Epigenetic and transcriptional single-cell profiling reveals CLL-like states in low-count MBL

Anja C. Rathgeber[1,2,3], Stacey M. Fernandes[4], Shuqiang Li[5], David M. Dorfman[6], Lars Bullinger[7], Mattew S. Davids[4], Jennifer R. Brown[4], Kenneth J. Livak[4,5], Leif S. Ludwig[1,2], Catherine J. Wu[4,8,9] and Livius Penter[4,7,10]

[1]Berlin Institute of Health at Charité - Universitätsmedizin Berlin, Germany; [2]Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association (MDC), Berlin Institute for Medical Systems Biology (BIMSB), Germany; [3]Department of Biology, Chemistry, Pharmacy, Freie Universität Berlin, Germany; [4]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, USA; [5]Translational Immunogenomics Lab, Dana-Farber Cancer Institute, Boston, USA; [6]Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, USA; [7]Department of Hematology, Oncology, and Tumorimmunology, Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Germany; [8]Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, USA; [9]Harvard Medical School, Boston, USA; [10]Berlin Institute of Health at Charité – Universitätsmedizin Berlin, BIH Biomedical Innovation Academy, BIH Charité Digital Clinician Scientist Program, Germany

Introduction

The premalignant conditions low-count (LC) and high-count (HC) monoclonal B cell lymphocytosis (MBL) precede chronic lymphocytic leukemia. Despite multiple genomic sequencing studies which showed a high genetic and epigenetic similarity between MBL and CLL[1-3], the mechanisms driving and characterizing the progression remain unknown. Taking advantage of technological advances, we performed an integrated single cell genomics analysis of transcriptional, epigenetic, and genetic profiles across the physiologic B cell to LC-/HC-MBL and CLL cell continuum. To also gain deeper insights into subclonal dynamics among the (pre-)malignant populations we combined single cell B cell receptor (scBCR) sequencing and mitochondrial single cell ATAC[4] (mtscATAC) sequencing data in our analysis of 24 patients.

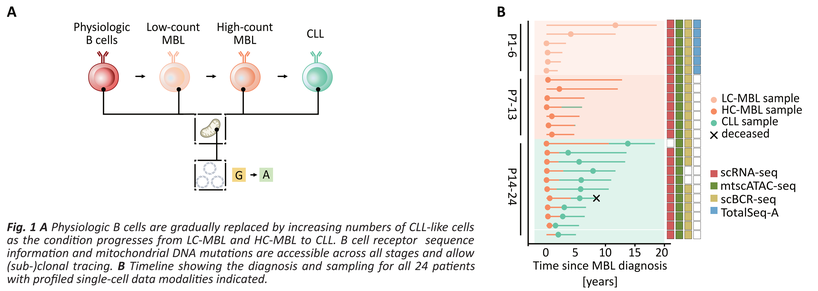

Fig. 1 *A Physiologic B cells are gradually replaced by increasing numbers of CLL-like cells as the condition progresses from LC-MBL and HC-MBL to CLL. B cell receptor sequence information and mitochondrial DNA mutations are accessible across all stages and allow (sub-)clonal tracing.* ***B*** *Timeline showing the diagnosis and sampling for all 24 patients with profiled single-cell data modalities indicated.*

Objectives

- Integrative characterization of epigenetic, transcriptional, and clonal changes in HC-MBL to CLL progression.
- Definition of epigenetic and transcriptional changes between B cells and LC-MBL/HC-MBL/CLL cells.
- Characterization of (sub-)clonal dynamics in LC-/HC-MBL to CLL progression using scBCR sequencing and mitochondrial DNA mutations.

(Immuno-)phenotypic CLL features in LC-MBL

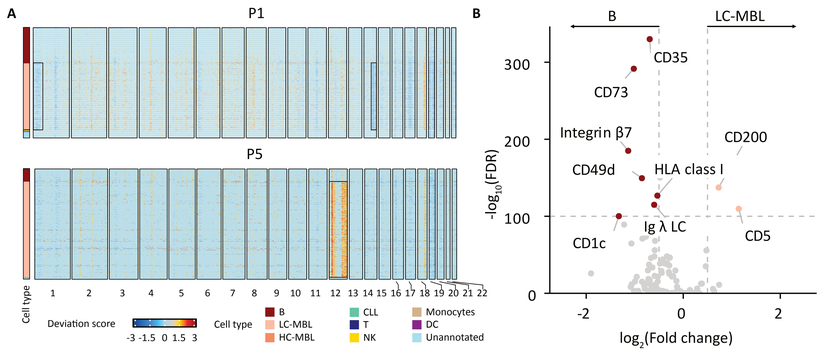

Fig. 3 *A Copy number changes calculated from scATAC-seq data for each cell type for patient P1 show deletions in chromosome 1 and a del(14q) as well as a tri(12) in P5. LC-MBL T cells were used as a reference offset.* ***B*** *Differential cell surface marker expresion between physiologic B cells of 6 LC-MBL patients and respective LC-MBL cell populations. Vertical and horizontal dashed lines indicate fold change thresholds and filtering of significance. Wilcoxon rank sum test with Benjamini-Hochberg correction*

