American Society of Hematology
2021 L Street NW, Suite 900,
Washington, DC 20036
Phone: 202-776-0544 | Fax 202-776-0545
editorial@hematology.org

A multiomic analysis of Waldenström macroglobulinemia defines distinct disease subtypes

Tracking no: BLD-2024-028164R1

Dylan Gagler (NYU Perlmutter Cancer Center, United States) Hussein Ghamlouch (CHU Amiens, France) Di Zhang (NYU Langone Health, United States) Patrick Blaney (NYU Perlmutter Cancer Center, United States) Avital Tenenbaum (NYU Perlmutter Cancer Center, United States) James Langton (NYU Perlmutter Cancer Center, United States) Marine Armand (Hopital Pitié Salpétrière, APHP, France) Alexandre Eeckhoutte (INSERM, France) Amina Joudat (INSERM, France) Michaël Degaud (Institut Gustave Roussy, France) Michela Esposito (INSERM, France) Gaurav Varma (NYU Langone Health, United States) Yubao Wang (New York University Langone Health, United States) Sanghoon Lee (NYU Perlmutter Cancer Center, United States) Sanxiong Liu (NYU Perlmutter Cancer Center, United States) Oscar Lahoud (NYU Langone - Perlmutter Cancer Center, United States) David Kaminetzky (Perlmutter Cancer Center at NYU Langone Health,) Marc Braunstein (NYU Langone Medical Center, United States) Louis Williams (Taussig Cancer Center, Cleveland Clinic Foundation, United States) Florence Nguyen-Khac (Hôpital Pitié-Salpêtrière, Sorbonne Université, France) Brian Walker (University of Miami, United States) Damien Roos-Weil (Pitié-Salpêtrière, France) Faith Davies (New York University Langone, United States) Olivier Bernard (INSERM, France) Gareth Morgan (NYU Langone Health Care, United States)

Abstract:
We carried out a single-cell (sc) multiomic analysis on a series of MYD88 mutated Waldenström macroglobulinemia (WM) cases and identified two distinct subtypes of disease, memory B-cell-like (MBC-like) and plasma cell-like (PC-like), based on their expression of key lineage defining genes. Biologically, the subtypes are characterized by their variable capacity to differentiate fully towards a plasma cell (PC) and exhibit unique transcriptomic, chromatin accessibility, and genomic profiles. The MBC-like subtype is unable to differentiate beyond the memory B-cell (MBC) stage, upregulates key MBC genes, and is characterized by upregulated BCR and AKT/mTOR signaling. In contrast, the PC-like subtype can partially differentiate towards a PC, upregulates key PC genes, has enhanced NF-kB signaling, and has an upregulated unfolded protein response. Pseudotime trajectory analysis of combined scRNA-sequencing and scATAC-sequencing supports the variable differentiation capacity of each subtype and implicate key transcription factors SPI1, SPIB, BCL11A, and XBP1 in these features. The existence and generalizability of the two disease subtypes were validated further using hierarchical clustering of bulk RNA-seq data from a secondary set of cases. The biological significance of the subtypes was further established using whole genome sequencing, where it was shown that CXCR4, NIK, and ARID1A mutations occur predominantly in the MBC-like subtype and 6q deletions in the PC-like subtype. We conclude that the variable differentiation blockade seen in WM manifests itself clinically as two disease subtypes with distinct epigenetic, mutational, transcriptional, and clinical features with potential implications for WM treatment strategies.

Conflict of interest: No COI declared

COI notes:

Preprint server: No;

Author contributions and disclosures: GJM and OAB supervised the research and the writing of the paper. GJM, OAG, DCG, and HG conceived, designed, and developed the methodology of the study. DCG performed research, bioinformatic analysis, and figure generation and participated in the writing of the paper. HG and DZ performed laboratory experiments and research, generated data, and participated in the writing of the paper. GJM, DCG, HG, PB, AE, GV, YW, SL, BW, and FED participated in the analysis and interpretation of data. DRW, PM, FNK, OL, LW, DK, MB, VL, FNK, DK, SL, ME, MD, AJ, MA, JL, AT, and FED contributed reagents, samples, and data. OAB contributed data, performed research, and contributed to the writing of the paper. All authors reviewed and edited the manuscript.

Non-author contributions and disclosures: No;

Agreement to Share Publication-Related Data and Data Sharing Statement: Public repository of single-cell raw data will be made available (through European Genome-Phenome Archive) with accession number TBD.

Clinical trial registration information (if any):

A multiomic analysis of Waldenström macroglobulinemia defines distinct disease subtypes

Dylan C. Gagler[1*], Hussein Ghamlouch[1,2*], Di Zhang[1], Patrick Blaney[1], Avital Tenenbaum[1], James Langton[1], Marine Armand[3,4,5], Alexandre Eeckhoutte[3], Amina Joudat[3], Michaël Degaud[3], Michela Esposito[3], Gaurav Varma[1], Yubao Wang[1], Sanghoon Lee[1], Sanxiong Liu[1], Oscar Lahoud[1], David Kaminetzsky[1], Marc Braunstein[1], Louis Williams[6], Florence Nguyen-Khac[4,5], Brian Walker[7], Damien Roos-Weil[4,5,8], Faith E. Davies[1], Olivier A. Bernard[3#,] Gareth J. Morgan[1#]

Author Affiliations

[1]Multiple Myeloma Research Program, Perlmutter Cancer Center, NYU Langone Health, 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]Université Paris-Saclay, Inserm, Gustave Roussy, Dynamique moléculaire de la transformation hématopoïétique, 94805, Villejuif, France.
[4]Centre de Recherche des Cordeliers, Sorbonne Université, Université Paris Cité, Inserm UMRS 1138, Drug Resistance in Hematological Malignancies Team, Paris, France.
[5]Service d'Hématologie Biologique, Hôpital Pitié-Salpêtrière, Sorbonne Université, AP-HP, Paris, France.
[6]Cleveland Clinic, Cleveland, Ohio, USA
[7]Simon Comprehensive Cancer Center, Indiana University, Indianapolis, Indiana, USA.
[8]Sorbonne Université, Service d'Hématologie Clinique, Hôpital Pitié-Salpêtrière, APHP, Paris, France.
[*]These authors contributed equally; [#]These authors co-directed this work.

Lead contact: Gareth Morgan

Corresponding Authors

Olivier A Bernard
INSERM U1170
Gustave Roussy
114 rue Edouard-vaillant
94805 Villejuif France
Email: Olivier.bernard@gustaveroussy.fr

Gareth Morgan
Director Myeloma Research
Perlmutter Cancer Center
550 First Avenue
New York, NY, 10016
Phone: +1 (305) 243-7687
Email: Gareth.morgan@nyulangone.org

Running Title

A Multiomic Analysis of WM Subtypes

Counts

Running Title	30 characters	(Max 50)
Abstract	241 words	(Max 250)
Body	3945 words	
Figures	5	
Tables	1	
References	63	

Data Sharing Statement

Raw and processed single-cell data used for this analysis is available on the Gene Expression Omnibus (GEO) under accession number GSE296167. Additional raw data, including WGS and bulk RNA-seq, are available upon request.

Key Points

1) Single-cell multiome analysis identifies two subtypes of WM with variable differentiation capacity and *SPI1*, *SPIB*, and *XBP1* activity.
2) The MBC-like have increased CXCR4 mutations, BCR, and PI3K/AKT/mTOR signaling, and the PC-like have increased del(6q) and NF-κB signaling.

Abstract

We carried out a single-cell (sc) multiomic analysis on a series of *MYD88* mutated Waldenström macroglobulinemia (WM) cases and identified two distinct subtypes of disease, memory B-cell-like (MBC-like) and plasma cell-like (PC-like), based on their expression of key lineage defining genes. Biologically, the subtypes are characterized by their variable capacity to differentiate fully towards a plasma cell (PC) and exhibit unique transcriptomic, chromatin accessibility, and genomic profiles. The MBC-like subtype is unable to differentiate beyond the memory B-cell (MBC) stage, upregulates key MBC genes, and is characterized by upregulated BCR and AKT/mTOR signaling. In contrast, the PC-like subtype can partially differentiate towards a PC, upregulates key PC genes, has enhanced NF-κB signaling, and has an upregulated unfolded protein response. Pseudotime trajectory analysis of combined scRNA-sequencing and scATAC-sequencing supports the variable differentiation capacity of each subtype and implicate key transcription factors *SPI1*, *SPIB*, *BCL11A*, and *XBP1* in these features. The existence and generalizability of the two disease subtypes were validated further using hierarchical clustering of bulk RNA-seq data from a secondary set of cases. The biological significance of the subtypes was further established using whole genome sequencing, where it was shown that *CXCR4, NIK,* and *ARID1A* mutations occur predominantly in the MBC-like subtype and 6q deletions in the PC-like subtype. We conclude that the variable differentiation blockade seen in WM manifests itself clinically as two disease subtypes with distinct epigenetic, mutational, transcriptional, and clinical features with potential implications for WM treatment strategies.

Introduction

Waldenström Macroglobulinemia (WM) is a rare clinicopathologic entity consisting of a lymphoplasmacytic lymphoma with an associated IgM paraprotein, that has been hypothesized to arise from clonal-expansion of post-germinal center (GC) B-cells.[1] Morphologically, the neoplastic cells have variable appearance, resembling lymphocytes, lymphoplasmacytes, and plasma cells, and most commonly express surface $CD19^{+}$, $CD22^{low}$, $CD23^{-}$, $CD25^{+}$, $CD27^{+}$, and $SmIgM^{+}$, with clonally restricted κ or λ light chains.[2] However, immunophenotype alone is not able to identify subtypes of WM or account for variability in the composition of the neoplastic infiltrate observed between patients. MYD88 L265P mutations are a hallmark of WM and are present in >90% of patients.[3] MYD88 functions as an adaptor protein in the interleukin and toll-like receptor (TLR) signaling pathways. When mutated, MYD88 recruits IRAK4 to IRAK1/IRAK2 to form the "myddosome" complex and increases pro-survival signaling via NF-κB, as well as HCK-mediated crosstalk with the B-cell receptor (BCR) pathway via BTK/NF-κB, PI3K/AKT, and MAPK/ERK.[4-7]

Recurrent mutations at *CXCR4* and *ARID1A* as well as copy number abnormalities at chromosome 6q are observed in 30-40%, 17%, and 30-55% of patients, respectively, and are considered secondary molecular events.[8-10] Mutations in CXCR4 promote constitutive PI3K/AKT and MAPK/ERK signaling, leading to improved tumor cell fitness and drive sub-clonal expansion. Copy number abnormalities are important drivers of disease progression, with interstitial loss of chromosome 6q being the most frequent. Chromosome 6q deletions are variable and known to affect several genes including *BCLAF1*, *BLIMP1*, and *IBTK*, but the full profile of deleted regions has not been fully characterized.[11-13]

The use of BTK inhibitors, which inhibit downstream NF-κB signaling and induce apoptosis in WM cells, has significantly improved the clinical management of WM.[14] However, BTK inhibition is less effective in *CXCR4* mutated patients, for imperfectly understood reasons, and complete responses, even with second-generation BTK inhibitors, are rare.[9,15] Further understanding the biology of the tumor cells in WM could enhance our strategies for treatment and enhance treatment to achieve complete responses.

Recent DNA methylation studies have identified two disease subtypes of WM with memory B-cell (MBC) and plasma-cell (PC) features.[16] The mechanisms driving these subtypes and the distinctive features of their biology, however, remain unknown. To improve our understanding of the biology of WM and to enhance disease sub-classification, we performed a single-cell (sc) multiome analysis (scRNA-seq and scATAC-seq) and bulk whole genome sequencing (WGS) of B-cells derived from a series of MYD88-mutated WM cases.

Methods

Single-cell multiome analysis

Written consent for BM samples and analyses was obtained in accordance with the Declaration of Helsinki and with ethical approval from local and national ethics committees (NYU Langone S21-01218 and S19-01384; Gustave Roussy CNIL 2212382 and CPP Ile-De-France 05/21/2014). Patients were identified using IRB-approved protocols with *MYD88* status identified by clinical testing. Patients were untreated at the time of sample collection. Reference scRNA-seq and scATAC-seq data were derived from healthy tonsils.[17] B-cells were flow sorted from bone marrow (BM) using CD3-/CD19+ from 13 patients with MYD88 mutated WM. Cells were stained with CD14-APC (Biolegend, Ref#367118), CD3-FITC (Biolegend, Ref#300306), CD19-PE/Cyanine7 (Biolegend, Ref#302216), and live cells were flow cytometric sorted. Approximately 70,000-100,000 cells were used to generate nuclei using the 10X protocol "Low Cell Input Nuclei Isolation" (10x Genomics, CG000365.Rev C). After combining barcoded gel beads, Gel Beads-in-emulsion (GEMs) were generated, forming a pool of transposed DNA and barcoded, full-length cDNA from poly-adenylated mRNA. After post-GEM cleanup, barcoded transposed DNA and cDNA pre-amplification PCR were carried out in a pool to fill gaps and generate sufficient mass for library construction. Nuclear suspensions were incubated in a Transposition Mix, including Transposase and adapters. 10,000 nuclei were loaded onto the 10X platform for scATAC-seq and scRNA-seq and were sequenced on an Illumina NovaSeq 6,000 to 100,000 reads.

