Mechanisms of Notch signaling specificity in lymphocytes and their leukemic counterparts
Gentek, R.

Citation for published version (APA):
Chapter 3

Identification of Innate Lymphoid Cell 2 like leukemia in T-ALL and AML cohorts

Rebecca Gentek¹, Ruud Delwel², Derk Amsen¹,³

¹Department of Cell Biology and Histology, Academic Medical Center, Amsterdam, The Netherlands, ²Department of Hematology, Erasmus University Medical Center, Rotterdam, The Netherlands, ³Department of Hematopoiesis, Sanquin Research and Landsteiner Laboratory, Amsterdam, The Netherlands

Manuscript in preparation
Abstract
Innate lymphoid cells (ILC) constitute a newly discovered family of immune cells with functions in border patrol, tissue homeostasis and organogenesis. Notch signaling controls differentiation of various ILC subsets and also regulates differentiation of T cells. Strong Notch signaling in thymic CD34+ progenitor cells favors generation of group 2 ILC (ILC2) over generation of T cells. Notch is also a major oncogene involved in various types of leukemia, most notably T-ALL. We hypothesized that the strong Notch signals resulting from oncogenic mutation could lead to the generation of a thus far unknown type of leukemia, consisting of ILC2 like cells. We demonstrate that leukemia derived oncogenic NOTCH1 mutants induce ILC2 differentiation from human thymic progenitors. Furthermore, through bioinformatic analysis of gene expression data bases, we identified candidate ILC2 malignancies in cohorts of T-ALL and, strikingly, AML patients. Notably, we identified a group of CEPBA silenced, atypical AML cases that bear activating mutations in NOTCH1 and exhibit an ILC2 gene expression signature. Mechanistic differentiation studies therefore led us to the identification of a new type of leukemia, which we propose to name ILC2-LL (ILC2-like leukemia).

Keywords
Leukemia, innate lymphoid cells (ILC), ILC2, gene expression profiling, Notch
Identification of innate lymphoid cell like leukemia

**Introduction**

Many types of leukemia, including acute myeloid leukemia (AML) and T-cell acute lymphoblastic leukemia (T-ALL), display a large degree of heterogeneity. Different subtypes are often associated with distinct clinical behavior, underlining the importance of correct classification for treatment outcome. Despite considerable progress, largely facilitated by genome wide expression profiling, some hematopoietic malignancies, including biphenotypic leukemia and/or leukemia of ambiguous lineage, remain difficult to classify. Enhanced understanding of the pathobiological events underlying different leukemia subsets will support more accurate classification and thus, ultimately, therapeutic decision making.

A signaling pathway recurrently deregulated in various types of leukemia is Notch. Physiological Notch signaling is activated upon engagement of a Notch receptor by a ligand of the Delta (Dll) or Jagged (Jag) families. Notch proteins are heterodimeric single pass transmembrane proteins that are non-covalently held together by an extracellular heterodimerization (HD) domain. Ligand binding induces a series of proteolytic cleavages catalyzed by ADAM (A Disintegrin And Metalloproteinase) metalloproteases and the γ-secretase (GS) complex. The latter mediates the final cleavage step and can be efficiently blocked by small molecule inhibitors known as γ-secretase inhibitors (GSI). GS mediated cleavage releases the Notch intracellular domain (NICD), which translocates to the nucleus and activates target gene expression through association with the transcription factor CSL. The turnover of this complex and thus, signaling duration, is regulated by phosphodegron sequences in the C-terminal PEST (Proline (P), Glutamate (E), Serine (S) and Threonine (T) rich) domain of the NICD.

Aberrant Notch signaling can have both tumor suppressive and oncogenic functions in hematopoietic cells. A tumor suppressive role for Notch has been described for chronic myelomonocytic leukemia (CMML), B cell ALL (B-ALL) and, more recently, AML. On the other hand Notch acts as a potent oncogene in T-ALL. The human NOTCH1 homologue was initially identified as the chromosomal breakpoint of a rare chromosomal translocation ((t(7;9)(q34;q34.3)) in T-ALL, which results in the expression of a constitutively activated NICD1. Ectopic expression of NICD1 induces T-ALL in a bone marrow transfer model, and mutations activating NOTCH1 were found in more than 60% of human and murine T-ALL, identifying NOTCH1 as the major oncogene in this disease. Two main groups of mutations have been described: while mutations affecting the HD domain render the heterodimer instable and thus induce ligand independent signaling, mutations that truncate the PEST domain stabilize the NICD and therefore increase signaling endurance. HD and PEST domain mutations can synergize when they occur in cis.

The dominant oncogenic role of the pathway in T-ALL reflects its importance...
during normal T cell development, where Notch signals are critically required at multiple stages\textsuperscript{21–24}. In addition, Notch signaling has also been implicated in the development of innate lymphoid cells (ILC)\textsuperscript{24}. These cells derive from the common lymphoid progenitor (CLP) and express the IL-7-R\(\alpha\) chain, identifying them as lymphocytes, yet they lack rearranged antigen receptors and are therefore considered part of the innate immune system\textsuperscript{25,26}. ILC are involved in immune protection, most notably at border surfaces, and contribute to lymphoid tissue homeostasis and repair after infections\textsuperscript{26}. With the exception of the prototypic members, Natural Killer (NK)\textsuperscript{27,28} and lymphoid tissue inducer cells (LTi)\textsuperscript{29,30}, ILC have only recently been described.

