Title:

Aberrant extrafollicular B cells, immune dysfunction, myeloid inflammation and MyD88-mutant progenitors precede Waldenstrom macroglobulinemia

[1]Akhilesh Kaushal, [1,2]Ajay K. Nooka, [1]Allison R. Carr, [3]Katherine E. Pendleton, [1]Benjamin G. Barwick, [1]Julia Manalo, [1,4]Samuel S. McCachren, [1,2]Vikas A. Gupta, [1,2]Nisha S. Joseph, [1,2]Craig C. Hofmeister, [1,2]Jonathan L. Kaufman, [1,2]Leonard T. Heffner, [5]Stephen M. Ansell, [1,2]Lawrence H. Boise, [1,2]Sagar Lonial, [2,3,4]Kavita M. Dhodapkar*, [1,2,4]Madhav V. Dhodapkar*

Affiliations:
[1]Department of Hematology/Oncology, Emory University, Atlanta, GA

[2]Winship Cancer Institute, Emory University, Atlanta, GA

[3]Aflac Cancer and Blood Disorders Center, Children's Healthcare of Atlanta, Department of Pediatrics, Emory University, Atlanta, GA

[4]The Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, GA

[5]Division of Hematology, Mayo Clinic, Rochester, MN

*share senior authorship

Running title: origins of WM

Keywords: macroglobulinemia, MGUS, MYD88, extrafollicular B cells, immune-surveillance

Correspondence: Madhav Dhodapkar (madhav.v.dhodapkar@emory.edu) or Kavita Dhodapkar (kavita.dhodapkar@emory.edu) 1760 Haygood Drive, Atlanta GA 30322

Abstract:

Waldenstrom macroglobulinemia (WM) and its precursor IgM gammopathy are distinct disorders characterized by clonal mature IgM-expressing B cell outgrowth in the bone marrow. Here we show by high-dimensional single-cell immunogenomic profiling of patient samples that these disorders originate in the setting of global B cell compartment alterations, characterized by expansion of genomically aberrant extrafollicular B cells of the non-malignant clonotype. Alterations in immune microenvironment preceding malignant clonal expansion include myeloid inflammation and naïve B and T cell depletion. Host response to these early lesions involves clone-specific T cell immunity that may include MYD88 mutation-specific responses. Hematopoietic progenitors carry the oncogenic MYD88 mutations characteristic of the malignant WM clone. These data support a model for WM pathogenesis wherein oncogenic alterations and signaling in progenitors, myeloid inflammation and global alterations in extrafollicular B cells create the milieu promoting extra-nodal pattern of growth in differentiated malignant cells.

Statement of Significance:

These data provide evidence that growth of malignant clone in WM is preceded by expansion of extrafollicular B cells, myeloid inflammation and immune dysfunction in the preneoplastic phase. These changes may be related in part to MYD88 oncogenic signaling in pre-B progenitor cells and suggest a novel model for WM pathogenesis.

Introduction:

Waldenstrom macroglobulinemia (WM) is a clinically distinct B cell malignancy characterized by progressive growth and accumulation of an IgM-expressing lymphoplasmacytic clone(1, 2). Clinical malignancy in WM is preceded by a precursor state termed as IgM monoclonal gammopathy of undetermined significance (MGUS)(1). Genetic analyses have demonstrated recurrent mutations in myeloid differentiation primary response-88 (MYD88) gene in the vast majority of WM patients, wherein they are found to involve the entire expanded B cell clone(3-5). However, mutations in the MYD88 gene are neither specific nor sufficient for the pathogenesis of WM, and can be detected in IgM MGUS as well as in other B cell lymphomas, including central nervous system or testicular lymphomas (4, 6-11). The clinical phenotype of both IgM MGUS and WM is distinct and characterized by the infiltration of the bone marrow as the dominant site of tumor growth. While patients with IgM MGUS are asymptomatic and exhibit low clonal burden, onset of clinical malignancy is often characterized by the development of anemia and progressive tumor infiltration. In spite of recent advances in the genetics of WM tumor cells, the mechanisms underlying the pattern of tumor growth as well as the cell of origin of this malignancy remain poorly understood(12-14).

Current therapy for WM is based on monoclonal antibodies targeting B cells, strategies targeting B cell receptor (BCR) signaling, as well as systemic chemotherapy(1). While these therapies lead to high rates of tumor regression, they are not curative. Immune system has emerged as a promising strategy to target tumors, including those in the lymphoid system. Properties of the host immune response and the nature of tumor immune microenvironment are an important determinant of outcome in several malignancies and may impact the evolution of premalignant states. However, current information about immune microenvironment in WM and IgM MGUS are limited to studies evaluating only a small number of parameters(12, 15-17), and data with application of newer high-content single cell approaches in WM and MGUS to probe early changes in the immune microenvironment are limited. Importantly, evidence that the immune system can mediate tumor-specific immune recognition of WM tumor cells is lacking. In this study, we combine several single cell proteomic and genomic tools with functional studies to gain insights into immunobiology and host response in WM and its precursor IgM MGUS.

Results:

In order to better understand early events in the pathogenesis of WM, we analyzed the bone marrow mononuclear cells of patients with IgM MGUS (n=8) or newly diagnosed / previously untreated WM (n=8) (suppl Table1 for patient characteristics) with a combination of mass cytometry and CITE-seq and compared it to those from age-matched healthy donors (HD, n=5)(Fig1a). Mass cytometry analyses revealed an increase in $CXCR5^{neg}$ B cells in WM, along with a relative decline in myeloid cells (Fig 1b, suppl Fig1). Importantly, these changes were observed as early as the MGUS stage, when the clonal burden is low. Mass cytometry analysis also revealed an increase in bone marrow T cells in patients with MGUS(Fig 1b, suppl Fig1). CITE-seq analyses on 84,128 single cells (HD 20946 cells, MGUS 29780 cells, WM 33402 cells) identified 42 distinct clusters, which were also broadly classified into B / lymphoplasmacytoid (LPC), T/NK, monocyte/myeloid and precursor cell types (Fig1c-1d, suppl Fig2-3), and again confirmed alterations in myeloid, and B and T cell lineages (Fig1e, suppl Fig4). Together, these data utilize complementary tools to show that bone marrow cells in both WM and IgM MGUS are characterized by alterations in several hematopoietic lineages compared to healthy donor counterparts.

Changes in B and lympho-plasmacytoid cells

As WM is a B cell malignancy, we first focused on analyses of CD19+ cells by mass cytometry. Changes in the B cell compartment were evident as early as IgM MGUS, with an increase in $CXCR5^{neg}$ B cells and decline in naïve B cells, although the clonal population at this stage is still small (Fig2a-2d, suppl Fig5). CXCR5 is well recognized as a critical gene for entry of B cells into lymphoid follicles(18) and $CXCR5^{neg}CD21^{lo}$ B cells have been recently appreciated as cells in the extrafollicular (EF) pathway of B cell differentiation(19). The $CXCR5^{neg}$CD19+ population consisted of two distinct populations, with one representing mature B cells ($CXCR5^{neg}$ population 1), and another population expressing CD38 that includes lympho-plasmacytoid cells ($CXCR5^{neg}$ population 2) (Fig2b). Malignant clone in WM is characterized by the expression of clonotypic kappa or lambda light chain, which secretes a clonal Ig clinically documented as a M spike. While the observed alterations in B cells were enriched in the malignant B cells expressing clonotypic light chain, they were also observed in the B cells expressing the opposite (or non-clonal) light chain (Fig2e). In order to validate these findings, we also analyzed CD19+ B cells by CITE-seq (overall strategy in suppl Fig2). UMAP clustering of B cells based on antibody staining identified 10 distinct clusters, which again revealed progressive loss of $IgD^{+}CD27^{-}$ naïve B cells (cluster 2) from HD donors to IgM MGUS and WM, along with an increase in clusters containing $CXCR5^{neg}$ B cells (cluster 3) (Fig3a-c). Several of these clusters also expressed plasma cell markers such as BCMA and CD138, consistent with lympho-plasmacytoid differentiation of B cells in WM (suppl Fig6). Trajectory analysis of CD19+ B lineage cells identified two distinct trajectories originating from pre-B cells, with one enriched in HD and another in WM (suppl Fig7). UMAP clustering of CD19+ cells based on transcriptome identified 24 distinct clusters (Fig3d-f) (suppl Fig8). Of note, B cells from WM patients as well as some MGUS patients formed distinct clusters, indicating that the malignant clone in

each patient is transcriptionally distinct, analogous to prior studies in other malignancies(20, 21) (Fig3e). Clonality of B cells in these patient-specific clusters was confirmed by B cell receptor (BCR) sequencing, as well as computational imputation of BCRs from single cell transcriptome data (suppl Fig9). Examples of top genes characterizing specific clones in individual patients include genes implicated in lymphomagenesis including EZH2, CD79b, CXCR4, BCL2, and LT-β(Fig3f). Overall, B cells from WM and IgM MGUS patients were transcriptionally distinct from those in HD. Top pathways increased in WM and IgM MGUS B cells included NFkB pathway and IL-6/STAT3 pathway, consistent with prior genomic studies with bulk tumor cells (suppl Fig10-11)(7).

Changes in non-clonal light chain expressing B cells and pre-B cells
Based on the findings from mass cytometry relating to B cells expressing non-clonal light chain (for example, kappa light chain in a patient with IgM lambda M spike), we further evaluated these cells by CiTE-seq. Polyclonal nature of these B cells and lack of clonal BCR was verified by BCR sequencing (shown as an example in supplFig9c). Transcriptome of these non-clonal light chain expressing B cells in WM as well as MGUS patients also revealed several changes compared to their HD counterparts, with increased expression of genes such as Pim1, TCF4, and decrease in HLA-DR previously implicated in lymphoma biology(suppl Fig12).

In addition to mature B cells, transcriptomes of CD19+CD10+ pre-B cells from WM or IgM MGUS patients also revealed several alterations compared to HD pre-B cells (supplFig13). Interestingly, signatures derived from differentially expressed genes in WM B cells compared to HD were also found to be enriched in WM pre-B cells (suppl Fig 14). Top pathway enriched in WM pre-B cells included IL6-STAT3 pathway, as also seen in WM B cells (suppl Fig15). Taken together, these data demonstrate that clonal populations in WM and IgM MGUS originate in the context of more global alterations in B as well as pre-B cells, with resultant expansion of EF B cells.

Changes in the myeloid compartment
Besides B cells, the most prominent changes in the immune microenvironment were in the myeloid compartment (Fig1b,1e). UMAP clustering of CD3 and CD19-depleted cells was utilized to identify distinct clusters that broadly fell into myeloid, NK and progenitor clusters and were analyzed separately (Fig4a) (suppl Fig16). Analysis of CD14+/CD11c+ myeloid compartment revealed that both IgM MGUS and WM were associated with marked decline in clusters of classical monocytes (cluster1) and a relative increase in a distinct cluster of myeloid cells (cluster 7) particularly in MGUS (Fig4a-4d) (Suppl Fig17). This myeloid population was characterized by a distinct genomic signature of inflammation-associated genes such as IL1β, CCL4, CCL3, IL6, NLRP3 and CXCL3 (Fig4e). Pathway analysis was consistent with inflammation-associated signaling, particularly in MGUS (Fig4f). Together these data suggest that activation of myeloid inflammation is an early feature of MGUS, before the evolution of the malignant clone.

Changes in innate cells

Evaluation of CD3-CD56+ NK clusters identified alteration in NK cells, most prominently in the WM cohort (Fig4a,4c)(suppl fig18). These changes were characterized by greater expression of lytic genes /exhaustion markers in WM NK cells, along with loss of interferon-signature in these cells (suppl Fig18) consistent with the concept of immune exhaustion and dysfunction in NK cells with evolution of WM. Pathway analysis of differentially expressed genes in NK cells from both MGUS and WM revealed decreased interferon signaling consistent with NK dysfunction (suppl Fig19).

Changes in T cells and evidence of tumor-specific immunity
Analysis of T cell compartment by mass cytometry revealed evidence of T cell activation with a decrease in naïve CD8 T cells and increased CD8+ T cells expressing granzyme (Fig5a-d). These cells also express high levels of Eomes and T-bet and low levels of TCF1(Fig5b). Notably, expression of PD1 was not a prominent feature of these cells, which expressed Tigit instead (Fig5b). Analysis of transcriptomes of T cells between WM or MGUS cohorts and HD revealed that as with mass cytometry, the most prominent changes were an increase in lytic genes and markers associated with T cell exhaustion in WM/MGUS CD8+ T cells, concurrent with loss of interferon-response genes (Fig5e,5f). Changes in interferon-response genes were also seen in CD4+ T cells(suppl Fig20). Pathway analysis confirmed changes in T cells, most notably as decline in interferon response (suppl Fig21-22). Together, these data show that changes in the T cell compartment begin early in MGUS, before the establishment of progressive malignant clone, and are characterized by progressive depletion of naïve T cells and instead enrichment of terminal effector T cells.

Although the T cell phenotypes are consistent with evidence of T cell activation in vivo, whether the immune system mediates specific recognition of WM cells in vivo is not known. In order to test this directly, we analyzed reactivity of freshly isolated blood/bone marrow T cells to autologous monocyte-derived dendritic cells (DCs) loaded with autologous tumor cells. In all patients tested, freshly isolated T cells could mediate tumor-specific immune recognition (Fig6a). In order to further evaluate if these T cells still retained proliferative capacity and could be enriched further, we cultured these cells with autologous tumor-loaded DCs. DC-mediated stimulation led to expansion of tumor-specific T cells in culture (Fig6b). Tumor-reactivity persisted following depletion of CD4+ T cells, indicating that it consisted predominantly of CD8+ T cells (Fig6c). MYD88 L265P is a common and recurrent mutation found in WM cells. In order to test whether anti-tumor T cell reactivity includes MYD88-specific T cells, we analyzed reactivity to mutant and wild-type peptides in blood/marrow mononuclear cells, with peptide-specific stimulation. Data from representative responder and non-responder patients are shown in Fig6d. Overall, reactivity specific for mutant MYD88 peptide was detected in 5 of 15 patients tested (Fig6e). Together, these data demonstrate that the immune system is capable of mediating tumor-specific recognition of WM cells, that includes reactivity to mutant MyD88 in a subset of patients.

Genetic alterations in precursor cells
The findings discussed above demonstrate that the although the expanded clone in WM has the phenotype of mature IgM+ B cells, the evolution of the malignant clone occurs

in the context of more global genomic and phenotypic aberrancies in hematopoiesis, which precede the expansion of the clone. In order to test whether genomic alterations (including MYD88 mutations) previously characterized in the mature B cell clone can also be detected in earlier progenitors, we flow-sorted CD19+CD10- mature B cells, as well as less mature progenitors ($Lin^{neg}CD19+CD10+$ pre-B cells, $Lin^{neg}CD19^{neg}CD34^{neg}$ and $Lin^{neg}CD19^{neg}CD34+$ progenitors) from 5 patients with WM and 2 HD and analyzed their genomes by exome sequencing (gating strategy in suppl Fig23). In all patients studied, the MYD88 L265P mutation could be detected in pre-B progenitor compartments and involved the entire mature B cell clone (Fig7a-c). In 4 of the 5 patients studied, the MYD88 L265P mutation was detected with variant allele frequency (VAF) of >10% in at least one progenitor compartment (>10% VAF in CD19+CD10+ cells in 3 of 5 patients), indicating its detection in over 20% of progenitor cells (Fig7c). In contrast, MYD88 L265P mutations were not detected in any fraction from HD marrow (suppl Fig24). The finding that MyD88 mutant fractions show clear expansion (>10% VAF / 20% of cells) in CD19+CD10+ preB fractions suggests that this is not a function of low level contamination during flow sorting. This is also supported by the finding that several SNVs were not shared between sorted fractions. The detection of MYD88 L265P by exome sequencing was verified by allele-specific qPCR, which yielded similar VAF estimates (supplFig25a) and confirmed the detection of MYD88 L265P mutations in precursor compartments (suppl Fig25b). MYD88 L265P was also detected by PCR at low levels in the CD19+CD10- cells and at least one other progenitor compartment in MGUS marrow but not in any of the sorted populations in HD (suppl Fig25c-d). We also tested whether MYD88 mutations could be detected in circulating B cells expressing the non-clonal light chain. Analysis of flow sorted B cells (based on expression of CXCR5 and Ig light chain) by PCR revealed that low levels of MYD88 mutation could be detected in the non-clonal IgL+ B cells in WM patients but not HD (suppl Fig25e-f). Together, these data demonstrate that MYD88 L265P mutations can occur early in lymphopoiesis in WM, before the expansion of the malignant B cell clone, and are not restricted to the expanded clonal population.

Discussion:

These data demonstrate that the preneoplastic phase of WM is characterized by a distinct immune landscape with marked alterations in non-clonal B cell compartment. Coupled with the detection of MYD88-L265P mutations in early pre-B progenitors, these data also suggest that although the expanded clonal population in WM has the phenotype of a mature B cell, the clone emerges in the context of alterations in earlier B progenitors that begin well before expansion of the malignant clone.

Changes in the B cell compartment are manifest not only as a decline in normal naïve B cells but marked increase in EF B cells with $CXCR5^{neg}$CXCR4+ phenotype. Importantly, while clonal B cells in WM are also often $CXCR5^{neg}$, changes in EF B cells are also observed in those expressing non-clonal immunoglobulin light chain. These non-clonal cells also overexpress several genes such as Pim1 and TCF4 previously implicated in lymphomagenesis(22, 23). These data are therefore consistent with the concept that WM and IgM MGUS clones develop in the context of global alterations of B cells with expansion of aberrant non-clonal EF B cells. Recent studies have demonstrated enhanced autoreactivity in EF B cells implicating them in several autoimmune states(19, 24). The observed expansion of EF B cells in WM and IgM MGUS may therefore underlie the increased risk of autoimmunity in these disorders(25). The observed decline in naïve B cells may lead to reduced response to pathogens and vaccines in these patients(26, 27).

The mechanism underlying the observed increase in EF B cells is not known but may be related in part to aberrant MYD88 signaling, based on our finding of MYD88 mutations in pre-B cells. MYD88 has been implicated in both B cell as well as myeloid differentiation, the latter being responsible for its name(28, 29). MYD88 plays a central role in toll-like receptor (TLR) signaling and this pathway was recently suggested to play a critical role in B cell differentiation along the EF pathway (30, 31). Transformation of cells in the EF pathway may also help explain the predominantly extra-nodal / bone marrow-restricted pattern of tumor growth in WM. It is also notable that in addition to WM, MYD88 L265P mutations are also highly prevalent in some other extra-nodal lymphomas such as central nervous system or testicular lymphoma(32).

Our studies also provide the first evidence of tumor-specific immune recognition in WM. T cell compartment in WM and IgM MGUS also exhibited depletion of naïve T cells and concomitant increase in effector populations consistent with in vivo activation and similar to findings in other hematologic malignancies(21, 33). Importantly, evidence of immune dysfunction with attrition of TCF1+ stem like memory T cells and increase in terminally differentiated T as well as NK cells begins early in the MGUS stage, before the malignant clone is established, similar to recent findings in IgG MGUS(33). Although these T cells express markers of exhausted T cells, tumor-specific T cells could be expanded in culture and retained capacity for tumor-reactive cytokine production. In

order to measure tumor-specific immunity, we prioritized dendritic cell (DC)-based cross-presentation assays, which allow detection of immunity against multiple antigens. Our studies also show that at least in some patients, the MYD88 L265P mutation is immunogenic. While WM-specific T cells were present in all patients tested, MYD88-specific reactivity was observed only in a subset, suggesting that other neoantigens may also be targets of WM-specific response in this setting. Immunogenicity of MYD88-L265P in some WM patients is consistent with recent studies priming mutation-specific T cell receptors (TCRs) in healthy donors (34); however further studies are needed to better understand the functional properties of MYD88-L265P-specific T cells in WM tumor bed and identify specific MHC haplotypes wherein high-affinity TCRs could be harnessed for immunotherapy. Further studies are also needed to quantify the degree to which mutant MYD88 is a target of global WM-specific immunity. This analysis was hampered in the current study by limitations of sample and HLA restriction. Harnessing pre-existing immune response in WM patents may in principle also be feasible via blockade of immune checkpoints, particularly Tigit, overexpressed in these cells.

Our data suggest a novel model for the pathogenesis of WM, wherein aberrant increase in polyclonal CXCR5^{neg}CXCR4+ EF B cell compartment create the milieu for the emergence of MGUS (Fig7d). MYD88 mutations are not only present in the malignant B cell clone, but also earlier pre-B progenitors. In some patients, the mutant preB compartment was expanded to >20% of the preB compartment. The presence of mutations in even earlier progenitors requires further study, however in at least one patient (WM patient 1 in Fig 7), the mutant fraction accounted for >20% of CD34+ CD19- cells, indicating inter-patient heterogeneity relating to detection and expansion of the mutant fraction in earlier progenitors in some WM patients. Humanized models(35) that permit efficient growth of WM cells may be needed to better understand the biology of progenitor compartments in WM. Evolution of precursor states to WM may depend both on acquisition of additional genomic changes in tumor cells and alterations in non-malignant cells. This model is similar to emerging data for polyclonal origins of non-IgM monoclonal gammopathies with the pre-MGUS phase beginning in early decades of life(36-38). It is also consistent with the view that MYD88 mutations are not sufficient for the development of WM or other MYD88-driven lymphomas and additional alterations in tumor cells and/or host response are needed(8, 9). Immune microenvironment in the precursor state is also characterized by myeloid inflammation. The inflammatory environment may promote the malignant clone but also induce immune exhaustion. Our data that the immune system can mediate WM-specific immune responses supports potential for immune-mediated control, also suggested in MYD88-driven murine lymphoma models(39). The concept that alterations in hematopoietic progenitors may underlie the development of mature B cell malignancies has also been previously suggested in the case of some other B cell malignancies such as chronic lymphocytic leukemia and hairy cell leukemia(40-42).