In brief, both WM (n = 13) and reference (n = 7) scRNA-seq data were preprocessed and quality filtered with standard thresholds prior to normalization and SCTransform-based integration in Seurat and automated cell type annotation in CellTypist.[18-20] The TRUST4 package was used for immune repertoire reconstruction using the scRNA-seq data.[21] The patient-matched scATAC-seq data were preprocessed, quality filtered, and underwent dimensionality reduction and clustering in ArchR using a standard protocol before multi-modal

integration with the scRNA-seq data in ArchR. Immature B-cells were removed from the dataset, which was then reclustered and batch corrected using ArchR's implementation of Harmony.[22,23] A UMAP was generated for data visualization, and the final cell count of the integrated object was 129,372 cells, of which 99,888 were WM and 29,484 were references. To determine transcription factor motifs active across cell types at the single-cell level, the chromVAR package, implemented within ArchR, was used. The ArchR package was used for pseudotime and trajectory analyses.[24,25] For more methodological details, see Supplementary Data.

Whole genome sequencing analysis and bulk RNA-seq

36 WM samples underwent WGS; 11 samples had corresponding single-cell data discussed above, and another 21 had bulk RNA-seq data, **Supplementary Table 1**. B-cells were flow-sorted (CD14-CD3-CD19+) from bone marrow samples. Simultaneously, the corresponding T-cells were isolated as normal controls via CD14-CD19-CD3+. Genomic DNA was extracted using the AllPrep DNA/RNA Micro Kit (Qiagen, Ref#80284) and quantified with the Invitrogen Qubit 4 Fluorometer. Genomic DNA from both B-cells and T-cells was sequenced using an Illumina NovaSeq 6000 to a depth of 80x. WGS data underwent harmonized preprocessing and somatic variant analysis using Myeloma Genome Pipeline 1000 (https://github.com/pblaney/mgp1000). In brief, coverage metrics were generated for each BAM file, and each tumor/normal pair was subjected to the determination of SNVs, Indels, and CNVs before generating consensus somatic calls. Bulk RNA sequencing was performed on CD19+ sorted cells from 39 WM bone marrow samples, as previously described, of which 21 also underwent WGS as described above.[16]

Results

WM cells comprise a clonally expanded population of MBC

CellTypist automated cell type annotation and manual review of scRNA-seq expression data identified 9 B-cell clusters, including MBC, PC, age-associated B-cells, naïve B-cells, GC B-cells, proliferative GC B-cells, small pre-B cells, large pre-B cells, and pro-B cells, **Supplementary Figure 1.** The clonal MBC expressed IgM, IgD, IgG, and low IgA, were *FCRL5* and *SOX5* moderate, and approximately 50% of cells were *CD27* positive. Furthermore, the cells were *CD11C* (*ITGAX*), *CD21* (*CR2*), and *DLL1* negative, **Supplementary Figure 2**. The annotated scRNA data was integrated with patient-matched scATAC data and, after excluding precursor/immature B-cells and reclustering, yielded a dataset of 129,372 cells, of which 99,888 were WM and 29,484 were reference cells, **Supplementary Table 2**. Using the TRUST4 software for BCR repertoire reconstruction, most cells in each WM case were found to be clonally related via shared CDR3 amino acid sequences, **Figure 1E, Supplementary Figure 3, and Supplementary Table 3**. A subset of patients (n = 6), however, contained distinct subpopulations of MBC which were not clonally related, and which formed a single cluster in the overall UMAP, which are subsequently referred to as "Polyclonal MBC".

Two subtypes of WM were identified by clustering multiomic data

The scATAC-derived UMAP showed that WM cells clustered near to but separate from the reference cell populations and that most clonal cells formed patient-specific clusters despite Harmony batch correction, **Supplementary Figure 4**. Within the WM cases, two broad subtypes of disease were identified, MBC-like and PC-like, based on clonal PC content and the gene expression of PC and MBC markers in their clonal MBC populations, **Figure 1**. The MBC-like were identified by a lack of PC markers and an expression profile of *CD74*, *BACH2*, *CD27low*, *CD20, SPI1, HLA-DRA,* and *RACK1*, but at lower levels than the reference MBC, **Supplementary Figure 5.** Moreover, the MBC-like samples clustered in proximity to each other on the UMAP and near the putatively normal polyclonal MBC.

In contrast, the PC-like samples each contained a population of plasma-related cells (1-11%), which strongly expressed *BLIMP1, XBP1, IRF4, MZB1, JCHAIN,* and *CD138,* but less than reference PC, **Figure 1**. Furthermore, the MBC of the PC-like were distinguished by greater expression of *JCHAIN*, *MZB1*, *XBP1*, *PRDM1,* and *CD27* than either the MBC-like MBC or reference MBC, but at lower levels than the clonal PC and reference PC. One sample (WM43) had a unique expression profile (*XBP1+/CD27-/BACH2+/CD74-)* and clustered away from all other cells and was therefore not characterized as either MBC-like or PC-like, **Supplementary Figure 6.** The existence of MBC-like and PC-like cells was validated using flow cytometry of CD19, CD38, and CD138, on a representative case from each of the subtypes, **Supplementary Figure 7**. Furthermore, the expression ratio of the key PC-MBC transcription factors *BLIMP1* and *BACH2* was derived, and the MBC-like case had a lower ratio than the PC-like, **Supplementary Table 4.**

Gene expression and chromatin accessibility profiles of WM subtypes are distinct

The gene expression and chromatin accessibility profiles of the WM cells were compared to reference populations, **Figure 2**. For each subtype, we isolated the clonal and reference MBC, reclustered them based on their peak matrices, and performed pairwise Wilcoxon testing on both gene expression and chromVAR-derived transcription factor motif deviations. Differential genes were considered significant with $\log_2FC > 0.5$ and $p < 5e^{-9}$, and differential motifs were considered significant with mean differences > 0.025 and $FDR < 5e^{-9}$.

By comparison to the reference MBCs, both subtypes of WM showed distinctive gene expression profiles, with each subtype significantly and uniquely upregulating between 157 and 556 genes, **Figure 2G**. The MBC-like subtype largely upregulated genes broadly involved in BCR signaling (*CD79B, SYK, PLCG2, FCHSD2, OSBPL10, AFF3),* whereas the PC-like subtype upregulated a set of PC differentiation genes (*XBP1, JCHAIN, CD9*) and protein homeostasis genes (*PELI1, EIF2AK3, ERN1*). Both subtypes expressed a similar set of 42 genes, 5.5% of the total differentially expressed genes (DEG). The shared upregulated genes included *BCL2, FOXP1, PLCG2, HDAC9, ZNF804A,* and *VOPP1* which are consistent with common biological features of WM including apoptosis resistance, NF-κB signaling, BCR signaling, histone modification, and disease progression.[26-30] In contrast to the gene expression, the subtypes had broadly similar transcription factor motif enrichment profiles when compared to reference MBC, sharing 171 (59%) of the differentially enriched motifs, **Figure 2H**. Among the shared enriched motifs were those associated with epigenetic function (*ARID3A, TOPORS, HMGA1, FUBP1*), NF-κB activation (*NFKB2, REL*), and control of GC and B-cell development (*TCF4, ID3*).[31-35]

We compared the clonal MBC between the subtypes and found significant differences, highlighting their B-cell differentiation state and disease biology, **Figure 3**. The MBC-like were significantly enriched in only a handful of transcription factor motifs, including *SPI1, SPIB, SPIC, BCL11A,* and *BCL11B*. Notably, *SPI1* and *SPIB* have been implicated in the inhibition of PC differentiation, including in the context of WM.[36-38] In contrast, the PC-like were enriched in a diverse set of motifs, including those associated with plasma cell differentiation and protein secretion (*XBP1*, *ATF6*), control of B-cell and GC development (*LMO2, OCT1, OCT2, TCF3, TCF4, ID3*), NF-κB signaling (*REL, RELA, NFKB1, NFKB2, FOXO1*), AP-1 transcription factors (*JUN, FOS, BATF*), and endothelial-to-mesenchymal-transition (*SNAI1, TWIST1*), **Figure 3C, Figure 3D, and Supplementary Figure 8**.[35,39-43] Differentially expressed genes between the subtypes reinforced the above motif patterns, with the MBC-like upregulating *SPI1*, *SPIB*, and *BCL11A* and the PC-like upregulating *XBP1, ZEB1, FOXO1, RELB*, and *JUND,* **Figures 3A and 3B**. Furthermore, Gene Set Enrichment Analysis (GSEA) between the two subtypes showed enrichment of PI3K/AKT/mTOR signaling and several proliferative pathways (MYC, G2M checkpoint, E2F targets, mitotic spindle) in the MBC-like and NF-κB signaling and an inflammatory response in the PC-like and, **Figures 3E and 3F.**

The MBC-like are blocked in their capacity for plasma cell differentiation

A differentiation trajectory analysis in ArchR, based on the scATAC dimensionality reduction, showed a continuous trajectory of differentiation from polyclonal MBC to MBC-like MBC to PC-like MBC to clonal PC, **Figure 4A**. Notably, each sample cluster fell along this trajectory from those more like MBC to those more like PC, consistent with their subtype annotations. Correlating the differentiation trajectory to both gene expression and chromatin accessibility data, we found a set of genes that best capture differentiation as a function of pseudotime, **Figure 4C-F**. This included *SPI1, SPIB, and BCL11A* for the MBC-like and a wide suite of genes for the PC-like, including *XBP1, NFKB1, RELB, ATF6, STAT3, FOXO1*, and *ZEB1*. This further supports the conclusion that the MBC-like subtype is frozen in its capacity to differentiate due to sustained activity of *SPI1*, *SPIB,* and *BCL11A* and that the PC-like program is driven by *XBP1* while also maintaining a GC chromatin accessibility profile and upregulating NF-κB signaling.

Clonal WM plasma cells do not complete normal terminal differentiation and exhibit NF-κB signaling

Using both pairwise Wilcoxon testing of gene expression and GSEA, we show that WM clonal PC are phenotypically distinct from reference PC, do not complete differentiation towards a normal PC, and have increased BCR and NF-κB signaling, **Supplementary Figure 9**. Differentially upregulated genes in the clonal PC compared to the reference PC indicate enhanced BCR signaling (*LYN, BANK1, CD79B, PLCG2, FOXP1, SOX5, AFF3, OSBPL10, SMCHD1*) and NF-κB signaling (*NFKB1*, REL) with downregulated genes consistent

with their failure to complete normal plasma cell differentiation (*BLIMP1, XBP1, MZB1*)**.** GSEA between clonal PC and reference PC showed that the clonal PC were enriched in NF-κB and IL2/STAT5 signaling**.** Comparing the clonal PC and the paired PC-like MBC, we observed that the clonal PC have decreased NF-κB signaling (*REL*, *RELB*, *NFKB1*, *NFKBID*, *TRAF3*, *TRAF5*) and show increased PC marker gene expression (BLIMP1, XBP1, MZB1, CD38, CD138), although at decreased levels compared to the reference PC, as previously mentioned, **Supplementary Figure 10.** GSEA analysis between these groups confirms that the MBC-like PCs are enriched in NF-κB signaling and further reveals that the clonal PC are significantly enriched in unfolded protein response (UPR) and protein secretion. We also observed a set of PCs which clustered together but was derived from multiple patients, dubbed "Polyclonal PC", which presented a profile consistent with normal plasma cell function including upregulation of *BLIMP1*, *PAG1*, *DUSP5*, and *SSR4*, **Supplementary Figure 11**. Compared to the polyclonal PC, the clonal PC exhibited a profile associated with NF-kB signaling, anti-apoptosis, and BCR signaling via upregulation of genes including *FOXP1*, *CD79B*, *LYN*, *BCL2*, *NFKB1*, and *CD9*.