Notch signaling regulates development of different ILC lineages, including ILC2\textsuperscript{31}. These cells provide protective immunity against helminths\textsuperscript{32}, mediate eosinophil homeostasis\textsuperscript{33} and respiratory tissue repair following viral infections\textsuperscript{34}. ILC2 have been implicated in a number of pathological conditions, including allergic asthma\textsuperscript{35–39}. Development and function of ILC2 depend on the transcription factors GATA3\textsuperscript{40–42} and ROR\(\alpha\)\textsuperscript{31,43}. We have previously demonstrated that Notch can induce thymic progenitors to differentiate into both T cells and ILC2 in a dosage dependent manner\textsuperscript{44}. In particular, strong Notch signals favor ILC2 development\textsuperscript{44}. On the basis of this, we hypothesized that oncogenic activation of Notch, (which characteristically results in exaggerated Notch signaling) in CLP may result in the generation of ILC2 like leukemia. Malignancies of ILC2 have not been described thus far. Considering that these cells have only recently been identified, it is conceivable that such malignancies exist, but have been misclassified as other lineages. Given the close developmental and genetic relationship of ILC2 and early T cell progenitors\textsuperscript{31,44,45}, we expected that candidate ILC2 leukemia would be contained within (immature) T-ALL cohorts.

In this study, we tested this hypothesis. We found that leukemia derived mutant NOTCH1 alleles indeed elicit sufficiently strong signals in human thymocytes to induce differentiation of ILC2 \textit{in vitro}. Using a bioinformatics approach, we analyzed the expression of genes associated with the ILC2 lineage in gene expression data from multiple series of primary leukemia samples. This approach allowed us to identify leukemias resembling ILC2 in different T-ALL cohorts. Furthermore, using Notch pathway activation as a selection criterion, we were able to identify candidate ILC2 leukemias across multiple AML cohorts, where activating mutations in Notch are rare. Collectively, the data presented here demonstrate that leukemic entities resembling malignant ILC2 exist in cohorts of T-ALL and AML patients and suggest that Notch signaling might be causally involved in their generation. We propose to name these malignancies ILC2-LL (ILC2 like leukemia).
Results

Oncogenic NOTCH1 mutants induce ILC2 development in human thymocytes.

Activation of Notch signaling induces ILC2 differentiation in murine and human thymic progenitors. We have recently demonstrated that human CD34+CD1a−, which have not yet committed to the T cell lineage, develop into T cells or ILC2, depending on the strength of the Notch signal. High dosage of Notch signals favor ILC2 differentiation, and ectopic expression of NICD1 can induce robust ILC2 differentiation even in T cell committed CD34+CD1a+ precursors. Since strong activation of Notch signaling has a potent oncogenic effect in CLP, we hypothesized that oncogenic Notch activation might give rise to a hitherto unknown type of leukemia, resembling ILC2.

To study the likelihood of Notch mediated ILC2 like leukemia, we made use of previously described T-ALL derived, oncogenic mutants of NOTCH1. The mutant NOTCH1 allele P12, established from the T-ALL line P12-Ichikawa, carries a 13 amino acid insertion which reduplicates the S2 cleavage site, allowing for spontaneous cleavage by ADAM metalloproteases and thus, ligand independent Notch signaling. In addition, the C-terminal PEST domain is truncated in the mutant P12ΔPEST allele as a consequence of a premature stop codon induced by a frame shift, a mutation derived from the T-ALL line ALL-SIL. The deleted residues encompass a recognition site of the FBXW7 E3 Ubiquitin ligase and phosphorylation sites for Cyclin dependent kinase 8. Deletion of the PEST domain therefore increases the stability of NICD1, and further enhances Notch signaling when combined with the P12 HD mutation in cis.

We confirmed that P12 and P12ΔPEST induce constitutive Notch signaling, as expression of these alleles results in activation of two different Notch responsive promoters in the absence of ligand stimulation. Consistent with their respective mutational status, transcriptional activation by P12ΔPEST is stronger than observed with the P12 mutation alone. The mutant NOTCH1 alleles are also functional in primary human thymocytes, as demonstrated by enhanced surface expression of CD7, which is a sensitive measure for Notch dosage. Furthermore, P12 and P12ΔPEST promote thymocyte proliferation in a signal strength dependent manner. Ectopic expression of P12 and P12ΔPEST induced development of CD4+CD8+ double positive (DP) T cells in uncommitted CD34+CD1a− thymic progenitors cultured with OP9 stromal cells, as expected. However, not all cells developed into T cells in these cultures. ILC2 are characterized by the absence of markers associated with other hematopoietic lineages, including CD1a, CD3, TCRαβ, TCRγδ (all T cell markers), CD19 (a B cell marker), CD11c, CD14, CD94, CD123, FceR1 and...
A

<table>
<thead>
<tr>
<th>NECD</th>
<th>NICD</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGF repeats</td>
<td>LNR HD-N</td>
</tr>
<tr>
<td>HD-C RAM ANK PEST</td>
<td></td>
</tr>
</tbody>
</table>

P12
P12ΔPEST

B

HES1

[Graph showing luciferase activity for HES1]

CBF1

[Graph showing luciferase activity for CBF1]

C

CD7

[Flow cytometry graph for CD7 with events [% of max]]

D

CellTrace Violet

[Flow cytometry graph for CellTrace Violet with events [% of max]]

E

CD4+CD8+ [% of transduced]

[Graph showing CD4+CD8+ % of transduced cells]

F

CBF1+ [% of transduced lineage-]

