blood Regular Article

LYMPHOID NEOPLASIA

Genomic landscape of IgM-MGUS and patients with stable or progressive asymptomatic Waldenström macroglobulinemia

Tina Bagratuni,[1,*] Ourania Theologi,[1,*] Christos Vlachos,[1] Ioannis Kollias,[1] Kylee Maclachlan,[2] Foteini Aktypi,[1] Nefeli Mavrianou-Koutsoukou,[1] Christine Liacos,[1] Konstantina Taouxi,[1] Alexandra Papadimou,[1] Katerina Chrisostomidou,[1] Maria Sakkou,[1] Irene Solia,[1] Foteini Theodorakakou,[1] Gianmarco Favrin,[3] Maria Gavriatopoulou,[1] Evangelos Terpos,[1] Marzia Varettoni,[4] Zachary R. Hunter,[5] Steven Treon,[5] Francesco Maura,[6] Meletios A. Dimopoulos,[1,†] and Efstathios Kastritis[1,†]

[1]Department of Clinical Therapeutics, School of Medicine, National and Kapodistrian University of Athens, Athens, Greece; [2]Myeloma Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY; [3]Department of Molecular Medicine, University of Pavia, Pavia, Italy; [4]Division of Hematology, Fondazione Istituto di Ricovero e Cura a Carattere Scientifico, Policlinico San Matteo, Pavia, Italy; [5]Bing Center for Waldenström's Macroglobulinemia, Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA; and [6]Myeloma Service, Sylvester Comprehensive Cancer Center, University of Miami, Miami, FL

KEY POINTS

- **Whole-exome sequencing demonstrates that genomic data can enhance the prediction of progression from asymptomatic to symptomatic disease.**
- **Patients with IgM-MGUS display 2 genetically distinct entities, with "benign" and "WM-like" genomic features.**

Immunoglobulin M (IgM) monoclonal gammopathy of undetermined significance (IgM-MGUS) and asymptomatic Waldenström macroglobulinemia (WM; aWM) are precursor conditions of symptomatic WM with an annual 1.5% to 12% risk of progression. Although clinical prognostic models exist for risk stratification, it remains challenging to distinguish asymptomatic patients who will eventually progress from those who will not. Hence, the characterization of genomic features that shape disease progressors could potentially improve risk stratification. We performed whole-exome sequencing on 229 samples from 139 patients, including 9 patients with sequential samples. We observed an increasing mutation burden through the stages of disease evolution. Genes such as *CD79B*, *ARID1A*, and *CREBBP* were more often mutated in the aWM progressed (aWMpr) compared with the stable aWM (aWMst) group, whereas *MYD88*^{L265} variant allele frequency was significantly higher in patients with aWMpr than those with aWMst. In addition, patients with IgM-MGUS with *MYD88*^{WT} genotype showed a distinct genomic profile compared with the *MYD88*^{MUT} patients. Furthermore, the presence of more aneuploidies showed a significant association with a higher risk of progression to the symptomatic disease. Overall, our study shows that genomic profiling of patients' tumor at the time of aWM diagnosis might represent an improved strategy for identifying patients at high risk to progression who could benefit from earlier intervention.

Introduction

Waldenström macroglobulinemia (WM) is a B-cell lymphoproliferative disorder characterized by bone marrow (BM) infiltration of small lymphoplasmacytic lymphoma cells that secrete monoclonal immunoglobulin M (IgM).[1,2] Genomically, besides the *MYD88* L265P mutation, which is present in >90% of patients and is regarded as a hallmark of WM,[3] recurrent mutations in the *CXCR4* (30%-40%) and *ARID1A* (17%) genes, along with deletions in chromosome 6q (30%-50%), are considered secondary molecular events in the disease.[4-6] Genomic[7-9] and transcriptional[10,11] alterations have also resulted in disrupted signaling pathways, including nuclear factor κB, phosphatidylinositol 3-kinase (PI3K)/ protein kinase B (AKT), and MAPK, which promote tumor cell growth and increase resistance to apoptosis.[12-15] Furthermore, immune deregulation in combination with distinct tumor-intrinsic gene expression patterns have been identified to be linked to disease progression.[16,17]

Patients with the "premalignant" condition known as IgM monoclonal gammopathy of undetermined significance (IgM-MGUS) and asymptomatic WM (aWM) are at higher risk of developing symptomatic WM (sWM),[18,19] and this transition is estimated at a rate of 1.5% to 12%, respectively, per year.[19] Current prognostic models for IgM monoclonal gammopathy risk stratification are based only on clinical variables,[19,20] and the underlying biology for the progression risk from the asymptomatic phases as well as the genomic and phenotypic landscape driving this malignant transformation have not been extensively studied so far.[21-28] In this study, we aimed to comprehensively characterize the genomic landscape of WM precursor conditions, and identify potential biomarkers of progression to WM by means of whole-exome sequencing (WES) in 139 patients with IgM monoclonal gammopathies.

Methods

WES analysis was performed on a total of 229 genomic DNA samples. Matched BM and peripheral blood samples (tumor-normal samples) were collected from 76 patients, which also included 9 sequential samples of patients who progressed from the asymptomatic phase to symptomatic disease. The validation cohort included 63 patients, of whom 22 patients had matched tumor and normal samples, whereas the remaining 41 were tumor-only samples. Before any study procedure, all study participants provided a written informed consent form for the collection and analysis of their biological samples according to the Declaration of Helsinki.

Time to progression (TTP) was calculated from the date of diagnosis to the date of the first assessment showing evidence of symptomatic disease requiring treatment. For survival analysis, the Kaplan-Meier method was used to estimate the cumulative incidence of progression, and differences between the curves were tested by the log-rank test. A univariate Cox proportional hazards model was used to identify risk factors for disease progression among all patients. We also used random survival forests (RSFs) R package,[29] a machine learning approach tailored for survival data effective for examining patient outcomes over time, while accounting for right-censored data. Detailed methods are described in supplemental Methods, available on the *Blood* website.

Results

Patients

We interrogated the WES profiles of 76 (34 with IgM-MGUS, 30 with aWM, and 12 with sWM) and 63 (21 with IgM-MGUS, 40 with aWM, and 2 with sWM), newly diagnosed previously untreated patients with IgM. Within the combined aWM cohort (n = 70), and after a median follow-up of 48 months (range 1-408), 19 patients (25%) had progressed (aWMpr) to sWM, and 51 (75%) remained stable (aWMst). The characteristics of the patients in the initial and validation cohorts at the time of initial diagnosis are summarized in Table 1, supplemental Table 1, and supplemental Figure 1A.

Genomic landscape of patients with IgM monoclonal gammopathies

Overall, the WES analysis of the entire (initial and validation) cohort of patients (n = 76 and n = 63, respectively) identified 57 109 single-nucleotide variants (SNVs), of which 8061 were somatic nonsynonymous; 20 536 insertions and deletions (indels), of which 220 were nonsynonymous small insertions; and 1019 nonsynonymous small deletions. Within the coding region, the frequency of mutations across genomic domains was variable, where missense mutations had the highest mutational load (supplemental Figure 1B). We observed a significantly increasing tumor mutation burden (defined as the total number of nonsynonymous mutations and indels per patient divided by the total exome size) through the stages of disease evolution in both the initial and the validation cohorts (Figure 1A; supplemental Figure 1C-D). Among the 139 patients, the mutation burden was significantly higher in $MYD88^{L265P}$ (n = 111) patients than in $MYD88^{WT}$ patients (n = 28; median, 0.9 vs 0.567, respectively; Wilcoxon rank-sum test, $P = 1.35 \times 10^{-5}$; supplemental Figure 1E-G). Among patients with IgM-MGUS (n = 55), those with $MYD88^{L265P}$ (n = 33) exhibited numerically higher mutation burden than those with $MYD88^{WT}$ IgM-MGUS (n = 22) in genes such as *KMT2C* (15% vs 9%), *IGLL5* (18% vs 0%), and *DNAH11* (18% vs 0%; supplemental Figure 2A). Importantly, within the IgM-MGUS cohort, we observed a subset of patients (8/34 patients in the initial and 4/21 in the validation cohort) who did not carry any *MYD88* or any other WM-related gene mutations while they had similar clinical characteristics (Figure 1B; supplemental Figure 2B; supplemental Table 2).