Clinical and mutational features of the disease subgroups

Using a series of clinically characterized WM cases with bulk RNA-seq, we performed a hierarchical clustering analysis. Clustering using a set of marker genes derived from the sc experiment recapitulated the two subtypes, **Figure 5A**. For the MBC-like, the signature included the genes *CD24, AFF3, ITGB1, BANK1, CD74, FCHSD2, AKT3, FCRL1, PYHIN1, SOX5,* and *SGSM1* whereas the PC-like signature included *CD9, EIF2AK3/PERK, MSI2, CD19, MS4A1/CD20, MZB1, XBP1, FCRL5, PELI1, PRDM1/BLIMP1, JCHAIN, SDC1/CD138, CD38,* and *JCHAIN*. We performed a second validation in another set of publicly available bulk RNA-seq data from WM patients (Mondello et al., 2023) and were again able to recapitulate the 2 subtypes via a hierarchical clustering analysis, **Supplementary Figure 12**. The relevance of surface marker genes including $CD138^+/CD38^+/CD27^+/FCRL5^+$ for the PC-like and $CD24^+/CD74^+$ for the MBC-like suggests that at least some of these markers may be useful for clinical identification of subtypes, but this requires future validation.

WGS analysis on a combined set of subtyped cases with subtype annotations (10 MBC-like and 22 PC-like) derived from either the single-cell analysis (n = 11) or bulk RNA-seq analysis (n = 21) revealed the genetic landscape of the subtypes**.** Overall, recurrent mutations were identified in *MYD88* (97%), *IGLL5* (53%), *CXCR4* (25%), *HIST1H1E* (16%), *ARID1A* (16%), *KMT2D* (12%), *MAP3K14* (*NIK)* (9%), *SP140* (9%), *TP53* (9%), *CD79B* (6%), *EZH2* (3%), and *CARD11* (3%), **Figure 5B**. Mutational distributions varied between the subtypes, however, with the MBC-like harboring the majority of *CXCR4, HIST1H1E (H1-4), and MAP3K14 (NIK)*, whereas the PC-like had all of the *SP140* mutations and an overall lower tumor mutational burden (0.65 per megabase vs. 0.79 per megabase), **Supplementary Figure 13**. For somatic copy number abnormalities, we identified del(6q) (27%), del(22p) (27%), trisomy 4 (21%), del(13p) (21%), trisomy 18 (12%), trisomy 12 (9%), and del(17p) (9%), **Figure 5C**. Most notably, del(6q) was more frequent in the PC-like (32% PC-like vs. 20% MBC-like).

Clinically, the MBC-like patients had greater circulating IgM levels, greater splenic enlargement, less adenopathy, and an elevated LDH consistent with what has been observed in *CXCR4* mutated cases previously. The PC-like were associated with higher levels of platelets and lymph node enlargement, **Table 1.**

Discussion

Prior studies on WM have identified broad categories of phenotypic heterogeneity and clonal expansion of lymphoplasmacytic cells but lacked the single-cell resolution to define distinct differentiation states and cellular biological characteristics.[16,44-47] While they recognized the diversity of clonal MBC populations and identified subtypes based on their mutational landscape (e.g., *MYD88* and *CXCR4* mutations), the cellular and molecular basis of the differences was not fully understood. To address these gaps, we used sc multiomic profiling to characterize WM and identified a key disease feature, which is its failure to complete normal PC differentiation, leading to 2 broad disease entities: MBC-like and PC-like. Our findings provide novel insights into WM pathogenesis by identifying a role for differentiation blockade and the specific transcriptional and genomic patterns associated with the disease subgroups.

Both major subgroups of WM are characterized by clonal expansions of MBC populations but are distinguished by their variable ability to differentiate towards a PC. The MBC-like group comprises an expanded population

of MBC blocked in their capacity to differentiate beyond the MBC stage via *SPI1*, *SPIB*, and *BCL11A*, have features of increased BCR and PI3K/AKT/mTOR signaling, and carry mutations in *CXCR4, HIST1H1E, NIK*, and *ARID1A.* In contrast, the PC-like group, while retaining a chromatin accessibility profile consistent with an activated B-cell/germinal center stage, can differentiate partially towards a PC via *XBP1* and has enhanced NF-κB signaling, upregulated UPR, and excess del(6q). Consistent with their unique biology, we identify notable clinical differences between the two groups, with the MBC-like having higher serum IgM levels, more splenic enlargement, less adenopathy, and elevated LDH, whereas the PC-like has higher platelet levels and more lymph node enlargement.

The potential drivers of the variable B-cell differentiation noted between the subtypes were suggested by a pseudotime differentiation trajectory analysis with correlated gene expression and transcription factor motif enrichment. The differentiation blockage in the MBC-like subtype is notable for its enrichment of the ETS transcription factors *SPI1* and *SPIB*, which repress PC differentiation by downregulating *BLIMP1*.[37,48] Indeed, *SPIB* overexpression in healthy donor B-cells inhibits differentiation into PC, while *SPIB* knockdown in primary WM cells restores some PC differentiation capacity.[37] Interestingly, *SPIB* also interacts with promoters of TLRs and *MYD88*, suggesting a feedback mechanism that could lock MBC-like cells into this blocked differentiation state.[49]

In contrast to the complete differentiation block characterizing the MBC-like, the PC-like cases contain MBC, which heterogeneously express *IRF4* and *BLIMP1* up to the level seen in clonal PC. The most obvious driver of PC differentiation is *XBP1*, a critical regulator of PC differentiation.[50,51] It is notable that these cells have low to moderate expression and no significant motif enrichment of the normally essential PC regulators *IRF4* and *BLIMP1*, consistent with abnormal and incomplete PC differentiation in WM.[52] Supporting this, several AP-1 transcription family members were enriched, including *JUN, FOS, BATF*, and *BATF3*, many of which have previously been identified with pre-plasmablasts and were associated with a PC-like subtype in another multiomic study of WM.[47,53] Another mechanism contributing to this differentiation blockade is *FOXO1*, which is upregulated in PC-like cells. *FOXO1* is known to upregulate *BATF*, which subsequently interacts with *IRF4* to inhibit *BLIMP1* expression, thereby impeding full differentiation into functional plasma cells.[54,55] Furthermore, the expression ratio of *BLIMP1* and *BACH2*, key drivers of the PC and MBC programs, respectively, has been shown to inform MBC fate and, consistent with this, we found these ratios corresponded to the extent of PC differentiation in the WM subtypes.[56] *BLIMP1* and *BACH2* are also located on 6q, but their involvement in del(6q) and potential role in WM pathogenesis requires further examination.

Consistent with their abnormal programming, both WM subtypes exhibit a mixed activated/GC B-cell state compared to healthy donor MBC. This state is supported by motif enrichment of transcription factors such as *NFKB1*, *NFKB2*, *REL*, *TCF3*, *ID3*, and *ID4*, alongside a relative reduction in factors like *RUNX1*, *STAT6*, and *ETV6*.[17] Importantly, each subtype demonstrates unique motif enrichment patterns within this mixed state: MBC-like cells show enrichment of Kruppel-like (KLF) transcription factors, such as *KLF2*, while PC-like cells are enriched for transcription factors associated with GC maintenance, including *OCT1*, *OCT2*, and *FOXO1.*[35] The observation that WM cells harbor a GC-like phenotype which is consistent with mutational signatures observed in WM and suggests an important role for the GC in WM.[57]

Beyond their extent of B-cell differentiation and activation states, the MBC-like and PC-like WM subtypes have distinct gene expression profiles and transcription factor motif enrichment, which underscores their unique biological features. The MBC-like subtype is characterized by relatively heightened BCR signaling and diminished NF-κB signaling, whereas the PC-like subtype exhibits the opposite pattern. In the MBC-like, elevated BCR signaling is marked by the upregulation of BCR-signaling kinases, including LYN, BLNK, and BTK, which are transcriptionally upregulated by *SPI1* and *SPIB*.[58] Additionally, these cells have increased expression of genes associated with BCR function, including *FCHSD2*, *ACTB*, *CD74*, *CD52*, *CR1* (C3b), and *CR2* (CD21).[59-62] The relevance of this signaling phenotype is further supported by GSEA enrichment of the PI3K/AKT/mTOR pathway, which might be driven by mutations in *CXCR4* and elevated BCR signaling activity. In contrast, the PC-like subtype is distinguished by amplified NF-κB signaling, evident through the upregulation of *NFKB1*, *REL*, and *RELB*, alongside GSEA enrichment for the NF-κB pathway. These molecular differences could stem from the mutational landscape within each subtype or be sustained by an inflammatory microenvironment. Notably, prior work has linked del(6q) with inflammation in WM, suggesting that del(6q) may

help perpetuate NF-κB activity. Indeed, genes located in the deleted 6q region, such as *IBTK*, *HIVEP2*, and *TNFAIP3*, are known negative regulators of NF-κB, and their loss likely contributes to the elevated NF-κB signaling and inflammatory response observed in PC-like cells by GSEA.[10] In line with this, we noted the differential expression of *PELI1,* a regulator of TLR signaling. PELI1 can interact with TRAF6, IKK, and NF-κB and mediate pro-inflammatory signaling via IL-1R, MAPK, and PI3K.[63]

Multiple therapeutic implications can be derived from our results. Both subtypes display overexpression of *BCL2*, a gene that supports pro-survival signaling and anti-apoptotic functions and which is a well-established therapeutic target in WM.[64] Additionally, the MBC-like subtype's elevated BCR signaling suggests that therapies targeting BCR signaling could be particularly beneficial in these cases. This may even extend to using alternative approaches to BTK inhibitors, which primarily disrupt NF-κB signaling but may be less effective in MBC-like cases due to their lower NF-κB activity and higher frequency of *CXCR4* mutations, which are known to reduce BTK inhibitor efficacy.[9,14] Functional validation of the role of BCR signaling in CXCR4-mutated WM cases will help to solidify these observations. Furthermore, iff it can be established that subtypes can be routinely identified by surface markers or microarrays, the results of this study suggest that each subtype may be targetable by distinct treatment options, opening the door for more personalized treatment options in WM.

Although this study provides a robust classification of WM subtypes, there are several potential limitations. The sample size is relatively small, so it was important to validate the results in a larger series. We were able to do this using hierarchical clustering analysis of 27 additional cases, but additional data may increase the resolution of subtype variability. Furthermore, one of the cases, WM43, did not cluster near either the MBC-like or PC-like samples and had a unique marker profile. This case minimally expressed MBC and PC genes but differentially upregulated a blend of subtype-specific genes. Interestingly, many of the upregulated genes were associated with neuronal function (*ZNF804A*, *BDNF*, *NRG1*, *ASTN2*) and cell adhesion (*CADM2*, *SDK1*, *CNTN5*, *PCDH9*). Additionally, the multiomic approach, while comprehensive, could be impacted by technical limitations, such as capture efficiency variations between the scATAC and scRNA-seq data. A previous epigenetic analysis of WM, which also identified a MBC-like and PC-like subset, identified similar transcription factors as relevant—including the upregulation of *REL* and *JUN* in PC-like cases—while also notably observing enrichment of *SPIB* and *BCL11A* in PC-like, not MBC-like cases[47]. This highlights the need for functional validation of these results, especially the role of the PC-inhibitors *SPI1/SPIB* in differentiation block in WM. Finally, while the enrichment of CXCR4 mutations in MBC-like cases and del6q in PC-like cases is now documented in multiple studies, the role of these aberrations in driving disease pathology and biology remains unclear.[16,47]

Overall, our data is consistent with a model whereby an immortalized clonal initiating cell is driven to proliferate and mature via MYD88-induced NF-κB signaling, setting in place an abnormal cellular program conferring some of the key hallmarks of WM. The fate of these initiating cells is further shaped after acquiring either a CXCR4 mutation or del(6q) in the GC, pushing the clonal initiating cells towards either a differentiation blocked MBC stage or a partially differentiated PC stage, respectively. The differences and commonalities of these two groups provide a model for understanding in more detail the mutational and epigenetic patterns which underlie treatment responses and disease progression in WM.

