therapy study ALL-BFM 83). Klin Padiatr. *199*, 151-160.
10. Schrappe, M., Aricò, M., Harbott, J., Biondi, A., Zimmermann, M., Conter, V., Reiter, A., Valsecchi, M.G., Gadner, H., Basso, G., Bartram, C.R., Lampert, F., Riehm, H., and Masera, G. (1998). Philadelphia chromosome-positive (Ph+) childhood acute lymphoblastic leukemia: good initial steroid response allows early prediction of a favorable treatment outcome. Blood *92*, 2730-2741.
11. Gajjar, A., Ribeiro, R., Hancock, M.L., Rivera, G.K., Mahmoud, H., Sandlund, J.T., Crist, W.M., and Pui, C.H. (1995). Persistence of circulating blasts after 1 week of multiagent chemotherapy confers a poor prognosis in childhood acute lymphoblastic leukemia. Blood *86*, 1292-1295.
12. Chessells, J.M., Bailey, C., and Richards, S.M. (1995). Intensification of treatment and survival in all children with lymphoblastic leukaemia: results of UK Medical Research Council trial UKALL X. Medical Research Council Working Party on Childhood Leukaemia. Lancet *345*, 143-148.
13. Steinherz, P.G., Gaynon, P.S., Breneman, J.C., Cherlow, J.M., Grossman, N.J., Kersey, J.H., Johnstone, H.S., Sather, H.N., Trigg, M.E., Chappell, R., Hammond, D., and Bleyer, W.A. (1996). Cytoreduction and prognosis in acute lymphoblastic leukemia--the importance of early marrow response: report from the Childrens Cancer Group. J Clin Oncol. *14*, 389-398.
14. Flohr, T., Schrauder, A., Cazzaniga, G., Panzer-Grumayer, R., van der Velden, V., Fischer, S., Stanulla, M., Basso, G., Niggli, F.K., Schafer, B.W., Sutton, R., Koehler, R., Zimmermann, M., Valsecchi, M.G., Gadner, H., Masera, G., Schrappe, M., van Dongen, J.J., Biondi, A., and Bartram, C.R. (2008). Minimal residual disease-directed risk stratification using real-time quantitative PCR analysis of immunoglobulin and T-cell receptor gene rearrangements in the international multicenter trial AIEOP-BFM ALL 2000 for childhood acute lymphoblastic leukemia. Leukemia *22*, 771-782.
15. van Dongen, J.J., Seriu, T., Panzer-Grumayer, E.R., Biondi, A., Pongers-Willemse, M.J., Corral, L., Stolz, F., Schrappe, M., Masera, G., Kamps, W.A., Gadner, H., van Wering, E.R., Ludwig, W.D., Basso, G., de Bruijn, M.A., Cazzaniga, G., Hettinger, K., van der Does-van den Berg, A., Hop, W.C., Riehm, H., and Bartram, C.R. (1998). Prognostic value of minimal residual disease in acute lymphoblastic leukaemia in childhood. Lancet *352*, 1731-1738.
16. Conter, V., Bartram, C.R., Valsecchi, M.G., Schrauder, A., Panzer-Grumayer, R., Moricke, A., Arico, M., Zimmermann, M., Mann, G., De Rossi, G., Stanulla, M., Locatelli, F., Basso, G., Niggli, F., Barisone, E., Henze, G., Ludwig, W.D., Haas, O.A., Cazzaniga, G., Koehler, R., Silvestri, D., Bradtke, J., Parasole, R., Beier, R., van Dongen, J.J., Biondi, A., and Schrappe, M. (2010). Molecular response to treatment redefines all prognostic factors in children and adolescents with B-cell precursor acute

lymphoblastic leukemia: results in 3184 patients of the AIEOP-BFM ALL 2000 study. Blood *115*, 3206-3214.

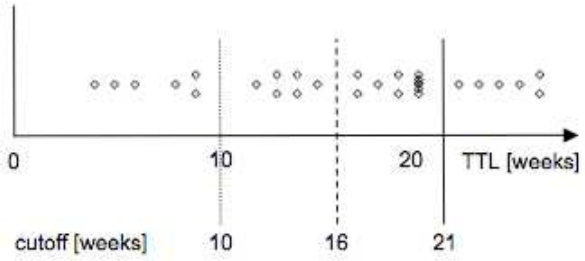
17. Lock, R.B., Liem, N., Farnsworth, M.L., Milross, C.G., Xue, C., Tajbakhsh, M., Haber, M., Norris, M.D., Marshall, G.M., and Rice, A.M. (2002). The nonobese diabetic/severe combined immunodeficient (NOD/SCID) mouse model of childhood acute lymphoblastic leukemia reveals intrinsic differences in biologic characteristics at diagnosis and relapse. Blood *99*, 4100-4108.
18. Baersch, G., Möllers, T., Hötte, A., Dockhorn-Dworniczak, B., Rübe, C., Ritter, J., Jürgens, H., and Vormoor, J. (1997). Good engraftment of B-cell precursor ALL in NOD-SCID mice. Klin Padiatr. *209*, 178-185.
19. Borgmann, A., Baldy, C., von Stackelberg, A., Beyermann, B., Fichtner, I., Nurnberg, P., and Henze, G. (2000). Childhood ALL blasts retain phenotypic and genotypic characteristics upon long-term serial passage in NOD/SCID mice. Pediatr Hematol Oncol *17*, 635-650.
20. Nijmeijer, B.A., Mollevanger, P., van Zelderen-Bhola, S.L., Kluin-Nelemans, H.C., Willemze, R., and Falkenburg, J.H. (2001). Monitoring of engraftment and progression of acute lymphoblastic leukemia in individual NOD/SCID mice. Exp Hematol *29*, 322-329.
21. Pui, C.H., Sandlund, J.T., Pei, D., Campana, D., Rivera, G.K., Ribeiro, R.C., Rubnitz, J.E., Razzouk, B.I., Howard, S.C., Hudson, M.M., Cheng, C., Kun, L.E., Raimondi, S.C., Behm, F.G., Downing, J.R., Relling, M.V., and Evans, W.E. (2004). Improved outcome for children with acute lymphoblastic leukemia: results of Total Therapy Study XIIIB at St Jude Children's Research Hospital. Blood *104*, 2690-2696.
22. Yeoh, E.J., Ross, M.E., Shurtleff, S.A., Williams, W.K., Patel, D., Mahfouz, R., Behm, F.G., Raimondi, S.C., Relling, M.V., Patel, A., Cheng, C., Campana, D., Wilkins, D., Zhou, X., Li, J., Liu, H., Pui, C.H., Evans, W.E., Naeve, C., Wong, L., and Downing, J.R. (2002). Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. Cancer Cell *1*, 133-143.
23. Opgen-Rhein, R., and Strimmer, K. (2007). Accurate ranking of differentially expressed genes by a distribution-free shrinkage approach. Stat Appl Genet Mol Biol *6*, Article 9.
24. Kraus, J.M., and Kestler, H.A. (2010). A highly efficient multi-core algorithm for clustering extremely large datasets. BMC Bioinformatics *11*, 169.
25. Marchand, M., and Shawe-Taylor, J. (2002). The set covering machine. Journal of Machine Learning Research *3*, 723-746.
26. Kestler, H.A., Lindner, W., and Müller, A. (2006). Learning and feature selection using the set covering machine with data-dependent rays on gene expression profiles. In Schwenker F., Marinai S. (eds.), Artificial Neural Networks in Pattern Recognition *vol. LNAI 4087*, 286-297. Springer.
27. Kamel-Reid, S., Letarte, M., Sirard, C., Doedens, M., Grunberger, T., Fulop, G., Freedman, M.H., Phillips, R.A., and Dick, J.E. (1989). A model of human acute lymphoblastic leukemia in immune-deficient SCID mice. Science. *1989 Dec 22*, 1597-1600.
28. Kamel-Reid, S., Letarte, M., Doedens, M., Greaves, A., Murdoch, B., Grunberger, T., Lapidot, T., Thorner, P., Freedman, M.H., Phillips, R.A., and et al. (1991). Bone marrow from children in relapse with pre-B acute lymphoblastic leukemia proliferates and disseminates rapidly in scid mice. Blood *78*, 2973-2981.
29. Uckun, F.M., Sather, H., Reaman, G., Shuster, J., Land, V., Trigg, M., Gunther, R., Chelstrom, L., Bleyer, A., Gaynon, P., and et al. (1995). Leukemic cell growth in SCID mice as a predictor of relapse in high-risk B-lineage acute lymphoblastic leukemia. Blood *85*, 873-878.
30. Uckun, F.M., Sather, H.N., Waurzyniak, B.J., Sensel, M.G., Chelstrom, L., Ek, O., Sarquis, M.B., Nachman, J., Bostrom, B., Reaman, G.H., and Gaynon, P.S. (1998). Prognostic significance of B-lineage leukemic cell growth in SCID mice: a Children's Cancer Group Study. Leuk Lymphoma *30*, 503-514.
31. Monaco, G., Konopleva, M., Munsell, M., Leysath, C., Wang, R.Y., Jackson, C.E., Korbling, M., Estey, E., Belmont, J., and Andreeff, M. (2004). Engraftment of acute myeloid leukemia in NOD/SCID mice is independent of CXCR4 and predicts poor patient survival. Stem Cells *22*, 188-201.
32. Pearce, D.J., Taussig, D., Zibara, K., Smith, L.L., Ridler, C.M., Preudhomme, C., Young, B.D., Rohatiner, A.Z., Lister, T.A., and Bonnet, D. (2006). AML engraftment in the NOD/SCID assay reflects the outcome of AML: implications for our understanding of the heterogeneity of AML. Blood *107*, 1166-1173.
33. Bhojwani, D., Kang, H., Moskowitz, N.P., Min, D.J., Lee, H., Potter, J.W., Davidson, G., Willman, C.L., Borowitz, M.J., Belitskaya-Levy, I., Hunger, S.P., Raetz, E.A., and Carroll, W.L. (2006). Biologic pathways associated with relapse in childhood acute lymphoblastic leukemia: a Children's Oncology Group study. Blood *108*, 711-717.

