

# Massively parallel flow-cytometry-based screening of hematopoietic lineage cell populations from up to 25 donors simultaneously

Jan Devan<sup>1,2\*</sup>, Michaela Sandalova<sup>1,2</sup>, Pamela Bitterli<sup>1,2</sup>, Nick Herger<sup>1,2</sup>, Tamara Mengis<sup>1,2</sup>, Kenta Brender<sup>1,2</sup>, Irina Heggli<sup>1,2</sup>, Oliver Distler<sup>1,2</sup>, Stefan Dudli<sup>1,2</sup>

<sup>1</sup>Center of Experimental Rheumatology, Department of Rheumatology, University Hospital Zurich, University of Zurich, Switzerland

<sup>2</sup>Department of Physical Medicine and Rheumatology, Balgrist University Hospital, Balgrist Campus, University of Zurich, Zurich, Switzerland

Correspondence to: Jan Devan, [jan.devan@balgrist.ch](mailto:jan.devan@balgrist.ch)

## Background:

A thorough understanding of biological processes and pathological conditions requires efficient extraction of information about individual cells in the tissue samples. In the last decade, single-cell RNA sequencing became a state-of-the-art technology to explore complex tissues. It allows the screening of thousands of transcripts on thousands of cells. However, biological samples are often very complex, and even though we have sufficient numbers of cells available in most situations, the limitation of the aforementioned currently available techniques to analyze only a few thousand cells hinders us from deeply exploring cellular tissue composition, particularly rare cell populations. This study aimed to develop a method allowing high-dimensional and technically uniform screening of surface markers on cells of hematopoietic origin.

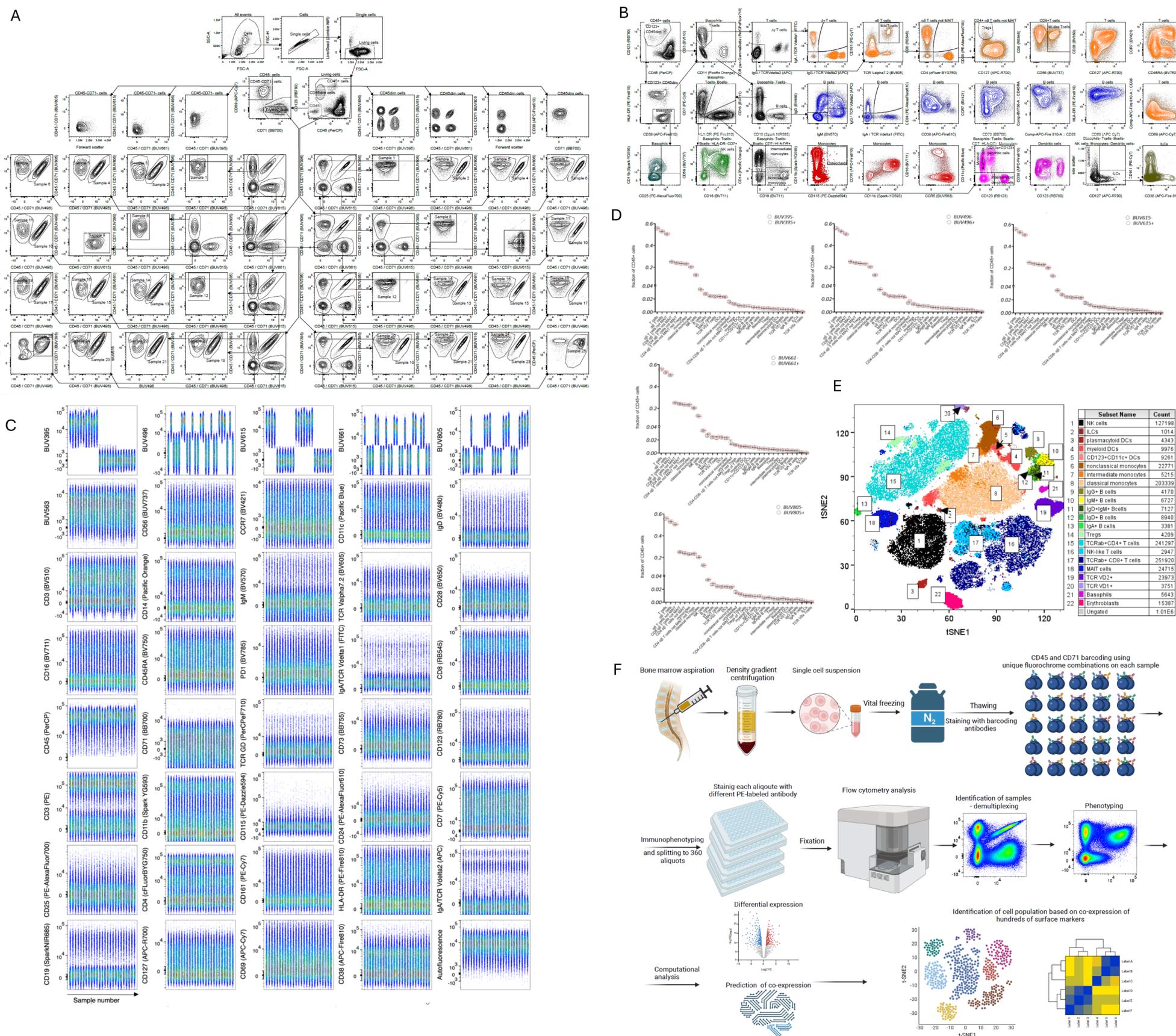
## Methods:

