

blood

Prepublished online December 23, 2013;
doi:10.1182/blood-2013-09-525808

The genomic landscape of Waldenstöm's Macroglobulinemia is characterized by highly recurring MYD88 and WHIM-like CXCR4 mutations, and small somatic deletions associated with B-cell lymphomagenesis

Zachary Hunter, Lian Xu, Guang Yang, Yangsheng Zhou, Xia Liu, Yang Cao, Robert J. Manning, Christina Tripsas, Christopher J. Patterson, Patricia Sheehy and Steven P. Treon

Information about reproducing this article in parts or in its entirety may be found online at:
http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at:
<http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#reprints>

Information about subscriptions and ASH membership may be found online at:
<http://bloodjournal.hematologylibrary.org/site/subscriptions/index.xhtml>

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036.

Copyright 2011 by The American Society of Hematology; all rights reserved.



The genomic landscape of Waldenström's Macroglobulinemia is characterized by highly recurring MYD88 and WHIM-like CXCR4 mutations, and small somatic deletions associated with B-cell lymphomagenesis.

Zachary Hunter^{1,2}, Lian Xu¹, Guang Yang¹, Yangsheng Zhou¹, Xia Liu¹, Yang Cao¹, Robert J. Manning¹, Christina Tripsas¹, Christopher J. Patterson¹, Patricia Sheehy¹, and Steven P. Treon^{1,3}

¹Bing Center for Waldenström's Macroglobulinemia, Dana-Farber Cancer Institute, Boston, MA 02115, USA; ²Department of Pathology and Laboratory Medicine, Boston University School of Graduate Medical Sciences, Boston, MA 02118, USA; ³Harvard Medical School, Boston, MA 02115, USA.

Corresponding author:

Steven P. Treon M.D., M.A., Ph.D.

Bing Center for Waldenström's Macroglobulinemia

Dana Farber Cancer Institute

M547, 450 Brookline Avenue, Boston MA 02215 USA

Tel: (617) 632-2681 Fax: (617) 632-4862

Email: steven_treon@dfci.harvard.edu

Short Title: Genomic Landscape of Waldenström's Macroglobulinemia.

Keywords: Waldenström's Macroglobulinemia, Lymphoplasmacytic Lymphoma, WHIM Syndrome, MYD88, CXCR4, ARID1A, BTG1, HIVEP2, LYN, PRDM2, ARID1B, PLEKHG1, MKLN1.

KEY POINTS

- Highly recurring mutations are present in WM, including MYD88 L265P, WHIM-syndrome like mutations in CXCR4, and ARID1A.
- Small previously undetected copy number alterations affecting B-cell regulatory genes are highly prevalent in WM.

ABSTRACT

The genetic basis for Waldenström's Macroglobulinemia (WM) remains to be clarified. 6q losses are commonly present, though recurring gene losses in this region remain to be defined. We therefore performed whole genome sequencing (WGS) in 30 WM patients, which included germline/tumor sequencing for 10 patients. Validated somatic mutations occurring in >10% of patients included *MYD88*, *CXCR4* and *ARID1A* that were present in 90%, 27%, and 17% of patients, respectively, and included the activating mutation L265P in MYD88, and WHIM-syndrome like mutations in CXCR4 which previously have only been described in the germline. WGS also delineated copy number alterations (CNA) and structural variants in the 10 paired patients. The CXCR4 and CNA findings were validated in independent expansion cohorts of 147 and 30 WM patients respectively. Validated gene losses due to CNAs involved *PRDM2* (93%), *BTG1* (87%), *HIVEP2* (77%), *MKLN1* (77%), *PLEKHG1* (70%), *LYN* (60%), *ARID1B* (50%), and *FOXP1* (37%). Losses in *PLEKHG1*, *HIVEP2*, *ARID1B*, and *BCLAF1* constituted the most common deletions within chromosome 6. While no recurrent translocations were observed, in 2 patients deletions in 6q corresponded with translocation events. These studies evidence highly recurring somatic events, and provide a genomic basis for understanding the pathogenesis of WM.

INTRODUCTION

Waldenström's Macroglobulinemia (WM) is a distinct indolent B-cell malignancy in the World Health Organization classification. WM shares many clinical and pathological features with other B-cell lymphomas and multiple myeloma (MM), which often complicates the diagnosis of this entity. IgM Monoclonal Gammopathy of Unknown Significance (MGUS) is a precursor state for WM. Approximately 2% of IgM MGUS patients evolve to a B-cell malignancy per year, with most of these individuals progressing to WM.¹⁻³ Since the initial description of Waldenström's Macroglobulinemia by Jan Waldenström in 1944, the genetic basis for WM has been elusive. We therefore employed whole genome sequencing (WGS) utilizing tumor derived DNA from 30 patients, which included paired germline/tumor sequencing for 10 patients. The initial findings of this study resulted in the identification of a recurrent somatic mutation, L265P, in *MYD88* in over 90% of WM patients that distinguished WM from other overlapping entities such as marginal zone lymphoma (MZL), chronic lymphocytic leukemia (CLL), and multiple myeloma wherein *MYD88* L265P was either absent or infrequently observed (<10%).^{4,5,6} Importantly, *MYD88* L265P was present in IgM MGUS with a frequency of 50-80% using allele specific (AS) polymerase chain reaction (PCR) suggesting an early oncogenic event for this mutation, and that other genomic events are likely to be responsible for WM disease progression. Activating *MYD88* mutations are not unique to WM and appear in other B-cell lymphomas including primary central nervous system lymphoma (50%), and activated B-cell (ABC) subtype diffuse large B-cell lymphoma (DLBCL) (8%-29%).^{7,8,9} A multi-hit model would also help explain why no significant differences in treatment response, progression free and overall survival based solely on *MYD88* mutation status determination were observed in a large series of WM patients.¹⁰

Little is known about genetic structural variant (SV) changes such as copy number alterations (CNA) or translocations in WM. By fluorescence in situ

hybridization (FISH) and array comparative genomic hybridization (aCGH) deletions in chromosome 6q21-23 have been identified in 40%-60% of WM patients, with concordant gains in 6p in 41% of those with 6q deletions.¹¹⁻¹³ Gains in chromosomes 3q, 4, 18, 8q and Xq as well as losses of 11q23, 13q14, and 17p have also been described in up to 20% of WM cases.^{14,15} We therefore sought to more fully characterize the molecular events responsible for WM pathogenesis using WGS.

METHODS

Sample Collection and Preparation

Bone marrow aspirates and peripheral blood (PB) samples were collected in heparinized syringes from thirty patients with the clinicopathological diagnosis of WM as defined by the Second International Workshop on WM.¹⁶ Participants provided informed consent prior to sample collection in accordance with the Dana-Farber Cancer Institute Institutional Review Board. This study was conducted in accordance with the Declaration of Helsinki. Relevant clinical characteristics and blood work are presented in Supplemental Table 1. Bone marrow and PB mononuclear cells were isolated by density gradient centrifugation using Ficoll-Paque (Amersham-Pharmacia Biotech, Piscataway, NJ). WM lymphoplasmacytic cells (LPC) were isolated from the BM mononuclear cells by CD19⁺ selection using immunomagnetic MACS micro-beads (Miltenyi-Biotech, Auburn, CA). To minimize potential contamination from circulating WM LPC, PB MCs were depleted of CD19⁺ cells using CD19⁺ MACS micro-beads.

Complete methodology for the WGS analysis and validation is available in the supplementary methods.

