



**HAL**  
open science

# Single-cell profiling identifies pre-existing CD19-negative subclones in a B-ALL patient with CD19-negative relapse after CAR-T therapy

Tracy Rabilloud, Delphine Potier, Saran Pankaew, Mathis Nozais, Marie Loosveld, Dominique Payet-Bornet

## ► To cite this version:

Tracy Rabilloud, Delphine Potier, Saran Pankaew, Mathis Nozais, Marie Loosveld, et al.. Single-cell profiling identifies pre-existing CD19-negative subclones in a B-ALL patient with CD19-negative relapse after CAR-T therapy. *Nature Communications*, 2021, 12, pp.865. 10.1038/s41467-021-21168-6 . hal-03215356

**HAL Id: hal-03215356**

**<https://hal.science/hal-03215356v1>**

Submitted on 3 May 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.

# Single-cell profiling identifies pre-existing CD19-negative subclones in a B-ALL patient with CD19-negative relapse after CAR-T therapy

Tracy Rabilloud<sup>1,3</sup>, Delphine Potier <sup>1,3</sup>, Saran Pankaew<sup>1</sup>, Mathis Nozais <sup>1</sup>, Marie Loosveld <sup>1,2,4</sup>  & Dominique Payet-Bornet <sup>1,4</sup> 

Chimeric antigen receptor T cell (CAR-T) targeting the CD19 antigen represents an innovative therapeutic approach to improve the outcome of relapsed or refractory B-cell acute lymphoblastic leukemia (B-ALL). Yet, despite a high initial remission rate, CAR-T therapy ultimately fails for some patients. Notably, around half of relapsing patients develop CD19 negative (CD19<sup>neg</sup>) B-ALL allowing leukemic cells to evade CD19-targeted therapy. Herein, we investigate leukemic cells of a relapsing B-ALL patient, at two-time points: before (T1) and after (T2) anti-CD19 CAR-T treatment. We show that at T2, the B-ALL relapse is CD19 negative due to the expression of a non-functional *CD19* transcript retaining intron 2. Then, using single-cell RNA sequencing (scRNAseq) approach, we demonstrate that CD19<sup>neg</sup> leukemic cells were present before CAR-T cell therapy and thus that the relapse results from the selection of these rare CD19<sup>neg</sup> B-ALL clones. In conclusion, our study shows that scRNAseq profiling can reveal pre-existing CD19<sup>neg</sup> subclones, raising the possibility to assess the risk of targeted therapy failure.

<sup>1</sup> Aix Marseille Université, CNRS, INSERM, CIML, Marseille, France. <sup>2</sup> APHM, Hôpital La Timone, Laboratoire d'Hématologie, Marseille, France. <sup>3</sup> These authors contributed equally: Tracy Rabilloud, Delphine Potier. <sup>4</sup> These authors jointly supervised this work: Marie Loosveld, Dominique Payet-Bornet. email: [loosveld@ciml.univ-mrs.fr](mailto:loosveld@ciml.univ-mrs.fr); [payet@ciml.univ-mrs.fr](mailto:payet@ciml.univ-mrs.fr)

**B**-cell acute lymphoblastic leukemia (B-ALL) is a hematologic malignancy arising from uncontrolled proliferation of B lymphoblasts. Although the 5-year overall survival rate reaches currently 80–90% for children, the prognosis of relapsed and refractory B-ALL is dismal, and those B-ALL remain the principal cause of cancer-related childhood mortality<sup>1,2</sup>. Hence, novel therapeutic options for these B-ALL cases are urgently warranted. In this context, one of the most innovative approach relies on chimeric antigen receptors (CARs)-T cell therapy<sup>3–6</sup>. CARs are artificial receptors including an extracellular antibody-derived domain recognizing a specific target fused to intracellular signaling domains derived from T cells<sup>7</sup>. Autologous T cells engineered to express CARs (CAR-T) acquire the ability to recognize and to kill efficiently cells harboring at their surface the target antigen. Regarding the treatment of relapsed/refractory B-ALL, adoptive T cells expressing a CAR directed against the B-cell-associated surface marker, CD19, have been used in several clinical trials<sup>8–12</sup>. Those studies showed that CAR-T therapies achieved high remission rates ranging from 70 to 90% at short-term. However, ~40% of responding B-ALL patients ultimately relapse. Two modes of relapse were observed, CD19-positive (CD19<sup>pos</sup>) and CD19-negative (CD19<sup>neg</sup>). The CD19<sup>pos</sup> relapse mainly results from the low potency and poor persistence of CAR-T cells<sup>13</sup>. Nevertheless, to overcome these issues, efforts are made to optimize clinical strategies and CAR-T cells engineering<sup>14,15</sup>. In CD19<sup>neg</sup> relapses, which account for around half of the cases (39% and 68% in respectively Gardner et al.<sup>9</sup> and Maude et al.<sup>8</sup> studies), leukemic cells lose CD19 epitope surface expression and thus escape CAR-mediated recognition, inhibiting B-ALL clearance. Recent studies showed that CD19 negativity results mainly from impaired expression of *CD19* mRNA (through de novo frameshift/missense *CD19* mutations, alternatively spliced CD19 mRNA species and hemizygous deletions spanning the *CD19* locus)<sup>16–19</sup>. However, no study has clearly demonstrated whether CAR-T cell therapy is directly responsible of CD19 transcription dysregulation or simply allows the emergence of minor CD19<sup>neg</sup> clones escaping CD19-targeted therapy. In the second hypothesis, detection of those CD19<sup>neg</sup> clones may represent a valuable predictor of CD19-negative relapse permitting to assess beforehand the risk for CAR-T treatment failure and to adapt accordingly the therapeutic strategy.

