Resolving inter- and intra-patient heterogeneity in NPM1-mutated AML at single-cell resolution

Karakaslar, E. Onur; Argiro, Eva M.; Struckman, Nadine E.; Shirali HZ, Ramin; Severens, Jeppe F.; Honders, M. Willy; Reinders, Marcel J.T.; Griffioen, Marieke; van den Akker, Erik B.; More Authors

DOI
10.1038/s41375-025-02745-w

Publication date
2025

Document Version
Final published version

Published in
Leukemia

ARTICLE

ACUTE MYELOID LEUKEMIA

Resolving inter- and intra-patient heterogeneity in *NPM1*-mutated AML at single-cell resolution

E. Onur Karakaslar[1,2,3], Eva M. Argiro[4], Nadine E. Struckman[4], Ramin Shirali HZ[1,2,3], Jeppe F. Severens[1,2,3], M. Willy Honders[4], Susan L. Kloet[5], Hendrik Veelken[4], Marcel JT Reinders[1,2,3], Marieke Griffioen[4]✉ and Erik B. van den Akker[1,2,3]✉

NPM1-mutated AML is one of the largest entities in international classification systems of myeloid neoplasms, which are based on integrating morphologic and clinical data with genomic data. Previous research, however, indicates that bulk transcriptomics-based subtyping may improve prognostication and therapy guidance. Here, we characterized the heterogeneity in *NPM1*-mutated AML by performing single-cell RNA-sequencing and spectral flow cytometry on 16 AML belonging to three distinct subtypes previously identified by bulk transcriptomics. Using single-cell expression profiling we generated a comprehensive atlas of *NPM1*-mutated AML, collectively reconstituting complete myelopoiesis. The three *NPM1*-mutated transcriptional subtypes showed consistent differences in the proportions of myeloid cell clusters with distinct patterns in lineage commitment and maturational arrest. In all samples, leukemic cells were detected across different myeloid cell clusters, indicating that *NPM1*-mutated AML are heavily skewed but not fully arrested in myelopoiesis. Same-sample multi-color spectral flow cytometry recapitulated these skewing patterns, indicating that the three *NPM1*-mutated subtypes can be consistently identified across platforms. Moreover, our analyses highlighted differences in the abundance of rare hematopoietic stem cells suggesting that skewing occurs early in myelopoiesis. To conclude, by harnessing single-cell RNA-sequencing and spectral flow cytometry, we provide a detailed description of three distinct and reproducible patterns in lineage skewing in *NPM1*-mutated AML that may have potential relevance for prognosis and treatment of patients with *NPM1*-mutated AML.

Leukemia (2025) 39:2916–2925; https://doi.org/10.1038/s41375-025-02745-w

HIGHLIGHTS

- *NPM1*-mutated AML shows strong intra- and interpatient heterogeneity with leukemic cells skewed rather than fully arrested at different maturation stages in myelopoiesis.
- Single-cell RNA sequencing and spectral flow cytometry revealed recurrent patterns in proportions of leukemic myeloid cells with distinct patterns in lineage commitment and maturational arrest.

INTRODUCTION

Acute Myeloid Leukemia (AML) is a heterogeneous malignant disease characterized by uncontrolled proliferation of immature myeloid precursor cells in the bone marrow. The latest WHO and ICC classifications stratify adult AML into various categories that are mainly based on distinct recurrent genetic aberrations, with one of the largest entities being *NPM1*-mutated AML [1, 2]. A 4 base pair frameshift mutation in *NPM1* is sufficient to drive malignant transformation. Although *NPM1*-mutated AML generally have a relatively favorable prognosis [3], other co-mutations are also important and may affect the prognosis and treatment response of AML patients. For example, patients who acquire *FLT3-ITD* and *DNMT3A*^{R882} co-mutations in addition to the *NPM1* mutation have a significantly worse prognosis [4].

In line with others who also reported on different gene signatures, we previously identified three distinct subtypes of *NPM1*-mutated AML by bulk-transcriptomics, each showing distinct ex-vivo drug sensitivity profiles, and enrichments for different co-mutations [5–8]. Patients with the transcriptional NPM1(1) subtype had mutually exclusive co-mutations with either *IDH1/2* or *TET2*, whereas AML samples with NPM1(2) or NPM1(3) subtypes showed strong enrichments for *FLT3-ITD* or *DNMT3A*^{R882} co-mutations, respectively. These observations raise the question whether the three distinct transcriptional subtypes in *NPM1*-mutated AML are the result of different co-mutational patterns or whether the three subtypes reflect progenitor cells arrested at

[1]Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands. [2]Pattern Recognition & Bioinformatics, Delft University of Technology, Delft, The Netherlands. [3]Leiden Center for Computational Oncology, Leiden University Medical Center, Leiden, The Netherlands. [4]Department of Hematology, Leiden University Medical Center, Leiden, The Netherlands. [5]Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands. ✉email: m.griffioen@lumc.nl; e.b_van_den_akker@lumc.nl

Received: 29 January 2025 Revised: 22 July 2025 Accepted: 12 August 2025
Published online: 15 September 2025

different stages in myelopoiesis with variable susceptibility to acquire certain type of co-mutations.

Most large-scale gene expression studies, including ours, leveraged RNA sequencing data in bulk of AML samples in which healthy and malignant hematopoiesis may co-exist [9, 10]. Recent single-cell RNA sequencing revealed that leukemic cells with different maturational arrests carry different gene mutations [10, 11], and that these leukemic cell subsets in AML typically exhibit phenotypes reminiscent of healthy hematopoietic cell types, albeit often accompanied with aberrant characteristics [10, 12–14]. Consequently, to advance our understanding of *NPM1*-mutated AML, single-cell analyses are essential for investigating and comparing intra-patient heterogeneity at a resolution that can highlight transcriptional subtypes identified in large-scale bulk sequencing studies.

In this study, we investigated whether the distinct *NPM1*-mutated AML subtypes as revealed by bulk transcriptomics are characterized by subtype-specific cell clusters or whether the same cell clusters are present in different proportions. To investigate and compare cellular heterogeneity, we employed single-cell RNA sequencing and spectral flow cytometry on 16 *NPM1*-mutated AML selected from the three distinct subtypes. Using both single-cell modalities, we identified various leukemic cell clusters resembling different maturation stages during normal hematopoiesis. These leukemic cell clusters were present in different proportions in all three transcriptional subtypes. Leukemic cells resembling early hematopoietic stem and multipotent progenitor cells dominated in two *NPM1* subtypes, whereas leukemic cell subsets with more mature myeloid phenotypes were enriched in the third subtype. Our data support a model in which the three *NPM1* transcriptional subtypes originate from different progenitor cells with variable capacity to produce more differentiated myeloid offspring. Since AML cells with different maturation phenotypes have shown variable ex-vivo drug responses, our data may be relevant for the prognosis and treatment of patients and provide preliminary insight into the etiology of *NPM1*-mutated AML.

MATERIALS AND METHODS

AML samples

All sixteen peripheral blood and bone marrow samples were obtained from patients with AML with high blast levels (77–99%) at diagnosis ($n = 15$) or relapse after chemotherapy (AML08, $n = 1$). Mononuclear cells were isolated by Ficoll-Isopaque density gradient centrifugation and cryopreserved in the Leiden University Medical Center (LUMC) Biobank for Hematological Diseases.

Sample preparation for single-cell RNA-seq

Single-cell RNA-seq was performed using the 10X Genomics Chromium Next GEM Single-cell 3′ Kit v3.1 (PN-1000268) and 10X Genomics Chromium Next GEM Chip G Single-cell Kit (PN-1000120) according to the manufacturer's instructions (User Guide CG000315 Revision D). The libraries were pooled and sequenced on a NovaSeq 6000 platform (Illumina) using a 300-cycle kit S4 flow cell with v1.5 chemistry.

