



HAL
open science

Linking the KIR phenotype with STAT3 and TET2 mutations to identify chronic lymphoproliferative disorders of NK cells

Cedric Pastoret, Fabienne Desmots-Loyer, Gaelle Drillet, Simon Le Gallou, Marie-Laure Boulland, Alexia Thannberger, Anne-Violaine Doncker, Véronique Salaun, Gandhi Damaj, Richard Veyrat-Masson, et al.

► To cite this version:

Cedric Pastoret, Fabienne Desmots-Loyer, Gaelle Drillet, Simon Le Gallou, Marie-Laure Boulland, et al.. Linking the KIR phenotype with STAT3 and TET2 mutations to identify chronic lymphoproliferative disorders of NK cells. *Blood*, 2021, 137 (23), pp.3237-3250. 10.1182/blood.2020006721 . hal-03134940

HAL Id: hal-03134940

<https://hal.science/hal-03134940>

Submitted on 19 Feb 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Linking the KIR phenotype with *STAT3* and *TET2* mutations to identify chronic lymphoproliferative disorders of NK cells

Tracking no: BLD-2020-006721R2

Cedric PASTORET (INSERM UMR1236, Rennes, France) Fabienne Desmots-Loyer (INSERM UMR U1236, CHU Rennes, France) Gaelle Drillet (CHU de Rennes, France) Simon Le Gallou (INSERM UMR1236, Rennes, France) Marie-Laure Boulland (CHU Pontchaillou, France) Alexia Thannberger (Hôpital de Saint Malo, France) Anne-Violaine Doncker (Hôpital Privé Sévigné, France) Véronique SALAUN (CHU de CAEN Normandie, Normandie université, UNICAEN, France) Gandhi Damaj (Normandy University, Hematology Institute, France) Richard Veyrat-Masson (CHU Clermont-Ferrand, France) Olivier Tournilhac (CHU Estaing, Clermont-Fd, Chelater EA7453, CIC-1405, Université Clermont Auvergne, France) Aline Moignet (CIC1414, Rennes, France, France) Celine PANGAULT (INSERM UMR1236, Rennes, France) Mikael ROUSSEL (INSERM UMR1236, Rennes, France) Thierry Fest (INSERM UMR1236, Rennes, France) Thierry Lamy (INSERM UMR1236, Rennes, France)

Abstract:

Distinguishing chronic lymphoproliferative disorders of NK cells (CLPD-NK) from reactive NK cell expansions is challenging.

We assessed the value of NK receptor phenotyping and targeted high-throughput sequencing in a cohort of 114 consecutive patients with NK cell proliferation, retrospectively assigned to a CLPD-NK group (N=46) and a reactive NK group (N=68). We then developed a NK-clonality score combining flow cytometry and molecular profiling with a positive predictive value of 93%.

STAT3 and *TET2* mutations were respectively identified in 27% and 34% of the CLPD-NK patients - constituting a new diagnostic hallmark for this disease. *TET2*-mutated CLPD-NK exhibited preferentially a CD16^{low} phenotype, displayed more frequently a lower platelet count, and were associated with other hematologic malignancies such as myelodysplasia. To explore the mutational clonal hierarchy of CLPD-NK, we performed a whole exome sequencing of sorted, myeloid, T, and NK cells and identified that *TET2* mutations were shared by myeloid and NK cells in 3 out of 4 cases. Thus, we hypothesized that *TET2* alterations occur early in CLPD-NK disease which could explain a potential link between NK-LGL leukemia and other myeloid malignancies. Finally, we analyzed the transcriptome by RNA-seq of 7 CLPD-NK and evidenced two groups of patients. The first group displayed *STAT3* mutations or *SOCS3* methylation and overexpressed *STAT3* target genes. The second group, including two *TET2*-mutated cases, significantly under-expressed genes known to be down-regulated in angioimmunoblastic T-cell lymphoma. Our results provide new insights into the pathogenesis of NK cell proliferative disorders and potentially new therapeutic opportunities.

Conflict of interest: No COI declared

COI notes:

Preprint server: No;

Author contributions and disclosures: C.Pas., M.R., T.F., and T.L. designed the study; G.D. and T.L. reviewed the clinical files for initial classification; M.R. and C.Pas. analyzed flow cytometry data; S.L.G. performed cell-sorting, C.Pas. and M.L.B. generated and analyzed high throughput sequencing and RNAseq data, F.D. designed and analyzed methylation experiments; C.Pan. managed the sample collection; G.D., A.T., A.V.D., V.S., G.L.D., R.V.M., O.T., A.M., and T.L. provided clinical information; C.Pas. and T.F. wrote the manuscript. All the authors critically revised the manuscript.

Non-author contributions and disclosures: Yes; David Fraser PhD (Biotech Communication SARL, Ploudalmézeau, France) was involved in the copy-editing of this manuscript.

Agreement to Share Publication-Related Data and Data Sharing Statement: For original data, please contact thierry.lamy@univ-rennes1.fr RNAseq data are deposited on GEO.

Clinical trial registration information (if any):

ACCEPTED MANUSCRIPT

Linking the KIR phenotype with *STAT3* and *TET2* mutations to identify chronic lymphoproliferative disorders of NK cells

Short Title: *TET2 mutations in CLPD-NK*

Cédric Pastoret^{1,2}, Fabienne Desmots^{1,2}, Gaëlle Drillet³, Simon Le Gallou^{1,2}, Marie-Laure Boulland¹, Alexia Thannberger⁴, Anne-Violaine Doncker⁵, Véronique Salaun⁶, Gandhi Laurent Damaj⁷, Richard Veyrat-Masson⁸, Olivier Tournilhac⁹, Aline Moignet^{3,10}, Céline Pangault^{1,2}, Mikaël Roussel^{1,2}, Thierry Fest^{1,2*}, Thierry Lamy^{2,3,10*}

Affiliations :

¹Centre Hospitalier Universitaire de Rennes, Laboratoire d'Hématologie, Rennes, France

²Institut national de la santé et de la recherche médicale, Unité mixte de Recherche U1236, Université Rennes 1, Etablissement Français du sang Bretagne, F-35000, Rennes, France

³Centre Hospitalier Universitaire de Rennes, Service d'Hématologie Clinique, Rennes, France

⁴Hôpital de Saint Malo, Service d'Hématologie clinique, Saint Malo, France

⁵Hôpital privé Sévigné, Cesson-Sévigné, France

⁶Centre Hospitalier Universitaire de Caen Normandie, Laboratoire d'Hématologie, Caen, France

⁷Centre Hospitalier Universitaire de Caen Normandie, Service d'Hématologie clinique, Caen, France

⁸Centre Hospitalier Universitaire de Clermont-Ferrand, Laboratoire d'Hématologie, Clermont-Ferrand, France

⁹Centre Hospitalier Universitaire de Clermont-Ferrand, Service d'Hématologie clinique, Clermont-Ferrand, France

¹⁰CIC1414, Rennes, France

* co-senior authorship

Corresponding authors:

Thierry Fest (thierry.fest@univ-rennes1.fr) and Thierry Lamy (thierry.lamy@univ-rennes1.fr) Hematology Department, CHU Pontchaillou, 2 rue Henri Le Guilloux, F-35033 Rennes cedex 9, France.

Tel: +33-299-284-291; Fax: +33-299-284-161

Key words: *TET2*, chronic lymphoproliferative disorders of NK cells, NK cells, KIR phenotype, *STAT3*, large granular lymphocyte leukemia

Key points:

- ***TET2* mutations present in 34% of chronic lymphoproliferative disorders of NK cells are associated with CD16^{low} phenotype.**
- ***TET2* mutations can occur in early hematopoietic progenitors and link the pathogenesis of CLPD-NK to other associated lymphoid or myeloid malignancies**

Abstract

Distinguishing chronic lymphoproliferative disorders of NK cells (CLPD-NK) from reactive NK cell expansions is challenging.

We assessed the value of NK receptor phenotyping and targeted high-throughput sequencing in a cohort of 114 consecutive patients with NK cell proliferation, retrospectively assigned to a CLPD-NK group (N=46) and a reactive NK group (N=68). We then developed a NK-clonality score combining flow cytometry and molecular profiling with a positive predictive value of 93%.

STAT3 and *TET2* mutations were respectively identified in 27% and 34% of the CLPD-NK patients – constituting a new diagnostic hallmark for this disease. *TET2*-mutated CLPD-NK exhibited preferentially a CD16^{low} phenotype, displayed more frequently a lower platelet count, and were associated with other hematologic malignancies such as myelodysplasia. To explore the mutational clonal hierarchy of CLPD-NK, we performed a whole exome sequencing of sorted, myeloid, T, and NK cells and identified that *TET2* mutations were shared by myeloid and NK cells in 3 out of 4 cases. Thus, we hypothesized that *TET2* alterations occur early in CLPD-NK disease which could explain a potential link between NK-LGL leukemia and other myeloid malignancies. Finally, we analyzed the transcriptome by RNA-seq of 7 CLPD-NK and evidenced two groups of patients. The first group displayed *STAT3* mutations or *SOCS3* methylation and overexpressed *STAT3* target genes. The second group, including two *TET2*-mutated cases, significantly under-expressed genes known to be down-regulated in angioimmunoblastic T-cell lymphoma. Our results provide new insights into the pathogenesis of NK cell proliferative disorders and potentially new therapeutic opportunities.

Introduction

Within the spectrum of large granular lymphocyte (LGL) leukemia, chronic lymphoproliferative disorder of NK cells (CLPD-NK) was listed provisionally as an individual clinical entity in the latest (2016) WHO classification.¹ This disorder is characterized by the persistent expansion of mature NK cells with a typical LGL morphology and a restricted killer immunoglobulin-like receptor (KIR) pattern.^{2,3,4,5} Unlike T-LGL leukemias which are characterized by a clonal T-cell receptor (TCR) rearrangement, NK cell proliferations do not display TCR clonality, making the differential diagnosis between clonal CLPD-NK and reactive NK cell expansion challenging.^{6,7} Indeed, after viral infection, autoimmune disease or organ transplantation, some patients have an elevated circulating LGL count, which raises the question of a reactive or neoplastic origin and highlights the need for an unambiguous diagnosis.

Flow cytometry (FCM) analysis of NK receptors is the most appropriate tool for diagnosing malignant NK cell proliferations.⁸ However, this technique is not widely available, and its results are sometimes difficult to interpret. Currently, *STAT3* and *STAT5B* mutations are considered as the sole molecular diagnostic markers for LGL leukemias, detected in about 30% and 5% of cases respectively.^{9,10,11} The exhaustive mutational profiling of CLPD-NK recently provided new insights in the knowledge of this disease¹². Notably, genes involved in (i) other signaling pathways (*MAPK*, *PI3K-AKT*), (ii) viral infections (*DDX3X*), (iii) DNA repair (*PAXIP1*), or (iv) epigenetic regulation (*TET2*, *SETD1B*), have been found altered in a cohort of 13 CLPD-NK, negative for *STAT3* mutations.¹² Mutations of *TET2* have also been reported in T-cell LGL leukemia associated with myeloid malignancies.¹³ Thus, the diagnostic value and clinical impacts of *TET2* mutations have to be validated in a larger cohort of CLPD-NK patients.

Herein, we retrospectively analyzed 114 consecutive patients with suspected CLPD-NK. We phenotyped by flow cytometry NK receptors and performed targeted high-throughput sequencing using a 76-gene panel designed for the diagnosis of hematologic malignancies.¹⁴ *STAT3* and *TET2* mutations were identified, respectively, in 27% and 34% of CLPD-NK patients, constituting a new diagnostic hallmark for this disease.