[Graph showing CBF1+ % of transduced lineage- cells]
Figure 1: Leukemia-derived mutant NOTCH1 alleles induce ILC2 differentiation in human thymic progenitors. (A) Cartoon representation of the protein domain structure of human NOTCH1 and mutations in the leukemic alleles used in this study. NECD: Notch extracellular domain, NICD: Notch intracellular domain, EGF repeats: epidermal growth factor like repeats, LNR: LIN-12/Notch repeats, HD(-N/C): (N/C-terminal) Heterodimerization domain, RAM: RBPJ association motif, ANK: Ankyrin repeats, PEST: Proline (P), Glutamate (E), Serine (S) and Threonine (T) rich region. The sites of mutation are indicated by grey (P12) and black (P12ΔPEST) triangles. Modified after[20,47]. (B) Activation of Notch responsive luciferase reporter genes by leukemic NOTCH1 mutants. U2OS cells were transfected with a mixture of either HES1-luciferase (left) or CBF1-luciferase (right) reporters, a plasmid constitutively expressing Renilla luciferase and MigRI empty vector, NOTCH1 P12-MigRI (grey bars) or NOTCH1 P12ΔPEST-MigRI (black bars). Firefly luciferase activities were normalized to Renilla luciferase activities and these are displayed as fold of the empty vector control. Data are representative of two independent experiments performed in triplicate and shown as mean + SD. (C, D) Primary human CD34+CD1a-thymocytes were transduced with MigRI retroviruses encoding GFP only (empty vector, grey filled histogram), or GFP and NOTCH P12 (dark grey) or NOTCH1 P12ΔPEST (black), respectively. After transduction, thymocytes were subjected to co-culture with OP9 cells. Data are representative of two individual experiments. (C) Flow cytometric analysis of CD7 surface expression by transduced (GFP+) thymocytes after 4 days of co-culture. Light grey filled histogram represents isotype matched control antibody staining. (D) Prior to transduction, thymocytes were labeled with a proliferation dye (CellTrace Violet, Invitrogen). After one week of co-culture with OP9 cells, dilution of this dye by transduced (GFP+) cells was measured flow cytometrically. Freshly labeled, undivided cells are shown as light grey filled histogram. (E, F) Human CD34+CD1a-thymocytes were retrovirally transduced with MigRI empty vector (white bars), NOTCH1 P12-MigRI (grey bars) or NOTCH1 P12ΔPEST-MigRI (black bars) and cultured on OP9 cells for two weeks in the presence of IL-7 only (F) or IL-7 and Flt3l (E). The fraction of double positive (CD4+CD8+) T cells (E) or ILC2 (defined as lineage− (CD1a−CD3−CD4−CD8−CD11c−CD14−CD19−CD34−CD123−TCRaβ−TCRγδ−BDCA2−FcεRI−) CRTH2+ cells) was determined by flow cytometry. Cumulative data from three (E) or two (F) experiments are shown as mean + SD.

BDCA2 (all myeloid cells markers). Collectively, these markers are referred to as ‘lineage’. Furthermore, these cells express CRTH2 (Chemoattractant receptor homologous molecule expressed on Th2 cells). Ectopic expression of P12 and P12ΔPEST resulted in the generation of a lineage−CRTH2+ ILC2 population (Figure 1F), which was similar in magnitude to the population of T cells in these cultures (Figure 1E). Earlier experiments had shown that such cells also express other markers of ILC2, including IL-7-Rα (CD127), IL-2-Rα (CD25), ICOS, RORα, IL-5, IL-13 as well as the receptors for TSLP, IL-25 and IL-33[44]. Thus, leukemia derived constitutively active NOTCH1 alleles induce development of ILC2, supporting the notion that aberrant Notch signaling might be involved in possible ILC2 malignancies.

Identification of candidate ILC2 malignancies requires further analyses in addition to immunophenotypical criteria.

As a first approach, we analyzed material obtained from two different tissue banks (Sanquin Blood Bank, VUMC Free University Medical Center; both Amsterdam, the Netherlands). We selected 13 patients with immature phenotypes that were denoted difficult to classify based on standard diagnostic immunophenotyping. Samples of these candidates were analyzed by flow cytometry for the expression ‘lineage’ markers (see above), markers
expressed by hematopoietic stem progenitors cells (CD34), as well as markers which have been reported to be expressed by ILC2 (CRTH2, IL-7-Rα, CD7, CD25, ICOS, CD161, cKit). CD45 was included in the analyses to identify leukemic blast cells, which were defined as CD45\textsuperscript{dim} (Supplementary Figure 1A)\textsuperscript{53}. Phenotyping for one patient is shown as an example in Supplementary Figure 1B. Although some patients largely lacked ‘lineage’ markers and did express lymphoid antigens such as the IL-7-Rα chain and CD7, none of the 13 patients analyzed phenotypically resembled ILC2 (Supplementary Table 1). We reasoned that our failure to identify ILC2 like leukemias using this strategy might be caused by low frequencies, consistent with the frequencies of their physiological counterpart under homeostatic conditions. We therefore resorted to an approach allowing more high throughput screening.

**Gene expression analyses for the identification of candidate ILC2 leukemia.**

The most potent and reliable method to identify ILC2-LL would involve large scale gene expression profiling. To this end, we defined a set of genes related to the ILC2 phenotype and analyzed publically available datasets from genome wide expression studies on samples from leukemia patients. As described above, normal ILC are characterized by the lack of markers expressed by other lineages (amongst these, CD11c, CD14, CD19, CD94, CD123, BDCA2), but they do express the IL-7-Rα chain, and have often been found to express CD161 and cKit\textsuperscript{25}. Furthermore, ILC2 express lineage specific transcription factors (RORA, GATA3), cytokine receptors (IL2RA, IL17RB, ST2) and additional surface markers (ICOS, CRTH2). Thus, the ILC2 gene set is comprised of genes whose expression should be absent or low in candidate ILC2-LL, as well as genes which should (abundantly) be expressed. ILC respond to pathogens with the production of lineage specific cytokines. Malignant ILC2 like cells might thus also express transcripts for type 2 immunity related cytokines (IL-4, IL-5, IL-13), which were therefore included in the analysis.