Downstream analysis of single-cell RNA-seq data

CellRanger v7.0.0 was run on all samples with the human reference genome hg38. For all QC *Seurat* v4 was used [15]. Our QC pipeline had three steps per sample: 1) soft filtering, 2) low quality cluster removal, and 3) doublet detection. In soft filtering, *Seurat* objects were created with cells expressing at least 200 genes and with the genes expressed at least in 3 cells. Then, standard Seurat command list with default parameters was run to detect low quality clusters. Clusters with >15% mitochondrial and <1000 mRNA in median were removed. Next, *DoubletFinder* v3 [16] was run for each sample with 7.5% detection rate and then singlets were selected. To integrate the samples, CCA mode of *Seurat* pipeline was used with most variable 2000 features. Finally, after integration we filtered cells with >15% mitochondrial mRNA.

We used standard *Seurat* commands to scale and normalize the data on integrated features. The first 30 principal components were used to create UMAP plots. We used *clustree* [17] to determine the optimal cluster number, based on *FindClusters* with resolutions sweeping from 0 to 1.2. We chose res = 0.5, as clusters became stable (Supplementary Fig. S1a, b). Next, we merged two clusters (CC5 and CC12) into one GMP-like cluster as one of these clusters (CC12) had high expression of HSP-genes yet still retained its cell-type specific properties (Supplementary Fig. S1c, d).

Spectral flow cytometry

AML patient derived bone marrow or peripheral blood mononuclear cells were thawed, purified by ficoll-paque and stained with a 23-color antibody panel on the same day as data acquisition. Cells were stained with Zombie Red viability dye (Biolegend) for 15 min at room temperature (RT), washed, incubated for 15 min in PBS supplemented with 2.5% human serum (Sanquin) and subsequently stained with fluorochrome-conjugated antibodies in Brilliant stain buffer plus (BD) for 30 min at 4 °C. The antibodies used in the spectral flow cytometry marker panel were included in Supplementary Table 1. The staining protocol to measure intracellular mutant NPM1 by flow cytometry has been described in the extended methods.

RESULTS

Single-cell atlas of *NPM1*-mutated AML reconstitutes myelopoiesis

To investigate the identity and cellular composition of the three *NPM1* transcriptional subtypes, we performed single-cell RNA sequencing on samples from 16 *NPM1*-mutated AML patients. All AML samples contained high blast levels (77–99%) and were specifically selected from each of the three major gene expression-based clusters as previously identified for *NPM1*-mutated AML by bulk RNA sequencing. In total, we selected 7, 5 (including relapse sample AML08), and 4 AML from the transcriptional NPM1(1), NPM1(2), and NPM1(3) subtypes, respectively (Fig. 1a). All patient-derived samples featured the hotspot frameshift mutation in *NPM1*, and co-mutation patterns varied across the three NPM1 subtypes as previously described [5] (Fig. 1b, Supplementary Table 1, see Methods).

After stringent quality control and doublet removal, single-cell RNA sequencing (scRNA-seq) data from 83,162 total cells were integrated and clustered into 15 cell clusters (Fig. 1c, Supplementary Figs. S1, S2a–d, see Methods). Subsequent annotation of these clusters was performed using various strategies. First, we used Azimuth [15] to project a healthy bone marrow atlas onto our data (Supplementary Fig. S2e). This allowed us to identify the most resembling cell subset in healthy hematopoiesis for each of our identified clusters. Given that particularly the myeloid lineage is affected by disease, we refer to '-like' subsets for clusters mostly resembling granulocyte-monocyte progenitors (GMPs), monocytes, dendritic cells (DCs), and erythroid cells. While clusters with mature myeloid cell populations, e.g. monocytes and erythrocytes, strongly resembled their healthy counterparts, most clusters with characteristics of progenitor cell populations could not be faithfully projected onto the healthy bone marrow atlas (Supplementary Fig. S2f). We next compared the obtained cell clusters to a previously published dataset by van Galen et al. [12] containing 16 AML with diverse genetic lesions, including 5 *NPM1*-mutated AML (Supplementary Fig. S2g). Many of our cell clusters resembled one of the 6 cell types previously described in this dataset, including hematopoietic stem cell (HSC)-like, progenitor-like, GMP-like, promonocyte-like, monocyte-like, and conventional dendritic cell (cDC)-like leukemic cells.

The combined analysis resulted in 11 myeloid cell clusters including one long-term HSC-like (LT-HSC-like) cluster expressing *CD34*, one lymphoid-primed multipotent progenitor-like (LMPP-like) cluster, one progenitor-like cluster, three GMP-like clusters in different cell cycle states (non-cycling, S, and G2), two monocyte-like clusters ($CD14^+$ and $CD16^+$ monocytes), one DC-like cluster

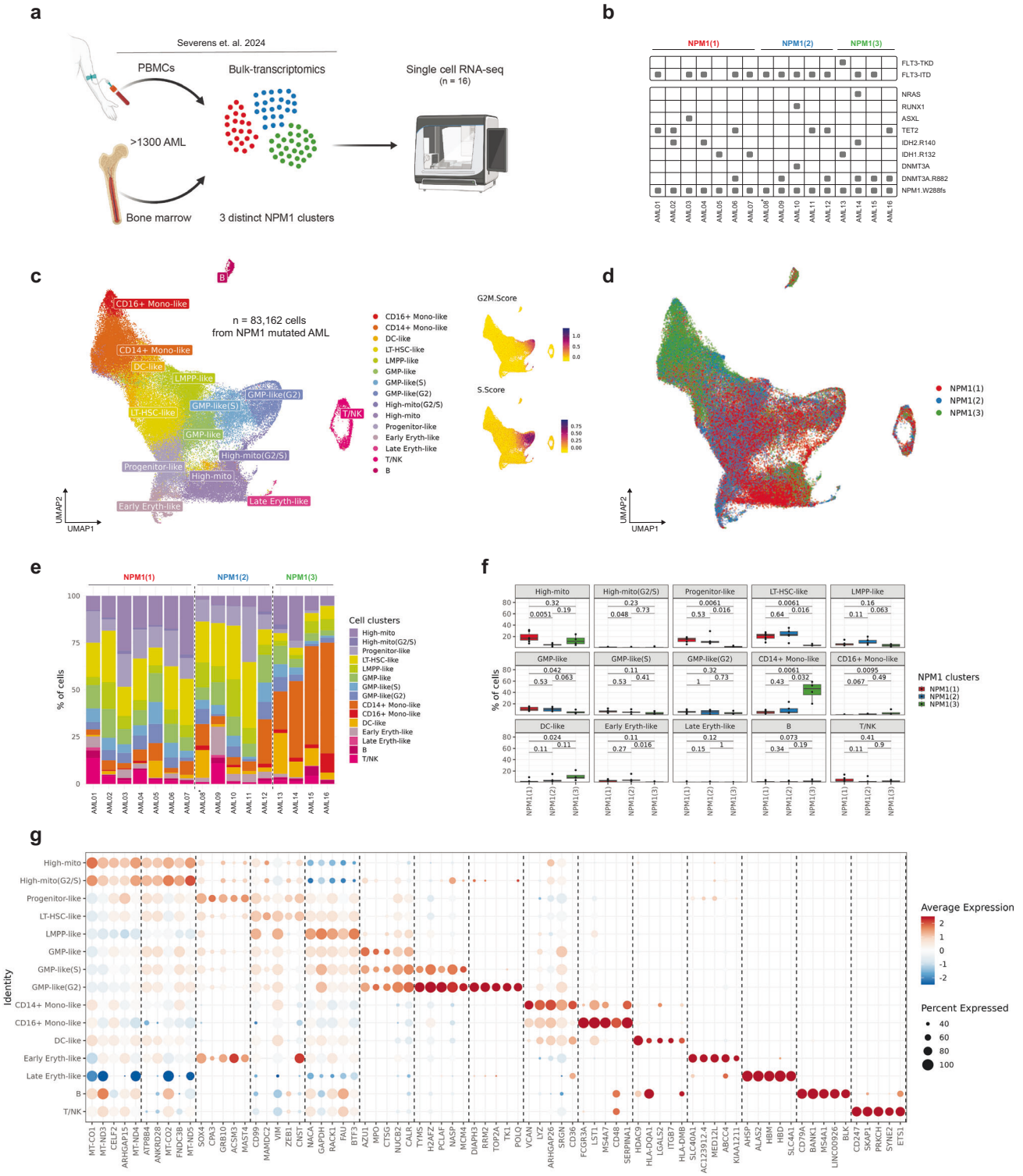

Fig. 1 AML heterogeneity between NPM1 subtypes by single-cell RNA-sequencing. a Experimental design – 16 AML cases representing the three distinct transcriptional NPM1 subtypes identified by bulk transcriptomics were selected for single cell RNA-sequencing. **b** Gene mutations detected in our cohort of 16 AML cases collected at primary diagnosis ($n = 15$) or relapse ($n = 1$; AML08*). **c** UMAP plot showing cells from 16 AML cases ($n = 83,162$) colored by annotated cell types accompanied by cell cycle information. **d** UMAP plot colored by NPM1 subtype. **e** Barplot showing cell cluster distribution per patient. Patients were grouped based on their NPM1 subtypes. Relapse sample AML08* belongs to the NPM1(2) subtype. **f** Boxplots showing percentages of cell clusters per sample grouped by NPM1 subtypes. Boxplots colored by NPM1 subtype. *p* values are calculated using the Wilcoxon rank-sum test. **g** Heatmap showing the top 5 up-regulated DEGs per cell cluster. Dot color indicates average expression per cluster and size shows the percentage of cells expressing the respective gene for that cluster.