To describe the involvement of *TET2* alterations in CLPD-NK's pathogenesis, we performed whole exome sequencing on sorted, myeloid, T and NK populations, and compared transcriptomic profile and gene methylation status on tumor NK cells. *TET2* mutations were

always detected in the NK cell compartment with a high allelic frequency, although some patients exhibited alterations in myeloid cells (probably as a result of clonal hematopoiesis). Finally, we proposed a practical algorithm to classify NK cell proliferations.

Materials and methods

Patients

We included all consecutive patients having NK receptors phenotyping between January 2011 and January 2019 in our institution because of suspected CLPD-NK. Collected samples from 26 hospitals throughout France were obtained after approval by the local institutional review board. All patients gave their written consent to participate to the study which was conducted in accordance with the Declaration of Helsinki. Samples were centralized in a single center for analyses. We extracted the patients' clinical and biological data at the time of diagnosis. We retrospectively classified the patients either in the CLPD-NK or reactive NK proliferation group, based on the standard criteria for LGL leukemia.¹⁵ CLPD-NK was defined by a persistent increase over 0.5 G/L of circulating LGL with (i) an evocative clinical presentation (neutropenia, anemia, autoimmune disorder, infections); (ii) a proven clonality by flow cytometry; (iii) a *STAT3* mutation (sanger sequencing when available); or (iv) a histological evidence of LGL involvement.

Flow cytometry analysis

Flow cytometry (FCM) was performed prospectively as part of our routine workflow. All FCM data generated during the study were retrospectively reanalyzed, and gating strategies were cross-validated by two local expert reviewers. Ficoll-purified mononuclear cells were stained with the labeled antibodies listed in Supplemental Table 1. The NK cell count was calculated from the absolute lymphocyte count (ALC, determined with a routine hematologic analyzer) and the FCM-defined percentage of CD3⁻/CD56⁺ or CD16⁺ cells. Next, NK cell-receptor expressions on NK cells were evaluated using consecutive PE-labeled CD158A, CD158B, NKB1, CD94, NKG2D, CD161, NKP44, NKP46, and NKP30 antibodies. Analyses were performed consecutively on FC500, Navios (Beckman Coulter, Brea, CA), or Lyric (Becton Dickinson, Franklin Lake, NJ) flow cytometers. The proportion of positive NK cells (in %) and

the mean fluorescent intensity of each marker were calculated with Kaluza software (Beckman Coulter).

Targeted high-throughput sequencing (HTS)

Using cell samples that had been frozen at diagnosis, we sequenced a panel dedicated to hematologic neoplasms and constituted by 76 genes (Supplemental Table 2) including a subset of genes recommended for T/NK lymphoid malignancies: *STAT3*, *STAT5B*, *RHOA*, *PLCG1*, *JAK3*, *CD28*, *CARD11*, *DNMT3A*, *TET2*, *IDH1*, and *IDH2*.¹⁴ Genomic DNA (gDNA) was extracted from ficoll-isolated mononuclear cells with a Maxwell Rapid Sample Concentrator (Promega, Madison, WI) and quantified with Quantus fluorimeter (Promega). Libraries were generated in duplicate using the amplicon-based Access ArrayTM strategy (Fluidigm, San Francisco, CA) from 125 ng of DNA with Advanta NGS Library Prep reagents on an Access ArrayTM 48.48 Integrated Fluidic Circuit and sequenced on a NextSeq550 system (Illumina, San Diego, CA). Data were analyzed with an in-house bioinformatic pipeline (Supplemental Method 1). For the whole cohort, the median depth of sequencing was 2,735 reads by amplicon, ranged from 1,362 to 6,432 reads, ensuring a reliable identification of mutations with a variant allele frequency (VAF) above 2%. We evaluated copy number variations affecting the targeted genes with CovCopCan tool.¹⁶

Sanger sequencing

To validate *STAT3* and *TET2* mutations, 100 ng of DNA were amplified by PCR using AmpliTaq Gold Polymerase (Applied Biosystems, Foster City, CA) and specific primers listed in Supplemental Table 3. PCR products were purified and sequenced using the Big Dye Terminator kit on the ABI 3130xl analyzer (Applied Biosystems).

Cell sorting of myeloid, T-cell and NK-cell populations

Subsequent whole exome sequencing, transcriptomic and methylation analyses were performed on sorted myeloid, T and NK cells. Ten million DMSO-frozen cells were stained using DAPI, CD3-BUV395, CD33 and CD36-PerCP-Cy5.5 and CD56 and CD16-FITC. From alive DAPI-negative cells, myeloid cells (SSC^{high}, CD33⁺ or CD36⁺), T cells (SSC^{low}, CD3⁺), and NK cells (CD16⁺ or CD56⁺ and CD3⁻, CD33⁻, CD36⁻) were sorted on a BD FASCARIATM Fusion platform (Becton Dickinson) (Supplemental Figure 1). The purity of each cell subpopulation was

measured with the same gating strategy. DNA and RNA extractions were performed using the AllPrep DNA/RNA Micro Kit (Qiagen, Hilden, Germany) according to manufacturer instructions.

Whole exome sequencing

The gDNA was captured using Twist Human Core Exome Enrichment System (Twist Bioscience, San Francisco, CA) without modification except for library preparation performed with NEBNext Ultra II kit (New England Biolabs, Ipswich, MA).¹⁷ Libraries were sequenced on an Illumina NovaSeq platform as Paired-end 100 reads. Mutational analysis was performed with MERCURY software (IntegraGen SA, Evry, France). Detailed methods and bioinformatic analyses are available in Supplemental Method 2.

RNA-sequencing and bioinformatic analysis of tumor NK cells

Libraries are prepared with NEBNext Ultra II Directional RNA Library Prep Kit (New England Biolabs) for Illumina protocol according supplier recommendations. Sequencing was carried out on Paired End 100b reads on an Illumina NovaSeq platform. Library preparations, sequencing and bioinformatic with Galileo software have been done with IntegraGen SA (detailed in Supplemental Method 3).

Sodium bisulfite treatment and methylation-specific polymerase chain reaction (MSP)

To detect methylation of selected genes, we used the quantitative MSP as described previously.¹⁸ Briefly, 500 ng of gDNA was sodium bisulfite–modified using EZ methylation-Gold Kit (Zymo Research, Irvin, CA). Then, 10 ng of converted gDNA was amplified using ZymoTaq PreMix (Zymo Research) and methylation- or unmethylation-specific primer pairs (Supplemental Table 3). After denaturation in a MasterCycler (Eppendorf, Hambourg, Germany), amplification was carried out for 45 cycles before final extension. To test conversion efficiency and specificity of primers, universal methylated and unmethylated Human DNA standards were used in order to generate positive and negative controls. The analysis of PCR products was done on 2% agarose gels after staining with ClearSight DNA Stain (Euromedex, Souffelweyersheim, France).

5hmC dosage

The levels of gDNA-5hmC were measured using Global 5-hmC DNA ELISA Kit (Active Motif, Carlsbad, CA). Briefly, 20 ng of gDNA were fragmented by enzymatic digestion then denatured to create single-stranded DNA (ssDNA). 96-stripwell plates were coated with a DNA binding agent. Following the addition of ssDNA to the coated wells, unbound DNA fragments were washed away and subsequently treated with a primary antibody specific to 5hmC and a secondary antibody conjugated to horseradish peroxidase. Bound hydroxymethylated fragments were quantified by spectrophotometry and an absolute quantification was performed using concentrations of positive controls by the standard curve method. The amount of 5hmC was calculated from 20 ng ssDNA samples and expressed as a percentage.

Statistical analysis

Statistical analyses were performed with Prism software (version 7, GraphPad Inc., San Diego, CA). We used Mann-Whitney test to compare quantitative variables and Fisher's exact or chi-squared test to compare categorical variables. We divided the whole cohort in two groups: the training set enabled to build the NK-clonality score and the validation set allowed to confirm its diagnosis value on independent patients. The optimal threshold for quantitative variables was defined with a receiver operating characteristic curve. Parameters showing a significant difference between reactive NK and CLPD-NK were included as qualitative variable in a logistic regression model. We assigned a weight of 2 for significant variable in the multivariate analysis and a weight of 1 for parameter significantly different only in the univariate analysis. Sensitivity, specificity, predictive positive and negative values were reported for each parameter as well as for the NK-clonality score.

Results

Patients' characteristics

From January 2011 to January 2019, 114 patients were included in the study. The median [interquartile range (IQR)] age was 68 [57–77], and the sex ratio was 1.47. All blood samples were analyzed with our workflow for suspected CLPD-NK. Molecular assessments and FCM were mainly performed in cases displaying hyperlymphocytosis with LGLs on a blood smear and neutropenia or autoimmune disease. Twenty-five patients (22%) had a past or

concomitant malignant disease, 20 (18%) had an autoimmune disease, and 5 (4%) displayed skin symptoms or an elevated blood eosinophil count. Twenty-eight patients (26%) displayed neutropenia (*i.e.* an absolute neutrophil count (ANC) < 1.5 G/L). Overall, the median [IQR] proportion of NK cells (as a percentage of lymphocytes) was 47.3% [37.5-61.1%].

According to standard criteria for LGL leukemia available at diagnosis, 46 patients were considered to have a true CLPD-NK and 68 classified as having a reactive NK cell proliferation. Patients' characteristics are summarized in Table 1. The two groups of patients did not differ significantly with regard to the age, hemoglobin level, and neutrophil counts. Platelet- and NK cell- counts were, respectively, significantly lower and higher for CLPD-NK than reactive NK patients.

Flow cytometry characteristics of CLPD-NK patients

We classified the phenotype of NK cells in CLPD-NK patients according to their cytolytic potential using the previously described classification.¹⁹ Among the CLPD-NK patients, 24% had a true cytotoxic CD16^{high}/CD57⁻ profile, 53% an intermediate CD16^{high}/CD57⁺ phenotype, and 22% were CD16^{low} with a low cytolytic potential. We evaluated the expression of three KIRs (*i.e.* CD158A [KIR2DL1], CD158B [KIR2DL2/DL3], and NKB1 [KIR3DL1]), CD94, NKG2A, CD161 and three natural cytotoxicity receptors (NCR) (*i.e.* NKP30 [NCR3], NKP46 [NCR1] and NKP44 [NCR2]) for the 114 patients included in this study (Figure 1). The expression of the three KIRs was significantly lower in CLPD-NK than NK reactive group. The expression of CD94 and NKG2A on NK cells was significantly higher in CLPD-NK than in reactive NK patients: 94.6% vs. 45.2% ($P < 0.001$) and 93.4% vs. 44.5% ($P < 0.001$), respectively. The median percentage of CD161⁺ NK cells was significantly lower in CLPD-NK (63.3%) than in the other group (76.4%), $P = 0.006$. Conversely, the two groups did not differ significantly with regard to the expression of the NCR.