These data have several implications for biology and therapy of WM. Homing properties of EF cells based on lack of CXCR5 expression may explain the predominantly extra-

nodal pattern of tumor involvement in WM. The presence of oncogenic mutations in earlier progenitors lacking clonal B cell receptor (BCR) suggests that oncogenic signaling may contribute to changes not just in tumor cells, but also non-malignant cells in the tumor microenvironment. Targeting MYD88 mutant progenitors as well as harnessing the ability of the immune system to target these tumors may be needed to improve therapy and curability of WM. With increasing appreciation of age-associated genomic alterations in hematopoiesis, this issue also needs broader evaluation across human tumors.

Methods:

Patients:
Bone marrow specimens from patients with a diagnosis of MGUS or WM (1) were obtained following written informed consent from patients. Studies were approved by Emory institutional review board and were conducted in accordance with recognized ethical guidelines. Specimens from healthy donors were purchased from All Cells, Inc. Patient characteristics are shown in supplemental Table 1.

Mass cytometry
Thawed bone marrow mononuclear cells (BMMNCs) were stained with custom panels of metal conjugated antibodies according to manufacturer-suggested concentrations (Fluidigm) (Antibodies as noted in supplemental Table 2). Cells were fixed, permeabilized and washed in accordance with the manufacturer's cell surface and nuclear staining protocol as described. After antibody staining, cells with incubated with intercalation solution, mixed with EQ Four Element Calibration Beads (catalogue number 201708), and acquired using a Helios mass cytometer (all from Fluidigm). Gating and data analysis was performed using Cytobank (https://www.cytobank.org). Viable cells and doublets were excluded using cisplatin intercalator and DNA content with iridium intercalator. Equal numbers of cells from each donor were utilized when data were concatenated prior to analysis.

CITE-Seq
After thawing bone BMMNCs, live cells were isolated using the EasySep™ Dead Cell Removal (Annexin V) Kit (STEMCELL Technologies) per the manufacturer's protocol. Cells were then stained using the TotalSeq-C antibody cocktail (supplemental Table 3) following 10X Genomics' protocol for Chromium Single Cell Immune Profiling with Feature Barcoding Technology (ver. 1.0). Gene expression and cell surface libraries were prepared according to the protocol from 10X Genomics. The quality of the prepared libraries was assessed using the Agilent HS Bioanalyzer 2100 and sequencing was conducted with the Illumina NovaSeq 6000. In order to minimize variability, all patients and controls were sequenced together.

Data Analysis for CITE-Seq
Reads were aligned to human reference sequence GRCh38/hg38, filtered, deduplicated and converted into a feature barcode matrix using Cell Ranger 4.0. Samples with low fraction reads (<70%), high reads mapped to antisense strand (>10%) and with problematic barcode rank plot were excluded. Final samples with feature barcode matrix containing data for genes and 55 surface proteins were analyzed using Seurat V3 (43). Cells with fewer than 200 unique sequenced genes, more than 10% mitochondrial gene and more than 70,000 sequenced features were filtered to exclude dead cells and doublets. The data was batch corrected using canonical correlation analysis (CCA) method using Seurat V3 data integration pipeline. Gene expression for each cell was

normalized using R package SCTransform(44) which fits regularized negative binomial (NB) regression model. For each gene unique molecular identifiers (UMI) counts are considered as the response and cellular sequencing depth as the predictor variable to obtain regularized parameter estimates. The estimated regularized regression parameter was used to transform the UMI counts for each gene into Pearson residuals serving as a scaled gene expression data. Regularized NB model in comparison to log-normalization technique has been suggested to prevent overestimation of true variance in low-abundance genes(44). The final data set used for analyses consisted of 84,128 cells from 18 samples (Healthy, n =4; IgM MGUS, n = 7; and WM, n =7). Protein expression data were treated as compositional data and normalized independently of transcriptomic data using centered log-ratio (CLR) normalization, where counts were divided by the geometric mean of the corresponding feature across cells, and log-transformed (45). Principal component analysis (PCA) was then performed using top 3000 variable genes ranked by residual variance. Elbow method was utilized to identify number of principal components that explained maximum variability in data which resulted in 20 as an optimal number of principal components. Local neighborhood for each cell was defined by taking the 20 nearest neighbors in the kNN-Graph calculated using Euclidean distance in the PCA space. Rare and non-convex cell populations were further identified using shared nearest neighbor (SNN) networks by calculating neighborhood overlap using Jaccard index. Finally, Louvain community detection algorithm with resolution of 0.5 was used to identify 42 clusters of cells with similar gene expression profiles. These 42 clusters were further classified into four major categories B, T/NK, myeloid and precursor cell types based on the protein expression of CD19, CD3, CD14 and confirmed by reference transcriptomic datasets of immune cells with SingleR (46). Cells in each of these four major categories were analyzed separately and sub-clustering within each category were performed based on transcriptomic and protein expression data. The sub-cluster identity within each category was determined by using significantly differentially expressed genes and proteins using Wilcoxon rank-sum test between each cluster and the rest. Differentially expressed genes were denoted as statistically significant for the Bonferroni adjusted $p < 0.05$ with a fold change exceeding 1.2x (i.e., ≥ 1.2 or ≤ 0.83). Pathway analysis for significantly differentially expressed genes was performed with DAVID Gene Functional Classification Tool (47) using over-representation analysis method. Gene set enrichment analysis (GSEA) (48) method was used for pathway analysis based on pre-ranked genes. Reactome (49) and Hallmark (50) gene sets were utilized and pathways with q-value<0.05 were retained. Data are uploaded to GEO at GSE179221.

B cell receptor (BCR) sequencing and BCR imputation from transcriptomic data

B cell receptor sequencing (BCR) reads from FASTQ were aligned to GRCh38-alts-ensembl using Cellranger vdj pipeline version 6.0.2 to generate single-cell V(D)J (variable, diversity and joining) segments for heavy and light chains. The unique combination of heavy and light chain genes is referred to as clonotype. The clonotype data for BCR sequencing were mapped to Seurat object containing CiTe-Seq data by matching cell barcodes. Clonotypic frequency obtained from Cellranger output were

depicted using pie-chart. We also imputed the BCR clonotype from transcriptomic data using Cellranger VDJ pipeline 6.0.2. Clonotypic frequency were generated similar to BCR sequencing data.

Cell sorting

Frozen Bone marrow mononuclear cells (BMNNCs) or peripheral blood mononuclear cells (PBMCs) from healthy donors (n=3) and patients with IgM MGUS (n=2) and WM (n=5) were thawed and stained prior to cell sorting using BD FACSAria. BMMNCs were stained with live dead dye (Thermoscientific) as well as the following antibodies to detect CD19 (clone HIB19, BD Bioscience) and CD3 (HIT3a), 14 (M5E2) CD56 (HCD56), CD10 (HI10a) and CD34 (581) all from Biolegend. Cells were gated to obtain single, living, Lin-(CD3,14,56-) cells and 4 different fractions were collected for exome sequencing: CD19+CD10-, CD19+CD10+, CD19-CD34- and CD19-CD34+. DNA was extracted using the AllPrep DNA/RNA Micro Kit (Qiagen) and used for exome sequencing or QPCR analysis to detect MYD88L265P mutation. PBMNCs were labeled with live dead dye, and antibodies to detect CD19 and CXCR5 (clone RF8B2, BD) as well as Igkappa light chain (TB28-2), Iglambda light chain (MHL-38) from Biolegend. Cells were gated to obtain single, live cells and 4 different fractions of CD19+ cells were sorted: Kappa+CXCR5+, Kappa+CXCR5-, Lambda+CXCR5+, Lambda+CXCR5- cells. DNA was extracted using AllPrep DNA/RNA Micro Kit (Qiagen) and used to perform QPCR for detection of MyD88L265P.

Exome sequencing and data analysis

Exome capture was performed using the IDT xGen v1 capture kit (Integrated DNA Technologies) according to manufacturer's instructions. Exome FASTQ sequencing files were quality and adapter trimmed using Trim Galore (v0.6.4) with Cutadapt (v2.8) (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) prior to mapping to the GRCh38 reference genome (GRCh38.d1.vd1.fa.tar.gz; https://gdc.cancer.gov/about-data/gdc-data-processing/gdc-reference-files) with BWA (v0.7.17-r1188)(51). Mapped SAM files were converted to BAM files with SAMtools (v1.10)(52)and read groups and duplicate reads were marked using the Genome Analysis Toolkit (v4.1.9.0)(53). Variant alleles were called with Mutect2 (v4.1.9.0)(53) with all BAM files for a given patient included and using the 1000 genomes project 'Panel of Normals'. VCF files were converted to a matrix format in R (v3.6.3) with the bedr (v1.0.7) package, and annotation of variants was performed with the VariantAnnotation (v1.32.0) package(54). Variants with at least 40x coverage in all samples were analyzed for different allele frequencies between any two cell types using a two-sided Fisher's exact test with a Benjamini-Hochberg FDR correction. Variant allele frequencies from Mutect2 were shown in figures and Venn diagrams of mutations required a minimum of two reads to call a mutation present in a given cell type. Overall, exome sequencing was performed to median depth of 162,486,965 reads with greater than 99% mappability resulting in a median exome coverage of 213x.

MYD88 L265P allele specific qPCR

Allele specific qPCR for MYD88 L265P was performed as previously described(55). Briefly the qBiomarker Somatic Mutation PCR Assay (Qiagen) MYD88_85940 was used to determine the total MYD88 fraction as compared to the MYD88 L265P alleles on a CFX96 Touch Real-Time PCR Detection System (Bio-Rad). Variant allele fraction was calculated as the ratio of variant allele to total MYD88 using the following formula as per manufacturer recommendations.

$$\mathrm{VAF} = 2^{(\mathrm{Ct\ MYD88\ total} - \mathrm{Ct\ MYD88\ L265P})}.$$

Evaluation of WM-specific immunity

The presence of WM-specific T cells in blood or bone marrow was quantified both directly ex vivo and after in vitro stimulation with autologous tumor loaded dendritic cells (DCs) using a 16h ELISPOT assay as described(56). This assay allows for detection of global tumor-specific immunity against multiple tumor-derived antigens. Briefly, autologous T cells were cultured with unpulsed DCs (as controls), or CD19+ tumor-loaded DCs at a T:DC ratio of 1:10. After 16h, the presence of tumor-specific IFN-γ spot forming cells was quantified as described(56). For in vitro expansion, T cells were cultured with autologous tumor-loaded DCs for 10-14 days at T:DC ratio of 1:30 in the presence of IL2 (Chiron) at 10 U/ml on days 4 and 7. In this assay, DCs are needed both for expansion of tumor-specific T cells as well as for readout of T cell immunity.

Detection of MYD88-speciifc immunity

Thawed MNCs were cultured overnight with peptides containing wild type or mutant (L265P) MYD88 mutation sequences. After overnight culture, culture supernatant was harvested and analyzed for the presence of IP-10, as previously described(57).

Statistics

Statistical analysis of mass cytometry data was performed using 2D graphing and statistics software GraphPad Prism. Nonparametric Mann-Whitney (for comparing 2 groups) and Kruskal-Wallis (for comparing 3 groups) tests with a significance threshold of $P < 0.05$ were used to compare different cohorts. Wilcoxon's rank-sum test with a significance threshold of $P < 0.05$ after Bonferroni's correction was used to identify differentially expressed genes between clusters and disease states in the scRNA-Seq data. We used χ^2 with a significance threshold of $P < 0.005$ to identify clusters with differential composition by disease state. Data in bar graphs were plotted as mean ± SEM.

List of supplementary materials:

Supplementary Figures 1-25

Supplementary Tables 1-3

References:

1. Gertz MA. Waldenstrom Macroglobulinemia: 2020 Update on Diagnosis, Risk Stratification, and Management. Am J Hematol. 2020.
2. Dhodapkar MV, Jacobson JL, Gertz MA, Rivkin SE, Roodman GD, Tuscano JM, et al. Prognostic factors and response to fludarabine therapy in patients with Waldenstrom macroglobulinemia: results of United States intergroup trial (Southwest Oncology Group S9003). Blood. 2001;98(1):41-8.
3. Treon SP, Xu L, Yang G, Zhou Y, Liu X, Cao Y, et al. MYD88 L265P somatic mutation in Waldenstrom's macroglobulinemia. N Engl J Med. 2012;367(9):826-33.
4. Xu L, Hunter ZR, Yang G, Zhou Y, Cao Y, Liu X, et al. MYD88 L265P in Waldenstrom macroglobulinemia, immunoglobulin M monoclonal gammopathy, and other B-cell lymphoproliferative disorders using conventional and quantitative allele-specific polymerase chain reaction. Blood. 2013;121(11):2051-8.
5. Treon SP, Xu L, Guerrera ML, Jimenez C, Hunter ZR, Liu X, et al. Genomic Landscape of Waldenstrom Macroglobulinemia and Its Impact on Treatment Strategies. J Clin Oncol. 2020;38(11):1198-208.
6. Ngo VN, Young RM, Schmitz R, Jhavar S, Xiao W, Lim KH, et al. Oncogenically active MYD88 mutations in human lymphoma. Nature. 2011;470(7332):115-9.
7. Hunter ZR, Yang G, Xu L, Liu X, Castillo JJ, Treon SP. Genomics, Signaling, and Treatment of Waldenstrom Macroglobulinemia. J Clin Oncol. 2017;35(9):994-1001.
8. Knittel G, Liedgens P, Korovkina D, Seeger JM, Al-Baldawi Y, Al-Maarri M, et al. B-cell-specific conditional expression of Myd88p.L252P leads to the development of diffuse large B-cell lymphoma in mice. Blood. 2016;127(22):2732-41.
9. Sewastianik T, Guerrera ML, Adler K, Dennis PS, Wright K, Shanmugam V, et al. Human MYD88L265P is insufficient by itself to drive neoplastic transformation in mature mouse B cells. Blood Adv. 2019;3(21):3360-74.
10. Wang JQ, Jeelall YS, Beutler B, Horikawa K, Goodnow CC. Consequences of the recurrent MYD88(L265P) somatic mutation for B cell tolerance. J Exp Med. 2014;211(3):413-26.
11. Wang JQ, Jeelall YS, Humburg P, Batchelor EL, Kaya SM, Yoo HM, et al. Synergistic cooperation and crosstalk between MYD88(L265P) and mutations that dysregulate CD79B and surface IgM. J Exp Med. 2017;214(9):2759-76.
12. Paiva B, Corchete LA, Vidriales MB, Garcia-Sanz R, Perez JJ, Aires-Mejia I, et al. The cellular origin and malignant transformation of Waldenstrom macroglobulinemia. Blood. 2015;125(15):2370-80.
13. Garcia-Sanz R, Jimenez C, Puig N, Paiva B, Gutierrez NC, Rodriguez-Otero P, et al. Origin of Waldenstrom's macroglobulinaemia. Best Pract Res Clin Haematol. 2016;29(2):136-47.
14. Sahota SS, Forconi F, Ottensmeier CH, Stevenson FK. Origins of the malignant clone in typical Waldenstrom's macroglobulinemia. Semin Oncol. 2003;30(2):136-41.
15. Jalali S, Ansell SM. The Bone Marrow Microenvironment in Waldenstrom Macroglobulinemia. Hematology/oncology clinics of North America. 2018;32(5):777-86.

16. Jalali S, Price-Troska T, Paludo J, Villasboas J, Kim HJ, Yang ZZ, et al. Soluble PD-1 ligands regulate T-cell function in Waldenstrom macroglobulinemia. Blood Adv. 2018;2(15):1985-97.
17. Hodge LS, Ziesmer SC, Yang ZZ, Secreto FJ, Gertz MA, Novak AJ, et al. IL-21 in the bone marrow microenvironment contributes to IgM secretion and proliferation of malignant cells in Waldenstrom macroglobulinemia. Blood. 2012;120(18):3774-82.
18. Pereira JP, Kelly LM, Cyster JG. Finding the right niche: B-cell migration in the early phases of T-dependent antibody responses. Int Immunol. 2010;22(6):413-9.
19. Jenks SA, Cashman KS, Woodruff MC, Lee FE, Sanz I. Extrafollicular responses in humans and SLE. Immunol Rev. 2019;288(1):136-48.
20. Ledergor G, Weiner A, Zada M, Wang S-Y, Cohen YC, Gatt ME, et al. Single cell dissection of plasma cell heterogeneity in symptomatic and asymptomatic myeloma. Nature Medicine. 2018;24(12):1867-76.
21. Bailur JK, McCachren SS, Pendleton KE, Vasquez JC, Lim HS, Duffy A, et al. Risk associated alterations in marrow T cells in pediatric leukemia. JCI Insight. 2020.
22. van Lohuizen M, Verbeek S, Krimpenfort P, Domen J, Saris C, Radaszkiewicz T, et al. Predisposition to lymphomagenesis in pim-1 transgenic mice: cooperation with c-myc and N-myc in murine leukemia virus-induced tumors. Cell. 1989;56(4):673-82.
23. Jain N, Hartert K, Tadros S, Fiskus W, Havranek O, Ma MCJ, et al. Targetable genetic alterations of TCF4 (E2-2) drive immunoglobulin expression in diffuse large B cell lymphoma. Sci Transl Med. 2019;11(497).
24. Rawlings DJ, Metzler G, Wray-Dutra M, Jackson SW. Altered B cell signalling in autoimmunity. Nat Rev Immunol. 2017;17(7):421-36.
25. Bockorny B, Atienza JA, Dasanu CA. Autoimmune manifestations in patients with Waldenstrom macroglobulinemia. Clin Lymphoma Myeloma Leuk. 2014;14(6):456-9.
26. Dhodapkar MV, Dhodapkar KM, Ahmed R. Viral immunity and vaccines in hematologic malignancies. Blood Cancer Discovery. 2020.
27. Branagan AR, Duffy E, Gan G, Li F, Foster C, Verma R, et al. Tandem high-dose influenza vaccination is associated with more durable serologic immunity in patients with plasma cell dyscrasia. Blood Adv. 2021.
28. Hua Z, Hou B. TLR signaling in B-cell development and activation. Cell Mol Immunol. 2013;10(2):103-6.
29. Cannova J, Breslin SJP, Zhang J. Toll-like receptor signaling in hematopoietic homeostasis and the pathogenesis of hematologic diseases. Front Med. 2015;9(3):288-303.
30. Jenks SA, Cashman KS, Zumaquero E, Marigorta UM, Patel AV, Wang X, et al. Distinct Effector B Cells Induced by Unregulated Toll-like Receptor 7 Contribute to Pathogenic Responses in Systemic Lupus Erythematosus. Immunity. 2018;49(4):725-39 e6.
31. Lam J, Baumgarth N. TLR adaptors MyD88 and TRIF are critical for extrafollicular B cell responses to influenza. J Immunol. 2020;204(1 supplement):151.
32. Lionakis MS, Dunleavy K, Roschewski M, Widemann BC, Butman JA, Schmitz R, et al. Inhibition of B Cell Receptor Signaling by Ibrutinib in Primary CNS Lymphoma. Cancer Cell. 2017;31(6):833-43 e5.

33. Bailur JK, McCachren SS, Doxie DB, Shrestha M, Pendleton K, Nooka AK, et al. Early alterations in stem-like/resident T cells, innate and myeloid cells in the bone marrow in preneoplastic gammopathy. JCI Insight. 2019;5.
34. Cinar O, Brzezicha B, Grunert C, Kloetzel PM, Beier C, Peuker CA, et al. High-affinity T-cell receptor specific for MyD88 L265P mutation for adoptive T-cell therapy of B-cell malignancies. Journal for immunotherapy of cancer. 2021;9(7).
35. Das R, Strowig T, Verma R, Koduru S, Hafemann A, Hopf S, et al. Microenvironment-dependent growth of preneoplastic and malignant plasma cells in humanized mice. Nat Med. 2016;22(11):1351-7.
36. Nair S, Sng J, Boddupalli CS, Seckinger A, Chesi M, Fulciniti M, et al. Antigen-mediated regulation in monoclonal gammopathies and myeloma. JCI Insight. 2018;3(8).
37. Dhodapkar MV. MGUS to myeloma: a mysterious gammopathy of underexplored significance. Blood. 2016;128(23):2599-606.
38. Rustad EH, Yellapantula V, Leongamornlert D, Bolli N, Ledergor G, Nadeu F, et al. Timing the initiation of multiple myeloma. Nat Commun. 2020;11(1):1917.
39. Reimann M, Schrezenmeier JF, Richter-Pechanska P, Dolnik A, Hick TP, Schleich K, et al. Adaptive T-cell immunity controls senescence-prone MyD88- or CARD11-mutant B-cell lymphomas. Blood. 2020.
40. Chung SS, Kim E, Park JH, Chung YR, Lito P, Teruya-Feldstein J, et al. Hematopoietic stem cell origin of BRAFV600E mutations in hairy cell leukemia. Sci Transl Med. 2014;6(238):238ra71.
41. Kikushige Y, Ishikawa F, Miyamoto T, Shima T, Urata S, Yoshimoto G, et al. Self-renewing hematopoietic stem cell is the primary target in pathogenesis of human chronic lymphocytic leukemia. Cancer Cell. 2011;20(2):246-59.
42. Husby S, Gronbaek K. Mature lymphoid malignancies: origin, stem cells, and chronicity. Blood Adv. 2017;1(25):2444-55.
43. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, 3rd, et al. Comprehensive Integration of Single-Cell Data. Cell. 2019;177(7):1888-902 e21.
44. Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Swerdlow H, et al. Simultaneous epitope and transcriptome measurement in single cells. Nat Methods. 2017;14(9):865-8.
45. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. 2019;20(1):296.
46. Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nat Immunol. 2019;20(2):163-72.
47. Wang J, Vasaikar S, Shi Z, Greer M, Zhang B. WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. Nucleic Acids Res. 2017;45(W1):W130-W7.
48. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545-50.
49. Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, et al. The reactome pathway knowledgebase. Nucleic Acids Res. 2020;48(D1):D498-D503.

50. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015;1(6):417-25.
51. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-60.
52. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9.
53. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1297-303.
54. Obenchain V, Lawrence M, Carey V, Gogarten S, Shannon P, Morgan M. VariantAnnotation: a Bioconductor package for exploration and annotation of genetic variants. Bioinformatics. 2014;30(14):2076-8.
55. Ansell SM, Hodge LS, Secreto FJ, Manske M, Braggio E, Price-Troska T, et al. Activation of TAK1 by MYD88 L265P drives malignant B-cell Growth in non-Hodgkin lymphoma. Blood Cancer J. 2014;4:e183.
56. Dhodapkar MV, Krasovsky J, Osman K, Geller MD. Vigorous premalignancy-specific effector T cell response in the bone marrow of patients with monoclonal gammopathy. J Exp Med. 2003;198(11):1753-7.
57. Spisek R, Kukreja A, Chen L-C, Matthews P, Mazumder A, Vesole D, et al. Frequent and specific immunity to the embryonal stem cell–associated antigen SOX2 in patients with monoclonal gammopathy. Journal of Experimental Medicine. 2007;204(4):831-40.

Acknowledgments:

This work was supported in part by funds from International Waldenstrom Macroglobulinemia Foundation and NIH R35CA197603 (to MVD). KMD is supported in part by funds from NIH CA238471, AR077926. MVD and KMD are also supported in part by funds from Specialized Center for Research (SCOR) award from the Leukemia and Lymphoma Society and from Riney Foundation. Authors acknowledge the help of Deon Doxie and Roman Radzievski in Winship mass cytometry resource and colleagues in the Emory integrated genomics core (both supported by P30CA138292 to Winship shared resources).

Figure legends:

Fig1. Changes in bone marrow microenvironment comparing healthy donors (HD), IgM MGUS and WM

1a. Overall strategy. Bone marrow mononuclear cells (BMMNCs) were obtained from patients with IgM MGUS (n=8), WM (n=8) as well as age-matched healthy donors (HD; n=5). Cite-Seq, single cell mass cytometry, BCR sequencing, and exome sequencing was performed on the samples. BMMNCs were also used for functional assays to test T cell reactivity to tumor.

1b. BMMNCs from HD(n=5), MGUS (n=8) and WM (n=8) were stained with metal conjugated antibodies. Data was analyzed using Cytobank software. Figure shows tSNE plots of concatenated live cell gated data from mass cytometry analysis. Concatenations were done with equal cell numbers of cells from each donor. The overlay plots show differences in different immune cell subsets including differences in B cell subsets (naïve B cells (brown) and $CXCR5^{neg}$ B cells in MGUS and WM (orange), myeloid/monocyte population (green) and T cells (pink) in patients with MGUS (pink).

1c. BMMNCs from HD (n=4), MGUS (n=7) and WM (n=7) were labeled with total seqC antibodies and processed using the 10X dropseq platform. Figure shows UMAP clustering of 84,128 single BMMNCs based on transcriptome. 42 distinct clusters could be identified, including B/lymphoplasmacytoid (LPC) cells (cluster 3,4,10,14,18,22,29,34,32,37,39), T/NK cells (0,1,6,7,9,11,12,19,20,27,31,33,36,41), Myeloid cells including monocytes and dendritic cell subsets (cluster 2,5,23,25,30) as well as progenitors/precursor cells (cluster 8,13,15,16,21,24,26,28,40).

1d. Feature plots showing surface expression of lineage antibodies CD3, CD56, CD19, CD14 on clusters in 1c.

1e. UMAP clustering of BMMNCs cells by cohort. Figure shows differences in distinct myeloid (cluster 2,5) and B cell populations (cluster 3) in MGUS and WM, as well as in B/LPC populations in MGUS and WM (e.g. clusters 4, 10, 17, 22).

Fig 2. Mass cytometry analysis of changes in B cells

Single cell mass cytometry was performed on BMMNCs from HD (n=5), or patients with IgM MGUS (n=8) and WM (n=8).

2a. tSNE plot gated on CD19+ B cells by cohort. Figure shows differences in different B cell subsets including differences in naïve B cells (green) and switched memory B cells

in MGUS and WM (brown) and in CXCR5^{neg} B cell subset in MGUS and WM patients compared to HD (purple).

2b. Phenotypes of major subpopulations of CD19+ cells in Fig 2a. Heatmap shows expression of cell surface markers (CXCR5, CD27, CD38, CD21, CD22, HLADR, CD40, CXCR4, C-kit, CD79b and CD20), immunoglobulins (IgD, IgM) in the different CD19+ cell populations shown in Fig 2a. Histogram shows expression of CD19 on the various B cell populations and T cells as control.

2c. Differences in CXCR5 positive B cells. Graphs shows percent of CXCR5 positive cells in HD, MGUS and WM. Each dot represents a unique patient.

2d. Differences in naïve B cells. Figure shows naïve B cells as percentage of total B cells in HD, MGUS and WM. Each dot represents a unique patient.

2e. Differences in B cells based on expression of clonal light chain. CD19+ B cells from MGUS and WM were gated based on the presence of clonal Ig light chain (kappa/lambda, depending on the light chain of the M spike). Figure shows differences in B cell populations including in naïve B cells (right panels) and CXCR5+ B cells (left panels) in both clonal light chain positive (clonal LC+) as well as B cells expressing the opposite light chain (termed as non-clonal LC+) in MGUS and WM.

p=0.06, * p<0.05, **p<0.01 Kruskal Wallis.

Fig 3. CITE-seq analysis of differences in CD19+ cells. B cells were identified based on their surface binding of anti-CD19 antibody and antibody based UMAP clustering analysis was performed on the B cell associated antibodies used in CITE-Seq (IgD, CD27, CXCR5, CD138, PD-1, CD10. CD16, CD21, CD79b, CD27, CD38 and CD20). B cells from patients with IgM MGUS (4,039 cells) and WM (13,077 cells) were compared to those from HD (3,403 cells).

3a. UMAP of B cells clustered based on binding to antibodies. Figure shows distribution of B cells in 10 distinct clusters within the three cohorts, showing differences in IgD+CD27- naïve B cells (cluster 2) and in CXCR5^{neg} B cells (cluster 3) in MGUS and WM.

3b. Feature plot showing binding of antibodies against IgD, CD27, CD10 and CXCR5 in B cells within the different clusters.

3c. Proportion of cell clusters by cohort. Bar graph shows distribution of B cells from HD, MGUS and WM within different clusters shown in 3a. MGUS and WM cohorts had decreased proportion of cells in cluster2 and a relative increase in cluster 3 when compared to HD B cells. Graph shows mean+/- SEM. *p<0.05, **p<0.01 Kruskal Wallis.

3d. UMAP of CD19+ B cells based on transcriptome. UMAP clustering analysis was performed on all B cells (n=20,519) from HD, MGUS and WM. This clustering analysis revealed 24 transcriptionally distinct B cell populations.

3e. UMAP of CD19+ B cells based on transcriptome by cohort. Figure shows decline in B cells in cluster 1 in MGUS and WM, while this it is the dominant cluster in HD. Transcriptome based UMAP also revealed unique (patient-specific) B cell clusters (circled) in WM.

3f. Heatmap of patient-specific clusters of clonal B cells.
Comparison of transcriptomes of unique B cell clusters identified in WM revealed distinct patient-specific patterns of gene expression. Data shown in the heatmap includes top 10 differentially expressed genes from each B cell cluster.

Fig 4. Changes in myeloid cells
BMMNCs that did not bind to either anti-CD3 or anti-CD19 antibody from healthy donors (11,160 cells), MGUS (10,376 cells) and WM (11,262 cells) were analyzed using transcriptome based UMAP clustering to evaluate changes in myeloid cells, NK cells and progenitors.

4a. UMAP of CD3-CD19- cells based on transcriptome expression revealed 24 different clusters including myeloid/monocyte clusters (0,1,7,13,14,17), NK cell clusters (5,9,19,22) and progenitor/precursor cell clusters (2,3,4,6,8,10,12,15,16,18, 21,23).

4b. Feature plot showing antibody binding for CD14, CD11c, CD16, CD56 and c-kit as well as hemoglobin expression (by transcript).

4c. UMAP of CD3-CD19- cells by cohort showing decreased proportion of cells in myeloid clusters 0 and 1 in WM.

4d. Proportion of myeloid clusters by cohort
Bar graph showing the proportions of all myeloid clusters in HD, MGUS or WM, with decline in cluster 1 and increase in cluster 7 in MGUS and WM compared to HD.

4e. Volcano plot of genes differentially expressed in myeloid cluster 7. Many of the genes overexpressed in this cluster include those associated with myeloid inflammation.

4f. Pathway analysis of differentially regulated genes enriched in myeloid cluster 7.

Fig 5. Changes in T cells

5a. tSNE plot of changes in T cells by cohort.
BMMNCs from HD, MGUS and WM were analyzed using mass cytometry. CD3 T cells from HD, MGUS and WM cohorts were concatenated using equal numbers of cells from

individual patients and analyzed using t-SNE analysis. Overlay plot shows distribution of T cells across different cohorts with progressive increase in CD8 granzyme+ T cells (red) in MGUS/WM and decline in naïve CD8+ T cells (brown) compared to HD.

5b. Phenotype of major CD4/CD8+ T cell subsets. Heat map shows protein expression of cell surface markers as well as transcription factors/lytic molecules in the different T cell subsets shown in 5a. The CD8+ granzyme+ T cells enriched in WM have a phenotype of $TCF1^{lo}$Tigit+ KLRG1+ T cells.

5c. Proportions of CD8 T cell subsets in HD and WM. Figure shows naïve (TN; CCR7, CD45RO-), effector memory (TEM; CCR7-CD45RO+), central memory (TCM; CCR7-, CD45RO+CD45RA+) and terminal effectors ($T_{ERM\ EFF}$; CCR7-CD45RO-) cells as percent of total CD8 T cells in healthy donors and WM patients. Each dot represents a unique patient. *p<0.05 Kruskall Wallis.

5d. Proportion of granzyme positive CD8 T cells as a proportion of total CD8 T cells in HD, MGUS and WM patients. Each dot represents a unique patient.
*p<0.05 Kruskall Wallis.

5e. Volcano plot of differentially expressed genes in CD8+ T cells in MGUS versus HD based on CITEseq.

5f. Volcano plot of differentially expressed genes in CD8+ T cells WM versus HD based on CITEseq.

Fig 6. Tumor-specific immunity in WM

6a. Detection of WM-specific immunity in freshly isolated T cells from blood/bone marrow. Freshly isolated blood or marrow derived T cells were cultured overnight in the presence of autologous monocyte-derived mature DCs loaded with autologous tumor cells (or unpulsed mature DCs as control). Tumor-specific interferon-γ producing T cells were quantified with an Elispot assay. Each dot represents mean of replicates from an individual patient.

6b. Ex-vivo expansion of WM-specific T cells by autologous tumor-loaded DCs. Blood or bone marrow T cells were stimulated for 10-14 days in the presence of autologous tumor-loaded (or control either unloaded or loaded with CD19- (non-tumor; NT) cells) mature DCs. The presence of tumor-specific T cells was quantified using IFN-γ Elispot with autologous tumor loaded DCs for detection of tumor-specific T cells. Each dot represents mean of replicates from an individual patient.

6c. CD4 T cell depletion: Responder T cells from experiment described in 6b were depleted of CD4+ T cells utilizing magnetic beads, prior to testing for reactivity using tumor-loaded DCs. The presence of tumor-specific T cells was quantified using IFN-γ Elispot. Each dot represents mean of replicates from an individual patient.

6d and 6e. Detection of MYD88 L265P-specific reactivity in WM patients. MNCs were cultured with peptides spanning wild type or mutant MyD88 L265P sequences. The presence of peptide reactivity was assayed based on the detection of IP-10/CXCL-10 in the supernatant. 6d shows data from representative patients with or without MYD88 reactivity. 6e shows summary of MyD88 L265P reactivity in all patients and healthy donors studied.

Fig 7. Exome sequencing of early B cell progenitors and proposed model for WM evolution.

7a-7c. MYD88 mutations are present throughout the hematopoietic system.

7a: Variant allele fractions (VAF) in CD19+CD10- cells (y-axis) as compared to CD19+CD10+ pre-B cells (x-axis; left), CD19-CD34- (middle), and CD19-CD34+ (right) cells. All variant alleles are shown and those with significantly different allelic fractions (FDR ≤0.05) between any two cell fractions are shown in red with select genes labelled in colored triangles.
7b: Venn diagrams of variants shared and distinct between cell fractions.
7c: VAF of $MYD88^{L265P}$ across CD19-CD34+, CD19-CD34-, CD19+CD10+, and CD19+CD10- cell fractions.

7d. Proposed model for WM development: Acquisition of MYD88 mutation in hematopoietic progenitors is an early event in the origin of WM and associated with several changes in immune microenvironment including increase in extrafollicular B cells, myeloid inflammation and alteration in immune function that begin as early as MGUS stage. B cell clone emerges in this milieu and undergoes progressive growth and evolution in the WM stage. Host immune system mediates tumor-specific recognition of the clone but undergoes immune exhaustion over time.

Figure 1

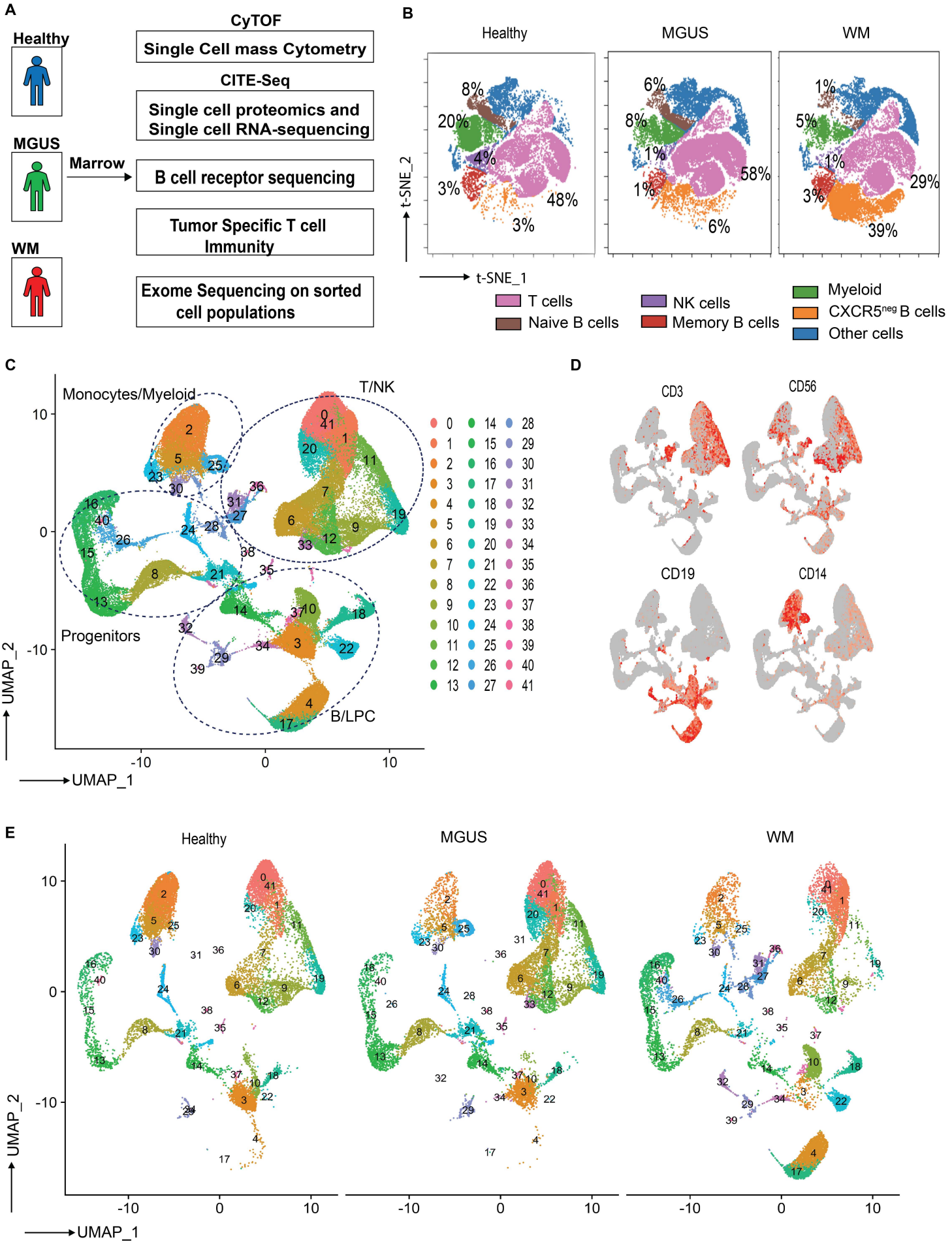

A
Healthy
MGUS
WM
Marrow
CyTOF
Single Cell mass Cytometry
CITE-Seq
Single cell proteomics and Single cell RNA-sequencing
B cell receptor sequencing
Tumor Specific T cell Immunity
Exome Sequencing on sorted cell populations
B
Healthy
MGUS
WM
t-SNE_2
t-SNE_1
T cells
Naive B cells
NK cells
Memory B cells
Myeloid
CXCR5neg B cells
Other cells
C
Monocytes/Myeloid
T/NK
Progenitors
B/LPC
UMAP_2
UMAP_1
D
CD3
CD56
CD19
CD14
E
Healthy
MGUS
WM
UMAP_2
UMAP_1

Figure 2

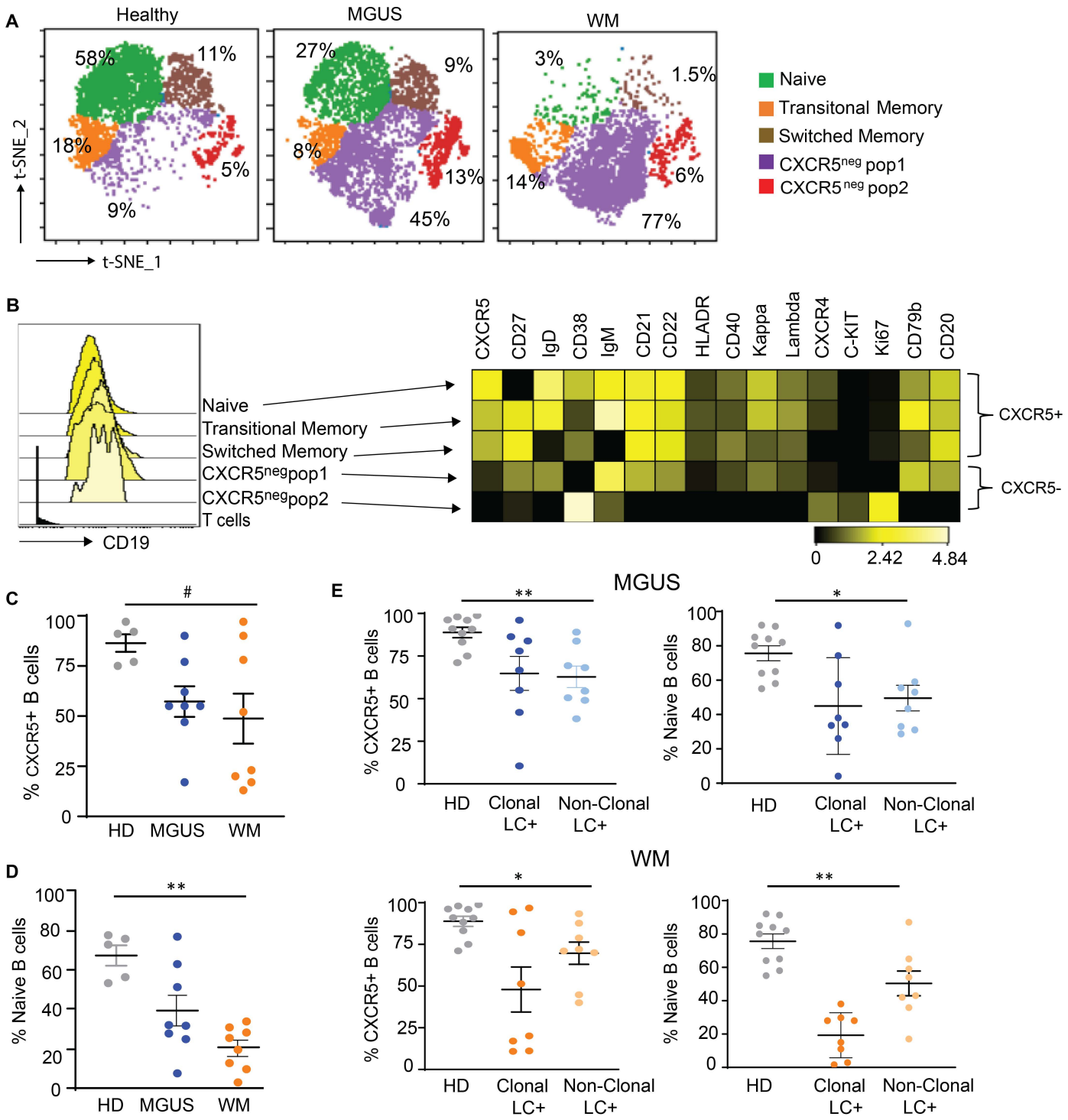

A
Healthy
MGUS
WM
58%
11%
18%
5%
9%
27%
9%
8%
13%
45%
3%
1.5%
14%
6%
77%
t-SNE_2
t-SNE_1
Naive
Transitonal Memory
Switched Memory
CXCR5neg pop1
CXCR5 neg pop2
B
CD19
Naive
Transitional Memory
Switched Memory
CXCR5negpop1
CXCR5negpop2
T cells
CXCR5
CD27
IgD
CD38
IgM
CD21
CD22
HLADR
CD40
Kappa
Lambda
CXCR4
C-KIT
Ki67
CD79b
CD20
CXCR5+
CXCR5-
0
2.42
4.84
C
% CXCR5+ B cells
HD
MGUS
WM
D
% Naive B cells
HD
MGUS
WM
E
MGUS
% CXCR5+ B cells
% Naive B cells
HD
Clonal LC+
Non-Clonal LC+
WM
% CXCR5+ B cells
% Naive B cells
HD
Clonal LC+
Non-Clonal LC+