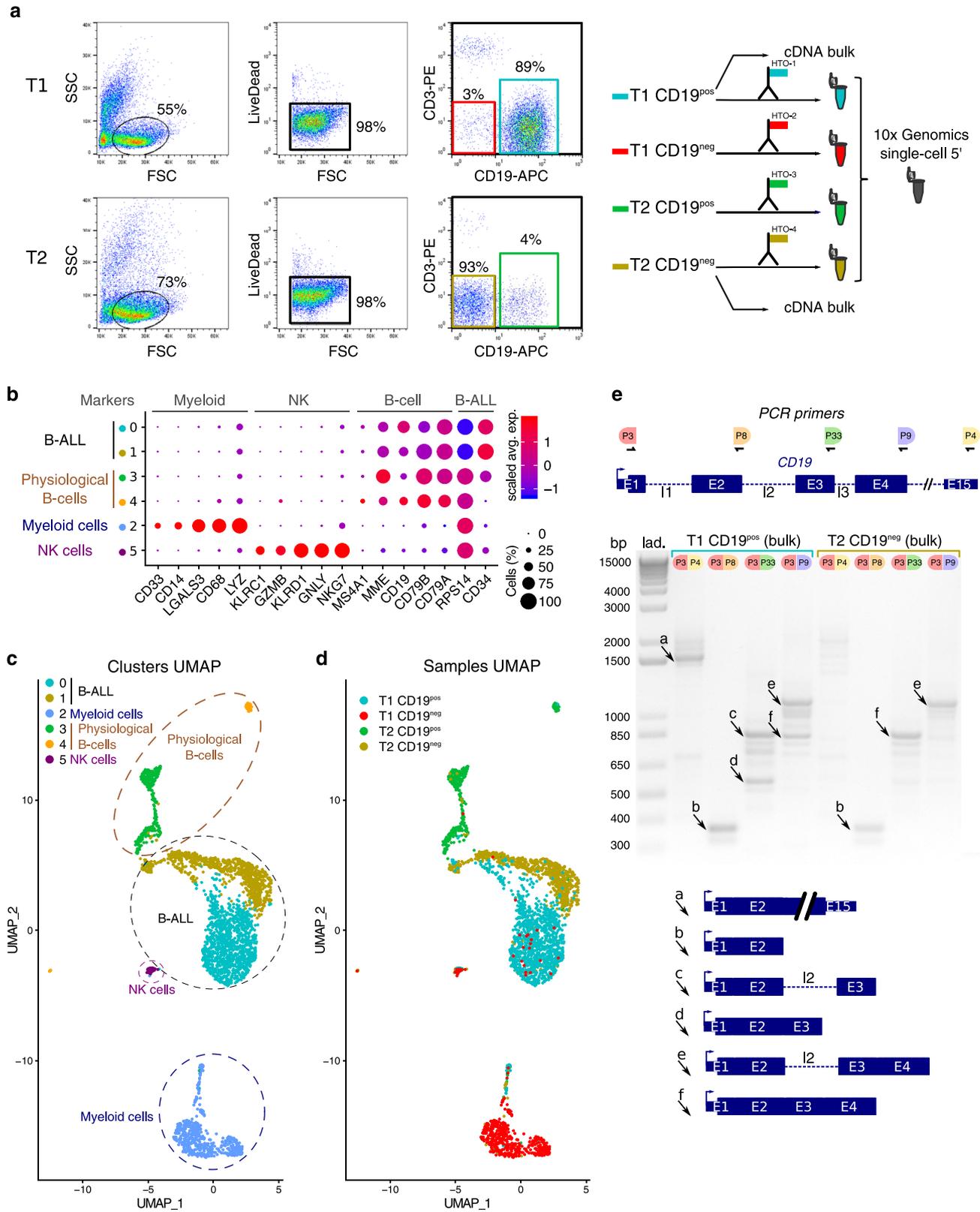
Here, we report the single-cell RNA-sequencing (scRNAseq) profiling of leukemic samples from a B-ALL patient, at two-time points, before and after CAR-T therapy. Our work reveals the presence of preexisting CD19<sup>neg</sup> B-ALL clones before the CAR-T treatment.