High-throughput sequencing profile

Eighty-six patients (41 CLPD-NKs and 45 reactive NK proliferations) were characterized by targeted HTS (Figure 2). Overall, we identified 85 clinically pertinent variants in 22 different genes (Supplemental table 4). In the CLPD-NK group, *STAT3* was mutated in 11 patients (27%), *STAT5* in two patients (5%); *TET2* in 14 patients (34%) with 26 different mutations (Figure 2A). Only one patient harbored a dual mutation profile, *TET2* and *STAT3*. Multiple

mutations for *TET2* were only detected in CLPD-NK patients. *TET2* truncating mutations were equally distributed across the whole gene, however, only missense mutations affecting functional domains of TET2 protein were considered as clinically relevant (Figure 2B). Mutations in *STAT3* and *STAT5B* were restricted to the SH2 domain. Both, *STAT3* and *TET2* mutations with a VAF higher than 10% were confirmed by Sanger sequencing (Supplemental Figure 2). In the CLPD-NK group, we also detected mutations in *TNFAIP3* for 4 patients (10%), *DNMT3A* for 3 patients (7%), and *NRAS* or *KRAS* for 3 patients (7%). In NK reactive proliferations, 6 (13%) presented a *DNMT3A* mutation and 3 (7%) a *TET2* mutation. Furthermore, we detected 19 mutations (5 in the CLPD-NK group and 14 in the reactive NK group) in genes described as clinically relevant in myeloid malignancies, namely, *JAK2*, *MPL*, *PPM1D*, *ASXL1*, *SF3B1*, *SRSF2*, *ZRSR2*, *RUNX1*, *U2AF1*, and *NPM1*. No copy number variations were detected for any explored genes. Overall, mutations in the *STAT3*, *TET2*, and *TNFAIP3* genes were significantly overrepresented in the CLPD-NK group suggesting their clinical interest in the CLPD-NK diagnosis (Figure 2A).

Clinical and phenotypical features of TET2 mutated CLPD-NK patients

The CLPD-NK group was divided in three subgroups based on their molecular profile for *STAT3* and *TET2*: *STAT3*-mutated (N=11), *TET2*-mutated (N=13), and wild-type (N=17) (Figure 3). Patients mutated for *STAT3* showed a significant decrease in hemoglobin concentration compared to the 2 other subgroups: median level of 9.6 g/dL [IQR 8.8-12.2] vs. 12.8 g/dL [11.5-14.5], respectively, $P = 0.012$. Moreover, the *STAT3*-mutated subgroup showed a lower count of neutrophil compared to patients with a reactive NK proliferation: 1.6 G/L [1.0-3.0] vs. 3.9 G/L [1.7-5.8], respectively, $P = 0.049$. Significant lower platelet counts were seen in *TET2* mutated patients compared to wild-type patients or reactive NK proliferations with: 165 G/L [93-229] vs. 218 G/L [195-300] or 232 G/L [184-322], $P = 0.047$ and $P = 0.014$, respectively (Figure 3A). On the phenotypical level, 6 (54%) patients mutated for *STAT3* harbored a cytotoxic CD16^{hi}/CD57⁻ profile and only 4 (13%) patients in the two other subgroups ($P = 0.012$). *TET2*-mutated cases preferentially exhibited a CD16^{low} phenotype (46% vs. 17%, $P = 0.041$). Finally, *TET2*-mutated patients presented another malignant hematologic disease more often than the other subgroups (38% vs. 7%, $P = 0.024$) (Figure 3B). Indeed, 5 *TET2*-mutated patients developed either acute myeloid leukemia, myelodysplastic syndrome, pure red cell aplasia, AITL, or Hodgkin lymphoma. In contrast,

only 2 out of 28 *TET2*-unmutated patients (7%) exhibited a second hematopoietic neoplasm (Supplemental table 5).

Establishment of a NK-clonality score

We established a novel NK-clonality score to diagnose CLPD-NK in the clinical practice (Figure 4). According to the data of the training cohort (N=76), NK cell count, CD158A, CD158B, NKB1, CD94, and NKG2A expression levels discriminated CLPD-NK from NK reactive proliferations with optimal cut-offs defined with the analysis of ROC curves. Instead, CD161 and NCR expression failed to separate these two entities (Supplemental Figure 3). Among the 48 patients assessed with both FCM and HTS analyzes, four criteria were significantly associated with the CLPD-NK diagnosis: (i) NK cell count > 1G/L; (ii) a KIR-restricted pattern defined by a loss of expression of at least two KIRs (CD158A < 9% of NK cells, CD158B < 12%, and/or NKB1 < 4%); (iii) high CD94 or NKG2A expression (>77%), and (iv) mutations in *STAT3*, *STAT5B*, *TET2*, and *TNFAIP3* genes (Table 2). We further included these 4 criteria in a logistic regression model and attributed a weight of 2 points for parameters retained as significant in this multivariate analysis: NK cell count, the KIR-restricted phenotype, and somatic mutations (Table 2). We assigned a weight of 1 point for CD94/NKG2A^{high} phenotype. This score ranged from 0 (lowest possible clonality) to 7 (highest possible clonality). We integrated the score into a dedicated diagnostic workflow for patients with clinical or laboratory data suggestive of CLPD-NK (Figure 4A). Twenty (42%) out of 48 patients (19 CLPD-NKs and 1 reactive NK) had a score of 4 or higher. Eight (17%) patients (3 CLPD-NKs and 5 reactive NK) had a score of 2 or 3. Finally, 20 (42%) patients (1 CLPD-NKs and 19 reactive NK) had a score less than 2. Thus, a score of 4 or greater had a sensitivity of 83% and a specificity of 96% for the diagnosis of CLPD-NK (Table 2). A score less than 2 could rule out the diagnosis of CLPD-NK with a negative predictive value of 95%. The diagnostic value of the NK clonality score was confirmed in the independent validation cohort (N=38), showing a positive predictive value of 100% (Figure 4B and Table 2). The score value among reactive NK and CLPD-NK groups was equally distributed in the training and the validation sets (Figure 4C). Overall, bone marrow assessment was routinely performed for 14 patients and evidence an LGL infiltrate in 10 cases (Supplemental Table 5). On the complete dataset, we retrospectively reanalyzed the medical records of the 2 patients with an NK-clonality score less than 2 but who presented a real CLPD-NK disease according to the standard

criteria. The first case had an inclusion body myositis without circulating LGL; the second patient was diagnosed with diffuse large B cell lymphoma and presented an NK cell count of 0.7 G/L with a low level of CD161 as a unique FCM marker, which argued for a final classification in the group reactive NK. Only one patient had a score of 5 but was initially classified in the reactive NK group due to a NK cell count at 0.4 G/L. However, he showed neutropenia with an autoimmune disorder, met both FCM criteria and carried a *TET2* mutation.

Clonal hierarchy in CLPD-NK

We next studied the clonal hierarchy in 6 *TET2*-mutated patients by sequencing FACS-sorted myeloid, T, and NK cell compartments. The purity was higher than 96% for all tested subpopulations (Figure 5A and 5B). All *TET2* mutations in NK cell samples were at a high allelic frequency. In 4 cases (67%), at least one *TET2* mutation was also detected in myeloid cells but no in T cells. We performed comparative whole exome sequencing (WES) for 4 cases which had sufficient DNA for all three compartments; one patient (#105) had a fourth sample issued from an angioimmunoblastic T-cell lymphoma (AITL) diagnosed two years before the CLPD-NK (Figure 5C). Overall, we identified 189 different variants: 86 were restricted to NK cells, 60 were shared between myeloid and NK cells, 9 restricted to myeloid compartments, and 29 were only detected in the AITL sample (Supplemental Table 6). For patient #70, the *TET2* mutation was only present in NK cells but associated with a *GNAS* mutation previously described as driving a STAT3 pathway activation.²⁰ For patients #98, #48, and #105, the 5 *TET2* variants were identified in NK and myeloid cells. Their NK compartments had at least one activating mutation in genes involved in signaling pathways described in NK cell diseases, namely: *STAT3*, *TNFAIP3*, or *IL24*. Patient #105 had two *TET2* mutations in the NK and myeloid compartments that were also detectable two years earlier in the AITL tumor (Figure 5C).

Methylation profile of CLPD-NK proliferations

We evaluated the functional impact of *TET2* enzyme loss-of-function. Thus, we measured the overall levels of 5hmC from sorted NK cells issued from 18 CLPD-NK (7 *STAT3*, 5 *TET2*, and 6 wild-type) and two reactive NK samples. Overall, 5hmC levels were significantly decreased in CLPD-NK compared to reactive NK samples with a median of 0.034% [IQR

0.030-0.039] and 0.066% [0.060-0.072], respectively, $P = 0.036$ (Figure 6A). However, no detectable differences were found between the 3 subgroups of CLPD-NK. CpG methylation of selected gene promoters involved in NK cells proliferations were explored for 7 CLPD-NK patients (2 *STAT3*, 3 *TET2*, and 2 wild-type). For *KIR2DL1*, *KIR2DL3*, and *KIR3DX1* genes, we detected a strong methylation for all tested CLPD-NK patients without specific modulations according to the patient molecular profile (data not shown). We found a methylation in *SOCS3* gene for one patient (14%) mutated for *TET2* (Patient #70 explored by WES and RNA-seq) (Figure 6 B).

Transcriptomic profile of CLPD-NK proliferations

Finally, we studied the transcriptome by RNA-seq of blood-purified NK cells in these 7 CLPD-NK patients. Both, unsupervised clustering and principal component analysis (PCA) of the 1,000 most variable genes of the dataset identified two groups of patients (G1 and G2) (Figure 7A and 7B). Expression of *IL10* was highly discriminant between them (Figure 7C). We then explored differentially expressed genes between these two groups (G1 vs. G2) and identified a set of 2,410 overexpressed and 3,604 underexpressed genes (fold change > 2 and q -value > 0.05) (Figure 7D and Supplemental table 7). Gene Set Enrichment Analysis (GSEA) identified 899 and 68 curated signatures in, respectively, G1 overexpressed and underexpressed genes (Figure 7E). Integrative analysis of the data showed that the three samples that constituted G1 involved the two *STAT3* mutated cases while the third case (Patient #70) was *SOCS3* methylated and mutated for *GNAS*, both abnormalities are functionally related to a *STAT3* pathway activation.^{20,21} Accordingly, two signatures related to *STAT3* pathway activation were significantly enriched for this group (Dauer *et al.* $P = 8.1E-07$ and Azare *et al.* $P = 5.1E-05$) (Figure 7E and 7F).^{22,23} On the other hand, among the G2-related signatures one was the most striking and connected to angioimmunoblastic lymphoma (Figure 7E).²⁴ Indeed, G2 underexpressed 70 genes that belong to a 131-gene signature called AITL-Down ($P = 5.8E-40$) and overexpressed 35 genes involved in an AITL-Up signature of 79 genes ($P = 6.3E-04$). The G2 group included two *TET2*-mutated patients and showed lower expression of *TET2* (FC 2.4; $P = 1.4E-04$) (Supplemental table 7).

Discussion

The objective of the present study was to optimize the classification of NK cell proliferations by leveraging data from NK receptor phenotyping and HTS. Proven clonality is a mandatory diagnostic criterion for CLPD-NK.¹⁵ The HUMARA assay (evaluating the methylation pattern for the human androgen receptor gene on chromosome X) is the standard test for NK cell clonality.²⁵ However, such exploration is not widely available in hospital laboratories and is only relevant for female patients. Thus, the FCM profile of NK receptors has been broadly accepted as a surrogate marker for NK cell clonality.⁸ In our study, a restrictive KIR pattern, CD161 loss, and elevated expression of the lectin-type heterodimer CD94/NKG2A were the most discriminant markers for CLPD-NK. These data from a large real-life experience confirmed the previous published data on FCM diagnostic value for NK cell proliferations.^{26,27,28,29} Conversely, few studies have been reported on the diagnostic value of the natural cytotoxicity receptor in LGL leukemia.²⁷ In our hands, this parameter did not allow differentiation of CLPD-NK from reactive conditions, as it was similar in both groups and therefore did not constitute strong evidence of NK clonality.