Furthermore, we examined the $MYD88^{L265P}$ mutational variant allele frequency (VAF) in the 4 patient groups and observed an increasing $MYD88^{L265P}$ VAF from IgM-MGUS to sWM disease stages, with a significantly higher $MYD88^{L265P}$ VAF in patients with aWMpr (median VAF = 0.38) than in those with aWMst

Table 1. Characteristics of patients in the initial cohort

	IgM-MGUS (n = 34)	aWMst (n = 21)	aWMpr (n = 9)	sWM (n = 12)
Light chain type, κ/λ	27/7	13/8	7/2	9/3
Age, median (range), y	68 (47-89)	63 (44-82)	72 (50-82)	74 (57-88)
BM infiltration, median (range), %	3 (0-20)	20 (0-60)	25 (0-65)	60 (20-80)
IgM levels, median (range), mg/dL	430 (123-1630)	819 (194-4010)	2050 (547-6180)	3090 (438-8390)
Hemoglobin, median (range), g/dL	13.4 (11-14.8)	13.3 (11-16.2)	11.55 (10.9-13.1)	10.45 (8.5-12)
β2-Microglobulin, median (range), mg/L	1.95 (1.29-4.62)	1.9 (1.56-4.12)	2.63 (1.81-3.57)	3.2 (2-5.13)
$MYD88^{L265P}$	20/34	17/21	9/9	12/12
Paired BM and PB samples	34/34	21/21	9/9	12/12

PB, peripheral blood.

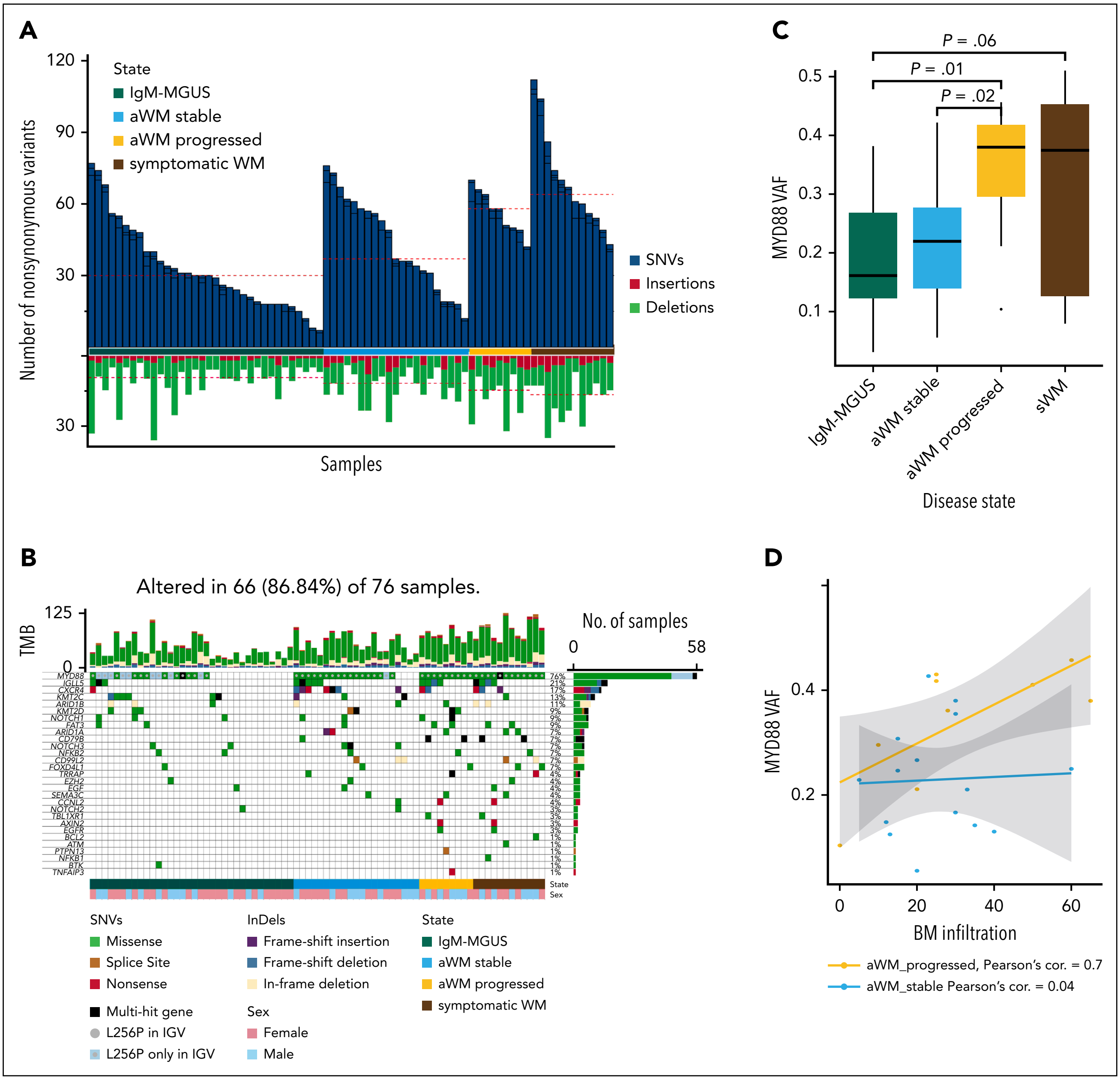

Figure 1. Overall mutational load in 76 patients with IgM monoclonal gammopathy at different clinical stages (IgM-MGUS, asymptomatic stable, asymptomatic progressed, symptomatic). (A) Bar plot showing the total number of nonsynomymous somatic alterations (y-axis) per patient. SNVs are shown on the top panel, indels are shown at the bottom panel for each patient, and samples are ordered from highest number of SNVs to lowest. (B) Oncoplot of nonsynonymous mutations (SNVs and indels) in the initial cohort (n = 76), highlighting selected genes relevant to WM biology, grouped by clinical stage. Variants are annotated with different colors representing mutation types. *MYD88* mutations, highlighted with light blue boxes, were manually identified using the IGV viewer. (C) Boxplot illustrating the MYD88^{L265P} mutational VAF across the 4 cohorts of patients. (D) Scatterplot of *MYD88*^{L265P} VAF (y-axis) vs BM infiltration (x-axis) between the cohorts of asymptomatic progressed (yellow) and stable (blue) patients. IGV, immunoglobulin variable region; TMB, tumor mutation burden.

(median VAF, 0.38 vs 0.21; *P* = .02; Figure 1C; supplemental Figure 2C), and a higher Pearson correlation between the BM infiltration and *MYD88*^{L265P} VAF in the aWMpr group (*r* = 0.7; *P* = .026), which was not observed in the aWMst group (*r* = 0.04; *P* = .87), suggesting that *MYD88*^{L265P} VAF could potentially become a more reliable marker for monitoring disease progression (Figure 1D; supplemental Figure 2D).

We further analyzed the most frequently mutated genes (SNVs and indels) relevant to WM biology (*CXCR4*, *KMT2C*, *NOTCH1*, *ARID1B*, *KMT2D*, *ARID1A*, and *CD79B*) in the entire cohort of patients (initial, n = 76; validation, n = 63; Figure 1B; supplemental Figure 2B).