## Results

**Analysis of B-ALL samples prior and after CAR-T cell treatment.** We used scRNAseq approach to investigate leukemic cells of a B-ALL patient at two-time points; before and after anti-CD19 CAR-T cell therapy. The B-ALL patient underwent a relapse after chemotherapy and CD19<sup>pos</sup> leukemic cells from bone marrow (BM) were harvested (T1 cells = before CAR-T cell therapy; Supplementary Fig. 1a). Then, the patient was treated with anti-CD19 CAR-T. After an initial complete remission, the patient experienced a frank relapse, as noted by circulating blasts and an infiltration (>90%) of CD10<sup>pos</sup>CD19<sup>neg</sup> leukemic cells in the BM (Supplementary Fig. 1b) from which cells were harvested (T2 cells = after CAR-T cell therapy). Therefore, we asked whether these CD19<sup>neg</sup> B-ALL clones were present before CAR-T cell therapy. To address this issue, T1 and T2 samples were sorted according to forward scatter and side scatter, a cell viability marker, CD3 and CD19 surface expression (Fig. 1a). We obtained

four tubes of sorted cells: T1-CD19<sup>pos</sup>, T1-CD19<sup>neg</sup>, T2-CD19<sup>pos</sup>, and T2-CD19<sup>neg</sup>. Each of those samples were marked with a distinct anti-CD45-hashtag oligonucleotide (HTO) antibody<sup>20</sup> (Supplementary Table 1), then mixed and analyzed by scRNAseq using the 5' 10× Genomics technology<sup>21</sup> (Fig. 1a). After libraries sequencing and data preprocessing, we applied the Seurat graph-based clustering algorithm and identified six main clusters. Then, we used the UMAP nonlinear dimensionality reduction method, to visualize cell transcriptome heterogeneity<sup>22</sup>. According to various gene markers, we assigned cell type to clusters (Fig. 1b, c). Non-B cells, i.e., myeloid and NK cells are located in clusters 2 and 5, respectively. Physiological B cells are in clusters 3 and 4, while B-ALL cells are in clusters 0 and 1. Sample demultiplexing assigned 1189, 524, 764, and 442 cells to T1-CD19<sup>pos</sup>, T1-CD19<sup>neg</sup>, T2-CD19<sup>pos</sup>, and T2-CD19<sup>neg</sup> samples, respectively, and allowed us to visualize sample of origin for each cell on the UMAP plot. Notably, T1-CD19<sup>pos</sup> and T2-CD19<sup>neg</sup> tumoral cells were detected mainly to clusters 0 and 1, respectively (Fig. 1d). CaSpER tool<sup>23</sup> and B-allele frequency (BAF) analysis indicate that cells from these two clusters displayed common copy number variants (CNV), such as chromosome 9q deletion, indicating that CD19<sup>pos</sup> and CD19<sup>neg</sup> B-ALL are related (Supplementary Fig. 2). The most differentially expressed transcription factor between clusters 0 and 1 was *KLF6* gene, a well-known tumor suppressor gene<sup>24,25</sup> (Supplementary Fig. 3). Strikingly, CD19 mRNA transcripts were detected in T2-CD19<sup>neg</sup> tumoral cells albeit at lower level than in T1-CD19<sup>pos</sup> tumoral cells (Supplementary Fig. 4a, b). The analysis of bulk cDNA from T1-CD19<sup>pos</sup> and T2-CD19<sup>neg</sup> sorted cells shows that T2-CD19<sup>neg</sup> cells do not express full-length transcripts (Fig. 1e). We observed that both samples express a previously described<sup>18</sup> nonfunctional CD19 transcript retaining intron 2 (Supplementary Fig. 4c). However, in T2-CD19<sup>neg</sup> cells only this nonfunctional isoform is expressed, explaining the absence of CD19 protein despite the presence of CD19 mRNA.

**Detection of preexisting CD19<sup>neg</sup> B-ALL cells.** Importantly, in the UMAP plot, we clearly distinguish in tumoral clusters, 20 cells belonging to T1-CD19<sup>neg</sup> sample (Figs. 1d and 2a, red dots) indicating that those cells are CD19<sup>neg</sup> B-ALL. Yet, to further establish that those cells were not mis-assigned, we compared the gene expression profiles of B-ALL cells versus physiological B cells, myeloid or NK cells. Resulting heatmaps of the most differentially expressed genes (Fig. 2b and Supplementary Fig. 5), as well as dot plot for marker genes expression (Fig. 2c) show that these 20 T1-CD19<sup>neg</sup> cells are alike B-ALL cells. Moreover, CaSpER tool succeeded to detect 9q deletion for half of the 20 T1-CD19<sup>neg</sup> cells (Supplementary Fig. 2a). In the same line, BAF analysis of *RPS14* SNP gave a similar result (loss of heterozygosity of *RPS14* allele) for T1-CD19<sup>neg</sup> B-ALL and the other T1/T2 B-ALL cells (Supplementary Fig. 2b). Altogether, these data confirm that T1-CD19<sup>neg</sup> cells in B-ALL clusters are true leukemic cells. However, as CD19 negativity was defined by FACS during cell sorting, we validated this observation at the molecular level. Thus, the 20 putative CD19<sup>neg</sup> B-ALL clones were backtracked using their cell identifier 10× barcode (BC), and CD19 mRNA expression was assessed by nested-PCR (Fig. 2d). As shown in Fig. 2e, one specific CD19 mRNA band is clearly detected for the three positive control cells (CD19<sup>pos</sup> B-ALL from T1-CD19<sup>pos</sup> samples), conversely only background levels of various PCR products were observed for negative control cells (CD19<sup>neg</sup> B-ALL from T2-CD19<sup>neg</sup> samples). Among the 20 CD19<sup>neg</sup> B-ALL clones spotted at T1, 10 clones were clearly similar to control CD19<sup>neg</sup> B-ALL cells (i.e., background level