Epigenetic and transcriptional changes arise at the stage of LC-MBL

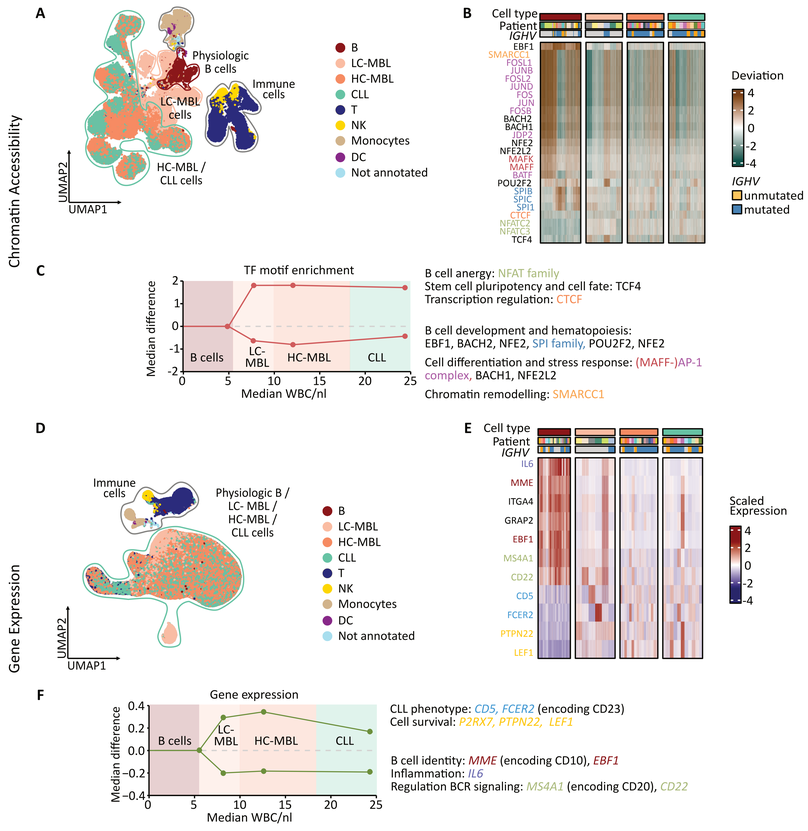

Fig. 2 *A Chromatin accessibility of 46,473 healthy and 161,968 (pre-)malignant cells of 23 patients and two healthy controls (H1-H2)[4-5]. Physiologic cell types were identified based on a healthy PBMC reference using label transfer. (Pre-)malignant cell types were annotated by imputed marker gene expression (gene scores) and diagnosis at sampling time.* ***B*** *Differential enrichment of transcription factor motifs between physiological B cells and LC-/HC-MBL or CLL cells. Deviation scores indicate enrichment or depletion over a GC-content matched set of background chromatin peaks in the indicated cell type. TF motif colors indicate protein complexes and families.* ***C*** *Median TF motif enrichment deviations across all up or down regulated TF motifs relative to the physiological B cells.* ***D*** *28,884 healthy and 298,009 (pre-)malignant cells of 24 patients and two publicly available healthy PBMC controls (H3-H4, 10X Genomics data sets). Physiologic cell types were identified based on marker gene expression and a healthy PBMC reference for label transfer. (Pre-)malignant cell types and physiologic B cells were annotated based on the diagnosis at sampling time as they could not be distinguished based on their gene expression.* ***E*** *Differential gene expression between physiologic B cells and LC-/HC-MBL or CLL cells. Colors indicate protein complexes and families (see Fig. 3).* ***F*** *Median gene expression changes across all differentially up or down regulated genes relative to the physiologic B cells. WBC - white blood cells*