Acknowledgements
The Morgan Lab was supported by grants from the International Waldenstrom's Macroglobulinemia Foundation (IWMF), the Myeloma Solutions Fund, the Paula & Rodger Riney Foundation, the NYU Langone Health's Genome Technology Center, and the Cancer Center Support Grant at the Laura and Isaac Perlmutter Cancer Center. The Bernard Lab was supported by grants from the Foundation pour la Recherche Médicale. We acknowledge support from Fondation ARC and Ligue Nationale contre le cancer (to DRW) and thank the financial support of INCa-DGOS-INSERM-ITMO Cancer (SiRIC Epicure) and INCa-DGOS-INSERM (SiRIC CURAMUS) and Paris Île-de-France Region – DIM "ImmunoTherapies, Auto-immunité et Cancer". MD is a recipient from fellowships from INSERM-ITMO Cancer and Ligue Nationale contre le cancer, AJ from université Paris Saclay and the Fundation pour la Recherche Médicale, and MA from the Philantropia Foundation. We also thank P. Meyn and Y. Zhang from the NYU Langone Genome Technology Center for their

invaluable help with scRNA-seq/scATAC-seq library design and preparation, and C. Quivoron from Gustave Roussy for help in sample processing.

Authorship Contributions

GJM and OAB supervised the research and the writing of the paper. GJM, OAG, DCG, and HG conceived, designed, and developed the methodology of the study. DCG performed research, bioinformatic analysis, and figure generation and participated in the writing of the paper. HG and DZ performed laboratory experiments and research, generated data, and participated in the writing of the paper. GJM, DCG, HG, PB, AE, GV, YW, SL, BW, and FED participated in the analysis and interpretation of data. DRW, PM, FNK, OL, LW, DK, MB, VL, FNK, DK, SL, ME, MD, AJ, MA, JL, AT, and FED contributed reagents, samples, and data. OAB contributed data, performed research, and contributed to the writing of the paper. All authors reviewed and edited the manuscript.

Disclosure of Conflict of Interest

The authors have no conflicts of interest to disclose.

References

1. Alaggio R, Amador C, Anagnostopoulos I, et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms. *Leukemia*. 2022;36(7):1720-1748.
2. San Miguel J, Vidriales M, Ocio E, et al. Immunophenotypic analysis of Waldenstrom's macroglobulinemia. Seminars in oncology: Elsevier; 2003:187-195.
3. Treon SP, Xu L, Yang G, et al. MYD88 L265P Somatic Mutation in Waldenström's Macroglobulinemia. *New England Journal of Medicine*. 2012;367(9):826-833.
4. Deguine J, Barton GM. MyD88: a central player in innate immune signaling. *F1000Prime Reports*. 2014;6.
5. Fitzgerald KA, Kagan JC. Toll-like Receptors and the Control of Immunity. *Cell*. 2020;180(6):1044-1066.
6. Munshi M, Liu X, Kofides A, et al. A new role for the SRC family kinase HCK as a driver of SYK activation in MYD88 mutated lymphomas. *Blood Advances*. 2022;6(11):3332-3338.
7. Munshi M, Liu X, Chen JG, et al. SYK is activated by mutated MYD88 and drives pro-survival signaling in MYD88 driven B-cell lymphomas. *Blood Cancer Journal*. 2020;10(1).
8. Hunter ZR, Xu L, Yang G, et al. The genomic landscape of Waldenström macroglobulinemia is characterized by highly recurring MYD88 and WHIM-like CXCR4 mutations, and small somatic deletions associated with B-cell lymphomagenesis. *Blood*. 2014;123(11):1637-1646.
9. Treon SP, Cao Y, Xu L, Yang G, Liu X, Hunter ZR. Somatic mutations in MYD88 and CXCR4 are determinants of clinical presentation and overall survival in Waldenström macroglobulinemia. *Blood*. 2014;123(18):2791-2796.
10. Forgeard N, Baron M, Caron J, et al. Inflammation in Waldenström macroglobulinemia is associated with 6q deletion and need for treatment initiation. *Haematologica*. 2022.
11. Schop RFJ, Kuehl WM, Van Wier SA, et al. Waldenström macroglobulinemia neoplastic cells lack immunoglobulin heavy chain locus translocations but have frequent 6q deletions. *Blood*. 2002;100(8):2996-3001.
12. Nguyen-Khac F, Lambert J, Chapiro E, et al. Chromosomal aberrations and their prognostic value in a series of 174 untreated patients with Waldenström's macroglobulinemia. *haematologica*. 2013;98(4):649-654.
13. Guerrera ML, Tsakmaklis N, Xu L, et al. MYD88 mutated and wild-type Waldenström's Macroglobulinemia: characterization of chromosome 6q gene losses and their mutual exclusivity with mutations in CXCR4. *haematologica*. 2018;103(9):e408.
14. Treon SP, Tripsas CK, Meid K, et al. Ibrutinib in previously treated Waldenstrom's Macroglobulinemia. *New England Journal of Medicine*. 2015;372(15):1430-1440.
15. Dimopoulos MA, Opat S, D'Sa S, et al. Zanubrutinib Versus Ibrutinib in Symptomatic Waldenstrom Macroglobulinemia: Final Analysis From the Randomized Phase III ASPEN Study. *Journal of Clinical Oncology*. 2023;41(33):5099-5106.
16. Roos-Weil D, Giacopelli B, Armand M, et al. Identification of two DNA methylation subtypes of Waldenström's macroglobulinemia with plasma and memory B cell features. *Blood*. 2020.
17. King HW, Wells KL, Shipony Z, et al. Integrated single-cell transcriptomics and epigenomics reveals strong germinal center–associated etiology of autoimmune risk loci. *Science immunology*. 2021;6(64):eabh3768.
18. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021;184(13):3573-3587. e3529.
19. Domínguez Conde C, Xu C, Jarvis L, et al. Cross-tissue immune cell analysis reveals tissue-specific features in humans. *Science*. 2022;376(6594):eabl5197.
20. Xu C, Prete M, Webb S, et al. Automatic cell-type harmonization and integration across Human Cell Atlas datasets. *Cell*. 2023;186(26):5876-5891. e5820.
21. Song L, Cohen D, Ouyang Z, Cao Y, Hu X, Liu XS. TRUST4: immune repertoire reconstruction from bulk and single-cell RNA-seq data. *Nature methods*. 2021;18(6):627-630.
22. Granja JM, Corces MR, Pierce SE, et al. ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nature Genetics*. 2021.

23. Korsunsky I, Millard N, Fan J, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nature Methods*. 2019;16(12):1289-1296.
24. Schep AN, Wu B, Buenrostro JD, Greenleaf WJ. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nature Methods*. 2017;14(10):975-978.
25. Cao J, Spielmann M, Qiu X, et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature*. 2019;566(7745):496-502.
26. Van Keimpema M, Grüneberg LJ, Mokry M, et al. FOXP1 directly represses transcription of proapoptotic genes and cooperates with NF-κB to promote survival of human B cells. *Blood*. 2014;124(23):3431-3440.
27. Agarwal N, Kim CH, Kunkalla K, et al. Smoothened (SMO) is an adaptor protein that recruits TRAF6 and phospholipase C gamma 2 (PLCg2) to enhance the activation of NF-Kb signaling pathway: American Society of Hematology Washington, DC; 2015.
28. Ingham RJ, Holgado-Madruga M, Siu C, Wong AJ, Gold MR. The Gab1 Protein Is a Docking Site for Multiple Proteins Involved in Signaling by the B Cell Antigen Receptor. *Journal of Biological Chemistry*. 1998;273(46):30630-30637.
29. Sun JY, Xu L, Tseng H, et al. Histone deacetylase inhibitors demonstrate significant preclinical activity as single agents, and in combination with bortezomib in Waldenström's macroglobulinemia. *Clinical Lymphoma Myeloma and Leukemia*. 2011;11(1):152-156.
30. Trojani A, Di Camillo B, Bossi LE, et al. Identification of a candidate gene set signature for the risk of progression in IgM MGUS to smoldering/symptomatic Waldenström macroglobulinemia (WM) by a comparative transcriptome analysis of B cells and plasma cells. *Cancers*. 2021;13(8):1837.
31. Ji L, Huo X, Zhang Y, Yan Z, Wang Q, Wen B. TOPORS, a tumor suppressor protein, contributes to the maintenance of higher-order chromatin architecture. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*. 2020;1863(5):194518.
32. Lin D, Ippolito GC, Zong R-T, Bryant J, Koslovsky J, Tucker P. Bright/ARID3A contributes to chromatin accessibility of the immunoglobulin heavy chain enhancer. *Molecular Cancer*. 2007;6(1):23.
33. Debaize L, Troadec M-B. The master regulator FUBP1: its emerging role in normal cell function and malignant development. *Cellular and Molecular Life Sciences*. 2019;76(2):259-281.
34. De Martino M, Nicolau-Neto P, Pinto LFR, et al. HMGA1 induces EZH2 overexpression in human B-cell lymphomas. *American journal of cancer research*. 2021;11(5):2174.
35. Laidlaw BJ, Cyster JG. Transcriptional regulation of memory B cell differentiation. *Nature Reviews Immunology*. 2021;21(4):209-220.
36. Horiuchi S, Koike T, Takebuchi H, et al. SpiB regulates the expression of B-cell-related genes and increases the longevity of memory B cells. *Frontiers in Immunology*. 2023;14.
37. Zhou Y, Liu X, Xu L, et al. Transcriptional repression of plasma cell differentiation is orchestrated by aberrant over-expression of the ETS factor SPIB in Waldenström macroglobulinaemia. *British journal of haematology*. 2014;166(5):677-689.
38. Roos-Weil D, Decaudin C, Armand M, et al. A Recurrent Activating Missense Mutation in Waldenström Macroglobulinemia Affects the DNA Binding of the ETS Transcription Factor SPI1 and Enhances Proliferation. *Cancer Discovery*. 2019;9(6):796-811.
39. Shaffer AL, Shapiro-Shelef M, Iwakoshi NN, et al. XBP1, Downstream of Blimp-1, Expands the Secretory Apparatus and Other Organelles, and Increases Protein Synthesis in Plasma Cell Differentiation. *Immunity*. 2004;21(1):81-93.
40. Gass JN, Gunn KE, Sriburi R, Brewer JW. Stressed-out B cells? Plasma-cell differentiation and the unfolded protein response. *Trends in immunology*. 2004;25(1):17-24.
41. Gilmore TD. The Rel/NF-κB signal transduction pathway: introduction. *Oncogene*. 1999;18(49):6842-6844.
42. Vetter G, Le Béchec A, Muller J, et al. Time-resolved analysis of transcriptional events during SNAI1-triggered epithelial to mesenchymal transition. *Biochemical and biophysical research communications*. 2009;385(4):485-491.

43. Zhu Q-Q, Ma C, Wang Q, Song Y, Lv T. The role of TWIST1 in epithelial-mesenchymal transition and cancers. *Tumor Biology*. 2016;37:185-197.
44. Sun H, Fang T, Wang T, et al. Single-cell profiles reveal tumor cell heterogeneity and immunosuppressive microenvironment in Waldenström macroglobulinemia. *Journal of Translational Medicine*. 2022;20(1):576.
45. Lim JH, Wang JQ, Webb F, et al. Plasma cells arise from differentiation of clonal lymphocytes and secrete IgM in Waldenström macroglobulinemia. *Iscience*. 2022;25(8).
46. Bagratuni T, Aktypi F, Theologi O, et al. Single-cell analysis of MYD88L265P and MYD88WT Waldenström macroglobulinemia patients. *HemaSphere*. 2024;8(2):e27.
47. Hunter ZR, Tsakmaklis N, Richardson K, et al. Changes in Methylation and Chromatin Accessibility Underlie Subtype Classification and Disease Evolution in Waldenström's Macroglobulinemia. *Blood*. 2023;142:755.
48. Carotta S, Willis SN, Hasbold J, et al. The transcription factors IRF8 and PU.1 negatively regulate plasma cell differentiation. *Journal of Experimental Medicine*. 2014;211(11):2169-2181.
49. Care MA, Cocco M, Laye JP, et al. SPIB and BATF provide alternate determinants of IRF4 occupancy in diffuse large B-cell lymphoma linked to disease heterogeneity. *Nucleic acids research*. 2014;42(12):7591-7610.
50. Reimold AM, Iwakoshi NN, Manis J, et al. Plasma cell differentiation requires the transcription factor XBP-1. *Nature*. 2001;412(6844):300-307.
51. Le TA, Lino AC, Schrezenmeier E, et al. Efficient CRISPR-Cas9-mediated mutagenesis in primary human B cells for identifying plasma cell regulators. *Molecular Therapy-Nucleic Acids*. 2022;30:621-632.
52. Trezise S, Nutt SL. The gene regulatory network controlling plasma cell function. *Immunological Reviews*. 2021;303(1):23-34.
53. Alaterre E, Ovejero S, Bret C, et al. Integrative single-cell chromatin and transcriptome analysis of human plasma cell differentiation. *Blood*. 2024;144(5):496-509.
54. Ochiai K, Maienschein-Cline M, Simonetti G, et al. Transcriptional regulation of germinal center B and plasma cell fates by dynamical control of IRF4. *Immunity*. 2013;38(5):918-929.
55. Scharer CD, Barwick BG, Guo M, Bally APR, Boss JM. Plasma cell differentiation is controlled by multiple cell division-coupled epigenetic programs. *Nature Communications*. 2018;9(1).
56. Shao W, Wang Y, Fang Q, Shi W, Qi H. Epigenetic recording of stimulation history reveals BLIMP1–BACH2 balance in determining memory B cell fate upon recall challenge. *Nature Immunology*. 2024;25(8):1432-1444.
57. Maclachlan KH, Bagratuni T, Kastritis E, et al. Waldenström macroglobulinemia whole genome reveals prolonged germinal center activity and late copy number aberrations. *Blood Advances*. 2023;7(6):971-981.
58. Willis SN, Tellier J, Liao Y, et al. Environmental sensing by mature B cells is controlled by the transcription factors PU. 1 and SpiB. *Nature communications*. 2017;8(1):1426.
59. Roberts AD, Davenport TM, Dickey AM, Ahn R, Sochacki KA, Taraska JW. Structurally distinct endocytic pathways for B cell receptors in B lymphocytes. *Molecular Biology of the Cell*. 2020;31(25):2826-2840.
60. Su H, Na N, Zhang X, Zhao Y. The biological function and significance of CD74 in immune diseases. *Inflammation Research*. 2017;66:209-216.
61. Erdei A, Kovács KG, Nagy-Baló Z, et al. New aspects in the regulation of human B cell functions by complement receptors CR1, CR2, CR3 and CR4. *Immunology Letters*. 2021;237:42-57.
62. Bhamidipati K, Silberstein JL, Chaichian Y, et al. CD52 is elevated on B cells of SLE patients and regulates B cell function. *Frontiers in immunology*. 2021;11:626820.
63. Choe J-Y, Park M, Yun JY, et al. PELI1 expression is correlated with MYC and BCL6 expression and associated with poor prognosis in diffuse large B-cell lymphoma. *Modern Pathology*. 2016;29(11):1313-1323.
64. Castillo JJ, Allan JN, Siddiqi T, et al. Venetoclax in previously treated Waldenström macroglobulinemia. *Journal of Clinical Oncology*. 2022;40(1):63-71.