34. Flotho, C., Coustan-Smith, E., Pei, D., Cheng, C., Song, G., Pui, C.H., Downing, J.R., and Campana, D. (2007). A set of genes that regulate cell proliferation predicts treatment outcome in childhood acute lymphoblastic leukemia. Blood *110*, 1271-1277.
35. Lugthart, S., Cheok, M.H., den Boer, M.L., Yang, W., Holleman, A., Cheng, C., Pui, C.H., Relling, M.V., Janka-Schaub, G.E., Pieters, R., and Evans, W.E. (2005). Identification of genes associated with chemotherapy crossresistance and treatment response in childhood acute lymphoblastic leukemia. Cancer Cell *7*, 375-386.
36. Cario, G., Stanulla, M., Fine, B.M., Teuffel, O., Neuhoff, N.V., Schrauder, A., Flohr, T., Schafer, B.W., Bartram, C.R., Welte, K., Schlegelberger, B., and Schrappe, M. (2005). Distinct gene expression profiles determine molecular treatment response in childhood acute lymphoblastic leukemia. Blood *105*, 821-826.
37. Holleman, A., Cheok, M.H., den Boer, M.L., Yang, W., Veerman, A.J., Kazemier, K.M., Pei, D., Cheng, C., Pui, C.H., Relling, M.V., Janka-Schaub, G.E., Pieters, R., and Evans, W.E. (2004). Gene-expression patterns in drug-resistant acute lymphoblastic leukemia cells and response to treatment. N Engl J Med *351*, 533-542.
38. Brown, E.J., Albers, M.W., Shin, T.B., Ichikawa, K., Keith, C.T., Lane, W.S., and Schreiber, S.L. (1994). A mammalian protein targeted by G1-arresting rapamycin-receptor complex. Nature *369*, 756-758.
39. Corradetti, M.N., Inoki, K., and Guan, K.L. (2005). The stress-inducted proteins RTP801 and RTP801L are negative regulators of the mammalian target of rapamycin pathway. J Biol Chem *280*, 9769-9772.
40. Long, X., Lin, Y., Ortiz-Vega, S., Yonezawa, K., and Avruch, J. (2005). Rheb binds and regulates the mTOR kinase. Curr Biol *15*, 702-713.
41. Brown, V.I., Fang, J., Alcorn, K., Barr, R., Kim, J.M., Wasserman, R., and Grupp, S.A. (2003). Rapamycin is active against B-precursor leukemia in vitro and in vivo, an effect that is modulated by IL-7-mediated signaling. Proc Natl Acad Sci U S A *100*, 15113-15118.
42. Teachey, D.T., Obzut, D.A., Cooperman, J., Fang, J., Carroll, M., Choi, J.K., Houghton, P.J., Brown, V.I., and Grupp, S.A. (2006). The mTOR inhibitor CCI-779 induces apoptosis and inhibits growth in preclinical models of primary adult human ALL. Blood *107*, 1149-1155.
43. Crazzolara, R., Cisterne, A., Thien, M., Hewson, J., Baraz, R., Bradstock, K.F., and Bendall, L.J. (2009). Potentiating effects of RAD001 (Everolimus) on vincristine therapy in childhood acute lymphoblastic leukemia. Blood *113*, 3297-3306.
44. Teachey, D.T., Sheen, C., Hall, J., Ryan, T., Brown, V.I., Fish, J., Reid, G.S., Seif, A.E., Norris, R., Chang, Y.J., Carroll, M., and Grupp, S.A. (2008). mTOR inhibitors are synergistic with methotrexate: an effective combination to treat acute lymphoblastic leukemia. Blood *112*, 2020-2023.
45. Wei, G., Twomey, D., Lamb, J., Schlis, K., Agarwal, J., Stam, R.W., Opferman, J.T., Sallan, S.E., den Boer, M.L., Pieters, R., Golub, T.R., and Armstrong, S.A. (2006). Gene expression-based chemical genomics identifies rapamycin as a modulator of MCL1 and glucocorticoid resistance. Cancer Cell *10*, 331-342.
46. Jiang, X., Li, J., Paskind, M., and Epstein, P.M. (1996). Inhibition of calmodulin-dependent phosphodiesterase induces apoptosis in human leukemic cells. Proc Natl Acad Sci U S A *93*, 11236-11241.
47. Ogawa, R., Streiff, M.B., Bugayenko, A., and Kato, G.J. (2002). Inhibition of PDE4 phosphodiesterase activity induces growth suppression, apoptosis, glucocorticoid sensitivity, p53, and p21(WAF1/CIP1) proteins in human acute lymphoblastic leukemia cells. Blood *99*, 3390-3397.
48. Moon, E.Y., and Lerner, A. (2003). PDE4 inhibitors activate a mitochondrial apoptotic pathway in chronic lymphocytic leukemia cells that is regulated by protein phosphatase 2A. Blood *101*, 4122-4130.
49. Meyers, J.A., Taverna, J., Chaves, J., Makkinje, A., and Lerner, A. (2007). Phosphodiesterase 4 inhibitors augment levels of glucocorticoid receptor in B cell chronic lymphocytic leukemia but not in normal circulating hematopoietic cells. Clin Cancer Res *13*, 4920-4927.
50. Raval, A., Tanner, S.M., Byrd, J.C., Angerman, E.B., Perko, J.D., Chen, S.S., Hackanson, B., Grever, M.R., Lucas, D.M., Matkovic, J.J., Lin, T.S., Kipps, T.J., Murray, F., Weisenburger, D., Sanger, W., Lynch, J., Watson, P., Jansen, M., Yoshinaga, Y., Rosenquist, R., de Jong, P.J., Coggill, P., Beck, S., Lynch, H., de la Chapelle, A., and Plass, C. (2007). Downregulation of death-associated protein kinase 1 (DAPK1) in chronic lymphocytic leukemia. Cell *129*, 879-890.
51. Kissil, J.L., Feinstein, E., Cohen, O., Jones, P.A., Tsai, Y.C., Knowles, M.A., Eydmann, M.E., and Kimchi, A. (1997). DAP-kinase loss of expression in various carcinoma and B-cell lymphoma cell lines: possible implications for role as tumor suppressor gene. Oncogene *15*, 403-407.
52. Katzenellenbogen, R.A., Baylin, S.B., and Herman, J.G. (1999). Hypermethylation of the DAP-kinase CpG island is a common alteration in B-cell malignancies. Blood *93*, 4347-4353.

