Supplementary files

**CombiFlow: Combinatorial AML-specific plasma membrane expression profiles allow longitudinal tracking of clones**

Roos Houtsma¹#, Nisha K. van der Meer¹#, Kees Meijer², Linde M. Morsink¹, Shanna M. Hogeling¹, Carolien M. Woolthuis¹, Emanuele Ammatuna¹, Marije T Nijk², Bauke de Boer¹,³, Gerwin Huls¹, André B. Mulder² and Jan Jacob Schuringa¹*

¹Department of Hematology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

²Department of Laboratory Medicine, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands.

³ Present address: The Finsen Laboratory, Biotech Research and Innovation Centre (BRIC), University of Copenhagen, Copenhagen, Denmark

# contributed equally

*corresponding author: j.j.schuringa@umcg.nl
**Supplementary Methods**

**Patient cohort**

The patient cohort consisted of 288 patients diagnosed with *de novo* acute myeloid leukemia (AML), secondary AML (sAML), therapy-related AML (tAML) or myelodysplastic syndrome (MDS) in the University Medical Center Groningen (UMCG) between 2014 and 2019 according to the 2016 WHO criteria (Supplementary table 1 and 2). 139 of the cohort of 288 patients were also included in our earlier study published by de Boer et al.[1] Patient samples were studied after informed consent and protocol approval by the Medical Ethical Committee of the UMCG in accordance with the Declaration of Helsinki. Patient characteristics at diagnosis such as age, white blood cell counts, karyotype and molecular status were obtained from the electronic patient records. For 72 patients from the AML cohort, longitudinal PM marker expression data were available at diagnosis and for the MRD time point. Furthermore, information was collected on morphologic remission status, immunophenotypic measurable residual disease (MRD) status, type of treatment received and additional allogeneic or autologous stem cell transplantation. Survival and relapse status were also available for the analysis.

**Flow cytometry analysis**

Seven plasma membrane (PM) markers were implemented in routine diagnostics since 2014 using multi-parameter flow cytometry on a FACSCanto II™ flow cytometer (BD Bioscience). The immunophenotype at diagnosis, during follow-up, and at time of relapse was acquired according to the Euroflow protocol. Samples were stained with a shared backbone consisting of CD34 (PerCP-Cy5-5), CD45 (HV500), CD19 (APC-H7),
CD117 (PE-Cy7), CD38 (APC) and HLA-DR (HV450). The seven PM markers were all stained with PE-labeled antibody. A list of antibodies used in this study is provided as supplemental table 5. Normal bone marrow (NBM) samples were included as healthy control (n=11). The standardized routine diagnostics protocol measures 100,000 CD45\(^+\) events per tube. In absence of a CD45 staining, 1.0x10\(^6\) events are measured. Raw flow data were analyzed in FlowJo V10\(^\oplus\) (Tree Star, Inc). The geometric mean, here referred to as the mean fluorescent intensity (MFI), of each marker was obtained within the CD34\(^+\) or CD34\(^-/CD117\(^+\) cell populations. Patients were considered to have increased expression (upregulation) for a marker if the MFI within the CD34\(^+\) cells was >1.5-fold elevated compared to the average MFI within NBM CD34\(^+\) cells. For NPM1 mutant (NPM1\(_{cyt}\)) patients, due to their low CD34 expression, the MFI was determined within the CD34\(^-/CD117\(^+\) cells. We excluded gates containing less than 100 cells since such a low amount of cells will be too susceptible to outliers. The gating strategy is outlined in Supplementary Figure 1A, and additional 10 examples are shown in Supplementary Figure 2.

