Check for updates

Received: 22 December 2025 | Accepted: 14 July 2026

DOI: 10.1002/hem3.70456

ARTICLE

HemaSphere eha

Clonotypic characterization defines B-cell drivers of clonal expansion and intratumor heterogeneity in IgM monoclonal gammopathies

David F. Moreno[1,2,^] | Ferran Nadeu[2,3,^] | Marta Casado-Peláez[4] | Caterina Mata[4] | Ian Márquez-López[2] | Marco A. Fernández[5] | Sergi Vaqué[2] | Elena Medina[1] | Helena Llobregat[2] | Marta Garrote[6] | Joan Mañé[2] | Oriol Cardús[2] | Luis Gerardo Rodríguez-Lobato[1,2] | Anna de Daniel[1,2] | María Teresa Cibeira[1,2] | Theo Gremen Oliveira[7] | Elizaveta Kulaeva[8,9] | Ester Lozano[1,2] | Dolors Colomer[3,6] | Laura Rosiñol[1,2] | Ruben van Boxtel[8,9] | Manel Esteller[3,4,10,11,12] | Carlos Fernández de Larrea[1,2,13]

Correspondence: David F. Moreno (dfmoreno@clinic.cat); Manel Esteller (mesteller@santpau.cat); Carlos Fernández de Larrea (cfernan1@clinic.cat)

Abstract

Waldenström macroglobulinemia (WM) and IgM monoclonal gammopathy of undetermined significance (MGUS) share the same cell of origin but differ in clonal size. Compared with other B-cell neoplasms, the lymphoplasmacytic clone in WM can be rather small, limiting our understanding of clonal expansion. We applied an integrative approach using single-cell RNA with B-cell receptor (BCR) sequencing, the assay for transposase-accessible chromatin, and whole-genome sequencing to characterize the tumor clone in patients with IgM MGUS, smoldering WM (SWM), and symptomatic WM (WM). IgM MGUS and low- or intermediate-risk SWM harbored multiple B-cell clones compared to WM. *CD9*, *JCHAIN*, *RASSF6*, and *DUSP22* were the main markers of the dominant B-cell clone at gene expression and chromatin activity levels, with *CD9* preferentially expressed in plasma cell-like tumor cells. *POU2F2* had high activity in the tumor clone and was linked to *CD9* regulatory regions. *MYD88* and *IGLL5* mutations, mainly associated with the mutational signature SBS5, were present in minor clones, whereas the *MYD88* mutation was also detected in nonexpanded B-cells. The 6q deletion was present in tumor cells from high-risk patients, which harbored fitness advantage over copy-neutral tumor cells. Coding mutations clustered tumor and minor clones from oligoclonal patients and were associated with abnormal transcriptional programs. The B-cell clones also showed enriched predicted interactions with monocytes. Our integrative single-cell approach reveals the importance of clone size in IgM gammopathy and identifies key markers promoting clonal expansion.

[1]Amyloidosis and Myeloma Unit, Department of Hematology, Hospital Clínic de Barcelona, University of Barcelona, Barcelona, Spain
[2]Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), University of Barcelona, Barcelona, Spain
[3]Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Madrid, Spain
[4]Josep Carreras Leukaemia Research Institute (IJC), University of Barcelona, Barcelona, Spain
[5]Clinical Flow Cytometry Facility, Biomedical Diagnostic Center, Hospital Clínic de Barcelona
[6]Hematopathology Unit, Department of Pathology, Hospital Clínic de Barcelona, University of Barcelona, Barcelona, Spain
[7]Fundação Pró-Sangue Hemocentro de São Paulo, São Paulo, Brazil
[8]Princess Máxima Centrum for Pediatric Oncology, Utrecht, The Netherlands
[9]Oncode Institute, Utrecht, The Netherlands
[10]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain
[11]Physiological Sciences Department, School of Medicine and Health Sciences, University of Barcelona, Barcelona, Spain
[12]Cancer Epigenetics Group, Sant Pau Research Institute (IR Sant Pau), Barcelona, Spain
[13]Department of Medicine, Facultat de Medicina i Ciències de la Salut, Universitat de Barcelona (UB), Barcelona, Spain
^David F. Moreno and Ferran Nadeu contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2026 The Author(s). *HemaSphere* published by John Wiley & Sons Ltd on behalf of European Hematology Association.

INTRODUCTION

Waldenström macroglobulinemia (WM) is a lymphoplasmacytic lymphoma that produces a monoclonal IgM in the serum.[1,2] The clinical course is highly heterogeneous, as some patients experience an asymptomatic phase.[3,4] The lymphoplasmacytic clones can also be detected in patients with the premalignant condition IgM monoclonal gammopathy of undetermined significance (MGUS) using multiparametric flow cytometry and next-generation sequencing,[5–7] suggesting a clonal expansion that can start years before an overt lymphoma.

The malignant clone in IgM gammopathy (IgM MGUS and WM) originates from a self-proliferating memory B-cell in the germinal center (GC) harboring the recurrent *MYD88* L265P mutation.[6,8] Recent studies identified the *MYD88* mutation in naïve and other B-cell precursors or in patients without clonal B-cells,[8–10] while murine models suggested that the mutation itself is insufficient to cause clonal expansion.[8] This explains why recent multicenter studies failed to incorporate the *MYD88* mutation in progression-free or overall survival models.[4,11] *CXCR4*, *ARID1A*, and *CD79B* mutations, and the 6q deletion (del6q) have been proposed as genomic alterations driving disease progression in WM[12–14]; however, because of their subclonal allelic frequency, assessment using bulk sequencing techniques is difficult.

At the single-cell level, the malignant clone in WM shows a plasma cell-like (PC-like) or memory B-cell-like (MBC-like) phenotype,[15] and the expression of *CD9*, *JCHAIN*, and *DUSP22* in B-cells has been reported.[16] These features highlight phenotypic complexity in WM, which is exacerbated by intra-tumor and inter-patient heterogeneity, particularly in low tumor burden cases with variable bone marrow (BM) infiltration.[5,17] In addition, small BM B-cell clones can cause significant symptomatic disease.[18] This variability complicates the assessment of tumor-specific biomarkers and underscores the need to characterize the clonal composition in IgM gammopathy. Evidence from other B-cell malignancies indicates a diverse B-cell receptor (BCR) repertoire, in which a predominant clone is selected.[19,20] Clonal expansion typically involves cells sharing the same BCR rearrangements that acquire additional genomic alterations, leading to aberrant clones with enhanced proliferative capacity and immune evasion.[8,21] This process remains poorly defined in IgM gammopathies.

In the present study, we tracked distinct B-cell clones by molecular profiling of their BCR and linked them with genomic and transcriptomic signatures across diverse clinical contexts. This approach allowed us to characterize the clonal composition in low tumor burden patients and how predominant clones emerge and gain advantage in patients with adverse clinical outcomes.

MATERIALS AND METHODS

Patient and sample processing

We collected BM and peripheral blood (PB) samples from the Biobank of Hospital Clínic de Barcelona. Patients were diagnosed with IgM MGUS, asymptomatic or smoldering WM (SWM), and symptomatic WM (WM) according to International Consensus.[2,18] Risk stratification in asymptomatic (IgM MGUS and SWM) and symptomatic patients followed previously reported criteria.[4,11] In cases with paired BM and PB samples, patients had detectable circulating disease by multiparametric flow cytometry at sampling.

All patients provided written informed consent according to the Declaration of Helsinki. The study was approved by the Ethics Committee from Hospital Clínic of Barcelona. In total, 16 samples (BM, PB, and longitudinal) were obtained from 3 IgM MGUS patients, 4 SWM patients, and 5 WM patients. Patients were selected based on sample availability and sufficient number of cells to perform multiomics experiments and to represent all clinical stages of IgM gammopathy. The flow-sorting process involved isolating purified mononuclear cells from all 16 samples for further encapsulation in the Chromium Controller (10× Genomics), whereas CD19+ and CD138+ cells were isolated as tumor and CD3+ cells were considered normal only for whole-genome sequencing (WGS).

Library preparation and sequencing

All samples were used for single-cell (sc-) 5′RNA with BCR sequencing, while 4 BM samples were used for a single-cell assay for transposase-accessible chromatin (sc-ATAC) sequencing (Supporting Information S3: Table S1). Single-cell libraries were generated using the Chromium Controller (10× Genomics). For sc-ATAC, nuclei were isolated before transposition and barcoding. Libraries were indexed with single-index primers following the manufacturer's protocols (Chromium Next GEM Single Cell ATAC v2, reference number 1000406). For sc-5′RNA and BCR sequencing, reverse transcription and cDNA amplification were performed after gel bead barcoding, followed by human BCR sequences' enrichment. Libraries were dual-indexed following the manufacturer's protocols (Chromium Next GEM Single Cell 5′ v2–reference number 1000263 and Chromium Single Cell Human BCR Amplification—reference number 1000253). Libraries were pooled, quantified, and sequenced on a NovaSeq X Plus instrument (150 bp paired-end; Illumina).

DNA extraction and sequencing

Fourteen tumor/normal samples from 10 patients were used for WGS. These included 2 paired BM–PB samples and 2 paired longitudinal samples (Supporting Information S3: Table S1). DNA extraction was performed according to the manufacturer's recommendations (QIAamp DNA Mini Kit; Qiagen). The DNA concentration was quantified using the Qubit 4 Fluorometer (Thermo Fisher Scientific). Library preparation was performed following Illumina recommendations and were sequenced on a NovaSeq X Plus instrument (150 bp paired-end; Illumina).

Single-cell sequencing analyses

Briefly, raw FASTQ files from sc-5′RNA with BCR and sc-ATAC sequencing were aligned, filtered, and processed using Cell Ranger to obtain feature-barcode matrices and fragment files, respectively. Downstream analyses were performed using Seurat version (v) 5.3.0[22] and Signac v1.15.0.[23] We followed the standard workflow recommended by Seurat to filter cell doublets, and normalize, scale, and transform data for further clustering and cell annotation. Cell annotation was performed using BM references. The BCR sequences reported by Cell Ranger were processed using IgScan (https://github.com/nadeulab/IgScan) to call clonotypes based on V-gene usage and CDR3 similarity. Relative clonotype abundance was used to label clones from C1 to C*n* (*n* = number of clonotypes identified). C1 represents the most abundant B-cell clone. Additional expanded clones (i.e., those accounting for ≥2 cells that were not C1) were labeled as C2, C3, C4, C5, and other expanded clones, and were grouped as "minor clones" in some analyses. Clones represented only by 1 cell were labeled as "not expanded". Gene set enrichment analyses (GSEA) were performed using Single Cell Pathway Analysis (SCPA)

25729241, 2026, 9, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hem3.70456 by Readcube (Labtiva Inc.), Wiley Online Library on [26/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

v1.6.2. InferCNV v.1.22.0 was used to infer copy number gains and losses from gene expression data.[24] SComatic was used to call somatic single-nucleotide variants (SNVs) from sc-RNA bam files.[25] Liana v0.1.14 was used to analyze cell-to-cell interactions between tumor microenvironment (TME) cells and the B-cell clones.[26]

Anchor-based integration of the RNA and ATAC data was performed following Seurat recommendations, using gene expression from the RNA data as the reference and ATAC gene activities as the query. Label transfer was used to add the BCR and other cell-type annotations to the ATAC data. The adjusted rand index was used to assess the correlation and agreement between both assays. Chromvar v1.30.1 and JASPAR 2020 motifs database were used to compute motif deviation scores at single-cell resolution in accessible chromatin regions.[27] The Wilcoxon rank-sum test was used for differential gene expression analyses. P values were adjusted using the false discovery rate (FDR). More details can be found in the Supporting Information: S2.