53. Esteller, M., Sanchez-Cespedes, M., Rosell, R., Sidransky, D., Baylin, S.B., and Herman, J.G. (1999). Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. Cancer Res *59*, 67-70.
54. Sanchez-Cespedes, M., Esteller, M., Wu, L., Nawroz-Danish, H., Yoo, G.H., Koch, W.M., Jen, J., Herman, J.G., and Sidransky, D. (2000). Gene promoter hypermethylation in tumors and serum of head and neck cancer patients. Cancer Res *60*, 892-895.
55. Deiss, L.P., Feinstein, E., Berissi, H., Cohen, O., and Kimchi, A. (1995). Identification of a novel serine/threonine kinase and a novel 15-kD protein as potential mediators of the gamma interferon-induced cell death. Genes Dev *9*, 15-30.
56. Cohen, O., Inbal, B., Kissil, J.L., Raveh, T., Berissi, H., Spivak-Kroizaman, T., Feinstein, E., and Kimchi, A. (1999). DAP-kinase participates in TNF-alpha- and Fas-induced apoptosis and its function requires the death domain. J Cell Biol *146*, 141-148.
57. Inbal, B., Cohen, O., Polak-Charcon, S., Kopolovic, J., Vadai, E., Eisenbach, L., and Kimchi, A. (1997). DAP kinase links the control of apoptosis to metastasis. Nature *390*, 180-184.
58. Meyer, L.H., Karawajew, L., Schrappe, M., Ludwig, W.D., Debatin, K.M., and Stahnke, K. (2006). Cytochrome c-related caspase-3 activation determines treatment response and relapse in childhood precursor B-cell ALL. Blood *107*, 4524-4531.
59. Meyer, L.H., Queudeville, M., Eckhoff, S.M., Creutzig, U., Reinhardt, D., Karawajew, L., Ludwig, W.D., Stahnke, K., and Debatin, K.M. (2008). Intact apoptosis signaling in myeloid leukemia cells determines treatment outcome in childhood AML. Blood *111*, 2899-2903.
60. Holleman, A., den Boer, M.L., de Menezes, R.X., Cheok, M.H., Cheng, C., Kazemier, K.M., Janka-Schaub, G.E., Gobel, U., Graubner, U.B., Evans, W.E., and Pieters, R. (2006). The expression of 70 apoptosis genes in relation to lineage, genetic subtype, cellular drug resistance, and outcome in childhood acute lymphoblastic leukemia. Blood *107*, 769-776.
61. Conter, V., Arico, M., Valsecchi, M.G., Basso, G., Biondi, A., Madon, E., Mandelli, F., Paolucci, G., Pession, A., Rizzari, C., Rondelli, R., Zanesco, L., and Masera, G. (2000). Long-term results of the Italian Association of Pediatric Hematology and Oncology (AIEOP) acute lymphoblastic leukemia studies, 1982-1995. Leukemia *14*, 2196-2204.
62. Gautier, L., Cope, L., Bolstad, B.M., and Irizarry, R.A. (2004). affy--analysis of Affymetrix GeneChip data at the probe level. Bioinformatics *20*, 307-315.
63. Strimmer, K. (2008). A unified approach to false discovery rate estimation. BMC Bioinformatics *9*, 303.
64. R-Development-Core-Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.
65. Friesen, C., Herr, I., Krammer, P.H., and Debatin, K.M. (1996). Involvement of the CD95 (APO-1/FAS) receptor/ligand system in drug-induced apoptosis in leukemia cells. Nat Med *2*, 574-577.

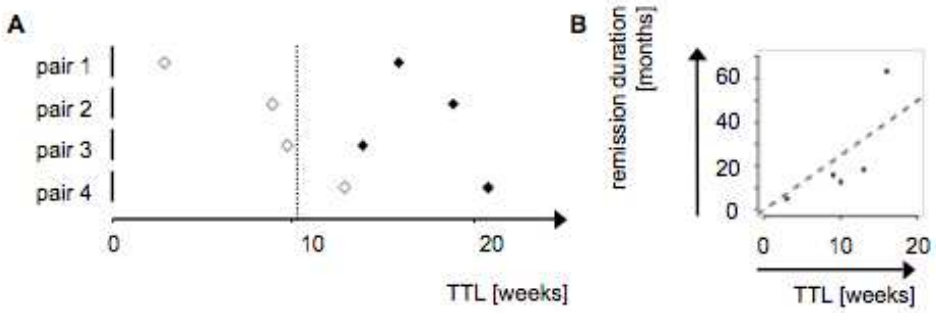
SUPPLEMENTARY FIGURES

SF 1 (referring to: Fig. 1B): Analysis of additional thresholds to discriminate TTL phenotypes



Alternative cutoff points to 10 weeks (dotted line) were analyzed: intervals without leukemia manifestation at 16 (dashed line) and 21 weeks (solid line) were additionally used to divide the cohort comparing time to overt leukemia between the resulting subgroups (Supplementary Table ST 1).

SF 2 (referring to: Fig. 2): TTL is confirmed in a independent subset of transplanted leukemia samples.

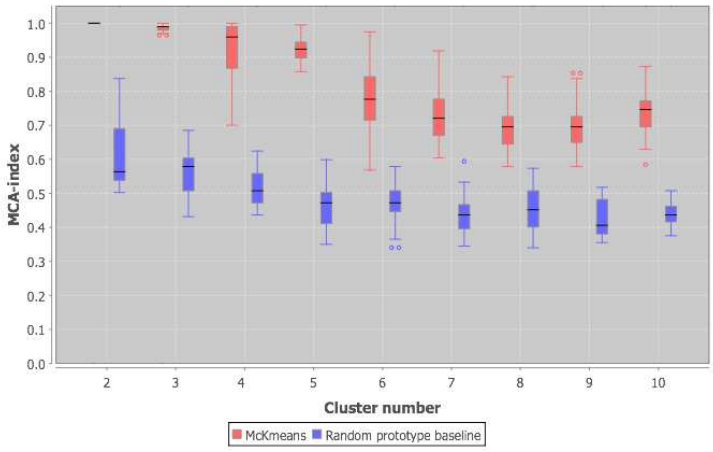


A. Four patient pairs (N=8 cell bank leukemia samples) matched for age, gender and risk groups but either relapsed early within 24 months after diagnosis (n=4, open rhombs) compared to late or no relapse (n=4, solid rhombs) were transplanted. Samples of early relapsed patients showed a significantly shorter TTL in contrast to late or non relapsing patients (Mann- Whitney U- test, P= .021).

B. Nonparametric correlation (Spearman's rho) of TTL and remission duration of relapsed patients. Cell bank samples, non-HR (N=5, rho= .9, P= .04).

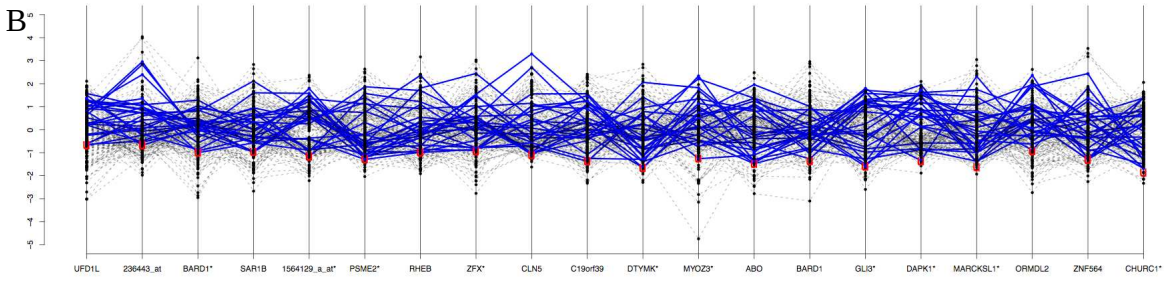
SF 3 (referring to: Fig. 3): Cluster number estimation analysis and generation of a classifier based on the TTL signature

A



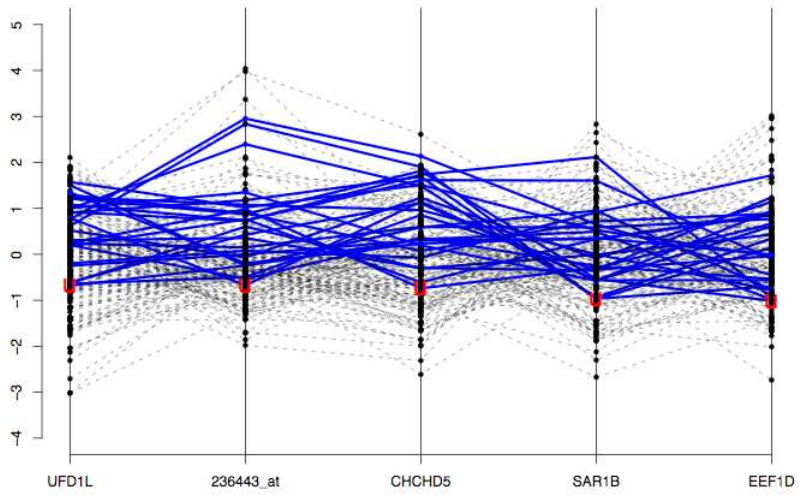
A. Cluster number estimation was performed based on the k-means cluster algorithm via repeated clusterings on resampled versions of the data using a parallelized implementation (McKmeans [1]). Clusterings are compared pairwise via the MCA index and corrected for random partitions [1]. The boxplots show the distribution of 4950 pairwise comparisons (100 clusterings) of two partitions for the cluster case (red) and for the random prototype case (blue) for an assumed number of groups between 2 and 10. According to the TTL signature (88 differentially regulated probe sets) two robust clusters were identified in the 197 patient cohort with this algorithm.

B



B. Based on the genes identified in the TTL signature (xenograft ALL samples) a classifier was generated using a set covering machine with data dependent rays. By conjunction of different threshold decisions based on subsets of genes of the TTL signature a decision rule identifying early relapse patients was obtained. The classification of 197 patients of the independent cohort according to threshold decisions is based on a subset of 20 genes of the TTL signature including *RHEB* and *DAPK1*. On the y-axis the decision threshold is depicted for each gene (x-axis). For each sample the decision is made if the corresponding expression levels are above each threshold (‘red u’). For axes with an asterisk the decision rule has to be multiplied by -1. Blue lines represent early relapse samples passing all threshold decisions, black dashed lines represent profiles of patient samples not meeting all 20 decision thresholds.

C



C. Further refinement of the classifier with increased cross-validation performance. Classification of 197 patients of the independent cohort according to threshold decisions based on a subset of 5 probe sets of the TTL signature (sensitivity: 100%, accuracy: 75%; cross-validation (leave-one-out) accuracy: 72%, sensitivity: 77%).