Analysis of 36 markers with flow cytometry was also performed in a research setting (see Supplementary table 4 for overview of included patients and Supplementary table 5 for an overview of the markers used). Cryopreserved mononuclear cell (MNC) fractions of patients were thawed, resuspended in new born calf serum (NCS) supplemented with DNase I (20 Units/mL), 4 \(\mu\)M MgSO4 and heparin (5 Units/mL), and incubated at 37 °C for 15 min. MNCs were spun down and resuspended in PBS, blocked with human FcR blocking reagent (Miltenyi Biotech) before being incubated with the backbone markers CD34 (APC), CD38 (FITC), CD45 (PE-Cy7), and CD45RA
(BV421) for 20 min at 4°C. Stained MNCs were divided over multiple tubes, each of which was stained for 30 min at 4°C with a PE-labelled PM marker panel antibody. Stained MNCs were washed once with PBS, centrifuged for 5 min at 450xg, and resuspended in 100µL PBS. Fluorescence was measured on the MACSQuant Analyzer 10 (Miltenyi Biotech). NBM (n=7) and healthy mobilized peripheral blood mononuclear cells (n=2) served as controls.

**Infinicyt analysis**

Expression data of the markers measured in different tubes were merged based on their expression of shared backbone markers using the Infinicyt™ software (Cytognos S.L.) as previously described [1]. Briefly, after thawing cells are stained with a backbone stain composed of CD34-APC, CD38-FITC, CD45-PECy7 and CD45RA-BV421. Following the backbone stain, cells are divided over multiple tubes, after which one PE-labeled marker antibody is added per tube. Expression data is acquired per tube and raw fcs files are combined using the Infinicyt™ software. The backbone stain, combined with the FSC-A and SSC-A parameters, provide sufficient information to merge cell expression data obtained in different tubes into one merged fcs file. A first selection of cells was made by gating out the erythrocytes within single, viable cells based on CD45 and SSC-A using FlowJo V10®. Exported fcs files were imported in Infinicyt and merged using the backbone markers. Merged files from a time point of an individual patient were again exported as merged fcs file for subsequent analyses.

**Combiflow**
The merged fcs files are used as input data for the Combiflow pipeline, which will be made available on Github (https://github.com/rhoutsma/rhoutsma/tree/main). In brief, the pipeline requires five files in order to run. Firstly, the flow data, which is composed of merged fcs files. Here, we mainly used a combination of diagnosis and relapse fcs files from an individual AML patient, with addition of MRD files when available. Merged fcs files from 1-3 NBMs were included as healthy control. The fcs files were zipped and loaded into the pipeline. For our analyses all viable, single cells except erythrocytes were included. It is also possible to zoom in on the blast or lymphocyte population by gating and exporting these populations. The remaining four files are excel sheets detailing which fcs files are in the zipped folder (metadata), which parameters are included (panel), the names of the clusters (clusters) and the sample each cluster originates from and the condition it is assigned to (samples). Important is that the included parameters need to be present in all included fcs files. The excel file assigning a cluster to a condition cannot be prepared prior to running the analysis as it is based on the output of the FlowSOM clustering step. After assigning each cluster to a condition, the analysis can continue till the principle component analysis. Here, the markers and clusters to include in the analysis can be selected. In our analyses we included all clusters to gain insight in differences between both diagnosis and relapse and healthy BM cells. Another possibility is to exclude the healthy clusters to zoom in further on the differences between diagnosis and relapse.

**Statistical analyses**
Statistical differences of marker expression were determined by Wilcoxon rank-sum test. Multiple testing correction according to the FDR method was performed for the extended marker panel analysis in Figure 7. Univariate survival analysis was performed with the Kaplan Meier method. Differences in relapse free survival and overall survival were analyzed with the Log rank test. Cox proportional-hazards model was used for multivariate analysis to identify independent prognostic predictors. Confounding factors with known prognostic influence (age, risk group, treatment etc.) were included in the model. Hazard ratios and 95% CI for incidence ratio were calculated. Statistical significance of each individual covariate was determined based on the Wald test. Due to the small sample size the likelihood-ratio test was used to determine overall significance of the model. A p-value <0.05 was considered significant.