WGS analyses

Tumor and normal FASTQ files were aligned to the reference human genome GRCh38 using the bwa MEM algorithm.[28] SNVs and small indels, copy number alterations (CNAs), and structural variants were called using a custom multi-caller approach.[29] Single base substitutions (SBS) signatures fitting was used to quantify the contribution of SBS1, SBS5, SBS8, and SBS9 previously reported in WM,[30–32] using MutationalPatterns.[33] Per mutation probability was computed on fitted SBS signatures using SigProfilerAssignment.[34] More details can be found in the Supporting Information: S2.

RESULTS

BCR clonotypic characterization of the tumor clone and other oligoclonal B-cells

As an overview of the study, Supporting Information S3: Table S1 shows the clinical characteristics of all patients, and Figure 1A summarizes the experimental approach. Using sc-RNA sequencing, 134,040 cells that passed quality control filters were annotated in distinct mononuclear cell populations (Supporting Information S1: Figure S1A,B). 71,332 (53.2%) and 2291 (1.7%) cells were annotated as B-cells and plasma cells, respectively (Figure 1B). The BCR was detected in 61,176 B-cells and plasma cells. The major clone detected (also called C1, $N = 57{,}313$) was larger in size in SWM and WM compared to IgM MGUS (Wilcoxon test, $P = 0.036$). Minor clones were categorized as C2 ($N = 1533$), C3 ($N = 90$), C4 ($N = 66$), C5 ($N = 35$), and other expanded clones ($N = 241$), while those that accounted for only one cell were annotated as "not expanded" ($N = 1898$) (Supporting Information S4: Table S2). Low-risk SWM (Patient [Pnt] 13) had oligoclonal B-cells and nonexpanded B-cells, with a similar profile as the IgM MGUS cases (Figure 1C). Memory B-cells comprised the vast majority of each B-cell clone, less than 5% were annotated as age-associated B-cells, and less than 12% of plasma cells were present (Figure 1D). Clonal plasma cells were identified across all patient groups and were most frequently detected in patients with symptomatic WM (Supporting Information S1: Figure S1C). Using principal component analysis (PCA), we observed that C1 cells differed from the remaining clones in most samples, except in 1 low-risk IgM MGUS (Pnt7) (Figure 1E). After dimensionality reduction, minor clones and nonexpanded B-cells co-localized closely (Figure 1F).

Using pseudobulk differential gene expression analyses (DGE) with FDR-adjusted p-values, C1 cells showed upregulation of genes previously associated with WM tumor cells, including *CD9*, *JCHAIN*, *IGHM*, *DUSP22*, and *RASSF6*. Minor clones preferentially upregulated *IGHG2*, *ADAM28*, *S1PR1*, *FAM30A*, *ZFP91*, and *ARID4B*, among others. In contrast, nonexpanded cells showed increased expression of *BCL7A*, *SOX4*, *DTX1*, and *IGHA1* (Supporting Information S5: Table S3 and Supporting Information S1: Figure 1G). Based on previous validated signatures,[15] we assigned MBC-like and PC-like scores to our population of B-cells from baseline samples ($N = 12$), and observed that *JCHAIN*, *CD27*, *PRDM1*, and *XBP1* were highly expressed in the PC-like compartment of C1 cells and other minor clones and nonexpanded B-cells, while *CD9* was preferentially upregulated only in PC-like C1 cells (Figure 1H), supporting a state-dependent rather than a universal expression pattern. *CD9*, a marker of germinal center (GC) B-cells, had a heterogeneous expression across patients, given that fewer *CD9*-positive C1 cells were present in a relapsed/refractory (R/R) WM (Pnt5) or a previously exposed WM (Pnt11), while treatment-naïve cases had more C1 cells expressing *CD9* (Supporting Information S1: Figure S1D).

We next performed *in silico* validation of the top upregulated genes identified in C1 cells using two publicly available data sets: WM B-cells analyzed by sc-RNA 3′ sequencing (WM-1)[15] and WM B-cells from a bulk RNA sequencing study (WM-2).[35] No significant differences were found between C1 and WM-1 cells. In contrast, only *JCHAIN* ($P = 0.020$) and *XBP1* ($P < 0.001$) had a differential expression between the single-cell studies and WM-2, likely reflecting a technical batch effect between assays (Supporting Information S1: Figure S1E). C1 cells also upregulated the Toll-like receptor pathway, along with NF-kB-, IL6-mediated JAK/STAT3 signaling and MAPK signaling, which were associated with an inflammatory response (Figure 1I). These results suggest that C1 cells are characterized by a predominant MBC-like identity with GC expression markers but transitioning to a plasmacytic phenotype, consistent with an intermediate or partially differentiated state.

Regarding specific targetable surface antigens in WM, we observed that *MS4A1* (CD20) was present in all B-cells, while *TNFRSF17* (BCMA) expression was low or absent, and this low expression was clone- and patient-independent. It is noteworthy that *TNFRSF13C* (BAFFR) and *TNFRSF13B* (TACI) expression was consistent across all samples and B-cell clones, with BAFFR showing comparatively higher expression (Supporting Information S1: Figure S1F–G). *CD9* expression was enriched in a subset of C1 cells in a patient-dependent manner, while BAFFR emerged as the second upregulated marker (pairwise Wilcoxon test, BAFFR vs. remaining genes, P-adjusted < 0.001). In contrast, from C2 to nonexpanded B-cells, BAFFR was the most consistent upregulated marker after CD20 (Supporting Information S1: Figure S1G).

Distinct chromatin accessibility patterns in B-cell clones

We further integrated the BCR clonotypic gene expression with sc-ATAC data in samples that had both assays available. In total, 30,033 B-cells from 1 IgM MGUS patient, 2 SWM patients, and 1 WM patient were successfully integrated and their cell annotations were transferred (Supporting Information S1: Figure S2A). The distribution of annotated B-cells was similar to the RNA data, in which memory B-cells were the most representative cell type (Figure 2A). Plasma cells constituted a minority within each annotated BCR clone (range 1.13%–4.47%) (Figure 2B). Similarly, the minor clones and nonexpanded cells co-localized closely and belonged to IgM MGUS and SWM patients (Figure 2C). Differential motif accessibility analysis using chromVAR across all B-cell clones revealed that the

25729241, 2026, 9, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hem3.70456 by Readcube (Labtiva Inc.), Wiley Online Library on [26/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

