A Multiomic Analysis of Waldenstrom's Macroglobulinemia Defines Three Subtypes of Disease

Dylan Gagler[1*], Hussein Ghamlouch[1,2*], Di Zhang[1], Patrick Blaney[1], Marine Armand[3,4,5], Alexandre Eeckhoutte[3,4], Amina Joudat[3,4], Michael Degaud[3,4], Michela Esposito[3,4], Gaurav Varma[1], Yubao Wang[1], Sanxiong Liu[1], Oscar Lahoud, David Kaminetzsky, Marc Braunstein, Florence Nguyen-Khac[5,6], Brian Walker[4], Damien Roos-Weil[5,7], Faith Davies[1], Olivier A. Bernard[3,4#], Gareth J Morgan[1#]

1. Myeloma Research Program, NYU Langone, Perlmutter Cancer Center, New York, NY, USA.
2. Service d'Hématologie Clinique et de Thérapie Cellulaire, CHU Amiens-Picardie, Amiens, France; HEMATIM, UR4666, Université Picardie Jules Verne, UPJV, Amiens, France.
3. INSERM U1170 ; équipe labélisée Fondation pour la Recherche Médicale ; Gustave Roussy, Villejuif, France.
4. Faculté de Médecine, Université Paris-Sud, Université Paris-Saclay, Orsay, France.
5. Centre de Recherche des Cordeliers, Sorbonne Université, Université Paris Cité, Inserm UMRS 1138, Drug Resistance in Hematological Malignancies Team, Paris, France.,
6. Service d'Hématologie Biologique, Hôpital Pitié-Salpêtrière, Sorbonne Université, AP-HP, Paris, France.
7. Sorbonne Université, Service d'Hématologie Clinique, Hôpital Pitié-Salpêtrière, APHP, Paris, France.
8. University of Indiana, Indianapolis.

Introduction

Waldenstrom's macroglobulinemia (WM) is a lymphoplasmacytic lymphoma which recent DNA methylation studies have shown to have a variable phenotype. To enhance disease classification and explore the features and potential mechanisms underlying these subtypes, we performed a single-cell multiomic analysis on a series of MYD88-mutated WM, bulk RNA-seq, and whole genome sequencing (WGS).

Methods

Single-cell (sc) multiomic analysis was carried out on flow sorted CD19+ mature B-cells from 13 MYD88-mutated WM patients and was analyzed alongside reference B-cell populations derived from healthy tonsils. Patient-matched scRNA-seq and scATAC-seq data were preprocessed in Seurat and ArchR prior to modality integration in ArchR. We employed a combined automated and manual cell type annotation approach using the CellTypist and TRUST4 packages to identify phenotype and clonality. A second series of WM patients (n = 37) underwent bulk RNA-seq and a subset (n = 32) had WGS data. WGS data underwent preprocessing and somatic variant analysis using our genomics pipeline, the MGP1000.

Results

We show that each patient phenotype was of an expanded memory B-cell (MBC) population which was clonally related via a shared CDR3 sequence. The analysis revealed three disease subtypes—MBC-like, PC-like, and intermediate—based on the presence of plasma cells (PC) and gene expression patterns of PC and MBC markers in the clonal MBC. Comparing the clonal MBC to healthy MBC showed that each subtype differentially expressed a substantial set of unique genes but largely shared a number of differential transcription factor motifs. Pseudotime trajectory analysis using ArchR suggests that WM is characterized by clonally expanded MBC with variably blocked capacity for plasma cell differentiation. Genes most correlated with pseudotime include *SPIB*, *BCL11A*, *STAT6*, *POU2F1*, *XBP1*, and *NFKB1*. We validated the existence of the three disease subtypes using hierarchical clustering of bulk RNA-seq. WGS for a subset of patients with expression data identified mutations in *MYD88* (90%), *CXCR4* (17%), *HIST1H1E* (14%), and *MAP3K14* (*NIK*) (11%), among others. The MBC-like patients

carried all of the *HIST1H1E* and *NIK* mutations in addition to the majority of *CXCR4* mutations (5 of 6). For somatic copy number abnormalities, we identified del 6q (43%), trisomy 4 (13%), and del 8p (13%), among others. The majority of del 6q were found in the PC-like and intermediate subtypes (11 of 15).

Conclusions

We show the existence of 3 subtypes of WM with distinct transcriptional programs and conclude that WM is a disease characterized by a failure to complete normal differentiation from an MBC to a PC. The MBC-like group have mutated *CXCR4*, *HIST1H1E*, and *NIK* with differentiation blockage at the MBC stage and is associated with the upregulation of the transcription factors *SPI1*, *BCL11A*, and *STAT6*. In contrast, the PC-like is unable to complete differentiation into a normal mature PC and is enriched in del 6q and characterized by upregulation of the transcription factors *POU2F1*, *XBP1*, and *NFKB1*.