and two erythrocyte-like clusters (early and late erythrocytes) (Supplementary Fig. S2g, h). In addition, two lymphoid cell subsets clustered separately from the myeloid cells (B-cells and NK/T cells), which is in line with previously published AML studies [5, 6].

Next, we performed additional analyses to explore the identity of two remaining cell clusters with a relatively high mitochondrial RNA content that passed our stringent quality control (Supplementary Fig. S2c). A similar cell cluster was observed in a recent single-cell transcriptomics study including 10 *NPM1*-mutated AML [13]. Since a relatively high content of mitochondrial RNAs may indicate poor data quality, we evaluated to what extent our clustering was affected by variability in mitochondrial content. After adjusting for mitochondrial content, the two high-mito clusters persisted as independent clusters (Supplementary Fig. S2i, j). Moreover, projecting the cells of each individual patient into a separate map revealed that high-mito cells were present within each AML sample (Supplementary Fig. S3). Collectively, these data suggest that the two cell clusters with a high content of mitochondrial RNAs are not the result of poor-quality data of one or a few samples but represent a consistent signal across all AML samples.

Finally, for each cell cluster, differentially expressed genes (DEGs) were identified by comparing its gene expression profile with the other cell clusters (Fig. 1g, Supplementary Table 2). We plotted the DEGs identified for high-mito cells using an independent healthy bone marrow atlas [18] and demonstrated that these cells exhibited signatures reminiscent of early progenitors (Supplementary Fig. S2k). Overall, expression of the DEGs was stronger in the more mature myeloid cell clusters than in the hematopoietic stem cell or progenitor cell clusters. Of note, the DEGs identified for the progenitor-like cell cluster were also expressed by early erythrocyte-like cells, suggesting a common lineage origin and early commitment of progenitor-like cells towards the erythrocyte lineage.

To conclude, our single-cell transcriptomics atlas revealed 13 myeloid cell clusters in *NPM1*-mutated AML representing different maturation signatures and states varying from early to late myeloid differentiation.

Transcriptional subtypes of *NPM1*-mutated AML exhibit distinct cell type compositions

To investigate whether the cell clusters were exclusively present in one of the three transcriptional *NPM1*-mutated AML subtypes, or whether these cells were present in all three subtypes in different proportions, we checked the distribution of the three *NPM1*-mutated AML subtypes over the UMAP plot (Fig. 1d). This analysis showed a clear difference in distribution between the three subtypes due to differences in abundance of specific cell clusters. This observation was further corroborated when the composition of each AML sample was quantified, with samples from the same subtype showing relatively similar cell cluster compositions, and samples from different subtypes exhibiting notable differences in cell cluster composition (Fig. 1e, Supplementary Fig. S2l). Moreover, in alignment with our previous observations, NPM1(1) and NPM1(2) samples had more abundant cell clusters with early progenitor phenotypes, whereas cell clusters resembling more mature monocytes and DCs were dominating in NPM1(3) samples (Fig. 1f). Of note, although both NPM1(1) and NPM1(2) subtypes were more abundant in cell clusters with early progenitor phenotypes, there were also differences between the two subtypes. For instance, the two high-mito clusters were more abundant in NPM1(1) samples (8–33%, Wilcoxon rank-sum test $P = 0.005$ and $P = 0.05$ for high-mito and -G2, respectively), whereas NPM1(2) samples had higher fractions of LT-HSC-like and LMPP-like cell clusters albeit not significant. Overall, the single-cell transcriptomics data revealed that all 13 cell clusters, representing different myeloid differentiation stages, were present in different proportions in the three *NPM1*-mutated AML subtypes.

NPM1(2) samples resemble multipotent hematopoietic stem cells

To further explore the difference in composition between NPM1(1) and NPM1(2) subtypes, we compared the expression of *CD34*, which has been used as marker to annotate LT-HSC-like cells. The data showed that *CD34* expression was mostly restricted to NPM1(2) samples (Fig. 2a). To validate this finding, *CD34* expression was analyzed in primary *NPM1*-mutated AML samples from BEAT-AML, an independent bulk-transcriptomics cohort [19]. This analysis confirmed that NPM1(2) samples had higher expression of *CD34* (Wilcoxon rank-sum $P = 0.0009$, Fig. 2b, Supplementary Fig. S4a). Since previous studies associated *FLT3-ITD* status with high *CD34* expression [12, 20], we also separately analyzed *NPM1*-mutated AML samples of BEAT-AML based on their *FLT3-ITD* status (Supplementary Fig. S4b). The data showed that *CD34* expression was higher in NPM1(2) samples compared to other subtypes, irrespective of *FLT3-ITD* status ($P = 0.09$).

To further characterize the differences between the *NPM1* clusters, we calculated the stemness score [21] for our AML samples per cell cluster by creating pseudo-bulks from single-cell data. While the stemness score was shown to decline along the myeloid maturation trajectory, NPM1(2) samples had consistently higher stemness scores across all cell clusters compared to NPM1(1) and NPM1(3) subtypes (Fig. 2c). Cross-checking the individual genes contributing to the stemness score revealed that *CD34* expression strongly correlated with stemness score (Supplementary Fig. S4c, d, Pearson $R = 0.62$).

Altogether, our data indicate that NPM1(2) samples have the highest stemness scores and that they seem to retain transcriptional characteristics of multipotent HSCs along differentiation towards more committed progenitor and mature myeloid cells.

Myeloid cell clusters in *NPM1*-mutated AML samples are leukemic

To investigate whether the identified cell clusters represent leukemic cells, we performed variant calling in single-cell RNA-seq data for the known hotspot mutations $NPM1^{W288Cfs}$ and $DNMT3A^{R882H}$. Mutated and wildtype *NPM1* reads were detected in all AML samples (Fig. 2d, Supplementary Fig. S4e). In the two lymphoid cell clusters (B and T/NK cell clusters), most single-cells had only wildtype *NPM1* transcripts, whereas mutant *NPM1* transcripts were detected in many single-cells in all myeloid cell clusters (Supplementary Fig. S4f). Since all samples were selected for high blast percentages (77–99%) by routine morphology and high VAFs (27–38%) for mutant *NPM1* reads by bulk RNA-Seq, we expect that all cells that share gene expression profiles with cells in which mutant *NPM1* reads have been detected are leukemic cells. However, it cannot be excluded that within the various cell clusters, healthy or preleukemic cells co-localize with leukemic cells. Furthermore, the call rates of transcripts across cell clusters were concordant with the expression of the *NPM1* gene (Supplementary Fig. S4g).