Supplementary Table and Figure legends

Supplementary table 1. Diagnoses of patient cohort
Supplementary table 2. Clinical characteristics of the AML cohort
Supplementary table 3. Patient characteristics of the longitudinal cohort
Supplementary table 4. Patient characteristics per figure
Supplementary table 5. Patient characteristics Combiflow cohort
Supplementary table 6. Antibodies used in this study

Supplementary Figure 1. PM marker expression correlations to patient characteristics and survival

A. Gating strategy for determining PM protein expression (MFI).
B. Schematic representation of the correlation between PM protein expression and mutational status, karyotype, or ELN risk. Colors indicate upregulation (red), downregulation (blue) and no significant difference (white). * p < 0.05, ** p < 0.01, *** p < 0.001 Mann Whitney U test (n=256).

C. Overall survival curve of AML cohort subdivided in risk groups. Risk groups based on ELN guidelines 2017 (n=256).

D. Hazard ratios of marker positivity for overall and relapse-free survival.

E. Pearson’s correlation coefficients of PM marker expression (MFI).

Supplementary Figure 2. CD33 expression NPM1wt and NPM1cyt AML patients

CD33 expression within different hematopoietic stem/progenitor compartment was determined in NPM1wt (n=5) and NPM1cyt (n=5) patients. One NBM is shown as healthy control.

Supplementary Figure 3. PM marker expression post induction therapy predict overall survival

Overall survival for PM marker positivity at diagnosis (n=72). A Log Rank p-value of p<0.05 was considered significant.

Supplementary Figure 4. Identified clusters in the tSNE landscape were assigned to a condition based on total cell numbers

A. tSNE landscape of patient 1 colored by clusters for all samples combined (left) or per sample (right)
B. Total cell count per cluster per sample. As an example, cluster 1 was assigned to healthy controls, cluster 33 to diagnosis, cluster 7 and 11 to MRD and cluster 27 to relapse.

Supplementary Figure 5. Further refinement of CombiFlow by including more PM markers

A. Fish plots depicting mutational changes between diagnosis and relapse for patient 22. B. tSNE landscapes of patient 22 created by the five markers (top) or aberrant markers selected from the 36-marker panel (bottom). C. PCAs depicting 40 clusters colored by condition: diagnosis (red circle), relapse (blue square) and healthy (green triangle). PCAs were based on the five markers (left) or aberrant markers selected from the 36-marker panel (right). D. Bar graphs depicting the ranking of the included markers for PC1 and PC2 based on the five markers (left)
Supplementary Figure 6. IL1RAP is the most defining marker in PCAs including aberrantly expressed markers selected from the 36-marker panel for 64 AML diagnosis samples

A. Bar graphs depict which marker was most often ranked first (top), second (middle), or third (bottom) for their contribution to principle component 1
B. Bar graphs depict which marker was most often ranked first (top), second (middle), or third (bottom) for their contribution to principle component 2

Reference List

1. de Boer B, Prick J, Pruis MG, Keane P, Imperato MR, Jaques J, et al. Prospective Isolation and Characterization of Genetically and Functionally Distinct AML Subclones. Cancer Cell. 2018;34:674-689 e678.
Supplementary figure 1

A. Flow cytometry panel showing viable, singlets, non-fluorescent, blasts, CD34+, SSC-A, FSC-A, CD19, CD45, and PE markers.

B. Heatmap showing upregulated (red) and downregulated (blue) markers for CD82, CD97, FLT3, IL1RAP, TIM3, CD25, and CD123. PM marker 1-sample t-test results are shown with ** for p < 0.01.

C. Kaplan-Meier survival analysis showing survival probability over time for poor (red), intermediate (blue), and good (green) strata. The p-value is 0.06.

D. Overall survival analysis showing hazard ratios (HR), 95% confidence intervals (95%CI), and p-values for PM markers CD82, CD97, FLT3, IL1RAP, CD25, and CD123.

E. Decision tree illustrating the probability of survival for different PM markers.
Supplementary figure 3

Survival probability

Time (days)

Strata

Number at risk

PM+

PM-

$\text{p} = 0.18$
Supplementary figure 4

A

Clusters

B
Supplementary figure 6

A

Position #1

Principle component 1

Position #2

Position #3
Position #1

Principle component 2

Position #2

Position #3

Marker

Count

Marker

Count

Marker

Count