Figure 3

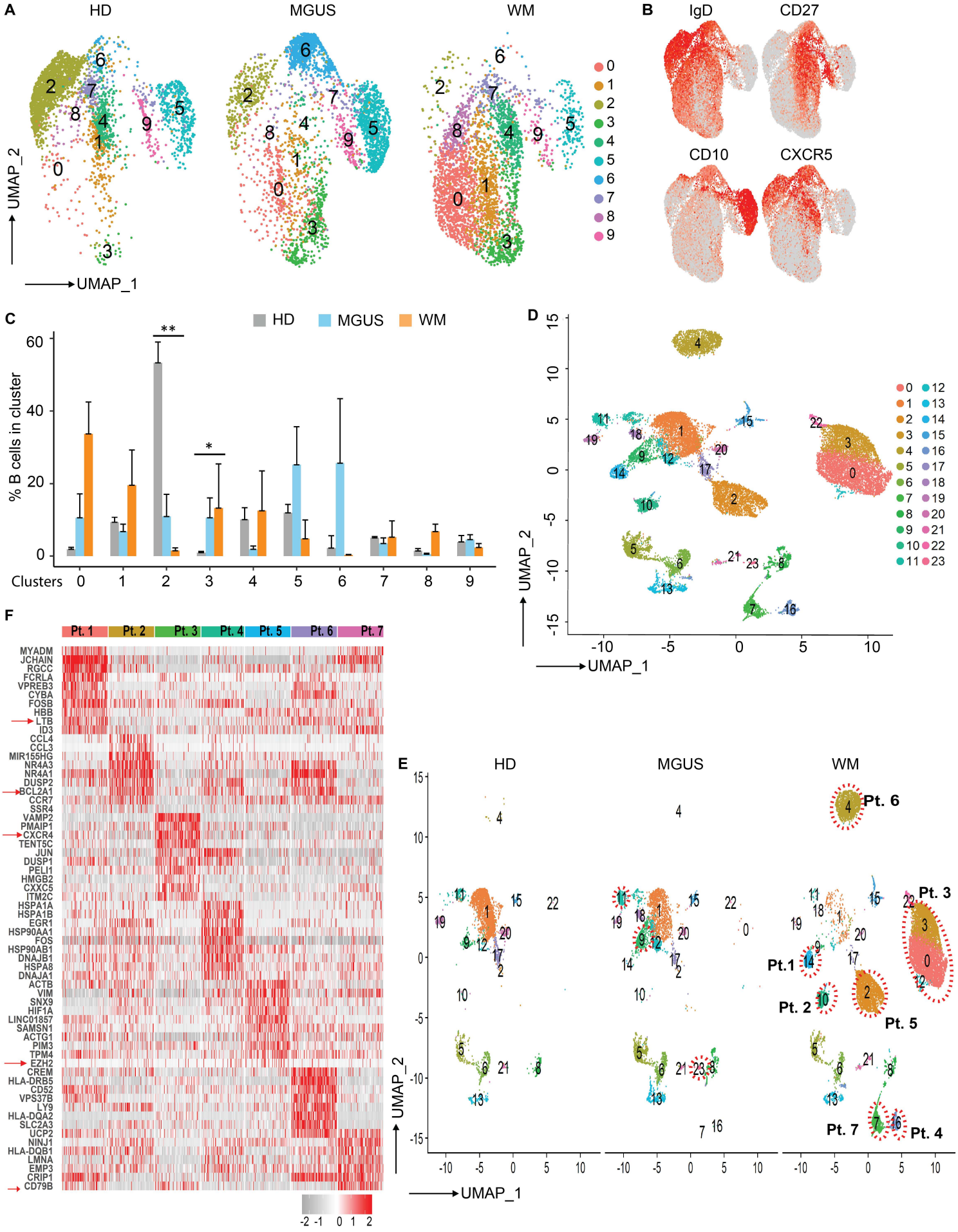

A
HD
MGUS
WM
UMAP_2
UMAP_1
B
IgD
CD27
CD10
CXCR5
C
HD
MGUS
WM
% B cells in cluster
Clusters
D
UMAP_2
UMAP_1
E
HD
MGUS
WM
Pt. 6
Pt. 3
Pt.1
Pt. 2
Pt. 5
Pt. 7
Pt. 4
UMAP_2
UMAP_1
F
Pt. 1
Pt. 2
Pt. 3
Pt. 4
Pt. 5
Pt. 6
Pt. 7
MYADM
JCHAIN
RGCC
FCRLA
VPREB3
CYBA
FOSB
HBB
LTB
ID3
CCL4
CCL3
MIR155HG
NR4A3
NR4A1
DUSP2
BCL2A1
CCR7
SSR4
VAMP2
PMAIP1
CXCR4
TENT5C
JUN
DUSP1
PELI1
HMGB2
CXXC5
ITM2C
HSPA1A
HSPA1B
EGR1
HSP90AA1
FOS
HSP90AB1
DNAJB1
HSPA8
DNAJA1
ACTB
VIM
SNX9
HIF1A
LINC01857
SAMSN1
ACTG1
PIM3
TPM4
EZH2
CREM
HLA-DRB5
CD52
VPS37B
LY9
HLA-DQA2
SLC2A3
UCP2
NINJ1
HLA-DQB1
LMNA
EMP3
CRIP1
CD79B

Figure 4

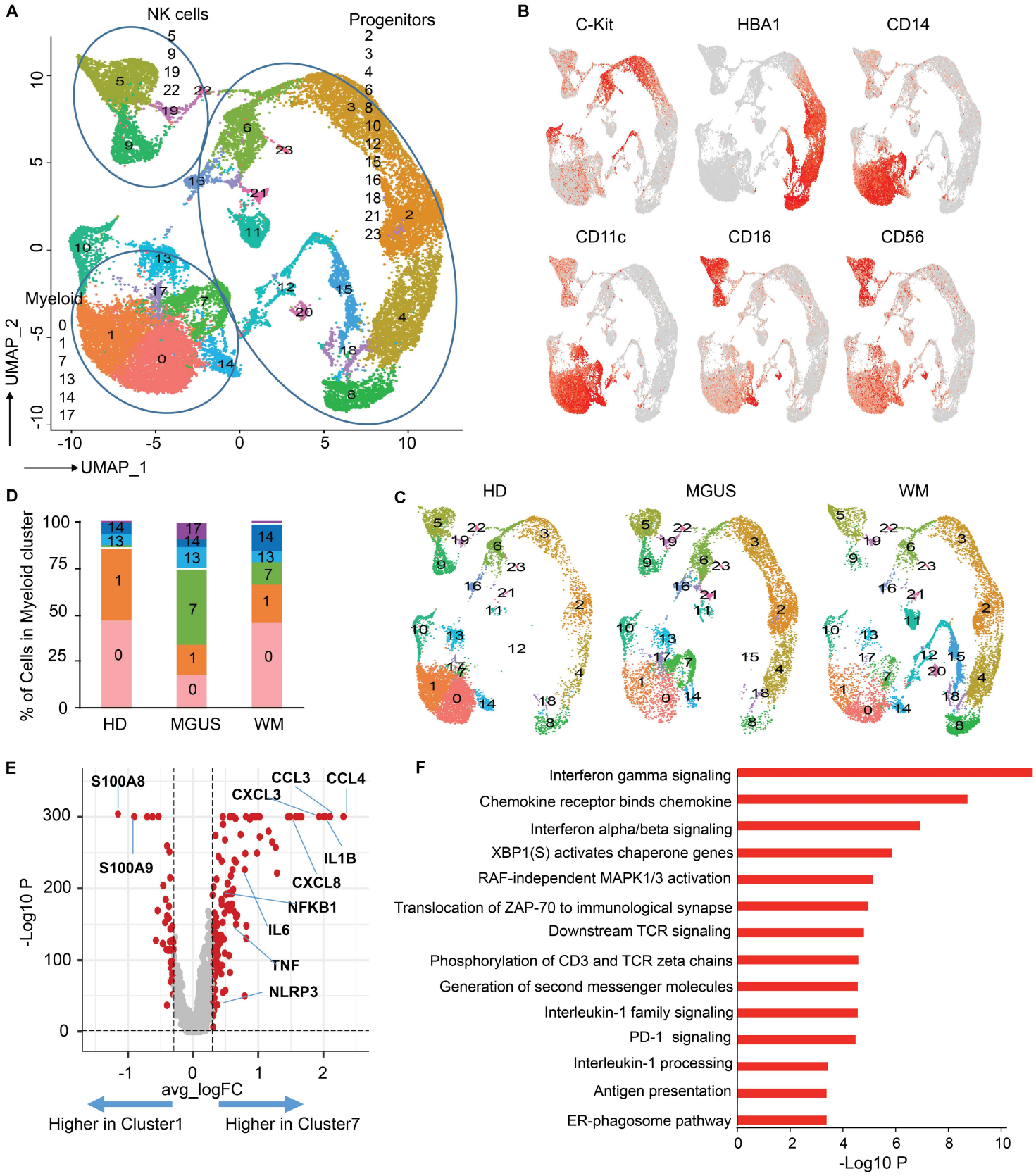

A
NK cells
Progenitors
Myeloid
UMAP_2
UMAP_1
B
C-Kit
HBA1
CD14
CD11c
CD16
CD56
C
HD
MGUS
WM
D
% of Cells in Myeloid cluster
HD
MGUS
WM
E
S100A8
S100A9
CCL3
CCL4
CXCL3
IL1B
CXCL8
NFKB1
IL6
TNF
NLRP3
-Log10 P
avg_logFC
Higher in Cluster1
Higher in Cluster7
F
Interferon gamma signaling
Chemokine receptor binds chemokine
Interferon alpha/beta signaling
XBP1(S) activates chaperone genes
RAF-independent MAPK1/3 activation
Translocation of ZAP-70 to immunological synapse
Downstream TCR signaling
Phosphorylation of CD3 and TCR zeta chains
Generation of second messenger molecules
Interleukin-1 family signaling
PD-1 signaling
Interleukin-1 processing
Antigen presentation
ER-phagosome pathway
-Log10 P

Figure 5

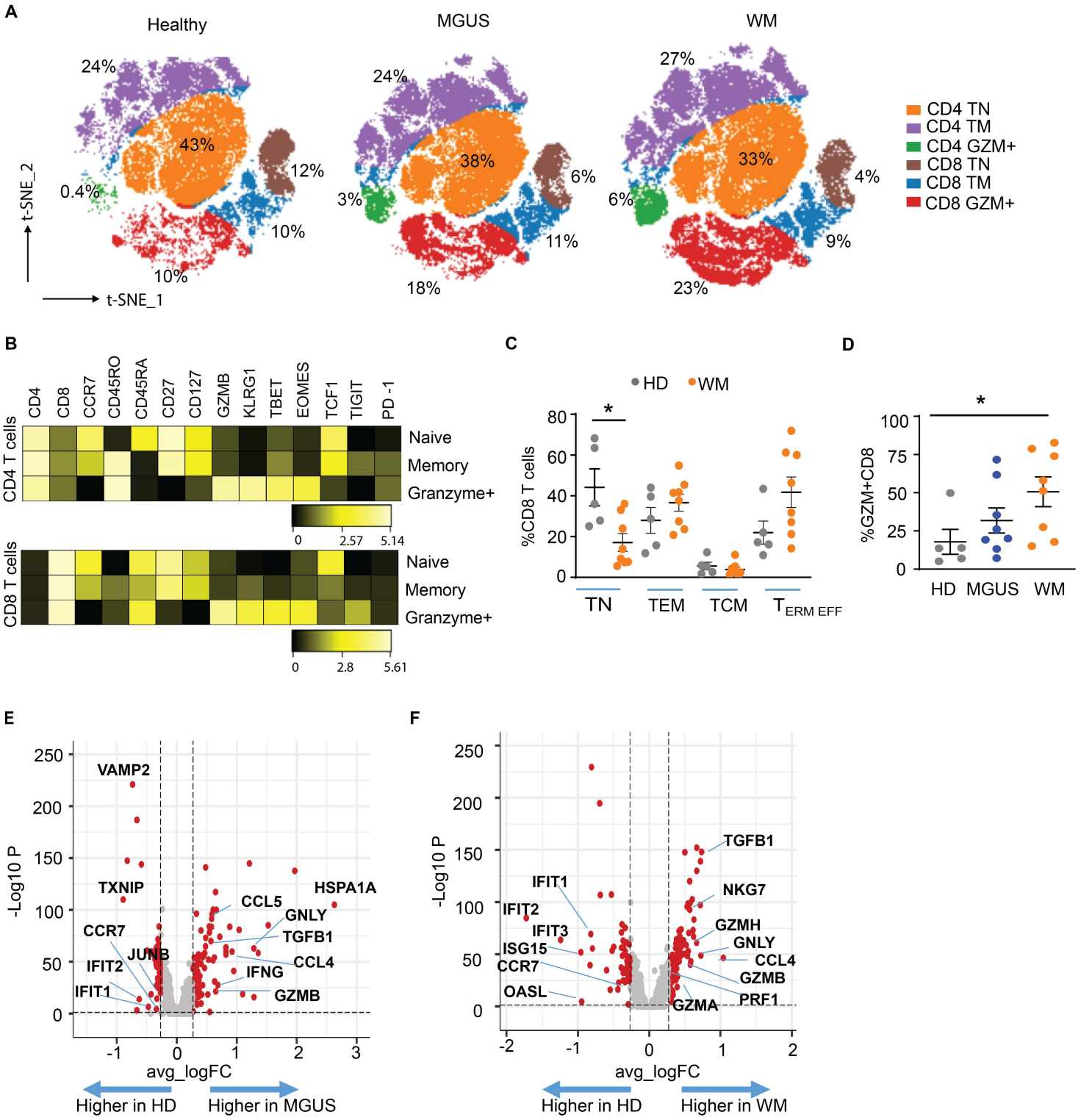

A
Healthy
MGUS
WM
24%
43%
12%
0.4%
10%
10%
24%
38%
6%
3%
11%
18%
27%
33%
4%
6%
9%
23%
t-SNE_2
t-SNE_1
CD4 TN
CD4 TM
CD4 GZM+
CD8 TN
CD8 TM
CD8 GZM+
B
CD4
CD8
CCR7
CD45RO
CD45RA
CD27
CD127
GZMB
KLRG1
TBET
EOMES
TCF1
TIGIT
PD -1
CD4 T cells
Naive
Memory
Granzyme+
0
2.57
5.14
CD8 T cells
Naive
Memory
Granzyme+
0
2.8
5.61
C
HD
WM
%CD8 T cells
TN
TEM
TCM
T ERM EFF
D
%GZM+CD8
HD
MGUS
WM
E
-Log10 P
VAMP2
TXNIP
CCR7
JUNB
IFIT2
IFIT1
CCL5
GNLY
TGFB1
IFNG
CCL4
GZMB
HSPA1A
avg_logFC
Higher in HD
Higher in MGUS
F
-Log10 P
IFIT1
IFIT2
IFIT3
ISG15
CCR7
OASL
TGFB1
NKG7
GZMH
GNLY
CCL4
GZMB
PRF1
GZMA
avg_logFC
Higher in HD
Higher in WM

Fig 6

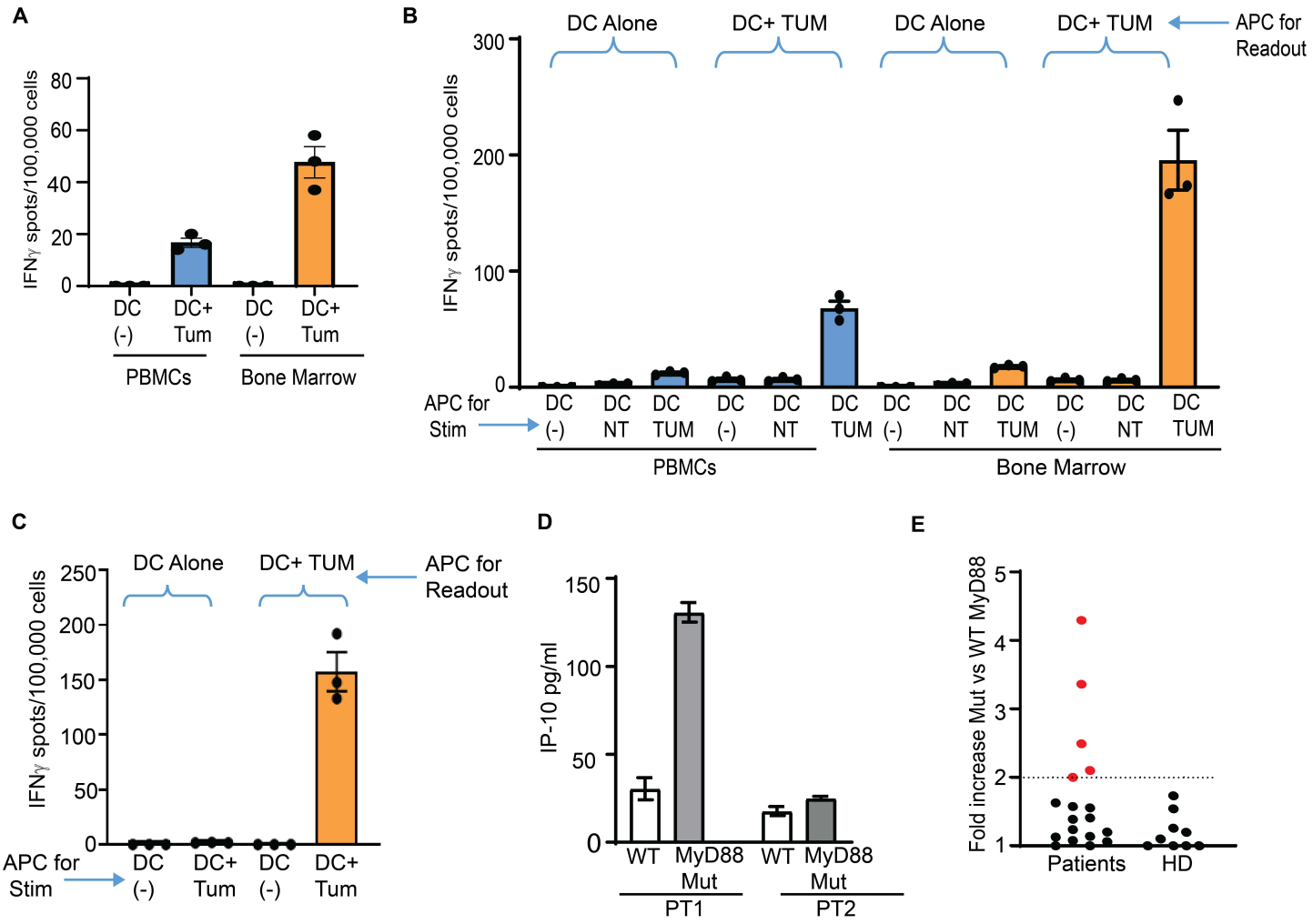
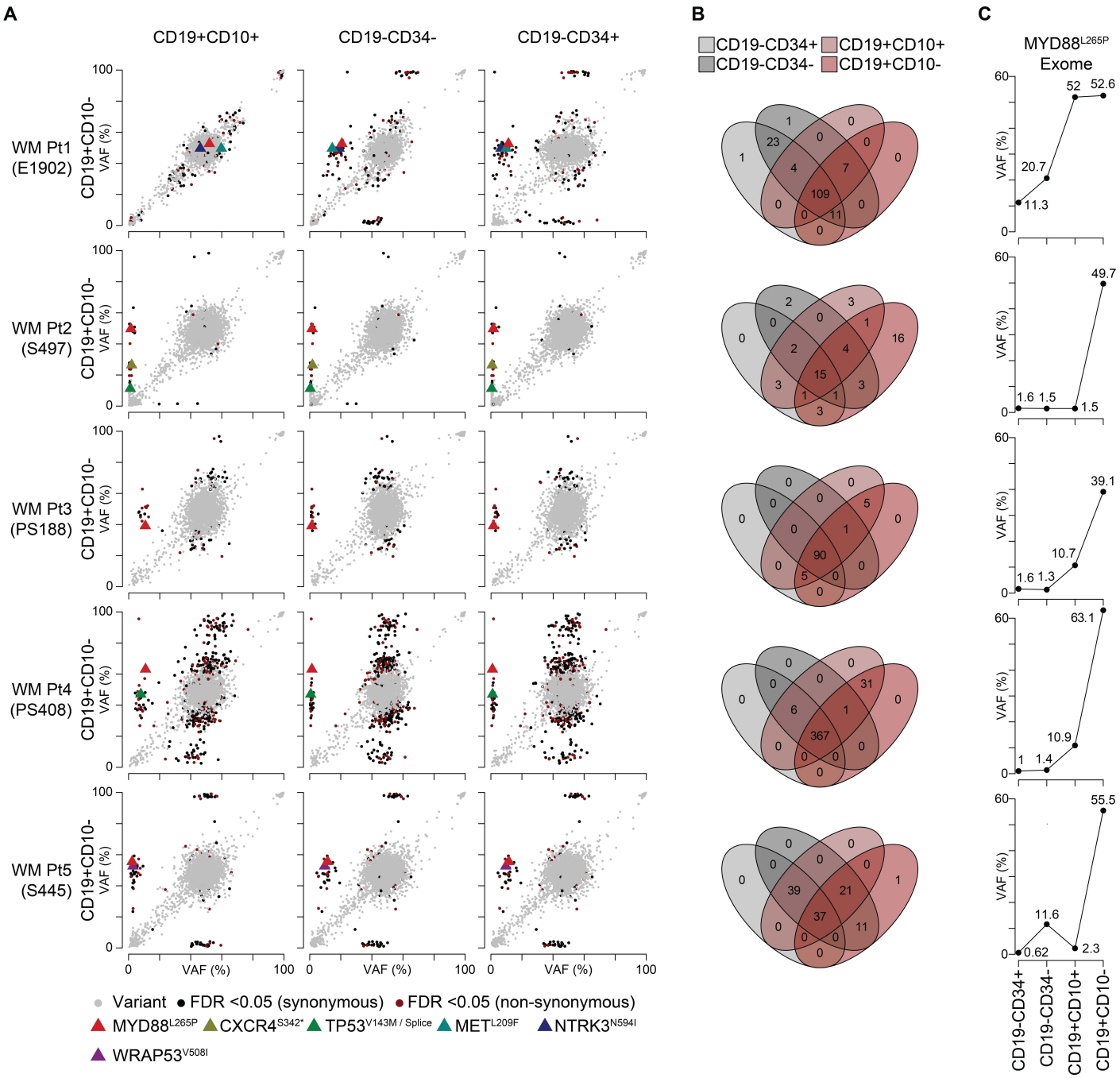
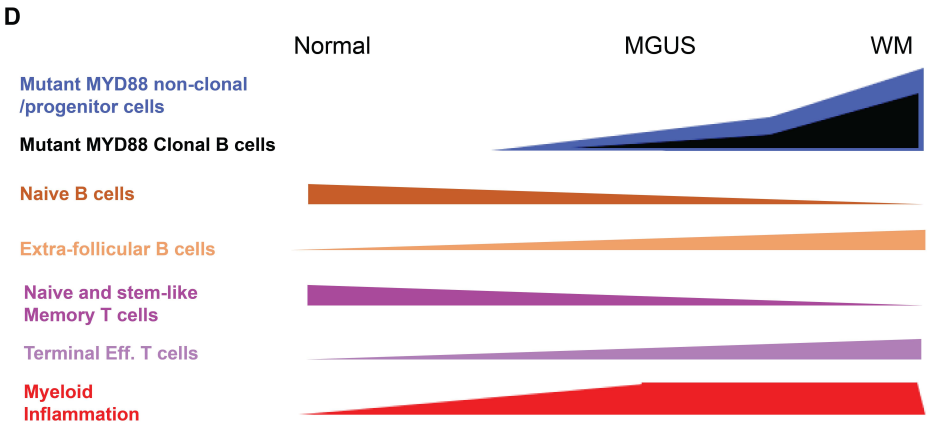