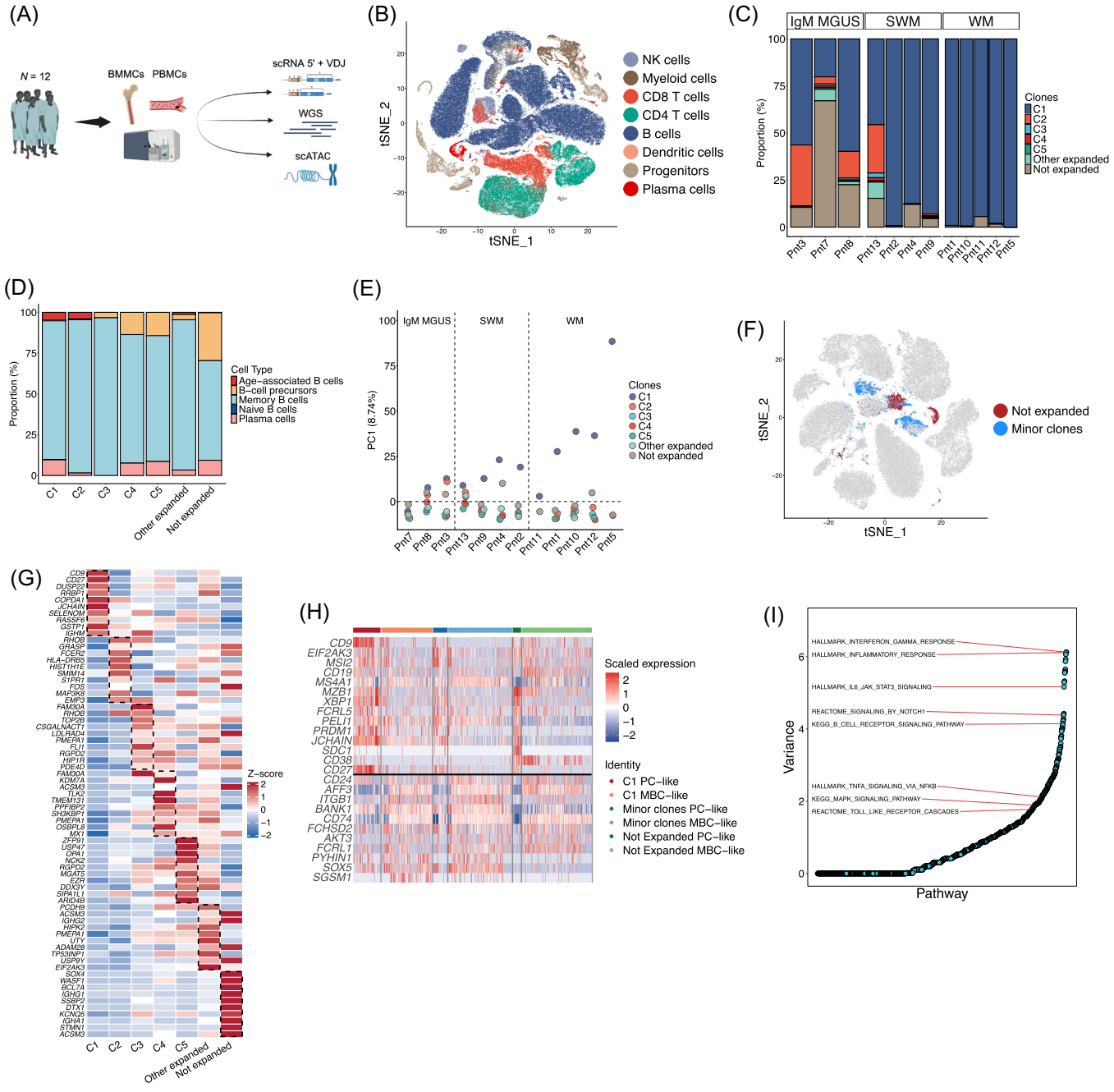


FIGURE 1 **B-cell receptor (BCR) repertoire characterization in IgM gammopathy.** **(A)** The design of the study included 16 samples from 12 patients. **(B)** Main annotated cell types after t-distributed stochastic neighbor embedding (tSNE). **(C)** Distribution of B-cell clonotypes across baseline or diagnostic samples showing oligoclonal cells in IgM monoclonal gammopathy of undetermined significance (MGUS) and low-risk smoldering Waldenström macroglobulinemia (SWM) (Pnt13). **(D)** Proportion of annotated B-cell subtypes within each B-cell clonotype group. A total of 61,176 B-cells had a BCR annotation for downstream analyses. **(E)** Principal component analysis across baseline or diagnostic samples taking into account whole-gene expression per clone. **(F)** Distribution of minor clones (C2, C3, C4, C5, and other expanded B-cells) and nonexpanded B-cells on tSNE dimensions. C1 cells are depicted in gray. **(G)** Pseudobulk differential gene expression analyses using the Wilcoxon rank-sum test and false discovery rate-adjusted P values identified the main markers of each B-cell clone group. **(H)** Heatmap showing the expression of memory B-cell-like (MBC-like) and plasma cell-like (PC-like) signature genes previously characterized in WM across C1 cells, minor clones (C2, C3, C4, C5, and other expanded), and nonexpanded B-cells. **(I)** Upregulated signature pathways comparing C1, minor clones, and nonexpanded B-cells. The most variable pathways across comparisons were shown.

octamer-binding transcription factors (OCT/POU family), *FOXP1* and *NFKB2*, had increased motif activity in C1 cells compared to the remaining B-cell clones. Conversely, motifs corresponding to *XBP1*, *KLF3/9*, *IRF4*, and the SMAD family of transcription factors (TFs) were enriched in minor clones, whereas *PRDM1*, *SPIB*, and *PAX5* were enriched in nonexpanded B-cells (Figure 2D and Supporting Information S6: Table S4). It is noteworthy that the POU family were the top TFs in C1 cells after comparing C1 cells and nonexpanded B-cells (Figure 2E). These results suggest that the C1 cells retained strong POU2F1/2 activity that may result in repression of the plasmablast program.

25729241, 2026, 9, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hem3.70456 by Readcube (Labtiva Inc.), Wiley Online Library on [26/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

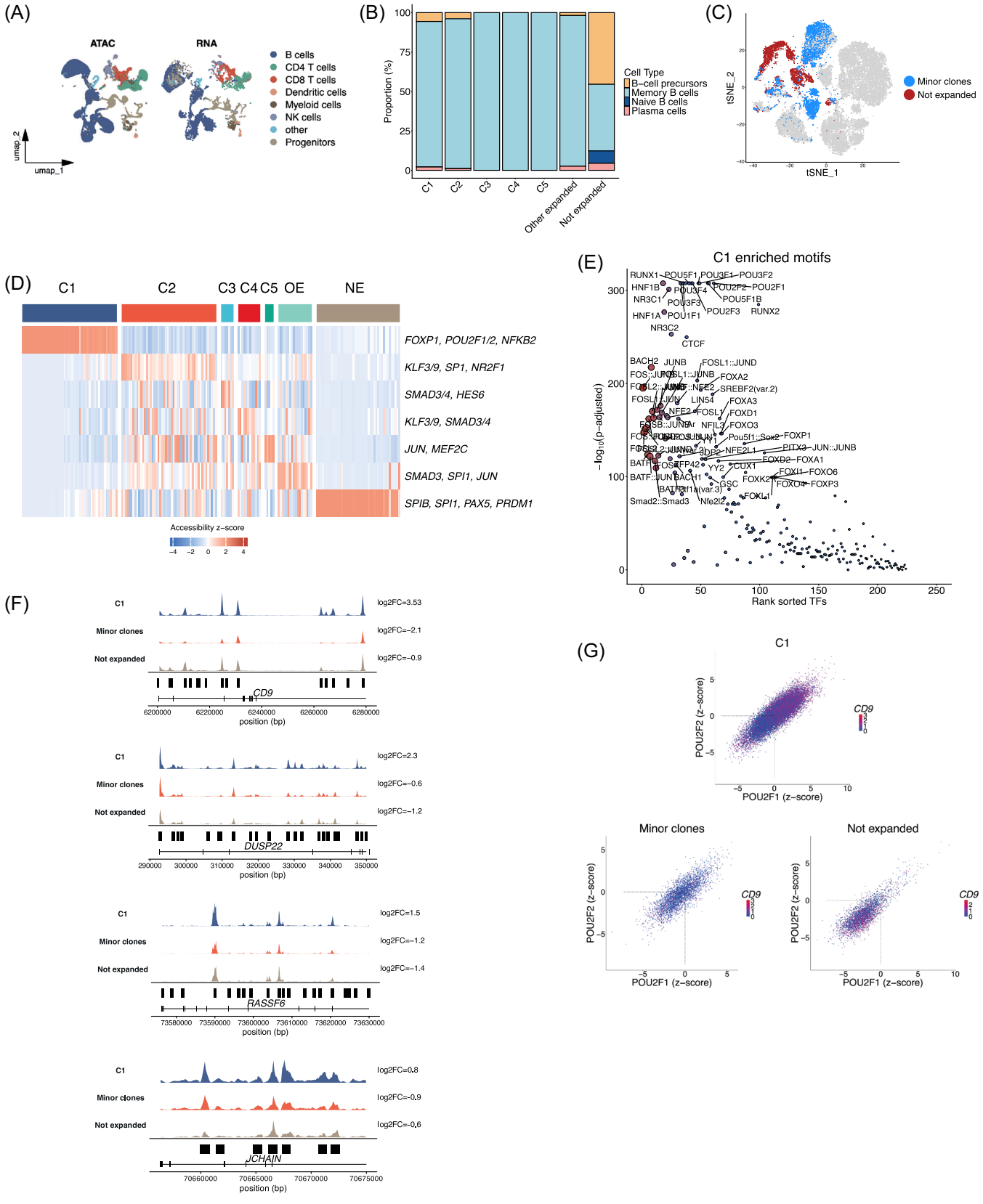


FIGURE 2 (See caption on next page).

25729241, 2026, 9, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hem3.70456 by Readcube (Labtiva Inc.), Wiley Online Library on [26/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

FIGURE 2 **Chromatin accessibility characterization of the B-cell clones**. All differential analyses were performed using the Wilcoxon rank-sum test with false discovery rate-adjusted P values. **(A)** UMAP dimension reduction after co-embedding single-cell RNA and single-cell ATAC data. **(B)** Main B-cell subtype distribution per clone. **(C)** t-distributed stochastic neighbor embedding (tSNE) of the integrated RNA and ATAC B-cells colored by clonotype group (C1, minor clones [C2, C3, C4, C5, and other expanded B-cells], and nonexpanded B-cells), showing the close proximity of minor clones and nonexpanded B-cells. **(D)** Differential transcription factor (TF) motif analyses were performed after computing ChromVAR single-cell activities in C1, C2, C3, C4, C5, other expanded (OE), and nonexpanded (NE) B-cells. **(E)** Top differential motifs identified in C1 cells. The size and color from each TF were plotted based on the average log2-fold change. **(F)** Differential accessibility analyses were performed in C1 cells, minor clones (C2, C3, C4, C5, and other expanded), and nonexpanded B-cells. Accessible peaks were linked with the nearest gene. Genomic regions display pseudobulk accessibility tracks for the *CD9*, *DUSP22*, *RASSF6*, and *JCHAIN* loci. Positive Log2 fold change (Log2FC) values represent an increased accessibility in C1 cells in contrast to minor clones and nonexpanded B-cells. **(G)** *POU2F1* and *POU2F2* motif deviation scores in C1, minor clones, and nonexpanded B-cells. Individual cells are colored according to scaled *CD9* expression. Dotted lines indicate the zero-centered thresholds used to distinguish lower versus higher motif deviation scores.

POU transcription factors' activity in the tumor clone

Given that POU TFs share similar DNA-binding motifs, we focused only on those with previously reported relevance in B-cell biology. *POU2F1* and *POU2F2* are TFs that regulate B-cell proliferation and differentiation within the GC in normal and malignant B-cells.[36,37] We performed chromVAR motifs pairwise correlation activities to delineate distinct B-cell states. In C1 cells, the *POU2F1/2* TFs were decoupled from the plasmablast program (near zero or negative correlations with *IRF4*, *XBP1*, and *PRDM1*) and showed opposite activity to *SPIB* ($r = -0.49$). Minor clones displayed a plastic state, in which the activity of *POU2F1/2* was correlated with *NFKB2* ($r = 0.32–0.36$) and weakly correlated with *IRF4* and *XBP1*. Nonexpanded cells had a more physiological GC to plasmablast transition, in which *POU2F1/2* were correlated positively with *IRF4*, *XBP1*, and *SP1*, while *IRF4* and *XBP1* were negatively correlated with *SPIB* and positively correlated with *SP1* (Supporting Information S1: Figure S2B–D). Based on diagnosis, we identified that *POU2F2* was highly active in WM in contrast to IgM MGUS and SWM, which showed more activity of *IRF4*, *XBP1*, and *PRDM1* at both chromatin and expression levels. The increased expression of *XBP1* reflected activation of a broader plasma cell-associated program (Supporting Information S1: Figure S2E).

To further investigate the top markers in C1 cells identified from sc-RNA data, we performed differential chromatin accessibility analyses and observed increased accessibility at the *CD9*, *DUSP22*, *RASSF6*, and *JCHAIN* loci in C1 cells compared to minor and nonexpanded B-cells (Figure 2F). Peak-to-gene linkage analyses revealed that *CD9* cis-regulatory elements located at chr12:6,44 Mb contained *POU2F1/2* motifs, which also lay in proximity to CD27, a memory B-cell marker (Supporting Information S1: Figure S2F). We also found that *CD9* activity correlated positively with *POU2F1* ($r = 0.23$, $P < 0.001$) and *POU2F2* ($r = 0.21$, $P < 0.001$) motif activities, but preferentially in C1 cells (Figure 2G), suggesting that these TFs are associated with the transcriptional activation of *CD9*. Using the RNA expression data from the remaining samples, we observed that *POU2F2* was consistently upregulated in C1 cells from WM patients. Footprinting analyses performed in each case demonstrated *POU2F2* occupancy in C1 cells from SWM patients preferentially (Supporting Information S1: Figure S2G). Altogether, C1 cells harbored high *POU2F2* activity and displayed features suggestive of an aberrant pre-plasmablast-like state.