STAT3 mutation, found in up to 30% of the cases, is now considered as a robust marker to distinguish CLPD-NK from NK reactive conditions.^{9,10,12} Recently, the mutational landscape of CLPD-NK has been characterized by WES and revealed in two *STAT3*-unmutated cases the presence of *TET2* mutations in a series of 13 patients.¹² In the context of T-cell LGL leukemia, two reports described *TET2* mutations related to clonal hematopoiesis of indeterminate potential.^{13,30} Our study on CLPD-NK confirmed the prevalence of *STAT3* and *STAT5B* mutations (respectively, 27% and 5%).^{10,11} We identified *TET2* mutations in 34% of the cases, constituting a new hallmark of CLPD-NK. We detected additional mutations in genes involved in other signaling pathways like *TNFAIP3*, as previously reported in small series of NK-cell and T-cell LGL leukemia.^{31,32} Overall, we found a strong molecular marker of CLPD-NK for 63% of the patients using a targeting HTS strategy dedicated to hematologic malignancies.

We propose that a combination of FCM and targeting HTS should be considered for the diagnosis of NK cell proliferation. We developed a NK-clonality score, integrating NK cell count, the two most discriminant phenotypic criteria (KIR restriction and elevated CD94/NKG2A expression), and the detection of at least one mutation in *STAT3*, *STAT5B*, *TET2* and/or *TNFAIP3* genes. Assessing clonality with the only FCM represents a complex multiparametric procedure, reserved to specialized centers. This score reduces the

complexity of the FCM interpretation by selecting the most relevant markers and providing optimal thresholds for the final decision. In addition, it includes a simple quantitative parameter (NK cell count) and the detection of mutations providing objective evidence of clonality. Although WES assay is readily available nowadays, targeted HTS is more widely used in routine practice and offers a higher sensitivity to detect variants in subclonal populations in unsorted samples. In our cohort, a score of 4 or more allows to diagnose a CLPD-NK with an excellent positive predictive value (>95%). Instead, patients with a score of 0-1 were almost exclusively diagnosed as having reactive NK proliferations. We suggest that patients with a score of 2 or 3 fall within a “gray area” that justifies the performance of a bone marrow biopsy to screen for an interstitial infiltrate of granzyme B-positive, TIA1⁺ LGL.^{4,26,33} The value of this score was confirmed in an independent set of 38 patients and is currently used to validate patient inclusions in an ongoing French national prospective clinical trial (clinicaltrials.gov identifier: NCT01976182).

STAT3- and *TET2*-mutated CLPD-NK have distinct clinical and phenotypical patterns. Traditionally, mature NK cells are divided into two entities: (i) CD56^{hi}CD16^{low} regulatory NK cells with greater cytokine-producing ability, and (ii) CD56^{dim}/CD16^{hi} NK cells with greater cytotoxic function. In the latter, the acquisition of CD57 defines a subset of terminally differentiated “memory” NK cell.^{34,35} Consistent with previous studies, the most cytotoxic CD16^{hi}CD57⁺ profile was preferentially observed in *STAT3*-mutated CLPD-NK characterized by deeper cytopenia compared to unmutated cases.^{9,19,31} Conversely, we show that *TET2*-mutated CLPD-NKs exhibit a regulatory CD16^{low} phenotype with thrombocytopenia and are more often subjected to other hematologic malignancies in their clinical history.

These findings led us to question the clonal hierarchy of mutations in CLPD-NK. Currently, CLPD-NK pathogenesis model postulates that a first event consisting of an expansion of polyclonal NK cells - induced by an exogenous stimulus - leads to the production of cytokines and then to the selection of a population of clonal NK cells. These expanded NK cells ultimately accumulate genetic alterations and cause clinical symptoms such as cytopenia and autoimmune disorders.³⁶ We proposed an alternative scenario for *TET2*-mutated cases. Indeed, our WES data found *TET2* mutations in NK and myeloid compartments while additional mutations, implicated in NK biology (*i.e.* *STAT3*, *TNFAIP3*, *IL24*, or *GNAS*), were restricted to NK cells. Our results suggest a multiple-hit model of leukemogenesis in which *TET2* mutations can occur in early hematopoietic progenitors and

give rise to different hematologic neoplasms. This scenario is somewhere illustrated by the patient who presented an AITL two years before his CLPD-NK, a disease in which preexisting TET2-clonal hematopoiesis has been postulated previously.^{37,38} *TET2* is one of the main gene involved in clonal hematopoiesis of intermediate potential (CHIP).³⁹ About 5% of the LGL leukemia patients suffer from a concurrent myelodysplastic syndrome.^{7,13} These patients display a high prevalence of *TET2*, *ASXL1*, and *DNMT3A* mutations – suggesting that myelodysplasia and LGL leukemia could arise from a common CHIP.^{38,40} Although NK ontogeny has been considered to be exclusively lymphoid, several studies have shown that NK cells may originate from myeloid progenitors.^{41–43} Specifically, common myeloid progenitors and granulocytic-monocytic precursors can give rise to mature NK cells in a cytokine environment supporting NK (IL-3, IL-15, IL-7, SCF, and FLT3-L).⁴² In accordance with this observation, we hypothesize that the occurrence of *TET2* mutations in these progenitors could ultimately lead to an expansion of myeloid and NK cells, highlighting the epidemiological link between CLPD-NK and myeloid malignancies.

The high prevalence of *TET2* mutations associated with the detection of multiple truncating or functionally damaging variants in the same sample (presumably associated with biallelic inactivation) suggest that TET2 loss-of-function could lead to the emergence of CLPD-NK. TET2 induces an oxidation of 5-methylcytosine (5mC) into 5-hydroxymethylcytosine (5hmC) in the first step of an active DNA demethylation process.⁴⁴ Accordingly, we found a significant decrease of 5hmC levels in CLPD-NK in accordance with previous studies in other hematologic malignancies.⁴⁵ Due to the paucity of samples, we could not observe any correlations between 5hmC residues and mutation profile of CLPD-NK as was done in acute myeloid leukemia.⁴⁵ Furthermore, we did not detect methylation modifications of CpG islands in *KIR* gene promoters according to the mutation profile, while such element was considered as a crucial event in the clonal selection of NK cells.^{46,47} We could speculate that our investigations, limited to some regions due the methyl specific PCR approach, do not allow a complete identification of methylation variations of *KIR* genes. However, we found a methylation of *SOCS3* promoter in one *TET2*-mutated patient. *SOCS3* is an important regulator of the STAT3 pathway, involved in a negative feedback loop.⁴⁸ Thus, the epigenetic *SOCS3* down-regulation might enhance the STAT3 pathway activation demonstrated by our RNA-seq analysis, in agreement with a previous report in T-LGL leukemia.²¹

The transcriptomic analysis of clonal NK cells issued from CLPD-NKs sheds new light on the pathogenesis of CLPD-NK. Two different groups of patients were identified. The first one, we called “STAT3-like”, was linked to a strong imprint of STAT3 pathway while the second group, linked to TET2 loss-of-function, displayed an “AITL-like” signature.²⁴ These data open up therapeutic perspectives in CLPD-NK. Currently, immunosuppressive therapy is the standard treatment for LGL leukemia¹⁵ but complete response is rare and long-term remission almost nonexistent – regardless of the patient’s mutation status. Thus, CLPD-NK represents an unmet clinical need and new therapeutic options are required for relapsing or treatment-refractory cases. A better understanding of the molecular features of T- or NK-LGL leukemia is likely to accelerate the development of targeted therapies. Specific inhibitors of JAK-STAT pathway should be preferred for *STAT3* mutated or “STAT3-like” patients. Although the specific JAK3 inhibitor tofacitinib has shown good clinical efficacy, this drug is currently only authorized for patients suffering from rheumatoid arthritis.⁴⁹ Chromosomal instability and DNA hypermethylation have been observed in NK cells from the IL-15 transgenic mouse model - suggesting that proteasome inhibitors and demethylating agents might be efficacious in CLPD-NK.⁵⁰ Hence, 5-azacytidine (already evaluated in the context of AITL) should be tested in cases *TET2* mutated or “AITL-like” patients.⁵¹

In conclusion, we confirm that *TET2* mutation is a new molecular marker and a potential therapeutic target in CLPD-NK. Our results also provided new insights into the pathogenesis and classification of NK cell proliferative disorders.

Acknowledgments

We thank the Centre de Ressource Biologique-Santé biological resource center (Rennes, France) for providing biological samples (reference: BB-0033-00056). This project was partly funded by the Hematology Laboratory at Rennes University Medical Center and the Association pour le développement de l'Hématologie Oncologie (ADHO). We also thank Cécile Thomas de la Pintièrre for technical support and David Fraser PhD (Biotech Communication SARL, Ploudalmézeau, France) for copy-editing support.

Contributions:

C.Pas., M.R., T.F., and T.L. designed the study; G.D. and T.L. reviewed the clinical files for initial classification; M.R. and C.Pas. analyzed flow cytometry data; S.L.G. performed cell-sorting, C.Pas. and M.L.B. generated and analyzed high throughput sequencing and RNAseq data, F.D. designed and analyzed methylation experiments; C.Pan. managed the sample collection; G.D., A.T., A.V.D., V.S., G.L.D., R.V.M., O.T., A.M., and T.L. provided clinical information; C.Pas. and T.F. wrote the manuscript. All the authors critically revised the manuscript.

Disclosure:

The authors report no conflicts of interest.

Data sharing:

Data are available in GEO under accession number GSE163093.