Focusing on the aWM group, we observed that patients with aWMpr were characterized by a higher number of coding mutations in genes previously described in WM[4,7,8] (n = 3 mutations per patient; range, 1-10) with a median mutation density of 0.5 per Mb, than those with aWMst (n = 2 mutations per patient; range, 0-6) with a median mutation density of 0.03

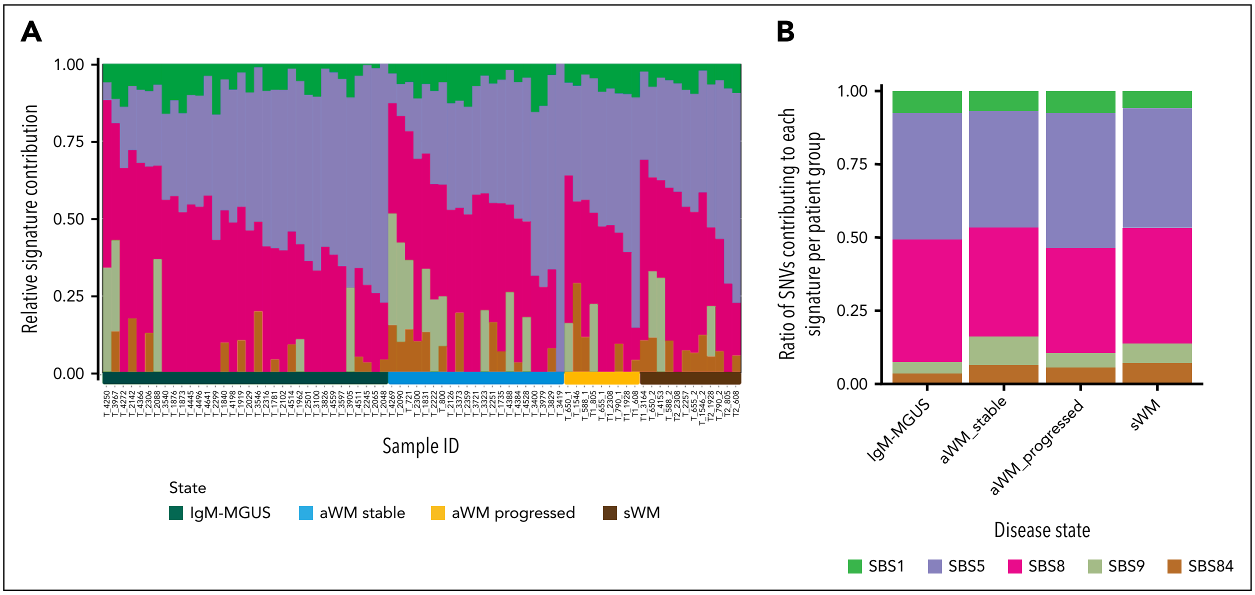

Figure 2. SBS signatures in the 4 cohorts of patients. (A) Mutational signature contributions for each patient across the 4 clinical stages of WM. (B) Mutational signature contributions per group of patients.

per Mb (Wilcoxon rank-sum test P = .483). Mutated genes mostly seen in patients with aWMpr (n = 19) than those with aWMst (n = 51) from the entire cohort (initial and validation) included *MYD88* (95% vs 92%), *CD79B* (21% vs 8%), *TTN* (21% vs 2%), *CREBBP* (16% vs 6%), *ARID1A* (16% vs 6%), and *BCL11B* (11% vs 0%); however, none of these differences reached statistical significance (supplemental Figure 3A). Interestingly, *CXCR4* was more frequently mutated in patients with aWMst compared with patients with aWMpr (27% vs 21%).

To identify potential driver mutations, we ran the dndscv[30] R package against both an unrestricted and a restricted hypothesis for 604 known oncogenic drivers (MSK-IMPACT Heme panel) as previously described.[29] This analysis was performed separately on the initial cohort, which included paired BM and peripheral blood samples, and also on the combined initial and validation cohorts to enhance the robustness of the findings. Our analysis using the Heme panel in both the initial and validation cohorts identified 5 potential driver genes, including *MYD88*, *CXCR4*, *CD79B*, *ARID1A*, and *H1-3*, whereas the unrestricted analysis identified 25 potential drivers, encompassing most of the aforementioned genes (supplemental Figure 3B; supplemental Tables 3 and 4). The aforementioned driver genes were mainly mutated in aWM (both stable and progressors) and sWM groups, although key alterations such as the *MYD88* ^{L265P} mutation also occur in patients with IgM-MGUS.

Overall, in our entire cohort (initial and validation), we found nonsynonymous alterations in the nuclear factor-κB pathway (in 70% with *MYD88* and 4% with *TNFAIP3* mutations), 20% had alterations in the MAPK pathway, 4% in the extracellular signal-regulated kinase (ERK) pathway (*MAPK1* and *EGFR*), and 7% in the p53 pathway (*BAX* and *BCL2*; supplemental Figure 4).

The contribution of mutational signatures in the progression to sWM

To investigate whether the mutation burden reflects the underlying mutational processes operative in precursor conditions of WM, we examined the single-base substitution (SBS) mutational signature landscape in each patient sample using only the initial cohort, downsampling mutational matrices to the exome regions of the genome. To estimate the activity of previously reported SBS mutational signatures,[8] we applied the fitting algorithm mmsig designed for multiple myeloma. These mutational signatures were SBS1 and SBS5 (associated with aging) with median contribution of 48% in the initial cohort (n = 76), SBS8 (genomic instability) with median contribution of 39% in the entire cohort, SBS9 (reflecting poly-eta activity in germinal center) active in 19 patients (mostly aWMst) with median contribution of 22%, and SBS84 (associated with the activation-induced cytidine deaminase activity) active in 35 patients with a median contribution of 10% (Figure 2A). The contribution of age-related signatures SBS1/SBS5 was weakly correlated with the age at diagnosis (Spearman correlation, ρ = 0.23; P = .05). Furthermore, in agreement with our previously published study,[8] we did not find APOBEC-enriched samples in our patients. Overall, no significant differences in mutational signatures were observed across the 4 patient cohorts (Figure 2B).

The landscape of CNAs during progression to sWM

When exploring the copy number alteration (CNA) landscape in the entire (initial and validation) cohort (n = 139), we found that patients with IgM-MGUS (n = 55) and aWMst (n = 51) had significantly fewer CN aberrations than those with aWMpr (n = 19) and sWM (n = 14; Kruskal-Wallis rank-sum test, P < .01), who were mainly characterized by recurrent chromosomal abnormalities, including del6q, gain18, and gain3q (Figure 3A-B; supplemental Figure 5A). In agreement with our previously published study,[8] del6q, one of the main chromosomal aberrations in WM, was significantly more frequent in patients with sWM compared with all precursor conditions and showed a gradual increase through disease evolution, being observed in 3 of 55 patients with IgM-MGUS (5.4%), 6 of 70 patients with