of PCR products and absence of a clear specific CD19 mRNA band; Fig. 2e). The PCR profiles of the remaining ten clones were ambiguous, as a faint CD19 mRNA band was detected (Supplementary Fig. 6). Overall, our results demonstrate that CD19<sup>neg</sup> B-ALL clones were present before anti-CD19 CAR-T therapy, which, per se, is thus not responsible for CD19 mRNA dysregulation.

**Discussion**

In addition to somatic mutations found in exons 2–5 (ref. 16), mis-splicing of CD19 mRNA is an alternative way to impede CD19 expression at the cell surface<sup>18,19</sup>. In this context, Sotillo et al.<sup>19</sup> showed that skipping of CD19 exon 2 is due to down-regulation of SRSF3 splicing factor, which is normally involved in exon 2 retention. Herein, our analysis of CD19 sequencing reads

**Fig. 1 CD19<sup>neg</sup> B-ALL relapse following CAR-T therapy.** **a** Cell sorting strategy. Cells before (T1) and after (T2) CAR-T treatment were gated according to FSC/SSC profile (left FACS plots). Then, live cells (middle FACS plots) were analyzed according to CD3 and CD19 expression (right FACS plots; see also Supplementary Fig. 9). Gates of the four sorted subpopulations T1-CD19<sup>pos</sup>, T1-CD19<sup>neg</sup>, T2-CD19<sup>pos</sup>, and T2-CD19<sup>neg</sup> are highlighted in cyan, red, green, and gold, respectively. Those sorted subpopulations were labeled with a specific anti-CD45-HTO antibody, then multiplexed and analyzed by scRNAseq using the 10× Genomics single-cell 5' technology. Some cells from T1-CD19<sup>pos</sup> and T2-CD19<sup>neg</sup> samples were used in bulk to prepare cDNA. **b** Seurat Dotplot showing the expression level of marker genes in each cluster. Dot size represents the percentage of cell expressing the gene of interest, while dot color represents the scaled average expression (Scaled Avg. Exp.) of the gene of interest across the various clusters (a negative value corresponds to an expression below the mean expression). We used CD34 and RPS14 expression as tumoral markers. Indeed FISH analysis revealed a 5q32 deletion, explaining RPS14 lower expression in B-ALL cells. **c** UMAP visualization of the six main clusters and their corresponding cell types of T1 and T2 sorted samples. **d** UMAP visualization of the four demultiplexed samples: T1-CD19<sup>neg</sup>, T1-CD19<sup>pos</sup>, T2-CD19<sup>neg</sup>, and T2-CD19<sup>pos</sup>. **e** Agarose gel of CD19 cDNA amplified products using exons-specific primer sets depicted on the top panel. PCR were performed with bulk cDNA from T1-CD19<sup>pos</sup> cells and T2-CD19<sup>neg</sup> cells. Agarose gel data are representative of two independent experiments. Lane "Lad" is the 1 kb DNA size marker. Schematic representations of PCR products indicated by "a"- "f" arrows are shown below the gel.

did not reveal any somatic mutation within exons 2–5. However, we showed that loss of CD19 is due to the expression of a non-functional *CD19* isoform retaining intron 2. This alteration was previously described as a recurrent event<sup>18</sup>, yet further investigations would be required to characterize the mechanism of intron 2 retention.

Here, using scRNAseq approach, we detected CD19<sup>neg</sup> B-ALL cells before CAR-T treatment. A previous study suggested the presence of such preexisting clones<sup>6</sup>. Yet, in this study CD19<sup>neg</sup> cells (~7%) were detected by FACS after a gating strategy based on live cells that were CD45<sup>pos</sup>, SSC low, and CD34<sup>pos</sup> (ref. 6). Thus, this analysis did not formally exclude some non-B-ALL cells, such as myeloblasts. We attempted to identify by FACS CD19<sup>neg</sup> B-ALL clones; however, those potential clones represent only 0.03% of live cells and were not clearly clustered (Supplementary Fig. 7), and thus require molecular biology to convincingly validate their leukemic origin.

In conclusion, our study demonstrates that a Darwinian-like selection of preexisting CD19<sup>neg</sup> B-ALL cells is a mechanism for CD19<sup>neg</sup> B-ALL relapse after CAR-T cell therapy. A scRNAseq approach can unambiguously detect minor CD19<sup>neg</sup> subclones before CAR-T cell therapy. Thus, analysis of additional B-ALL patients might further demonstrate that such approach represents a useful tool to identify patients at high risk of relapse after CAR-T cell therapy.