A
IFNγ spots/100,000 cells
80
60
40
20
0
DC (-)
DC+ Tum
DC (-)
DC+ Tum
PBMCs
Bone Marrow
B
DC Alone
DC+ TUM
DC Alone
DC+ TUM
APC for Readout
IFNγ spots/100,000 cells
300
200
100
0
APC for Stim
DC (-)
DC NT
DC TUM
DC (-)
DC NT
DC TUM
DC (-)
DC NT
DC TUM
DC (-)
DC NT
DC TUM
PBMCs
Bone Marrow
C
DC Alone
DC+ TUM
APC for Readout
IFNγ spots/100,000 cells
250
200
150
100
50
0
APC for Stim
DC (-)
DC+ Tum
DC (-)
DC+ Tum
D
IP-10 pg/ml
150
100
50
0
WT
MyD88 Mut
WT
MyD88 Mut
PT1
PT2
E
Fold increase Mut vs WT MyD88
5
4
3
2
1
Patients
HD

Fig 7


A
CD19+CD10+
CD19-CD34-
CD19-CD34+
WM Pt1 (E1902)
WM Pt2 (S497)
WM Pt3 (PS188)
WM Pt4 (PS408)
WM Pt5 (S445)
CD19+CD10- VAF (%)
VAF (%)
Variant
FDR <0.05 (synonymous)
FDR <0.05 (non-synonymous)
MYD88L265P
CXCR4S342*
TP53V143M / Splice
METL209F
NTRK3N594I
WRAP53V508I
B
CD19-CD34+
CD19+CD10+
CD19-CD34-
CD19+CD10-
C
MYD88L265P Exome
VAF (%)
CD19-CD34+
CD19-CD34-
CD19+CD10+
CD19+CD10-

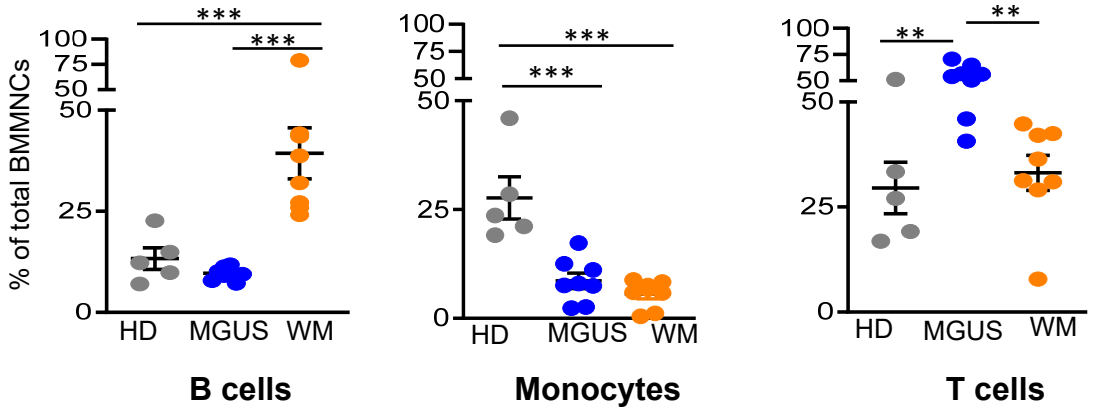

D
Normal
MGUS
WM
Mutant MYD88 non-clonal /progenitor cells
Mutant MYD88 Clonal B cells
Naive B cells
Extra-follicular B cells
Naive and stem-like Memory T cells
Terminal Eff. T cells
Myeloid Inflammation

Supplementary Materials

Kaushal et al.

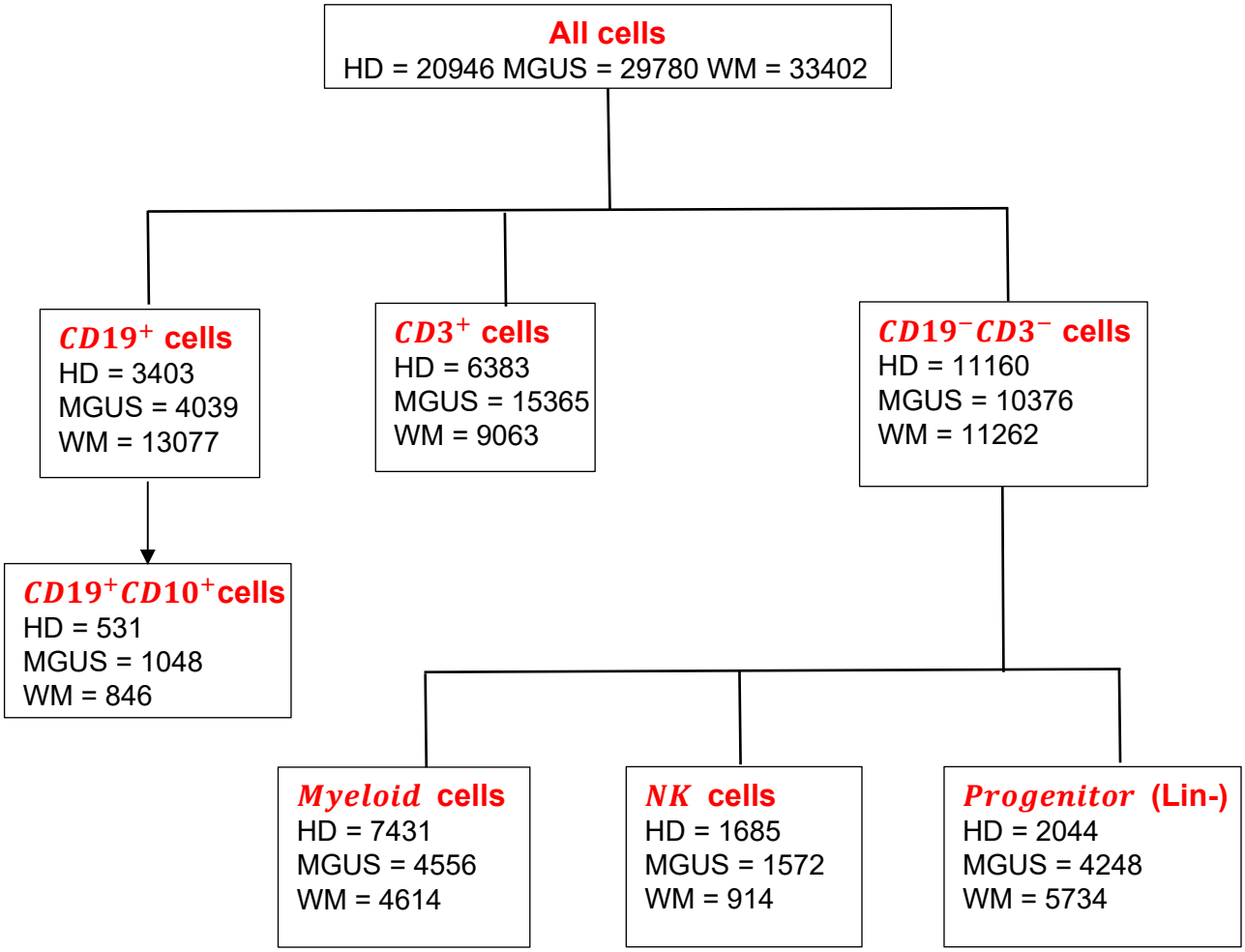
Table of Contents:

Supplementary Figure 1



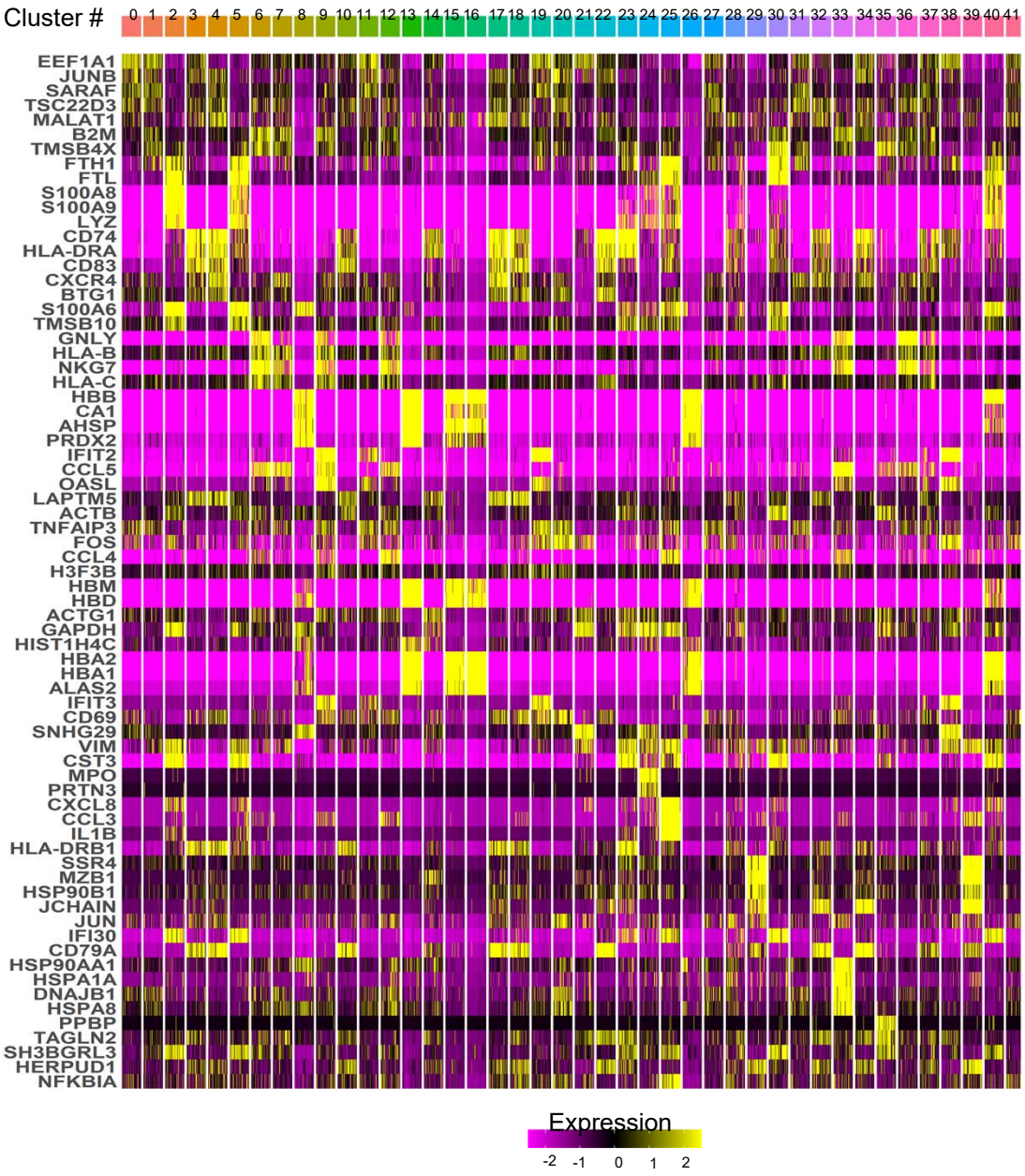
Supplementary Fig 1. Proportion of bone marrow B cells, T cells and monocytes by mass cytometry.
Single cell mass cytometry was performed on bone marrow mononuclear cells (BMMNCs) obtained from patents with IgM MGUS (MGUS; n=8), Waldenstrom macroglobulinemia (WM; n=8) and age-matched healthy donors (HD; n=5). Figure shows CD19+ B cells, CD14+ monocytes and CD3+ T cells as percent of total bone marrow mononuclear cells (BMMNCs) in each sample. Bar graph represents mean +/- SEM. (** $p<0.01$, ***$p<0.001$

Supplementary Figure 2



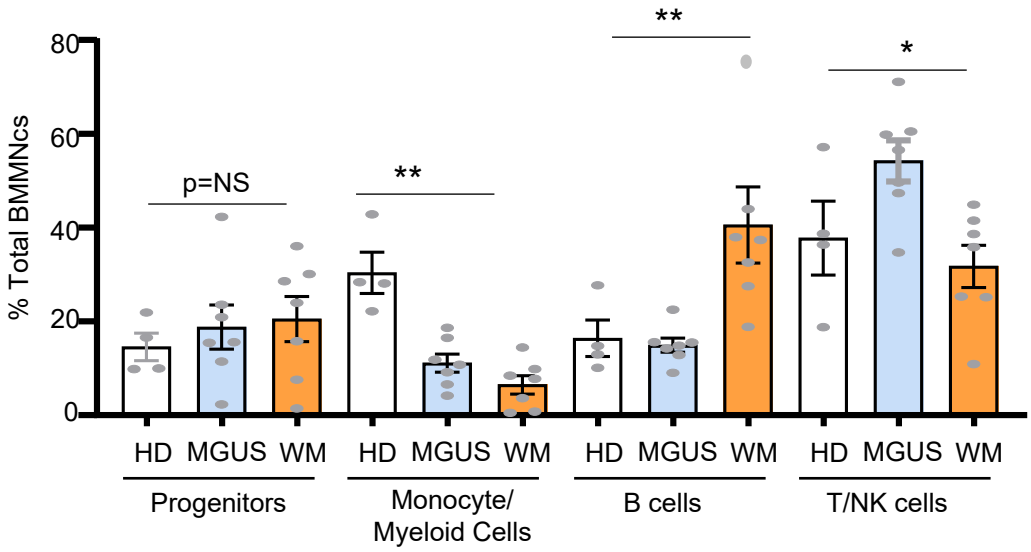
Supplementary Fig 2. Overview of CITE-Seq analysis plan for cell types. Cellular Indexing of Transcriptomes and Epitopes by Sequencing (CITE-Seq) was performed on 84,128 bone marrow cells from patients with MGUS (29,780 cells), WM (33,402 cells) and age matched HD (20,946 cells). Figure shows schema for analysis as well as the number of cells analyzed by CITESeq.

Supplementary Figure 3



Supplementary Fig 3. Cluster identity heatmap.
Transcriptome based UMAP clustering analysis was performed on scRNA on bone marrow mononuclear cells from IgM MGUS (MGUS; n=7), Waldenstrom (WM; n=7) and age-matched healthy donors (HD; n=4). Figure shows top 5 differentially regulated genes for each of the 42 clusters shown in Figure 1c.

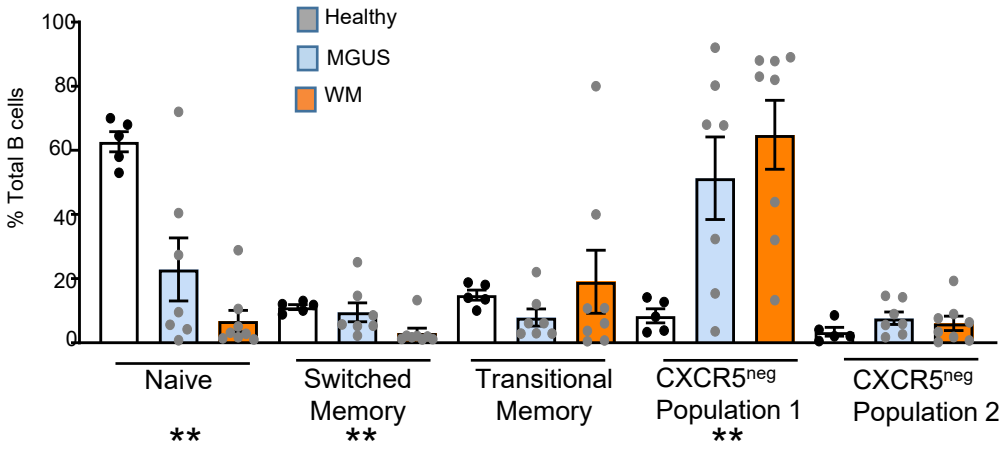
Supplementary Figure 4



Supplementary Fig 4. Annotation of cell proportions based on transcriptome in CITE-Seq.