RESULTS

Thirty patients meeting the International Consensus Criteria for WM diagnosis were evaluated in this study.¹⁶ The clinical characteristics were typical of WM of this cohort are presented in Supplemental Table 1. Tumor DNA from CD19⁺ BM lymphoplasmacytic cells (LPL) cells was isolated from 30 patients, along with DNA from CD19-depleted peripheral blood mononuclear cells (PBMC) for use in germline analysis. All 30 tumor and 10 PBMC samples were submitted for WGS with Complete Genomics as previously described.⁴

Coding variant analysis

WGS findings were filtered for novel non-synonymous variants as described in the Supplementary Methods. This identified 5,835 variants in 4,427 genes. The 10-paired genomes were used to identify genes containing at least one somatic variant. Variant data from all 30 WM tumor genomes were then fit to this model wherein each affected gene was categorized as potentially somatic, germline, or unclassified if variants were only observed within the 20 unpaired samples (Supplementary Table S2). Genes were further categorized by mutational frequency and validated based on bioinformatic strength and biological relevance, using Sanger sequencing of both tumor and germline DNA (Table 1). These cut off scores were significantly lower than our previous analysis allowing for the detection of a number of additional mutations. Primers and representative chromatograms are provided in Supplementary Table S3. A full list of somatic, and germline mutations affecting genes in the COSMIC cancer gene census are shown in Supplementary Tables S4 and S5, respectively.¹⁷

As previously described, a somatic T/C mutation in *MYD88* resulting in an L265P substitution was observed in 27/30 samples. In addition to the L/P activating substitution in the primary transcript, this mutation results in a stop loss in two of

the alternative isoforms. The potential functional significance for these shorter transcripts in MYD88 L265P is not known. A subclonal C/G mutation was found in MYD88 resulting in an amino acid substitution (S219C) in one patient who also had the L265P mutation.⁸ No other mutations in *MYD88* were observed in the study population.

To better understand the *MYD88* wild type (WT) population, we identified somatic mutations found exclusively in the 3 *MYD88* WT patients. Cross listing these genes with COSMIC, we observed that 2 of these 3 (67%) patients had damaging mutations in *MLL2* while mutations in the genes *JAK2*, *PRCC*, *SUZ12*, *BCL10*, *KDM6A*, and *SETD2* were each observed once. Sanger sequencing confirmed the somatic presence of *MLL2* mutations in these patients as previously described.⁴ While a larger sample size is needed to determine if the high frequency of *MLL2* mutations is typical of the *MYD88* WT WM population, it nonetheless represents a gene of great interest since *MLL2* mutations are highly recurrent in follicular lymphoma and DLBCL patients.¹⁸

The next most common somatic mutation was observed in the G-protein coupled receptor (GPCR) *CXCR4*. *CXCR4* is a chemokine receptor that promotes WM survival, migration, and adhesion to the bone marrow (BM) stroma through interactions with its ligand CXCL12.¹⁹ Eight of 30 (27%) patients harbored one of 5 somatic variants in *CXCR4*, each of which were identical or functionally similar to mutations associated with *Warts*, *Hypogammaglobulinemia*, *Infection*, and *Myelokathexis* (WHIM) syndrome (Figure 1).^{20,21} WHIM syndrome is a rare, autosomal dominant genetic disorder that is caused by frame shift or nonsense mutations in the carboxyl-terminal cytoplasmic tail (c-tail) of *CXCR4*. These mutations occur after the last of the seven transmembrane helices, destroying the c-tail but leaving the regions responsible for ligand binding and down stream signaling via g-proteins intact.²² These c-tail mutations result in the loss of regulatory serines resulting in impaired internalization and prolonged activation.^{23–25} To validate this finding, we screened an independent cohort of

147 WM patients for CXCR4 mutations by Sanger sequencing. Combined with the 30 patients in the WGS cohort whose results were validated by Sanger sequencing, 51/177 (28.8%) patients had c-terminal mutations in CXCR4. Among these 177 patients, 160 demonstrated the MYD88 L265P mutation. Importantly, 50/51 (98%) of the patients with CXCR4 c-terminal mutations were also MYD88 L265P positive. Conversely, 16/17 (94.1%) of the MYD88 wild-type patients were wild-type for CXCR4 ($p=0.026$).

Other genes with validated somatic mutations in this series included *ARID1A* (5/30; 17%), *CD79B* (2/30; 7%), *TP53* (2/30; 7%), *MYBBP1A* (2/30; 7%), *MUC16* (1/30; 3%), *TRAF2* (1/30; 3%), *TRAF3* (1/30; 3%), *RAG2* (1/30; 3%), and *NOTCH2* (1/30; 3%) of patient samples (Figure 2).

Structural Variant Analysis

Structural variants were inferred from high confidence discordant mate-pair and relative coverage data supplied by Complete Genomics. To gain further insight into the SV events underling WM, we combined individual patient data for lesser allele fraction (LAF), heterozygous call rate, normalized relative coverage, somatic CNAs, and somatic SV calculations with cohort wide small variant distribution and CNA regions of statistical significance (Supplementary Figure S1). These analyses aid in the detection of complex events such as acquired uniparental disomies (aUPD), which can be observed in chromosomes 3p and 21 in Supplementary Figure S1. The most common aUPD was at 3p, occurring in 4/30 (13%) of patients and whose location and size were highly conserved between the samples (Figure 3). This locus included the *MYD88* L265P locus, resulting in homozygous somatic mutation in these patients. Additional large aUPDs were observed on chromosomes 1, 2, 5, 9, 17, 21, and X with many of these affecting the validated small variant findings (Table 2).

Translocations were rare events, and we were unable to demonstrate any recurrent translocations in the 10-paired samples (Supplementary Table S5). One patient had two distinct t(2;17) translocations disrupting both alleles of *RNF213*. A t(6;X) translocation disrupting the gene *BIA3* was responsible for the chromosome 6q deletion observed in another patient (Supplementary Figure S1). A third patient had a chromothriptic event centered on chromosome 6 which made up over half of all translocations observed in this study (Supplementary Figure S2), and contributed to 6q deletions observed in this patient's tumor cells. In this patient, we performed PCR for t(6;7) predicted to disrupt *BMP5* and *ANKRD7*, as well as t(6;11) predicted to disrupt *GRIA4* and *PKHD1* respectively, and validated both of these findings. For 2 of 5 patients with paired normal/tumor genomes and in whom 6q deletions were identified, translocations were responsible for these regional losses. Macro-level analysis of normalized coverage for all 30 WM samples demonstrated that large deletions in 6q were present in 13/30 (43%) with corresponding gains in 6p in 3/13 (23%) cases. The next most common alteration observed was an amplification of chromosome 4 which occurred in 7/30 (23%) patients confirming previous FISH and aCGH findings.^{14,15}

Small Somatic Copy Number Alterations

To detect small somatic deletions and amplifications, we calculated 100Kb blocks of the log-transformed ratio of tumor to germline relative coverage in each of the paired genomes. Significant CNAs were then compared across all 10 patients in order to identify statistically significant regions. CNAs consisted mostly of isolated 100Kb blocks both within and outside of chromosome 6 (Supplementary Table S6 and Supplementary Figure S1E). To investigate the biological relevance of these non-chromosome 6 CNAs, we performed functional enrichment analysis on the gene set affected by these alterations. Unique functional annotation results are listed in order of significance in Figure 4A.

To better characterize these highly recurrent deletions, consecutive blocks of equal significance were merged together to form 194 statistically significant regions with a median length 100Kb (range 100Kb–4,900Kb). Of these regions, 134 (69%) were found outside of chromosome 6 with a median length of 100Kb (range 100Kb–300Kb), of which 126 of 134 (94%) were restricted to a single 100Kb region. Within chromosome 6, CNAs were larger with only 30/60 (50%) CNAs constituting single 100Kb regions ($p=7.709 \times 10^{-12}$). CNAs targeted COSMIC census genes for 14 of 172 (8.14%) genes outside of chromosome 6, versus 4 of 239 (1.67%) genes within chromosome 6 respectively ($p=0.0024$). To determine the probability of generating these results as a random by-product of genomic instability, we calculated the number of total genes affected by these deletions ($N=411$) and the number of these genes found in the COSMIC database ($N=18$). We randomly distributed matching deletion sets across the genome and enumerated the total number of overlapping genes and the number of those genes listed in the COSMIC census (Figure 4B). These simulations revealed a statistically significant increase in targeting of COSMIC genes by CNAs outside of chromosome 6. Affected genes in the COSMIC census were *BTG1* (9/10; 90%), *FOXP1* (7/10; 70%), *FNBP1* (7/10; 70%), *CD74* (7/10; 70%), *TOP1* (6/10; 60%), *MYB* (5/10; 50%), *CBLB* (5/10; 50%), *ETV6* (5/10; 50%), *TNFAIP3* (5/10; 50%), *FBXW7* (5/10; 50%), *PRDM1* (5/10; 50%), *TFE3* (4/10; 40%), *JAK1* (4/10; 40%), *MAML2* (4/10; 40%), *FAM46C* (4/10; 40%), *EBF1* (4/10; 40%), *STL* (4/10; 40%), and *BIRC3* (4/10; 40%). Other affected genes of interest included *PRDM2* (8/10; 80%), *HIVEP2* (8/10; 80%), *ARID1B* (7/10; 70%) as well as *LYN* (7/10; 70%).