To test the applicability of this gene set and avoid bias towards a certain type of leukemia, we initially analyzed mixed patient datasets, including a large (n = 2004) dataset from the MILE (Microarray innovations in leukemia) study containing patients diagnosed with T-ALL, B-ALL, AML, CML and CLL (GSE 13519)\textsuperscript{54} and a mixed T-ALL/B-ALL cohort (GSE 10255)\textsuperscript{55}. Indeed, leukemia samples resembling ILC2 could be identified in mixed cohorts by intensity based clustering for expression of the genes contained within the ILC2 gene set as can be appreciated for the mixed T-ALL/B-ALL cohort (Figure 2A): The heat map representation displays log transformed expression values of the indicated genes across all samples of the dataset, and reveals clear clustering of multiple subgroups. Patients in of these clusters largely lack expression of genes encoding ‘lineage’ markers (including myeloid genes such as CD11C, CD14, CD19, CD94, CD123) displays (high) expression of CRTH2, GATA3,
CD3E, KIT and IL7RA and some of the samples in this cluster also express transcripts for cytokines (IL-4, IL-13) and cytokine receptors (IL17RB, IL2RA, ST2) associated with the ILC2 lineage (Figure 2A, candidates annotated as red dots). We confirmed these findings using a different type of analysis: Traditional clustering represents multidimensional data and can thus complicate the identification of small groups, especially in a forward approach like we have taken here. Principle component analysis (PCA) reduces data dimensions by identifying directions, so-called principle components, which describe the maximum variation in the data, and are used as surrogates to represent each sample. Principle components are uncorrelated in that they may represent different features of a given sample, which makes PCA a powerful tool to visualize similarities and differences between samples and identify subgroups. Indeed, when we performed PCA for the ILC2 gene set in the mixed T-ALL/B-ALL dataset in which we had identified a subgroup...
of candidate ILC2 patients by intensity based clustering (Figure 2A), this
group (shown as red dots) could be separated from the rest of the patients
(shown as black dots) (Figure 2B). Considering the close genetic proximity of
ILC2 and developing T cells\textsuperscript{45}, we postulated that candidate ILC2 leukemia
might be specifically found within T-ALL cohorts, in particular within
groups of immature, early T cell precursor (ETP) T-ALL. However, when we
analyzed the ILC2 gene set in several T-ALL datasets using intensity based
clustering (data not shown) and PCA (two representative datasets shown in
Figures 2C and 2D) we could not reliably identify (sub)groups of candidate
patients. It is possible that this reflects specific characteristics of the selected
datasets rather than these diseases, although we consider this unlikely, as
we have screened multiple different cohorts, some of which were large in
size. It is also conceivable that the frequency of ILC2 malignancies is still too
low within these cohorts to allow generation of a clear cluster. However, a
fundamentally different possibility is that the gene expression programs of
ILC2 and early T cells are too close to allow clear separation based on the
limited set of genes in the ILC2 set used here. Indeed, ILC2 express multiple
genes which have previously been associated with an ‘early T progenitor’
signature\textsuperscript{45}. Importantly, Notch pathway activation, which we hypothesize
to be relevant to ILC2 leukemia, is a prominent feature also of T-ALL\textsuperscript{18} and
can therefore not be used to distinguish malignancies of T cell precursors
from putative ILC2 neoplasms.

\textbf{Early T progenitor signature and Notch pathway activation are powerful
distinguishing criteria for candidate ILC2 leukemia within (atypical) AML
cohorts.}

Notch signaling is now generally believed to have a tumor suppressor
function in myeloid cells\textsuperscript{10,14}. In line with this, activating Notch mutations
are rare\textsuperscript{58}. However, the few activating Notch mutations described in AML
 correlate with poor prognosis\textsuperscript{59} and seem to be limited to unusual cases of
multilineage or ‘lineage switch’ leukemia\textsuperscript{58,60}. It is tempting to speculate that
such entities contain candidates for ILC(2) like leukemia.

To further address this possibility, we expanded our bioinformatic approach
by analyzing additional gene sets in cohorts of (atypical) AML patients. As
mentioned above, ILC2 express a number of genes which have long been
regarded as characteristic for (early) T cell precursors, including \textit{Lck}, \textit{Lat} and
\textit{Tcf7}\textsuperscript{45}. In contrast to T-ALL datasets, analysis of these genes might improve
selection of ILC2 like candidates in cohorts of AML malignancies, where
these signatures should be restricted to atypical cases. Indeed, expression
analysis of ‘(early) T cell progenitor’ genes (\textit{BCL11B}, \textit{CD1A}, \textit{CD3E}, \textit{CD3D},
robustly separates subgroups within AML datasets and identifies a clear
cluster highly expressing these genes (Figure 3A, see cluster on the right
side).
We further aimed to analyze AML datasets for signs of Notch pathway activity. Target gene expression by Notch is critically dependent on the cellular context, the molecular basis of which is currently under intense investigation\textsuperscript{61,62}. Defining Notch pathway gene sets therefore is not trivial. Notch has been studied extensively in the hematopoietic system, though, in particular during lymphocyte development\textsuperscript{24}, where a number of robust Notch targets have been described.

Figure 3: Identification of candidate ILC2 like leukemia in an AML cohort. (A, B, C) Samples from a cohort of AML patients (GSE 12417) were analyzed for the expression of (A) genes associated with (early) T cell progenitors, (B) selected target genes of the Notch signaling pathway and (C) genes characterizing ILC2. Data are hierarchically clustered according to intensity and shown as heat maps. Log transformed (base 2), mean centered expression data are represented by color codes (bright green: -3, black: no change to mean, bright red: 3). Individual samples are shown in columns, different genes as annotated on the left are shown as rows. ILC2-like candidates are marked with red dots above the heat map. (D) Meta analysis of the gene classifiers used in (A, B, C). The gene set map was created by simultaneously analyzing all samples for the different gene sets and is displayed as intensity based clusters. Notch: Notch target gene set, ILC2: gene set with ILC2 characteristic genes, (early) T: classifier based on genes associated with (early) T cell progenitors. Average gene set intensity is displayed in color code, ranging from (bright blue: -1,5, bright red: 1,5).
We selected a number of genes strongly and positively associated with Notch pathway activity in hematopoietic cells (DTX1, DTX3, DTX3L, DTX4, HES1, HES5, TRIB2, NOTCH1, NOTCH3) and performed gene expression analyses in AML datasets (Figure 3B). Indeed, expression of these genes was observed in subsets of AML patients across multiple datasets (see Figure 3B, clusters on the right side). We proposed that samples displaying expression of genes associated with T cell progenitors and Notch pathway activity are candidates for ILC2 like leukemia, which we then tested by analyzing expression of the ILC2 gene set also used in T-ALL datasets (Figure 3C). This approach identified a number of possible ILC2 like malignancies, which, based on the analysis of ILC2 related genes alone, often scattered in small groups throughout cohorts (indicated as red dots in Figure 3C). However, importantly, the combined analysis of all three characteristics, expression of genes related to T cell precursors, Notch activity and ILC2, robustly identifies subgroups of candidate patients, as indicated by clustering in gene set maps combining these gene sets (Figure 3D, see cluster on the left). These candidates co-expressed the genes shared by ILC2 and early T cell progenitors, but also expressed genes selectively associated with the ILC2 phenotype, which have not been found on immature T cells (for instance CRTH2, IL-13, IL-5). These gene expression patterns therefore support the notion that these malignancies resemble ILC2-LL more closely than T-ALL.