Genomic landscape of IgM MGUS and WM

We analyzed the genomic alterations using WGS from sorted CD19+ and CD138+ cells in available samples. All harbored the *MYD88* L265P mutation. Coding mutations in *IGLL5* and *PRDM1* were present in at least 3 patients. *CXCR4* mutations were detected in 2 patients and were small indels (Figure 3A and Supporting Information S7 and S8 Table S5 and S6). del6q and gain of 6p (gain6p) were detected in 4 patients. *TNFAIP3*, *PRDM1*, *IBTK*, *FOXO3*, and *ARID1B* were involved in chr6 deletions, whereas *IRF4* and *FOXP4* were involved in chr6 gains. Additional copy-number gains affected *MYC* (chr8) and *BCL2* (chr18). Translocations of chr7, chr8, chr6, and chr2 were observed in an R/R WM (Pnt10) and a high-risk SWM (Pnt9). An intermediate-risk IgM MGUS patient who progressed after 5 years (Pnt3) carried deletions in chr11 and chr19, and a gain of chr18, while Pnt10 additionally harbored a chr17 deletion encompassing *TP53* (Figure 3B and Supporting Information S9 and S10 Tables S7 and S8). Mutational signature analyses revealed dominant activity of SBS1, SBS5, SBS8, and SBS9 across all cases, except Pnt7 (Figure 3C), in whom the C1 clone represented a minor fraction of the total B-cell repertoire (Figure 1C). These results reflect aging-related processes as the predominantly mutational force in IgM gammopathy, followed by the activation-induced cytidine deaminase (AID) poly-eta activity.

Integrative genomic characterization of B-cell clones

We next characterized the clonal expansion using an integrative approach. We called SNVs from sc-RNA aligned data on B-cells based on their clonotypic features (C1, minor clones, and nonexpanded cells). Given the presence of false-positive calls, we kept only those coding RNA-based SNVs that were also detected by WGS from the entire cohort. In total, 114 mutations were identified across all samples and B-cells. C1 cells harbored the majority of the somatic mutations, with *MYD88* L265P present in all samples, followed by a synonymous variant in *KIAA2013* (Figure 4A). It is noteworthy that *MYD88* L265P was also present in minor clones from 1 IgM MGUS (Pnt3) and 1 SWM (Pnt13) patients, and even in nonexpanded B-cells from the same AWM patient. *IGLL5* mutations were also detected in minor clones from two other cases. Other SNVs identified in nonexpanded B-cells were missense mutations in *PRDM1* and *RPS20*, and a synonymous mutation in *NFATC2*. All mutations found in minor clones and nonexpanded cells were also detected in C1 cells, except from *IGLL5* (Pnt3 and Pnt4), *MSL2* (Pnt13), and a single missense mutation in *H1-3* (Pnt3). *TP53* mutations were identified only in C1 cells from two patients (Pnt5 and Pnt10) (Figure 4A and Supporting Information S11: Table S9).

The *MYD88* L265P has been previously reported in nonclonal B-cells from patients with WM.[8,10] Because we also identified *IGLL5* mutations in minor clones (Pnt3 and Pnt4), we analyzed the mutational processes underlying both alterations and found SBS5 to be the predominant signature (Supporting Information S1: Figure S3A–D). Visual inspection using the Integrative Genomics Viewer (IGV) of individual cases confirmed that all *IGLL5* mutations did not map to other immunoglobulin loci. These results suggest that *IGLL5* mutations may be associated with a similar aging mutagenic process as *MYD88*.

25729241, 2026, 9, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hem3.70456 by Readcube (Labtiva Inc.), Wiley Online Library on [26/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

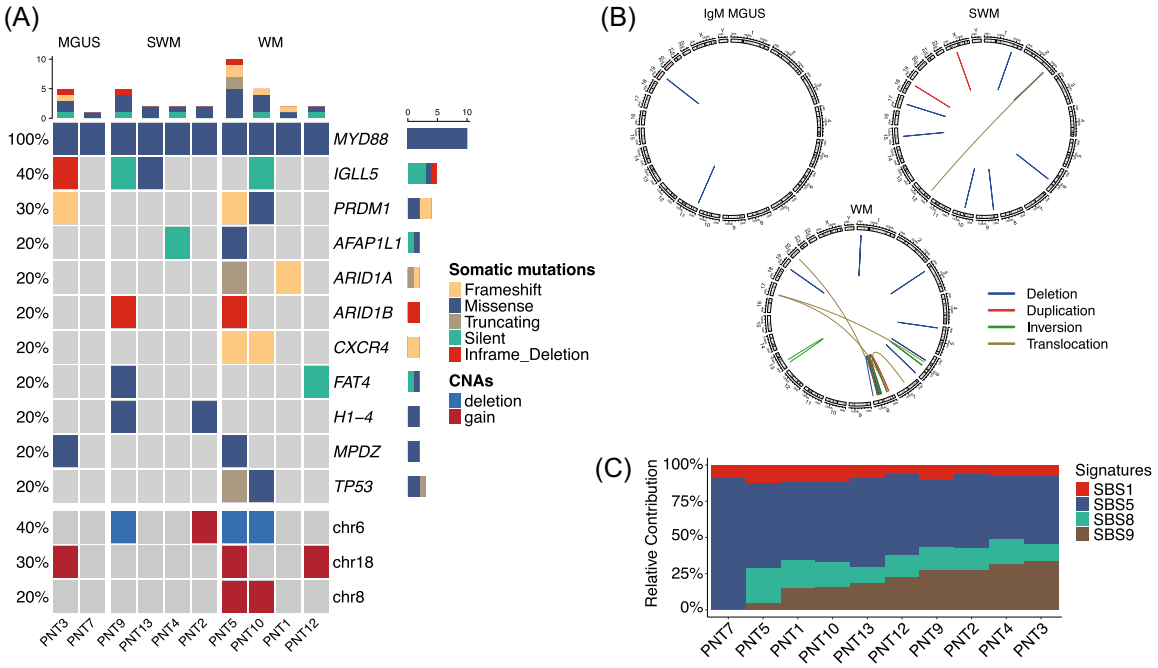


FIGURE 3 **Genomic landscape of the cohort. (A)** Mutations in *MYD88* were present in all samples, followed by *IGLL5*, *H1-4*, and *PRDM1*. Copy number alterations (CNAs) involving chromosomes 6, 8, and 18 were the most frequent and prevalent in more than one case. **(B)** Circos plots summarizing structural variants (SVs) identified by whole-genome sequencing across diagnostic groups. Deletion, duplication, inversion, and translocations are shown. Each internal line corresponds to an SV connecting two breakpoint positions. All SVs detected in our cohort are based on data presented in Supporting Information S10: Table S8. **(C)** Distribution of single-base substitution (SBS) mutational signatures per case, inferred from somatic mutations identified by whole-genome sequencing. Signature contributions were estimated using a fitting approach on SBS1, SBS5, SBS8, and SBS9 based on prior studies in Waldenström macroglobulinemia and other mature lymphoid malignancies.

To investigate whether somatic mutations cluster together within B-cell clones, we analyzed patients who had WGS data and sufficient number of distinct B-cell clones from sc-RNA data. In Pnt9, C1 cells carrying *CD83* 139G>A clustered separately from the *H1-4* 127 C > T mutant cells, while most C1 cells also harbored del6q, in contrast to C2 and nonexpanded B-cells (Figure 4B). Gain of the entire chr18 was detected mostly in C1 cells from Pnt3 (Figure 4C). Pnt7, Pnt4, and Pnt1 had few cells with callable somatic mutations due to genes with low expression or did not have CNAs identified from sc-RNA or WGS. In Pnt7, nonexpanded B-cells were the most prominent cells, which limited the capacity to identify somatic mutations from C1 cells (Supporting Information S1: Figure S4A–C). In Pnt13, which had a more oligoclonal context, the *IGLL5* mutation was detected preferentially in C1 cells (Figure 4D).

Intratumor heterogeneity is associated with genomic alterations and aberrant gene expression programs

To gain more insight into intratumor heterogeneity, a high number of tumor cells is needed. Thus, we analyzed patients (Pnt10 and Pnt12) with both BM and PB samples to characterize in depth how the C1 clone behaves using somatic mutation data and CNAs. By inferring somatic mutations clusters from WGS, we detected 4 subclones that emerged from a *MYD88*- and *CXCR4*-mutated clone in Pnt10. *IGLL5* c.153C>T-cells did not show a clear trend of grouping together; however, inferred del6q or del17p cells co-localized each other and clustered into another C1 subcluster (Figure 4E and Supporting Information S1: Figure S4D). In the case of Pnt12, *RPS20* and *PLEKHJ1* mutations, identified as subclonal by bulk WGS, clustered apart in the C1 clone using sc-RNA in both BM and PB (Figure 4F). In contrast, inferred gain18q was present in all C1 cells (Supporting Information S1: Figure S4E). In both Pnt10 and Pnt12, the main tumor clone was detected in BM and PB (Supporting Information S1: Figure S5A–B). After combining both BM and PB from Pnt12 as a single group of C1 cells, we performed DGE based on the presence of *RPS20* and *PLEKHJ1* mutations and identified that *CD9* cells clustered apart from *VPREB3* cells, which is a less mature B-cell marker (Figure 4G and Supporting Information S12: Table S10).

6q deletion impacts in disease progression or transformation

We further analyzed patients with longitudinal samples (Pnt2 and Pnt5) at diagnosis and disease progression or transformation. Pnt2, a high-risk SWM, progressed after 1 year of follow-up. At both time points, there were mutations in *MYD88*, *IGLL5*, and *H1-4*, along with del6q and gain6p (Figure 5A). In addition, we observed that the *KLHL6* c.125C>T mutation spread between multiple C1 cells at the single-cell level, but *TMED4* c.35T>A was confined within a C1 cluster. Gain6p was present in nearly all C1 cells; meanwhile, del6q C1 cells clustered together as the *TMED4* mutation, and was associated with lower *CD9* expression (Figure 5A).