Tables

Table 1
Clinical Characteristics at Diagnosis

	MBC-like	*PC-like*	*p-value*
Number of patients	10	24	
Female sex, n (%)	4 (40%)	7 (29%)	0.451
Age (median, range)	65 (52-82)	69 (40-80)	0.663
BM Infiltration (median %, range)	54 (23-88)	26.5 (11-75)	0.151
Spleen Enlargement (% patients)	37.5	5.6	0.072
Lymph Node Enlargement (% patients)	25	27.8	1
IgM (median g/dL, range)	1.6 (0.83-5.3)	1.2 (0.49-5.75)	0.228
Hemoglobin (median g/dL, range)	10.9 (7.3-13.8)	12.4 (8.5-14.5)	0.061
Platelets (median 10^3/L, range)	237.5 (50-365)	273 (95-424)	0.279
Elevated LDH (% patients)	30	8.7	0.28
Overall Survival (median, in months)	84.5	104	0.281

Table 1. Clinical characteristics of WM patients at diagnosis, by subtype. Median values are shown, unless referring to reference ranges. BM infiltration is equivalent with disease involvement and elevated LDH refers to LDH values greater than UNL. P-values for BM infiltration, IgM, hemoglobin, platelets, age, and overall survival derived from two-sample T-tests and p-values for spleen enlargement, lymph node enlargement, sex, and elevated LDH derived from Fisher's exact tests.

Figure Legends

Figure 1. Identifying cell populations in the bone marrow of WM patients using single-cell RNA-seq and single-cell ATAC-seq. A. Cartoon overview of the workflow involved in this study. Bone marrow aspirates were taken from 13 patients with Waldenstrom's Macroglobulinemia, flow-sorted to select mature B cells based on $CD19^+/CD3^-$, and underwent 10x Multiome sequencing to generate RNA-seq and ATAC-seq information. **B.** UMAP showing annotated cells in the final data object after quality control, processing, and integration of the scRNA-seq and scATAC-seq datasets (n cells = 129,372) cells). **C.** Stacked bar plot showing relative abundance of annotated cell populations across the 13 WM patients (01-076, 01-115, 01131, 01-163, 01-190, 04-003, 04-006, WM25, WM43, WM46, WM47, WM54, WM65) and seven reference samples used in this study. **D.** Box plots showing normalized gene expression of plasma cell and memory B-cell marker genes across the annotated cell types. **E.** UMAPs showing patient annotations (left) and cells deemed to be monoclonal as a result of shared CDR3 amino acid sequences within a given patient, as per TRUST4 (right). Cells which did not have the same CDR3 amino acid sequence as the dominant clone were left uncolored. **F.** UMAPs showing normalized gene expression of plasma cell and memory B-cell marker genes.

Figure 2. Gene expression profiles of WM subtypes are distinct from reference memory B-cells. A. UMAP of reclustered MBC-like MBC, polyclonal MBC, and reference MBC. **B.** Volcano plot of differentially expressed genes between the MBC-like MBC and reference MBC. Colored points represent differential genes deemed to be statistically significant with absolute $\log_2FC > 0.5$ and FDR < $5e^{-9}$. **C.** Volcano plot of differentially enriched transcription factor motifs between the MBC-like and reference MBC. Colored points represent differential motifs deemed to be statistically significant with absolute MeanDiff > 0.025 and FDR < $5e^{-9}$. **D**. UMAP of reclustered PC-like MBC and reference MBC. **E.** Volcano plot of differentially expressed genes between the PC-like MBC and reference MBC. Colored points represent differential genes deemed to be statistically significant with absolute $\log_2FC > 0.5$ and FDR < $5e^{-9}$. **F.** Volcano plot of differentially enriched transcription factor motifs between the PC-like and reference MBC. Colored points represent differential motifs

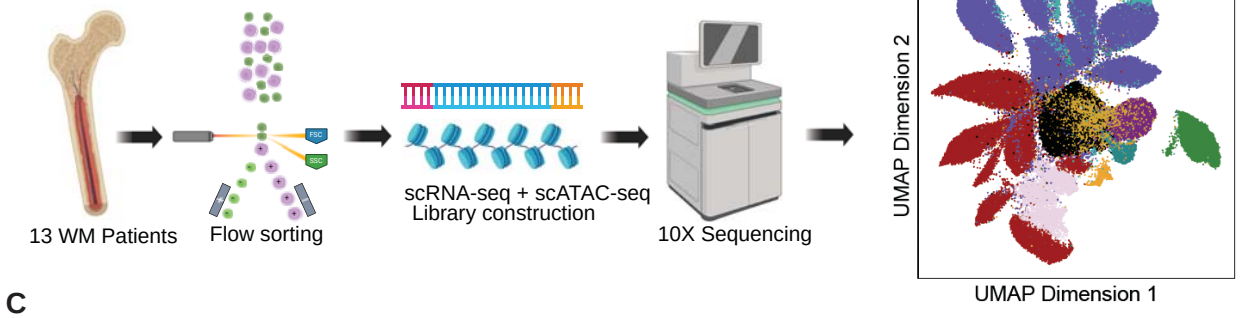
deemed to be statistically significant with absolute MeanDiff > 0.025 and FDR $5e^{-9}$. **G.** Venn diagram showing the number of shared and unique genes upregulated by the two subtypes compared to the reference MBC. Only includes statistically significant genes with absolute $log_2FC > 0.5$ and FDR < 0.05. **H.** Venn diagram showing the number of shared and unique motifs enriched in the two subtypes compared to the reference MBC. Only includes statistically significant genes with absolute MeanDiff > 0.025 and FDR < $5e^{-9}$.

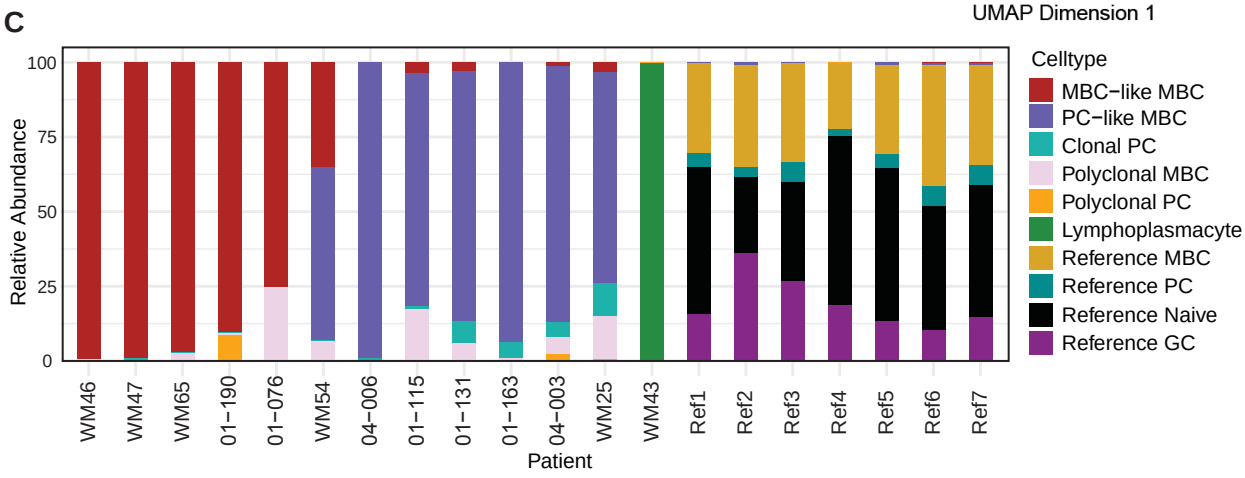
Figure 3. WM subtypes exhibit distinct disease phenotypes. A. Volcano plot showing genes differentially expressed between the MBC-like MBC and PC-like MBC. Colored points represent statistically significant differential genes with $log_2FC > 0.5$ and FDR < $5e^{-9}$. **B.** UMAPs showing gene expression of select genes which exhibited significant statistical expression in each the subtypes. The first 2 rows are genes uniquely expressed in each subtype compared to each other and the third row reflects genes differentially expressed by both subtypes compared to the reference MBC. **C.** Volcano plot showing single-cell transcription factor motifs differentially enriched between the MBC-like MBC and PC-like MBC. Colored points represent significantly enriched motifs with MeanDiff > 0.025 and FDR < $5e^{-}9$. **D.** UMAPs showing single-cell transcription factor motif enrichment of select motifs which exhibited significant statistical enrichment in each of the subtypes. The first 2 rows are motifs uniquely enriched in each subtype compared to each other and the third row reflects motifs differentially enriched by both subtypes compared to the reference MBC. **E.** Dot plot showing pathways enriched in the MBC-like MBC via pre-ranked Gene Set Enrichment Analysis (GSEA). Points are colored by FDR q-value. Pathways with < 0.25 FDR q-values are considered statistically significant. It is noteworthy that the androgen response, mTORC1 signaling, and IL2/STAT5 signaling pathways were nearly significant with FDR q-values < 0.26, and **F.** Dot plot showing pathways enriched in the PC-like MBC via pre-ranked GSEA. Points are colored by FDR q-value, with values < 0.25 considered to be significant. All shown pathways met this criteria.