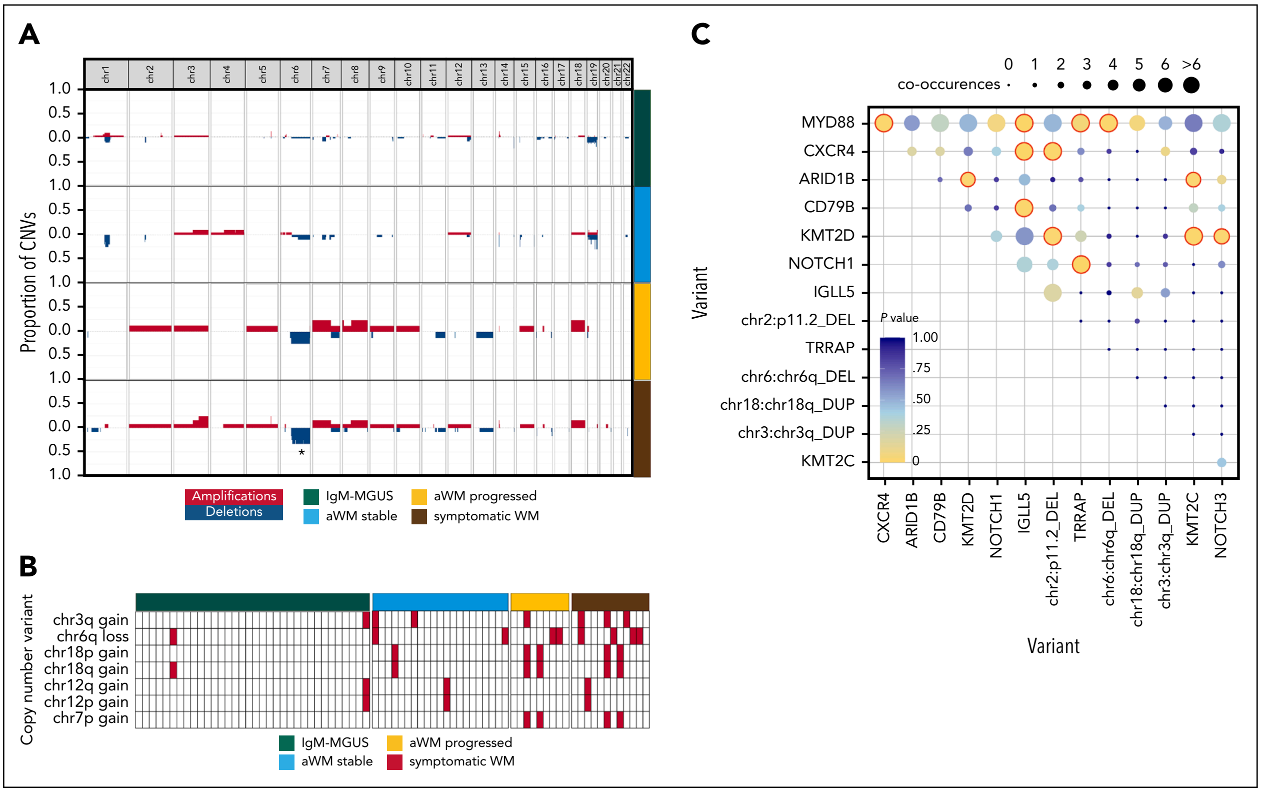

Figure 3. Copy number profiles and correlation matrix across the 4 clinical stages of WM. (A) Cumulative copy number profiles across the 4 clinical stages of WM. Red represents gain; blue represents deletion. *Peak is significantly different between groups according to *GISTIC* analysis. (B) Heat map of large chromosomal aberrations grouped by clinical stage. (C) Correlation matrix of the 76 patients who demonstrated associations and co-occurrence of selected WM-related genes. The size of the bubble corresponds to the number of co-occurrences. The red circle around the bubble indicates the significance value. Chr, chromosome.

aWMst (8.5%), 7 of 19 patients with aWMpr (37%), and 4 of 14 patients with sWM (28%) of the entire cohort. Similar CNA patterns were observed among the initial and validation cohorts.

In terms of SNV and CNA co-occurrences (including both nonsynonymous and synonymous alterations) in WM precursor conditions, we observed that *MYD88*^{L265P} significantly co-occurred with most WM relevant genes as well as most frequently seen CNAs. As such, *MYD88*^{L265P} significantly co-occurred with SNVs and indels in *CXCR4*, *IGLL5*, and *TRRAP*, as well as with del6q (Figure 3C).

Clonal architecture and evolution in precursor WM conditions

We next sought to determine the clonal architecture of mutations that affected recurrently mutated genes relevant to WM as well as mutations in genes that were identified in this analysis using only the initial cohort (supplemental Figure 6A). We observed heterogeneity of mutations that affected recurrently mutated genes in WM.As expected, SNVs in *MYD88* gene were clonal, as were SNVs in genes such as *CD79B* and *IGLL5*. In contrast, SNVs in genes such as *CXCR4*, *EZH2*, *KMT2D*, *NFKB1*, *ARID1A*, *ARID1B*, and *NOTCH3* were subclonal, suggesting that they are more likely to represent later events in the disease evolution (Figure 4A; supplemental Figure 6B). Regarding the most frequent CNAs, del6q and dup18q were clonal in 64% and 75% of the patients in our initial cohort (n = 76), respectively, suggesting that CNAs are early events of the precursor conditions (Figure 4C).

To evaluate the time of activation of specific mutational processes during the development of WM, we examined the differences in signature contribution in clonal vs subclonal mutations. The contributions of most mutational signatures were comparable between total clonal and subclonal mutations, with the exception of the SBS9 signature, which was initially observed only in the clonal compartment (Figure 4B). However, SBS9 was present in only 19 of 76 patients of the initial cohort, and when we reanalyzed the data by aggregating all exome-wide mutations from these 19 patients by separating them into clonal and subclonal categories, we detected the SBS9 signature in both compartments (supplemental Figure 6C). This is probably due to the fact that WES data are primarily active in noncoding genomic regions,[31] and when mutations are collapsed across all samples, more dominant signatures overshadow SBS9, making it difficult to detect at the subclonal level. We next plotted the contribution of each mutational signature from clonal to subclonal mutations, and our results showed that SBS8 signature increased through subclonal evolution (Figure 4D).

The clonal evolution of 9 patients asymptomatic patients who progressed to symptomatic disease was analyzed using sequential samples. All showed clonal heterogeneity, with distinct mutation clusters evolving over time. Two evolutionary patterns were observed; a linear evolution in most patients (6/9; Figure 5A-B,D), and a branching evolution in the rest (3/9; Figure 5C,E-I). The most common CNA, del6q, was present in 2 patients during the asymptomatic stage and was acquired as

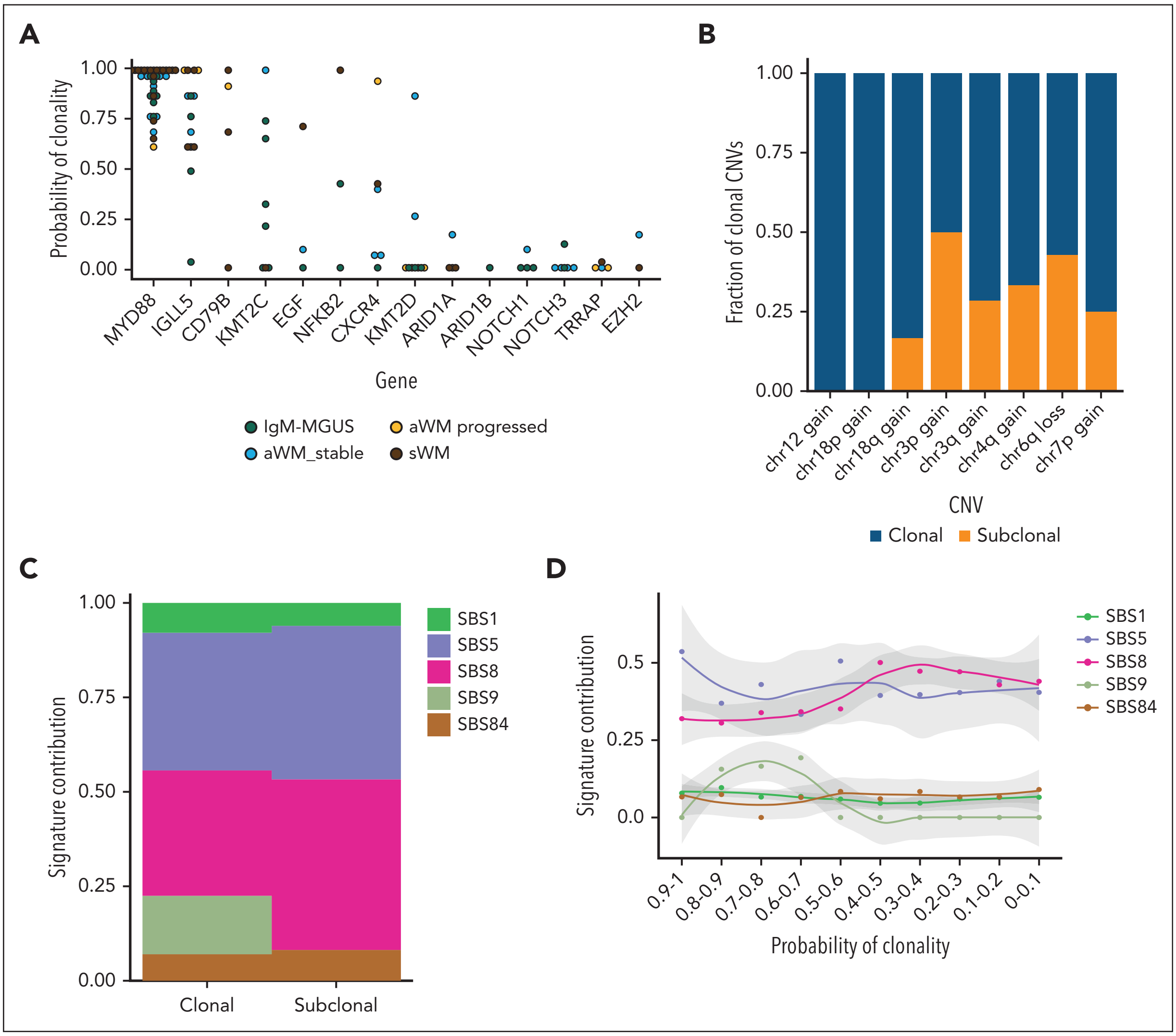

Figure 4. Clonal and subclonal contributions of SNVs and mutational signatures across the 4 cohorts of patients. (A) The clonal proportions of recurrent somatic nucleotide variants across the samples. Colors represent each group of patients. (B) Clonal architecture of recurrent CNAs in the entire cohort. (C) SBS mutational signatures obtained after collapsing all exome-wide mutations at the clonal and subclonal level. (D) Contribution (y-axis) of each mutational signature (color-coded lines) from clonal mutations to subclonal mutations (x-axis). Colors represent each SBS signature. Chr, chromosome.

a clonal event in 1 patient during progression, suggesting it may play a significant role in disease advancement. Similarly, dup18q was present as a clonal event in 2 patients at the asymptomatic stage (Figure 5F,I). Regarding SNVs, we identified several alterations in genes such as *CXCR4*, *CD79B*, *FAT3*, *NFKB2*, and *EGFR*, present already at the asymptomatic stage. In contrast, many subclonal mutations acquired later involved genes such as *ARID1A*, *ARID1B*, *KMT2C*, *MAP3K*, *CD99L*, *NOTCH4*, *TNFRSF1A*, and HLA-related genes. In 1 patient with sequential sampling after treatment, we observed the persistence of the *FAT3* clone, whereas subclones with *ARID1B* and *CD99L* mutations, acquired at disease progression, were lost (Figure 5I).