SUPPLEMENTARY TABLES

ST 1 (referring to: Fig. 1B): Discrimination of time to leukemia (TTL) into subgroups according to different classifiers.

distinction of times to overt leukemia by		N	mean TTL [weeks]	95% CI	incidence fraction	adjusted P value	rank
cutoff 10 weeks	TTL < 10 weeks	6	6.8	5.1 - 8.5	1	3.084063 e-10	1
	TTL > 10 weeks	23	19.0	17.3 - 20.7	1		
cutoff 16 weeks	TTL < 16 weeks	12	10.2	8.0 - 12.3	1	2.880554 e-9	2
	TTL > 16 weeks	17	20.9	19.6 - 22.3	1		
cutoff 21 weeks	TTL < 21 weeks	23	14.4	12.2 - 16.6	1	7.834243 e-5	3
	TTL > 21 weeks	6	24.3	23.0 - 25.6	1		
age	unfavorable	9	12.6	9.5 - 15.6	1	.003135248	4
	favorable	20	18.3	15.5 - 21.0	1		
ALL-BFM risk groups	high risk	4	10.5	4.7 - 16.3	1	.00977492	5
	non- high risk	25	17.4	15.1 - 19.7	1		
TEL/AML1 fusion *	absent	17	13.7	10.7 - 16.7	1	.02374496	6
	present	11	20.6	18.4 - 22.9	1		
prednisone response	poor	3	12.0	4.9 - 19.1	1	.07074063 #	7
	good	26	17.0	14.6 - 19.4	1		
hyperleuko cytosis	> 50 [1000/µl]	14	15.3	12.1 - 18.5	1	.1615348 #	8
	< 50 [1000/µl]	15	17.6	14.3 - 20.9	1		

* TEL/AML1 fusion was not available for n=1 patient; # not significant.

A. Different cutoff points and leukemia characteristics were used to group the cohort of engrafted leukemia samples (N=29). Time to overt leukemia was compared between the respective subgroups (e.g. cutoff 10 weeks: TTL < 10 weeks versus TTL > 10 weeks) by Kaplan Meier analysis and log rank test.

impact on patient outcome (RFS) according to		N	mean RFS [months]	95% CI	incidence fraction	adjusted P value	rank
cutoff 10 weeks	TTL < 10 weeks	6	12.1	5.1 – 19.1	.833		
	TTL > 10 weeks	44	35.4	31.5 – 39.3	.114	2.879047 e-10	1
cutoff 16 weeks	TTL < 16 weeks	12	16.3	9.8 – 22.8	.583		
	TTL > 16 weeks	38	37.8	34.0 – 41.5	.079	4.955158 e-7	2
cutoff 21 weeks	TTL < 21 weeks	23	27.1	20.6 – 33.6	.391		
	TTL > 21 weeks	27	37.3	32.8 – 41.9	.037	.001137039	3
TEL/AML1 fusion	absent	29	27.5	21.9 – 33.0	.310		
	present	19*	39.0	34.0 – 44.0	.052	.01967564	4
hyperleuko cytosis	> 50 [1000/µl]	16	29.0	20.7 – 37.4	.375		
	< 50 [1000/µl]	34	34.3	29.8 – 38.9	.118	.02825113	5
prednisone response	poor	6	25.6	9.9 – 41.2	.500		
	good	44	33.6	29.4 – 37.8	.160	.03264534	6
age	unfavorable	18	27.1	19.0 – 35.1	.333		
	favorable	32	35.6	31.5 – 40.0	.125	.03617227	7
ALL-BFM risk groups	high risk	7	24.6	11.3 – 37.9	.429		
	non- high risk	43	33.9	29.7 – 38.1	.163	.05821112 #	8

* TEL/AML1 fusion was not available for n=2 patients; # not significant.

B. Different cutoff points and leukemia characteristics were tested to separate the cohort of all leukemia samples analyzed (N=50) with respect to impact on patient outcome. Relapse free survival (patients) was compared between the respective subgroups (e.g. cutoff 10 weeks: TTL < 10 weeks versus TTL > 10 weeks) by Kaplan Meier analysis and log rank test.

ST 2 (referring to: Fig. 1C): Stability and consistency of the TTL-phenotype.

no.	subtype	TTL [weeks] recipient 1	TTL [weeks] recipient 2	**mean TTL [weeks]**
D1	pre-B	4.0	4.3	**4.1**
D2	pro-B MLL/AF4	6.1	6.1	**6.1**
D3	pro-B MLL/ENL	5.1	7.7	**6.4**
D4	pre-B	6.6	6.6	**6.6**
D5	pre-B	8.4	8.4	**8.4**
D6	pre-B	8.7	8.9	**8.8**
D7	pre-B	8.9	8.9	**8.9**
D8	pre-B	9.0	9.0	**9.0**
D9	c-ALL	11.9	11.9	**11.9**
D10	pre-B	12.3	12.3	**12.3**
D11	pre-B	12.6	15.1	**13.9**
D12	c-ALL TEL/AML1	15.1	16.4	**15.8**
D13	c-ALL	17.4	17.9	**17.6**
D14	pre-B	18.0	18.4	**18.2**
D15	c-ALL	19.0	19.1	**19.1**
D16	c-ALL	20.1	22.0	**21.1**
D17	c-ALL TEL/AML1	21.0	21.7	**21.4**
D18	c-ALL TEL/AML1	20.3	23.4	**21.9**
D19	c-ALL TEL/AML1	25.9	25.9	**25.9**
D20	c-ALL	24.4	27.6	**26.0**
D21	pre-B TEL/AML1	26.0	26.3	**26.1**

A TTL of primary patient samples obtained at diagnosis (D) transplanted in parallel onto two recipients ($1x10^7$ cells per transplant).

no.	subtype	TTL [weeks] recipient 1	TTL [weeks] recipient 2	**mean TTL [weeks]**
D1	pre-B	4.0	4.3	**4.1**
R D1	pre-B	3.6	3.7	**3.6**
D3	pro-B MLL/ENL	5.1	7.7	**6.4**
R D3	pro-B MLL/ENL	5.0	5.0	**5.0**

B. TTL of primary patient samples obtained at diagnosis (D) or relapse (R) transplanted in parallel onto two recipients ($1x10^7$ cells per transplant).

no.	subtype	TTL [weeks] recipient 1	TTL [weeks] recipient 2	**mean TTL [weeks]**
D3 P0	pro-B MLL/ENL	5.1	7.7	**6.4**
D3 P1	pro-B MLL/ENL	4.9	6.0	**5.4**
D4 P0	pre-B	6.6	6.6	**6.6**
D4 P1	pre-B	7.6	7.9	**7.7**
D4 P2	pre-B	5.9	5.9	**5.9**
D5 P0	pre-B	8.4	8.4	**8.4**
D5 P1	pre-B	6.1	7.9	**7.0**
D10 P0	pre-B	12.6	15.1	**13.9**
D10 P1	pre-B	15.6	18.6	**17.1**

C. Consistency of TTL phenotypes between diagnosis and following passages in the NOD/SCID/huALL model. TTL of primary patient samples obtained at diagnosis (primary recipients, P0) and of consecutive secondary (P1) or tertiary (P2) recipients. Each leukemia sample was transplanted in duplicates injecting $1x10^7$ cells per recipient.

no.	cell number injected	transplantation of cells isolated from	TTL [weeks]
TTL^{long} 1	**$1x10^7$**	**diagnosis**	**25 (= TTL^{long})**
	$1x10^5$	1^{rst} recipient (P0)	> 16
	$1x10^6$	1^{rst} recipient (P0)	> 16
	$1x10^7$	**1^{rst} recipient (P0)**	**> 16**
	$3x10^7$	1^{rst} recipient (P0)	> 16
TTL^{long} 2	**$1x10^7$**	**diagnosis**	**18 (= TTL^{long})**
	$1x10^5$	6^{th} recipient (P5)	> 16
	$1x10^6$	6^{th} recipient (P5)	> 16
	$1x10^7$	**6^{th} recipient (P5)**	**16**
	$3x10^7$	6^{th} recipient (P5)	16
TTL^{short} 1	**$1x10^7$**	**diagnosis**	**9 (= TTL^{short})**
	$1x10^5$	3^{rd} recipient (P2)	8
	$1x10^6$	3^{rd} recipient (P2)	8
	$1x10^7$	**3^{rd} recipient (P2)**	**7**
	$3x10^7$	3^{rd} recipient (P2)	6
TTL^{short} 2	**$1x10^7$**	**diagnosis**	**9 (= TTL^{short})**
	$1x10^5$	6^{th} recipient (P5)	7
	$1x10^6$	6^{th} recipient (P5)	5

	1x10^7	**6th recipient (P5)**	**5**
	3x10^7	6th recipient (P5)	5

D. Independency of the TTL phenotype on cell dose and consistency in following passages.

ST 3 (referring to: Tab. 1): Prognostic impact of different classifiers including different thresholds defining TTL.