There were no singular regions of statistical significance in 6q. Some patients had multifocal deletions in this region suggesting multiple target genes (Figure 5). Neither of the previously suspected target genes for 6q loss, i.e. *PRDM1* and *TNFAIP3* were included in the regions of highest statistical significance.^{11,12} The relative coverage data from all 10-paired patients is shown in Figure 6b

illustrating that only 3 patients had highly clonal deletions in 6q. Two additional patients had subclonal deletions in 6q, therefore 5/10 (50%) paired patients had at least subclonal loss. Deletions in *HIVEP2* (8/10; 80%) as well as *ARID1B* (7/10; 70%) and *BCLAF1* (7/10; 70%) constituted the most common deletions in chromosome 6, and were present in patients with and without visible 6q losses.

To test our validation system and examine the accuracy of the GC normalized relative coverage data, we used a PCR based copy number assay to validate the established 6q22 deletion in a subset of patients (Figure 6A). Having recapitulated the expected results, we conducted additional validation studies for eight of our top targets in the same five patients using commercially available assays. While many results appeared subclonal, we observed a strong correlation between the PCR relative copy number and WGS coverage predictions (Figure 6D). To establish the frequency of these CNAs, we validated these eight findings in an independent cohort of 30 WM patients revealing somatic losses in *PDRM2* (28/30; 93%) at 1p36.21, *BTG1* (26/30; 87%) in Chr. 12q21.33, *HIVEP2* (23/30; 77%) at 6q24.2, *MKLN1* (23/30; 77%) at 7q32, *PLEKHG1* (21/30; 70%) at 6q25.1, *LYN* (18/30; 60%) at 8q12.1, *ARID1B* (15/30; 50%) at 6q25.1, and *FOXP1* (11/30; 37%) at 3p13, and (Figure 2). For *MKLN1* and *HIVEP2*, germline CNAs were significant, and therefore tumor to germline relative coverage should be interpreted accordingly (Figure 6C).

CXCR4 and MYD88 L265P mutation status was determined for in bone marrow samples for all 30 patients by Sanger sequencing and AS-PCR respectively. There were fewer median validated deletions in CXCR4 mutated patients (5) compared to WT (7; $p=0.002$) and with the median total number of 6q CNAs (2) compared to WT (3; $p=0.007$; Figure 2). This was also true for any combination of two of the three validated 6q genes ruling out possible biasing by a single gene ($p<0.023$ for all).

DISCUSSION

The most pronounced finding from this work was the discovery of a somatic mutation in MYD88 (L265P), which was present in 90% of patients with WM. The details for this discovery were previously published following the preliminary examination of our whole genome sequencing results.⁴ MYD88 serves as an adaptor molecule in Toll-like receptor (TLR) and interleukin-1 receptor (IL-1R) signaling mediating IRAK4 and IRAK1 activation and down stream signaling through NFkB.^{4,8,26} Mutations in MYD88 are significant given the importance of NFkB signaling in WM cell growth and survival.²⁷ We recently demonstrated that MYD88 L265P can trigger NFkB signaling through an IRAK-independent pathway by direct interaction with Bruton's Tyrosine Kinase (BTK) in WM cells, suggesting parallel pathways for NFkB activation.²⁸ In contrast to the findings by Ngo et al who observed multiple MYD88 mutations (L265P, V217F, S219C, M232T, S243N, T294P) in ABC subtyped DLBCL patients, our findings were limited to the identification of L265P, and S219C as a subclonal event in one WM patient with a L265P mutation.⁸ We also did not observe mutations in CARD11 in WM patients, demonstrating differences in potential oncogenic drivers and dependence on TLR signaling for WM versus ABC subtyped DLBCL.

The next most common somatic variant after MYD88 L265P was in *CXCR4*, present in 27% of the patients. The mutations identified in WM tumor cells recapitulated those found in the germline of patients with WHIM syndrome. To our knowledge, this is the first time WHIM like mutations have ever appeared as somatic mutations, and also associated with a malignant disease. In WHIM syndrome, loss of regulatory serines in the c-tail of CXCR4 are known to impair receptor internalization, thereby prolonging G-protein and β -arrestin signaling.²⁵ CXCR4 stimulation by its ligand CXCL12 is known to activate AKT1 and MAPK family signaling, as well as facilitate cell migration and homing in WM cells.¹⁹ The prolonged activation of CXCR4 signaling due to WHIM mutations may therefore exaggerate these effects in WM cells, and deserves further study. Nearly all

patients with CXCR4 mutations also had carried the MYD88 L265P mutation, and comprehensive studies will be required to delineate the relative impact of these mutations on WM clinical and treatment response characteristics. The presence of CXCR4 mutations may also offer a targeted approach to therapy of WM by use of CXCR4 antagonists given their successful employment in the treatment of WHIM syndrome patients.²⁹ Several antagonists to CXCR4 have been developed and are in clinical trials including plerixafor, BMS-936564, AMD-070, TG-0054 and others and warrant investigation alone and/or in combination in WM patients.

While most WM cases in our series did not have mutations in CXCR4, it is possible that CXCR4 signaling may still be critical for many of these patients and offer an opportunity for targeted therapy with CXCR4 antagonists.²¹ In WM, polymorphisms of the CXCR4 ligand, CXCL12, have been associated with poor post treatment clinical outcomes.³⁰ Copy loss of *RGS17* that was observed in 5 of 10 (50%) of patients in our series may also affect this pathway. CXCR4 signals downstream through $G_{i\alpha}$ G-proteins increasing phosphorylated AKT1 levels and promoting cell survival while *RGS17* inhibits $G_{i\alpha}$ G-proteins by promoting GTP hydrolysis.^{31,32} In an analogous case, *RGS17* inactivation was observed in ovarian cancer affecting $G_{i\alpha}$ GPCR downstream signaling in response to lysophosphatidic acid.³¹

The other major pathway identified in this study was the loss of the chromatin remodeling proteins, *ARID1A* and *ARID1B*.³³ *ARID1A* was the third most common SNV variant target in WM, with 5/30 (17%) of patients having validated nonsense or frame shift mutations. In one patient, a mutation in *ARID1A* (Y551 frameshift) was homozygous as a result of an aUPD, whereas in another patient a nonsense mutation (Q2037*) was opposite CNA loss resulting in biallelic inactivation (Table 2). Loss of the alternate family member *ARID1B* was present in 7/10 (70%) of the paired patient samples making it a more frequent 6q target than either *PRDM1* or *TNFAIP3* (5/10; 50% for both).³³ One potential target for

this pathway is the B-cell development regulator *EBF1* which itself was affected by CNV deletions in 4/10 (40%) patients.^{34,35} Both *ARID1A* and *ARID1B* are members of the SWI/SNF family of proteins which are known to be mutated in other neoplastic malignancies wherein they are thought to exert their effects via p53 and *CDKN1A* regulation.^{36,37,38} *TP53* itself was mutated in 2/30 (7%) of the sequenced genomes including one case of biallelic mutation. *PRDM2* and *TOP1* both participate in *TP53* related signaling and were deleted in 8/10 (80%; 28/30 in validation) and 6/10 (60%) of patients respectively.^{39,40} Taken together, these findings would imply that multiple mutations exist in the genome of WM patients which dysregulate the DNA damage response.