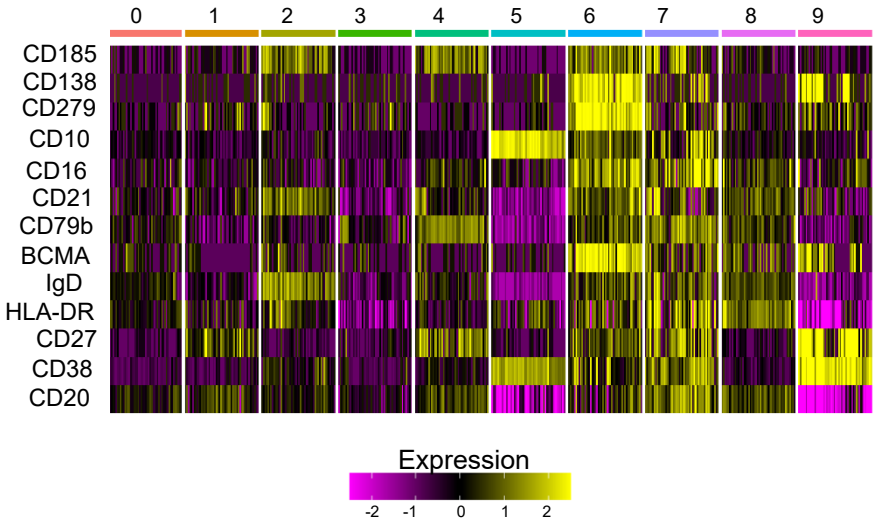
Patient level data showing proportions of T/NK cells, B cells, monocyte/myeloid cells and hematopoietic progenitors in heathy donor bone marrow (n=4) as well as bone marrow from patients with MGUS (n=7) and Waldenstrom (n=7) shown in Figure 1c. Figure shows mean +/- SEM. (* $p<0.05$, ** $p<0.01$)

Supplementary figure 5



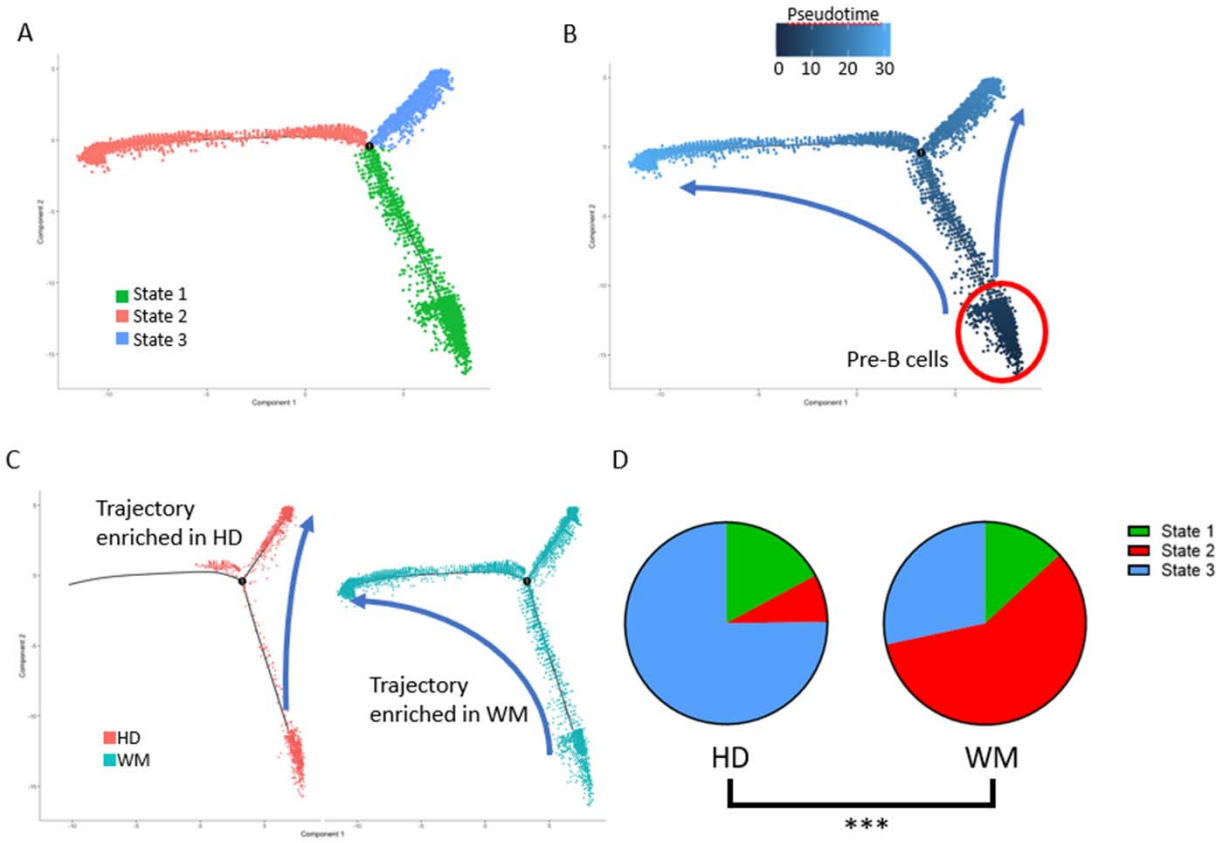
Supplementary Fig 5. Patient level data for B cell subsets shown in Fig2a. Single cell mass cytometry was performed on bone marrow mononuclear cells (BMMNCs) obtained from patents with IgM MGUS (MGUS; n=8), Waldenstrom (WM; n=8) and age-matched healthy donors (HD; n=5). Figure shows patient level data for naïve B cells (IgD+, CD27+), switched memory B cells (IgD-, CD27+, IgM-), transitional memory (IgD+, CD27+, IgM+), CXCR^{neg} population-1 and CXCR5^{neg} population-2 as shown in the tSNE plot in Figure 2a. Figure shows mean +/- SEM. (**p<0.01)

Supplementary Figure 6



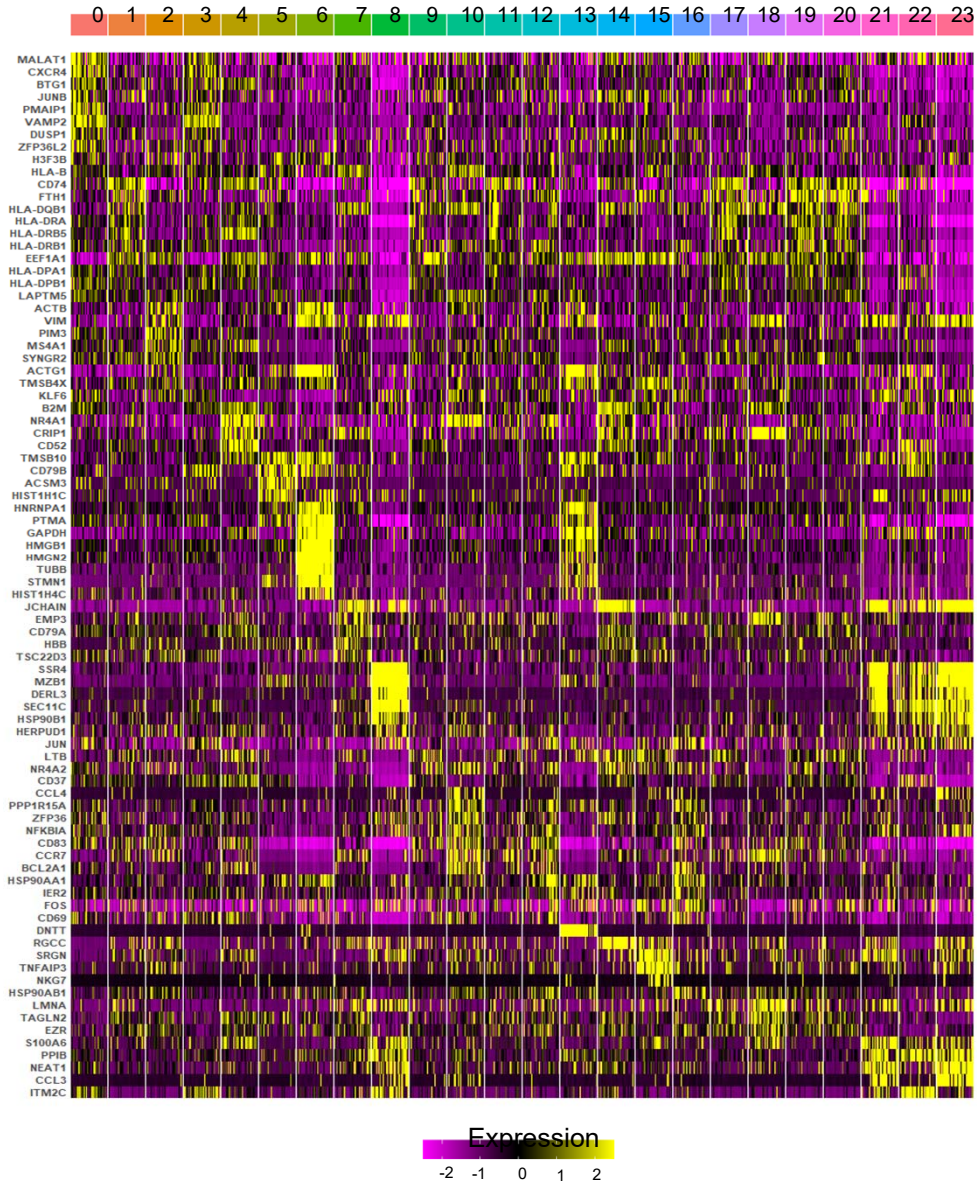
Supplementary Fig 6. Cluster identity heatmap (clustering based on antibody staining) for B cell clusters in Fig 3a.
B/lymphoplasmacytoid cells were identified based on their surface binding of anti-CD19 antibody and antibody based UMAP clustering analysis was performed using B cell-associated antibodies used in CITE-Seq. Cluster identity heatmap of differential antibody expression for the 10 different B cell clusters identified by antibody based UMAP clustering shown in Figure 3a.

Supplementary Figure 7



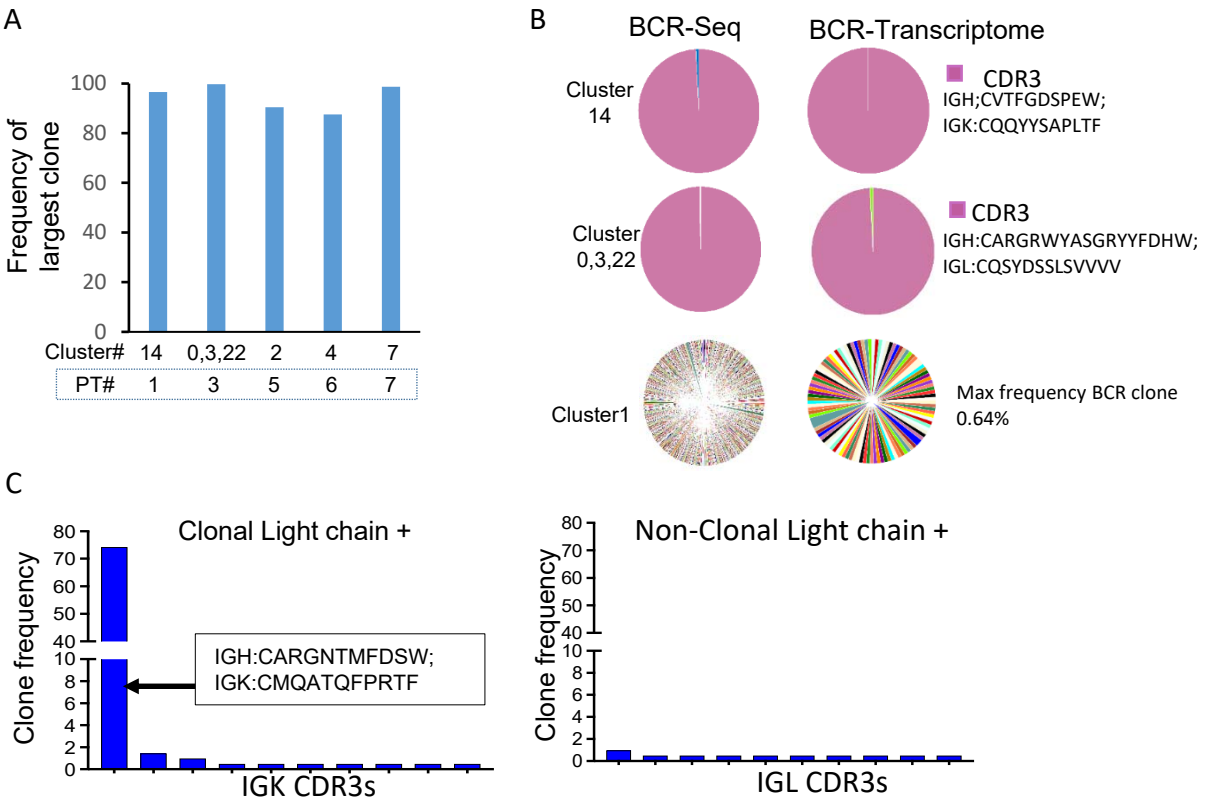
Supplementary Figure 7: Monocle single-cell trajectory analysis reveals differences in B cell trajectories between WM and HD. CD19+ cells from all WM (n=13077 CD19+ cells) and HD (n=3403 CD19+ cells) single-cell RNA sequencing data were isolated, and trajectories were calculated using Monocle. A) Monocle trajectory analysis identified 3 states with one branchpoint. B) Pseudotime analysis identified trajectories from pre-B cells to mature B cells. The pseudotime of 0 corresponds to the pre-B cell population at the beginning of State 1, and later pseudotimes correspond to the two terminal states post-branch point with mature B cells. C) Distribution of cell populations along Monocle trajectories by disease state visualized on the trajectory plots (left=HD, right=WM). D) Proportion of total cells in each state for both HD and WM. The two terminal states had differential distributions, with the left state (State 2) enriched for WM B cells and the right state (State 3) enriched for HD B cells. ***p<0.001, chi square test.

Supplementary Figure 8



Supplementary Fig 8. Cluster identity heatmap of B cells clusters in Fig 3d. Transcriptome based UMAP clustering analysis performed on CD19+ cells (n=20,519) from HD, MGUS and WM revealed 24 transcriptionally distinct B/lymphoplasmacytoid cell populations. Figure shows top 10 differentially regulated genes in each of the 24 distinct clusters shown in figure 3d.

Supplementary Figure 9



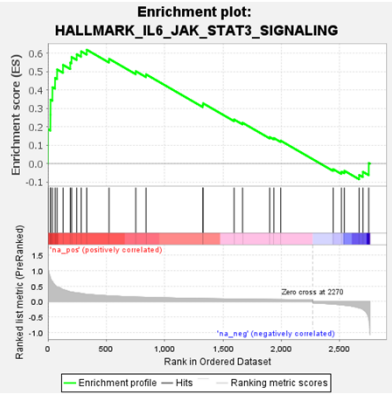
Supplementary Fig 9. B cell receptor analysis of B cell clusters and clonal versus non-clonal IgL+ B cells.
Transcriptome based UMAP clustering on CD19+ cells revealed that clonal B cells from each of the WM patients formed distinct clusters. To access clonal nature of B cells within these distinct clusters, B cell receptor sequencing was performed using the 10X single cell sequencing platform from bone marrow cells of WM patients (n=5) and healthy donors (n=4). The BCR sequences were mapped on the UMAP clusters shown in figure 3D-E. A. Bar graph shows frequency of the largest clone in the distinct clusters. Also shown are the cluster numbers and patient sample ID for the clusters. B. We also imputed BCR sequencing using the 10X VDJ cell ranger pipeline. Pie chart shows frequency as well as sequence of the clone as determined by direct BCR sequencing (BCR-Seq) and VDJ transcriptome analysis (BCR-Transcriptome). Cluster 14, 0/3/22 are clusters from WM patients and cluster 1 is a cluster of non-clonal B cells from HD and WM patients. C. Detection of clonal BCRs in clonal versus non-clonal IgL+ B cells. Bar graph shows the frequency of top25 clones (unique CDR3 regions) identified in the clonal light chain (kappa) expressing B cells as well as non-clonal light chain expressing B cells (lambda) from a WM patient as a representative example.

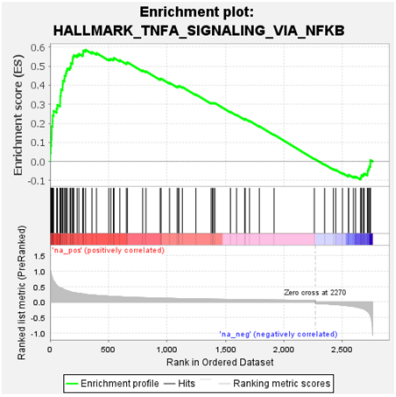
Supplementary Figure 10

A. Top differentially regulated Pathways in mature B cells from WM compared to HD

HALLMARK pathways	NES	FDR q-val
TNFA_SIGNALING_VIA_NFKB	2.15	0.0000
IL2_STAT5_SIGNALING	2.11	0.0010
EPITHELIAL_MESENCHYMAL_TRANSITION	1.94	0.0037
IL6_JAK_STAT3_SIGNALING	1.94	0.0028
HYPOXIA	1.93	0.0024
TGF_BETA_SIGNALING	1.89	0.0030

B. Enrichment Plots





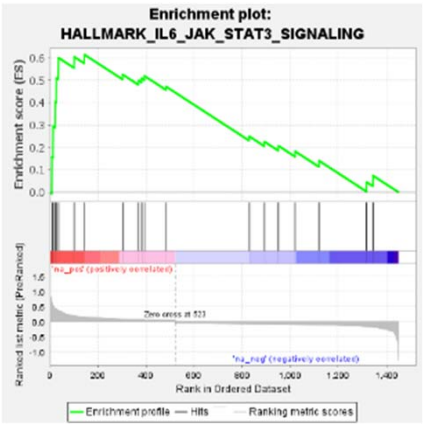
Supplementary Fig 10. WM B cell pathway analysis and enrichment plots
Gene set enrichment analysis(GSEA) was performed using fold change based pre ranked genes from WM mature B cells compared to mature B cells from healthy donors. A. Top differentially up-regulated pathways in mature B cells from WM compared to HD. B. Enrichment plot for TNFa/NFKB and IL6/Jak/STAT3 pathway in WM. NES: Normalized enrichment score based on permutation-based testing from (GSEA).

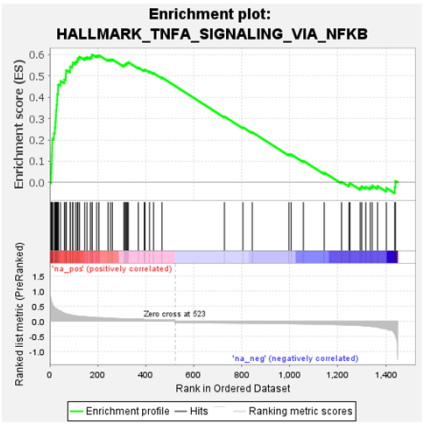
Supplementary Figure 11

A. Pathways upregulated in mature B cells from MGUS compared to HD

HALLMARK Pathway	NES	FDR q-val
TNFA_SIGNALING_VIA_NFKB	2.62	0.0000
EPITHELIAL_MESENCHYMAL_TRANSITION	2.33	0.0000
P53_PATHWAY	2.29	0.0000
HYPOXIA	2.14	0.0015
APOPTOSIS	2.06	0.0045
IL6_JAK_STAT3_SIGNALING	1.98	0.0112
INFLAMMATORY_RESPONSE	1.94	0.0142

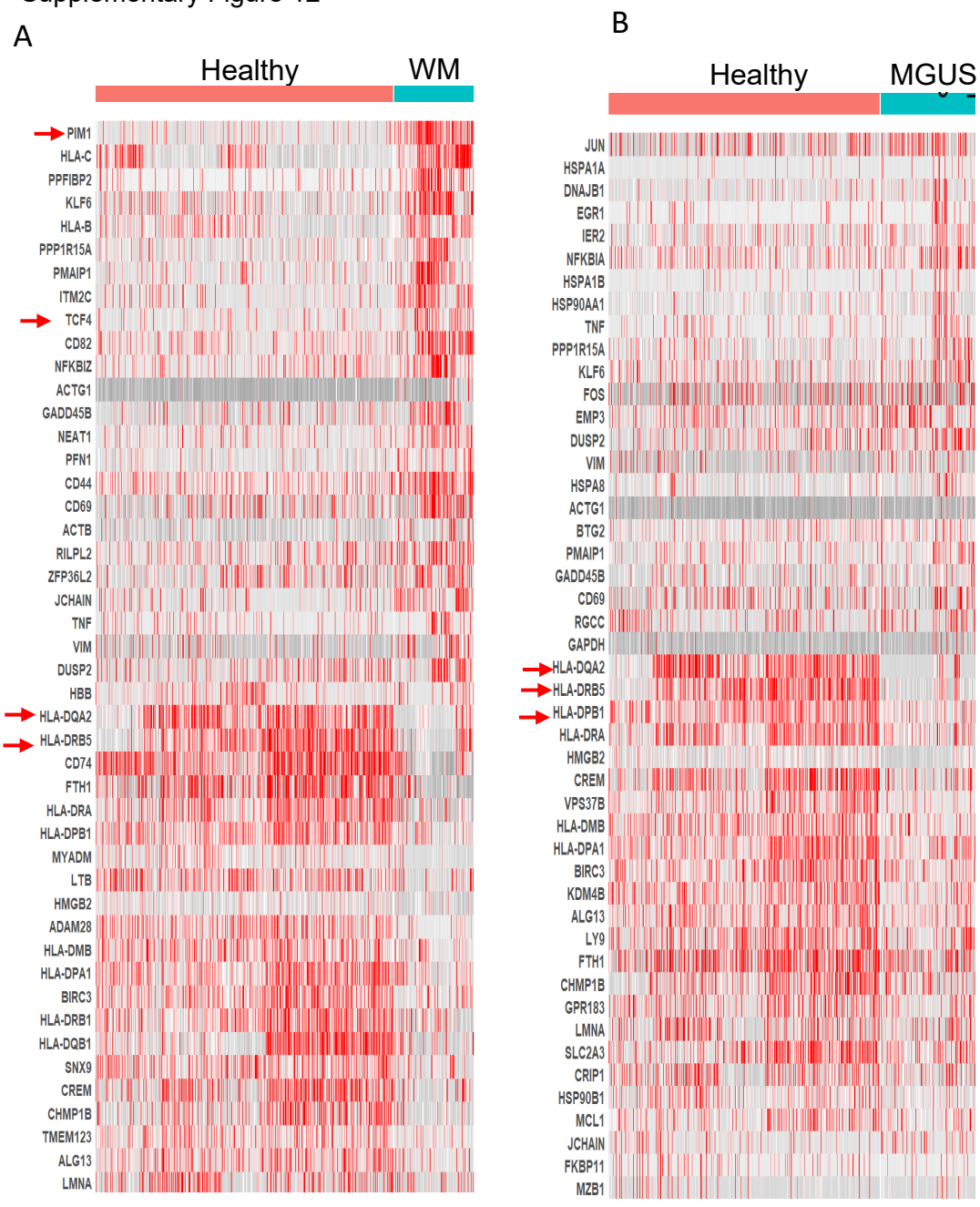
B. Enrichment Plots





Supplementary Fig 11. MGUS B cells pathway analysis and enrichment plots
Gene set enrichment analysis(GSEA) was performed using fold change based pre ranked genes from IgM MGUS mature B cells compared to mature B cells from healthy donors. A. Top differentially up-regulated pathways in mature B cells from MGUS compared to HD. B. Enrichment plot for TNFa/NFKB and IL6/Jak/STAT3 pathway in MGUS. NES: Normalized enrichment score based on permutation-based testing from (GSEA).