Genomic events for the prediction to progression to symptomatic disease

We next sought to determine whether independent genomic features could serve as predictors for the progression from the asymptomatic to the symptomatic disease. In our initial cohort of asymptomatic patients (n = 64), the median TTP from the aWM to sWM was 48 months (95% confidence interval, 15-70). The main clinical features of patients with aWMst and aWMpr are shown in Table 2. We initially aimed to assess whether the presence of the *MYD88* mutation could serve as an independent genomic predictor of disease progression. However, Kaplan-Meier analysis revealed no statistically significant association between *MYD88* mutation status and progression (P = .09; supplemental Figure 7A). Next, we evaluated whether the *MYD88*^{L265P} VAF could serve as a risk factor, especially as it showed a significant correlation with BM infiltration in the patients with aWMpr. We found that the *MYD88*^{L265P} VAF was indeed a significant predictor of disease progression (P = .016), indicating that VAF levels may mirror the clonal tumor burden, and could potentially serve as a potential biomarker for monitoring disease progression and identifying high-risk patients (Figure 6A). This association has also been supported by previous studies.[9,32,33]

To provide a more comprehensive framework for predicting disease progression, we used RSFs, a machine learning

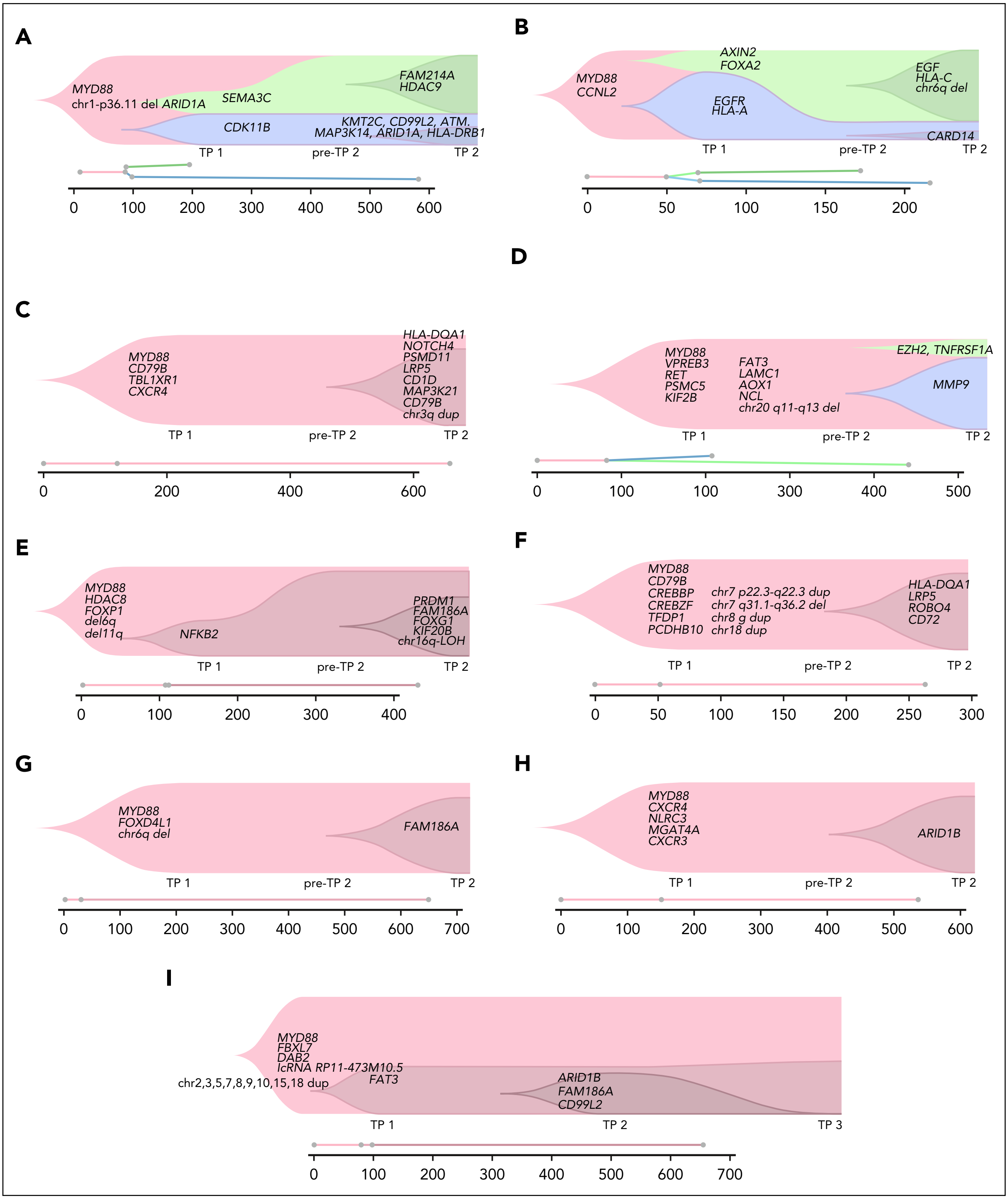

Figure 5. Fish plots of 9 patients with aWM who progressed to the symptomatic disease. (A-B) Patients with branching evolution. (C-H) Patients with linear evolution. Time points correspond to the time of aWM diagnosis (TP 1) and the TTP to sWM (TP 2). (I) Patient at the time of aWM diagnosis (TP 1), TTP (TP 2), and after treatment (TP 3). TP, time point.

Table 2. Risk factors for disease progression in patients with aWM (initial cohort)

Univariate analysis		
Predictive variable	**HR (95% CI)**	***P* value**
Age, y	1.01 (0.94-1.07)	.8
Sex, male	2.29 (0.54-9.80)	.3
Baseline IgM level of <3000 mg/dL	0.08 (0.02-0.37)	.001
Baseline hemoglobin of <13 g/dL	14.3 (1.77-115)	.013
Lymphocyte count of <1.5 x 10^3/μl	0.7 (0.15-3.40)	.7
WBC count of <6 x 10^3/μl	1.13 (0.22-5.87)	>.9
Albumin level of <4.0 g/dL	9.19 (2.02-41.7)	.004
MYD88 VAF of <0.4	0.21 (0.05-0.84)	.028
BMI <10	0.39 (0.08-1.93)	.2

BMI, bone marrow infiltration; CI, confidence interval; HR, hazard ratio; WBC, white blood cell.

approach tailored for survival data. RSFs construct hundreds to thousands of decision trees, each of which uses a random subset of clinical and mutational predictors to stratify patients into subgroups with similar survival outcomes. This ensemble-based method aggregates the outputs of all trees developed by the RSF model and generates a "risk of progression" score for each patient (supplemental Figure 7B). For this, we integrated as input the presence of the 42 most recurrently WM-related genes and 10 copy number variations (CNVs) in addition with the basic clinical parameters (IgM, BM infiltration, beta-2 microglobulin [b2M], and albumin) for each patient. Receiver operative characteristic was used to determine an optimal risk score threshold, allowing us to stratify the initial cohort into high- and low-risk groups based on clinical and genomic features (Figure 6B; supplemental Table 5; supplemental Figure 8), which assigned all patients with aWMpr as high risk, along with 3 patients with aWMst, and 3 patients with IgM-MGUS. Patients with no WM relevant mutations demonstrated low prediction risk scores (low risk), possibly suggesting that they may retain their precursor stage for longer time. According to the clinical scoring system developed by Bustoros et al,[34] 4 of 9 patients with aWMpr were characterized as low risk, and 1 patient as an intermediate-risk group, possibly suggesting that the addition of the genomic events provides an improved tool for stratification of patients with aWM. In fact, focusing on the genomic features only, we observed that patients with a high number of genomic events (>4 SNVs and CNVs) had a significantly shorter TTP (P < .0001), with no events observed in the low-risk group (Figure 6C). These genomic events were mainly driven by SNVs in genes such as *CD79B*, *COLEC12*, *CTNND2*, *NOTCH3*, and *ARID1B*, and CNAs in chromosomes 6, 7, 18, and 22. Furthermore, Kaplan-Meier survival analysis and log-rank test showed statistically significant differences in TTP between the 2 risk groups (P < .0001; Figure 6D; supplemental Figure 8), whereas Cox proportional hazards modeling confirmed a significant difference in risk between the groups (likelihood ratio test, χ^2 = 18.12; degrees of freedom [df] = 1; $P = 2.07 \times 10^{-5}$); however, due to small numbers of events and quasi-complete separation, reliable estimation of the hazard ratio and confidence intervals was not possible.