Loss of additional cancer associated genes was observed including *ETV6*, which was deleted in 6/10 (60%) of the samples analyzed. Deletions of *ETV6* have been observed in acute lymphoblastic leukemia (ALL), acute myeloid leukemia, and myelodysplastic syndromes.⁴¹⁻⁴³ *ETV6* is a member of the ETS transcription factor family and acts primarily as a transcriptional repressor involved in hematopoiesis and functions as a tumor suppressor in myeloid leukemias.⁴⁴ *FOXO3*, which was deleted in 6/10 (60%) of patients, has been shown to negatively regulate growth and survival in mantle cell lymphoma, B-CLL, and natural killer cell neoplasms.⁴⁵⁻⁴⁷ In B-CLL, phosphorylation of *FOXO3* by *AKT1* down stream of *CXCL12* and *CXCR4* prevents *FOXO3* nuclear translocation and signaling.⁴⁶ Upon nuclear translocation, *FOXO3* can induce transcription of *BTG1* which was deleted in 9/10 (90%; 26/30 in validation) paired patients.⁴⁸ While *FOXO3* (6q21) was not explicitly validated, 6q deletions were significantly less common in *CXCR4* mutated patients and there was a trend for fewer *BTG1* deletions as well. *BTG1* is a nuclear co-activator that modulates transcription, a member of the anti-proliferative TOB/BTG protein family, and was recently demonstrated to be recurrently mutated in several studies of DLBCL.^{18,49,50} Small deletions affecting *BTG1* have been reported in up to 12% of the B-cell precursor subtype of ALL (BCP-ALL).^{43,51,52} *BTG1* deletions in BCP-ALL were noted to coincide with deletions in *ETV6* and *EBF1* both of which were recurrently deleted

in our study and suggests that these events may be functionally related.⁵² Preclinical studies of BTG1 in ALL have shown that loss of BTG1 is associated with glucocorticoid resistance. These findings may explain the poor responses observed to single agent steroids in WM, and warrant further investigation of BTG1 loss in treatment outcomes in patients undergoing glucocorticoid inclusive therapy.^{53,54} Finally, biallelic loss of *RNF213* resulting from two distinct *t*(2;17)s was observed in 1 of 10 paired patients, and represents an interesting finding since *RNF213* is a fusion partner of *ALK* and *MYC* in acute anaplastic large cell lymphoma. In the WM patient with the *t*(2;17) translocation, *RNF213* was translocated to the intergenic space indicating that the loss of *RNF213* itself may play an important role in oncogenesis.^{55,56} It is important to note that most CNAs were consistent with heterozygous loss and classical tumor suppressors require biallelic loss for oncogenesis. However, these losses could lead to pathway modulation due to altered protein levels or creation of dominant negatives by partial gene loss.

Many of the mutations identified in this study impact NF κ B signaling distal to the TLR4/MYD88 pathway. *MYBBP1A*, mutated in 2/30 (7%) patients, is thought to inhibit NF κ B activity by repressing *RELA*.⁵⁷ Copy loss of *HIVEP2* (8/10; 80%; 23/30 in validation) and *TNFAIP3* in 5/10 (50%) of the paired patients is of interest since their loss results in the removal of NF κ B negative regulators in WM.^{12,58} Chromosome 6q is often deleted in WM patients, and the apparent loss of *HIVEP2* in 6q intact patients is particularly compelling for a role of this gene in the pathogenesis of WM. Likewise, *BIRC3* loss in 4/10 (40%) of patients is associated with splenic MZL and can increase activation of non-canonical NF κ B signaling.⁵⁹ Interestingly, both of *BIRC3*'s partner proteins, *TRAF2* and *TRAF3*, that regulate non-canonical NF κ B signaling were each found to be mutated in 1/30 (3%) patients. Biallelic loss of *TRAF3* has been reported by Braggio et al¹² in WM patients. These findings may therefore provide an impetus for studying non-canonical NF κ B signaling in WM.

The loss in 7/10 (70%; 18/30 in validation) paired patients of *LYN*, a kinase that plays a regulatory role for B-cell receptor (BCR) signaling along with mutations in *CD79B* (2/30; 7%) indicates a possible role for BCR signaling in this disease.⁶⁰⁻⁶² The loss of *IBTK* in 4/10 (40%) of patients, and the interaction between activated MYD88 and BTK, raises the possibility for BTK mediated cross-talk with the TLR/MYD88 pathway wherein BTK plays an important role in NFκB activation.^{28,63-65} Moreover, *CBLB*, a gene disrupted in chronic myelogenous leukemia and lost in 5/10 (50%) of the paired patients has been shown to inhibit the TLR4 response during inflammation by controlling TLR4 and MYD88 association, and subsequent NFκB activation.^{66,67} However, there is still limited data regarding the structure of the MYD88 L265P mutant protein. Signaling pathways that are well documented in the wild type setting will need to be carefully re-examined to clarify the contribution of the L265P mutation to downstream MYD88 signaling.

Our WGS studies in WM patients have therefore identified mutations in genes involving TLR, NFκB, and CXCR4 signaling, chromatin remodeling, and cell cycle regulation. These findings may denote a multistep process for WM evolution from IgM MGUS to asymptomatic and symptomatic WM, which invariably will require comparative, and even prospective longitudinal sequencing studies. These efforts are particularly warranted since the principal mutation (MYD88 L265P) identified in these studies is also present in 50-80% of IgM MGUS patients using AS-PCR, signifying its role as an early oncogenic event.^{10,68,69} Other mutations therefore are likely to be acquired in the evolution of MGUS to symptomatic WM, and may offer opportunity to identify those patients at high risk for disease evolution. Lastly, these studies have also identified novel targets for rationale approach to WM treatment, including potentially the use of inhibitors targeting TLR and CXCR4 signaling.

ACKNOWLEDGEMENTS

This study would not have been possible without the generous support of the Peter S. Bing, the International Waldenström's Macroglobulinemia Foundation, the Coyote Fund for WM, the D'Amato Family Fund for Genomic Discovery, the Edward and Linda Nelson Fund for WM Research, and the WM patients who provided their samples. The authors would like to thank Yaoyu Wang and John Quackenbush at the Center for Cancer Computational Biology at the Dana-Farber Cancer Institute for their assistance in developing the copy number analysis. Portions of this study including the identification of CXCR4 WHIM-like mutations in WM were first presented at the 48th Annual Meeting of the American Society of Clinical Oncology held June 1-5, 2012 in Chicago IL.

AUTHOR CONTRIBUTIONS

ZRH and SPT designed the study wrote the manuscript. ZRH performed the data analysis and conducted the copy number validation. ZRH and XL designed Sanger sequencing primers. GY, YZ, YC, XL, and LX prepared the samples and performed the Sanger validation studies. SPT and PS provided patient care, obtained consent and samples. RM, CT, and CP selected samples and provided clinical data analysis.

AUTHOR DISCLOSURES

The authors have no competing interests to declare.