Read distributions showed that most cells had only one read covering either the mutant or wildtype NPM1 allele (Supplementary Fig. S4h). We therefore determined pseudo variant allele frequencies (pVAF) by calculating the ratio of mutant reads to the total number reads for a given number of cells. All 16 *NPM1*-mutated AML samples (Fig. 2e, $R^2 = 0.96$, $P = 5.6 \times 10^{-11}$) and all myeloid cell clusters (Fig. 2f) had similar pVAF. We also re-analyzed *NPM1*-mutated AML samples from Sergi-Beneyto et al. [10], who used targeted genotyping to increase the coverage of *NPM1*. pVAF distributions after targeted genotyping closely resembled those in our data, supporting the robustness of pVAF values (Supplementary Fig. S4m–p). Cell clusters with more mature myeloid phenotypes had slightly lower pVAF compared to early progenitor cell clusters, but these differences were not significant (Supplementary Fig. S4i, $P = 0.75$). This slight decrease in pVAF was also observed when samples were compared between the three NPM1 subtypes, with NPM1(3) samples having overall lower pVAF

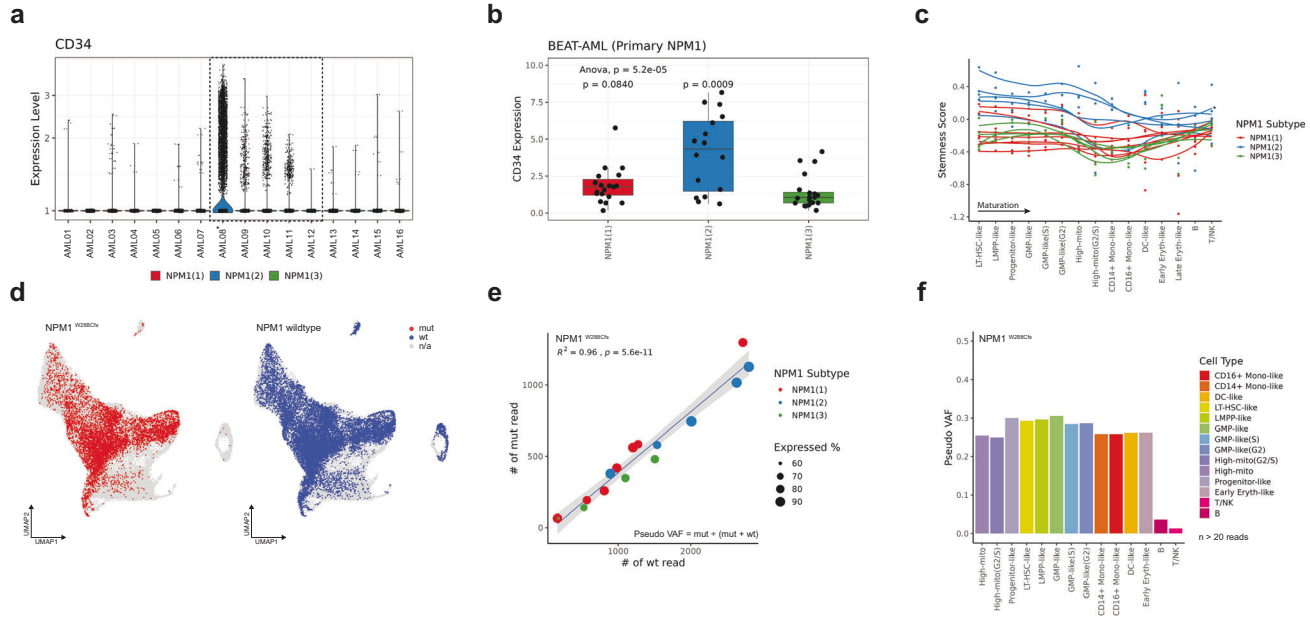

Fig. 2 AML in NPM1(2) subtype resemble hematopoietic stem cells or early progenitor cells. a Violin plot for *CD34* gene expression per sample. NPM1(2) samples, including one relapse sample (AML08*) are highlighted. **b** CD34 gene expression of NPM1-mutated AML at diagnosis of the BEAT-AML cohort. p values are calculated using the Wilcoxon rank-sum test with NPM1(3) samples as reference. **c** Stemness score of each cell cluster calculated per sample. To visualize per sample trends smoothed lines were fitted across cell types. Cell clusters are ordered based on their maturation stage. **d** UMAPs plots annotated with cells expressing mutated NPM1^{W288Cfs} (left) or wild type NPM1 allele (right). **e** Scatter plot of mutant (mut) and wildtype (wt) NPM1 allele counts per sample. Samples are colored by their respective NPM1 subtype. **f** Barplot of NPM1^{W288Cfs} pVAF for all cell types, except for late erythrocytes in which less than 20 reads were detected. All myeloid cell clusters are malignant, whereas lymphoid cell clusters mostly lack the mutation.

(Supplementary Fig. S4j). Also, NPM1(3) samples in BEAT-AML had lower VAF_{DNA} than NPM1(1) and NPM1(2) subtypes, but the difference was not statistically significant (Supplementary Fig. S4k). Despite the low number of single cells with *DNMT3A*^{R882H}, mutant reads for this hotspot mutation were detected in all cell clusters present in AML samples with this mutation (Fig. 1b, Supplementary Fig. S4l). Overall, for each sample, we detected mutant *NPM1* transcripts in all myeloid cell clusters, thereby confirming the leukemic origin of the single cells in these clusters.

Heterogeneity of *NPM1*-mutated AML captured by spectral flow cytometry

To investigate whether *NPM1*-mutated AML from the three transcriptional subtypes also displayed different cell surface markers, we performed spectral flow cytometry using an antibody panel against 19 cell surface markers (Fig. 3a, Supplementary Fig. S5a, see Methods). For this purpose, antibodies against well-established hematopoietic stem or progenitor cell and myeloid differentiation markers were used (CD33, CD45, CD34, CD38, CD117, CD45RA, CLL-1, GPR56, CD14, CD16, CD61, CD11c, CD163, CD68, CD85k-LILRB4) as well as antibodies against HLA class I and II molecules (HLA-DR, -DQ, -DP), for which low gene expression was observed particularly in the NPM1(1) subtype [5] (Supplementary Fig. S5b).

We used this spectral flow cytometry antibody panel to stain the same set of 16 *NPM1*-mutated AML cases as analyzed by single-cell transcriptomics. After removal of doublets, dead cells and lymphocytes, $CD33^+$ and/or $CD34^+$ cells were gated, and data were analyzed for a total of 433,646 cells. On average, lower cell numbers were acquired for NPM1(1) samples than NPM1(2) and NPM1(3) samples (Supplementary Fig. S6a), and NPM1(3) samples had higher number of cells showing strong expression of multiple markers (Supplementary Fig. S6b).

Similar to single-cell RNA-seq data, we projected all 433,646 single-cells from the 16 *NPM1*-mutated AML samples acquired by spectral flow cytometry onto a UMAP plot and annotated the cells based on the *NPM1* subtype as measured by bulk transcriptomics. Like single-cell RNA-seq results, the data showed that cells from the three *NPM1*-mutated AML subtypes were differently distributed over the UMAP plot (Fig. 3b), and that single-cells from the same *NPM1* subtype co-localized (Supplementary Fig. S6c). To further characterize the phenotype of the AML cells, we performed clustering with PhenoGraph [22], revealing 33 different cell clusters (Fig. 3c, see Methods). Cell cluster abundances differed between the three NPM1 subtypes (Supplementary Fig. S6c, d). Various cell clusters were significantly more abundant in NPM1(1) samples (clusters 4, 7, 11, 14), whereas other cell clusters were mainly present in NPM1(3) samples (clusters 2, 10, 12, 17, 21, 22, 23) (Supplementary Fig. S6e), illustrating that the AML sample composition differed between the three NPM1 subtypes.

NPM1 subtypes display different cell surface markers

Based on what is known about marker expression during normal hematopoiesis from hematopoietic stem cells to more mature myeloid cell populations [23–25], we grouped the 33 flow cytometry clusters into 5 major AML phenotypes, i.e. Hematopoietic Stem Cell-like (HSC-like) ($CD34^+CD38^-Lin^-$), Early Progenitor-like ($CD34^-CD117^+CD45RA^+Lin^-$, $CD34^-CD117^+CD45RA^-Lin^-$, $CD34^-CD117^-CD45RA^+Lin^-$), Late Progenitor-like ($CD34^-CD117^-CD45RA^-Lin^-$), Mono-like ($CD14^+CD11c^+$, $CD16^+CD11c^+$, $CD68^+CD11c^+$), and DC-like ($HLA\text{-}DQ^+CD11c^+$) AML cells. The Mono- and DC-like clusters were defined by expression of at least one lineage (Lin) marker, including CD14, CD16, CD68 and/or CD11c, whereas all progenitor cell clusters were negative for these lineage markers (Fig. 3d).