Internal and external validation of developed prediction model

To validate our genomic prediction model using RSF with clinical and genomic features, we used an internal validation cohort of 34 patients with IgM monoclonal gammopathy (21 with IgM-MGUS and 13 with aWM) and an external validation cohort of 27 patients with aWM from Pavia, Italy. The main clinical features of patients with aWMst and aWMpr are shown in supplemental Table 6. Three of 34 patients and 7 of 27 patients progressed to symptomatic disease from the internal and external cohorts, respectively. Median TTP was 36 months (range, 24-60) for the internal cohort, whereas median TTP was 96 months (range, 61.5-134) for the external cohort. We found that within our internal validation cohort, 2 of 3 progressing patients were characterized as high risk, carrying >4 high-risk genomic features and with a median TTP of 2.5 years. One progressing patient was characterized as low risk, but was close to the threshold of the high-risk group with a TTP of ~80 months. Within the external validation cohort, 4 of 7 patients were characterized as high risk, whereas the remaining were characterized as low risk; however, 2 of them transformed directly to diffuse large B-cell lymphoma (DLBCL). In agreement with our initial cohort, we observed that patients with a high number of genomic events (>4 SNVs and CNVs) had a significantly shorter TTP (60 vs 120 months, respectively; P < .001; Figure 6E). Cox proportional hazards modeling in the entire validation cohort showed that grouping of patients derived from the risk scores was significantly associated with TTP (hazard ratio, 0.27; P = .029; Figure 6F; supplemental Table 7; supplemental Figure 9). Individuals classified as low risk had a 73% lower chance of disease progression than high-risk individuals. The model showed moderate accuracy, with a concordance index of 0.67 (standard error [SE], 0.088). Combining genomic data with clinical factors improved predictive performance beyond clinical data alone (C-statistic, 0.67 vs 0.52; P = .038). Moreover, in a multivariable Cox analysis accounting for existing clinical risk stratification (International Prognostic Scoring System for Waldenström Macroglobulinemia [IPSSWM] 2019), genomic classification remained a significant independent predictor of progression in both the initial and validation cohorts, underscoring its strong prognostic value (Table 3).

Discussion

The genetic composition of WM clones is likely to determine disease progression, as well as response to targeted treatments. Because we cannot accurately predict which patients with IgM-MGUS or aWM will progress to sWM, further understanding of disease progression using comprehensive molecular characterization will deconvolute part of the genomic landscape of this neoplasm.

In this study, we used WES to analyze CD19-selected samples from 139 patients, primarily with IgM-MGUS and aWM. To our knowledge, our findings are the first to demonstrate that evolving (from asymptomatic to symptomatic) patients have distinct genomic landscapes compared with clinically stable IgM-MGUS and aWM. Although the cohort of patients with aWM is

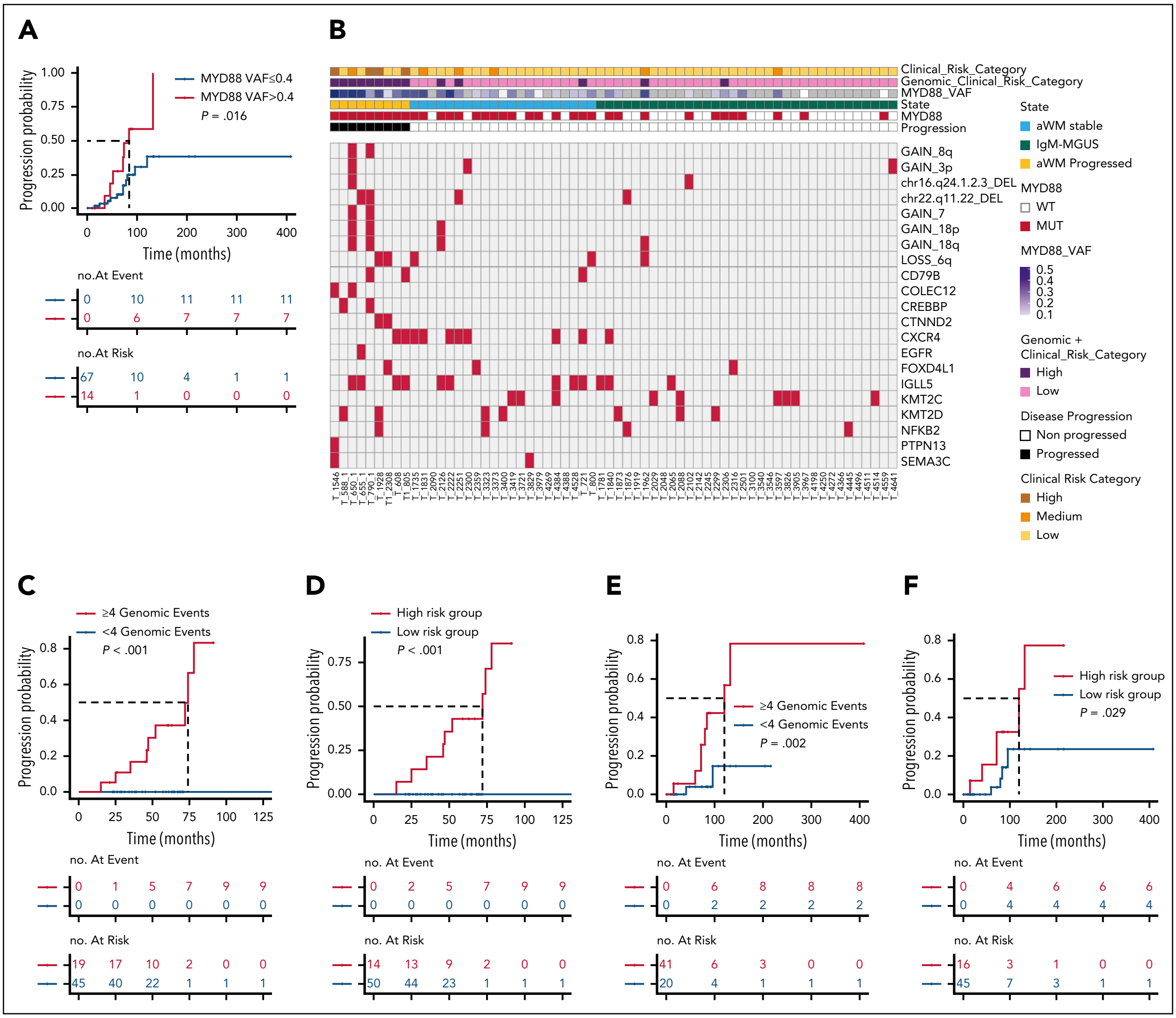

Figure 6. TTP probability in patients with precursor WM. (A) Kaplan-Meier curves analysis of TTP determined by *MYD88*^{L265P} VAF. (B) Heat map of selected genetic events in our initial cohort of 76 patients. Purple illustrates high-risk patients (>4 genetic events), whereas pink illustrates the patients at low risk. Dark purple illustrates high *MYD88*^{L265P} VAF, with a gradual color decrease represented by decreasing MYD88^{L265P} VAF. Bustoros et al risk score application determined by high- (brown), medium- (orange), and low-risk (yellow) patients. (C) Kaplan-Meier curves for the analysis of TTP determined by the presence of high genomic events in the initial cohort. (D) Kaplan-Meier curves analysis of TTP determined patients in high-risk and low-risk groups in the initial cohort. (E) Kaplan-Meier curves for the analysis of TTP determined by the presence of high genomic events in the validation cohort. (F) Kaplan-Meier curves analysis of TTP determined in high-risk and low-risk groups in the validation cohort. MUT, mutated; WT, wild-type.

