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

Running Head: Genomic analysis of DS-ALL

Libi Hertzberg1-3,*, Elena Vendramini4,*, Ithamar Ganmore1-2,*, Gianni Cazzaniga5, Maike Schmitz6, Jane Chalker7, Ruth Shiloh1, Ilaria Iacobucci8, Chen Shochat1,2,9, Sharon Zeligson1, Gunnar Cario10, Martin Stanulla10, Sabine Strehl11, Lisa J Russell12, Christine J Harrison12, Beat Bornhauser6, Akinori Yoda13, Gideon Rechavi1,2, Dani Bercovich9, Arndt Borkhardt14, Helena Kempski7, Geertruy te Kronnie4,8, Jean-Pierre Bourquin6,§, Eytan Domany3,§, Shai Izraeli1,2,§

* Co-first authors
§ Co-senior authors

Affiliations:

1Pediatric Hemato-Oncology and the Cancer Research Center, Sheba Medical Center, Tel Hashomer, Ramat-Gan, Israel
2Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel
3Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot, Israel
4Hematology Laboratory, Department of Pediatrics, University of Padova, Padova, Italy
5Centro Ricerca Tettamanti, Clinica Paediatrica Università di Milano-Bicocca, Ospedale San Gerardo, Monza, Italy
6Department of Pediatric Oncology, University Children’s Hospital, University of Zurich, Zurich, Switzerland
7Paediatric Malignancy Cytogenetics Unit, Camelia Botnar Laboratories, Great Ormond Street Hospital & Institute of Child Health, London, UK
8Department of Hematology/Oncology "L. and A. Seràgnoli", University of Bologna, Bologna, Italy

9Human Molecular Genetics and Pharmacogenetics Laboratory, Migal–Galilee Biotechnology Center, Kiryat Shmona, and Tel-Hai Academic College, Israel

10Department of Pediatrics, University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany

11Children’s Cancer Research Institute, St Anna Kinderkrebsforschung, Vienna, Austria

12Leukaemia Research Cytogenetics Group, Northern Institute for Cancer Research, Newcastle University, UK;

13Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA, USA

14Clinic of Pediatric Oncology, Hematology and Clinical Immunology, Children’s University Hospital, Heinrich Heine University, Düsseldorf, Germany

**Corresponding Author:** Shai Izraeli, MD Sheba Medical Center, Tel Hashomer Ramat Gan 52621 Israel. sizraeli@sheba.health.gov.il  Tel +97235305943 Fax +97235305942
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-B-cells. 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 (DS-ALL). 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-AML1 ALLs, the two most common genetic subtypes of childhood ALL. Given that these cytogenetic abnormalities are less frequent in DS-ALL, the existence of unique collaborating somatic genetic events in DS-ALL, similar to the GATA1 mutation in DS-AMKL has been postulated.

We and others reported the presence of somatic activating mutations in JAK2 in approximately 20% of DS-ALL. Similar mutations are present in about 10% of high-risk ALL in non-DS children corresponding to approximately 3% of unselected childhood ALLs. We hypothesized that the mutated JAK2 may cooperate with a type I cytokine receptor that is aberrantly expressed in DS-ALL.

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 the ethics committees of all participating institutions in accordance with the Declaration of Helsinki. 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. 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. 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 NCBI's 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. 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\textsuperscript{10}.

Flow Cytometry analysis (Becton-Dickinson Canto-II, FlowJo software) was performed on primary cryopreserved ALL cells after the first xenotransplantation in Nod/LtSzScid IL2\textgammanull 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 pHRSINC\textgammawild lentivirus\textsuperscript{11} 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 (QuickChange\textsuperscript{TM}-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-hCRLF2\textsuperscript{12} and hCRLF2 expressing cells were selected with puromycin (2\textmu g/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 described before. Antibodies used were anti-JAK2 (C-20, Santa Cruz), anti-STAT5, anti-phospho-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 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.