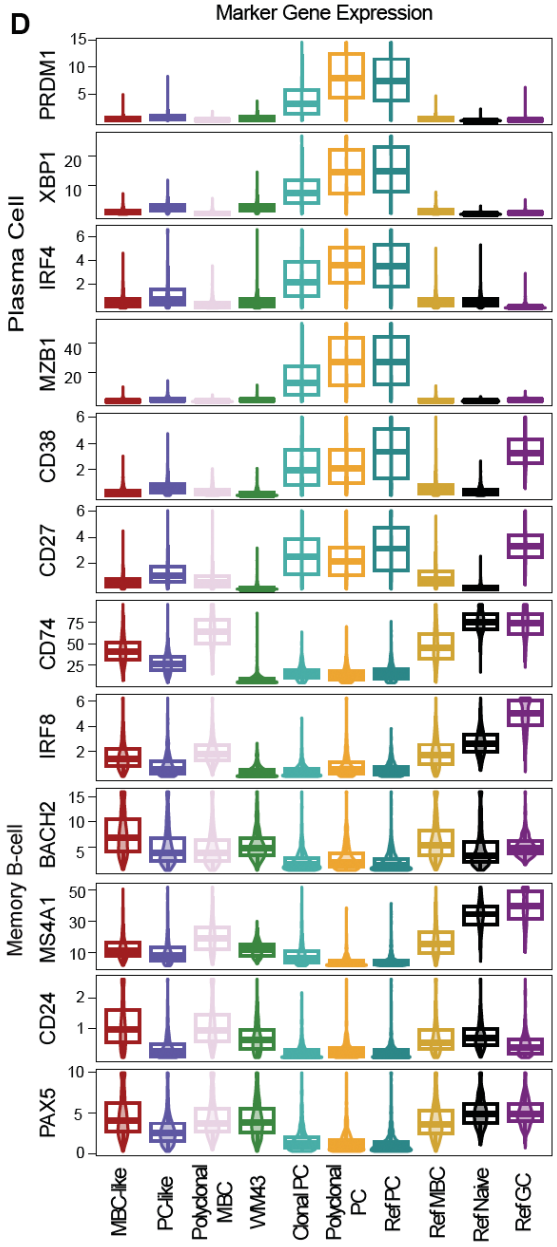
Figure 4. **Differentiation blockade affects WM subtypes. A.** UMAP overlain with a differentiation trajectory created in ArchR, using the polyclonal MBC, MBC-like MBC, PC-like MBC, and clonal PC as the trajectory backbone. **B.** UMAP overlain with cell annotations for reference. **C.** Heatmap of genes associated with pseudotime in both the gene expression and motif enrichment data. Heatmap features were identified by correlating the gene expression and motif deviation matrices in ArchR and this heatmap shows gene expression of these features. **D.** Gene expression (y-axis) as a function of pseudotime (x-axis) for select genes. **E.** Heatmap of genes correlated with pseudotime in both the gene expression and motif deviation data. Heatmap features were identified by correlating the gene expression and motif deviation matrices in ArchR and this heatmap shows motif deviations of these features. **F.** Motif deviation (y-axis) as a function of pseudotime (x-axis) for select genes.

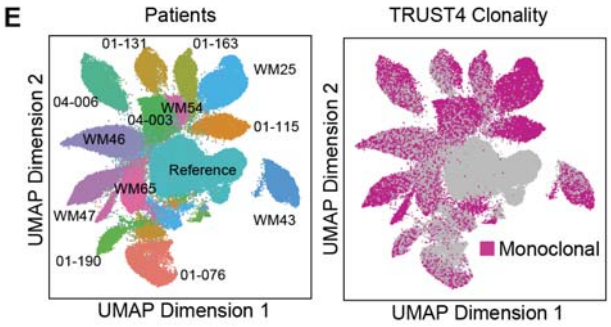
Figure 5. Bulk RNA-seq and whole-genome sequencing analysis of WM patients. A. Gene expression heatmap of 37 WM patients with bulk RNA-seq data based on subtype-specific gene expression derived from the single-cell data. **B.** Mutated gene frequencies of known canonically mutated WM genes in MBC-like subtype patients (n = 10) and the PC-like (n = 22) patients, annotated by WM subtype and the presence of 6q deletions. **C.** Aggregate relative frequency of CNVs across all patients. H1-4 = HIST1H1E; MAP3K14 = NIK.

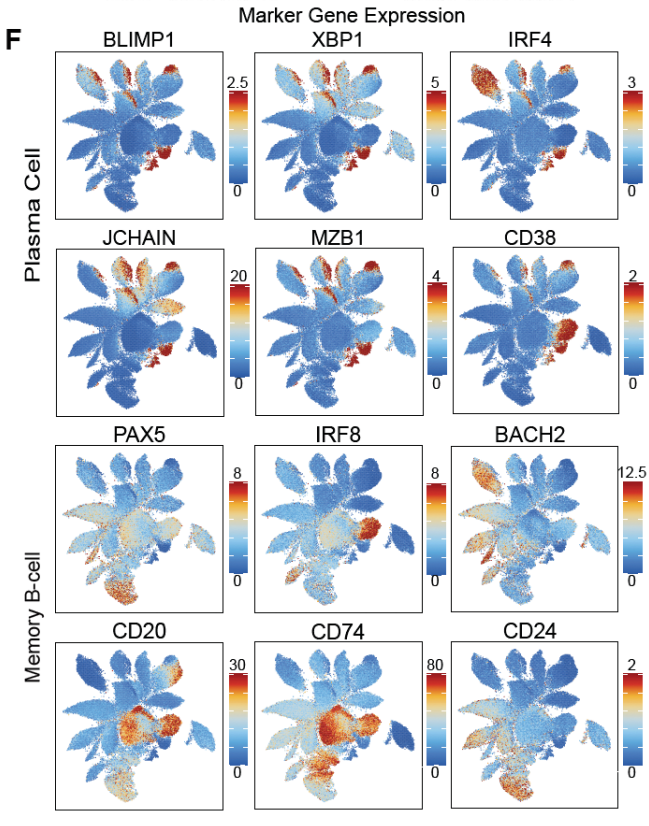
Figure 1

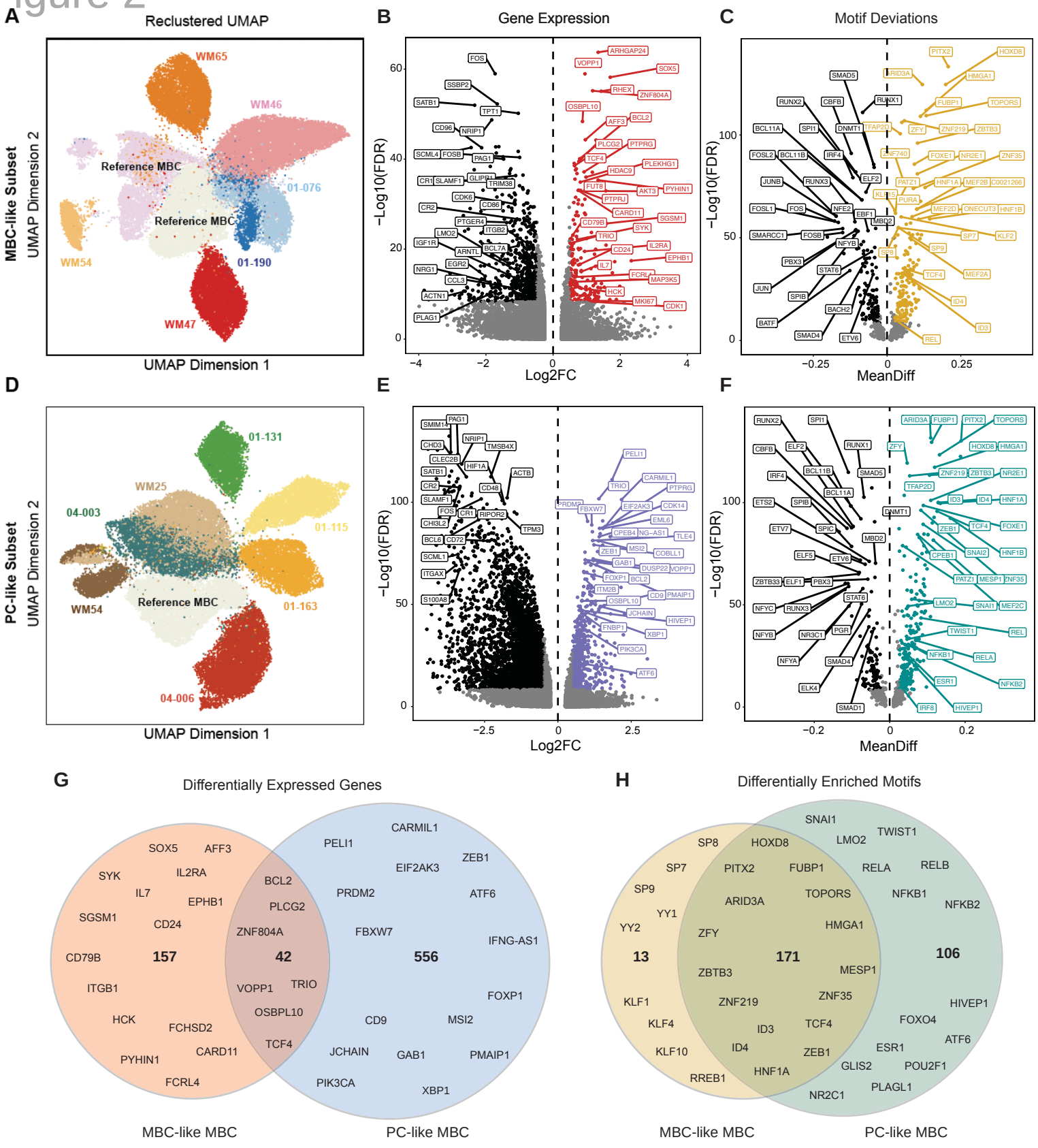











A
Reclustered UMAP
WM65
WM46
Reference MBC
01-076
Reference MBC
WM54
01-190
WM47
MBC-like Subset
UMAP Dimension 2
UMAP Dimension 1
B
Gene Expression
−Log10(FDR)
Log2FC
C
Motif Deviations
−Log10(FDR)
MeanDiff
D
01-131
WM25
04-003
01-115
WM54
Reference MBC
01-163
04-006
PC-like Subset
UMAP Dimension 2
UMAP Dimension 1
E
−Log10(FDR)
Log2FC
F
−Log10(FDR)
MeanDiff
G
Differentially Expressed Genes
157
42
556
MBC-like MBC
PC-like MBC
H
Differentially Enriched Motifs
13
171
106
MBC-like MBC
PC-like MBC