Although the dominant tumor clone appeared broadly stable during the 1-year follow-up (Supporting Information S1: Figure S5C), del6q remained restricted to a distinct subcluster despite the widespread distribution of gain6p across the clone. We next examined Pnt5, an intermediate-risk R/R WM that transformed into a diffuse large B-cell lymphoma (DLBCL). In this case, we observed that the *SRRT* c.16G>A mutation and gain8p spread across all C1 cells at the first time point, but then clustered together after two years of follow-up. In contrast, del6q C1 cells appeared to expand at the time of

25729241, 2026, 9, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hem3.70456 by Readcube (Labtiva Inc.), Wiley Online Library on [26/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

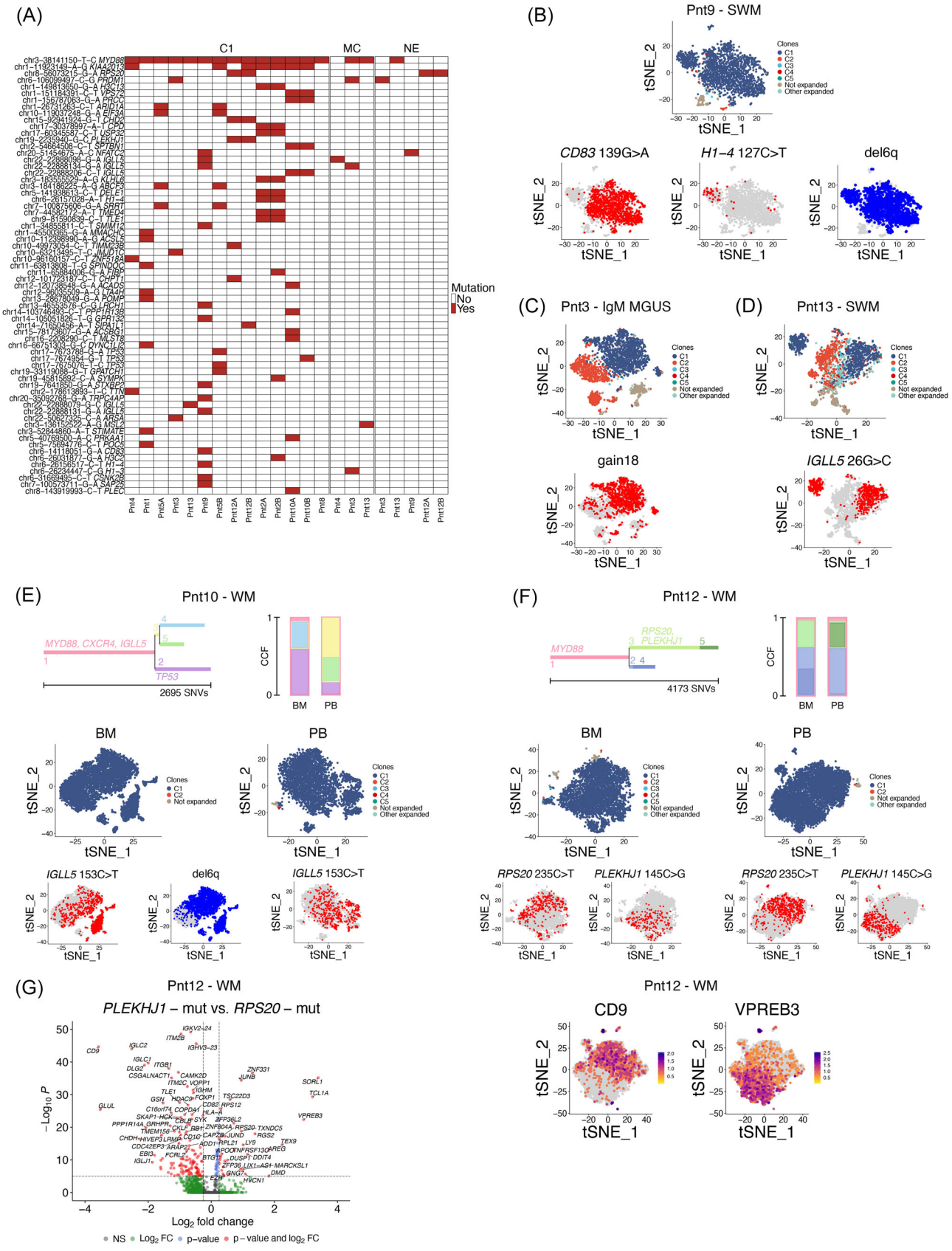


FIGURE 4 (See caption on next page).

25729241, 2026, 9, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hem3.70456 by Readcube (Labtiva Inc.), Wiley Online Library on [26/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

FIGURE 4 **B-cell genomic landscape after integration of whole-genome and single-cell RNA sequencing**. **(A)** Distribution of somatic mutations (in red) in C1 cells, minor clones, and nonexpanded B-cells. Rows represent individual mutations with their respective genomic region, while columns represent patient-specific clone groups. **(B–D)** Representative t-distributed stochastic neighbor embedding (tSNE) plots of individual cases with smoldering Waldenström macroglobulinemia (SWM) (Patient 9—Pnt9 and Patient 13—Pnt13) and IgM monoclonal gammopathy of undetermined significance (MGUS) (Patient 3—Pnt3). The lower plots show the distribution of somatic mutations or copy number alterations projected onto the same tSNE space, with alteration-positive cells highlighted and cells without the alteration shown in gray. **(E)** Integrative overview of Patient 10 (Pnt10, symptomatic WM) with matched bone marrow (BM) and peripheral blood (PB) samples collected at the same time point. The upper panel shows the subclonal architecture inferred from bulk WGS analyses. The middle panels show tSNE plots of BM and PB cells colored by clonotype group. The lower feature plots show the distribution of the indicated alterations on the same tSNE embeddings, with cells positive for the *IGLL5* mutation or the 6q deletion (del6q) highlighted, and the remaining cells shown in gray. **(F)** Similarly, Patient 12 (Pnt12, symptomatic WM) had matched BM and PB samples. Highlighted mutations affect *RPS20* and *PLEKHJ1* genes. G. Differential gene expression analyses performed on *RPS20* and *PLEKHJ1* mutant cells showing upregulated *VPREB3* and *CD9*, respectively.

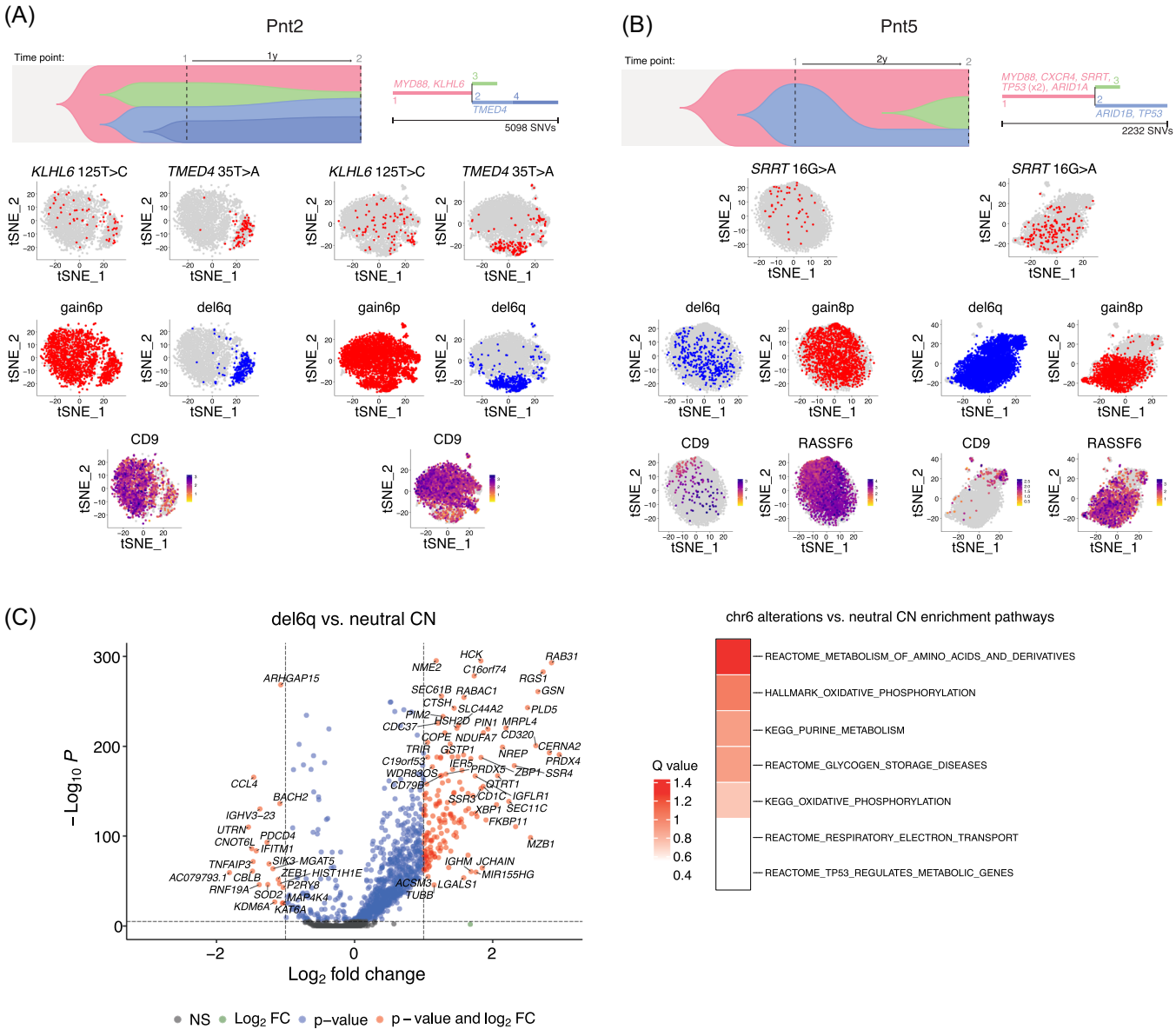


FIGURE 5 **Longitudinal analysis in high-risk cases**. **(A)** Patient 2 (Pnt2) had smoldering Waldenström macroglobulinemia (SWM) and required treatment 1 year later. Fish plot shows the subclonal architecture and dynamics inferred from WGS. Feature plots show the distribution of mutations (in red), copy number alterations (CNAs) (in red gains, in blue deletions), and selected gene expression patterns across C1 cells projected onto the tSNE embedding at each time point. The *TMED4* mutation clustered apart from the remaining C1 cells (in gray), which also localized closer to a lower expression of CD9 and 6q deletion (del6q). **(B)** Patient 5 (Pnt5) had symptomatic WM that transformed into high-grade lymphoma after 2 years. del6q C1 cells expanded during follow-up, while the number of C1 cells with *CD9* expression was small and clustered apart from *RASSF6* and 8p gain (gain8p). **(C)** Differential gene expression and pathway enrichment analyses in C1 cells comparing del6q versus neutral copy number (CN). To the right, positive log2-fold change for C1-del6q cells is shown.