**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 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),\textsuperscript{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,\textsuperscript{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 (Figure 1A-B). 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 (Figure 1A-B). 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 and 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 heterozygosity (LOH) analyses (Supplementary Figure 3S; Table
10S) generally confirm previous reports\textsuperscript{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.\textsuperscript{17} Recently, deletions in the *IKZF1* gene were reported in the majority of patients with *BCR/ABL* and "*BCR/ABL* like" ALL.\textsuperscript{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.\textsuperscript{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 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 DS-ALL profile we interrogated the DAVID database of Gene Ontology functional categories. 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 ($P=5\times10^{-4}$) 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-cells28 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 t-test P values of \(3.7 \times 10^{-11}\), \(2.2 \times 10^{-7}\), \(5.6 \times 10^{-5}\) for AIEOP, BFM and ICH datasets, respectively.

CRLF2 is known to dimerize with IL7RA to form the heterodimeric receptor for TSLP (thymic stromal lymphopoietin). While CRLF2 is aberrantly expressed in DS-ALLs (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 over-expression of SOX531. 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 findings10, the interstitial deletion is more common than the IgH@ translocation.

To explore the effect of CRLF2 on gene expression we compared the 30% of DS-ALLs 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 mJak2-FLAG, R683S mJak2-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 signaling 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\textsuperscript{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 (\textit{TEL-AML1}, Hyperdiploidy, \textit{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. \textit{CRLF2} is one of the three genes most differentiating between DS and non-DS ALLs. Confirming the expression of \textit{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 \textit{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 \textit{CRLF2} in 5\% of non selected childhood ALL, and in 35 of 68 DS-ALL (52\%) consecutively enrolled in UK treatment protocols\textsuperscript{10}.

We report that the interstitial deletion results in fusion transcript in which the first non-coding exon of \textit{P2YR8} fuses to the coding region of \textit{CRLF2}, thereby driving \textit{CRLF2} expression by the \textit{P2YR8} promoter. A similar mechanism was reported in a single patient with splenic lymphoma and \textit{P2YR8-SOX5} fusion\textsuperscript{31} and is reminiscent of the common \textit{SIL-SCL (STIL-TAL1)} rearrangement in T-ALL\textsuperscript{32}. Cloning of the genomic breakpoints is required to determine if, like \textit{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 al10 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 al10 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 al40 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 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 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 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 pre-leukemic 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.
Acknowledgments

We thank Dr T Enver for supplying the pHRSINCWSGW lentivirus plasmid and Dr Kitamura for providing the CRLF2 expression construct. We thank N Amariglio, G Basso, V Binder, A Biondi, S Bresolin, O Haas, J Harbott, M Grazia, G Martinelli, M Schrappe, S Orkin, D Weinstock, A Yitzhaky, Y Shinar, A Ovadia and the members of E Domany's and S Izraeli's groups, in particular Y Birger and E Ben-Meir, for their important contributions to the study. Funding for this study was provided by the Israel Ministry of Science (to ED + SI), Israeli Ministry of trade (DB + SI), Israeli Ministry of Health (SI), German Israeli Foundation (SI+AB), Israeli Science Foundation and Morasha program (SI), US Israel Binational Science Foundation (SI), Children with Leukaemia UK (SI), Curtis Katz (SI), Wolfson Foundation (ED+SI), Foulkes Foundation (LH+IG), The Ridgefield Foundation (ED), Converging Technologies Program (LH), Constantiner Institute for Molecular Genetics (IG), Hans Altschüler-Stiftung (SI+JPB), Swiss National Science Foundation (JPB), Leukemia Research Fund, UK (CH+LJR), PRIN/Programmi di ricerca di Relevante Interesse Nazionale, Rome (GteK), European LeukemiaNet, AIRC, AIL, Fondazione Del Monte di Bologna e Ravenna, FIRB 2006 (II). Reagents for the AIEOP GEP study were supplied by Roche Molecular Diagnostics in the framework of the MILE study (Haferlach et al, JCO 2009, in-press).