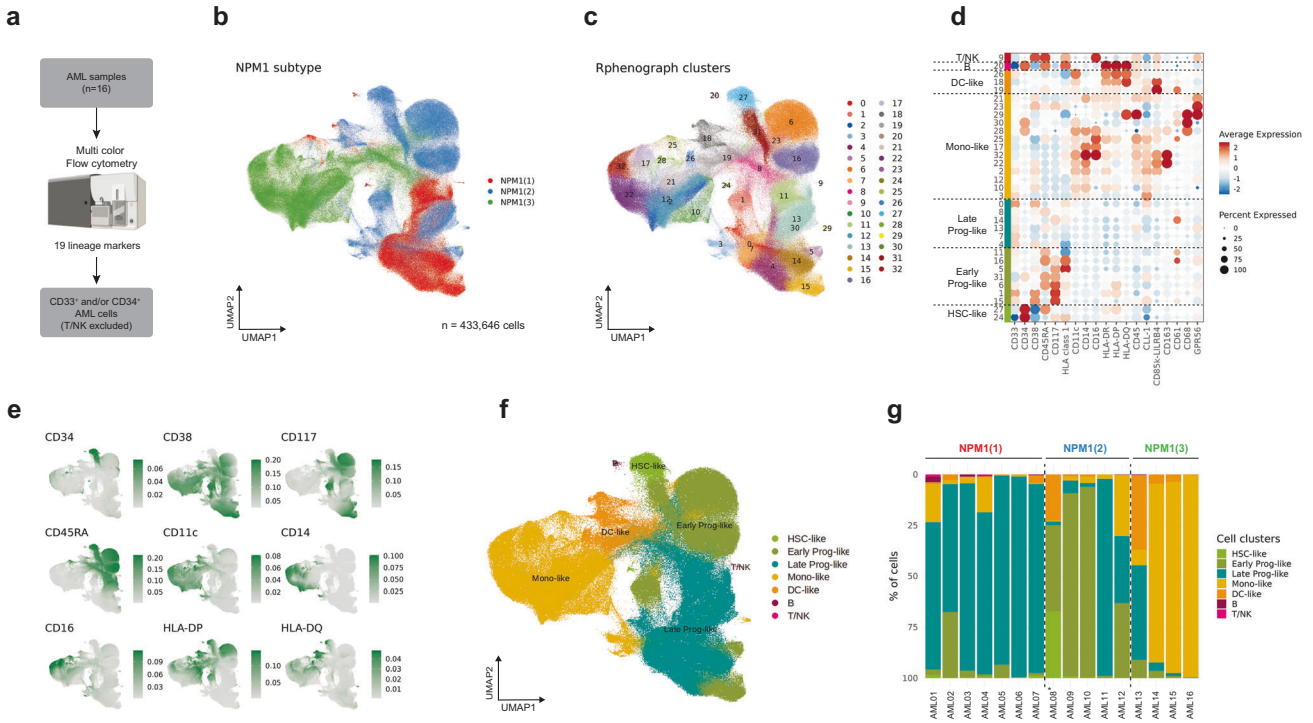

Fig. 3 AML heterogeneity between NPM1 subtypes by spectral flow cytometry. a Chart illustrating the strategy for spectral flow cytometry. The same 16 AML samples analyzed by single-cell RNA-sequencing were also measured by spectral flow cytometry. **b** UMAP plot of AML cells ($n = 433,646$) measured by spectral flow cytometry data colored by NPM1 subtypes. **c** UMAP plot for 33 *PhenoGraph* clusters ($k = 60$). **d** Dot plot showing all 19 markers. Each marker is scaled across clusters. Colored bars on the left indicate cell type annotations. **e** Expression of 9 cell surface markers (CD34, CD38, CD117, CD45RA, CD11c, CD14, CD16, HLA-DP and HLA-DQ). **f** UMAP plot colored with cell type annotations. **g** Barplot showing cell type distribution per patient. Patients are grouped based on their NPM1 subtypes. The relapse sample belongs to the NPM1(2) subtype (AML08*).

To corroborate these cluster annotations and as a data quality check, we selected $CD33^+$ or $CD34^+$ cells from our single-cell transcriptomics data and associated the flow cytometry markers with their corresponding gene expression (Supplementary Fig. S6f, see Methods). All markers used to annotate the 5 major AML phenotypes showed significant correlations between gene and protein expression (Fig. 3e, Supplementary Fig. S6f). Having ensured that marker expression correlated with their corresponding genes, we next correlated the markers of the 33 flow cytometry clusters with the clusters found in scRNA-seq data (Supplementary Fig. S6g). This analysis yielded similar patterns with the same 5 major AML cell subsets in Fig. 3d. Of note, high-mito clusters from single-cell RNA-seq data correlated with late progenitor-like cells ($CD33^+CD34^-CD117^-CD45RA^-Lin^-$). Next, we projected the 5 major cell subsets onto the UMAP plot of the flow cytometry data (Fig. 3f) and demonstrated their distribution within each of the 16 AML samples (Fig. 3g).

The flow cytometry data supported single-cell RNA-seq results, demonstrating that AML cells with early progenitor-like phenotype dominated in NPM1(2) samples, whereas NPM1(1) samples mainly contained AML cells with late progenitor-like phenotypes, and AML cells with more mature monocyte- or DC-like phenotypes were abundantly present in NPM1(3) samples. The only discrepant sample was AML11, which has been identified as NPM1(2) sample by bulk transcriptomics. Further inspection of this sample confirmed lower *KIT* (CD117) gene expression in this sample as compared to other NPM1(2) samples (Supplementary Fig. S6h), which may explain its resemblance to NPM1(1) by spectral flow cytometry. Overall, the spectral flow cytometry data confirmed that there is intra-patient heterogeneity in *NPM1*-mutated AML samples and that inter-patient heterogeneity largely overlapped with our identified transcriptional *NPM1* subtypes.

HSC in *NPM1*-mutated AML can be leukemic, but also healthy or preleukemic

Two cell clusters in the flow cytometry data resemble HSCs characterized by $CD34^+CD38^-CD45RA^-$ (cluster 24) or $CD34^+CD38^-CD45RA^+$ (cluster 27) expression (Fig. 4a). Cluster 24 was mainly present in NPM1(1) samples (Wilcoxon rank-sum $P = 0.048$ and $P = 0.024$), whereas cluster 27 was more dominant in NPM1(2) samples, reaching the significance threshold when compared to NPM1(1) (Fig. 4b, Wilcoxon rank-sum $P = 0.05$). Since we observed a relatively high contribution of the relapse sample to cluster 27, we repeated the analyses after excluding AML08, which confirmed that these leukemic cells can be readily identified in AML samples at diagnosis (Supplementary Fig. S6i–l).

To explore whether these cell clusters represent leukemic or non-leukemic cells, we made use of Cellular Indexing of Transcriptomes and Epitopes by Sequencing (CITE-seq) data from Sergi-Beneyto et al. [10]. We first selected the 7 *NPM1*-mutated AML samples from their cohort and re-integrated these samples. Next, we projected the cells on a UMAP plot colored by author's original cell type annotations, which resulted in various clusters including lymphoid (B-, T- and NK-cells) and AML cells, which clearly clustered separately (Fig. 4c). Besides lymphoid cell clusters, there was another healthy or preleukemic cell cluster identified (cluster 16), which consists of cells mostly lacking the *NPM1* mutation (Fig. 4d, Supplementary Fig. S7a). This cluster had high CD34 protein expression (Fig. 4e, Supplementary Fig. S7b), and was present in all 7 *NPM1*-mutated AML samples (Supplementary Fig. S7c). Since this cell cluster had the same phenotype as cell cluster 24 in our flow cytometry data ($CD34^+CD38^-CD45RA^-$), our analyses suggest that cluster 24 represents healthy or preleukemic HSCs that are mainly present in NPM1(1) samples.