Using this approach, candidates for ILC2 like leukemia were found in 9 independent datasets from different AML cohorts. In contrast to T-ALL cohorts, the combined analysis of these gene sets appears to be a powerful tool to stratify (atypical) AML cases and select candidates for ILC2-LL. Although the mutational status of NOTCH1 and Notch pathway components is not known for these candidates, our analyses further support the notion that Notch signaling might be activated in ILC2 malignancies, and provide a distinguishing criterion from myeloid leukemia.

**Activating mutations in NOTCH1 characterize candidate ILC2 like leukemia within a group of CEBPA silenced AML.**

A recent study has described a particularly interesting subset of biphenotypic AML, which was initially found in a cohort of 285 patients, and later confirmed in an additional dataset, indicating that this group indeed comprises a distinct entity.\(^\text{63}\)

These cases show simultaneous expression of myeloid (CD13, CD33) and T lymphoid markers (CD7, partially cytoplasmic CD3) and genes. However, they were tested negative for expression of CD1a, CD4, CD8, TCRαβ and TCRγδ.\(^\text{63}\)

Their gene expression profile resembles that of AML cases carrying mutations in the myeloid transcription factor C/EBPα (CCAAT/enhancer binding protein alpha, encoded by CEBPA), yet, these cells lack CEBPA mutations.
Figure 4: Mutational activation of NOTCH1 identifies candidate ILC2 leukemia within a CEBPA silenced subgroup of AML patients. A subgroup of patients from a large cohort of 460 AML cases (GSE 6891) was preselected based on a genome wide signature corresponding to a CEBPA mutated phenotype. Seven of these patients have previously been shown to not bear CEBPA mutations, but display CEBPA inactivation caused by epigenetic silencing. Mutational status of CEBPA in these samples is displayed on top, mut: mutated, sil: silenced. Intensity basted clustering is shown in heat map representation. Log transformed (base 2) gene expression data were mean centered and expressed in color codes (bright green: -3, black: no difference to mean, bright red: 3). Columns correspond to individual samples and rows represent genes as indicated on the left. (A) Analysis of a 5 gene signature previously demonstrated to predict CEBPA mutated and CEBPA silenced subgroups. (B) Samples were analyzed for the expression of a gene set associated with (early) T cell progenitors. (C, D, E) Known activating mutations in NOTCH1 are indicated as stars. (C) Analysis of selected Notch target gene expression. (D) Gene expression analysis using the ILC2 classifier. (E) Geneset map constructed from the meta analysis of the classifiers used in (B), (C) and (D). ILC2: ILC2 gene set, Notch: Notch target gene set, (early) T: (early) T cell progenitor classifier. Average gene set intensity is color coded (bright blue: -1,5, bright red: 1,5).
Instead, these cells specifically bear activating mutations in NOTCH1, causing enhanced levels of TRIB2 (Tribbles homolog 2), which in turn epigenetically silences CEBPA\(^{63,64}\). Therefore, this group is referred to as CEBPA silenced AML. Collectively, the CEBPA silenced AML subgroup contains promising candidates for ILC2 like leukemia.

To test this, we analyzed expression of our gene sets in these leukemia samples. We preselected cases of CEBPA mutated and CEBPA silenced AML based on the published report\(^{63}\) and confirmed their identity by analyzing a small set of genes which has been described to be predictive for the CEBPA silenced subgroup\(^{63}\) (Figure 4A). CEBPA mutated patients express CEBPA and CTNNA1, but lack expression of the lymphoid marker CD7, NOTCH1 and TRIB2. Conversely, CEBPA silenced AML are characterized by expression of CD7, NOTCH1 and TRIB2, but mostly lack expression of CEBPA and CTNNA1.

We then performed gene expression analyses of our gene sets associated with (early) T cell progenitors (Figure 4B) and Notch pathway activity (Figure 4C). Intensity based clustering of both gene sets robustly separated the CEBPA silenced from the CEBPA mutated group, which was remarkable, considering that these two groups were initially found to cluster together on a genome wide scale\(^{63}\). Importantly, these two groups could also be separated according to the expression of genes associated with ILC2 (Figure 4D). Thus, as illustrated by the meta analysis of the three different gene sets used in this study (Figure 4E), the CEBPA silenced AML subgroup could be identified as candidate ILC2 like malignancies.

Importantly, although activating mutations in the HD and/or PEST domain of NOTCH1 were found in these cells, CEBPA silenced leukemia significantly differ from T-ALL both genetically and epigenetically in that their genome wide DNA methylation profiles are distinct\(^{65}\). Interestingly, this was true in comparison with representative T-ALL cases of various differentiation stages ranging from very immature to mature, strongly suggesting that these are not misclassified T-ALL cases, but might resemble an entirely different leukemic entity, such as ILC2 like leukemia.

Discussion

ILC are a hematopoietic family of which most members have only recently been described. Despite their recent discovery, ILC have already been implicated in a variety of pathological conditions such as allergic asthma\(^{66}\), inflammatory bowel disease\(^{67}\) and atopic dermatitis\(^{68}\), reminiscent of their roles in protective immunity at border surfaces and (lymphoid) tissue remodeling. However, (rare) neoplasms resulting from the transformation and uncontrolled expansion of developing ILC have thus far only been described for the prototypic member of the ILC family, NK cells\(^{69}\). Malignancies caused
by transformation of other ILC probably exist, as for any other hematopoietic lineage, but have likely not been identified as such, given that their normal counterparts have only just been described.