REFERENCES

1. Kyle RA, Therneau TM, Rajkumar SV, et al. Long-term follow-up of IgM monoclonal gammopathy of undetermined significance. *Blood*. 2003;102(10):3759–64.
2. Baldini L, Goldaniga M, Guffanti A, et al. Immunoglobulin M monoclonal gammopathies of undetermined significance and indolent Waldenstrom's macroglobulinemia recognize the same determinants of evolution into symptomatic lymphoid disorders: proposal for a common prognostic scoring system. *J. Clin. Oncol.* 2005;23(21):4662–8.
3. Morra E, Cesana C, Klersy C, et al. Clinical characteristics and factors predicting evolution of asymptomatic IgM monoclonal gammopathies and IgM-related disorders. *Leukemia*. 2004;18(9):1512–7.
4. 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.
5. Chapman M a., Lawrence MS, Keats JJ, et al. Initial genome sequencing and analysis of multiple myeloma. *Nature*. 2011;471(7339):467–472.
6. Puente XS, Pinyol M, Quesada V, et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2011;475(7354):3–7.
7. Montesinos-Rongen M, Godlewska E, Brunn A, et al. Activating L265P mutations of the MYD88 gene are common in primary central nervous system lymphoma. *Acta Neuropathol.* 2011;122(6):791–2.
8. Ngo VN, Young RM, Schmitz R, et al. Oncogenically active MYD88 mutations in human lymphoma. *Nature*. 2011;470(7332):115–9.
9. Pasqualucci L, Trifonov V, Fabbri G, et al. Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat. Genet.* 2011;43(9):830–7.
10. Jiménez C, Sebastián E, Del Carmen Chillón M, et al. MYD88 L265P is a marker highly characteristic of, but not restricted to, Waldenström's macroglobulinemia. *Leukemia*. 2013;27(8):1722–8.
11. Treon SP, Hunter ZR, Aggarwal a, et al. Characterization of familial Waldenstrom's macroglobulinemia. *Ann. Oncol.* 2006;17(3):488–94.
12. Braggio E, Keats JJ, Leleu X, et al. Identification of copy number abnormalities and inactivating mutations in two negative regulators of

- nuclear factor-kappaB signaling pathways in Waldenstrom's macroglobulinemia. *Cancer Res.* 2009;69(8):3579–88.
13. Schop RFJ, Van Wier SA, Xu R, et al. 6q deletion discriminates Waldenström macroglobulinemia from IgM monoclonal gammopathy of undetermined significance. *Cancer Genet. Cytogenet.* 2006;169(2):150–3.
 14. Braggio E, Keats JJ, Leleu X, et al. High-resolution genomic analysis in Waldenström's macroglobulinemia identifies disease-specific and common abnormalities with marginal zone lymphomas. *Clin. Lymphoma Myeloma.* 2009;9(1):39–42.
 15. Nguyen-Khac F, Lambert J, Chapiro E, et al. Chromosomal aberrations and their prognostic value in a series of 174 untreated patients with Waldenstrom's macroglobulinemia. *Haematologica.* 2013;98(4):649–54.
 16. 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):110–5.
 17. Futreal PA, Coin L, Marshall M, et al. A census of human cancer genes. *Nat. Rev. Cancer.* 2004;4(3):177–83.
 18. Morin RD, Mendez-Lago M, Mungall AJ, et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature.* 2011;476(7360):298–303.
 19. Ngo HT, Leleu X, Lee J, et al. SDF-1/CXCR4 and VLA-4 interaction regulates homing in Waldenstrom macroglobulinemia. *Blood.* 2008;112(1):150–8.
 20. Hernandez PA, Gorlin RJ, Lukens JN, et al. Mutations in the chemokine receptor gene CXCR4 are associated with WHIM syndrome, a combined immunodeficiency disease. *Nat. Genet.* 2003;34(1):70–4.
 21. Balabanian K, Lagane B, Pablos JL, et al. WHIM syndromes with different genetic anomalies are accounted for by impaired CXCR4 desensitization to CXCL12. *Blood.* 2005;105(6):2449–57.
 22. Wu B, Chien EYT, Mol CD, et al. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science (80-.).* 2010;330(6007):1066–71.

23. Orsini MJ. Trafficking of the HIV Coreceptor CXCR4. Role of arrestins and identification of residues in the c-terminal tail that mediate receptor internalization. *J. Biol. Chem.* 1999;274(43):31076–31086.
24. Haribabu B, Richardson RM, Fisher I, et al. Regulation of human chemokine receptors CXCR4. Role of phosphorylation in desensitization and internalization. *J. Biol. Chem.* 1997;272(45):28726–31.
25. Lagane B, Chow KYC, Balabanian K, et al. CXCR4 dimerization and beta-arrestin-mediated signaling account for the enhanced chemotaxis to CXCL12 in WHIM syndrome. *Blood.* 2008;112(1):34–44.
26. Poulain S, Roumier C, Decambron A, et al. MYD88 L265P mutation in Waldenstrom's macroglobulinemia. *Blood.* 2013;121(22):4504–4511.
27. Leleu X, Eeckhoutte J, Jia X, et al. Targeting NF-kappaB in Waldenstrom macroglobulinemia. *Blood.* 2008;111(10):5068–77.
28. Yang G, Zhou Y, Liu X, et al. A mutation in MYD88 (L265P) supports the survival of lymphoplasmacytic cells by activation of Bruton tyrosine kinase in Waldenström macroglobulinemia. *Blood.* 2013;88(7):1222–32.
29. Dale DC, Bolyard AA, Kelley ML, et al. The CXCR4 antagonist plerixafor is a potential therapy for myelokathexis, WHIM syndrome. *Blood.* 2011;118(18):4963–6.
30. Poulain S, Ertault M, Leleu X, et al. SDF1/CXCL12 (-801GA) polymorphism is a prognostic factor after treatment initiation in Waldenstrom macroglobulinemia. *Leuk. Res.* 2009;33(9):1204–7.
31. Hooks SB, Callihan P, Altman MK, et al. Regulators of G-Protein signaling RGS10 and RGS17 regulate chemoresistance in ovarian cancer cells. *Mol. Cancer.* 2010;9(1):289.
32. Busillo JM, Benovic JL. Regulation of CXCR4 signaling. *Biochim. Biophys. Acta.* 2007;1768(4):952–63.
33. Nagl NG, Wang X, Patsialou A, Van Scoy M, Moran E. Distinct mammalian SWI/SNF chromatin remodeling complexes with opposing roles in cell-cycle control. *EMBO J.* 2007;26(3):752–63.
34. Choi J, Ko M, Jeon S, et al. The SWI/SNF-like BAF Complex Is Essential for Early B Cell Development. *J. Immunol.* 2012;188(8):3791–803.

35. Heltemes-Harris LM, Willette MJL, Ramsey LB, et al. Ebf1 or Pax5 haploinsufficiency synergizes with STAT5 activation to initiate acute lymphoblastic leukemia. *J. Exp. Med.* 2011;208(6):1135–49.
36. Jones S, Li M, Parsons DW, et al. Somatic mutations in the chromatin remodeling gene ARID1A occur in several tumor types. *Hum. Mutat.* 2012;33(1):100–103.
37. Guan B, Wang T-L, Shih I-M. ARID1A, a factor that promotes formation of SWI/SNF-mediated chromatin remodeling, is a tumor suppressor in gynecologic cancers. *Cancer Res.* 2011;71(21):6718–27.
38. Inoue H, Giannakopoulos S, Parkhurst CN, et al. Target genes of the largest human SWI/SNF complex subunit control cell growth. *Biochem. J.* 2011;434(1):83–92.
39. Humbert N, Martien S, Augert A, et al. A Genetic Screen Identifies Topoisomerase 1 as a Regulator of Senescence. *Cancer Res.* 2009;69(10):4101–4106.
40. Shadat NMA, Koide N, Khuda II-EI-E, et al. Retinoblastoma Protein-Interacting Zinc Finger 1 (RIZ1) Regulates the Proliferation of Monocytic Leukemia Cells via Activation of p53. *Cancer Invest.* 2010;28(8):806–12.
41. Wall M, Rayeroux KC, MacKinnon RN, Zordan A, Campbell LJ. ETV6 deletion is a common additional abnormality in patients with myelodysplastic syndromes or acute myeloid leukemia and monosomy 7. *Haematologica.* 2012;97(12):1933–6.
42. Stegmaier K, Pendse S, Barker GF, et al. Frequent loss of heterozygosity at the TEL gene locus in acute lymphoblastic leukemia of childhood. *Blood.* 1995;86(1):38–44.
43. Kuiper RP, Schoenmakers EFPM, van Reijmersdal S V, et al. High-resolution genomic profiling of childhood ALL reveals novel recurrent genetic lesions affecting pathways involved in lymphocyte differentiation and cell cycle progression. *Leukemia.* 2007;21(6):1258–66.
44. Bohlander SK. ETV6: a versatile player in leukemogenesis. *Semin. Cancer Biol.* 2005;15(3):162–74.
45. Obrador-Hevia A, Serra-Sitjar M, Rodríguez J, Villalonga P, Fernández de Mattos S. The tumour suppressor FOXO3 is a key regulator of mantle cell lymphoma proliferation and survival. *Br. J. Haematol.* 2012;156(3):334–45.