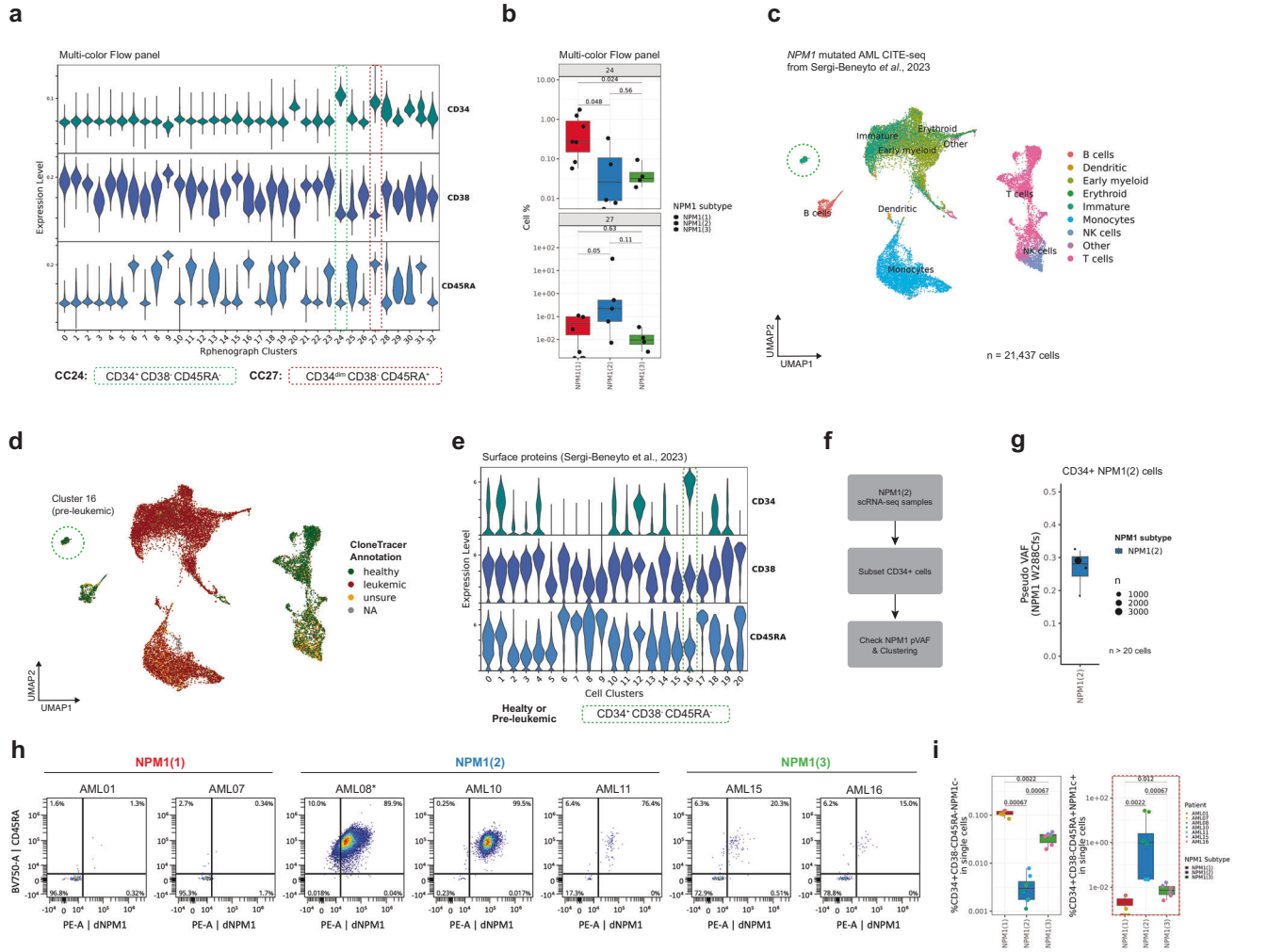

Fig. 4 AML in NPM1(1) subtype contain healthy or preleukemic HSCs. a Cell surface expression of CD34, CD38 and CD45RA on all cell clusters as identified by spectral flow cytometry. Two $CD34^{+}CD38^{-}$ cell clusters resembling HSC or early progenitor cells (CC24 and CC27) are highlighted. **b** Percentages of cells within each sample for CC24 and CC27 grouped by NPM1 subtypes. *p*-values were calculated using the Wilcoxon rank-sum test. **c** UMAP plot of integrated cells from 7 *NPM1*-mutated AML at diagnosis of Sergi-Beneyto et al. [1] analyzed by CITE-seq. Cells colored by reported cell type annotations. **d** Same UMAP colored with CloneTracer annotations indicating malignant cells. Cluster 16 is highlighted as healthy or preleukemic cell subset. **e** Surface expression of CD34, CD38 and CD45RA as measured by CITE-Seq for all cell clusters. **f** Strategy for single-cell RNA-seq samples to show pVAF in $CD34^{+}$ cells of NPM1(2) subtype. **g** $NPM1^{W288Cfs}$ pVAF for $CD34^{+}$ NPM1(2) cells. **h** Intracellular mutant NPM1 (x-axis) and surface CD45RA (y-axis) expression for $CD34^{+}CD38^{-}$ gated cells in 7 AML including relapse sample AML08 by flow cytometry. NPM1(1) samples have been selected specifically for clear presence of cluster 24. **i** Percentage of leukemic ($CD34^{+}CD38^{-}CD45RA^{+}NPM1c^{+}$) and healthy or preleukemic ($CD34^{+}CD38^{-}CD45RA^{-}NPM1c^{-}$) cells in total viable cells. Note that healthy or preleukemic HSCs are present at small frequencies in all AML samples, whereas leukemic HSC-like AML cells are detectable in NPM1(2) and NPM1(3) samples, but not or barely in analyzed NPM1(1) samples.

To show that early progenies indeed carry the *NPM1* mutation in NPM1(2) samples, we selected $CD34^{+}$ cells within this subtype ($n = 3668$) and demonstrated that the majority of these cells exhibit LT-HSC-like, LMPP-like or Progenitor-like phenotypes (Fig. 4f, Supplementary Fig. S7d), and that the mutant *NPM1* allele was detected across all different clusters and patients, thereby confirming that cell cluster 27 in the flow cytometry data ($CD34^{+}CD38^{-}CD45RA^{+}$) represents leukemic cells (Fig. 4g, Supplementary Fig. S7e). Lastly, we additionally stained two AML samples from each of the three transcriptional NPM1 subtypes as well as relapse sample with a polyclonal antibody against human mutant NPM1 (Supplementary Fig. S7f–g, Supplementary Table 3). The data confirmed that cells with a $CD34^{+}CD38^{-}CD45RA^{+}$ phenotype are leukemic, as they stained positive for the intracellular mutant NPM1 protein, whereas cells with a $CD34^{+}CD38^{-}CD45RA^{-}$ phenotype were negative for intracellular mutant NPM1, supporting that these cells are healthy or preleukemic HSCs (Fig. 4h, i). The data also showed that healthy or preleukemic HSC populations were detectable at low frequencies in all AML samples, whereas leukemic HSC-like AML cells were present in NPM1(2) and NPM1(3) samples, but absent or barely detectable in analyzed NPM1(1) samples.

DISCUSSION

NPM1-mutated AML accounts for ~30% of adult AML patients and it is a single entity in recent WHO and ICC 2022 classification [1, 2]. Nevertheless, previous studies demonstrated that this genetic subclass harbors significant phenotypic heterogeneity with potential prognostic and therapeutic significance [5, 6, 9]. Here, we followed up our recent work in which analyzed >1200 bulk transcriptomics data of AML cases and identified three major

transcriptional subtypes of *NPM1*-mutated AML. In the current manuscript, we leveraged same-sample AML single-cell RNA sequencing and spectral flow cytometry data to show that these transcriptional subtypes can be traced back to differences in proportions of cell subsets, ranging from early hematopoietic stem or progenitor cells to more committed progenitors and mature myeloid cell types. Overall, the results suggest that in *NPM1*-mutated AML, myelopoiesis continues with leukemic cells skewed towards different myeloid cell types.