N=50	n	adj. P	risk ratio (relapse)	CI
cutoff 10 weeks	6	1.42 e-5	26.58	6.04 - 116.90
cutoff 16 weeks	12	1.05 e-4	16.73	4.03 - 69.43
cutoff 21 weeks	23	.013	13.87	1.75 - 109.90
hyperleukocytosis	16	.041	3.76	1.06 - 13.30
PPR	6	.048	3.93	1.01 - 15.21
unfavorable age	18	.050	3.57	1.00 - 12.72

Univariate analysis (Cox's regression) on relapse free survival (N=50) including different TTL- thresholds: 10, 16 and 21 weeks and different risk factors: prednisone poor response (PPR), hyperleukocytosis (leukocyte count at diagnosis higher than 50 [1000/µl]), and unfavorable age (0-1 and > 9 years). CI: confidence interval.

ST 4 (referring to: Fig. 3A): Patient characteristics of xenograft samples used for gene expression analysis.
Table A. categorical variables

		total		TTL^{long}		TTL^{short}	
		N	*%*	N	*%*	N	*%*
total		12	*100*	7	*100*	5	*100*
gender	female	3	*25*	3	*42*	0	-
	male	9	*75*	4	*58*	5	*100*
prednisone response (PR)	good	12	*100*	7	*100*	5	*100*
immunophenotype	B-cell precursour ALL	12	*100*	7	*100*	5	*100*
fusion gene	no	12	*100*	7	*100*	5	*100*
risk groups (RG)	non-HR	12	*100*	7	*100*	5	*100*

early relapse/event *	no	7	*58,4*	7	*100*	0	-
	yes	5	*41.6*	0	-	5	*100*
NOD/SCID passage	P 0	6	*50*	4	*58*	2	*40*
	P 1	1	*8.3*	1	*14*	0	-
	P 2	1	*8.3*	1	*14*	0	-
	P 3	2	*16.8*	1	*14*	1	*20*
	P 4	1	*8.3*	0	-	1	*20*
	P 6	1	*8.3*	0	-	1	*20*

* follow up time > 24 months, n=1 patient died while on therapy.

Table B. continuous variables

	total			TTL^{long}			TTL^{short}			T-test
	N	range	mean	N	range	mean	N	range	mean	P
age	12	1.4-15.6	7.9	7	2.00-15.6	7.5	5	1.4-15.2	8.4	.8
WBC at diagnosis [1000/µl]	12	1.2-132.0	42.6	7	1.2-71.8	34.3	5	10.5-132.0	54.2	.5
PB blast cells at diagnosis [%]	12	1.0-97.0	64.5	7	1.0-97.0	57.5	5	39.5-94.0	39.5	.4
WBC at day 8 [1000/µl]	12	0.6-5.3	2.0	7	1.0-4.1	2.0	5	0.6-5.3	2.1	.9
PB blast cells at day 8 [%]	12	0-34.0	5.5	7	0-34.0	6.6	5	0-7.0	4.1	.6
BM blast cells at day 33 [%]	12	-	0	7	-	0	5	-	0	-

ST 5 (referring to: Fig. 3B): Patient characteristics of the AIEOP-LLA cohort.

(Patients diagnosed from January 2004 to December 2006)

		N	*%*
total		197	*100*
gender	female	93	*47.2*
	male	104	*52.8*
infant leukemia		11	*5.6*
immunophenotype	B-cell precursour ALL	197	*100*
genetic abnormalities	MLL- rearrangement	17	*8.6*
	BCR/ABL- fusion	8	*4.1*
	TEL/AML1- fusion	50	*25.4*
	t(1;19)	12	*6*
	hyperdiploid karyotype	23	*11.7*
prednisone response	poor	32	*16.3*
	good	162	*82.2*
	not available	3	*1.5*
risk group	HR	46	*23.4*
	non-HR	146	*74.1*
	not available	5	*2.5*
relapse	total	40	*20.3*
	< 24 months (early)	26	*13.2*
	> 24 months (late)	14	*7.1*

ST 6 (referring to: Fig. 3B-E): Gene expression of mTOR regulators is associated with TTL.

N=24	β	P	exp(β)	95% CI of exp(β)
PDE4A	1.355	.0674	3.879	.907 - 16.59
DDIT4L	- 2.808	.0324	.0603	.0046 - 0.790
RHEB	7.417	.0005	1664	25.82 - 107 258
FRAP1	5.833	.0016	341.5	9.176 - 12 709

Univariate analyses (Cox's regression) on time to leukemia (TTL) according to gene expression levels of *PDE4A*, *DDIT4L*, *RHEB* and *FRAP1* (quantitative RT-PCR analyses; N=24 xenograft samples).

SUPPLEMENTARY EXPERIMENTAL PROCEDURES

PATIENTS AND LEUKEMIA SAMPLES

All diagnostic leukemia samples included in our studies were obtained before treatment from pediatric de novo B cell precursor (BCP ALL) ALL patients. All patients were enrolled and treated according to the ALL-BFM 2000 study protocol [2] or the AIEOP LLA 2000 study protocol [3] after informed consent was given. Infants were included but treated according to the Interfant-99 protocol [4]. Both studies are performed in accordance to the Declaration of Helsinki and registered at http://clinicaltrials.gov (ALL BFM 2000: NCT00430118 and AIEOP LLA 2000: NCT00613457). Diagnosis and diagnosis of relapse were accomplished following the protocols' criteria. The immunophenotype was assessed according to standards of the European Group for the Immunological Characterization of Leukemias (EGIL) [5]. Patients with B- cell precursor leukemia (pro- B, common and pre- B ALL) were included in this study. Patients are divided in high risk- (HR) or non-HR groups. The stratification is based on different prognostic factors such as response to treatment and genetic characteristics known to be associated with poor outcome. High risk stratification is based on poor initial response to treatment (prednisone poor response: failure of reduction of peripheral leukemia cells after 7 days of prednisone systemically and one intrathecal dose of methotrexate) or non-achievement of remission on day 33 as well as prevalence of translocations t(9;22), t(4;11) or fusion of BCR/ABL or MLL/AF4, respectively [2, 6, 7]. In addition minimal residual disease (MRD) is analyzed by detection of leukemia- clone specific immunoglobulin gene rearrangements [8, 9]. Patients with high levels of MRD on day 78 of therapy are assigned to the HR- group [10]. Patients showing none of the HR- criteria are classified as non- high risk. The patient characteristics are summarized in Table 2 and Supplementary Tables ST 4 and ST 5.

XENOTRANSPLANT MODEL OF PEDIATRIC LEUKEMIA

All experimental animal studies were conducted according to the national animal welfare law (Tierschutzgesetz) and were approved by the appropriate authority (Regierungspräsidium Tübingen, Germany, experiment number 775). NOD/SCID (NOD/LtSz-scid/scid) mice (Charles River Laboratories, Research Models and Services, Germany) were kept in a pathogen-free environment. Autoclaved water and food were provided without restriction. Patient leukemia cells were isolated by density gradient centrifugation (Biocoll separating

solution; Biochrom AG, Germany) from diagnostic bone marrow aspirates or peripheral blood. Cells were washed in Hank's balanced salt solution (HBSS, Gibco Invitrogen, Germany) with 2% FCS. All samples used for transplant consisted of more than 90% leukemia cells. Unconditioned NOD/SCID mice [11, 12] with a median age of 10 weeks were transplanted by injection into the lateral tail vein of $1x10^7$ cells per recipient throughout the study. As an exception, different numbers of leukemia cells were injected in the experiment analyzing the impact of the cell dose.

DETECTION OF LEUKEMIA ENGRAFTMENT

After transplant blood samples were evaluated regularly at 2 weeks intervals for the presence of human leukemia cells. About 50 µl of peripheral blood were taken from the tail vein and analyzed for human leukemia cells by flow cytometry [13]. Briefly, heparinized blood samples (Liquemin, Hoffman-La Roche, Grenzach-Wyhlen, Germany; 500 IU per 50 µl blood) were stained with allophycocyanin (APC)- conjugated anti- human B-cell antigen (CD19) and phycoerythrin (PE)- conjugated anti- murine leukocyte common antigen (Ly-5) (both BD Bioscience, Heidelberg, Germany) and analyzed by multiparameter flow cytometry (FACSCalibur flow cytometer and cell quest software, BD Bioscience, Heidelberg, Germany). The proportion of human CD19- positive BCP- leukemia cells was calculated reflecting the leukemic burden in the mouse as described [13, 14]. At autopsy cell suspensions from spleen and bone marrow were prepared. Spleen tissue was minced and strained through nylon cell sieves (70 µm, Falcon, BD Bioscience, Heidelberg, Germany). Bone marrow cells were isolated flushing the femoral cavity with PBS. Mononuclear cells were isolated by density gradient centrifugation (Biocoll separating solution; Biochrom AG, Berlin, Germany) and subsequently analyzed by multiparameter flow cytometry as described above.