25729241, 2026, 9, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hem3.70456 by Readcube (Labtiva Inc.), Wiley Online Library on [26/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

transformation. *CD9* expression was low and tended to cluster in nongain8p/SRRT mutated C1 cells on transformation to DLBCL; meanwhile, *RASSF6* expression localized similarly as gain8p (Figure 5B). During the 2-year period in the transformation case, the tumor clone did not undergo a major shift (Supporting Information S1: Figure S5D).

To further assess the impact of del6q on C1 cells, we performed a DGE comparing C1 cells with the deletion and those with neutral copy number from patients with del6q detected from both gene expression and WGS (Pnt2, Pnt5, Pnt9, and Pnt10). In total, we analyzed 4,309 del6q-C1 cells versus 7144 neutral-C1 cells. *TNFAIP3*, *BACH2*, *FOXO1*, *CCL4*, and *LYN* were downregulated in del6q, while *JCHAIN*, *RASSF6*, *IGHM*, *PIM2*, *RGS1*, and *HCK* were upregulated (Supporting Information S13: Table S11). Given that some cells harbored del6q and gain6p, we performed GSEA and observed that OXPHOS and metabolic-related signatures were consistently present in these cells in contrast to copy neutral C1 cells (Figure 5C). These results suggest that tumor cells harboring alterations in chr6 represent subclones with high metabolic activity, and that these alterations may contribute to disease evolution in a subset of patients.

Immune cells and B-cell clones' communication

Previous studies reported impaired immune responses against the tumor clone in WM.[16,21,38] To investigate whether the TME was associated with the clonal architecture, we analyzed the predominant immune cells (CD4 T-cells, $n = 21{,}860$; CD8 T-cells, $n = 14{,}935$; NK cells, $n = 4075$; and monocytes, $n = 5056$) from all samples and performed cell-to-cell communication analyses across the C1 cells, minor clones, and nonexpanded B-cells. The inferred frequency of cell interactions was higher between C1 and monocytes, followed by C1 and CD8 T-cells (Figure 6A). Considering the immune cells as the source and B-cells as targets, the dominant monocyte-related interactions included *CXCL8-CD79A*, *S100A8-CD69*, *S100A9-CD68*, *VCAN-CD44/ITGA4/SELL*, *TNFSF13B/TNFSF13-TNFRSF13B*, and *PTGS2-ALOX5* (Figure 6B). The logFC expression was higher in interactions involving C1 cells with *CXCL8* and *S100A9* (ligand–receptor score [LRs] > 0.85). In addition, the interaction between *TNFSF13* (APRIL) or *TNFSF13B* (BAFF) with *TNFRSF13B* (TACI) was either enriched in C1 cells (*TNFSF13* logFC 1.29, LRs 0.87; *TNFSF13B* logFC 1.23, LRs 0.88) or more abundant in this B-cell subtype compared to minor clones (*TNFSF13* logFC 0.65, LRs 0.85; *TNFSF13B* logFC 0.74, LRs 0.85) and nonexpanded cells (*TNFSF13* logFC 0.40, LRs 0.81; *TNFSF13B* logFC 0.49, LRs 0.83). In contrast to these monocytic-specific interactions, *GNAI2–CNR1* was also enriched in C1 cells, but shared across immune cells (Figure 6B).

Considering the B-cells as a source, monocyte-related interactions again predominated, being enriched in C1 cells, and included *SERPINF1–PLXDC2* (logFC 1.50, LRs 0.84) and *CCN2–LRP1* interactions (logFC 1.35, LRs 0.81). Notably, the interaction between *ANGPTL1-LILRB2* was higher in C1 cells (logFC 1.30, LRs 0.85), compared to minor clones (logFC 0.95, LRs 0.82) and nonexpanded B-cells (logFC 0.85, LRs 0.83) (Figure 6C). Among other positive interactions, C1 cells dominated most top upregulated cell communications, especially as a source (Figure 6D,E), suggesting preferential enrichment of monocyte-related signaling programs in the dominant B-cell clone.

DISCUSSION

We characterized the BCR repertoire of patients with IgM MGUS and WM using an integrative approach. We first observed that an oligoclonal state was predominant in IgM MGUS and low-risk SWM, where multiple B-cell clones were present, in contrast to high tumor burden WM patients. Memory B-cells were the most frequent in each B-cell clone. However, only a relatively small number of plasma cells shared similar BCR sequences with memory B-cells, and very few overlapped with B-cell precursors. Our results, similar to others,[5,6,39] suggest that memory B-cells have a clonal expansion and that a subset of them can still differentiate toward plasma cells.

We observed and validated that *JCHAIN*, *IGHM*, *RASSF6*, *DUSP22*, and *CD9*, and the TF *POU2F2* had high activity in the tumor clone, in accordance with two other recent studies.[15,16] Previous evidence showed that a subset of GC B-cells can express *CD9*, and it increases during differentiation toward plasma cells in a manner dependent on Bruton's tyrosine kinase signaling.[40,41] We show that *CD9* expression was enriched in the tumor clone compared to the remaining B-cells, particularly within cells displaying previously described PC-like features. We also found a potential mechanistic association between *POU2F2* and *CD9* regulatory regions. In line with previous studies in WM,[15,42] these findings support a transcriptional program associated with the dominant tumor clone, although its functional contribution remains to be established. Pre-clinical studies in myeloma and other B-cell malignancies showed tumor growth inhibition using protein degraders targeting this complex.[43] However, *CD9* was not expressed in all tumor samples in our study. Among *CD9*-positive cases, expression was generally uniform, except for one case (Pnt12), which showed a subclonal pattern. Patients with low or absent *CD9* expression had previously been exposed to treatment and did not harbor genomic alterations affecting its expression. Previous *in vitro* studies using the L3055 cell line showed that *CD9* loss was associated with increased proliferative capacity and lymphoma progression.[44] The heterogeneous distribution of *CD9* across baseline, treated, and genomically distinct tumor subclones in our study suggests that *CD9* reflects a context-dependent tumor cell state rather than a universal marker of all malignant cells. Its confirmation at the protein level needs to be determined in larger cohorts.

Although recent reports suggested that higher soluble BCMA levels were predictive of worst prognosis in WM,[45,46] we observed relatively lower BCMA expression compared with BAFFR and TACI. In contrast, BAFF and APRIL produced by monocytes showed enriched predicted interactions with the dominant B-cell clone, supporting a potential association between myeloid-derived signaling in WM. Given the current interest in BAFFR-targeted therapies,[47,48] these findings warrant validation at the protein and spatial levels in WM.

After integrating genomic and transcriptomic data, we observed that the tumor clone harbored most somatic mutations. Some of these mutations clustered within B-cell clones in patients with an oligoclonal background, supporting oligoclonal competition models described in other B-cell malignancies.[19] In cases dominated by the tumor clone, mutation-defined subclusters suggested marked intratumor heterogeneity. Still, somatic mutations alone are limited to cause clonal expansion in WM.[8] We observed that CNAs tended to cluster preferentially in the tumor clone. For instance, del6q, sometimes accompanied by gain6p, was present in high-risk patients, including an R/R WM patient in whom transformation into a high-grade lymphoma occurred later. Given that previous studies reported that del6q was associated with worst outcomes in IgM MGUS and WM,[49] our data suggest that chr6 alterations may represent late fitness-associated events in a subset of patients. Still, further validation is needed in larger longitudinal cohorts.

It is noteworthy that some minor clones had mutations in *MYD88* and *IGLL5*, but only the *MYD88* mutation was present in nonexpanded cells, while the tumor clone displayed kappa or lambda light-chain restriction. The presence of *MYD88* L265P in non-clonal

25729241, 2026, 9, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hem3.70456 by Readcube (Labtiva Inc.), Wiley Online Library on [26/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