Notch appears to be commonly involved in development of different ILC lineages, and aberrant Notch signaling has been implicated in both solid and hematological malignancies, exerting oncogenic as well as tumor suppressive functions. It appears that the role of Notch in transformation of a certain tissue reflects its role during development of this particular cell type. Thus, Notch is a tumor suppressor in B-ALL, while it is a powerful oncogene in T-ALL. Activating mutations in NOTCH1 are in fact the most common genetic lesion in T-ALL. Considering that strong Notch activation induces ILC2 development, we hypothesized that Notch signaling, as elicited by oncogenic mutations, might be involved in generation of ILC2-LL. Conversely, we hypothesized that Notch activation might be a criterion by which candidate ILC2-LL can be identified.

In this study, we demonstrated that T-ALL derived oncogenic NOTCH1 alleles do induce ILC2 differentiation from primary human thymocytes. These observations further support the notion that Notch activation could lead to ILC2 progenitor transformation and prompted us to investigate approaches by which possible ILC2 like malignancies could be identified.

The identification of candidate ILC2 leukemia is hampered by several factors: ILC have generally mostly been defined by what they are not, i.e. the lack of rearranged antigen receptors and ‘lineage’ markers. Lineage characteristic factors, such as CRTH2 for ILC2, are by no means exclusively expressed by ILC2, but also by other cell types, such as basophils. Not surprisingly, therefore, manually screening tissue banks for candidate patients is laborious and highly inefficient.

Therefore, we resorted to large scale genetic analyses of genome wide expression data from leukemic cohorts. Although human ILC have not fully been characterized genetically, and most of what is known stems from studies in mice, we were able to define a gene set containing both genes which should be (highly) expressed by ILC2 like malignancies, as well as genes that should not (abundantly) be expressed. Using this approach, we were able to identify candidate ILC2-LL across different cohorts of (T-)ALL and AML.

It was recently shown that ILC2 share expression of genes which were classically attributed to T cell progenitors. We demonstrate that expression of such genes enables further subgroup stratification in AML datasets, where these genes are normally not expressed. Most notably, however, we show that the combined analysis of genes associated with ILC2 and (early) T cell progenitors as well as Notch pathway activity provides powerful means to identify candidate ILC2 like malignancies in AML cohorts. These findings are in line with a possible oncogenic role for Notch in ILC2 progenitors.
Although our studies have identified groups of leukemias genetically resembling ILC2, it will still be important to determine to what degree these cells resemble functional ILC2. Immunophenotypic as well as functional analyses (responses to IL-25, IL-33, TSLP and secretion of signature cytokines) should be informative for this purpose. Furthermore, also the genetic analyses proposed in this study require further refinement, especially for T-ALL datasets. Gene expression signatures from physiological ILC subsets and early T cell subsets will provide comparison gene sets, which will optimize the distinction between closely related (early) T cell progenitors and ILC2 and their respective malignant counterparts.

It is important to note that the CEBPA silenced leukemic entity which we here identify as ILC2-LL expresses the myeloid antigens CD13 and CD33, indicating that this group might still represent bilineal malignancies. Expression of CD13 and CD33 has to our knowledge not been tested in normal human ILC2. Aberrant expression of myeloid antigens might stem from the transformation of early progenitors. CLP and early thymic progenitors retain myeloid potential\(^{70,71}\), and, interestingly, a CD13\(^{+}\)CD33\(^{+}\) myeloid like population can give rise to mature functional NK cells, another ILC family member\(^{72}\). It stands to argue, however, that strong oncogenic stimuli, such as Notch activation, can even override lineage fidelity in later progenitors, perhaps by mobilizing differentiation programs from epigenetically poised loci. In line with this, we have recently demonstrated that strong, sustained Notch activation can induce ILC2 development from T cell committed CD34\(^{+}\)CD1a\(^{+}\) thymic progenitors\(^{44}\). Generally, there is increasing evidence that lineage specification is a much more plastic process than previously anticipated, for instance in T helper cells, but likely also in ILC\(^{73-75}\). Thus, lineage infidelity is an emerging theme also during normal hematopoiesis, and is likely reflected in leukemias. However, our genetic analyses demonstrate that the CEBPA silenced leukemic entity most closely resembles ILC2. In the light of lineage infidelity, it is important to note that this group is genetically and epigenetically distinct from both typical AML\(^{63}\) (immature) T-ALL samples\(^{65}\). This is at least consistent with the notion that these leukemia are in fact neither malignancies of myeloid, nor immature T cell precursors, but instead, ILC2 progenitors, which might however have retained some myeloid potential causing aberrant expression of CD13 and CD33.

We propose that similar genetic signatures and analyses as the ones used here will also guide the identification of leukemia resembling the other ILC lineages. Interestingly, reminiscent of its role in the differentiation of various T helper cell subsets, Notch signaling has also been implicated in the development and/or expansion of other ILC lineages, including LTi cells\(^{76}\) and NCR\(^{+}\) ILC3\(^{77}\).

In addition to improved candidate selection, profound analyses of such signatures will increase our understanding of the pathobiological
mechanisms underlying their transformation. The data presented here strongly support the notion that mutational activation of Notch is one such mechanism in ILC2 oncogenesis. Activating mutations of the types known from T-ALL have been found in the CEBPA silenced leukemia subgroup, but are very rare in AML overall, consistent with the more general tumor suppressive function of the pathway in myeloid cells\(^{10,14}\). Notch activation might be a distinguishing criterion between \textit{bona fide} AML and misclassified entities. This has immediate clinical relevance, considering that treatment choice considerably differs between distinct types of leukemia. Intriguingly, primary CEBPA silenced leukemic cells failed to expand \textit{in vitro} in response to the myeloid growth factors GM-CSF, G-CSF and IL-3\(^{65}\), and it would be important to test whether these cells can be expanded under lymphoid conditions and/or in the presence of ILC2 stimulating cytokines.