Based on our single-cell analyses, we hypothesize that AML belonging to the NPM1(2) subtype originate from early progenitor cells producing myeloid offspring. Flow cytometry data showed that CD45RA and CD117, which are early progenitor markers in normal hematopoiesis [23], are often present in leukemic cell populations in NPM1(2) subtype. scRNA-seq data also demonstrated that early progenitor-like AML cells are most frequent in NPM1(2) subtype, and that all leukemic cell populations in this subtype have higher stem cell scores than leukemic cell populations in AML of the NPM1(1) and NPM1(3) subtypes. Since monocyte- and DC-like AML cells have also been detected in the NPM1(2) subtype, our data suggest that at least a proportion of the early progenitor-like AML cells retain the capacity to differentiate into more mature myeloid cells.

In contrast to AML belonging to NPM1(2), CD34, CD45RA and CD117 were absent or low on all or the majority of leukemic cell populations in NPM1(1) and NPM1(3) samples. Flow cytometry and scRNA-seq data showed that AML cells with more mature phenotypes, expressing high levels of maturation markers (CD11c, CD14, CD16, HLA-DR/DQ/DP, LILRB4, CD163, CD68), are most abundant in samples of the NPM1(3) subtype. This suggests that these AML may have developed from progenitor cells that are more primed to the monocytic lineage, such as GMP cells, monoblasts or promonocytes [23], and that these progenitor cells, upon acquiring the *NPM1* mutation, retain the ability to differentiate into mature myeloid cells. An alternative possibility is that NPM1(3) samples originate from the same early progenitor-like cells as NPM1(2) samples, but that the progenitor cells in NPM1(3) samples are less arrested and therefore produce more mature myeloid offspring. This possibility may be supported by our scRNA-seq data showing that besides mature monocyte-, macrophage- and DC-like AML cells, leukemic cells with earlier progenitor-like phenotypes are also present in samples of the NPM1(3) subtype, albeit in smaller proportions than in NPM1(2) samples. Due to the abundant presence of mature AML cells, NPM1(3) cases could be more resistant to treatment with the BCL2-inhibitor Venetoclax, as a monocytic phenotype in AML has been associated with higher resistance to Venetoclax and lower overall survival rates [9, 26–29].

Flow cytometry data also shows absence of CD34, CD45RA and CD117 on all or the majority of cell populations in AML of the NPM1(1) subtype, suggesting that these AML may have originated from cells that are more committed than the early progenitor -like AML cells in the NPM1(2) subtype. Our scRNA-seq data shows that in particular cells with a high mitochondrial RNA content are abundantly present in the NPM1(1) subtype. A similar subset was recently reported by Naldini et al. for *NPM1*-mutated AML samples [13]. Our data shows evidence that AML cells with a high content of mitochondrial RNAs resemble progenitor cells that are more committed, such as CMP, GMP or myeloblasts, yet lack the characteristic markers to define the cell type of origin. Interestingly, the functionally complementary IDH1/2 and TET2 co-mutations, that are enriched in NPM1(1) subtype, have previously been linked to a disturbed alpha-ketoglutarate metabolism which leads to genome-wide hypermethylation [30]. This might explain upregulation of Fat mass and obesity-associated protein (FTO) gene expression as previously reported for the NPM1(1) subtype [5], encoding a mRNA demethylase potentially counteracting this global hypermethylation. This suggests that cells with a high content of mitochondrial RNAs, which are notably abundant in NPM1(1) samples, may exhibit reduced expression of various cell surface markers as a result of extensive hypermethylation.

As described above, AML cells with a high content of mitochondrial RNA may represent late progenitor cells such as CMP- or GMP-like cells or myeloblast-like cells that are more committed to the granulocytic lineage [23]. The hypothesis that AML belonging to the NPM1(1) subtype may have originated from more differentiated progenitor cells than AML of the NPM1(2) subtype is supported by flow cytometry data showing a small healthy or preleukemic cell cluster similar in phenotype to healthy HSC ($CD34^+CD38^-CD45RA^-$) in AML of all three NPM1 subtypes, whereas another small cell cluster resembling leukemic stem cells ($CD34^+CD38^-CD45RA^+$) is present in AML of the NPM1(2) and NPM1(3) subtypes, but barely detectable in NPM1(1) samples. Kersten et al. [31] demonstrated that $CD34^+CD38^-CD45RA^+$ cells are leukemic stem cells (LSCs) that are present in CD34-positive AML, but absent in CD34-negative AML. We demonstrated that healthy (or preleukemic) HSCs as well as LSCs are often present in *NPM1*-mutated AML, which are typically negative for CD34, but that frequencies are extremely low and that LSCs can be absent or barely detectable in NPM1(1) samples. However, our cohort is small, and our observations need to be confirmed by deep analysis of a larger panel of *NPM1*-mutated AML.

The current study also has some limitations. First, our panel of AML includes both bone marrow and peripheral blood samples. Although bone marrow and peripheral blood specimens are largely similar in cellular compositions [32], it cannot be excluded that small differences between the three NPM1 subtypes are attributed to sample source. Secondly, in our previous analysis with bulk-transcriptomics, the size of the compendium allowed us to subdivide NPM1(2) and NPM1(3) each in two further subtypes [5]. For instance, NPM1(3) could be distinguished into two clusters dominated by $CD14^+$ or $CD16^+$ monocyte-like AML cells, indicating that there is still underlying heterogeneity, and that future single-cell studies are warranted to deduce the mechanisms underlying these additional NPM1 subtypes. Thirdly, for variant calling, we recommend interpreting results from read-level pseudo VAF (pVAF) values on cell clusters, since variant calling on single-cells relies on only a few reads and may therefore be heavily biased and lead to wrong interpretations. Yet, pVAF values in our experiments were not affected by variable *NPM1* gene expression between different cell types. Furthermore, although we convincingly demonstrated that AML cells with a high content of mitochondrial RNAs have similar *NPM1* pVAF levels as other leukemic cell clusters, thereby confirming their leukemic origin, our antibody panel used for spectral flow cytometry lacked any strong positive lineage markers for these cells. Our single-cell transcriptomics data provide an opportunity to refine our flow cytometry panel, and since the high mito clusters lack most cell surface markers, other flow cytometry protocols may be considered requiring cell fixation and permeabilization to allow inclusion of antibodies against RUNX1, FOXP1, PBX3, and MEIS1, which have been identified as transcription factors that are highly expressed. (see Supplementary Table 2).

In conclusion, our data showed that AML in the three transcriptional NPM1 subtypes previously identified by bulk transcriptomics contained different proportions of the same leukemic cell types resembling early progenitor or more mature myeloid cells. Our data favor a model in which AML of the three NPM1 subtypes originate from different progenitor cells with variable capacity to produce more differentiated offspring. This model suggests that *NPM1*-mutated AML belonging to different transcriptional subtypes may differ in susceptibility to standard or new anti-cancer therapies, and our model may thus have potential relevance for prognosis and treatment of patients with *NPM1*-mutated AML.

DATA AVAILABILITY

All single-cell RNA-sequencing data was submitted to EGA with the accession number EGAS50000000332, and they are accessible upon request (due to privacy rules). Processed version of the scRNA-seq data for all 16 *NPM1*-mutated AML samples were uploaded to Figshare under the https://doi.org/10.6084/m9.figshare.26189771.

CODE AVAILABILITY

All source codes are available at https://github.com/eonurk/scNPM1. Our AML variant calling pipeline for single-cell RNA-seq data is also open-sourced and available under the same repository.