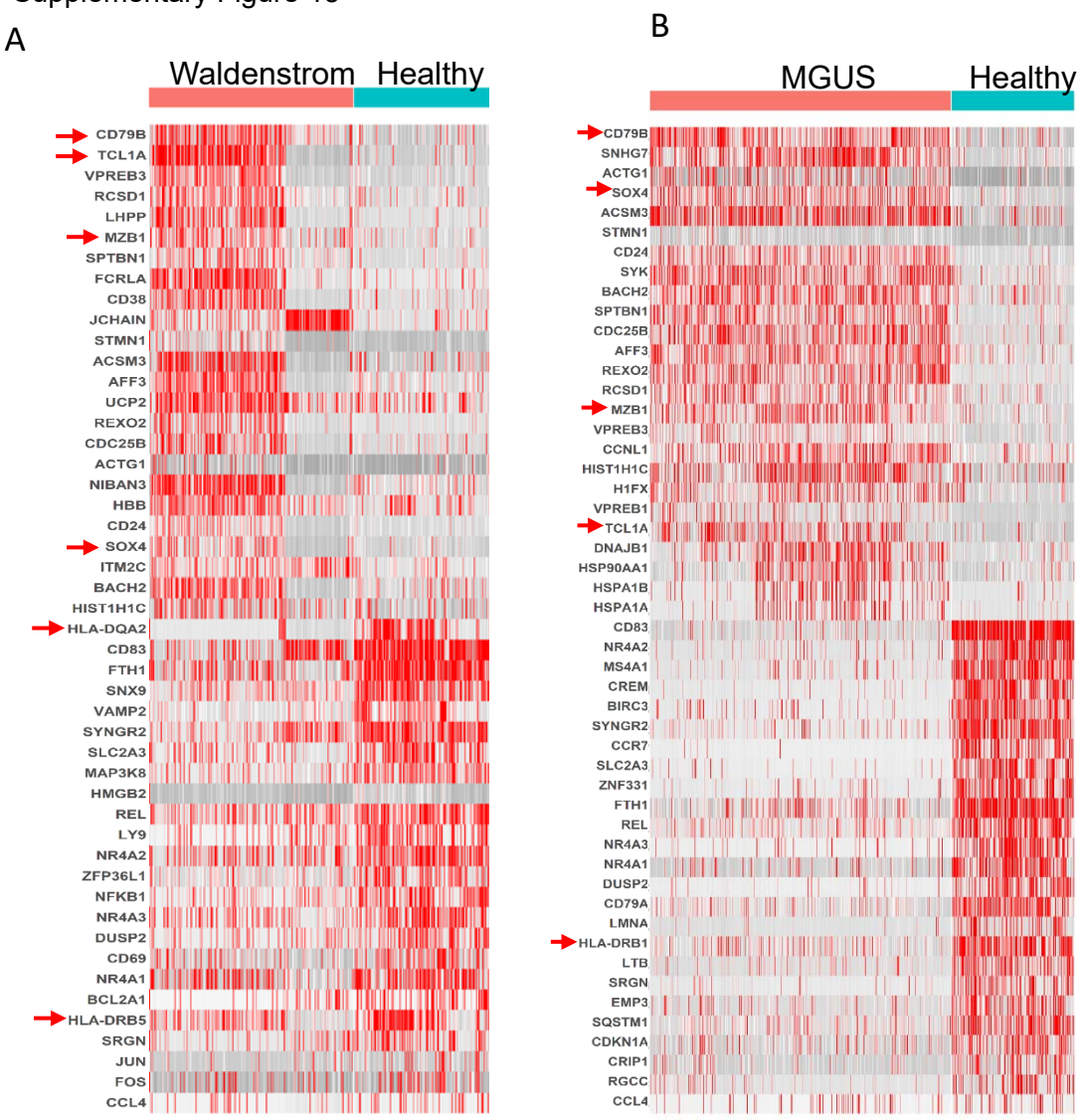
Supplementary Figure 12



Supplementary Fig 12. Aberrant gene expression in non-clonal Ig light chain+ B cells.

A. Heat map of 25 top differentially expressed genes between WM B cells expressing non-clonal light chain and HD B cells. Some of the genes previously implicated in lymphomagenesis are highlighted. B. Heat map showing 25 top differentially expressed genes between MGUS B cells expressing non-clonal light chain and HD B cells. Some of the genes previously implicated in lymphomagenesis are highlighted.

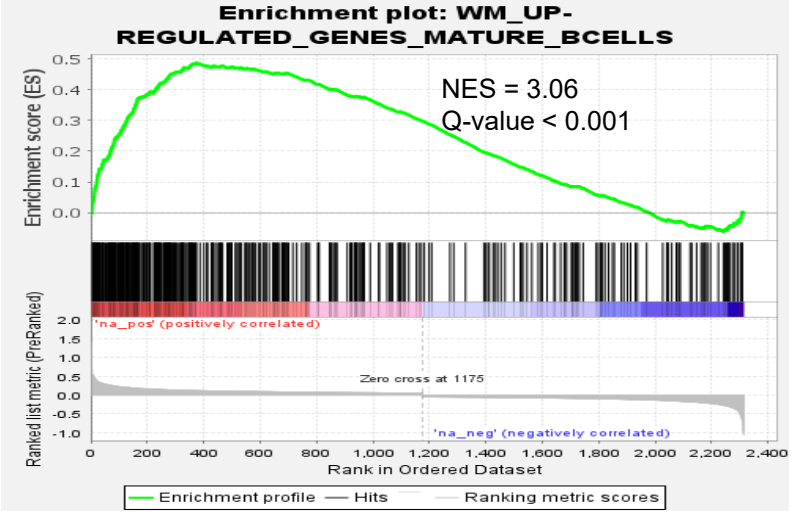
Supplementary Figure 13



Supplementary Fig 13. Heatmap comparing transcriptomes of CD19+CD10+ pre-B cells between MGUS, WM and HD.

A. Heatmap of differentially expressed genes (top 25 genes) in Waldenstrom CD19+CD10+ pre-B cells compared to healthy donor pre B cells B. Heatmap of differentially expressed genes (top 25 genes) in MGUS pre B cells compared to healthy donor pre-B cells.

Supplementary Figure 14



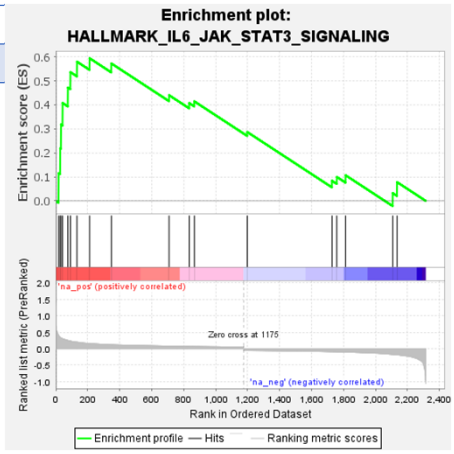
Supplementary Fig 14. Enrichment of B cell derived differential signature in pre-B cells

Genes upregulated in mature B cells from WM patients compared to those from HD (fold change > 1.2 and FDR p-value < 0.05) were used to create a differentially regulated gene set. The plot shows enrichment of this differentially regulated gene set in CD19+CD10+ pre-B cells from WM patients.

Supplementary Figure 15

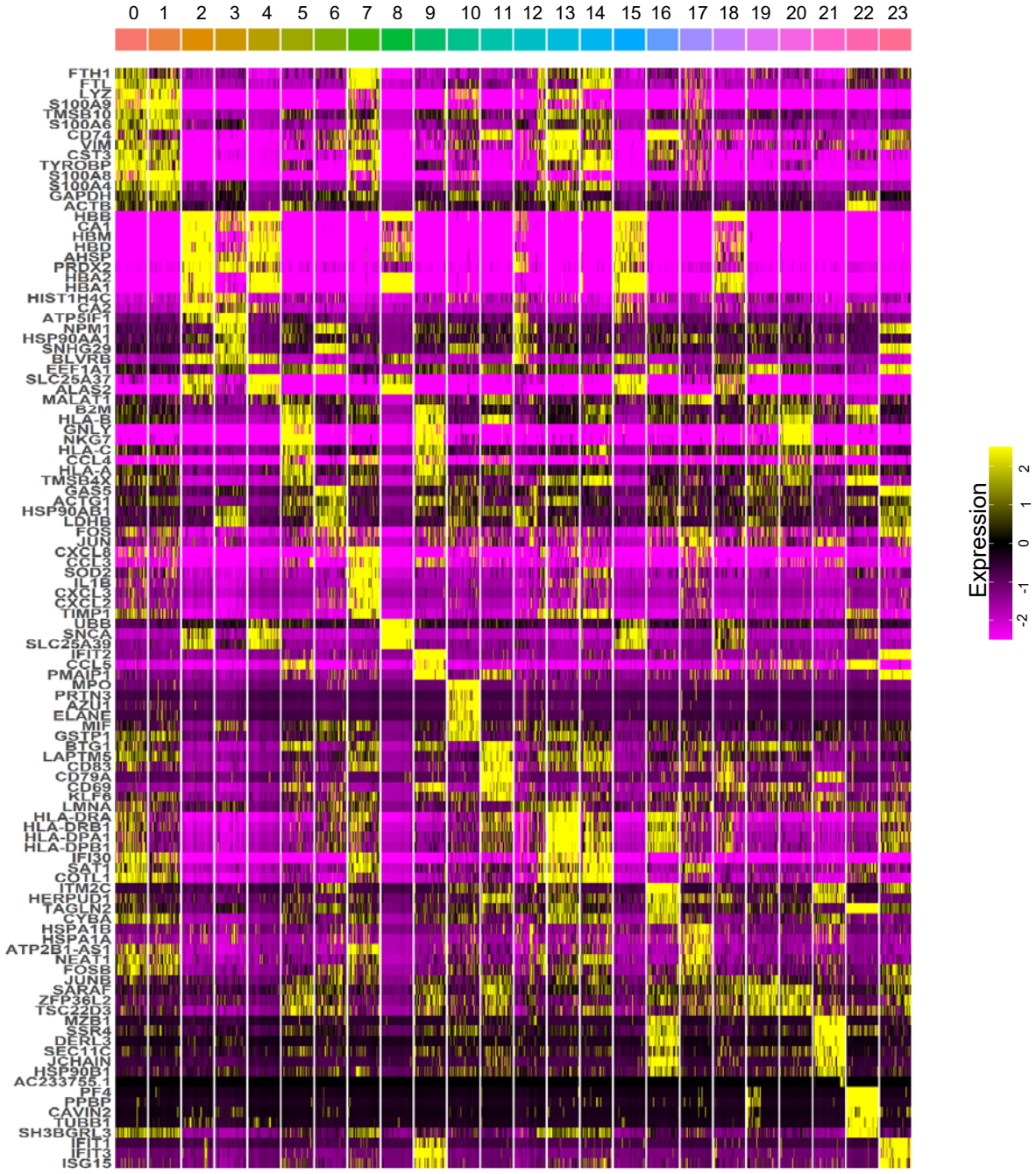
Top differentially regulated Pathways in preB cells from WM compared to HD

HALLMAR Pathway	NES	FDR q-val
IL6_JAK_STAT3_SIGNALING	2.00	0.0635



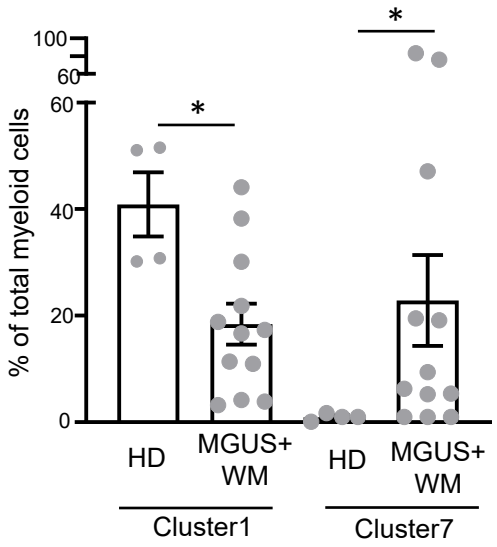
Supplementary Fig 15. Pathway analysis of top differentially regulated pathways in WM versus HD pre-B cells.
CD19+CD10+ pre-B cells from WM patients were compared to HD pre-B cells. Figure shows enrichment plot for IL6/JAK/STAT3 signaling pathway upregulated in WM pre-B cells.

Supplementary Figure 16



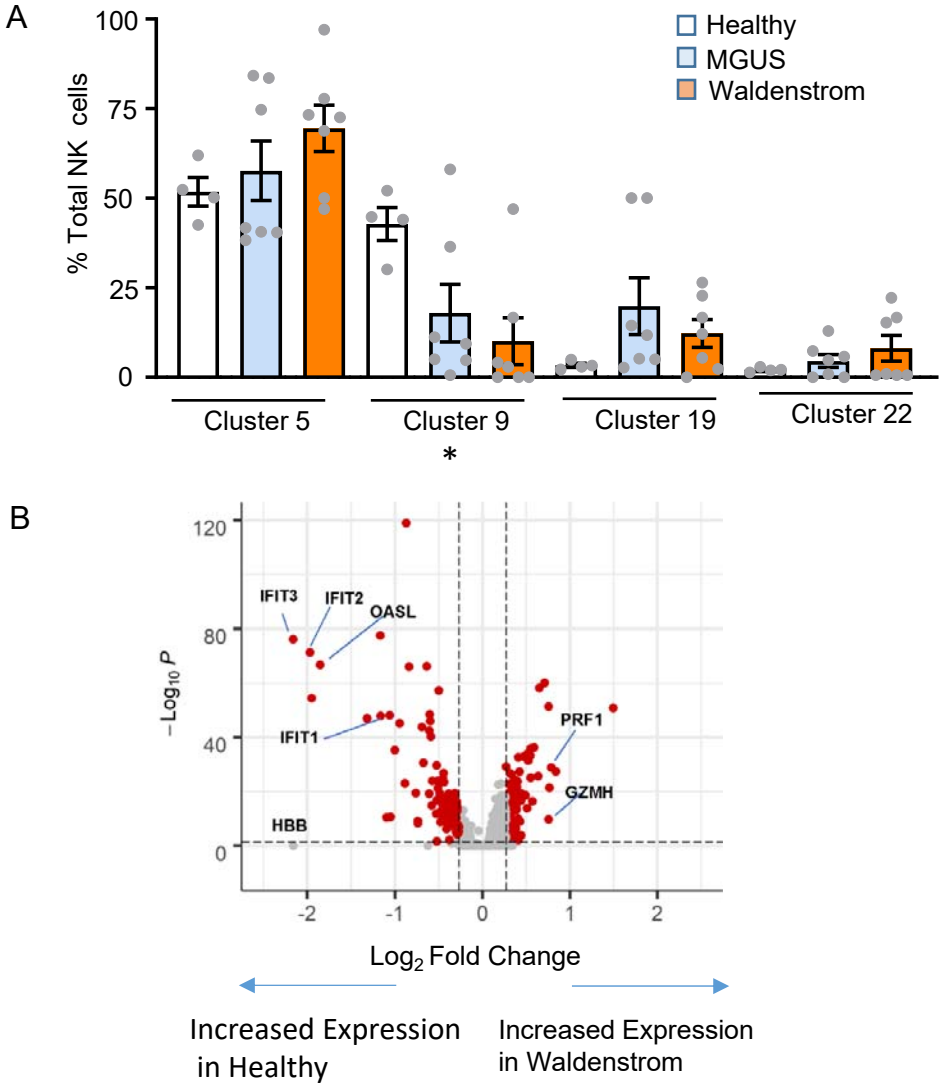
Supplementary Fig 16. Cluster identity map for Fig 4a.
BMMNCs that did not bind to either anti-CD3 or anti-CD19 antibody from healthy donors (11160 cells), MGUS (10376 cells) and WM (11262 cells) were analyzed using transcriptome based UMAP clustering shown in fig4a. Heat map shows top 10 genes differentially regulated in each of the 24 clusters.

Supplementary Figure 17



Supplementary Fig 17. Patient level data for clusters 1 and 7 in Fig4d.
Transcriptome based UMAP clustering of CD3-CD19- BMMNCs cells revealed 6 distinct myeloid clusters, with HD myeloid cells enriched in cluster 1 and WM and MGUS myeloid cells enriched in cluster 7 as shown in fig 4D. Figure shows proportions of myeloid cells in cluster 1 and 7. Each dot is a single sample. Figure shows mean +/-SEM (*$p<0.05$)

Supplementary Figure 18

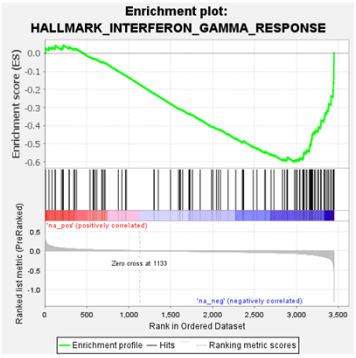


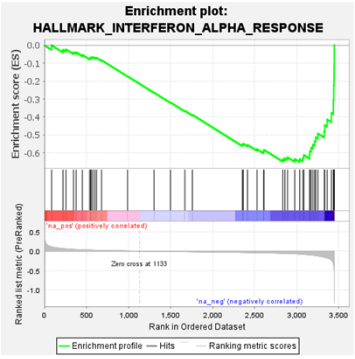
Supplementary Fig 18. Cluster distribution of NK clusters in main Fig 4a by cohort and volcano plot for NK cells HD HD versus WM.
A. Bar graphs show distribution of NK cells within the four different clusters identified by transcriptome based UMAP clustering shown in figure 4a and 4c. The figure shows distribution for NK cells in healthy donor (HD), MGUS and Waldennstrom (WM) bone marrow. Each dot represents a distinct sample. Figure shows mean +/-SEM (*p<0.05). B. The volcano plot shows differentially regulated genes between WM NK cells and NK cells from HD.

Supplementary Figure 19

A. Pathways downregulated in NK cells from MGUS patients compared to HD.

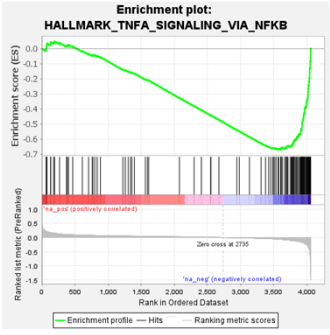
HALLMARK Pathway	NES	FDR q-val
INTERFERON_GAMMA_RESPONSE	-1.90	0.0026
INTERFERON_ALPHA_RESPONSE	-1.88	0.0018
G2M_CHECKPOINT	-1.72	0.0232

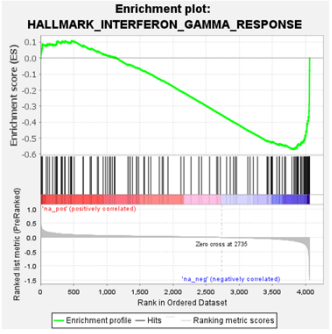




B. Pathways downregulated in NK cells from WM patients compared to HD.

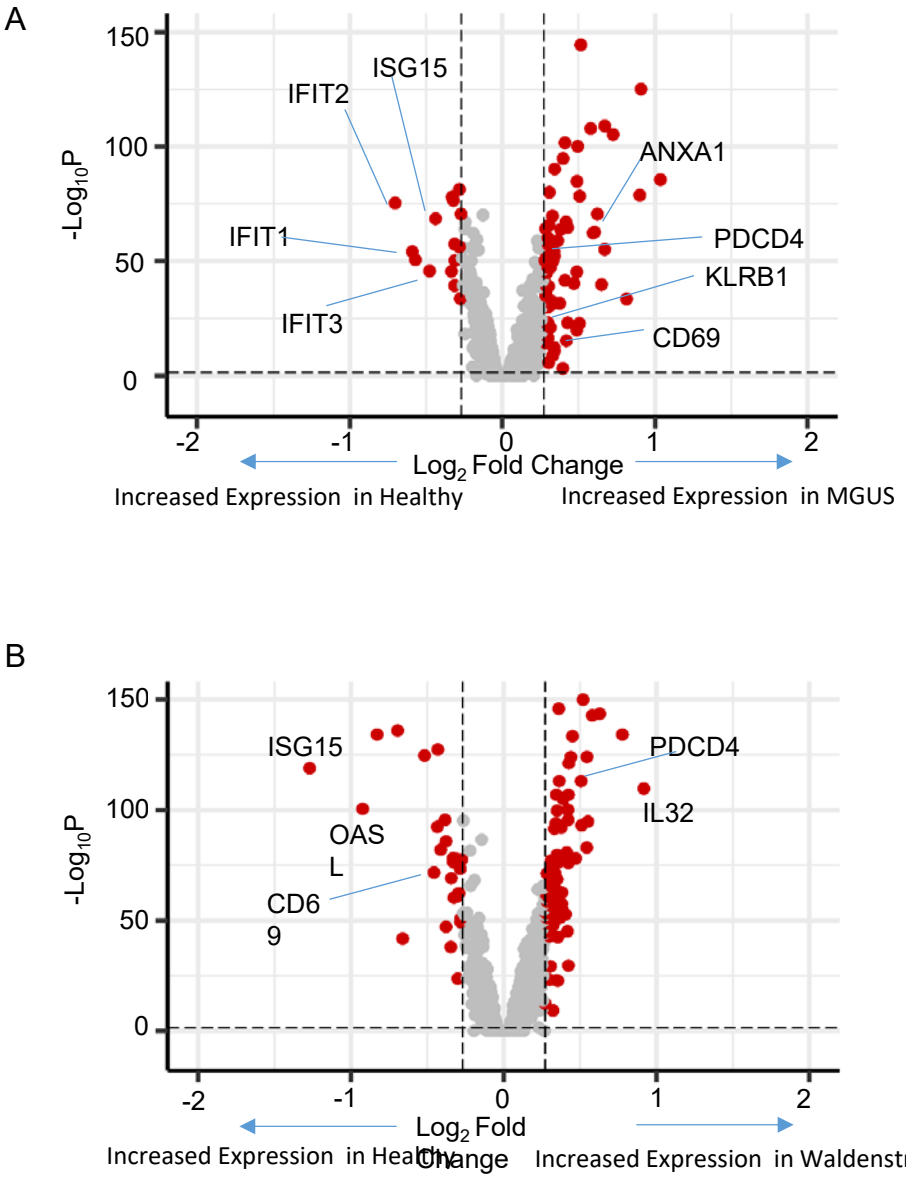
NAME	NES	FDR q-val
HALLMARK_TNFA_SIGNALING_VIA_NFKB	-2.53	0.0000
HALLMARK_INTERFERON_GAMMA_RESPONSE	-2.13	0.0015
HALLMARK_INTERFERON_ALPHA_RESPONSE	-1.89	0.0153
HALLMARK_APOPTOSIS	-1.82	0.0189
HALLMARK_G2M_CHECKPOINT	-1.82	0.0232





Supplementary Fig 19. Pathway analysis of differentially expressed genes in NK cells. Gene set enrichment analysis(GSEA) was performed using fold change based pre ranked genes from WM NK cells compared to NK cells from healthy donors. A. Top differentially down-regulated pathways in NK cells from WM compared to HD. None of the upregulated pathways were statistically significant (not shown). B. Enrichment plot for interferon gamma response and interferon alpha response pathway in WM. NES: Normalized enrichment score based on permutation-based testing from (GSEA).