This study has been performed as partial fulfilment of the requirement for a PhD degree from the Tel Aviv University Faculty of Medicine to Libi Hertzberg, Ithamar Ganmore and Chen Shochat
Authorship Contributions and Disclosure of Conflicts of Interest

The authors declare no conflicts of interests

Libi Hertzberg¹-³, *, Elena Vendramini⁴, *, Ithamar Ganmore¹-², *, Gianni Cazzaniga⁵,
Maike Schmitz⁶, Jane Chalker⁷, Ruth Shiloh¹, Ilaria Iacobucci⁸, Chen Shochat¹-²,⁹,
Sharon Zeligson¹, Gunnar Cario¹⁰, Martin Stanulla¹⁰, Sabine Strehl¹¹, Lisa J Russell¹²,
Christine J Harrison¹², Beat Bornhauser⁶, Akinori Yoda¹³, Gideon Rechavi¹-², Dani
Bercovich⁹, Arndt Borkhardt¹⁴, Helena Kempski⁷, Geertruy te Kronnie⁴,⁸, Jean-Pierre
Bourquin⁶, §, Eytan Doman⁷, §, Shai Izraeli¹-², §

SI, ED, JPB, HK, GC, GtK, GR, MSt - Conceived and planned the study
GC, GC, MSt, SS, AY, AB - Provided reagents and/or important data
EV, IG, GC, MS, JC, RS, II, CS, SZ, LJR, CH, BB, DB, JPB, GtK, SI – Performed
and/or analyzed experiments
LH and ED – Performed all the bioinformatics analysis
LH, IG, SI, ED wrote the paper with contributions by MS, II, RS, CS, SZ, GC, SS, LJR
SI assumes full responsibility for the content of the paper including supplementary
files.
References


Proceedings of the 50th Annual Meeting of the American Society of Hematology.
2008:Abstract Nr 295.


Tables

Table 1. Description of the gene expression datasets analyzed

<table>
<thead>
<tr>
<th>Dataset Symbol</th>
<th>No. samples</th>
<th>DS-ALL</th>
<th>HD</th>
<th>TEL-AML1</th>
<th>Other</th>
<th>platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIEOP</td>
<td>97</td>
<td>25</td>
<td>26</td>
<td>29</td>
<td>4 E2A-PBX1</td>
<td>Affymetrix HG-U133 Plus 2.0</td>
</tr>
<tr>
<td>ICH</td>
<td>15</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td></td>
<td>Affymetrix HG-U133 Plus 2.0</td>
</tr>
<tr>
<td>BFM</td>
<td>7†</td>
<td>7</td>
<td>12</td>
<td>17</td>
<td></td>
<td>Affymetrix HG-U133A</td>
</tr>
<tr>
<td>+ St Jude 15</td>
<td>29†</td>
<td>12</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL</td>
<td>27†</td>
<td>11</td>
<td>10</td>
<td>6</td>
<td></td>
<td>Affymetrix Exon 1.0 ST</td>
</tr>
</tbody>
</table>

*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
†We used a subgroup of HD and TEL-AML1 samples from St Jude dataset

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

<table>
<thead>
<tr>
<th>DS ALL up-regulated genes</th>
<th>ES (enrichment score)</th>
<th>Nominal p-value</th>
<th>FDR q-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICH dataset</td>
<td>0.437993</td>
<td>0.01002</td>
<td>0.008979</td>
</tr>
<tr>
<td>IL dataset</td>
<td>0.361621</td>
<td>0.035124</td>
<td>0.022851</td>
</tr>
<tr>
<td>BFM dataset</td>
<td>0.344479</td>
<td>0.001996</td>
<td>0.009101</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DS ALL down-regulated genes</th>
<th>ES (enrichment score)</th>
<th>Nominal p-value</th>
<th>FDR q-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICH dataset</td>
<td>-0.50601554</td>
<td>0.005617978</td>
<td>0.012959</td>
</tr>
<tr>
<td>IL dataset</td>
<td>-0.5008136</td>
<td>0.01192843</td>
<td>0.006397</td>
</tr>
<tr>
<td>BFM dataset</td>
<td>-0.55240697</td>
<td>0.001976285</td>
<td>0.001131</td>
</tr>
</tbody>
</table>
Table 3. Gene ontology pathways over-represented in the differential DS-ALL signature. For details see Table 6S