MANIFESTATION OF LEUKEMIA

After transplant the general condition and well-being of the mice was examined regularly. Manifestation of disease was assessed by clinical signs of overt leukemia such as severely impaired general condition, lethargy, ruffled fur and impaired posture. Time to leukemia (TTL) was determined for each patient sample transplanted as weeks from the date of transplant to the date of clinical manifestation of the disease. Upon clear evidence for leukemia related morbidity, mice were killed and autopsy was performed. Leukemia was

confirmed detecting leukemia cells in bone marrow, spleen and peripheral blood. Mice without evidence for disease and without evidence for human leukemia cells in peripheral blood analyzed at week 20 after transplant were killed. Autopsy was performed and absent leukemia cell infiltration of spleen or bone marrow confirmed no overt leukemia.

RNA PREPARATION AND GENE EXPRESSION DATA

Gene expression analysis was carried out on xenograft leukemia samples with sufficient follow up time and negative for known cytogenetic abnormalities (N=12; TTL^{short} n=5 and TTL^{long} n=7) which were isolated from recipients with overt leukemia. Cell suspensions containing more than 90% leukemia cells as estimated by flowcytometry were prepared from infiltrated spleens of leukemia bearing mice. RNA was isolated using a standardized procedure (RNeasy kit, Qiagen, Hiden, Germany) following the manufacturers instructions. Expression analysis was performed using the Human Genome U133 Plus 2.0 Array (Affymetrix, Santa Clara, USA). 2 µg of total RNA were labeled using the GeneChip® One-Cycle Target Labelling Assay Kit (Affymetrix) according to the manufacturer's instructions. After hybridization on a Hybridization Oven 640 (Affymetrix) arrays were stained and washed in a FS 450 Fluidics Station (Affymetrix) before imaging on an Affymetrix GeneChip 3000 scanner.

Primary patient leukemia cells were isolated from diagnostic specimens obtained from patients treated according to the AIEOP LLA 2000 protocol. Patients with available information on early relapse (minimal follow up time 24 months) and material obtainable at diagnosis were included (N=197). Preparation of RNA and gene expression analysis was carried out using an Human Genome U133 Plus 2.0 Array (Affymetrix, Santa Clara, USA) as described previously [15]. Gene expression microarray files (Affymetrix CEL files) were generated using the GCOS 1.4 or 1.2 software (Affymetrix). For each gene expression profile a detailed data quality report has been generated to define the overall quality of each experiment. The quality parameters that were monitored besides cRNA total yield and cRNA A260/A280 ratio included: (i) background noise (Q value), (ii) percentage of present called probe sets, (iii) scaling factor, (iv) information about exogenous *Bacillus subtilis* control transcripts from the Poly-A control kit (lys, phe, thr, and dap), and (v) the ratio of intensities of 3' probes to 5' probes for a housekeeping gene (GAPDH).

Raw data have been deposited in NCBI's Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo) and are accessible through GEO Series accession number GSE13576.

QUANTITATIVE RT-PCR

To validate the array result, we performed quantitative real-time PCR for *PDE4A*, *DDTI4L*, *RHEB* and additionally *FRAP1*. We used random hexamer priming and Superscript reverse transcriptase (both Invitrogen, Karlsruhe, Germany) to generate cDNA. PCR was carried out on a LightCycler (Roche Diagnostics, Mannheim, Germany) using the LightCycler Fast Start DNA Master SYBR Green I kit as described by the manufacturer's instructions (Roche, Mannheim, Germany). Real-time PCR conditions were as follows: 95°C for 10 minutes, then 45 cycles of 95°C for 10 seconds, 65°C for 10 seconds and 72°C for 20 seconds. Melting curve analysis and product sequencing were performed to verify the amplification specificity. Relative transcript levels were determined by a comparative threshold cycles for amplification (CT) method with normalization to *TBP*, *HPRT* and *GAPDH* . The most stably expressed control genes were determined independently using GeNorm [16] and Normfinder [17]. The primer sequences are as follows (s indicating 'sense, a 'antisense'):
PDE4A-14s: 5'-GCAGTGTTCACGGACCTGGAGATTCTCGC-3',
PDE4A-16a: 5'-GCGGTCGGAGTAGTTATCTAGCAGGAGGACCCC-3';
DDIT4L-1s: 5'-CGAGCGCGCAGGCCCCCGCGAAC-3',
DDIT4L-3a: 5'-GGGTCAGTTTCTCAGGGACAAGGACCTTTG-3';
RHEB-1s 5'-CCGGAAGATCGCGATCCTGGGCTAC-3',
RHEB-6a: 5'-CCATATGCAGGTCTTTCTTATTCCCAACCAACATAATAGG-3';
FRAP1-31s: 5'-CAACTCCACCAGCAGTGCTGTGAAAAGTGGAC-3',
FRAP1-34a: 5'-GTACTGGATAACCTCCTCCAGCTCGGACAGC-3';
TBP-2s: 5'-GAGGAAGTTGCTGAGAAGAGTGTGCTGGAG-3',
TBP-3a: 5'-GTCAGTCCAGTGCCATAAGGCATCATTGG-3';
HPRT1-3s: 5'-GAGATGGGAGGCCATCACATTGTAGCCCTC-3',
HPRT1-4a: 5'-CTCCACCAATTACTTTTATGTCCCCTGTTGACTGGTC-3';
GAPDH-8s: 5'-CAGAACATCATCCCTGCCTCTACTGGC-3',
GAPDH-9a: 5'-GGTCTTACTCCTTGGAGGCCATGTGGG-3'.

EX VIVO TREATMENT OF XENOGRAFT SAMPLES

Leukemia cells were isolated from recipient animals presenting with overt leukemia as described above. Cells were incubated following standard culture conditions in RPMI 1640 (Life Technologies, Eggenstein, Germany) supplemented with 10% fetal calf serum (FCS; Sigma Chemicals, Deisenhofen, Germany), together with penicillin, streptomycin, and glutamine (Biochrome, Berlin, Germany) at 37°C in humidified air/5% CO_2. Cells were treated with either rapamycin or rolipram diluted in dimethyl sulphoxide (DMSO, all Sigma Chemicals, Deisenhofen, Germany) at concentrations of 10 and 100 nM (rapamycin) or 10 and 100 µM (rolipram) or with DMSO alone. Viability of leukemia cells was assessed by flowcytometry according to forward and side scatter criteria. Specific cell death was calculated as described previously [18]: 100 x (experimental dead cells [%] - dead cells in DMSO control [%]) / (100% - dead cells in DMSO control [%]).