REFERENCES

1. Khoury JD, Solary E, Abla O, Akkari Y, Alaggio R, Apperley JF, et al. The 5th edition of the World Health Organization classification of haematolymphoid tumours: myeloid and histiocytic/dendritic neoplasms. Leukemia. 2022;36:1703–19.
2. Arber DA, Orazi A, Hasserjian RP, Borowitz MJ, Calvo KR, Kvasnicka H-M, et al. International Consensus Classification of Myeloid Neoplasms and Acute Leukemias: integrating morphologic, clinical, and genomic data. Blood. 2022;140:1200–28.
3. Falini B, Brunetti L, Sportoletti P, Martelli MP. NPM1-mutated acute myeloid leukemia: from bench to bedside. Blood. 2020;136:1707–21.
4. Papaemmanuil E, Gerstung M, Bullinger L, Gaidzik VI, Paschka P, Roberts ND, et al. Genomic classification and prognosis in acute myeloid leukemia. N Engl J Med. 2016;374:2209–21.
5. Severens JF, Karakaslar EO, Van Der Reijden BA, Sánchez-López E, Van Den Berg RR, Halkes CJM et al. Mapping AML heterogeneity - multi-cohort transcriptomic analysis identifies novel clusters and divergent ex-vivo drug responses. Leukemia. 2024. https://doi.org/10.1038/s41375-024-02137-6.
6. Mer AS, Heath EM, Madani Tonekaboni SA, Dogan-Artun N, Nair SK, Murison A, et al. Biological and therapeutic implications of a unique subtype of NPM1 mutated AML. Nat Commun. 2021;12:1054.
7. Cheng W-Y, Li J-F, Zhu Y-M, Lin X-J, Wen L-J, Zhang F, et al. Transcriptome-based molecular subtypes and differentiation hierarchies improve the classification framework of acute myeloid leukemia. Proc Natl Acad Sci USA. 2022;119:e2211429119.
8. Giacopelli B, Wang M, Cleary A, Wu Y-Z, Schultz AR, Schmutz M, et al. DNA methylation epitypes highlight underlying developmental and disease pathways in acute myeloid leukemia. Genome Res. 2021;31:747–61.
9. Karakaslar EO, Severens JF, Sánchez-López E, van Veelen PA, Zlei M, van Dongen JJM, et al. A transcriptomic based deconvolution framework for assessing differentiation stages and drug responses of AML. npj Precis Oncol. 2024;8:105.
10. Beneyto-Calabuig S, Merbach AK, Kniffka J-A, Antes M, Szu-Tu C, Rohde C, et al. Clonally resolved single-cell multi-omics identifies routes of cellular differentiation in acute myeloid leukemia. Cell Stem Cell. 2023;30:706–721.e8.
11. Nam AS, Dusaj N, Izzo F, Murali R, Myers RM, Mouhieddine TH, et al. Single-cell multi-omics of human clonal hematopoiesis reveals that DNMT3A R882 mutations perturb early progenitor states through selective hypomethylation. Nat Genet. 2022;54:1514–26.
12. van Galen P, Hovestadt V, Wadsworth II MH, Hughes TK, Griffin GK, Battaglia S, et al. Single-cell RNA-Seq reveals AML hierarchies relevant to disease progression and immunity. Cell. 2019;176:1265–1281.e24.
13. Naldini MM, Casirati G, Barcella M, Rancoita PMV, Cosentino A, Caserta C, et al. Longitudinal single-cell profiling of chemotherapy response in acute myeloid leukemia. Nat Commun. 2023;14:1285.
14. Triana S, Vonficht D, Jopp-Saile L, Raffel S, Lutz R, Leonce D, et al. Single-cell proteo-genomic reference maps of the hematopoietic system enable the purification and massive profiling of precisely defined cell states. Nat Immunol. 2021;22:1577–89.
15. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184:3573–3587.e29.
16. McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. Cell Syst. 2019;8:329–337.e4.
17. Zappia L, Oshlack A. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. GigaScience. 2018;7:https://doi.org/10.1093/gigascience/giy083.
18. Suo C, Dann E, Goh I, Jardine L, Kleshchevnikov V, Park J-E, et al. Mapping the developing human immune system across organs. Science. 2022;376:eabo0510.
19. Tyner JW, Tognon CE, Bottomly D, Wilmot B, Kurtz SE, Savage SL, et al. Functional genomic landscape of acute myeloid leukaemia. Nature. 2018;562:526–31.
20. Pianigiani G, Rocchio F, Peruzzi S, Andresen V, Bigerna B, Sorcini D, et al. The absent/low expression of CD34 in NPM1-mutated AML is not related to cytoplasmic dislocation of NPM1 mutant protein. Leukemia. 2022;36:1931–4.
21. Ng SWK, Mitchell A, Kennedy JA, Chen WC, McLeod J, Ibrahimova N, et al. A 17-gene stemness score for rapid determination of risk in acute leukaemia. Nature. 2016;540:433–7.
22. Levine JH, Simonds EF, Bendall SC, Davis KL, Amir ED, Tadmor MD, et al. Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. Cell. 2015;162:184–97.
23. Lewis JE, Hergott CB. The immunophenotypic profile of healthy human bone marrow. Clin Lab Med. 2023;43:323–32.
24. Foster BM, Zaidi D, Young TR, Mobley ME, Kerr BA. CD117/c-kit in cancer stem cell-mediated progression and therapeutic resistance. Biomedicines. 2018;6:31.
25. Córbi AL, Lopéz-Rodríguez C. CD11c integrin gene promoter activity during myeloid differentiation. Leuk Lymphoma. 1997;25:415–25.
26. Zeng AGX, Bansal S, Jin L, Mitchell A, Chen WC, Abbas HA, et al. A cellular hierarchy framework for understanding heterogeneity and predicting drug response in acute myeloid leukemia. Nat Med. 2022;28:1212–23.
27. Kuusanmäki H, Kytölä S, Vänttinen I, Ruokoranta T, Ranta A, Huuhtanen J, et al. Ex vivo venetoclax sensitivity testing predicts treatment response in acute myeloid leukemia. Haematologica. 2023;108:1768–81.
28. Pei S, Pollyea DA, Gustafson A, Stevens BM, Minhajuddin M, Fu R, et al. Monocytic subclones confer resistance to venetoclax-based therapy in patients with acute myeloid leukemia. Cancer Discov. 2020;10:536–51.
29. White BS, Khan SA, Mason MJ, Ammad-ud-din M, Potdar S, Malani D, et al. Bayesian multi-source regression and monocyte-associated gene expression predict BCL-2 inhibitor resistance in acute myeloid leukemia. npj Precis Onc. 2021;5:71.
30. Figueroa ME, Abdel-Wahab O, Lu C, Ward PS, Patel J, Shih A, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer Cell. 2010;18:553–67.
31. Kersten B, Valkering M, Wouters R, Van Amerongen R, Hanekamp D, Kwidama Z, et al. CD 45 RA, a specific marker for leukaemia stem cell sub-populations in acute myeloid leukaemia. Br J Haematol. 2016;173:219–35.
32. Caliskan G, Pawitan Y, Vu TN. Similarities and differences of bone marrow and peripheral blood samples from acute myeloid leukemia patients in terms of cellular heterogeneity and ex-vivo drug sensitivity. eJHaem. 2024;5:721–7.

ACKNOWLEDGEMENTS

Authors would like to thank Peter van Balen for his help on data acquisition and transfer, and Leon Mei and Davy Cats for their help on data upload. This study received financial support from a strategic investment from the Leiden University Medical Center, integrated within the Leiden Oncology Center and conducted within the Leiden Center for Computational Oncology, and partly by the Dutch Cancer Society (project number 15152). EBA was supported by a personal grant from the Dutch Research Council (NWO; VENI: 09150161810095). The funding entities played no part in determining the study's design, data collection, analysis, interpretation, manuscript composition, or the decision to submit it for publication. *BioRender* icons were used in graphical abstract, Figs. 1a and 3a.

AUTHOR CONTRIBUTIONS

EBA, MG, MJTR designed the project; EBA acquired funding; EMA, NES, MG, SK and HV helped with data generation and transfer; EOK, EMA performed computational and statistical analyses; EOK created all figures; EOK, RSHZ implemented variant calling pipeline; EOK, EMA, RSHZ, JFS, EBA, MG, and MJTR supported data exploration and interpretation; MJTR, MG and EBA provided supervision and scientific direction; EOK, EBA and MG wrote the manuscript; and all authors critically reviewed the manuscript and figures.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

All methods were performed in accordance with the relevant guidelines and regulations, including the Declaration of Helsinki. Approval to use *NPM1*-mutated AML patient samples for this research was obtained from the Leiden University Medical Center (LUMC) Institutional Review Board (protocol no. RP24.056). Written informed consent to participate was obtained from all participants prior to sample collection.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41375-025-02745-w.

Correspondence and requests for materials should be addressed to Marieke Griffioen or Erik B. van den Akker.