Reference

1. Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016;127(20):2375–2390.
2. Lamy T, Loughran TP. Current concepts: large granular lymphocyte leukemia. *Blood Rev*. 1999;13(4):230–240.
3. Semenzato G, Marino F, Zambello R. State of the art in natural killer cell malignancies: ART IN NATURAL KILLER CELL MALIGNANCIES. *International Journal of Laboratory Hematology*. 2012;34(2):117–128.
4. Lamy T, Moignet A, Loughran TP. LGL leukemia: from pathogenesis to treatment. *Blood*. 2017;129(9):1082–1094.
5. Morice WG. The immunophenotypic attributes of NK cells and NK-cell lineage lymphoproliferative disorders. *Am J Clin Pathol*. 2007;127(6):881–886.
6. Poullot E, Zambello R, Leblanc F, et al. Chronic natural killer lymphoproliferative disorders: characteristics of an international cohort of 70 patients. *Ann. Oncol*. 2014;25(10):2030–2035.
7. Barilà G, Calabretto G, Teramo A, et al. T cell large granular lymphocyte leukemia and chronic NK lymphocytosis. *Best Practice & Research Clinical Haematology*. 2019;32(3):207–216.
8. Bárcena P, Jara-Acevedo M, Tabernero MD, et al. Phenotypic profile of expanded NK cells in chronic lymphoproliferative disorders: a surrogate marker for NK-cell clonality. *Oncotarget*. 2015;6(40):42938-42951.
9. Jerez A, Clemente MJ, Makishima H, et al. STAT3 mutations unify the pathogenesis of chronic lymphoproliferative disorders of NK cells and T-cell large granular lymphocyte leukemia. *Blood*. 2012;120(15):3048–3057.
10. Koskela HL, Eldfors S, Ellonen P, et al. Somatic STAT3 mutations in large granular lymphocytic leukemia. *New England Journal of Medicine*. 2012;366(20):1905–1913.
11. Rajala HLM, Eldfors S, Kuusanmäki H, et al. Discovery of somatic STAT5b mutations in large granular lymphocytic leukemia. *Blood*. 2013;121(22):4541–4550.
12. Gasparini VR, Binatti A, Coppe A, et al. A high definition picture of somatic mutations in chronic lymphoproliferative disorder of natural killer cells. *Blood Cancer J*. 2020;10(4):42.
13. Durrani J, Awada H, Kishtagari A, et al. Large granular lymphocytic leukemia coexists with myeloid clones and myelodysplastic syndrome. *Leukemia*. 2020;34(3):957–962.
14. Sujobert P, Le Bris Y, de Leval L, et al. The Need for a Consensus Next-generation Sequencing Panel for Mature Lymphoid Malignancies. *HemaSphere*. 2019;3(1):e169.
15. Lamy T, Loughran TP. How I treat LGL leukemia. *Blood*. 2011;117(10):2764–2774.
16. Derouault P, Chauzeix J, Rizzo D, et al. CovCopCan: An efficient tool to detect Copy Number Variation from amplicon sequencing data in inherited diseases and cancer. *PLoS Comput. Biol*. 2020;16(2):e1007503.
17. Gnirke A, Melnikov A, Maguire J, et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nature Biotechnology*. 2009;(2):182.
18. Herman JG, Graff JR, Myöhänen S, Nelkin BD, Baylin SB. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc. Natl. Acad. Sci. U.S.A.* 1996;93(18):9821–9826.
19. Barilà G, Teramo A, Calabretto G, et al. Dominant cytotoxic NK cell subset within CLPD-NK patients identifies a more aggressive NK cell proliferation. *Blood Cancer Journal*. 2018;8(6):51.
20. Ding H, Zhang X, Su Y, Jia C, Dai C. GNAS promotes inflammation-related

- hepatocellular carcinoma progression by promoting STAT3 activation. *Cell Mol Biol Lett.* 2020;25:8.
21. Teramo A, Gattazzo C, Passeri F, et al. Intrinsic and extrinsic mechanisms contribute to maintain the JAK/STAT pathway aberrantly activated in T-type large granular lymphocyte leukemia. *Blood.* 2013;121(19):3843–3854.
 22. Dauer DJ, Ferraro B, Song L, et al. Stat3 regulates genes common to both wound healing and cancer. *Oncogene.* 2005;24(21):3397–3408.
 23. Azare J, Leslie K, Al-Ahmadie H, et al. Constitutively activated Stat3 induces tumorigenesis and enhances cell motility of prostate epithelial cells through integrin beta 6. *Mol. Cell. Biol.* 2007;27(12):4444–4453.
 24. Piccaluga PP, Agostinelli C, Califano A, et al. Gene expression analysis of angioimmunoblastic lymphoma indicates derivation from T follicular helper cells and vascular endothelial growth factor deregulation. *Cancer Res.* 2007;67(22):10703–10710.
 25. Boudewijns M, van Dongen JJM, Langerak AW. The Human Androgen Receptor X-Chromosome Inactivation Assay for Clonality Diagnostics of Natural Killer Cell Proliferations. *The Journal of Molecular Diagnostics.* 2007;9(3):337–344.
 26. Morice WG, Jevremovic D, Olteanu H, et al. Chronic lymphoproliferative disorder of natural killer cells: a distinct entity with subtypes correlating with normal natural killer cell subsets. *Leukemia.* 2010;24(4):881–885.
 27. Zambello R. Expression and function of KIR and natural cytotoxicity receptors in NK-type lymphoproliferative diseases of granular lymphocytes. *Blood.* 2003;102(5):1797–1805.
 28. Epling-Burnette PK, Painter JS, Chaurasia P, et al. Dysregulated NK receptor expression in patients with lymphoproliferative disease of granular lymphocytes. *Blood.* 2004;103(9):3431–3439.
 29. de Mel S, Li JB, Abid MB, et al. The utility of flow cytometry in differentiating NK/T cell lymphoma from indolent and reactive NK cell proliferations. *Cytometry B Clin Cytom.* 2018;94(1):159–168.
 30. Raess PW, Cascio MJ, Fan G, et al. Concurrent STAT3, DNMT3A, and TET2 mutations in T-LGL leukemia with molecularly distinct clonal hematopoiesis of indeterminate potential. *Am. J. Hematol.* 2017;92(1):E6–E8.
 31. Kawakami T, Sekiguchi N, Kobayashi J, et al. STAT3 mutations in natural killer cells are associated with cytopenia in patients with chronic lymphoproliferative disorder of natural killer cells. *Int J Hematol.* 2019;109(5):563–571.
 32. Johansson P, Bergmann A, Rahmann S, et al. Recurrent alterations of TNFAIP3 (A20) in T-cell large granular lymphocytic leukemia. *Int. J. Cancer.* 2016;138(1):121–124.
 33. Evans H. Utility of immunohistochemistry in bone marrow evaluation of T-lineage large granular lymphocyte leukemia. *Human Pathology.* 2000;31(10):1266–1273.
 34. Smith SL, Kennedy PR, Stacey KB, et al. Diversity of peripheral blood human NK cells identified by single-cell RNA sequencing. *Blood Advances.* 2020;4(7):1388–1406.
 35. Poznanski SM, Ashkar AA. What Defines NK Cell Functional Fate: Phenotype or Metabolism? *Front Immunol.* 2019;10:1414.
 36. Zambello R, Teramo A, Barilà G, Gattazzo C, Semenzato G. Activating KIRs in Chronic Lymphoproliferative Disorder of NK Cells: Protection from Viruses and Disease Induction? *Frontiers in Immunology.* 2014;5:72.
 37. Couronné L, Bastard C, Bernard OA. TET2 and DNMT3A mutations in human T-cell lymphoma. *N. Engl. J. Med.* 2012;366(1):95–96.
 38. Tiacci E, Venanzi A, Ascani S, et al. High-Risk Clonal Hematopoiesis as the Origin of AITL and NPM1-Mutated AML. *N. Engl. J. Med.* 2018;379(10):981–984.
 39. Genovese G, Kähler AK, Handsaker RE, et al. Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence. *New England Journal of Medicine.*

2014;371(26):2477–2487.

40. Hirsch CM, Nazha A, Kneen K, et al. Consequences of mutant TET2 on clonality and subclonal hierarchy. *Leukemia*. 2018;32(8):1751–1761.
41. Oshimi K. Progress in understanding and managing natural killer-cell malignancies. *British Journal of Haematology*. 2007;139(4):532–544.
42. Grzywacz B, Kataria N, Kataria N, et al. Natural killer-cell differentiation by myeloid progenitors. *Blood*. 2011;117(13):3548–3558.
43. Cichocki F, Grzywacz B, Miller JS. Human NK Cell Development: One Road or Many? *Frontiers in Immunology*. 2019;10:2078.
44. Feng Y, Li X, Cassady K, Zou Z, Zhang X. TET2 Function in Hematopoietic Malignancies, Immune Regulation, and DNA Repair. *Front Oncol*. 2019;9:210.
45. Ahn J-S, Kim H-J, Kim Y-K, et al. 5-Hydroxymethylcytosine correlates with epigenetic regulatory mutations, but may not have prognostic value in predicting survival in normal karyotype acute myeloid leukemia. *Oncotarget*. 2017;8(5):8305–8314.
46. Gattazzo C, Teramo A, Miorin M, et al. Lack of expression of inhibitory KIR3DL1 receptor in patients with natural killer cell-type lymphoproliferative disease of granular lymphocytes. *Haematologica*. 2010;95(10):1722–1729.
47. Santourlidis S, Trompeter H-I, Weinhold S, et al. Crucial role of DNA methylation in determination of clonally distributed killer cell Ig-like receptor expression patterns in NK cells. *J. Immunol*. 2002;169(8):4253–4261.
48. Croker BA, Krebs DL, Zhang J-G, et al. SOCS3 negatively regulates IL-6 signaling in vivo. *Nat Immunol*. 2003;4(6):540–545.
49. Bilori B, Thota S, Clemente MJ, et al. Tofacitinib as a novel salvage therapy for refractory T-cell large granular lymphocytic leukemia. *Leukemia*. 2015;29(12):2427–2429.
50. Mishra A, Liu S, Sams GH, et al. Aberrant Overexpression of IL-15 Initiates Large Granular Lymphocyte Leukemia through Chromosomal Instability and DNA Hypermethylation. *Cancer Cell*. 2012;22(5):645–655.
51. Cheminant M, Bruneau J, Kosmider O, et al. Efficacy of 5-azacytidine in a TET2 mutated angioimmunoblastic T cell lymphoma. *Br. J. Haematol*. 2015;168(6):913–916.

Table 1: characteristics of the studied population

| | CLPD-NK N=46 | | Reactive NK N= 68 | | P-value |
|--|-----------------|-------------|----------------------|-------------|---------|
| Age (years) | 68 | [57-77] | 67 | [57-77] | ns |
| Sex ratio (M/F) | 28/18 | 1.55 | 42/26 | 1.62 | ns |
| <i>Clinical manifestations</i> | | | | | |
| Neutropenia | 12 | 26% | 16 | 24% | ns |
| Auto-immune disorders | 9 | 20% | 11 | 16% | ns |
| Cutaneous symptoms | 0 | 0% | 5 | 7% | ns |
| Other hematologic malignancies | 8 | 17% | 17 | 25% | ns |
| Other solid neoplasia | 1 | 2% | 4 | 6% | ns |
| Organ transplantation | 0 | 0% | 2 | 3% | ns |
| <i>Complete blood count</i> | | | | | |
| Hemoglobin (g/dL) | 12.5 | [10.4-14.2] | 13.0 | [11.5-14.5] | ns |
| Platelets (G/L) | 195 | [131-264] | 231 | [184-317] | 0.016 |
| Leukocytes (G/L) | 10.8 | [7.7-10.8] | 7.1 | [4.0-9.3] | <0.001 |
| Neutrophils (G/L) | 3.0 | [1.2-4.4] | 3.9 | [1.8-6.3] | ns |
| NK cells (G/L) | 2.3 | [0.9-3.2] | 0.6 | [0.4-0.9] | <0.001 |
| <i>FCM clonality assessment at diagnosis</i> | | | | | |
| Clonality suspected | 42 | (91%) | 6 | (9%) | <0.001 |
| Polyclonal | 4 | (9%) | 62 | (91%) | |
| <i>Sanger sequencing</i> | | | | | |
| STAT3 | 10 | (24%) | 0 | (0.0%) | <0.001 |

Continuous variables are quoted as the median [IQR] and categorical variables are quoted as number of patients (percentage of tested patients). Sanger sequencing data were available for 41 patients in the CLPD-NK group and 45 in the reactive NK proliferation group. Ns: not significant

Table 2: diagnostic performance of flow cytometry and molecular profiles for NK clonality assessment, the NK-clonality score.

| Training set (N=48) | Se | Sp | PPV | NPV | Univariate analysis | | Multivariate analysis | | Scoring |
|--|-------|-------|-------|-------|---------------------|---------|-----------------------|---------|---------|
| | | | | | Khi2 | P-value | Khi2(LR) | P-value | |
| NK cell count | | | | | | | | | |
| NK cell count>1G/L | 69.6% | 84.0% | 80.0% | 75.0% | 12.023 | < 0.001 | 6.581 | 0.010 | 2 |
| Flow cytometry | | | | | | | | | |
| KIR-restricted phenotype* | 82.6% | 92.0% | 90.5% | 85.2% | 24.149 | < 0.001 | 9.077 | 0.003 | 2 |
| CD94 or NKG2A ^{hi**} | 69.6% | 76.0% | 72.7% | 73.1% | 8.267 | 0.004 | 1.125 | 0.289 | 1 |
| STAT3, STAT5B, TET2 or TNFAIP3 genes status | | | | | | | | | |
| Mutated | 60.9% | 96.0% | 93.3% | 72.7% | 15.483 | < 0.001 | 5.222 | 0.022 | 2 |

| NK clonality score ≥4 | Se | Sp | PPV | NPV | AUC | P-value |
|------------------------------|-------|--------|--------|-------|-------|---------|
| Training set (N=48) | 82.6% | 96.0% | 95.0% | 95.7% | 0.937 | < 0.001 |
| Validation set (N=38) | 83.3% | 100.0% | 100.0% | 87.0% | 0.946 | < 0.001 |

Se: sensitivity; Sp: specificity; PPV: positive predictive value; NPV: negative predictive value; LR: logistical regression; AUC: Receiver operating characteristic area under curve.