relatively small (70 patients, including 51 stable and 19 who progressed), our data reveal significant differences in the distribution of genetic alterations between these subgroups. Notably, patients with progressing WM exhibit a higher mutation burden, suggesting that these genomic features may be associated with disease progression, even in cases that tend to be indolent and clinically stable. We clearly observed that the aWMpr group had a genetic makeup very similar to patients with sWM, including similar mutation burden of nonsynonymous SNVs and similar clonal architecture of the SNVs. Genes such as *CD79B*, *ARID1A*, *TTN*, and *CREBBP* were more often mutated in the aWMpr group compared with the aWMst group. Mutation counts in the *TTN* gene have been shown to serve as a reliable surrogate marker for tumor mutation burden in various malignancies, including lymphomas[35,36]; however, it is not considered as a biologically relevant gene in the context of WM or IgM plasma cell dyscrasia. SNVs in genes relevant to WM disease such as *CXCR4*, *EZH2*, *KMT2D*, *NFKB1*, *ARID1A*, *ARID1B*, and *NOTCH3* were mainly characterized as subclonal, suggesting that they were acquired later during disease progression. As expected, *MYD88*^{L265P} was clonal, indicating that it is probably a driver mutation acquired at early stages of the disease. Similarly, alterations in *IGLL5* gene exhibited clonal architecture indicating that may also be an early event in the WM natural history. The higher (although not significant) frequency of *CXCR4* mutations observed in patients with aWMst than in those with aWMpr appears to contradict current understanding of their functional role in promoting BM homing of malignant cells,[37,38] and being associated with shorter progression-free survival[39] and poorer responses to bruton tyrosine kinase (BTK) inhibitors.[40,41] However, this finding suggests that *CXCR4* alterations may not be critical drivers of disease progression, although larger patient cohorts are needed to validate this hypothesis and clarify

Table 3. Performance of the models on the initial cohort and on the validation cohort

Model	Likelihood ratio test statistic	Degrees of freedom	χ^2 *P* value	C-statistic (SE)
Initial cohort				
Clinical-IPSWM 2019	9.92	2	**.007019**	0.6447 (0.1766)
Genomic	15.35	1	**.00008956**	0.7211 (0.09708)
Clinical-IPSWM 2019 + genomic	18.02	3	**.000436**	0.6718 (0.1831)
RSFclin	13.17	1	**.0002841**	0.7669 (0.108)
RSFgenomic	20.76	1	**5.195e-06**	0.7677 (0.1025)
RSFclin + genomic	18.12	1	**2.074e-05**	0.7574 (0.09907)
Validation cohort				
Clinical-IPSWM 2019	0.9	2	.6	0.5972 (0.07201)
Genomic	8.67	1	**.003241**	0.7333 (0.0795)
Clinical-IPSWM 2019 + genomic	8.97	3	**.02974**	0.7621 (0.07338)
RSFclin	0.62	1	.4298	0.5265 (0.09111)
RSFgenomic	0.56	1	.4538	0.5441 (0.07566)
RSFclin + genomic	4.27	1	**.03881**	0.6726 (0.08811)

IPSWM, International Prognostic Scoring System for Waldenström Macroglobulinemia; RSFclin, random survival forest clinical; RSFgenomic, random survival forest genomic.
Bold indicates *P* value <.05.

the true biological impact of these mutations. We noted that the prevalence of CNAs increases in patients with aWMpr and sWM compared to aWMst and IgM-MGUS, with most CNAs showing a clonal architecture. This suggests that most CNAs are acquired early in disease development. The most common CNA, del6q, was identified in 3 of 9 patients with serial samples; in 1, del6q expanded clonally from the asymptomatic to symptomatic stage, in another it appeared as an acquired event during disease progression, and in the third, it was an early clonal alteration. These findings highlight the clonal heterogeneity characteristic of WM.

Our clonal evolution analysis revealed that nearly all patients with serial samples from both the asymptomatic and symptomatic stages showed the acquisition of genetic alterations over time, either as branching or linear events, along with changes in clonal composition between the disease stages. Although some new mutations arise during progression to symptomatic disease, most patients' tumor genetics are largely established during the asymptomatic phase, indicating that genetic biomarkers could potentially be identified at this early stage for better disease prediction and management.

The study also identified 2 genetically distinct subgroups among patients with IgM-MGUS: 1 with MYD88^{L265P} or other WM-related gene mutations, and another without these mutations, despite similar clinical features. This raises the possibility that the distinct biological profile of the second group may represent a more benign form of IgM-MGUS or a group that would potentially progress to another "lymphoma-like" or "myeloma-like" disease, highlighting the need for long-term follow-up, to evaluate whether our risk stratification model aligns with their eventual prognostic outcomes.

Given the aforementioned observations, we tried to implement a genomic scoring system to identify possible independent genomic predictors of progression that might underlie a more aggressive biology. We specifically analyzed patients with WM precursor, and found that a higher number of genomic events is strongly associated with a higher risk of disease progression. Furthermore, we validated our model using 2 groups (an internal and an external cohort), which doubled the number of patients from the initial cohort and showed that the integration of genomic events improved the prediction of progression when added to the existing clinical model. We acknowledge the limitations of the small number of progression events, but considering the nature and the rarity of this disease we believe these findings deliver a better understanding in the biology of WM. Furthermore, we have developed a targeted assay panel encompassing 78 SNVs and indels, along with selected CNAs identified in our study and in conjunction with lymphoma-related genes (supplemental Table 8) to screen all our newly diagnosed patients with IgM monoclonal gammopathies.

In conclusion, our study highlights the genetic evolution of WM precursors, and suggests that these findings could enhance predictive models for disease progression, emphasizing the importance of regular monitoring of the affected patients. This may enable better risk stratification, with high-risk patients requiring earlier intervention to enhance current management approaches.

Acknowledgment

The authors thank all the patients and their families for their contributions to this study.

F.M. was supported by a grant from the National Cancer Institute, National Institutes of Health (P30 CA 008748).

Authorship

Contribution: T.B. designed and performed research and wrote the manuscript; O.T., C.V., and F.M. analyzed data and wrote of the manuscript; K.M. and F.A. analyzed data; I.K., N.M.-K., C.L., K.T., A.P., K.C., and M.S. performed research; I.S. and F.T. analyzed clinical data; G.F. and M.V. performed research and analyzed clinical data; M.G. analyzed clinical data; E.T. wrote the manuscript; and Z.R.H., S.T., M.A.D., and E.K. wrote the manuscript.

Conflict-of-interest disclosure: E.K. receives honoraria and research support, through his institution, from Johnson and Johnson, GlaxoSmithKline, and Pfizer. The remaining authors declare no competing financial interests.

ORCID profiles: T.B., 0000-0001-8695-0493; O.T., 0009-0008-5229-4430; C.V., 0000-0002-9132-2878; I.K., 0000-0001-9853-2459; K.M., 0000-0001-7873-4854; F.A., 0000-0003-3209-4286; K.T., 0009-0007-0954-3908; A.P., 0009-0009-3659-0149; K.C., 0009-0002-4676-0016; M.S., 0000-0003-3616-3277; I.S., 0000-0001-9021-3824; F.T., 0000-0001-6926-0351; G.F., 0009-0005-5467-6019; M.G., 0000-0002-6244-1229; E.T., 0000-0001-5133-1422; M.V., 0000-0001-7304-1629; Z.R.H., 0000-0002-1689-1691; S.T., 0000-0001-6393-6154; F.M., 0000-0002-5017-1620; M.A.D., 0000-0001-8990-3254; E.K., 0000-0001-8191-5832.

Correspondence: Tina Bagratuni, National and Kapodistrian University of Athens, Alexandra Hospital, Vas Sofias 80, 11528 Athens, Greece; email: tbagratuni@med.uoa.gr.