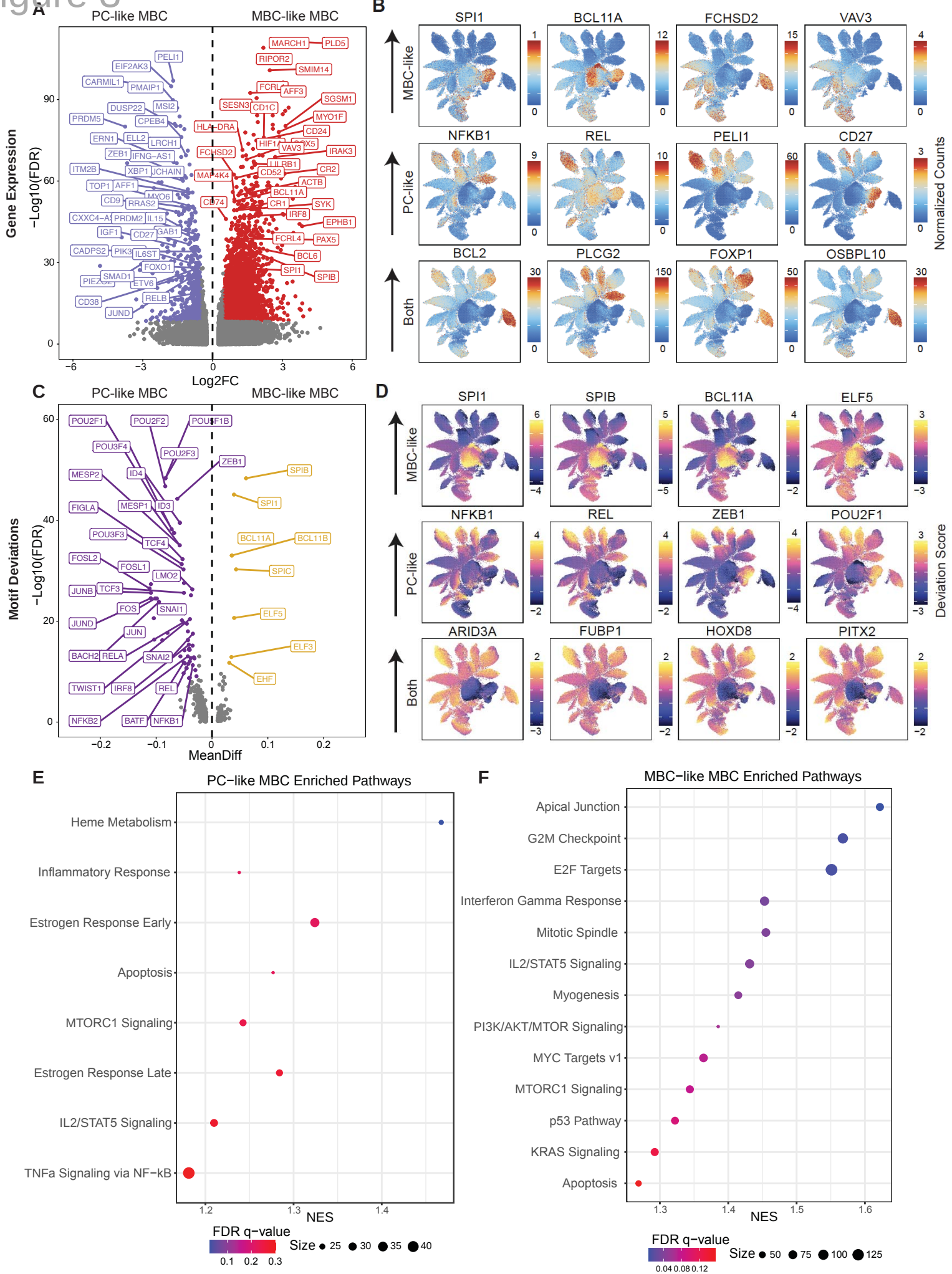
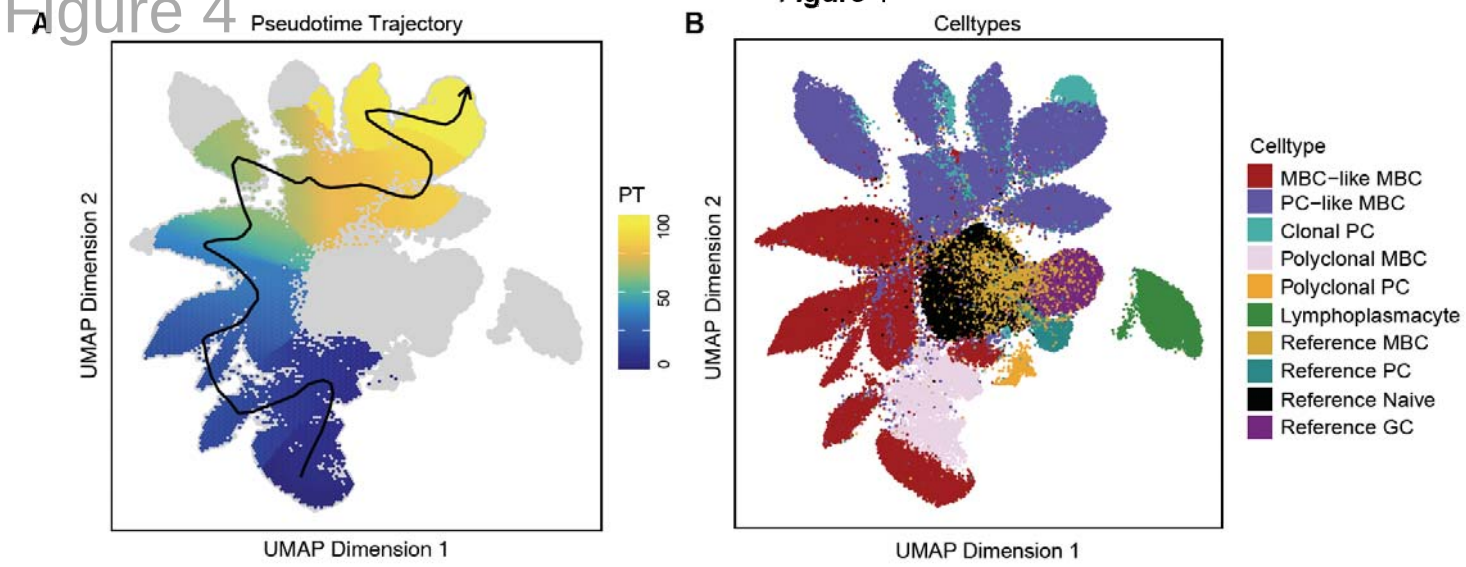
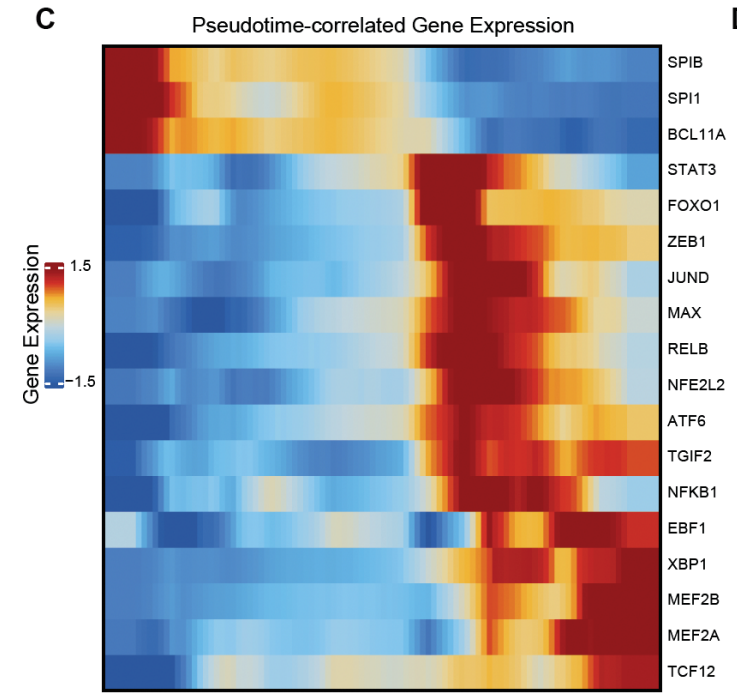
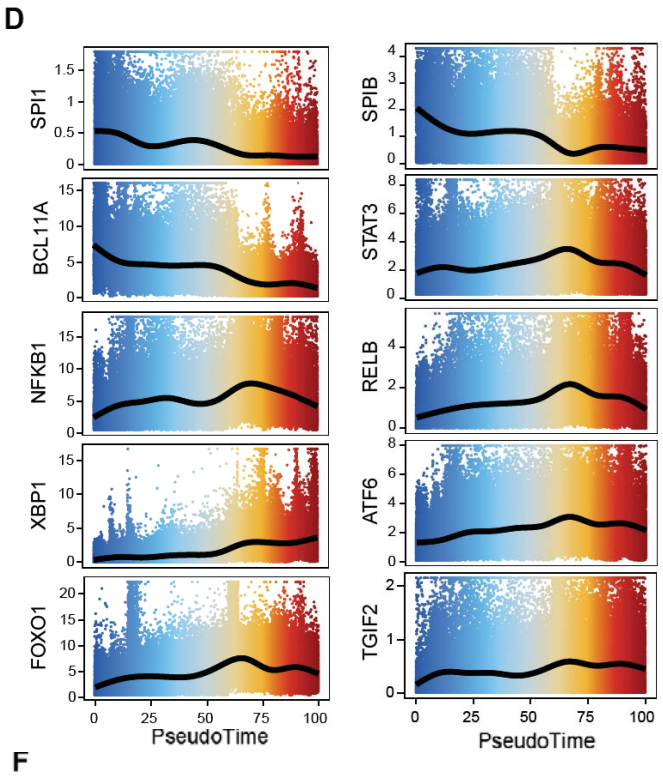


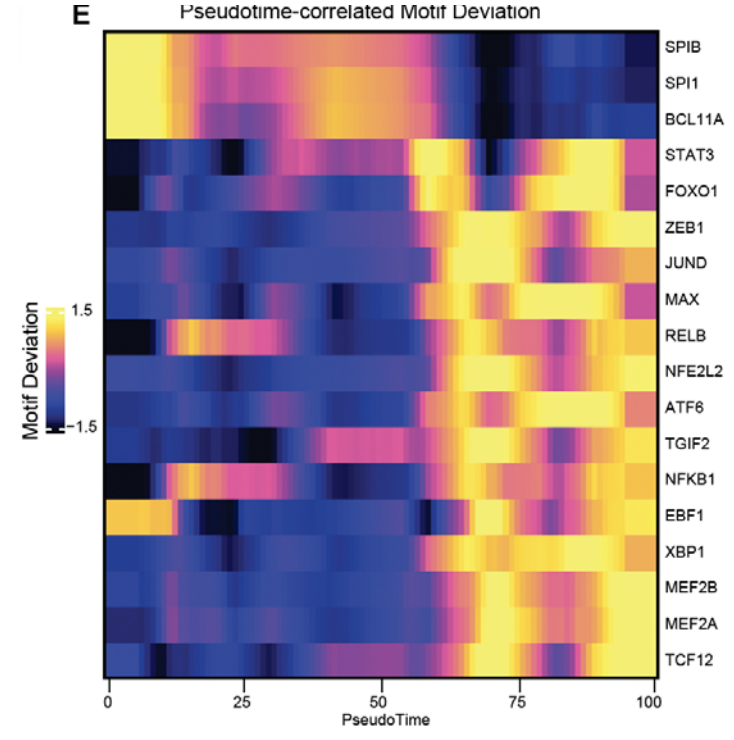
Figure 3
A
PC-like MBC
MBC-like MBC
Gene Expression
−Log10(FDR)
Log2FC
B
SPI1
BCL11A
FCHSD2
VAV3
NFKB1
REL
PELI1
CD27
BCL2
PLCG2
FOXP1
OSBPL10
MBC-like
PC-like
Both
Normalized Counts
C
PC-like MBC
MBC-like MBC
Motif Deviations
−Log10(FDR)
MeanDiff
D
SPI1
SPIB
BCL11A
ELF5
NFKB1
REL
ZEB1
POU2F1
ARID3A
FUBP1
HOXD8
PITX2
MBC-like
PC-like
Both
Deviation Score
E
PC−like MBC Enriched Pathways
Heme Metabolism
Inflammatory Response
Estrogen Response Early
Apoptosis
MTORC1 Signaling
Estrogen Response Late
IL2/STAT5 Signaling
TNFa Signaling via NF−kB
NES
FDR q−value
Size
F
MBC−like MBC Enriched Pathways
Apical Junction
G2M Checkpoint
E2F Targets
Interferon Gamma Response
Mitotic Spindle
IL2/STAT5 Signaling
Myogenesis
PI3K/AKT/MTOR Signaling
MYC Targets v1
MTORC1 Signaling
p53 Pathway
KRAS Signaling
Apoptosis
NES
FDR q−value
Size

Figure 4


A
Pseudotime Trajectory
UMAP Dimension 2
UMAP Dimension 1
PT
100
50
0
B
Celltypes
UMAP Dimension 2
UMAP Dimension 1
Celltype
MBC−like MBC
PC−like MBC
Clonal PC
Polyclonal MBC
Polyclonal PC
Lymphoplasmacyte
Reference MBC
Reference PC
Reference Naive
Reference GC


C
Pseudotime-correlated Gene Expression
Gene Expression
1.5
−1.5
SPIB
SPI1
BCL11A
STAT3
FOXO1
ZEB1
JUND
MAX
RELB
NFE2L2
ATF6
TGIF2
NFKB1
EBF1
XBP1
MEF2B
MEF2A
TCF12


D
SPI1
BCL11A
NFKB1
XBP1
FOXO1
SPIB
STAT3
RELB
ATF6
TGIF2
0
25
50
75
100
PseudoTime


E
Pseudotime-correlated Motif Deviation
Motif Deviation
1.5
−1.5
SPIB
SPI1
BCL11A
STAT3
FOXO1
ZEB1
JUND
MAX
RELB
NFE2L2
ATF6
TGIF2
NFKB1
EBF1
XBP1
MEF2B
MEF2A
TCF12
0
25
50
75
100
PseudoTime