* A KIR-restricted phenotype was defined as the loss of at least 2 KIR (CD158A < 12%, CD158B < 9% or NKB1 < 4% of NK cells) **CD94 or NKG2A expressed by more than 77% of NK cells

Figure legends

Figure 1: Flow cytometry analysis of NK receptors in the CLPD-NK and reactive NK groups.

For the reactive NK group, the expression distribution of each FCM marker is represented by a box plot. The whiskers correspond to the 10th-90th percentiles. Individual CLPD-NK samples are plotted as triangles. The optimal cutoff defined with the analysis of ROC curves for the NK clonality score was plotted in red for overexpressed antigens or in green for underexpressed antigens. Comparisons were performed using Mann-Whitney test.

Figure 2: Mutational status of CLPD-NK patients

A: Heat map of flow cytometry profile and mutations detected by targeted high throughput sequencing for the 41 CLPD-NK and 45 reactive NK patients. The percentages of mutation for each gene were compared using Fisher exact t-test (* $P < 0.05$). **B:** Lolliplots showing the distribution of *STAT3* and *TET2* mutations detected in this study, with their functional interpretation.

Figure 3: Clinical and phenotypical characteristics of patients according mutational status

A: Comparison of Complete blood count parameters according to the mutational profile (Mann-Whitney test). **B:** Heatmap of the prevalence (in percentage) of cytopenia, phenotypical profile and associated events in patients with CLPD-NK or reactive NK proliferations. * P -value for *STAT3*-mutated vs. the two other CLPD-NK groups and ** P -value for *TET2*-mutated vs. the two other CLPD-NK groups comparisons (Fisher exact t-test).

Figure 4: NK clonality score

A: Flow chart for NK clonality score development based on training set data (N=76). The pie charts show the distributions of patients in CLPD-NK (blue) vs. reactive NK group (gray) according to standard criteria. Comparison with final assessment according NK clonality score. **B:** Validation of NK clonality score on an independent set of patients. **C:** Distribution of the NK clonality score among the two groups.

Figure 5: Whole exome sequencing of sorted myeloid, T and NK cells

A: Evaluation by flow cytometry of the purity of sorted NK, T and myeloid cells from one representative case (Patient #105). **B:** Purity evaluation of FACS-sorted cells for all samples. *) For 2 patients, the number of sorted myeloid cells was too low (< 100,000) to perform purity control analysis. **C:** Whole exome sequencing results for 4 *TET2*-mutated patients. Venn diagrams represents the number of mutations in NK cells and myeloid cells as compared to the reference sample (T cell). Patient #105 analysis also includes a tumoral sample (angioid immunoblastic lymphoma, AITL) collected 2 years before the occurrence of the CLPD-NK. Mentioned mutations correspond to pathogenic variants and/or variants related in OncoKB database.

Figure 6: Methylation analysis of CLPD-NK according to molecular profile

A: Comparison of 5-hydroxymethyl-cytosine (5hmC) level in CLPD-NK vs. reactive NK group (Mann-Whitney test). **B:** Evaluation of *SOCS3* methylation by methyl-specific-polymerase chain reaction (MSP) of 7 CLPD-NK patients.

Figure 7: Transcriptional analysis by RNA-sequencing of 7 CLPD-NK patients

A: Hierarchical clustering and **B:** Principal component analysis (PCA) performed from the 1000 most variable genes of the dataset. Samples are clustered into 2 groups (G1 vs. G2) **C:** PCA gene contribution. **D:** Volcano plot representing the significantly overexpressed (red) and underexpressed (green) gene in G1 vs. G2 group (Q -value > 0.05 , Fold change > 2). **E:** Gene Set Enrichment Analysis (GSEA): *STAT3*-targets related signatures and top 10 overrepresented signatures among the differentially expressed genes in G1 vs. G2, ranked by Q -value **F:** GSEA plot for *STAT3* target signatures in G1 group **G:** GSEA plot for Angioimmunoblastic lymphoma signature in G2 group. (NES: normalized enrichment score)

Fig1

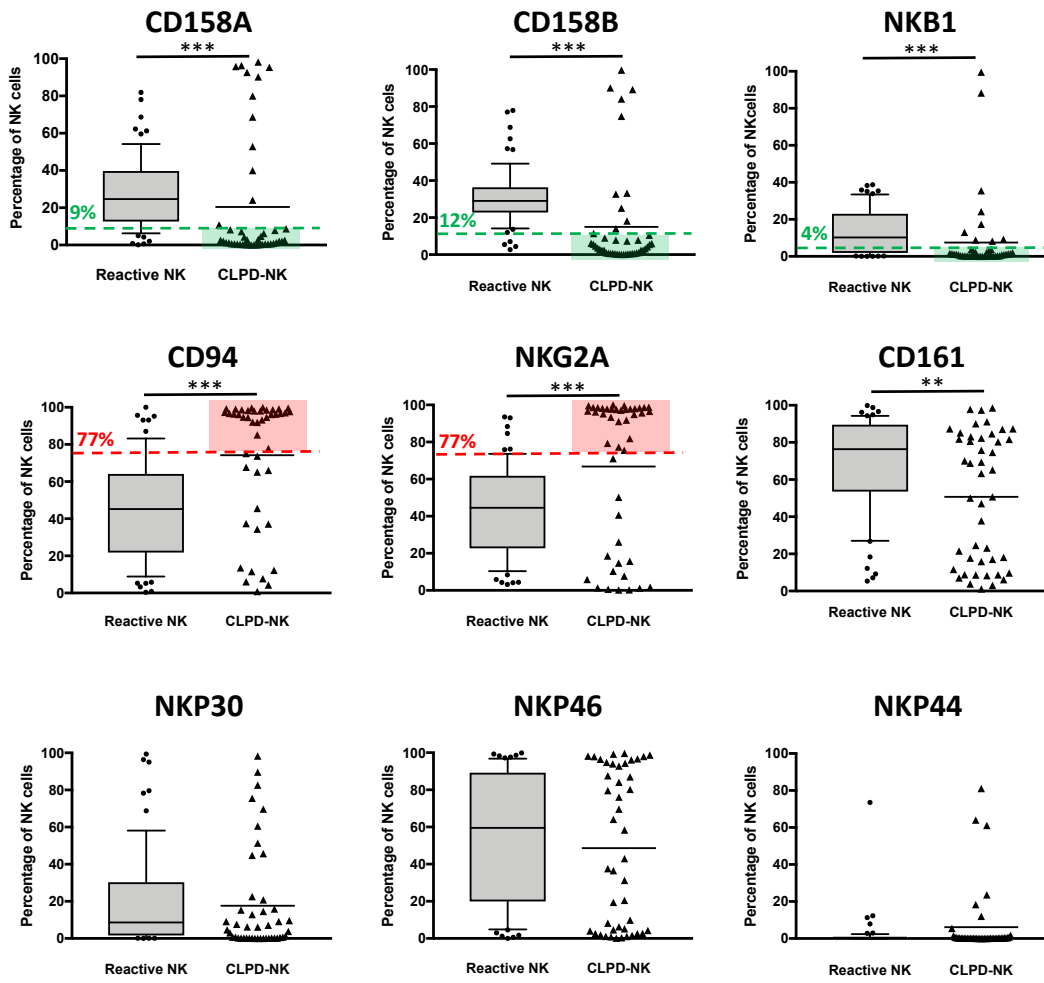
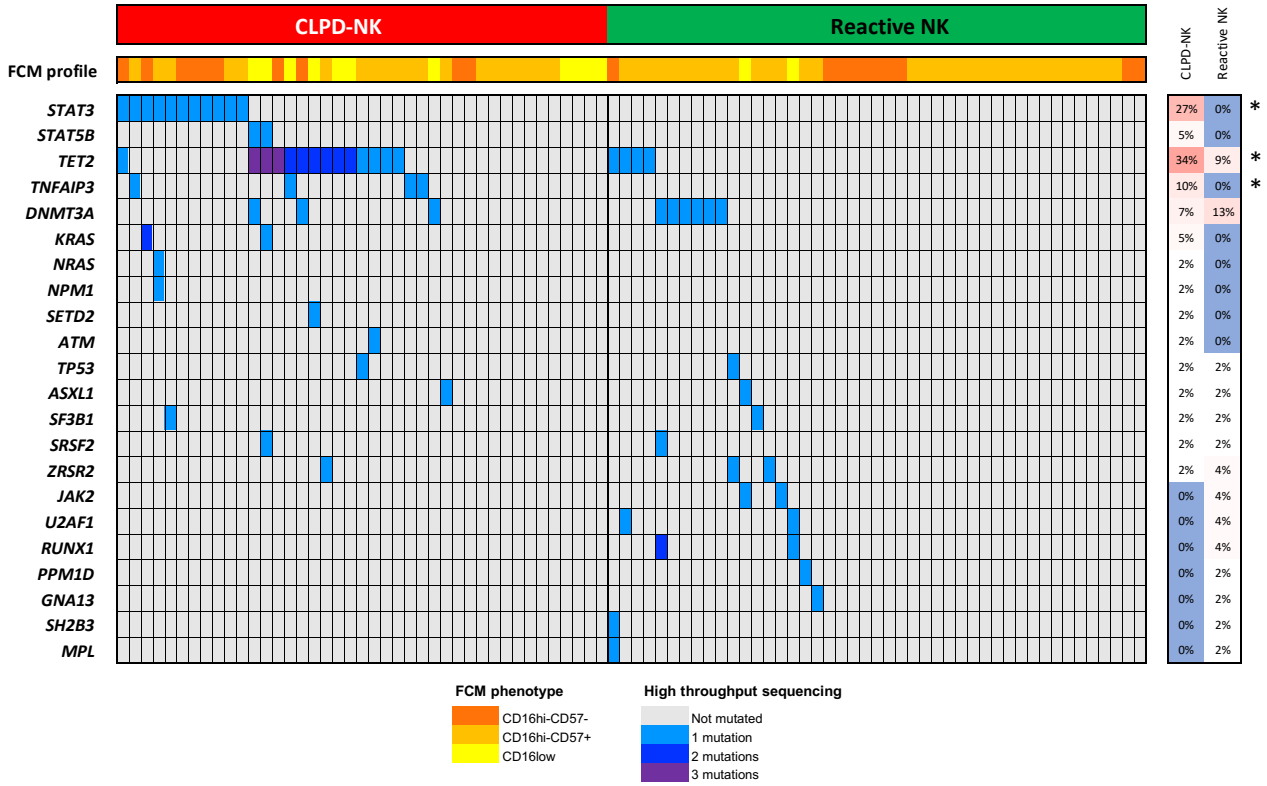