The flow-cytometry panel was developed on the 5-laser Cytek Aurora machine (355nm, 405nm, 488nm, 561nm, and 640nm) equipped with 64 detectors. Vertebral bone marrow aspirates from patients undergoing lumbar spinal fusion surgery were taken before screw insertion through the pedicle screw trajectory. Peripheral blood was taken from the same patients prior to surgery. All patients provided written informed consent. Mononuclear cells were isolated by density gradient centrifugation, and single-cell suspensions were cryopreserved until the final analysis. Marker and fluorochrome selection were performed in order to design a biologically informative panel with sufficient technical quality. Assignment of fluorochromes to antibodies was done based on the known co-expression of markers, level of their expression, known affinity of antibody to target molecule, similarity indexes.

## Results:

We utilize the highly expressed markers CD45 on immune cells and CD71 on erythroid progenitors to create unique fluorescent barcodes on each of the 25 samples (Fig.1A). CD71 is the marker of erythroid precursors, while CD45 is highly and selectively expressed on nucleated cells of hematopoietic origin except for erythroid lineage. Cells are stained with combinations of 1-3 antibodies against CD45 and CD71 from the pool of 5 (Fig.1A, C). Each sample is stained by the same fluorochrome combination for both antibodies. This creates a unique barcode on the surface of nucleated cells of hematopoietic origin from each sample, allowing their multiplexing and simultaneous analysis (Fig.1A). This is the first method describing such double-barcoding in an experiment. Additional antibodies against CD71 (BB700-P2) and CD45 (PerCP) are used to distinguish cells positive for CD71 and CD45 within the sample (Fig.1A). 34 other antibodies against surface markers allow the identification of T cells, B cells, NK cells, Monocytes, Dendritic cells, and ILCs, their lineage subsets, and activation and functional maturation stages (Fig.1B). Fluorescent barcoding was performed to reach technical uniformity of sample analysis. This goal can only be achieved if the barcode does not in any way impact the staining of markers in the immunophenotyping panel. By performing a whole experiment on one bone marrow sample divided into 25 fractions, we demonstrated that fluorescent barcodes do not affect staining by individual antibodies used for immunophenotyping. We observed no visual impact of fluorochromes used for barcoding on the intensity of fluorescence from other fluorochromes (Fig.1C). More importantly, when comparing the frequency of gated cell populations in samples stained with each of the fluorochromes used for barcoding with the rest of the samples, we observed no significant influence of any of the fluorochromes (Fig.1D). Multidimensionality reduction using t-distributed stochastic neighbor embedding (tSNE) revealed clusters overlapping with phenotypical populations achieved by manual gating (Fig.2E).

To sum up, here we present a methodology that allows simultaneous analysis of almost all nucleated cells of hematopoietic origin from up to 25 donors simultaneously. The combination of CD45 and CD71 for sample barcoding overcomes the limitation of previously used barcoding techniques and allows simultaneous analysis of changes in immune cells and changes in erythropoiesis, which are often closely related. Finally, this is the first protocol that combines the barcoding of samples using abundant surface molecules with over 40-parameter spectral flow cytometry. Notably, the phenotyping panel is designed for massively parallel flow cytometry screening using PE-labelled antibodies. This allows highly sensitive analysis of hundreds of markers on populations defined by the immunophenotyping panel. Lastly, the results of screening can be analyzed by machine learning algorithms such as InfinityFlow, which allows the prediction of co-expression of all screened markers with high accuracy, resulting in data containing several hundreds of parameters for each of the millions of analyzed cells. Such data can serve as input for the differential discovery analysis using packages like CATALYST or Seurat, which has also been successfully applied to flow cytometry data analysis (Fig.1F). The presented technique thus enables an analysis of changes in the populations of nucleated cells of hematopoietic origin defined by several hundred surface markers from up to 25 donors simultaneously.



**Figure 1. A.** A representative example shows the gating of CD45<sup>+</sup> cells and erythroid populations followed by the demultiplexing of individual samples in both populations based on the fluorescence barcode. CD45<sup>+</sup> (PerCP) and CD71<sup>+</sup> (BB700-P2) cells are not stained by any of the fluorochromes used for barcoding, suggesting the gating of CD45<sup>+</sup> and CD71<sup>+</sup> effectively includes all barcoded cells (left on top). The presence of barcoding antibodies on CD45<sup>dim</sup> CD123<sup>+</sup> cells (basophils), together with the absence of CD71 expression, suggests they should be included among CD45<sup>+</sup> cells (right on top). **B.)** A representative example of gating on CD45<sup>+</sup> cell lineage subsets and plots showing the expression of important activation and functional maturation markers for selected populations. **C.)** The fluorescence intensity of 40 parameters was plotted against 25 samples defined by fluorescent barcode (the same sample split into 25 aliquots). **D.)** Comparison of frequencies of CD45<sup>+</sup> cell subsets presented in Fig.1B between samples positive (n=11) and negative (n=14) for each of the 5 fluorochromes used for barcoding. **E.)** Manually gated populations from Fig.1B plotted over tSNE clustering plots of living single cells calculated based on the expression of all markers used in the analysis. **F.)** Schematic representation of full methodological pipeline.