## Methods

**Patient's sample.** A child diagnosed with B-ALL was first enrolled in a conventional chemotherapy protocol (CAALL-F01 protocol). During maintenance phase, the patient experienced a frank relapse (relapse 1). Then, anti-CD19 CAR-T therapy was initiated. At day 28 post-CAR-T cells infusion, the patient was in complete remission; however, at day 90, the patient relapsed as noted by an important infiltration of CD10<sup>pos</sup>CD19<sup>neg</sup> leukemic cells in the BM (relapse 2). BM samples were collected before (relapse 1) and after (relapse 2) CAR-T therapy. Informed consent for use of diagnostic and relapse specimens for future research was obtained in accordance with the Declaration of Helsinki. The study was approved by the local ethic committee of "Assistance Publique des Hôpitaux de Marseille".

**Flow cytometry.** BM samples were tested within 24 h of collection. Briefly, following NH<sub>4</sub>Cl red blood cell lysis, cells were stained with conjugated antibodies (CD45-KrO, CD19-PC7, CD34-ECD, CD10-FITC, CD38-APC, CD58-PC5, CD66c-PE, and CD123-PE) for 30 min at 4 °C and washed twice with FACS buffer (PBS, 2% FCS, and 1 mM EDTA). Data were acquired on Navios cytometer (Beckman Coulter) and analysis was made using Kaluza software (Beckman Coulter). For cell sorting, cells stained with CD19-APC, CD3-PE, and LIVE/DEAD™ Fixable Aqua Dead Cell Stain Kit (Invitrogen) were FACS-sorted using BD-Influx™ cytometer (Becton Dickinson). Antibodies used for flow cytometry are listed in Supplementary Table 2.

**Antibody-HTO conjugation.** Anti-CD45 antibody (clone HI30) provided by BioLegend was irreversibly conjugated to HTO (sequences shown in Supplementary Table 1), using an inverse electron-demand Diels–Alder reaction between methyltetrazine (mTz) and *trans*-cyclooctene (TCO) as previously described<sup>26</sup>, and

according to the protocol found at [www.cite-seq.com](http://www.cite-seq.com). Briefly, 10 nmol of HTO oligonucleotides were incubated in BBS buffer (50 mM borate, 150 mM NaCl, pH 8.5) containing 10% DMSO and 10 mM TCO-PEG4-NHS (Click Chemistry Tools); while anti-CD45 antibodies were incubated in BBS buffer containing 0.1 mM mTz-PEG4-NHS (Click Chemistry Tools). After 30 min at room temperature, reactions were quenched with 20 mM glycine pH 8.5. TCO-PEG4-HTO and mTz-PEG4-antibodies were then mixed (30 pmol of HTO per 1 µg of antibody) and incubated overnight at 4 °C. At the end of reaction, TCO-PEG4-glycine (1 mM final concentration) was added. TCO-PEG4-glycine was obtained by mixing in water 10 mM TCO-PEG4-NHS with 20 mM glycine pH 8.5. Finally, antibody-HTO samples were concentrated and cleaned-up using an Amicon Ultra-0.5 Centrifugal Filter Unit with 50 kDa cutoff membrane (Merck Millipore). Antibodies-HTO were validated by FACS (Supplementary Fig. 8).

**Single-cell RNA-sequencing experiment.** FACS-sorted cells were labeled with anti-CD45 conjugated to HTO (Supplementary Table 1). Cells were isolated using the Chromium technology from 10× Genomics. Briefly, 5000 cells of each sorted population were pooled, loaded on a Chromium Chip A and droplet-encapsulated with a chromium controller. Single-cell cDNA synthesis and sequencing libraries were prepared with Single-Cell 5' Library & Gel Bead kit (10× Genomics), according to manufacturer's instructions. Libraries were sequenced using a Next-seq500 and the following parameters, Read1: 26 cycles, i7: 8 cycles, and Read2: 57 cycles.