REFERENCES TO SUPPLEMENTARY INFORMATION

1. Kraus, J.M., and Kestler, H.A. (2010). A highly efficient multi-core algorithm for clustering extremely large datasets. BMC Bioinformatics *11*, 169.
2. Schrappe, M., Reiter, A., Ludwig, W.D., Harbott, J., Zimmermann, M., Hiddemann, W., Niemeyer, C., Henze, G., Feldges, A., Zintl, F., Kornhuber, B., Ritter, J., Welte, K., Gadner, H., and Riehm, H. (2000). Improved outcome in childhood acute lymphoblastic leukemia despite reduced use of anthracyclines and cranial radiotherapy: results of trial ALL-BFM 90. German-Austrian-Swiss ALL-BFM Study Group. Blood *95*, 3310-3322.
3. Conter, V., Arico, M., Valsecchi, M.G., Basso, G., Biondi, A., Madon, E., Mandelli, F., Paolucci, G., Pession, A., Rizzari, C., Rondelli, R., Zanesco, L., and Masera, G. (2000). Long-term results of the Italian Association of Pediatric Hematology and Oncology (AIEOP) acute lymphoblastic leukemia studies, 1982-1995. Leukemia *14*, 2196-2204.
4. Pieters, R., Schrappe, M., De Lorenzo, P., Hann, I., De Rossi, G., Felice, M., Hovi, L., LeBlanc, T., Szczepanski, T., Ferster, A., Janka, G., Rubnitz, J., Silverman, L., Stary, J., Campbell, M., Li, C.K., Mann, G., Suppiah, R., Biondi, A., Vora, A., and Valsecchi, M.G. (2007). A treatment protocol for infants younger than 1 year with acute lymphoblastic leukaemia (Interfant-99): an observational study and a multicentre randomised trial. Lancet *370*, 240-250.
5. Bene, M.C., Castoldi, G., Knapp, W., Ludwig, W.D., Matutes, E., Orfao, A., and van't Veer, M.B. (1995). Proposals for the immunological classification of acute leukemias. European Group for the Immunological Characterization of Leukemias (EGIL). Leukemia *9*, 1783-1786.
6. Schrappe, M., Reiter, A., Zimmermann, M., Harbott, J., Ludwig, W., Henze, G., Gadner, H., Odenwald, E., and Riehm, H. (2000). Long-term results of four consecutive trials in childhood ALL performed by the ALL-BFM study group from 1981 to 1995. Berlin-Frankfurt-Münster. Leukemia *14*, 2205-2222.
7. Viehmann, S., Borkhardt, A., Lampert, F., and Harbott, J. (1999). Multiplex PCR--a rapid screening method for detection of gene rearrangements in childhood acute lymphoblastic leukemia. Ann Hematol. *78*, 157-162.
8. van Dongen, J.J., Seriu, T., Panzer-Grumayer, E.R., Biondi, A., Pongers-Willemse, M.J., Corral, L., Stolz, F., Schrappe, M., Masera, G., Kamps, W.A., Gadner, H., van Wering, E.R., Ludwig, W.D., Basso, G., de Bruijn, M.A., Cazzaniga, G., Hettinger, K., van der Does-van den Berg, A., Hop, W.C., Riehm, H., and Bartram, C.R. (1998). Prognostic value of minimal residual disease in acute lymphoblastic leukaemia in childhood. Lancet *352*, 1731-1738.
9. Cave, H., van der Werff ten Bosch, J., Suciu, S., Guidal, C., Waterkeyn, C., Otten, J., Bakkus, M., Thielemans, K., Grandchamp, B., and Vilmer, E. (1998). Clinical significance of minimal residual disease in childhood acute lymphoblastic leukemia. European Organization for Research and Treatment of Cancer--Childhood Leukemia Cooperative Group. N Engl J Med *339*, 591-598.
10. Flohr, T., Schrauder, A., Cazzaniga, G., Panzer-Grumayer, R., van der Velden, V., Fischer, S., Stanulla, M., Basso, G., Niggli, F.K., Schafer, B.W., Sutton, R., Koehler, R., Zimmermann, M., Valsecchi, M.G., Gadner, H., Masera, G., Schrappe, M., van Dongen, J.J., Biondi, A., and Bartram, C.R. (2008). Minimal residual disease-directed risk stratification using real-time quantitative PCR analysis of immunoglobulin and T-cell receptor gene rearrangements in the international multicenter trial AIEOP-BFM ALL 2000 for childhood acute lymphoblastic leukemia. Leukemia *22*, 771-782.
11. Baersch, G., Möllers, T., Hötte, A., Dockhorn-Dworniczak, B., Rübe, C., Ritter, J., Jürgens, H., and Vormoor, J. (1997). Good engraftment of B-cell precursor ALL in NOD-SCID mice. Klin Padiatr. *209*, 178-185.
12. Spiegel, A., Kollet, O., Peled, A., Abel, L., Nagler, A., Bielorai, B., Rechavi, G., Vormoor, J., and Lapidot, T. (2004). Unique SDF-1-induced activation of human precursor-B ALL cells as a result of altered CXCR4 expression and signaling. Blood *103*, 2900-2907.
13. Lock, R.B., Liem, N., Farnsworth, M.L., Milross, C.G., Xue, C., Tajbakhsh, M., Haber, M., Norris, M.D., Marshall, G.M., and Rice, A.M. (2002). The nonobese diabetic/severe combined immunodeficient (NOD/SCID) mouse model of childhood acute lymphoblastic leukemia reveals intrinsic differences in biologic characteristics at diagnosis and relapse. Blood *99*, 4100-4108.
14. Nijmeijer, B.A., Mollevanger, P., van Zelderen-Bhola, S.L., Kluin-Nelemans, H.C., Willemze, R., and Falkenburg, J.H. (2001). Monitoring of engraftment and progression of acute lymphoblastic leukemia in individual NOD/SCID mice. Exp Hematol *29*, 322-329.
15. Campo Dell'Orto, M., Zangrando, A., Trentin, L., Li, R., Liu, W.M., te Kronnie, G., Basso, G., and Kohlmann, A. (2007). New data on robustness of gene expression signatures in leukemia: comparison of three distinct total RNA preparation procedures. BMC Genomics *8*, 188.

16. Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol *3*, RESEARCH0034.
17. Andersen, C.L., Jensen, J.L., and Orntoft, T.F. (2004). Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res *64*, 5245-5250.
18. Friesen, C., Herr, I., Krammer, P.H., and Debatin, K.M. (1996). Involvement of the CD95 (APO-1/FAS) receptor/ligand system in drug-induced apoptosis in leukemia cells. Nat Med *2*, 574-577.

CONCLUSIONS

Acute lymphoblastic leukaemia (ALL) is the most common and the most successfully treated malignancy in children, with long-term survival rates reaching 75–80% as the result of tailored treatment strategies[1]. However, up to a quarter of children suffer a recurrence of the disease and the outcome for these patients remains dismal[2, 3, 4]. Notably, many relapses occur in low risk stratified patients, who initially present with favourable prognostic features[5] indicating the need to improve the identification of patients at increased risk of treatment failure.

Many studies have demonstrated that gene expression profiles (GEP) can aid classification and risk stratification in ALL[6, 7, 8, 9, 10, 11]. GEP was successfully applied to study the mechanism related to drug resistance, to predict response to therapeutic compounds, to find biomarkers for monitoring of minimal residual disease. Using Gene Ontology tools information coming from thousand of genes permit to understand many biological nuances specific of leukemia subgroups. GEP technology has been also extended to the study of small RNA expression offering a much wider vision on genomic regulation related to diseases.

In this thesis, GEP was applied to several research approaches aiming to dissect the heterogeneity and to find novel therapeutic targets in B cell precursor ALL (BCP ALL).

In BCP ALL, individual chromosomal abnormalities are strong independent indicators of outcome[12] and lead to disruption of genes involved in the regulation of B-lymphocyte differentiation. Twenty-five percent of BCP ALL cases lack major genetic aberrations such as balanced chromosomal translocations or aneuploidy, and are general known as B-others[13, 14, 15]. In the last few years improvement in research tools such as FISH, GEP and analysis of genome-wide DNA copy number alterations have led to the discovery of new recurrent genetic lesions. Chromosome 21 amplification (iAmp21) are found in 2% of ALL, *CRLF2* aberrations regard 6% of ALL and *ERG* aberrations occur in 7% of ALL. These aberrations are found for the major part among B-others BCP ALL and recently these recurrent lesions are employed to define new ALL subtypes reducing the the group of B-others to 7% of ALL[16]. With exception of iAmp21, that is a defined subgroup associated to unfavourable outcome the other two subgroups (*CRLF2* aberrations and *ERG* deletions) are not yet

extensively studied. The incidence of *CRLF2* aberrations remains to be defined and type of rearrangements involving *CRLF2*, prognostic impact associated to different aberrations (overexpressioin or different rearrangements) and deregulated pathway involved in the disease are currently subject of collaborative studies. Also dissection of patients with *ERG* aberrancies are ongoing. Patients in this group are reported to share gene expression profiles and to present a truncated form of ERG protein, but at the genetic level only a subgroup of patients carries an *ERG* intragenic deletions. Together these open issues led me to regards B-others as a large heterogeneous group where new subtype are under investigation.

A consistent number of relapses currently occurs in the B-others group indicating the need for new biological insights and treatment options for these patients[17]. Gene expression profiling of B-other groups confirmed their heterogenic nature and revealed that some subgroups signatures resemble those of genetically defined groups (such as *BCR-ABL*), suggesting for common deregulated pathways.

Twenty-five percent of B-others patients (5% of BCP ALL) fitted in a group with favourable outcome and a unique signature associated to frequent *ERG* intragenic deletion. MicroRNAs expression profiling revealed in this group an unique miRNAs signature characterized by over expression of hsa-miR-125b, -125b-2*, -99a, -100, -125a-3p and has-miR-491-5p. Integration of data from gene and miRNAs profiling revealed over expression of miRNAs and genes in 21q21.1 chromosomal region. Genome-wide analysis excluded copy number alterations of this deregulated region. Unique of this group a specific snoRNAs expression profile with up regulation of many snoRNAs in 15q11.2 locus was observed.
These findings confirm GEP efficacy in the class discovery, underscore involvement of miRNAs in the diseases and the portent of a gene-miRNA integrated research approaches. Further investigations are need to elucidate the role of miR-125b-2 cluster in BCP ALL, testing its employments as class marker and to assess the association with favourable outcome (miR-125b-2 was also found up-regulated in *ETV6-RUNX1* BCP ALL, a subtype associate to favourable outcome[18]) . Further it will be great interest to dissect the role for 21q21.1 chromosomal region, not previuosly reported to be deregulated in leukemia.

Actually, human chromosome 21 (Hsa21) is frequently involved in ALL genomic aberrations (e.g. hyperdiploidy (HD), t(12;21) or iAmp21), and the high incidence of ALL in children with Down Syndrome (DS) reinforce the hypothesis of a directly and functional contribution of Hsa21 genes to the malignant transformation of hematopoietic cells. However, gene expression data shows large differences between DS ALL and HD ALL leukemias,

reflecting the fundamental differences between constitutional and acquired trisomy and suggesting distinct ways of Hsa21 involvement. Genomic analysis of a large group of DS ALLs reveals highly heterogeneity of the group, and an enrichment with DNA damage and BCL6 responsive genes suggests the possibility of B-cell lymphocytic genomic instability among the causes leading to leukemia in constitutional trisomy 21.

Two-thirds of the DS ALL patients shared the aberrant expression of the cytokine receptor *CRLF2* and data on cooperation between CRLF2 and mutated JAK2 in conferring cytokine independent growth to pro-B cells suggest that the DS ALL children with *CRLF2* aberrant expression may benefit from therapy blocking the CRLF2-JAK2 pathway.