Taken together, we present evidence for the existence of ILC2-LL, which can be identified by gene expression analyses. Our findings show that such cells are characterized by mutational activation of the Notch pathway, presumably reflecting the ability of strong Notch signaling to instruct differentiation of the ILC2 lineage.

**Material & Methods**

**Retroviral expression constructs.** The retroviral MigRI expression vectors encoding mutant human NOTCH1 alleles derived from T-ALL patients were described before\(^{47}\) and kindly provided by Dr. Warren Pear (University of Pennsylvania). NOTCH1 P12 carries an insertion in the C-terminal heterodimerization domain (1722 V/A, 1723 insRLGSKNIPYKIEA), which reduplicates the S2 metalloprotease cleavage site. NOTCH1 P12ΔPEST carries an additional C-terminal deletion of the PEST domain (encompassing aa 2473 to 2555) caused by a premature stop codon, which increases protein stability.

**Reporter gene assays.** The human osteosarcoma cell line U2OS was cultured in IMDM supplemented with FCS (10%), GlutaMax (2mM), penicillin (100 U/ml) and streptomycin (100 µg/ml) (all Gibco). HES1-luciferase\(^{78}\) and CBF1-luciferase\(^{79}\) were used as Notch responsive reporter constructs and have been described before. U2OS cells were transiently transfected using FuGene HD (Promega) according to the manufacturer’s instructions. A mixture containing a reporter construct, a control vector constitutively expressing Renilla luciferase (pRL-CMV) and retroviral plasmids encoding mutant NOTCH1 alleles derived from T-ALL patients. Differences in transfection efficiency were corrected for by normalization of Firefly luciferase activities to Renilla luciferase activities. These relative Firefly luciferase activities were then normalized to the respective empty vector controls and data are displayed as fold of these controls. Approximately 48 hours after transfection, cells were lysed. Luciferase activities were measured using the Dual Luciferase Reporter
Assay System (Promega) on a Synergy HT microplate reader (Syntek).

**Virus production and transduction.** The packaging cell line Phoenix GALV was maintained in IMDM containing FCS (10%), GlutaMax (2mM), penicillin (100 U/ml) and streptomycin (100 µg/ml) (all Gibco). To produce high-titer viruses, Phoenix GALV cells were transfected using FuGene HD (Promega). Virus containing supernatants were harvested 48 hours after transfection, snap frozen on dry ice and stored at -80°C until use. Primary human thymocytes were incubated with virus supernatant in Retronectin coated plates (30µg/ml, Takara Biomedicals) for 6 to 8 hours at 37°C. Transduction efficiencies were measured by flow cytometry 48 hours post transduction.

**Flow Cytometry.** Fixable live/dead dyes and the CellTrace Violet dye (both Invitrogen) were used according to the manufacturer’s instructions to distinguish viable from dead cells and determine cell divisions, respectively. Antibodies specific for the following human antigens were used: CD1a, CD3, CD4, CD7, CD8α, CD11c, CD14, CD19, CD25, CD34, CD45, CD94, CD123, CD127 (IL-7R-α), CD294 (CRTH2), CD303 (BDCA2), TCRαβ, TCRγδ and FcER1. These were directly labeled with Fluorescein Isothiocyanate (FITC), Phycoerythrin (PE), Phycoerythrin-Cyanine 5 (PE-Cy5), PE-Cy5.5, PE-Cy7, PerCP-Cy5.5, Allophycocyanin (APC)/Alexa Fluor 647, APC-Cy7, AF700 (BD Bioscience, Biologend or MACS Miltenyi), Horizon V500 (HV500, BD Bioscience), Brilliant Violet (BV) 421, BV711 and BV785 (all Biolegend). Surface staining was performed on single cell suspensions for 20-30 minutes at 4°C. Flow cytometry was performed on a LSR Fortessa instrument (BD Bioscience) and data were analyzed using FlowJo software (TreeStar).

**Phenotyping of patient material.** Patient samples were obtained from the VUMC (Free University Medical Center) or Sanquin Bloodbank (both Amsterdam, the Netherlands) and thawed under constant, gentle shaking in a waterbath at 37°C. Upon complete thawing, cells were transferred into a fresh vial and fresh medium containing DNase (100µg/ml) and MgCl₂ (10mM) was added dropwise. Cells were allowed to adjust to room temperature for 30 minutes, and harvested by gentle centrifugation (400g). Cells were then washed one more time in fresh medium and subjected to staining and FACS analysis as described above.

**Isolation of thymic hematopoietic progenitors.** Specimens of human postnatal thymus were obtained from pediatric heart surgeries at the LUMC (Leiden, the Netherlands). Use of these tissues was approved by the AMC ethical committee in accordance with the declaration of Helsinki. Cell suspensions were prepared by mechanical disruption of the tissues using a Stomacher 80 Biomaster (Seward). After overnight incubation at 4°C, viable cells were isolated from a Ficoll-Hypaque (Lymphoprep; Nycomed Pharma) density gradient. Thymocytes were then enriched for CD34+ cells by MACS (Miltenyi Biotec), stained with fluorescently labeled antibodies and subsequently sorted by flow cytometry on a FACS Aria instrument.
(BD Bioscience) as CD34+CD1a−CD3−CD56−BDCA2− (here referred to as CD34+CD1a−) to > 99% purity.

**In vitro differentiation of thymic progenitors.** In vitro differentiation of thymic progenitors was performed as described before. Briefly, sorted progenitors were cultured overnight in Yssel’s medium containing 5% normal human serum, SCF (20 ng/ml) and IL-7 (10 ng/ml, both PeproTech) and transduced the following day. After transduction, thymocytes were added to pre-seeded OP9 cells. Co-cultures were maintained in MEMα (Invitrogen) containing FCS (20%, FetalClone I, Hyclone) and IL-7 (5 ng/ml) or IL-7 and Flt3l (5 ng/ml, PeproTech) and were refreshed every 3 to 4 days. Cells were harvested weekly by forceful pipetting, passed through 70μm nylon mesh filters (Spectrum Labs) and analyzed or transferred to fresh OP9 cells.