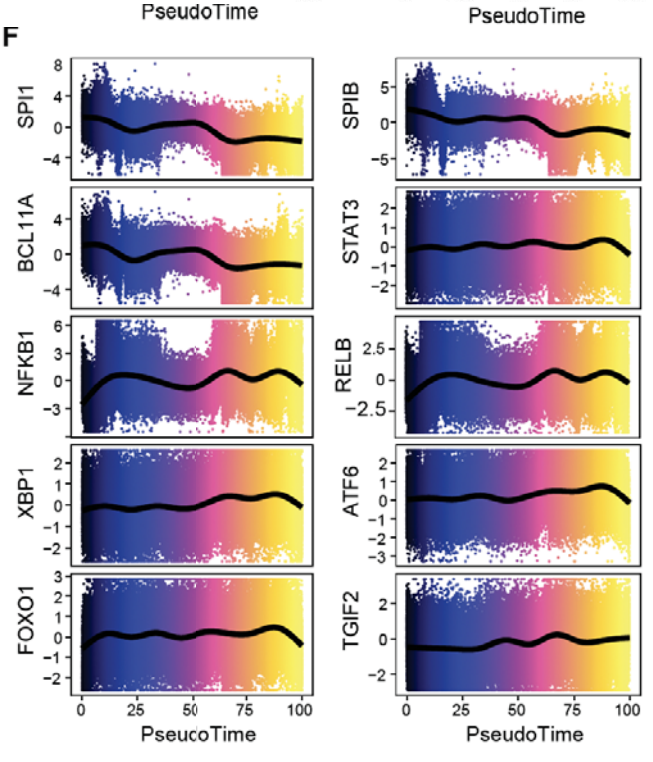
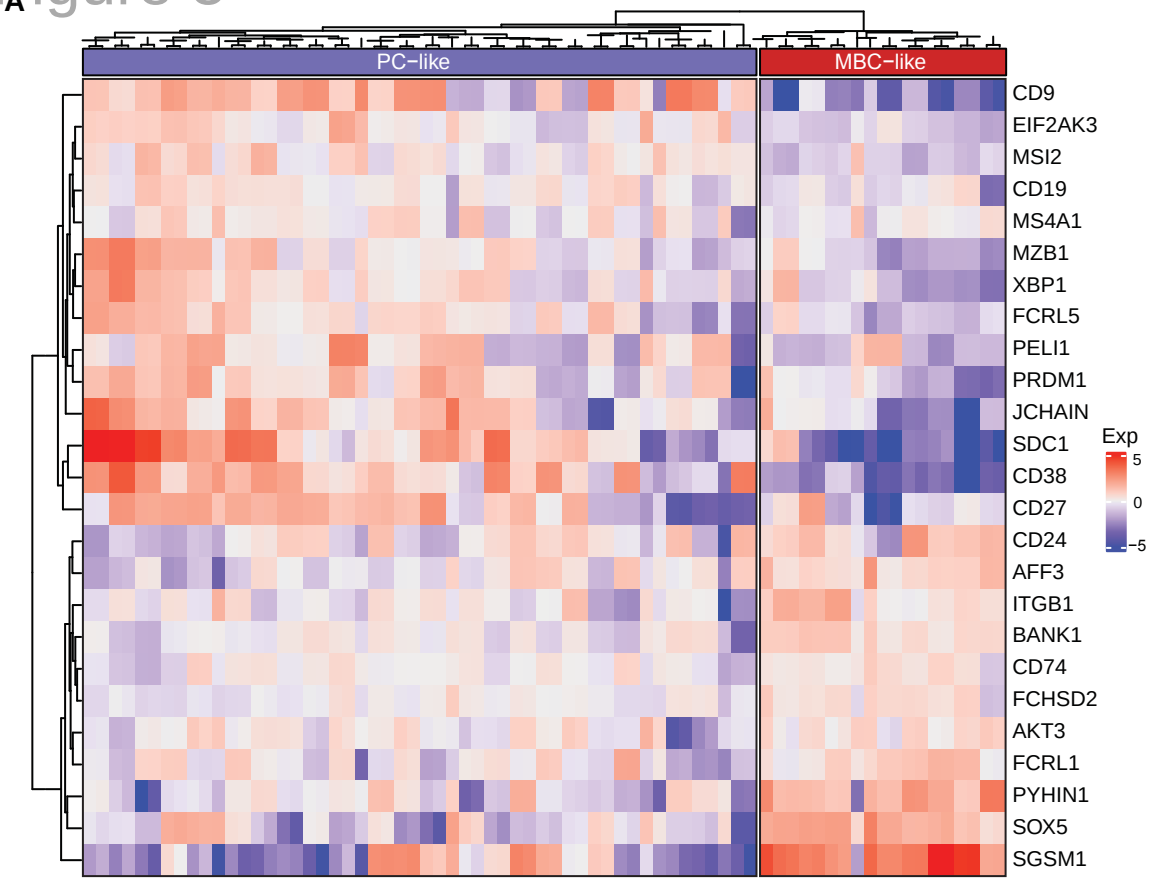
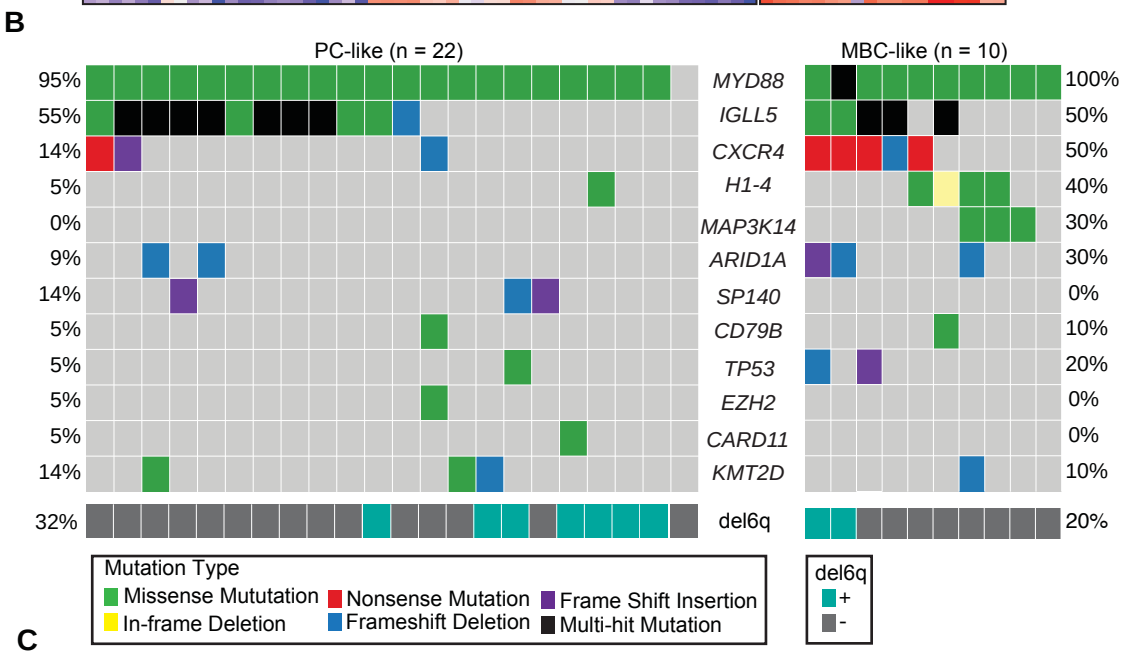
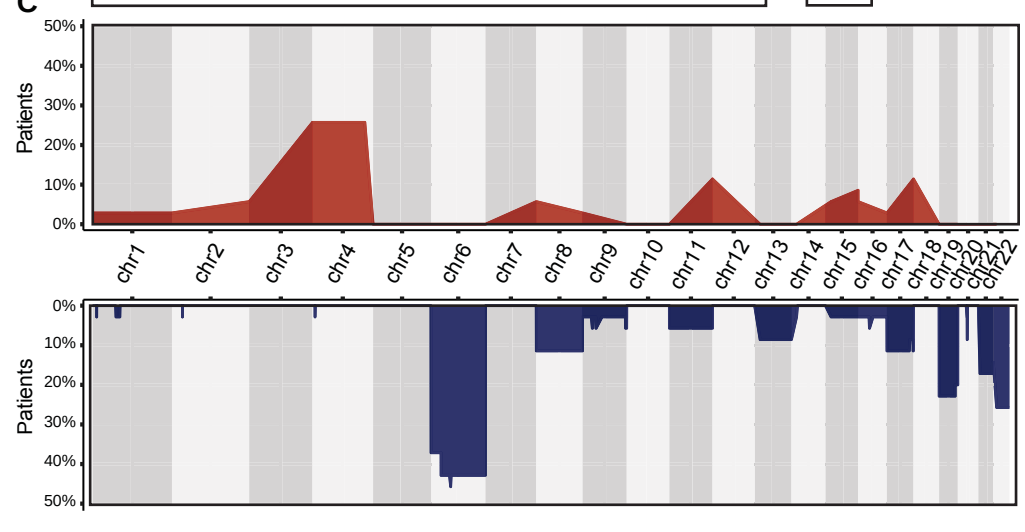

F
SPI1
BCL11A
NFKB1
XBP1
FOXO1
SPIB
STAT3
RELB
ATF6
TGIF2
0
25
50
75
100
PseucoTime
PseudoTime

Figure 5


A
PC−like
MBC−like
CD9
EIF2AK3
MSI2
CD19
MS4A1
MZB1
XBP1
FCRL5
PELI1
PRDM1
JCHAIN
SDC1
CD38
CD27
CD24
AFF3
ITGB1
BANK1
CD74
FCHSD2
AKT3
FCRL1
PYHIN1
SOX5
SGSM1
Exp
5
0
−5


B
PC-like (n = 22)
MBC-like (n = 10)
95%
55%
14%
5%
0%
9%
14%
5%
5%
5%
5%
14%
32%
MYD88
IGLL5
CXCR4
H1-4
MAP3K14
ARID1A
SP140
CD79B
TP53
EZH2
CARD11
KMT2D
del6q
100%
50%
50%
40%
30%
30%
0%
10%
20%
0%
0%
10%
20%
Mutation Type
Missense Mututation
Nonsense Mutation
Frame Shift Insertion
In-frame Deletion
Frameshift Deletion
Multi-hit Mutation
del6q
+
-


C
Patients
50%
40%
30%
20%
10%
0%
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
chr15
chr16
chr17
chr18
chr19
chr20
chr21
chr22
Patients
0%
10%
20%
30%
40%
50%

Single-Cell Multiomic Analysis for Identifying Distinct Subtypes of Waldenström Macroglobulinemia

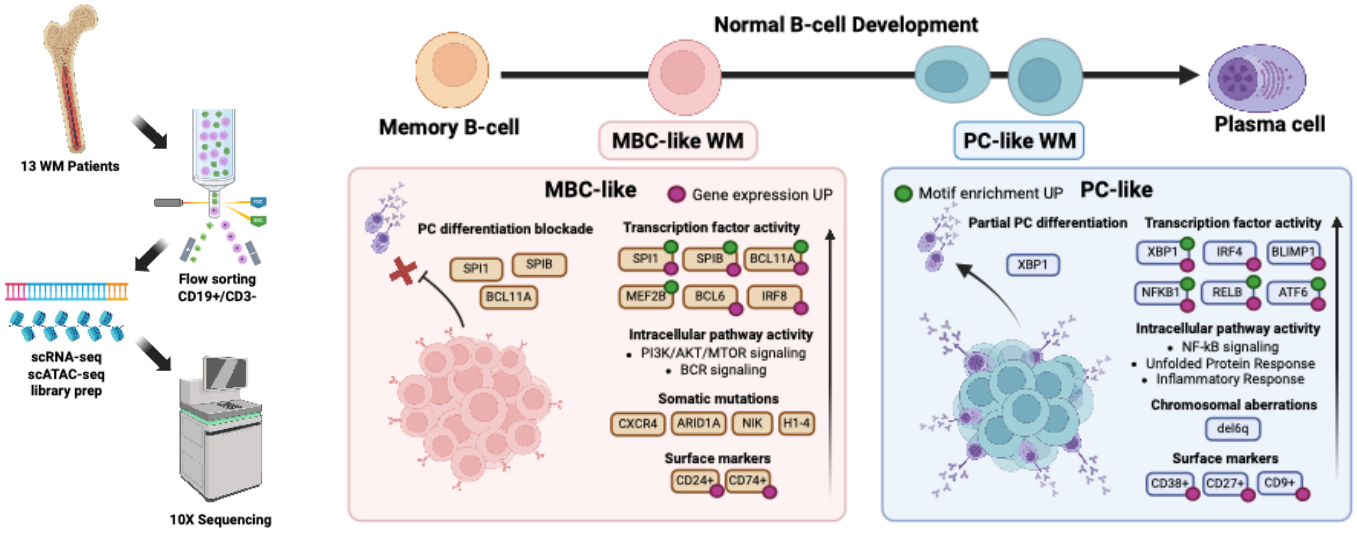
Context of Research

Waldenström macroglobulinemia (WM) is defined by an expansion of post-GC lymphocytes, but the cell of origin and nature of the malignant population are uncertain

Aim of This Study

To identify the transcriptional, epigenetic, and B-cell differentiation states of the malignant single-cell populations in *MYD88*-mutated WM

Findings



Conclusions: WM comprises a clonal expansion of memory B-cells (MBC) with variable capacities for plasma cell (PC) differentiation, forming two disease subtypes: MBC-like and PC-like. The MBC-like have enhanced *SPI1/SPIB* activity, more CXCR4 mutations, and increased BCR signaling whereas the PC-like have enhanced *XBP1* activity, more del(6q), and increased NF-κB activity.

*Gagler et al. DOI: 10.xxxx/**blood**.2025xxxxxx*

blood Visual Abstract