46. Tichioni M, Essafi M, Jeandel PY, et al. Homeostatic chemokines increase survival of B-chronic lymphocytic leukemia cells through inactivation of transcription factor FOXO3a. *Oncogene*. 2007;26(50):7081–91.
47. Karube K, Nakagawa M, Tsuzuki S, et al. Identification of FOXO3 and PRDM1 as tumor-suppressor gene candidates in NK-cell neoplasms by genomic and functional analyses. *Blood*. 2011;118(12):3195–204.
48. Bakker WJ, Blázquez-Domingo M, Kolbus A, et al. FoxO3a regulates erythroid differentiation and induces BTG1, an activator of protein arginine methyl transferase 1. *J. Cell Biol.* 2004;164(2):175–84.
49. Rouault JP, Rimokh R, Tessa C, et al. BTG1, a member of a new family of antiproliferative genes. *EMBO J.* 1992;11(4):1663–70.
50. Lohr JG, Stojanov P, Lawrence MS, et al. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc. Natl. Acad. Sci. U. S. A.* 2012;109(10):3879–84.
51. Mullighan CG, Goorha S, Radtke I, et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature*. 2007;446(7137):758–64.
52. Waanders E, Scheijen B, van der Meer LT, et al. The origin and nature of tightly clustered BTG1 deletions in precursor B-cell acute lymphoblastic leukemia support a model of multiclonal evolution. *PLoS Genet.* 2012;8(2):e1002533.
53. Van Galen JC, Kuiper RP, van Ernst L, et al. BTG1 regulates glucocorticoid receptor autoinduction in acute lymphoblastic leukemia. *Blood*. 2010;115(23):4810–9.
54. Treon SP. How I treat Waldenström macroglobulinemia. *Blood*. 2009;114(12):2375–85.
55. Moritake H, Shimonodan H, Marutsuka K, et al. C-MYC rearrangement may induce an aggressive phenotype in anaplastic lymphoma kinase positive anaplastic large cell lymphoma: Identification of a novel fusion gene ALO17/C-MYC. *Am. J. Hematol.* 2011;86(1):75–8.
56. Cools J, Wlodarska I, Somers R, et al. Identification of novel fusion partners of ALK, the anaplastic lymphoma kinase, in anaplastic large-cell lymphoma and inflammatory myofibroblastic tumor. *Genes. Chromosomes Cancer*. 2002;34(4):354–62.

57. Owen HR, Elser M, Cheung E, et al. MYBBP1a is a novel repressor of NF-kappaB. *J. Mol. Biol.* 2007;366(3):725–36.
58. Iwashita Y, Fukuchi N, Waki M, Hayashi K, Tahira T. Genome-wide repression of NF-kB target genes by transcription factor MIBP1 and its modulation by O-linked β -N-acetylglucosamine (O-GlcNAc) transferase. *J. Biol. Chem.* 2012;287(13):9887–900.
59. Rossi D, Deaglio S, Dominguez-Sola D, et al. Alteration of BIRC3 and multiple other NF-kB pathway genes in splenic marginal zone lymphoma. *Blood.* 2011;118(18):4930–4.
60. Craxton A. Syk and Bruton's Tyrosine Kinase Are Required for B Cell Antigen Receptor-mediated Activation of the Kinase Akt. *J. Biol. Chem.* 1999;274(43):30644–30650.
61. Niir H, Allam A, Stoddart A, et al. The B Lymphocyte Adaptor Molecule of 32 Kilodaltons (Bam32) Regulates B Cell Antigen Receptor Internalization. *J. Immunol.* 2004;173(9):5601–5609.
62. Gauld SB, Cambier JC. Src-family kinases in B-cell development and signaling. *Oncogene.* 2004;23(48):8001–8006.
63. Liu W, Quinto I, Chen X, et al. Direct inhibition of Bruton's tyrosine kinase by IBtk, a Btk-binding protein. *Nat. Immunol.* 2001;2(10):939–46.
64. Jefferies CA. Bruton's Tyrosine Kinase Is a Toll/Interleukin-1 Receptor Domain-binding Protein That Participates in Nuclear Factor B Activation by Toll-like Receptor 4. *J. Biol. Chem.* 2003;278(28):26258–26264.
65. Liu X, Zhan Z, Li D, et al. Intracellular MHC class II molecules promote TLR-triggered innate immune responses by maintaining activation of the kinase Btk. *Nat. Immunol.* 2011;12(5):416–24.
66. Makishima H, Jankowska AM, McDevitt MA, et al. CBL, CBLB, TET2, ASXL1, and IDH1/2 mutations and additional chromosomal aberrations constitute molecular events in chronic myelogenous leukemia. *Blood.* 2011;117(21):e198–206.
67. Bachmaier K, Taya S, Gao X, et al. E3 ubiquitin ligase Cblb regulates the acute inflammatory response underlying lung injury. *Nat. Med.* 2007;13(8):920–6.
68. Xu L, Hunter ZR, Yang G, et al. MYD88 L265P in Waldenstrom's Macroglobulinemia, IgM Monoclonal Gammopathy, and other B-cell

Lymphoproliferative Disorders using Conventional and Quantitative Allele-Specific PCR. *Blood*. 2013;121(11):2051–2058.

69. Varettoni M, Arcaini L, Zibellini S, et al. Prevalence and clinical significance of the MYD88 (L265P) somatic mutation in Waldenstrom's macroglobulinemia and related lymphoid neoplasms. *Blood*. 2013;121(13):2522–8.

Tables

Table 1: Validated somatic variants identified by WGS in WM patients.

List of all variants validated by Sanger sequencing ordered by chromosome and position.

Gene	N	Chr	Position	Reference	Variant	Zygosity	Protein	Note
ARID1A	1	1	27057944	C	-	hom	Y551Frame Shift	Homozygous by aUPD
ARID1A	1	1	27099946	C	T	het	R1276Stop	
ARID1A	1	1	27101665	C	T	het	Q1650Stop	
ARID1A	1	1	27106497	C	T	het	Q2037Stop	Opposite Copy Loss
ARID1A	1	1	27107136	-	A	het	S2249Frame Shift	
NOTCH2	1	1	120478125	A	C	het	F1209V	
CXCR4	1	2	136872467	GAAGACTCAG	AC	het	S344Frame Shift	
CXCR4	1	2	136872481	AGAT	-	het	S339Frame Shift	
CXCR4	2	2	136872485	G	C	het	S338STOP	WHIM associated rs104893626
CXCR4	3	2	136872485	G	T	het	S338STOP	WHIM associated rs104893626
CXCR4	1	2	136872570	-	T	het	T311Frame Shift	
MAP2	1	2	210557805	G	C	het	W304C	

MYD88	1	3	38182032	C	G	het	S219C	
MYD88	27	3	38182640	T	C	het	L265P	Homozygous by aUPD x 4
MED23	1	6	131915434	C	T	het	E1013K	
SYNE1	1	6	152708319	C	T	hom	R2792H	Opposite Copy Loss
TRRAP	1	7	98530988	A	T	het	D1326V	
TRAF2	1	9	139815501	G	-	hom	L324Frame Shift	Homozygous by aUPD
RAG2	1	11	36615549	A	G	hom	L57P	Opposite Copy Loss
MLL2	1	12	49425598	A	G	het	S4297L	
MLL2	1	12	49448412	C	-	het	G101Frame Shift	
TRAF3	1	14	103371839	-	T	het	L448Frame Shift	Opposite Copy Loss
MYBBP1A	1	17	4457136	G	A	het	P177S	
MYBBP1A	1	17	4457323	C	T	het	R148Q	
TP53	1	17	7577114	C	T	het	C275Y	
TP53	1	17	7577569	A	C	hom	C238G	Homozygous by aUPD
CD79B	1	17	62006799	A	T	het	Y197N	
CD79B	1	17	62007129	C	G	het	Exon Boundry	
MUC16	1	19	9014687	G	T	het	A12763E	

Figure Legends

Figure 1: Somatic *CXCR4* mutations in WM are similar to those found in WHIM syndrome.