<table>
<thead>
<tr>
<th>Gene Ontology Group</th>
<th>Size</th>
<th>P value</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0006974~ response to DNA damage stimulus</td>
<td>16</td>
<td>0.000503</td>
<td>0.77%</td>
</tr>
<tr>
<td>GO:0006397~ mRNA processing</td>
<td>13</td>
<td>0.001857</td>
<td>2.80%</td>
</tr>
<tr>
<td>GO:0015031~ protein transport</td>
<td>24</td>
<td>0.00216</td>
<td>3.30%</td>
</tr>
<tr>
<td>GO:0008104~ protein localization</td>
<td>26</td>
<td>0.002438</td>
<td>3.70%</td>
</tr>
<tr>
<td>GO:0046907~ intracellular transport</td>
<td>24</td>
<td>0.002913</td>
<td>4.40%</td>
</tr>
<tr>
<td>GO:0065003~ macromolecular complex assembly</td>
<td>20</td>
<td>0.003825</td>
<td>5.80%</td>
</tr>
<tr>
<td>GO:0051649~ establishment of cellular localization</td>
<td>27</td>
<td>0.005078</td>
<td>7.60%</td>
</tr>
<tr>
<td>GO:0043067~ regulation of programmed cell death</td>
<td>19</td>
<td>0.006188</td>
<td>9.10%</td>
</tr>
</tbody>
</table>
Table 4. Enrichment of BCL6 related gene expression signatures and direct targets in DS-ALL profile genes (FDR < 15%).

<table>
<thead>
<tr>
<th>Oncomine &quot;Molecular Concepts&quot; enriched in DS-ALL up-regulated genes</th>
<th>Size</th>
<th>P value</th>
<th>Q-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EREB Lymphoblastoid CellLine BCL6 transfection top 10% over-expressed&lt;sup&gt;28&lt;/sup&gt;</td>
<td>17</td>
<td>0.003733</td>
<td>0.093</td>
</tr>
<tr>
<td>Ramos Burkitt Lymphoma CellLine BCL6 Pest mutant top 5% under-expressed in Anti IgM&lt;sup&gt;28&lt;/sup&gt;</td>
<td>11</td>
<td>0.01138</td>
<td>0.138</td>
</tr>
<tr>
<td>Lymphoma BCL6 break top 5% over-expressed&lt;sup&gt;29&lt;/sup&gt;</td>
<td>11</td>
<td>0.01866</td>
<td>0.138</td>
</tr>
<tr>
<td>Lymphoma BCL6 break top 10% over-expressed&lt;sup&gt;29&lt;/sup&gt;</td>
<td>17</td>
<td>0.02213</td>
<td>0.138</td>
</tr>
<tr>
<td>BCL6 direct targets&lt;sup&gt;27&lt;/sup&gt;</td>
<td>11</td>
<td>0.02951</td>
<td>0.147</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Oncomine &quot;Molecular Concepts&quot; enriched in DS-ALL down-regulated genes</th>
<th>Size</th>
<th>P value</th>
<th>Q-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lymphoma BCL6 break top 10% under-expressed&lt;sup&gt;29&lt;/sup&gt;</td>
<td>36</td>
<td>7.89×10&lt;sup&gt;-5&lt;/sup&gt;</td>
<td>0.001</td>
</tr>
<tr>
<td>EREB Lymphoblastoid CellLine BCL6 transfection top 10% under-expressed&lt;sup&gt;28&lt;/sup&gt;</td>
<td>26</td>
<td>0.00011</td>
<td>0.001</td>
</tr>
<tr>
<td>Lymphoma BCL6 break top 5% under-expressed&lt;sup&gt;29&lt;/sup&gt;</td>
<td>20</td>
<td>0.001007</td>
<td>0.008</td>
</tr>
</tbody>
</table>
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.