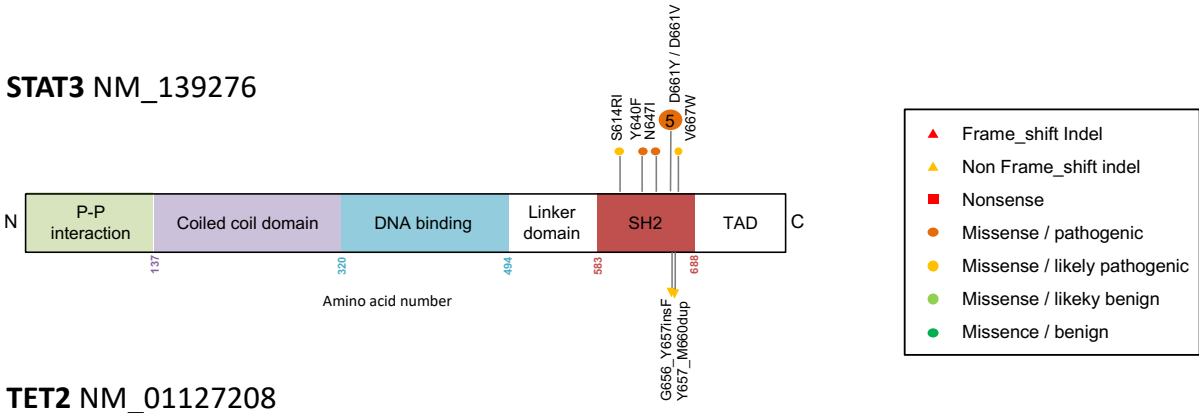
Fig2

A

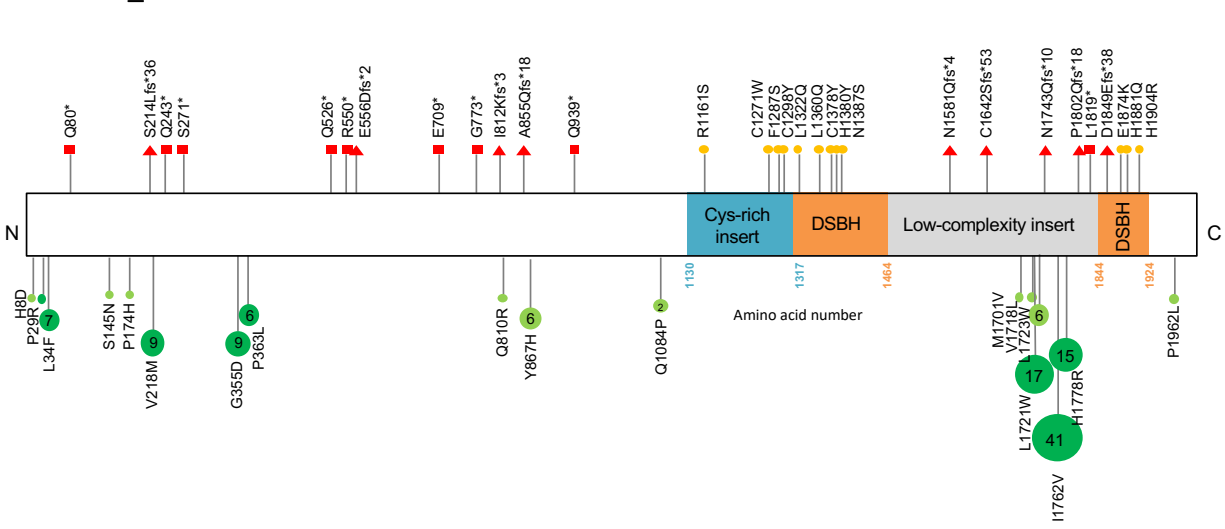


B

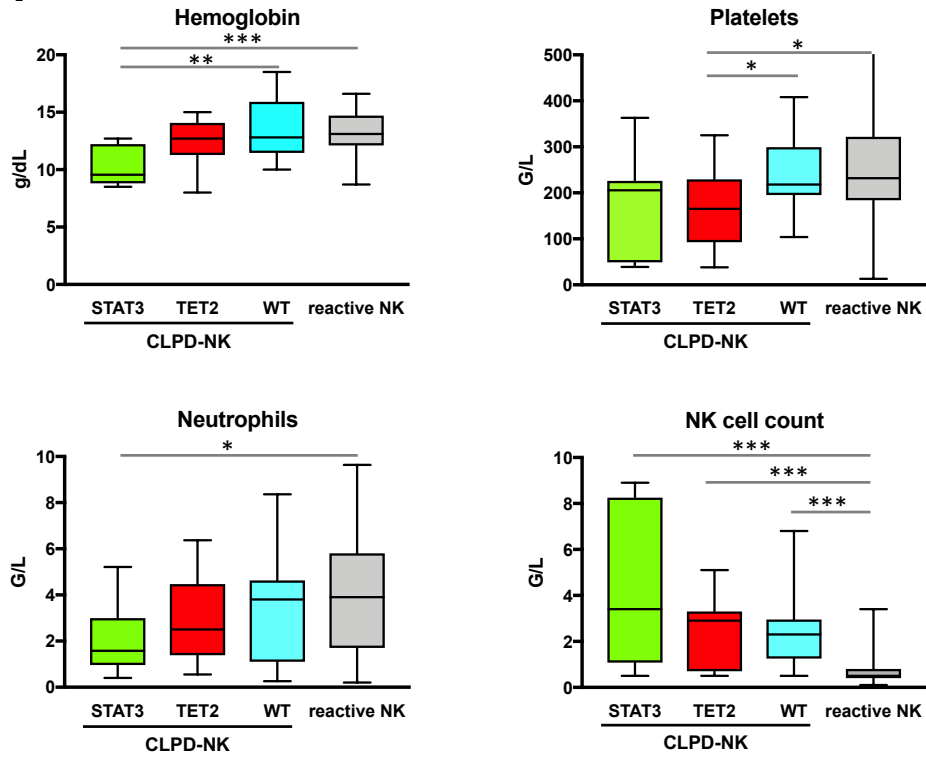
STAT3 NM_139276



TET2 NM_01127208



A



B

| | CLPD-NK | | | | |
|------------------------------------|---------------|--------------|------------|------------------|----------------|
| | STAT3 N=11 | TET2 N=13 | WT N=17 | Reactive N=45 | |
| Complete blood count | | | | | |
| Anemia (Hb < 12g/dL) | 36% | 31% | 24% | 16% | |
| Thrombocytopenia (Plt < 150G/L) | 27% | 38% | 12% | 4% | |
| Neutropenia (Neutrophils < 1.5G/L) | 36% | 23% | 24% | 22% | |
| NK cells > 2G/L | 45% | 46% | 41% | 2% | |
| Phenotype | | | | | |
| "NK Cytotoxic" CD16hi CD57- | 54% | 15% | 12% | 19% | $P = 0.012$ * |
| "NK Memory" CD16hi CD57+ | 45% | 38% | 59% | 72% | |
| CD16low | 0% | 46% | 29% | 0% | $P = 0.041$ ** |
| Clinical presentation | | | | | |
| Autoimmune disorders | 9% | 23% | 29% | 20% | |
| Other hematologic malignancies | 0% | 38% | 12% | 22% | $P = 0.024$ ** |