Supplementary Figure 20

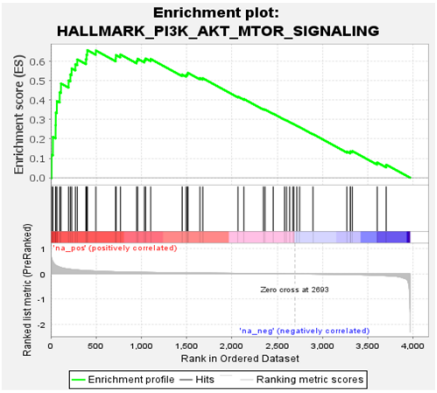


Supplementary Fig 20 . Volcano plots of differentially expressed genes in CD4+ T cells WM versus HD and MGUS versus HD.
Panel A: Differentially expressed genes in CD4+ T cells in MGUS versus HD.
Panel B: Differentially expressed genes in CD4+ T cells in WM versus HD.

Supplementary Figure 21

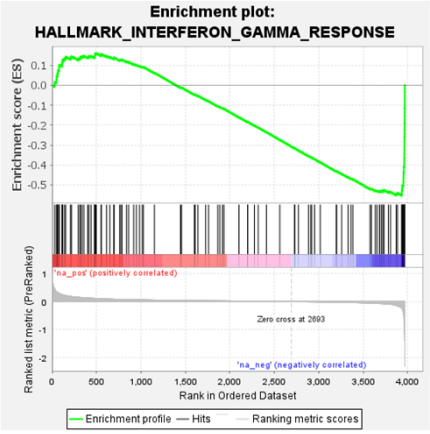
A. Pathways upregulated in T cells from WM compared to HD

HALLMARK Pathway	NES	FDR q-val
PI3K_AKT_MTOR_SIGNALING	1.78	0.0024
MTORC1_SIGNALING	1.67	0.0123
GLYCOLYSIS	1.65	0.0148



B. Pathways downregulated in T cells from WM compared to HD

HALLMARK Pathway	NES	FDR q-val
INTERFERON_GAMMA_RESPONSE	-1.98	0.0142
INTERFERON_ALPHA_RESPONSE	-1.91	0.0142
TNFA_SIGNALING_VIA_NFKB	-1.77	0.0285



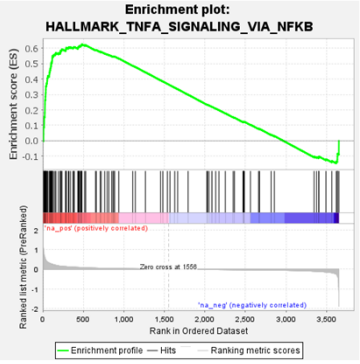
Supplementary Fig 21 . Differentially regulated pathways in WM T cells.

GSEA enrichment analysis of fold change based pre ranked gene list in T cells from waldenstrom patients (WM) compared to those from healthy donors (HD). A. Select pathways upregulated in T cells from WM compared to HD. B. select pathways downregulated in WM T cells compared to HD.

Supplementary Figure 22

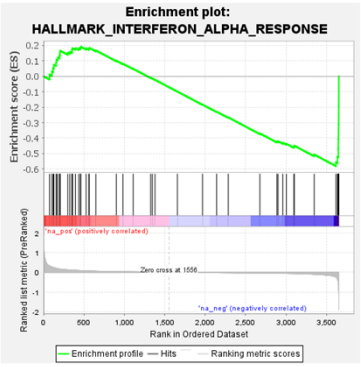
A. Pathways upregulated in T cells from MGUS compared to HD

HALLMARK Pathway	NES	FDR q-val
MTORC1_SIGNALING	1.91	0.0046
TNFA_SIGNALING_VIA_NFKB	1.91	0.0037
ALLOGRAFT_REJECTION	1.65	0.0406



B. Pathways downregulated in T cells from MGUS compared to HD

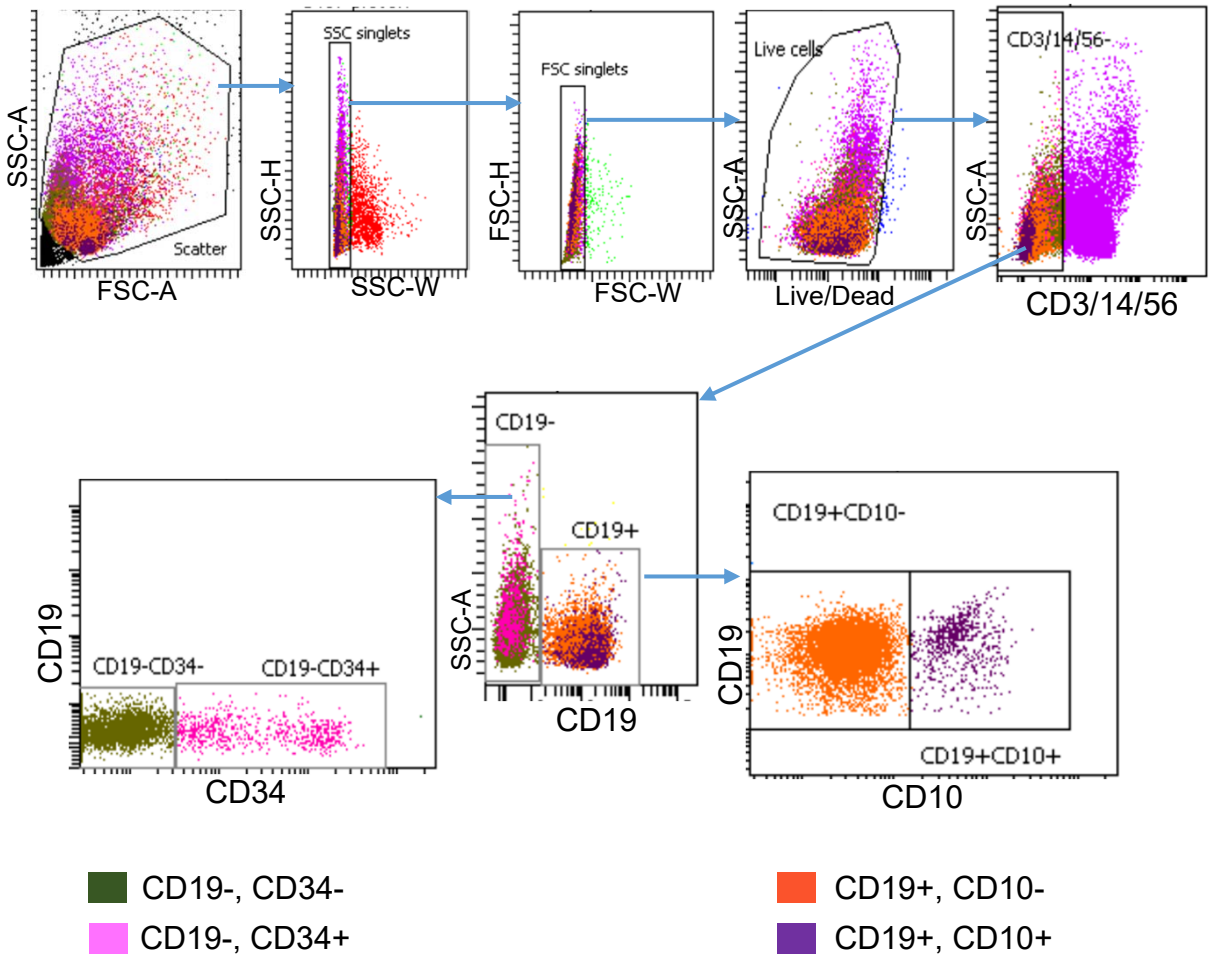
HALLMATK Pathway	NES	FDR q-val
NTERFERON_ALPHA_RESPONSE	-1.81	0.0359
OXIDATIVE_PHOSPHORYLATION	-1.80	0.0189
INTERFERON_GAMMA_RESPONSE	-1.61	0.0480



Supplementary Fig 22 . Differentially regulated pathways in MGUS T cells. GSEA enrichment analysis of fold change based pre ranked gene list T cells from MGUS patients (MGUS) compared to those from healthy donors (HD). A. Select pathways upregulated in T cells from MGUS compared to HD. B. Select pathways downregulated in T cells from MGUS patients compared to HD.

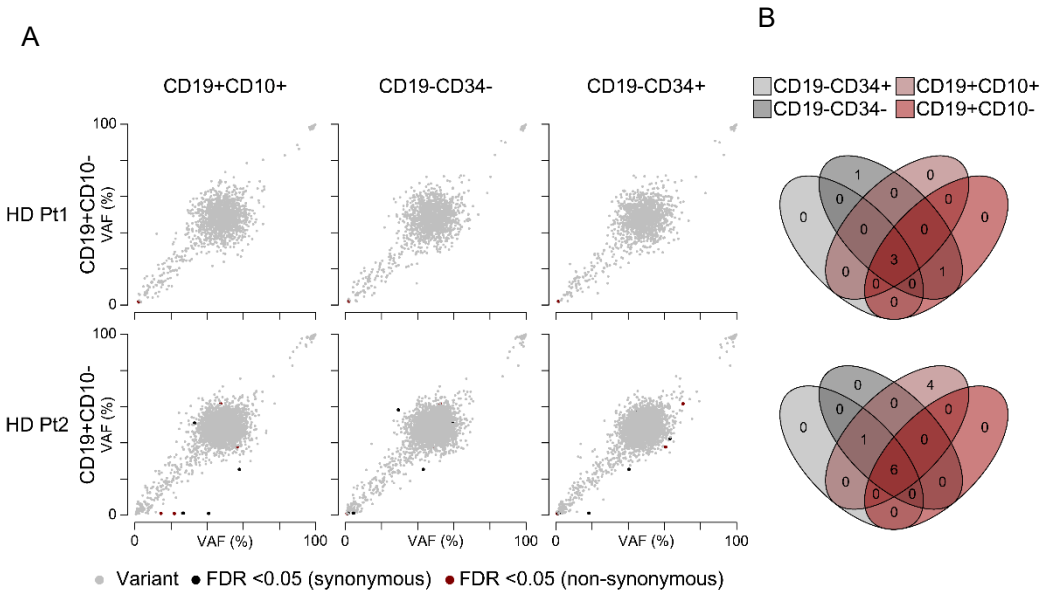
Supplementary Figure 23

Gating strategy used for flow sorting



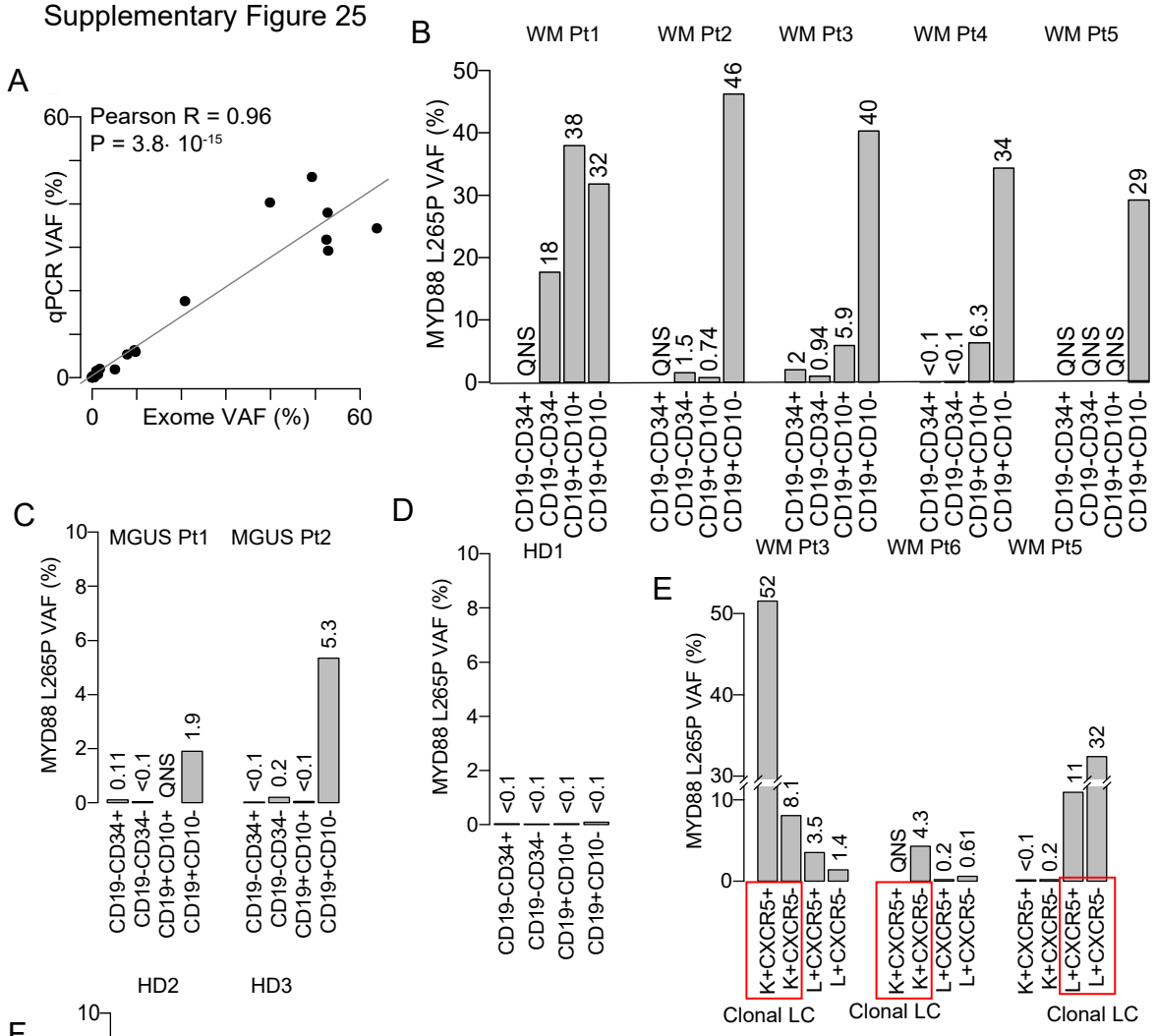
Supplementary Fig 23 . Gating strategy for flow sorting in Fig7.
Lineage negative (based on CD3/CD14/CD56) bone marrow mononuclear cells were sorted into CD19+CD10-, CD19+CD10+, CD19-CD34-, and CD19-CD34+ fractions.

Supplementary Figure 24



Supplementary Fig 24. Mutations in B cell subsets in healthy donors.
A. Variant allele fractions in CD19+CD10- cells (y-axis) as compared to CD19+CD10+ pre-B cells (x-axis; left), CD19-CD34- (middle), and CD19-CD34+ (right) cells in two healthy donors (HD). All variant alleles are show and those with significantly different allelic fractions (FDR ≤ 0.05) between any two cell types are shown in black (non-synonymous mutations) or red (synonymous mutations). B. Venn diagrams of mutations shared and distinct between cell types. A minimum of two reads were required to call a mutation in any given cell type.

Supplementary Figure 25



Supplementary Fig 25 . Validation of Exome sequencing data by qPCR.

A, MYD88 L265P mutations were validated by qPCR. Shown is the Variant Allele Fraction (VAF) determined by Exome-sequencing (x-axis) versus that by qPCR (y-axis). The Pearson correlation (R) value is shown as well as the P-value of correlation. B-D, Validation of MYD88L265P allelic fractions by qPCR in Waldenstrom Macroglobulinemia (WM) patients (B), IgM MGUS (C), and healthy donor (HD) (D). E-F, Analysis of MYD88 L265P allelic fractions in peripheral blood (PB) cells by qPCR. CD19+ B cells were sorted on kappa (K) or lambda (L) light chain and CXCR5 expression. The clonal light chain is denoted below each patient. Samples w/ less than the limit of detection are denoted by <0.1 and samples with no data i.e. insufficient sample are noted as QNS or quantity not sufficient.

Supplementary Table 1. Patient characteristics

	IgM MGUS (n=8)*	WM (n=8)*	Healthy Donor (n=5)	
Median Age (range)	56 years (46-70)	61 years (49-79)	53 years (52-55)	P=NS
Male	50%	50%	40%	P=NS
M spike IgL Kappa	2 (25%)	6 (75%)	NA	
IgM M spike	8 (100%)	8 (100%)	NA	
Median M spike g/dl(range)	0.2 (0.1-0.5)	1.35 (0.7-2.9)	NA	
Median clonal bone marrow infiltration (range)	<5 (<5-5%)	20 (15-40)		
MYD88 L265P	Not tested	100%^		

*MGUS: monoclonal gammopathy of undetermined significance; WM: Waldenstrom macroglobulinemia; Diagnosis of WM and MGUS was based on updated Mayo criteria and required ≥10% bone marrow infiltration in the case of WM (see Gertz M. Am J Hem 96:258-269, 2021). All patients were analyzed before initiation of therapy. NA: not applicable; NS: not statistically significant.

^ in 5 patients with available data

Supplementary Table 2: Antibodies used for CyTOF analysis

	Antibody	Clone		Antibody	Clone
1	CD45	HI30	27	CD8	RPA-T8
2	CD19	HIB19	28	CD14	RMO52
3	CD123	6H6	29	CD127	A019D5
4	CD4	RPA-T4	30	CD45RA	*HI100
5	IgD	1A6-2	31	TIGIT	MBSA43
6	CD20	2H7	32	TIM3	F382E2
7	IgA	Polyclonal	33	PD-L1	29E.2A3
8	CD138	*MI15	34	41BB	4B4-1
9	CD21	BL13	35	CCR7	G043H7
10	CXCR5	RF8B2	36	CD45RO	UCHL1
11	CD27	L128	37	NKG2D	ON72
12	CXCR4	12G5	38	CD25	2A3
13	CD11c	*Bu15	39	KLRG1	*SA231A2
14	CD22	HIB22	40	CD56	HCD56
15	CD79b	CB3-1	41	CD16	3G8
16	CD40	5C3	42	#TCF1	7F11A10
17	c-kit	*104D2	43	#FOXP3	PCH101
18	CD38	HIT2	44	#EOMES	**WD1928
19	CD24	ML5	45	#Gata3	TWAJ
20	CD3	UCHT1	46	#Granzyme	GB11
21	IgM	MHM-88	47	#lambda	MHL38
22	CD141	1A4	48	#kappa	MHK49
23	HLADR	L243	49	#tbet	4B10
24	PD-1	EH12.2H7	50	#BCL6	K11291
25	CD57	HNK-1	51	#ki67	Ki-67
26	CD69	FN50			

All metal conjugated antibodies were purchased from Fluidigm
Purified antibodies purchased from *Biolegend and ** invitrogen were metal tagged using fluidigm metal conjugation kit following manufacturers methods.
#Intracellular markers

Supplementary Table 3: Antibodies used for CiTE-Seq analysis

	Antibody	Clone		Antibody	Clone
1	CD117	104D2	31	CD10	HI10a
2	DNAM-1	11A8	32	CD45	HI30
3	LAG-3	11C3C65	33	CD19	HIB19
4	BCMA	19F2	34	CD38	HIT2
5	NKG2D	1D11	35	CD161	HP-3G10
6	BDCA-2	201A	36	IgD	IA6-2
7	PD-L1	29E.2A3	37	CD11b	ICRF44
8	KLRG1	2F1/KLRG1	38	CXCR5	J252D4
9	CD20	2H7	39	CD1c	L161
10	CD16	3G8	40	HLA-DR	L243
11	CD66b	6/40c	41	CCR4	L291H4
12	CD14	63D3	42	CD141	M80
13	CD123	6H6	43	Ig light chain κ	MHK-49
14	CLEC9A	8F9	44	Ig light chain λ	MHL-38
15	DC-SIGN	9E9A8	45	CD138	MI15
16	CD127 (IL-7Ra)	A019D5	46	CD27	O323
17	TIGIT (VSTM3)	A15153G	47	CD33	P67.6
18	CD25	BC96	48	CD57	QA17A04
19	CD294 (CRTH2)	BM16	49	NCAM	QA17A16
20	CD21	Bu32	50	CD4	RPA-T4
21	CD244 (2B4)	C1.7	51	CD8a	RPA-T8
22	CD79b (Igβ)	CB3-1	52	CD11c	S-HCL-3
23	CD28	CD28.2	53	CD155 (PVR)	SKII.4
24	PD-1	EH12.2H7	54	CD45RO	UCHL1
25	CXCR3	G025H7	55	CD3	UCHT1
26	CCR6	G034E3			
27	29	CCR7			
28	30	CD45RA			

TotalSeq-C antibodies were purchased from Biolegend as used following manufacturers methods.