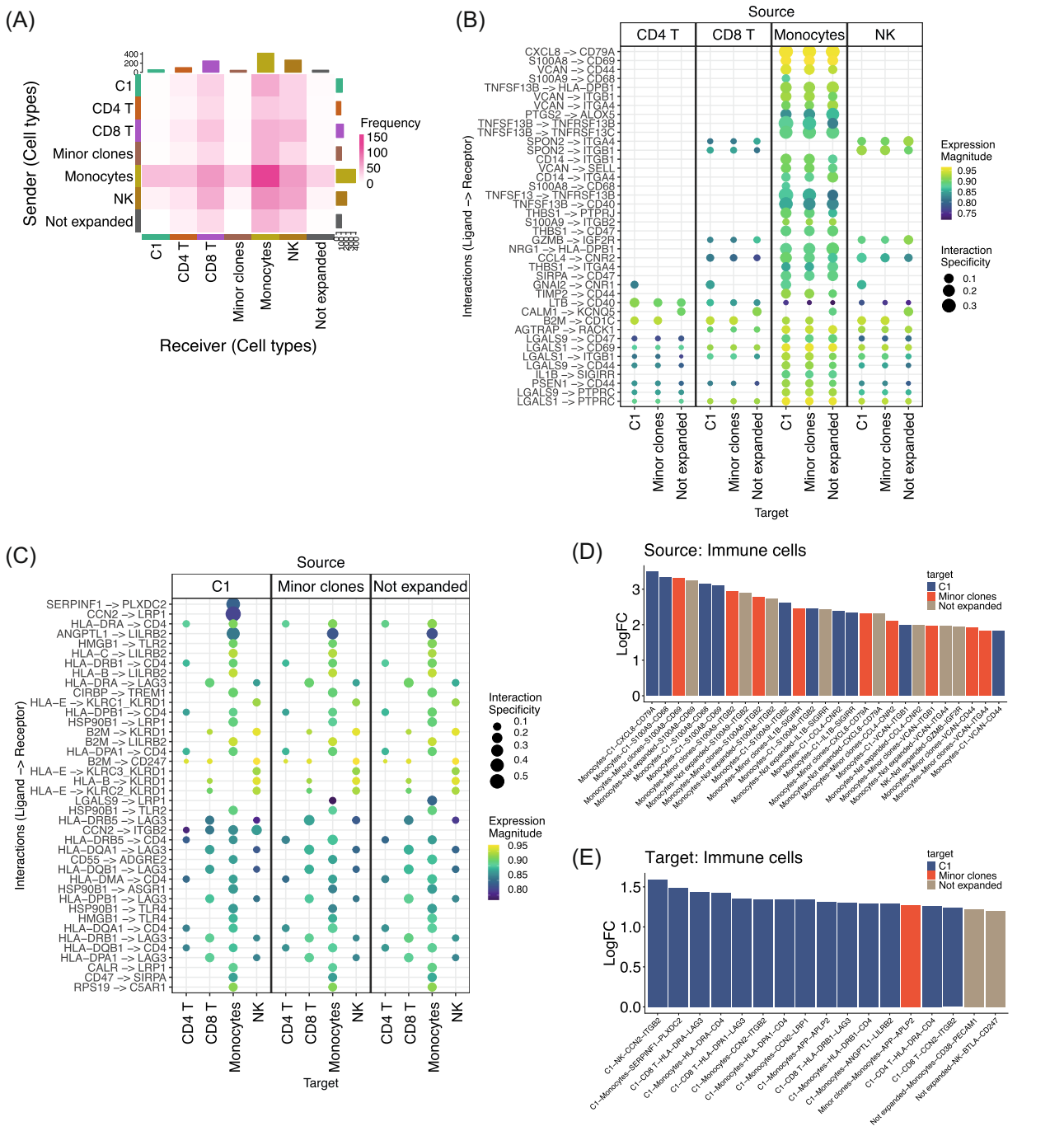


FIGURE 6 **Cell-to-cell communication between B-cell clones and tumor microenvironment cells (CD4 T-cells, CD8 T-cells, monocytes, and NK cells).** **(A)** Ligand–receptor interactions were inferred using LIANA against a Consensus ligand–receptor resource. Statistical significance and interaction prioritization scores were ranked using false discovery rate-adjusted P values. The most frequent interactions were observed with monocytes. **(B, C)** Either as a source or as a target, preferential specific interactions between C1 cells and monocytes are shown. **(D, E)** Distribution of the magnitude (log fold change—LogFC) of the interaction using the immune cells as the source or the target.

B-cells has been reported by others,[8–10] likely associated with age-related processes. In fact, mutations in *MYD88* and *IGLL5* were recently identified as lymphoid clonal hematopoiesis events and were related to aging.[50] However, these mutations were not detectable in minor or nonexpanded B-cells in all cases, which may indicate that they could be acquired at a later developmental cell stage in some patients. However, this could also reflect the limited sensitivity of single-cell RNA-based mutation calling, inter-patient differences in clonal composition, and our conservative filtering. Although *IGLL5* mutations can arise from off-target somatic hypermutation in other B-cell lymphomas[51–53] or may occasionally be confounded by misalignment from other immunoglobulin loci, we observed that all *IGLL5* mutations in our cohort fell in *IGLL5*-specific regions rather than neighboring or cross-mapped immunoglobulin genes. Validation at DNA and single-cell levels is needed to confirm the presence of mutations rather than *MYD88* L265P in minor clones and nonexpanded B-cells.

25729241, 2026, 9, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hem3.70456 by Readcube (Labtiva Inc.), Wiley Online Library on [26/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Our study has limitations regarding the sample size, the possibility that some B-cell clones were missed during integration, and the low sequencing depth at a single-cell resolution, which can limit our detection of mutations in low-expressed genes, such as *MYD88*. Still, the high number of cells from the chromatin accessibility assay and the strong correlation with the RNA data could mitigate this limitation. In addition, we retained only exonic mutations and CNAs detected from WGS data to filter out false-positive genomic alterations at a single-cell resolution. *CXCR4* mutations were not included in our integrative approach because they were small indels and not suitable for the detection of somatic SNVs. Furthermore, the number of longitudinal cases is limited to generalize observations related to the clonal stability over time.

Overall, our study supports a framework in which WM may emerge from an oligoclonal B-cell landscape carrying early mutations such as *MYD88* L265P, with subsequent clonal predominance being associated with genomic, transcriptional, epigenetic, and TME features.

AUTHOR CONTRIBUTIONS

David F. Moreno: Conceptualization; methodology; software; data curation; supervision; resources; funding acquisition; writing—review and editing; writing—original draft; formal analysis. **Ferran Nadeu**: Conceptualization; methodology; software; data curation; formal analysis; writing—original draft; writing—review and editing. **Marta Casado-Peláez**: Methodology; software; data curation; formal analysis. **Caterina Mata**: Methodology; investigation. **Ian Márquez-López**: Software; data curation; formal analysis. **Marco A. Fernández**: Methodology; investigation. **Sergi Vaqué**: Methodology; investigation. **Elena Medina**: Methodology; investigation. **Helena Llobregat**: Methodology; investigation. **Marta Garrote**: Methodology; investigation. **Joan Mañé**: Methodology; investigation; visualization. **Oriol Cardús**: Methodology; visualization; investigation. **Luis Gerardo Rodríguez-Lobato**: Methodology; visualization; investigation. **Anna de Daniel**: Methodology; visualization; investigation. **María Teresa Cibeira**: Methodology; visualization; investigation. **Theo Gremen Oliveira**: Methodology; visualization; investigation. **Elizaveta Kulaeva**: Software; formal analysis. **Ester Lozano**: Methodology; visualization. **Dolors Colomer**: Methodology; investigation; visualization. **Laura Rosiñol**: Methodology; investigation; visualization. **Ruben van Boxtel**: Software; formal analysis. **Manel Esteller**: Conceptualization; methodology; funding acquisition; writing—original draft; writing—review and editing. **Carlos Fernández de Larrea**: Conceptualization; methodology; funding acquisition; writing—original draft; writing—review and editing; project administration.

CONFLICT OF INTEREST STATEMENT

D.F.M. received honoraria from Janssen and BeOne for speaking at educational activities, and travel grants from Janssen and Sanofi. F.N. received honoraria from Janssen, AbbVie, AstraZeneca, and SOPHiA GENETICS, and research funding from Gilead. L.G.R.L. received Honoraria and travel grants from Janssen, Amgen, BMS, GSK, Menarini, and Sanofi. M.T.C. received honoraria for lectures and/or advisory committees from Janssen, Sanofi, GSK, Amgen and Pfizer. D.C. received honoraria from AstraZeneca, Sophia and Thermofisher for speaking at educational activities. R.v.B. is a scientific advisor for Bioskryb Genomics and the Hartwig Medical Foundation. L.R. received honoraria for lectures from Johnson&Johnson, BMS, Amgen, Sanofi, GSK, and Pfizer, and served on advisory boards from Johnson&Johnson, BMS, Amgen, Sanofi, GSK, Pfizer, Menarini, and Oncopeptides. C.F.d.L. served in an advisory board role for Janssen, BMS, Amgen, Pfizer, Sanofi, BeOne, GSK, Roche, and Menarini; received honoraria from Amgen, Janssen, BMS, GSK, Sanofi, Pfizer, and BeOne; and received grants/research support from BMS, Janssen, Amgen, and GSK. The remaining authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Sequencing data from WGS and single-cell experiments are available from the European Genome-phenome Archive (http://www.ebi.ac.uk/ega/) under accession no. EGAS50000002033. IgScan for B-cell receptor annotation analyses is available in https://github.com/nadeulab/IgScan. Downstream data are available within this article as supplementary material. Other remaining data are available on request from the corresponding authors.

FUNDING

This work was supported by grants from the International Waldenström's Macroglobulinemia Foundation (Seed Money and Kyle Award Grants). In addition, this work was supported by grants from the University of Barcelona Chair on Innovation in Precision Oncology and the Spanish Institute of Health Carlos III (cofunded by the EU; PI22/00647 and PI25/00571). D.F.M. acknowledges the European Hematology Association (Mobility Grant—2024) and the Spanish Institute of Health Carlos III (Rio Hortega Grant—2022).

SUPPORTING INFORMATION

Additional supporting information can be found in the online version of this article.

ORCID

David F. Moreno https://orcid.org/0000-0002-1752-3081
Carlos Fernández de Larrea https://orcid.org/0000-0003-4930-9255

REFERENCES

1. Owen RG, Treon SP, Al-Katib A, et al. Clinicopathological definition of Waldenstrom's macroglobulinemia: Consensus Panel Recommendations from the Second International Workshop on Waldenstrom's Macroglobulinemia. *Semin Oncol*. 2003;30(2):110-115. doi:10.1053/sonc.2003.50082
2. Treon SP, Tedeschi A, San-Miguel J, et al. Report of consensus Panel 4 from the 11th International Workshop on Waldenstrom's macroglobulinemia on diagnostic and response criteria. *Sem Hematol*. 2023;60(2):97-106. doi:10.1053/j.seminhematol.2023.03.009
3. Bustoros M, Sklavenitis-Pistofidis R, Kapoor P, et al. Progression Risk Stratification of Asymptomatic Waldenström Macroglobulinemia. *J Clin Oncol*. 2019;37(16):1403-1411. doi:10.1200/JCO.19.00394
4. Moreno DF, Jiménez C, Escalante F, et al. Prognostic risk and survival of asymptomatic IgM monoclonal gammopathy: results from a Spanish Multicenter Registry. *HemaSphere*. 2024;8(11):e70029. doi:10.1002/hem3.70029
5. Paiva B, Montes MC, García-Sanz R, et al. Multiparameter flow cytometry for the identification of the Waldenström's clone in IgM-MGUS and Waldenström's Macroglobulinemia: new criteria for differential diagnosis and risk stratification. *Leukemia*. 2014;28(1):166-173. doi:10.1038/leu.2013.124
6. Paiva B, Corchete LA, Vidriales MB, et al. The cellular origin and malignant transformation of Waldenström macroglobulinemia. *Blood*. 2015;125(15):2370-2380. doi:10.1182/blood-2014-09-602565
7. Bagratuni T, Theologi O, Vlachos C, et al. Genomic landscape of IgM-MGUS and patients with stable or progressive asymptomatic Waldenström macroglobulinemia. *Blood*. 2025;146(25):3086-3097. doi:10.1182/blood.2025030177

25729241, 2026, 9, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hem3.70456 by Readcube (Labtiva Inc.), Wiley Online Library on [26/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