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

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

<table>
<thead>
<tr>
<th></th>
<th>No/low CRLF2 expression (n=20)</th>
<th>High/medium CRLF2 expression (n=33)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>11</td>
<td>17</td>
<td>1†</td>
</tr>
<tr>
<td>Female</td>
<td>9</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SE)</td>
<td>9.87 (1.21)</td>
<td>5.56 (0.67)</td>
<td>0.004†</td>
</tr>
<tr>
<td>Median (range)</td>
<td>11.62 (1.74-18.7)</td>
<td>4.06 (1.99-20.22)</td>
<td></td>
</tr>
<tr>
<td><strong>White blood cell count at diagnosis (cell/μl)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SE)</td>
<td>41613 (8733)</td>
<td>39792 (9738)</td>
<td>0.392‡</td>
</tr>
<tr>
<td>Median (range)</td>
<td>23900 (2400-130000)</td>
<td>18930 (1500-259000)</td>
<td></td>
</tr>
<tr>
<td><strong>JAK2 R683 mutations</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>0</td>
<td>10</td>
<td>0.008†</td>
</tr>
<tr>
<td>No</td>
<td>20</td>
<td>22</td>
<td></td>
</tr>
</tbody>
</table>

†According to Fisher's exact test
‡According to Mann-Whitney U test
Figure Legends

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 x-axis 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: BCR-ABL; 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. (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: MLL-AF4;

**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 $3.7 \times 10^{-11}$, $2.2 \times 10^{-7}$, $5.6 \times 10^{-5}$ for AIEOP (i), BFM (ii) and ICH (iii), respectively.

Abbreviations: DS: Down Syndrome ALL; HD: High Hyperdiploid; TEL: TEL-AML1; 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 bars- patients 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 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 boundaries, 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) 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.


(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.
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
FIGURE 4

A

B

C

i

CRLF2  

IL3RA  

ASMTL  

P2RY8

SLC25A6

CRLF2  

P2RY8

ex6  ex5  ex4  ex3  ex2  ex1  ex1

ii

P2RY8F01

GAAAGGCAACCCCTGTGGCGCC  

CACGAGACCTTCTTCAAGAAGCTT

TGATGACCACAGGCTTGGCAACCTGCTCGCTTGGCCCCCGAAGCTCC

CGCTCTGACTGACGCCGTCTCAGTTAGCTGGACTGGCGCTCTCAGTT

CTGAGACCGGTCTCTGCTCCTGCTCTCGCTTGGGAGCTGCCCTTTCC

TGCACAGCATGGGGAGCCTGTTGGCTGGCCTCGTGGGAAGCTGCCCTTCC

CRLF2R03

TTTCCTGCTGGAAGGCTGGATGGCTTGGGCGAGGAGGAGGAGCA

iii

A  B  C  A  B  C

1  2  3  1  2  3

474 bp  219 bp

390 bp

iv

AC6CACCCCGCTTACCCCGAACTGCTGGGATGGCTTATATCC
FIGURE 6

A

DNA- Diagnosis

Wild-type sequence

DNA- Remission

B

Expression of the mutations

RNA- Diagnosis

C

CRLF2

E

BaF3 Jak2 wt CRLF2 F232C

BaF3 Jak2 wt CRLF2 wt

Jak2

p-Jak2

STAT5

p-STAT5

FLAG

CRLF2

α-tubulin
FIGURE 7

- CRLF2
- Mutated JAK2 R683
- Wt JAK2
- Other kinase mutations?
- Mutated CRLF2

Arrows indicate:
- 60% to CRLF2
- 30-50% to Mutated JAK2 R683
- 40-60% to Wt JAK2
- 10% to Mutated CRLF2
- Other kinase mutations?
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, Giovanni Cazzaniga, Maike Schmitz, Jane Chalker, Ruth Shiloh, Ilaria Iacobucci, Chen Shochat, Sharon Zeligson, Gunnar Cario, Martin Stanulla, Sabine Strehl, Lisa J Russell, Christine J Harrison, Beat Bornhauser, Akinori Yoda, Gideon Rechavi, Dani Bercovich, Arndt Borkhardt, Helena Kemptski, Geertruy te Kronnie, Jean-Pierre Bourquin, Eytan Domany and Shai Izraeli