Fig4

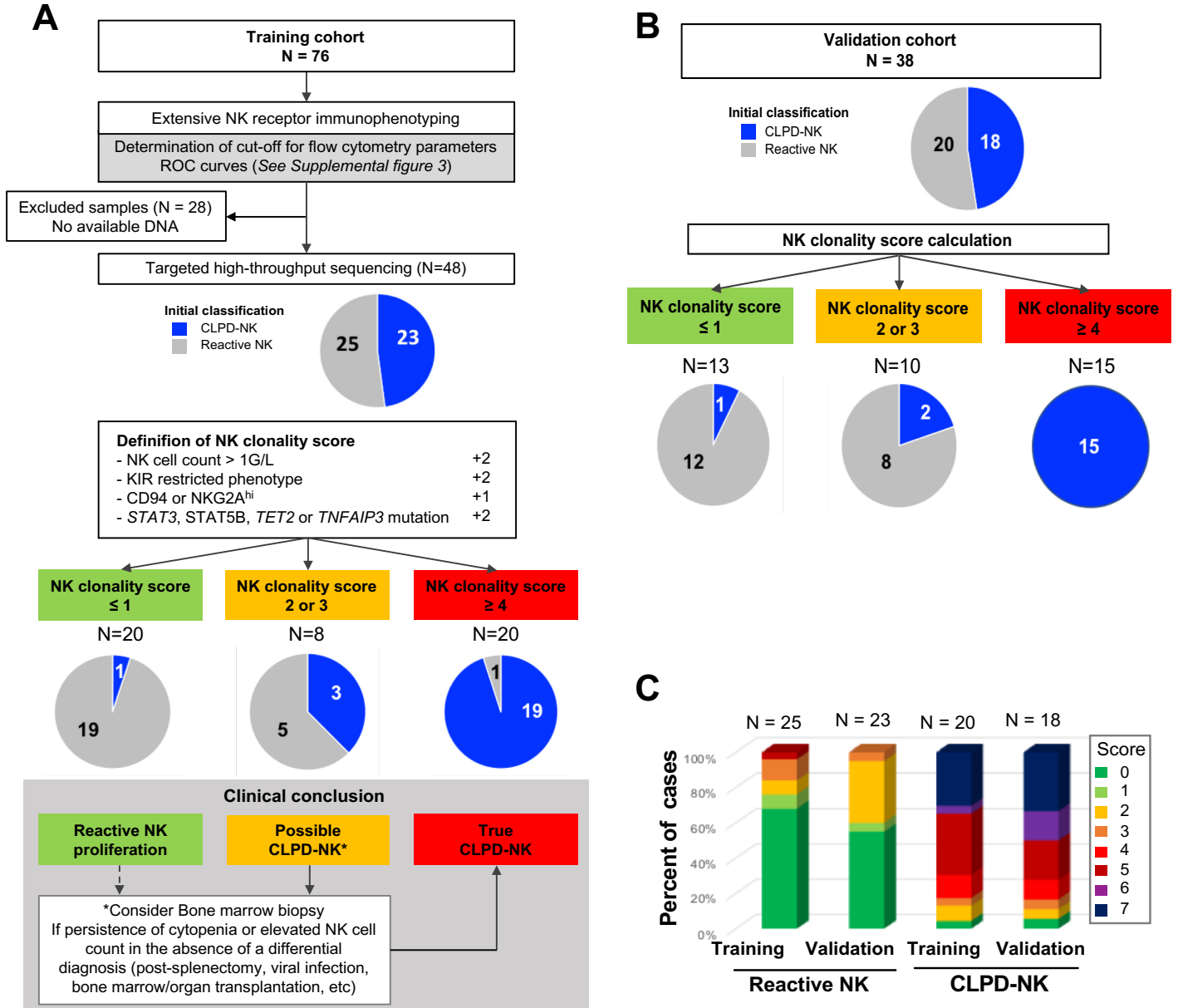


Fig5

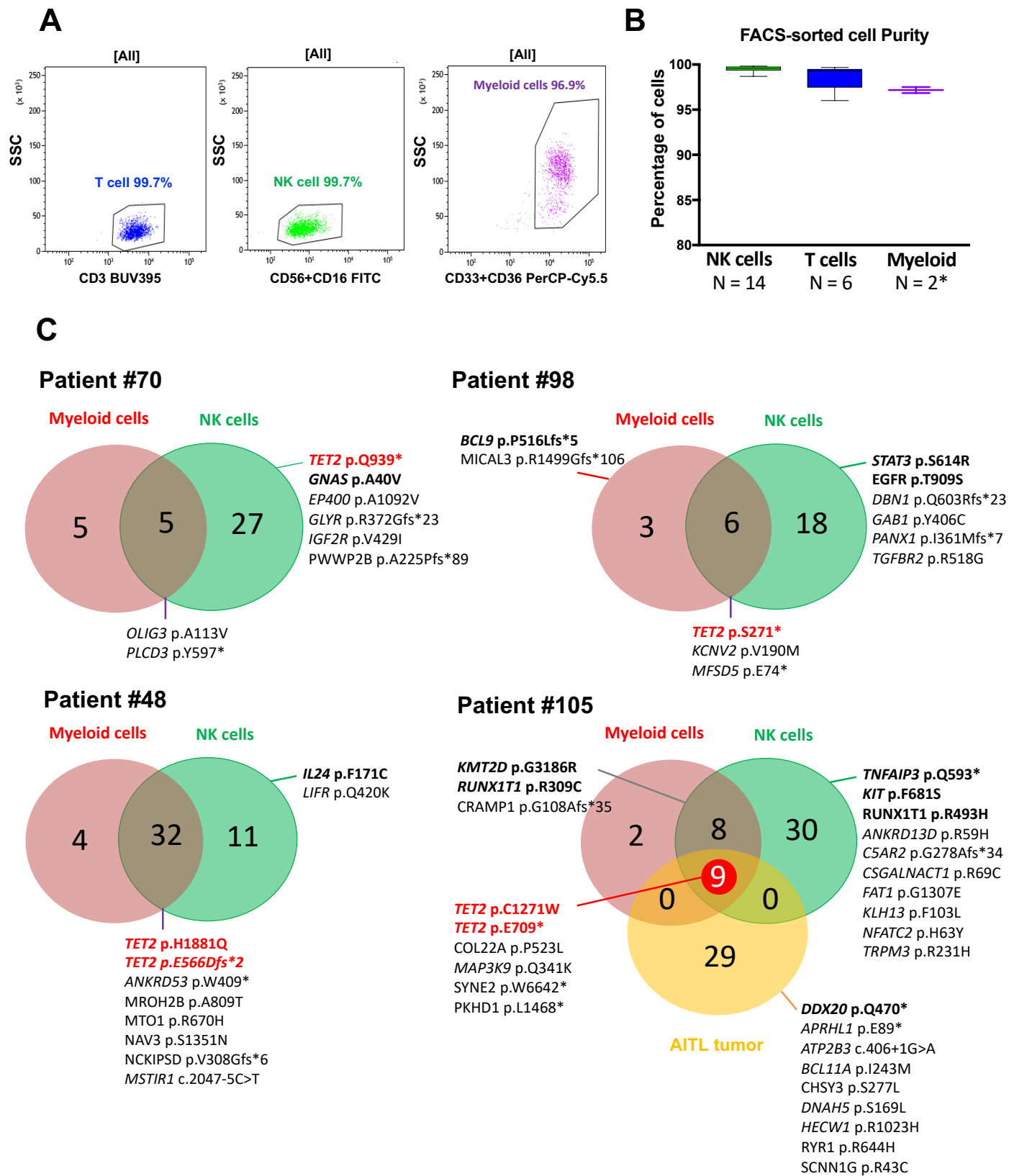


Fig6

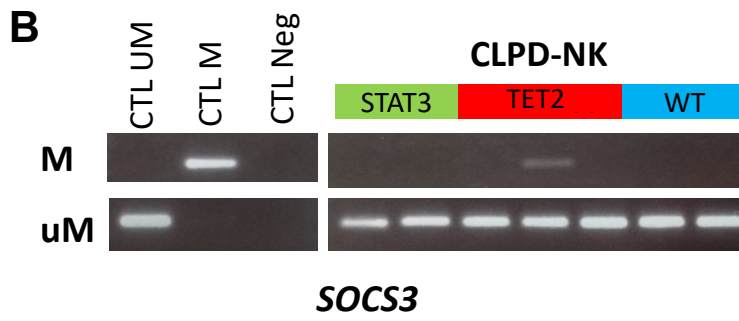
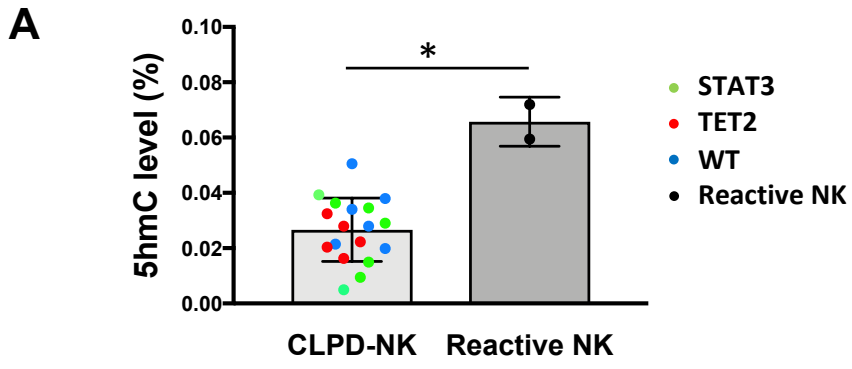
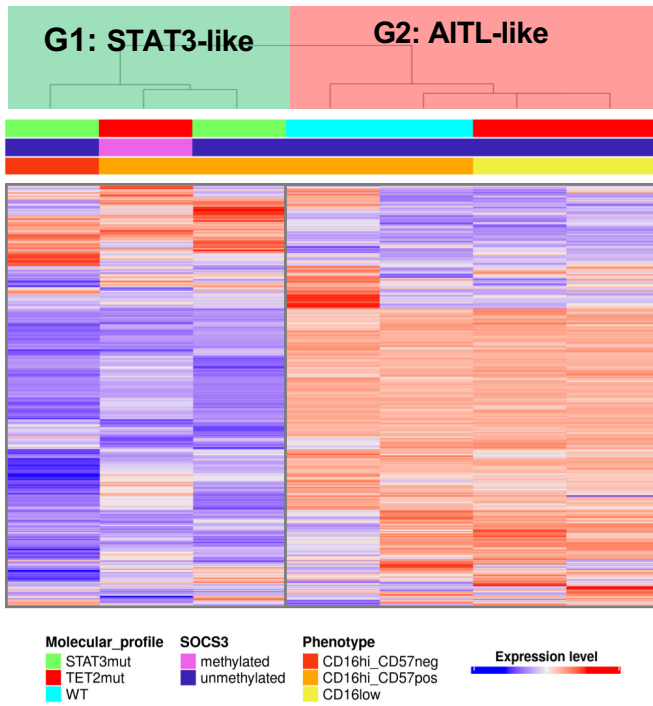
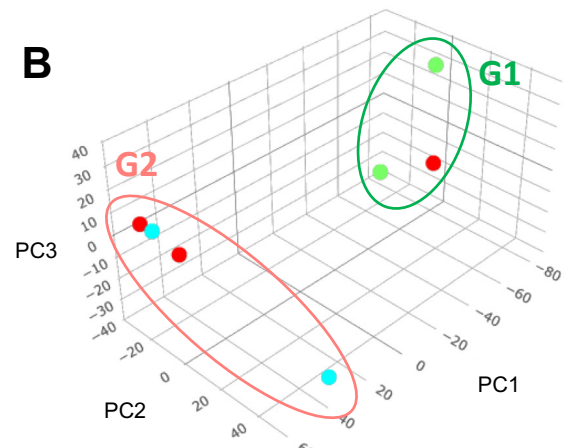


Fig7

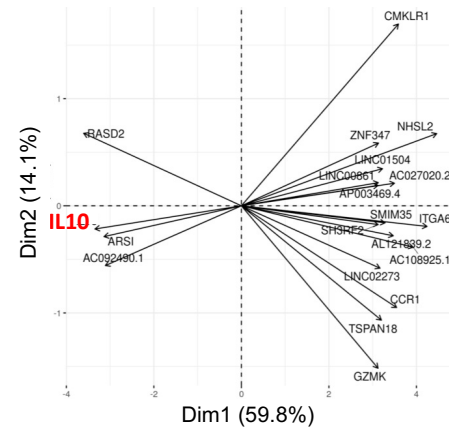
A



B



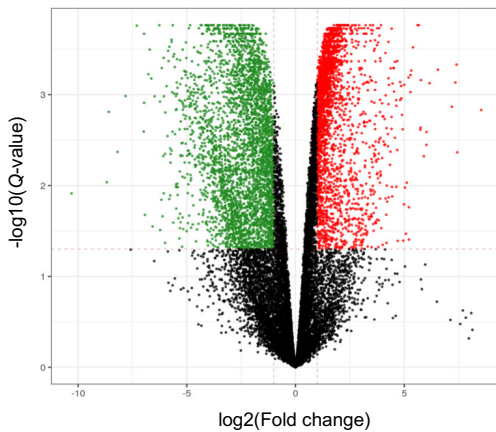
C



D

G1_STAT3-like vs. G2_AITL-like

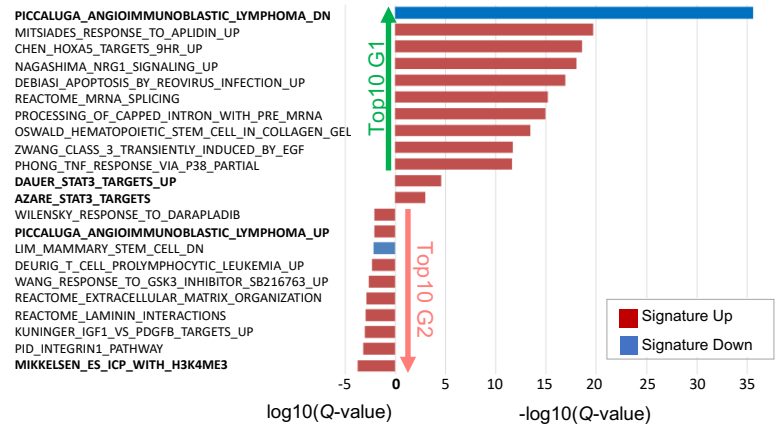
2,401 overexpressed and 3,461 underexpressed



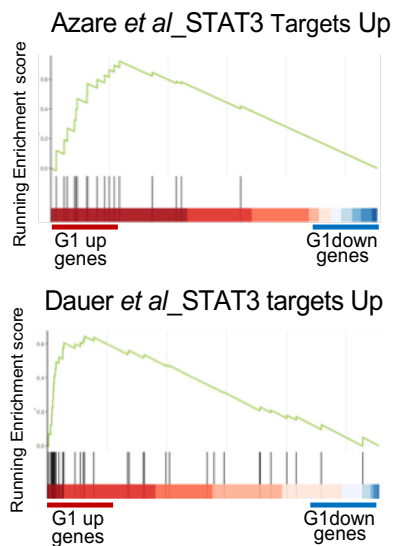
E

G1_STAT3-like vs. G2_AITL-like

Enrichment signature



F



G