8. Rodriguez S, Celay J, Goicoechea I, et al. Preneoplastic somatic mutations including *MYD88*^{L265P} in lymphoplasmacytic lymphoma. *Sci Adv*. 2022;8(3):eabl4644. doi:10.1126/sciadv.abl4644
9. Moreno DF, López-Guerra M, Paz S, et al. Prognostic impact of *MYD88* and *CXCR4* mutations assessed by droplet digital polymerase chain reaction in IgM monoclonal gammopathy of undetermined significance and smouldering Waldenström macroglobulinaemia. *Br J Haematol*. 2023;200(2):187-196. doi:10.1111/bjh.18502
10. Kaushal A, Nooka AK, Carr AR, et al. Aberrant extrafollicular B cells, immune dysfunction, myeloid inflammation, and MyD88-mutant progenitors precede Waldenstrom Macroglobulinemia. *Blood Cancer Discovery*. 2021;2(6):600-615. doi:10.1158/2643-3230.BCD-21-0043
11. Zanwar S, Le-Rademacher J, Durot E, et al. Simplified risk stratification model for patients with Waldenström Macroglobulinemia. *J Clin Oncol*. 2024;42(21):2527-2536. doi:10.1200/JCO.23.02066
12. 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. doi:10.1182/blood-2013-09-525808
13. 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;136(5):585-595. doi:10.1182/blood.2020005081
14. Moreno DF, Nadeu F, Brasó-Maristany F, et al. Genomic and immune profiling of prognostic risk groups in IgM gammopathy reveals novel biomarkers beyond *MYD88* L265P. *Front Immunol*. 2025;16:1604089. doi:10.3389/fimmu.2025.1604089
15. Gagler DC, Ghamlouch H, Zhang D, et al. A multiomic analysis of Waldenström macroglobulinemia defines distinct disease subtypes. *Blood*. 2025;146(10):1225-1238. doi:10.1182/blood.2024028164
16. Sklavenitis-Pistofidis R, Konishi Y, Heilpern-Mallory D, et al. Single-cell RNA sequencing defines distinct disease subtypes and reveals hypo-responsiveness to interferon in asymptomatic Waldenstrom's Macroglobulinemia. *Nat Commun*. 2025;16(1):1480. doi:10.1038/s41467-025-56323-w
17. Kriangkum J, Taylor BJ, Treon SP, Mant MJ, Belch AR, Pilarski LM. Clonotypic IgM V/D/J sequence analysis in Waldenstrom macroglobulinemia suggests an unusual B-cell origin and an expansion of polyclonal B cells in peripheral blood. *Blood*. 2004;104(7):2134-2142. doi:10.1182/blood-2003-11-4024
18. Dogliotti I, Jiménez C, Varettoni M, et al. Diagnostics in Waldenström's macroglobulinemia: a consensus statement of the European Consortium for Waldenström's Macroglobulinemia. *Leukemia*. 2023;37(2):388-395. doi:10.1038/s41375-022-01762-3
19. Bagnara D, Mazzarello AN, Cardente N, et al. Chronic lymphocytic leukemia often arises by a multiclonal selection process. *Haematologica*. 2025;110(8):1747-1757. doi:10.3324/haematol.2024.286380
20. Da Vià MC, Lazzaroni F, Matera A, et al. Aberrant single-cell phenotype and clinical implications of genotypically defined polyclonal plasma cells in myeloma. *Blood*. 2025;145(26):3124-3138. doi:10.1182/blood.2024025643
21. Sun H, Fang T, Wang T, et al. Single-cell profiles reveal tumor cell heterogeneity and immunosuppressive microenvironment in Waldenström macroglobulinemia. *J Transl Med*. 2022;20(1):576. doi:10.1186/s12967-022-03798-6
22. Hao Y, Stuart T, Kowalski MH, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol*. 2024;42(2):293-304. doi:10.1038/s41587-023-01767-y
23. Stuart T, Srivastava A, Madad S, Lareau CA, Satija R. Single-cell chromatin state analysis with Signac. *Nat Methods*. 2021;18(11):1333-1341. doi:10.1038/s41592-021-01282-5
24. Tirosh I, Izar B, Prakadan SM, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*. 2016;352(6282):189-196. doi:10.1126/science.aad0501
25. Muyas F, Sauer CM, Valle-Inclán JE, et al. De novo detection of somatic mutations in high-throughput single-cell profiling data sets. *Nat Biotechnol*. 2024;42(5):758-767. doi:10.1038/s41587-023-01863-z
26. Dimitrov D, Türei D, Garrido-Rodriguez M, et al. Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data. *Nat Commun*. 2022;13(1):3224. doi:10.1038/s41467-022-30755-0
27. Schep AN, Wu B, Buenrostro JD, Greenleaf WJ. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat Methods*. 2017;14(10):975-978. doi:10.1038/nmeth.4401
28. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2009;25(14):1754-1760. doi:10.1093/bioinformatics/btp324
29. Nadeu F, Royo R, Massoni-Badosa R, et al. Detection of early seeding of Richter transformation in chronic lymphocytic leukemia. *Nat Med*. 2022;28(8):1662-1671. doi:10.1038/s41591-022-01927-8
30. Maclachlan KH, Bagratuni T, Kastritis E, et al. Waldenström macroglobulinemia whole genome reveals prolonged germinal center activity and late copy number aberrations. *Blood Adv*. 2023;7(6):971-981. doi:10.1182/bloodadvances.2022008876
31. Rustad EH, Nadeu F, Angelopoulos N, et al. *mmsig*: a fitting approach to accurately identify somatic mutational signatures in hematological malignancies. *Commun Biol*. 2021;4(1):424. doi:10.1038/s42003-021-01938-0
32. Maura F, Degasperi A, Nadeu F, et al. A practical guide for mutational signature analysis in hematological malignancies. *Nat Commun*. 2019;10(1):2969. doi:10.1038/s41467-019-11037-8
33. Manders F, Brandsma AM, De Kanter J, et al. MutationalPatterns: the one stop shop for the analysis of mutational processes. *BMC Genomics*. 2022;23(1):134. doi:10.1186/s12864-022-08357-3
34. Díaz-Gay M, Vangara R, Barnes M, et al. Assigning mutational signatures to individual samples and individual somatic mutations with SigProfilerAssignment. *Bioinformatics*. 2023;39(12):btad756. doi:10.1093/bioinformatics/btad756
35. Mondello P, Paludo J, Novak JP, et al. Molecular clusters and tumor-immune drivers of IgM monoclonal gammopathies. *Clin Cancer Res*. 2023;29(5):957-970. doi:10.1158/1078-0432.CCR-22-2215
36. Hodson DJ, Shaffer AL, Xiao W, et al. Regulation of normal B-cell differentiation and malignant B-cell survival by OCT2. *Proc Natl Acad Sci*. 2016;113(14):E2039-46. doi:10.1073/pnas.1600557113
37. Song S, Cao C, Choukrallah MA, et al. OBF1 and Oct factors control the germinal center transcriptional program. *Blood*. 2021;137(21):2920-2934. doi:10.1182/blood.2020010175
38. Sun H, Liu X, Guerrera ML, et al. Single-cell RNA analysis of bone marrow mononuclear cells reveals abnormal activation of T cells and monocytes, and identifies CD35 as a novel biomarker and therapeutic target in Waldenstrom Macroglobulinemia. *Blood*. 2024;144(Suppl 1):743. doi:10.1182/blood-2024-211488
39. 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):104856. doi:10.1016/j.isci.2022.104856
40. Yoon SO, Zhang X, Lee IY, Spencer N, Vo P, Choi YS. CD9 is a novel marker for plasma cell precursors in human germinal centers. *Biochem Biophys Res Commun*. 2013;431(1):41-46. doi:10.1016/j.bbrc.2012.12.102
41. Yoon SO, Lee IY, Zhang X, Zapata MC, Choi YS. CD9 may contribute to the survival of human germinal center B cells by facilitating the interaction with follicular dendritic cells. *FEBS Open Bio*. 2014;4:370-376. doi:10.1016/j.fob.2014.04.001
42. 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. *Br J Haematol*. 2014;166(5):677-689. doi:10.1111/bjh.12936

25729241, 2026, 9, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hem3.70456 by Readcube (Labtiva Inc.), Wiley Online Library on [26/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

43. He T, Xiao L, Qiao Y, et al. Targeting the mSWI/SNF complex in POU2F-POU2AF transcription factor-driven malignancies. *Cancer Cell*. 2024;42(8):1336-1351.e9. doi:10.1016/j.ccell.2024.06.006
44. Yoon SO, Zhang X, Freedman AS, et al. Down-regulation of CD9 expression and its correlation to tumor progression in B lymphomas. *Am J Pathol*. 2010;177(1):377-386. doi:10.2353/ajpath.2010.100048
45. Oura M, Ikeda D, Aikawa S, et al. Soluble BCMA as a biomarker reflecting tumor volume and treatment response in Waldenström macroglobulinemia. *Leuk Lymphoma*. 2026;67(1):85-96. doi:10.1080/10428194.2025.2569759
46. Goldwater M, Moore J, Bujarski S, et al. Serum B-Cell Maturation Antigen as a Biomarker for Predicting Outcomes and Monitoring Patients With Waldenström Macroglobulinemia. *Eur J Haematol*. 2026;116(1):31-36. doi:10.1111/ejh.70029
47. Qin H, Dong Z, Wang X, et al. CAR T cells targeting BAFF-R can overcome CD19 antigen loss in B cell malignancies. *Sci Transl Med*. 2019;11(511):eaaw9414. doi:10.1126/scitranslmed.aaw9414
48. Wong DP, Roy NK, Zhang K, et al. A BAFF ligand-based CAR-T cell targeting three receptors and multiple B cell cancers. *Nat Commun*. 2022;13(1):217. doi:10.1038/s41467-021-27853-w
49. García-Sanz R, Dogliotti I, Zaccaria GM, et al. 6q deletion in Waldenström macroglobulinaemia negatively affects time to transformation and survival. *Br J Haematol*. 2021;192(5):843-852. doi:10.1111/bjh.17028
50. Bernstein N, Spencer Chapman M, Nyamondo K, et al. Analysis of somatic mutations in whole blood from 200,618 individuals identifies pervasive positive selection and novel drivers of clonal hematopoiesis. *Nat Genet*. 2024;56(6):1147-1155. doi:10.1038/s41588-024-01755-1
51. Kasar S, Kim J, Improgo R, et al. Whole-genome sequencing reveals activation-induced cytidine deaminase signatures during indolent chronic lymphocytic leukaemia evolution. *Nat Commun*. 2015;6:8866. doi:10.1038/ncomms9866
52. Knisbacher BA, Lin Z, Hahn CK, et al. Molecular map of chronic lymphocytic leukemia and its impact on outcome. *Nat Genet*. 2022;54(11):1664-1674. doi:10.1038/s41588-022-01140-w
53. Ye X, Ren W, Liu D, et al. Genome-wide mutational signatures revealed distinct developmental paths for human B cell lymphomas. *J Exp Med*. 2021;218(2):e20200573. doi:10.1084/jem.20200573

25729241, 2026, 9, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hem3.70456 by Readcube (Labtiva Inc.), Wiley Online Library on [26/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License