Clonal expansion occurs as early as LC-MBL stage

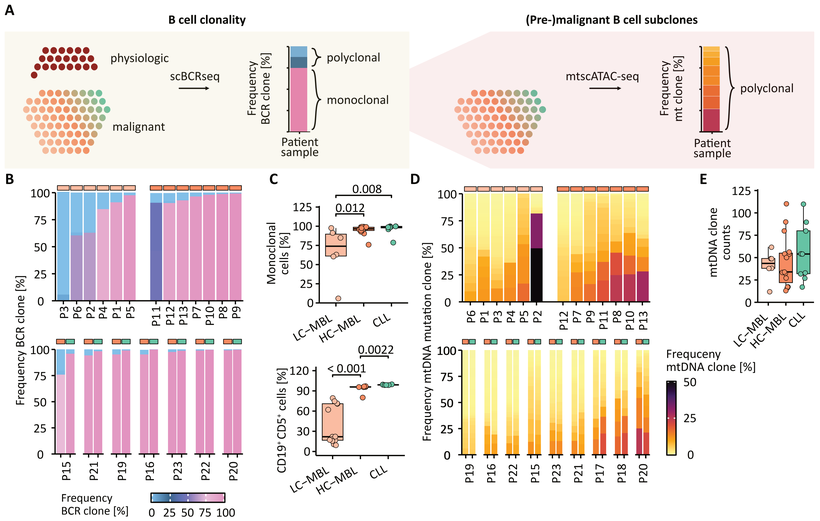

Fig. 4 *A B cell receptor sequence-based clones were called from physiologic B cells and (pre-)malignant B cells (LC-/HC-MBL and CLL) using scBCR-seq to determine the clonal repertoire and frequency. Subsequently mtDNA mutations were called from mtscATAC sequencing data and subclones were identified based on mtDNA mutations' heteroplasmy to determine their repertoire and frequency of (pre-)malignant cell types.* ***B*** *BCR clone frequencies of LC-MBL patients (top left), individual HC-MBL samples (top right) and patient-matched HC-MBL and CLL samples (bottom).* ***C*** *Quantification of frequency of major BCR clone per patient and condition (top). Flow cytometry based quantification of CD19+CD5+ B cells (bottom).* ***D*** *mtDNA mutation based clone frequencies of LC-MBL samples (top left), individual HC-MBL samples (top right) and patient-matched HC-MBL and CLL samples (bottom).* ***H*** *Quantification of mtDNA mutation based (pre-)malignant subclones.*

Subclonal stability in HC-MBL to CLL progression

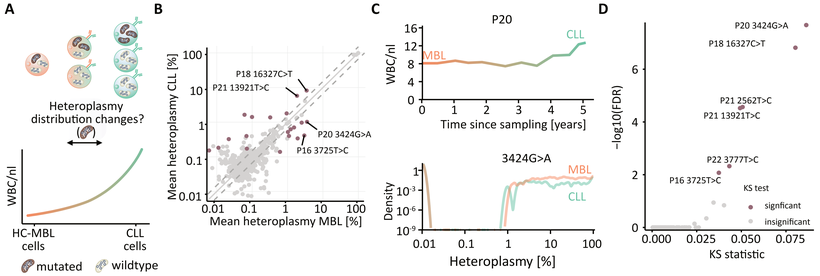

Fig. 5 *A Schematic of heteroplasmy changes during progression from HC-MBL to CLL.* ***B*** *Mean heteroplasmy for serial HC-MBL/CLL samples.* ***C*** *WBC counts of patient P20 (top) betwenn longitudinal samples. Heteroplasmy distributions of mtDNA mutation 3424G>A in both samples of P20 (bottom).* ***D*** *Detection of changes in mtDNA mutation distributions. Kolmogorov-Smirnov (KS) test and Benjamini-Hochberg FDR ≤ 0.05*

Summary

Our single cell omics analyses confirm high genetic, epigenetic, transcriptional, and (sub-)clonal similarity between HC-MBL and CLL suggesting an earlier disease onset.

Although physiologic B cells are more abundant in LC-MBL individuals than in advanced stages, LC-MBL cells already show a distinct epigenetic profile that phenocopies HC-MBL and CLL, which is reflected in their gene expression.

Large expanded clones are already present in a majority of LC-MBL individuals and increase in frequency during disease progression. Despite heterogenous (pre-)malignant subclone counts, the subclonal landscape in HC-MBL to CLL progression remains stable.

References

[1]Puente, X., Beà, S., Valdés-Mas, R. *et al.* Nature (2015)
[2]Barrio, S., Shanafelt, T., Ojha, J. *et al.* Leukemia (2017)
[3]Kretzmer, H. *et al.* Blood Cancer Discovery (2021)
[4]Lareau, C.A., Ludwig, L.S., *et al.* Nature Biotechnology (2021).
[5]Mimitou, E.P., Lareau, C.A., Chen, K.Y. *et al.* Nature Biotechnology (2021)

Contact
Anja C. Rathgeber
anja.rathgeber@bih-charite.de

Ludwig Lab

Penter Lab