Footnotes

Submitted 30 May 2025; accepted 30 July 2025; prepublished online on *Blood* First Edition 25 September 2025. https://doi.org/10.1182/blood.2025030177.

*T.B. and O.T. contributed equally to this study.

†M.A.D. and E.K. are joint senior authors.

Whole-exome sequencing data have been deposited at the European Genome-phenome Archive (EGA), which is hosted by the EBI and the CRG, under accession number EGAS50000001249.

Requests for additional information should be directed to the corresponding author, Tina Bagratuni (tbagratuni@med.uoa.gr).

The online version of this article contains a data supplement.

There is a *Blood* Commentary on this article in this issue.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

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.

2. Campo E, Swerdlow SH, Harris NL, Pileri S, Stein H, Jaffe ES. The 2008 WHO classification of lymphoid neoplasms and beyond: evolving concepts and practical applications. *Blood*. 2011;117(19):5019-5032.

3. Treon SP, Xu L, Yang G, et al. MYD88 L265P somatic mutation in Waldenström's macroglobulinemia. *N Engl J Med*. 2012;367(9):826-833.

4. Hunter ZR, Xu L, Yang G, et al. The genomic landscape of Waldenstrom 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.

5. Treon SP, Cao Y, Xu L, Yang G, Liu X, Hunter ZR. Somatic mutations in MYD88 and CXCR4 are determinants of clinical presentation and overall survival in Waldenstrom macroglobulinemia. *Blood*. 2014;123(18):2791-2796.

6. Forgeard N, Baron M, Caron J, et al. Inflammation in Waldenström macroglobulinemia is associated with 6q deletion and need for treatment initiation. *Haematologica*. 2022;107(11):2720-2724.

7. Hunter ZR, Xu L, Tsakmaklis N, et al. Insights into the genomic landscape of MYD88 wild-type Waldenström macroglobulinemia. *Blood Adv*. 2018;2(21):2937-2946.

8. 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.

9. Varettoni M, Zibellini S, Defrancesco I, et al. Pattern of somatic mutations in patients with Waldenström macroglobulinemia or IgM monoclonal gammopathy of undetermined significance. *Haematologica*. 2017;102(12):2077-2085.

10. Hunter ZR, Xu L, Yang G, et al. Transcriptome sequencing reveals a profile that corresponds to genomic variants in Waldenström macroglobulinemia. *Blood*. 2016;128(6):827-838.

11. Bagratuni T, Aktypi F, Theologi O, et al. Single-cell analysis of MYD88(L265P) and MYD88(WT) Waldenström macroglobulinemia patients. *Hemasphere*. 2024;8(2):e27.

12. Deguine J, Barton GM. MyD88: a central player in innate immune signaling. *F1000Prime Rep*. 2014;6:97.

13. Fitzgerald KA, Kagan JC. Toll-like receptors and the control of immunity. *Cell*. 2020;180(6):1044-1066.

14. Munshi M, Liu X, Kofides A, et al. A new role for the SRC family kinase HCK as a driver of SYK activation in MYD88 mutated lymphomas. *Blood Adv*. 2022;6(11):3332-3338.

15. Munshi M, Liu X, Chen JG, et al. SYK is activated by mutated MYD88 and drives pro-survival signaling in MYD88 driven B-cell lymphomas. *Blood Cancer J*. 2020;10(1):12.

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.

17. 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.

18. Kyle RA, Benson JT, Larson DR, et al. Progression in smoldering Waldenstrom macroglobulinemia: long-term results. *Blood*. 2012;119(19):4462-4466.

19. Kyle RA, Therneau TM, Rajkumar SV, et al. Long-term follow-up of IgM monoclonal gammopathy of undetermined significance. *Semin Oncol*. 2003;30(2):169-171.

20. Kyle RA, Rajkumar SV. Monoclonal gammopathy of undetermined significance and smoldering multiple myeloma. *Curr Hematol Malig Rep*. 2010;5(2):62-69.

21. 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.

22. Treon SP, Hunter ZR, Aggarwal A, et al. Characterization of familial Waldenstrom's macroglobulinemia. *Ann Oncol*. 2006;17(3):488-494.

23. Poulain S, Braggio E, Roumier C, et al. High-throughput genomic analysis in Waldenström's macroglobulinemia. *Clin Lymphoma Myeloma Leuk*. 2011;11(1):106-108.

24. Braggio E, Philipsborn C, Novak A, Hodge L, Ansell S, Fonseca R. Molecular pathogenesis of Waldenstrom's macroglobulinemia. *Haematologica*. 2012;97(9):1281-1290.

25. Paiva B, Corchete LA, Vidriales MB, et al. The cellular origin and malignant transformation of Waldenström macroglobulinemia. *Blood*. 2015;125(15):2370-2380.

26. San Miguel JF, Vidriales MB, Ocio E, et al. Immunophenotypic analysis of Waldenstrom's macroglobulinemia. *Semin Oncol*. 2003;30(2):187-195.

27. Morice WG, Chen D, Kurtin PJ, Hanson CA, McPhail ED. Novel immunophenotypic

features of marrow lymphoplasmacytic lymphoma and correlation with Waldenström's macroglobulinemia. *Mod Pathol*. 2009;22(6):807-816.

28. Remstein ED, Hanson CA, Kyle RA, Hodnefield JM, Kurtin PJ. Despite apparent morphologic and immunophenotypic heterogeneity, Waldenstrom's macroglobulinemia is consistently composed of cells along a morphologic continuum of small lymphocytes, plasmacytoid lymphocytes, and plasma cells. *Semin Oncol*. 2003;30(2):182-186.

29. Ishwaran H, Kogalur U, Blackstone E, Lauer M. Random Survival Forests. *Ann Appl Stat*. 2008;2(3):841-860.

30. Martincorena I, Raine KM, Gerstung M, et al. Universal patterns of selection in cancer and somatic tissues. *Cell*. 2017;171(5):1029-1041.e21.

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.

32. Xu L, Hunter ZR, Yang G, et al. MYD88 L265P in Waldenström macroglobulinemia, immunoglobulin M monoclonal gammopathy, and other B-cell lymphoproliferative disorders using conventional and quantitative allele-specific polymerase chain reaction. *Blood*. 2013;121(11):2051-2058.

33. Bagratuni T, Markou A, Patseas D, et al. Determination of MYD88L265P mutation fraction in IgM monoclonal gammopathies. *Blood Adv*. 2022;6(1):189-199.

34. Bustoros M, Sklavenitis-Pistofidis R, Kapoor P, et al. Progression risk stratification of asymptomatic Waldenström macroglobulinemia. *J Clin Oncol*. 2019;37(16):1403-1411.

35. Li Y, Liu L, Sun H, et al. Complete remission of Hodgkin's lymphoma in a pediatric patient with TTN gene mutation treated with brentuximab vedotin combined chemotherapy without anthracyclines: a case report. *Front Oncol*. 2022;12:1006166.

36. Oh JH, Jang SJ, Kim J, et al. Spontaneous mutations in the single TTN gene represent high tumor mutation burden. *NPJ Genom Med*. 2020;5:33.

37. Alouche N, Bonaud A, Rondeau V, et al. Hematologic disorder-associated Cxcr4 gain-of-function mutation leads to uncontrolled extrafollicular immune response. *Blood*. 2021;137(22):3050-3063.

38. Poulain S, Roumier C, Venet-Caillault A, et al. Genomic landscape of CXCR4 mutations in Waldenström macroglobulinemia. *Clin Cancer Res*. 2016;22(6):1480-1488.

39. Gustine JN, Xu L, Yang G, et al. Bone marrow involvement and subclonal diversity impairs detection of mutated CXCR4 by diagnostic next-generation sequencing in Waldenström macroglobulinaemia. *Br J Haematol*. 2021;194(4):730-733.

40. Cao Y, Hunter ZR, Liu X, et al. CXCR4 WHIM-like frameshift and nonsense mutations promote ibrutinib resistance but do not supplant MYD88(L265P)-directed survival signalling in Waldenström macroglobulinaemia cells. *Br J Haematol*. 2015;168(5):701-707.

41. Cao Y, Hunter ZR, Liu X, et al. The WHIM-like CXCR4(S338X) somatic mutation activates AKT and ERK, and promotes resistance to ibrutinib and other agents used in the treatment of Waldenstrom's macroglobulinemia. *Leukemia*. 2015;29(1):169-176.