Figure

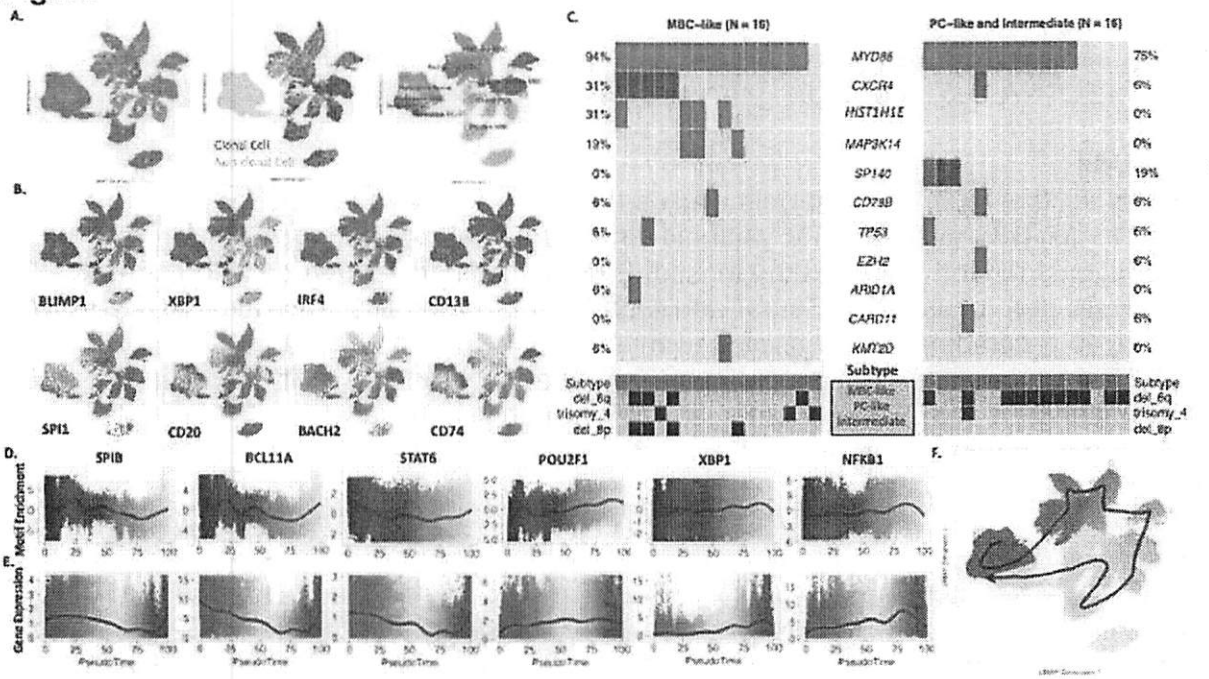

Figure 1. Single-cell multiome analysis of MYD88 Waldenstrom's macroglobulinemia patients reveals 3 subtypes of disease characterized by differentiation blockage at various stages. A. UMAP generated from clustering of sc-ATAC-seq data of WM patients (n = 13) and reference populations (n = 7) using ArchR, annotated by patient; same UMAP wherein each cell is colored by whether that cell's TRUST4 CDR3 assignment was part of the most abundant clone in that cell's patient of origin; UMAP of all cells used in this study, both WM and reference, annotated by cell type and WM subtype. **B.** UMAPs overlain with gene expression of key marker genes for plasma cells (*BLIMP1*, *XBP1*, *IRF4*, *CD138*) and memory B cells (*SPI1*, *CD20*, *BACH2*, *CD74*). **C.** Mutated gene frequencies of known canonically mutated WM genes in MBC-

like subtype patients (n = 16) compared to combined PC-like (n = 11) and intermediate subtype (n = 5) patients, annotated by WM subtype and the presence of select CNVs. **D.** Motif enrichment as a function of pseudotime for select genes putatively involved in aberrant MBC differentiation. SPIB and BCL11A are known repressors of PC differentiation and *POU2F1* (*OCT1*) and *XBP1* are known to promote PC differentiation. **E.** Gene expression as a function of pseudotime select genes putatively involved in aberrant MBC differentiation. **F.** scATAC-seq based UMAP overlain with ArchR-derived pseudotime values, showing differentiation progressing from reference MBC, to the MBC-like, then intermediates and PC-like, and finally the reference PC.