**Gene expression datasets and analysis.** Publically available datasets of different leukemia cohorts were studied using the R2 microarray analysis and visualization platform (http://r2.amc.nl). The following datasets were acquired from the Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo) or the EMBL-EBI databank (http://www.ebi.ac.uk) and used for this study: T-ALL datasets (GEO accession numbers GSE 8879, 10609, 26713, 42328); AML datasets (GSE 12417, 10358, 17855, 4608, 15210, 21261, 15434, 30377, 6891); Mixed leukemia datasets (GSE 10255, 131519 and EMBL-EBI accession number E-MEXP-313).

**Acknowledgements**

We thank Dr. Warren Pear (University of Pennsylvania) for DNA expression constructs and the surgical team of Dr. Mark Hazekamp from the LUMC for thymus specimens. We thank Marieke von Lindern (Department of Hematopoiesis, Sanquin) for useful suggestions. We acknowledge Ingrid Lommerse and Christa Homburg (Immunocytology lab, Sanquin) as well as Dr. Linda Smit and Dave de Leeuw (VUMC) for providing patient material. We further thank the R2 support team (Department of Oncogenomics, AMC), in particular Richard Volckman, for maintenance of the R2 platform and support with data analyses.

**References**

21. Radtke, F., Wilson, A. & MacDonald, H. R. Notch signaling in T-
30. Mebius, R. E., Rennert, P. & Weissman, I. L. Developing lymph nodes collect CD4+CD3- LTbeta+ cells that can differentiate to APC, NK cells, and follicular cells but not T or B cells. Immunity 7, 493–504 (1997).
37. Klein Wolterink, R. G. J. et al. Pulmonary innate lymphoid cells are major producers of IL-5 and IL-13 in murine models of allergic asthma.
52. Fryer, C. J., White, J. B. & Jones, K. A. Mastermind recruits CycC:CDK8...
68. Imai, Y. et al. Skin-specific expression of IL-33 activates group 2 innate
Supplementary information

Supplementary Table 1: Immuno phenotyping of 13 leukemia patients. Leukemic blast cells from 13 different patients (rows) were analyzed by flow cytometry for surface expression of the indicated antigens (columns). Presence (+) or absence (-) of the markers was scored. Some blasts consisted of heterogeneous populations (indicated as +/-).

<table>
<thead>
<tr>
<th>patient #</th>
<th>cKit</th>
<th>CD161</th>
<th>CD127</th>
<th>CRTH2</th>
<th>CD7</th>
<th>CD25</th>
<th>CD1a</th>
<th>CD3</th>
<th>CD4</th>
<th>CD8</th>
</tr>
</thead>
<tbody>
<tr>
<td>VUMC #1</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+/-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VUMC #2</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VUMC #3</td>
<td>-</td>
<td>-</td>
<td>+/-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Sanquin #1</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+/-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Sanquin #2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Sanquin #3</td>
<td>-</td>
<td>+</td>
<td>+/-</td>
<td>+/-</td>
<td>+</td>
<td>-</td>
<td>+/-</td>
<td>+</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td>Sanquin #4</td>
<td>-</td>
<td>+/-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td>Sanquin #5</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+/-</td>
<td>+</td>
<td>+/-</td>
<td>+</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td>Sanquin #6</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sanquin #7</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sanquin #8</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td>Sanquin #9</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sanquin #10</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>patient #</th>
<th>CD11c</th>
<th>CD14</th>
<th>CD19</th>
<th>CD34</th>
<th>CD94</th>
<th>CD123</th>
<th>BDCA2</th>
<th>FcεRI</th>
<th>TCRαβ</th>
<th>TCRγδ</th>
</tr>
</thead>
<tbody>
<tr>
<td>VUMC #1</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+/-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VUMC #2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VUMC #3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+/-</td>
<td>-</td>
<td>+/-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sanquin #1</td>
<td>+/-</td>
<td>-</td>
<td>-</td>
<td>+/-</td>
<td>-</td>
<td>+/-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+/-</td>
</tr>
<tr>
<td>Sanquin #2</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sanquin #3</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td>Sanquin #4</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Sanquin #5</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td>Sanquin #6</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sanquin #7</td>
<td>-</td>
<td>+/-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sanquin #8</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td>Sanquin #9</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sanquin #10</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Supplementary Figure 1: Phenotyping of candidate leukemia cells. Flow cytometrical analysis of one candidate leukemic patient (VUMC #2, see supplementary table). (A) Gating strategy to identify leukemic blast. Viable cells were pre-gated based on staining with a fixable live/dead dye (Invitrogen) and doublets were gated out based on FSC and SSC. Blast cells were defined as CD45dim cells. (B) Immuno phenotyping of leukemic blast cells as defined in (A). Surface staining was performed with directly labeled antibodies against antigens associated with hematopoietic stem and progenitor cells (CD34), B cells (CD19), T/NK cells (CD1a, CD3, CD4, CD8, TCRαβ, TCRγδ, CD7, CD94), myeloid cells (CD14, CD11c, CD123, FcεRI, BDCA2) and markers previously associated with the ILC2 lineage (CD25, cKit, IL-7-Rα, CD161 and CRTH2). Filled grey histogram: isotype matched control antibody staining, black histogram: leukemic blast cells.
<table>
<thead>
<tr>
<th>CD45</th>
<th>FSC</th>
<th>CD45</th>
<th>blast cells</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Events (% of max)
- CD45
- CD19
- CD34
- CD1a
- CD3
- CD8
- CD4
- TCRαβ
- TCRγδ
- CD94
- CD7
- CD11c
- CD123
- FCεRI
- CD14
- BDCA2
- CD25
- IL-7-Rα
- cKit
- CD161
- CRTH2