Somatic *CXCR4* nonsense and frame shift mutations found in 177 WM patient samples. (A) Protein sequence for the canonical full-length transcript (NP_003458.1) demonstrates that similar to the variants responsible for WHIM syndrome, these mutations result in a truncation of the cytosolic tail containing the regulatory phospho-serines leaving the seven transmembrane helix region involved in signaling and ligand binding intact. (B) The crystal structure of homodimeric *CXCR4* with the carboxyl terminal tail highlighted in yellow and indicated by the white arrows. (C) Precise location and number of WHIM-like somatic mutations in WM at transcript level and detailed mutation type summary of the most frequently mutated amino acid S338.

Figure 2: Summary of study results.

(A) Copy number results per patient for the independent 30 patient validation cohort as determined by qPCR. Somatic *MYD88* L265P and WHIM-like *CXCR4* mutations were assessed in this population and annotated here for reference. The eight CNA targets selected for validation were chosen based on the 10 paired patient WGS analysis and technical validation studies. (B) Validated mutations in the 30 patient WGS cohort. Somatic status of all findings was confirmed by germline Sanger sequencing. (C) Overall frequency of validated somatic mutation (above) or CNA (below) for the independent 30 patient WGS and validation cohorts respectively.

Figure 3: Acquired uniparental disomy (aUPD) in chromosome 3p in four patients with WM.

(A) Representative excerpts from paired tumor and germline tissue documenting an aUPD on chromosome 3 for a single WM patient. The red histogram below the ideogram corresponds to the heterozygous call rate. Below that, copy number (black line) and normalized relative coverage (orange) is shown above the lesser allele fraction (LAF) measurements (blue). LAF values more than 2 standard deviations from the mean are shown in green. (B) Sanger sequencing of *MYD88* L265P for this patient with a near homozygous mutant signal. © Heterozygous call rate, LAF, and relative coverage for the 4 patients with aUPDs in chromosome 3p demonstrating the near identical location of the aUPD for the 4 affected patients.

Figure 4: Characterization of somatic copy number alterations (CNA) in WM

(A) Functional annotation for genes affected by CNA found outside of chromosome 6. The list is ordered by statistical significance and filtered only for duplicated functional annotations matching more than one category. (B) Deletions of matching size were randomly distributed across the genome in 10,000 trials. The number of affected total RefSeq and COSMIC genes was calculated for each group. Results represent mean values with empirical 95% confidence intervals.

Figure 5: Somatic deletions identified on chromosome 6 by WGS in WM patients.

(A) Frequency of statistically significant chromosome 6 deletions from the 10 paired patients highlighting genes of interest. The positions of the deletions are mapped against chromosome 6 cytogenetic bands. (B) Relative coverage across chromosome 6 for each of the 10 paired samples. Losses in 6q were not always single contiguous deletions and some patients had deleted segments restricted to a subclone. The frequency of deletions for genes including *HIVEP2* and *ARID1B* were higher than the corresponding number of large block deletions.

Figure 6: Validation results using quantitative polymerase chain reaction (qPCR) for most frequent somatic deletions identified by WGS in WM patients.

Five patient samples, 3 from the paired and 2 from the unpaired WGS cohorts were selected for validation studies using qPCR copy number assays. All assays were run in at least triplicate. Results represent median values and ranges. (A) Validation of deletion in the known 6q deletion in two patients at *HINT3* (6q22.32). (B) Representative validation results normalized to germline as determined by qPCR. Deletions deemed significant by Welch's t-test are marked with *. (C) Somatic relative copy number needs to be interpreted in context. Significant germline copy number variation was noted in both *HIVEP2* and *MKLN1*. (D) Comparison of whole genome and qPCR validation estimates of relative somatic copy number demonstrating similar clonal estimates.

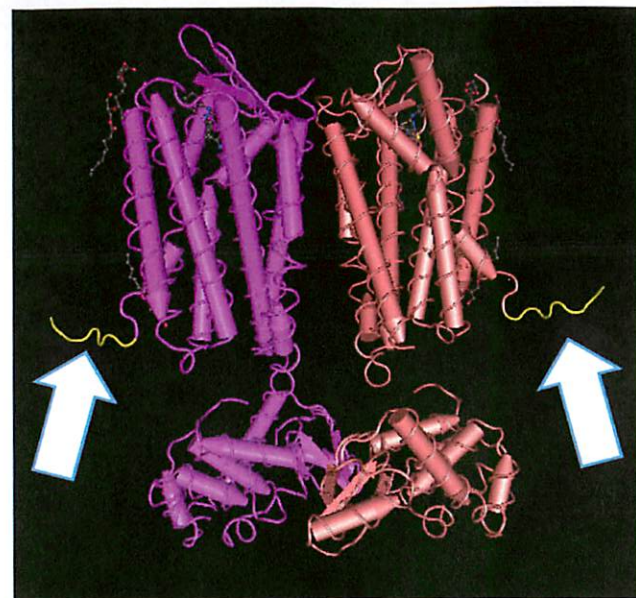
A

MEGISIYTS DNYTEEMGSGDYDSMKEPCFREENANFNK**IFLPTI**
YSIIFLTGIVGNGLVILVMGYQKKLRSM TD KYRL**LHLSVADLLFVI**
TLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSSVLILAFISLD
 RYLAIVHATNSQRPRKLLAEK**VVYVGWIPALLTIPDFIFANVS**
 EADDRYICDRFYPNDLW**VVFQFQHIMVGLILPGIVILSCYCIIS**
 KLSHSGKHQKRKALKTT**VILILAFFACWLPYYIGIS**DSFILLEIK
 QGCEFENTVHK**WISITEALAFFHCCLNPILYAFLGAKFKT**SAQH
 ALTSMVRGSSSLKILSKGK**R**GGH**SS**MT**SE**SSSSFHSS

LEGEND

- A** - Germline variant in WHIM syndrome **A** - Transmembrane helix
□ - Somatic frame shift or nonsense WM variant

B

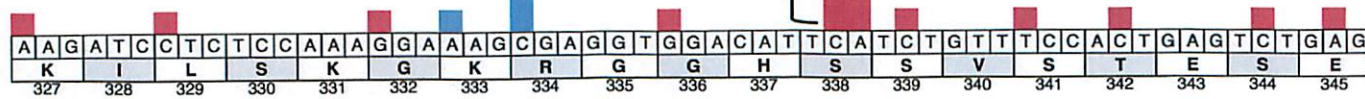
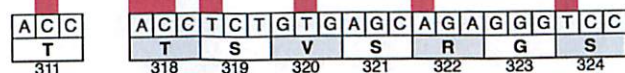
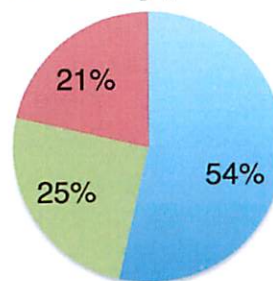