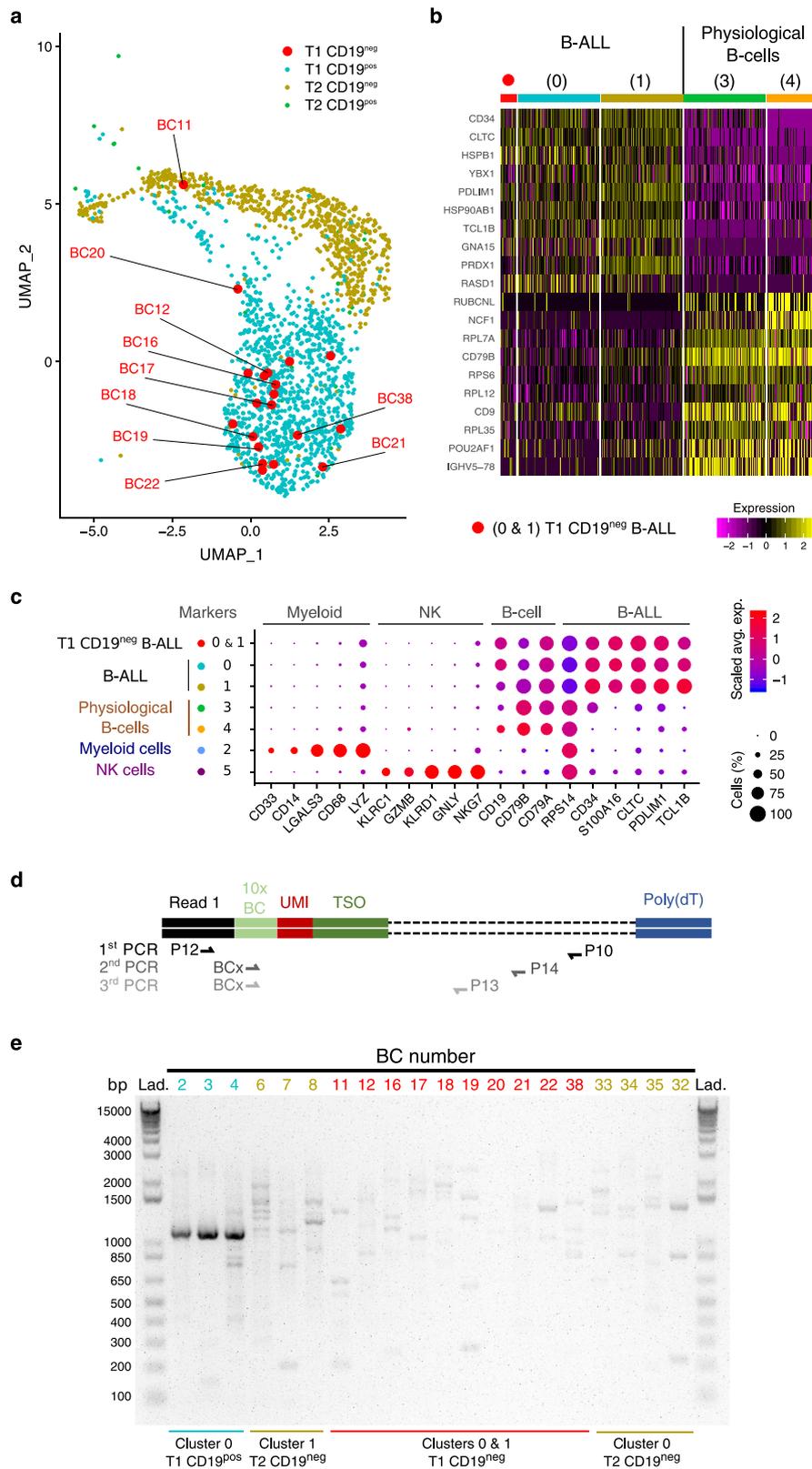
**scRNAseq data analysis.** For data preprocessing, mRNA library reads were aligned to the GRCh38 version of the human genome and quantified using Cell Ranger count (version 3.0.1). Antibody counts for cell hashing were quantified using CITE-seq-count (version 1.4.3, NYGCTech, <https://github.com/Hooohm/CITE-seq-Count>), using default parameters. The produced mRNA and HTO data matrices were imported into R (v3.5.3), and downstream analysis were performed with the Seurat package (v3.0.1)<sup>22,27</sup>.

In order to perform samples demultiplexing, HTOs for each cell were normalized using a centered log ratio transformation across cells. Cells were demultiplexed using the Seurat HTODemux function, and cell doublets and background empty droplets subsequently removed. Before mRNA expression analysis, we filtered out low quality cells from the mRNA matrix (i.e., cells expressing < 200 genes or >10% of mitochondrial-associated genes). Genes expressed in less than three cells were removed. Then, expression raw counts were normalized and scaled using respectively the Seurat NormalizeData and ScaleData functions. The top 2000 variable genes were selected (FindVariableFeatures function; selection method = "vst") to perform a dimensionality reduction using the principal component analysis method (RunPCA function), and a UMAP was computed using the 20 first components (RunUMAP function). We used the Seurat graph-based clustering algorithm to identify clusters of cells (using 20 dimensions and a coarse grained resolution of 0.1). The nonparametric Wilcoxon rank-sum test (FindMarkers seurat fonction) was used for all differential gene expression analysis between two groups of cells.

To investigate large-scale CNV, we used CaSPER tool<sup>23</sup>. Briefly, the scRNAseq experiment bam file was subdivided using 10× Genomics subset-bam, to obtain three bam files, respectively, containing the reads for cells in cluster 0, in cluster 1 and in clusters 2–5. Then BAFExtract was used to calculate BAF for each bam, and finally CaSPER was run using physiological cells (clusters 2–5) as control.

To infer gene regulatory network responsible for clonal transcriptional changes underlying tumoral evolution, we applied SCENIC method<sup>28</sup>, using the filtered matrix raw UMI counts.

**Reverse transcription-PCR.** RNA from T1-CD19<sup>pos</sup> and T2-CD19<sup>neg</sup> samples were extracted using RNAeasy mini kit (Qiagen), and bulk cDNA was synthesized using high capacity cDNA reverse transcription kit (Applied Biosystems). Then CD19 transcripts were assessed by PCR using specific primers (Supplementary Table 1) and



Herculase II enzyme (Agilent technologies). The following PCR program was applied: 95 °C 1 min, 35 cycles (95 °C 20 s, 56 °C 20 s, and 68 °C 90 s), 68 °C 4 min. Sanger sequencing of PCR products was performed by Eurofins Genomics.