A recent work on primary cells have established the critical roles of JAK1 and JAK2 for TSLP-mediated STAT5 activation[19]. The GEP data on enrichment in BCL6 responsive gene (activated by STAT5[20]) and the finding in patients with high *CRLF2* expressing of a signature similar to *BCR-ABL1* patients (BCR–ABL1 kinase constitutively activate STAT5) are in line with a constitutional STAT5 activation and underscore the power of GEP to catch the biological features of disease.

Further investigation of incidence and prognostic impact of *CRLF2* aberrations in non-DS ALL patients enrolled into the AIEOP-BFM ALL2000 study revealed that the *P2RY8-CRLF2* rearrangement in association with 20 times over expression of *CRLF2* identifies BCP ALL patients with a very poor prognosis and, among them, an important subset of patients currently stratified in the intermediate risk. In the light of this result, combined *P2RY8-CRLF2* fusion and *CRLF2* overexpression will be proposed as a new marker of poor prognosis for children with intermediate risk BCP ALL.

Developing the issues above and going forward on deregulated pathways investigation will require the use of model studies such as mouse xenotransplantation a valid tool for amplifying leukemia cells and for in vivo testing of inhibitors. The GEP study on the NOD/SCID/huALL transplant model proved its capability to recapitulate the primary human leukaemia, mimicking the features of the primary malignancy and retaining these characteristics over serial passages without selection for a subclone of leukaemia initiating cells. Furthermore, gene signatures reflecting the features of those xenotransplants that showed rapid engraftment attitude, identified a cluster of early relapsed patients in an independent cohort. Gene expression profiling studies of rapid engraftment revealed a set of genes associated with this aggressive phenotype providing a potential strategy to identify patients with high risk of early relapse. Most importantly, pathways regulating cell growth

were identified. Indeed, xenotransplants that showed rapid engraftment are characterized by an activated mTOR pathway (*DDIT4L*low/*RHEB*high/*FRAP1*high) and *ex vivo* treatment of xenograft cells with mTOR inhibitor showed to increase cells death. In line with the involvement of mTOR pathway in aggressiveness of the disease, there is the finding of overexpression of miRNAs (miR-100 and miR-99a) know to inhibits mTOR signalling in the group associated to a favourable outcome. Furthermore in this group down-regulation of genes related to mTOR signalling are observed. Taken together this finding suggest that mTOR pathway could provide targets for alternative therapeutic strategies for high risk patients.

Concluding, ten years after its introduction in oncohematology, GEP continues to be a valuable research tool, efficacious in subtype discovery and leading researchers to discover deregulated pathways and biomarkers identification in BCP ALL.

REFERENCES

[1] Pui CH, Relling MV, Downing JR. Acute lymphoblastic leukemia. *N Engl J Med.* 2004;350:1535–48.

[2] Chessells, J.M. (1998) Relapsed lymphoblastic leukaemia in children: a continuing challenge. British Journal of Haematology, 102, 423–438

[3] Gaynon, P.S., Qu, R.P., Chappell, R.J., Willoughby, M.L., Tubergen, D.G., Steinherz, P.G. & Trigg, M.E. (1998) Survival after relapse in childhood acute lymphoblastic leukaemia: impact of site and time to first relapse – the Children's Cancer Group Experience. Cancer, 82, 1387–1395.

[4] Uderzo, C., Dini, G., Locatelli, F., Miniero, R. & Tamaro, P. (2000) Treatment of childhood acute lymphoblastic leukaemia after the first relapse: curative strategies. Haematologica, 85, 47–53.

[5] Seeger, K., von Stackelberg, A., Taube, T., Buchwald, D., Korner, G., Suttorp, M., Dorffel, W., Tausch, W. & Henze, G. (2001) Relapse of TEL-AML1 – positive acute lymphoblastic leukaemia in childhood. Journal of Clinical Oncology, 19, 3188–3193.

[6] Yeoh, E.J., Ross, M.E., Shurtleff, S.A., Williams, W.K., Patel, D., Mahfouz, R., Behm, F.G., Raimondi, S.C., Relling, M.V., Patel, A., Cheng, C., Campana, D., Wilkins, D., Zhou, X., Li, J., Liu, H., Pui, C.H., Evans, W.E., Naeve, C., Wong, L. & Downing, J.R. (2002) Classification, subtype discovery, and prediction of outcome in paediatric acute lymphoblastic leukaemia by gene expression profiling. Cancer Cell, 1, 133–143.

[7] Willenbrock, H., Juncker, A.S., Schmiegelow, K., Knudsen, S. & Ryder, L.P. (2004) Prediction of immunophenotype, treatment response, and relapse in childhood acute lymphoblastic leukaemia using DNAmicroarrays. Leukaemia, 18, 1270–1277.

[8] Winter, S.S., Jiang, Z., Khawaja, H., Griffin, T., Devidas, M., Asselin, B.L. & Larson, R.S. (2007) Identification of genomic classifiers that distinguish induction failure in T-lineage acute lymphoblastic leukaemia: a report from the Children's Oncology Group. Blood, 110, 1429–1438.

[9] Ross, M.E., Zhou, X., Song, G., Shurtleff, S.A., Girtman, K., Williams, W.K., Liu, H.C., Mahfouz, R., Raimondi, S.C., Lenny, N., Patel, A. & Downing, J.R. (2003) Classification of paediatric acute lymphoblastic leukaemia by gene expression profiling. Blood, 102, 2951–2959.

[10] Chiaretti, S., Li, X., Gentleman, R., Vitale, A., Vignetti, M., Mandelli, F., Ritz, J. & Foa, R. (2004) Gene expression profile of adult T-cell acute lymphocytic leukaemia identifies distinct subsets of patients with different response to therapy and survival. Blood, 103, 2771–2778.

[11] Gottardo, N.G., Hoffmann, K., Beesley, A.H., Freitas, J.R., Firth, M.J., Perera, K.U., de Klerk, N.H., Baker, D.L. & Kees, U.R. (2007) Identification of novel molecular prognostic markers for paediatric T-cell acute lymphoblastic leukaemia. British Journal of Haematology, 137, 319–328.

[12] Moorman, A.V., Ensor, H.M., Richards, S.M., Chilton, L., Schwab, C., Kinsey, S.E., Vora, A., Mitchell, C.D. & Harrison, C.J. (2010) Prognostic effect of chromosomal abnormalities in childhood B-cell precursor acute lymphoblastic leukaemia: results from the UK Medical Research Council ALL97/99 randomised trial. The Lancet Oncology, 11, 429–438.

[13] Pui CH, Relling MV, Downing JR. Acute lymphoblastic leukemia. *N Engl J Med.* 2004;350:1535–48.

[14] Yeoh, E.J., Ross, M.E., Shurtleff, S.A., Williams, W.K., Patel, D., Mahfouz, R., Behm, F.G., Raimondi, S.C., Relling, M.V., Patel, A., Cheng, C., Campana, D., Wilkins, D., Zhou, X., Li, J., Liu, H., Pui, C.H., Evans, W.E., Naeve, C., Wong, L. & Downing, J.R. (2002) Classification, subtype discovery, and prediction of outcome in paediatric acute lymphoblastic leukaemia by gene expression profiling. Cancer Cell, 1, 133–143.

[15] Pieters R, Carroll WL. Biology and treatment of acute lymphoblastic leukemia. Pediatr Clin North Am. 2008;55:1–20.

[16] J Clin Oncol. 2011 Jan 10. [Epub ahead of print] Biology, Risk Stratification, and Therapy of Pediatric Acute Leukemias: An Update. Pui CH, Carroll WL, Meshinchi S, Arceci RJ.

[17] Möricke A, Reiter A, Zimmermann M, et al. Risk-adjusted therapy of acute lymphoblastic leukemia can decrease treatment burden and improve survival: treatment results of 2169 unselected pediatric and adolescent patients enrolled in the trial ALL-BFM 95. Blood. 2008;111:4477–89.

[18] Gefen N, et al. (2010) Hsa-mir-125b-2 is highly expressed in childhood ETV6/RUNX1 (TEL/AML1) leukemias and confers survival advantage to growth inhibitory signals independent of p53. Leukemia 24(1):89–96

[19] Thymic stromal lymphopoietin-mediated STAT5 phosphorylation via kinases JAK1 and JAK2 reveals a key difference from IL-7–induced signalling. Yrina Rochmana, Mohit Kashyapa, Gertraud W. Robinsonb, Kazuhito Sakamotoc, Julio Gomez-Rodriguezd, Kay-Uwe Wagnerc, and Warren J. Leonarda. PNAS,November 9, 2010, vol. 107, no. 45, 19455–19460.

[20] Scheeren FA, Naspetti M, Diehl S, Schotte R, Nagasawa M, Wijnands E, Gimeno R, Vyth-Dreese FA, Blom B, Spits H: STAT5 regulates the self-renewal capacity and differentiation of human memory B cells and controls Bcl-6 expression. Nat Immunol 2005, 6:303-313.