**Backtracking of CD19<sup>neg</sup> clones.** Investigation of CD19 cDNA of individual cells was performed by nested-PCRs using amplified 10x barcoded full-length cDNA. Three successive PCR were carried out with Herculase II enzyme

(Agilent technologies) and with the following PCR program: 95 °C 1 min, 20 cycles (95 °C 20 s, 56 °C 20 s, and 68 °C 90 s), 68 °C 4 min (for the third PCR only 15 cycles were performed). In the first PCR, CD19 transcript was amplified with “partial Read1” primer (or P12) and CD19 primer P10. Cell-specific CD19 cDNA amplifications were carried out with primers comprising the 10x cell BC (BC primers) and P14 or P13 for respectively second and third PCR (Supplementary Table 1). For these nested-PCR, 1 µl aliquots of primary or secondary PCR were used. PCR products were analyzed by 1% agarose gel electrophoresis.

**Fig. 2 Detection of CD19<sup>neg</sup> B-ALL clones before CAR-T treatment.** **a** UMAP plot focused on tumoral cells (clusters 0 and 1) and colored according to sample of origin. The 20 cells from T1-CD19<sup>neg</sup> samples are highlighted by larger red dots. Sequences of 10× cell barcode (BC) of interest are indicated in Supplementary Table 1. **b** Heatmap showing normalized and scaled expression level of 20 genes differentially expressed between tumoral clusters (0 and 1), and either immature B cells (cluster 3) or mature B cells (cluster 4). Gene expression profiles of the 20 T1-CD19<sup>neg</sup> cells are shown independently to the other B-ALL cells. Clusters 0, 1, and 3 were down-sampled to 100 cells for a better readability. Number of expressed genes detected in the various clusters are shown in Supplementary Fig. 10. **c** Dot plot of marker genes expression, same as Fig. 1b except that T1-CD19<sup>neg</sup> cells from clusters 0 and 1 were analyzed independently, and that some differentially expressed genes between tumoral versus physiological cells were added. **d** Nested-PCR strategy to detect CD19 transcript in single cell. **e** Agarose gel of CD19 cDNA amplified by nested-PCR. Lanes are labeled according to the cell BC number for which the sample of origin is indicated at the bottom of the gel. Data are representative of independent backtracking experiments that were performed twice for cells with BC number 32, 33, 34, and 35; and three times for all other cells. Localizations on the UMAP plot of backtracked T2-CD19<sup>neg</sup> cells are shown in Supplementary Fig. 11.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### Data availability

Fastq raw data are available in SRA (<https://www.ncbi.nlm.nih.gov/sra?term=SRP269742>). Processed data can be accessed from the NCBI Gene Expression Omnibus database (accession code GSE153697). Human transcriptome reference used for our analysis is available at 10× Genomics website (<https://cf.10xgenomics.com/supp/cell-exp/refdata-cellranger-GRCh38-3.0.0.tar.gz>) or in zenodo (<https://zenodo.org/record/4114854/files/refdata-cellranger-GRCh38-3.0.0.tar.gz?download=1>). The cisTarget database (hg19-tss-centered-10kb-7species.mc9nr.feather) used for SCENIC analysis is available at [https://resources.aertslab.org/cistarget/databases/homo\\_sapiens/hg19/refseq\\_r45/mc9nr/gene\\_based/hg19-tss-centered-10kb-7species.mc9nr.feather](https://resources.aertslab.org/cistarget/databases/homo_sapiens/hg19/refseq_r45/mc9nr/gene_based/hg19-tss-centered-10kb-7species.mc9nr.feather).

### Code availability

For reproducibility, docker images and detailed scripts used for preprocessing and further analysis are available respectively at zenodo (<https://doi.org/10.5281/zenodo.4114854>) and github (<https://github.com/Delphine-Potier/B-ALL-CAR-T>).

Received: 23 June 2020; Accepted: 14 January 2021;

Published online: 08 February 2021

### References

- Hunger, S. P. & Mullighan, C. G. Acute lymphoblastic leukemia in children. *N. Engl. J. Med.* **373**, 1541–1552 (2015).
- Pui, C.-H. Precision medicine in acute lymphoblastic leukemia. *Front. Med.* **14**, 689–700 (2020).
- Kansagra, A. J. et al. Clinical utilization of chimeric antigen receptor T cells in B cell acute lymphoblastic leukemia: an expert opinion from the european society for blood and marrow transplantation and the american society for transplantation and cellular therapy. *Biol. Blood Marrow Transplant.* **25**, e76–e85 (2019).
- Brentjens, R. J. et al. CD19-targeted T cells rapidly induce molecular remissions in adults with chemotherapy-refractory acute lymphoblastic leukemia. *Sci. Transl. Med.* **5**, 177ra38 (2013).
- Porter, D. L. et al. Chimeric antigen receptor T cells persist and induce sustained remissions in relapsed refractory chronic lymphocytic leukemia. *Sci. Transl. Med.* **7**, 303ra139 (2015).
- Grupp, S. A. et al. Chimeric antigen receptor–modified T cells for acute lymphoid leukemia. *N. Engl. J. Med.* **368**, 1509–1518 (2013).
- Sadelain, M., Rivière, I. & Riddell, S. Therapeutic T cell engineering. *Nature* **545**, 423–431 (2017).
- Maude, S. L. et al. Tisagenlecleucel in children and young adults with B-cell lymphoblastic leukemia. *N. Engl. J. Med.* **378**, 439–448 (2018).
- Gardner, R. A. et al. Intent-to-treat leukemia remission by CD19 CAR T cells of defined formulation and dose in children and young adults. *Blood* **129**, 3322–3331 (2017).
- Davila, M. L. et al. Efficacy and toxicity management of 19-28z CAR T cell therapy in B cell acute lymphoblastic leukemia. *Sci. Transl. Med.* **6**, 224ra25–224ra25 (2014).
- Lee, D. W. et al. T cells expressing CD19 chimeric antigen receptors for acute lymphoblastic leukaemia in children and young adults: a phase 1 dose-escalation trial. *Lancet* **385**, 517–528 (2015).
- Park, J. H. et al. Long-term follow-up of CD19 CAR therapy in acute lymphoblastic leukemia. *N. Engl. J. Med.* **378**, 449–459 (2018).
- Maude, S. L. et al. Chimeric antigen receptor T cells for sustained remissions in leukemia. *N. Engl. J. Med.* **371**, 1507–1517 (2014).

- Greenbaum, U., Mahadeo, K. M., Kebriyai, P., Shpall, E. J. & Saini, N. Y. Chimeric antigen receptor T-cells in B-acute lymphoblastic leukemia: state of the art and future directions. *Front. Oncol.* **10**, 1594 (2020).
- Schultz, L. Chimeric antigen receptor T cell therapy for pediatric B-ALL: narrowing the gap between early and long-term outcomes. *Front. Immunol.* **11**, 1985 (2020).
- Orlando, E. J. et al. Genetic mechanisms of target antigen loss in CAR19 therapy of acute lymphoblastic leukemia. *Nat. Med.* **24**, 1504–1506 (2018).
- Fischer, J. et al. CD19 isoforms enabling resistance to CART-19 immunotherapy are expressed in B-ALL patients at initial diagnosis. *J. Immunother.* **40**, 187–195 (2017).
- Asnani, M. et al. Retention of CD19 intron 2 contributes to CART-19 resistance in leukemias with subclonal frameshift mutations in CD19. *Leukemia* **34**, 1202–1207 (2020).
- Sortillo, E. et al. Convergence of acquired mutations and alternative splicing of CD19 enables resistance to CART-19 immunotherapy. *Cancer Discov.* **5**, 1282–1295 (2015).
- Stoekius, M. et al. Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol.* **19**, 224 (2018).
- Zheng, G. X. Y. et al. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017).
- Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
- Serin Harmanci, A., Harmanci, A. O. & Zhou, X. CaSpER identifies and visualizes CNV events by integrative analysis of single-cell or bulk RNA-sequencing data. *Nat. Commun.* **11**, 89 (2020).
- Narla, G. KLF6, a candidate tumor suppressor gene mutated in prostate cancer. *Science* **294**, 2563–2566 (2001).
- Tetreault, M.-P., Yang, Y. & Katz, J. P. Krüppel-like factors in cancer. *Nat. Rev. Cancer* **13**, 701–713 (2013).
- van Buggenum, J. A. G. L. et al. A covalent and cleavable antibody-DNA conjugation strategy for sensitive protein detection via immuno-PCR. *Sci. Rep.* **6**, 22675 (2016).
- Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e21 (2019).
- Aibar, S. et al. SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).

### Acknowledgements

High-throughput sequencing was performed at the TGML Platform, supported by grants from Inserm, GIS IBISA, Aix-Marseille Université, and ANR-10-INBS-0009-10. This study was partly supported by research funding from the Canceropôle PACA, Institut National du Cancer and Région Sud. S.P. received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 713750, and from the Regional Council of Provence-Alpes-Côte d'Azur, A\*MIDEX (No. ANR-11-IDEX0001-02). D.P.-B. received financial support from ITMO Cancer of AVIESAN (Alliance Nationale pour les Sciences de la Vie et de la Santé, National Alliance for Life Sciences & Health) within the framework of the Cancer Plan (Project No. C19046S) and from CNRS "Osez l'interdisciplinarité!" program—"DMATH" project. The authors acknowledge the "genomic instability and human hemopathies" team and the bioinformatic platform of CIML for technical support.

### Author contributions

T., R., D.P., M.L., and D.P.-B. performed the experiments, analyzed the data, and wrote the manuscript. D.P., S.P., and M.N. performed bioinformatics analysis. All authors read and approved the final manuscript.

### Competing interests

The authors declare no competing interests.

**Additional information**

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-21168-6>.

**Correspondence** and requests for materials should be addressed to M.L. or D.P.-B.

**Peer review information** *Nature Communications* thanks Jan Melenhorst and Zohar Sachs for their contribution to the peer review of this work. Peer review reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021