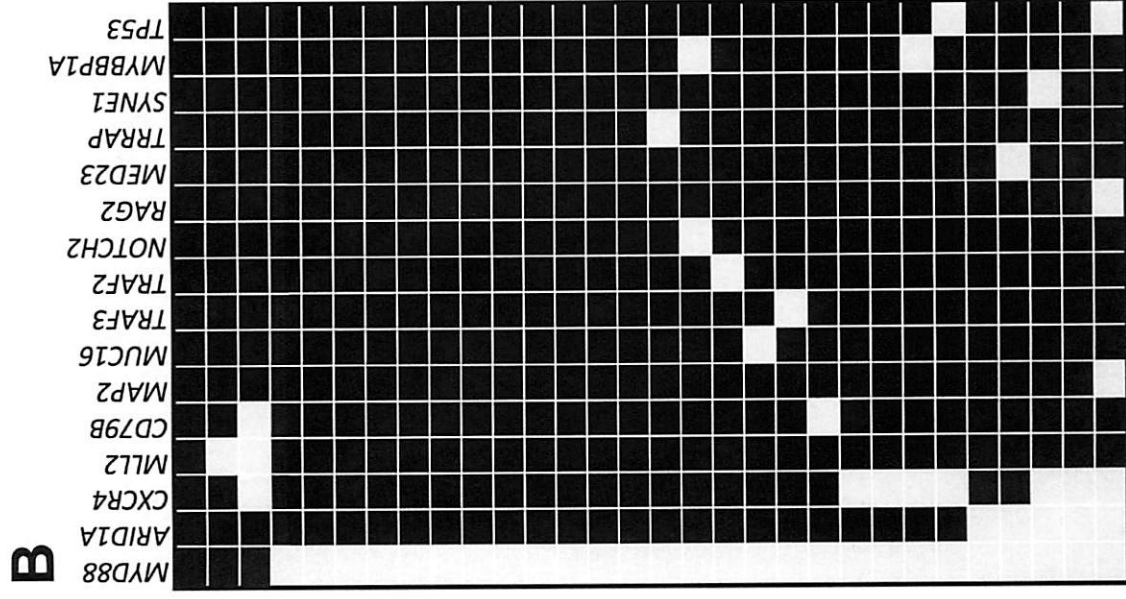
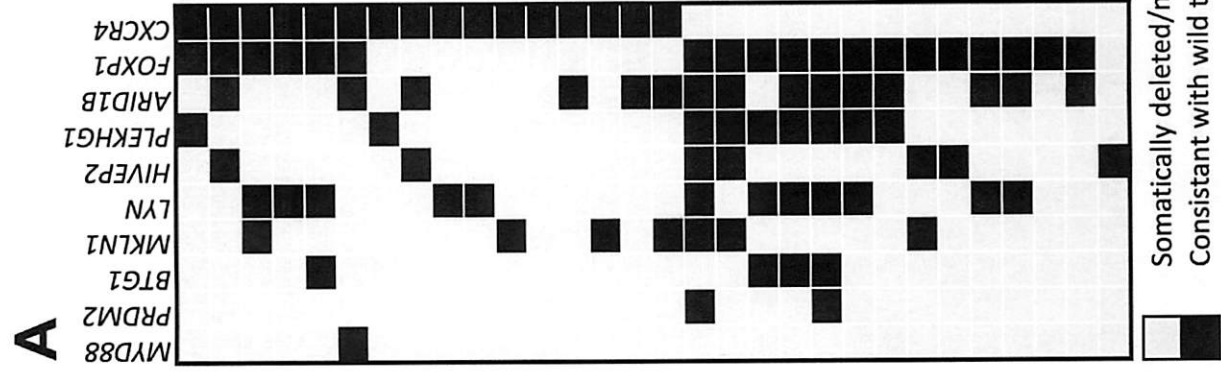
C

- Frame shift mutation
■ Nonsense mutation

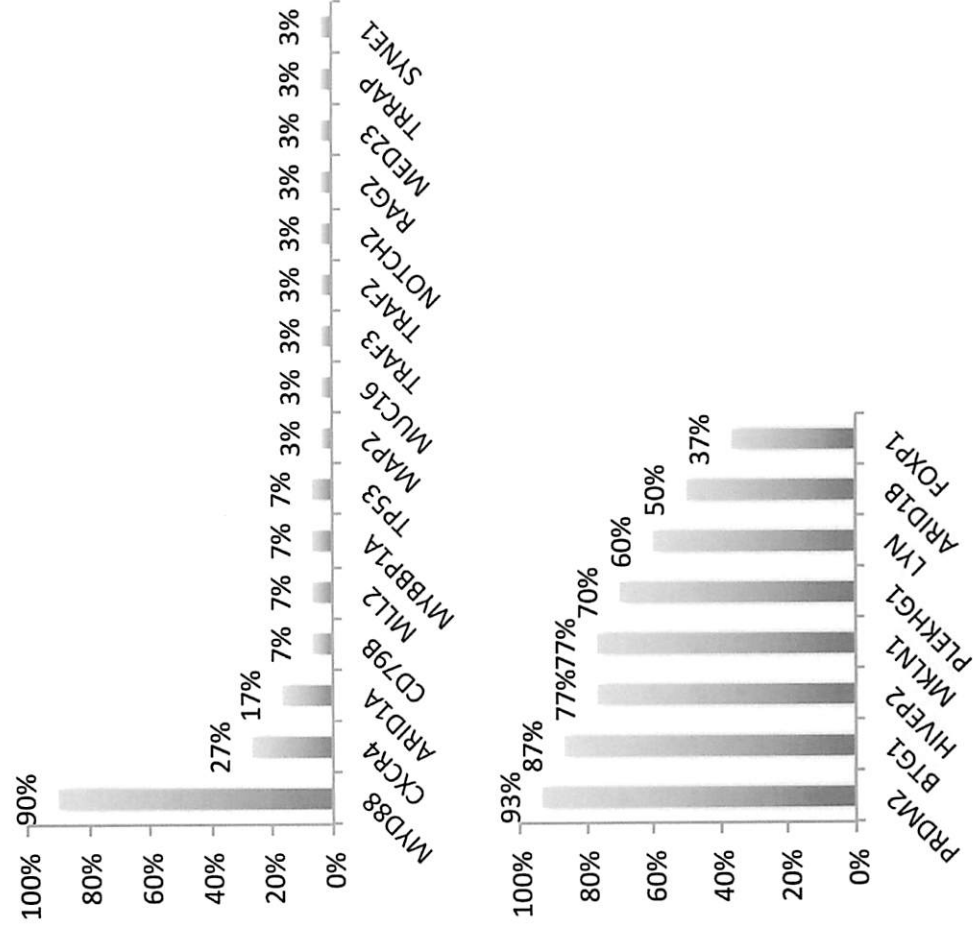
S338 Mutation Types

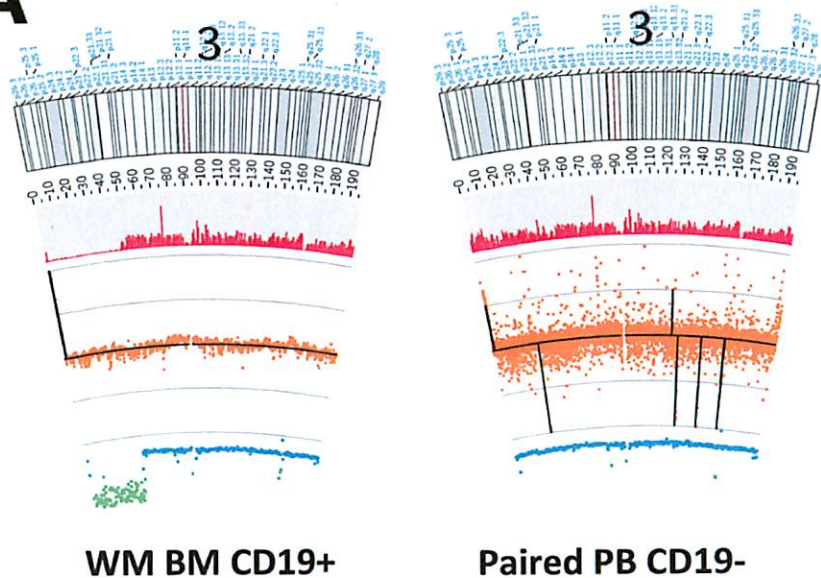
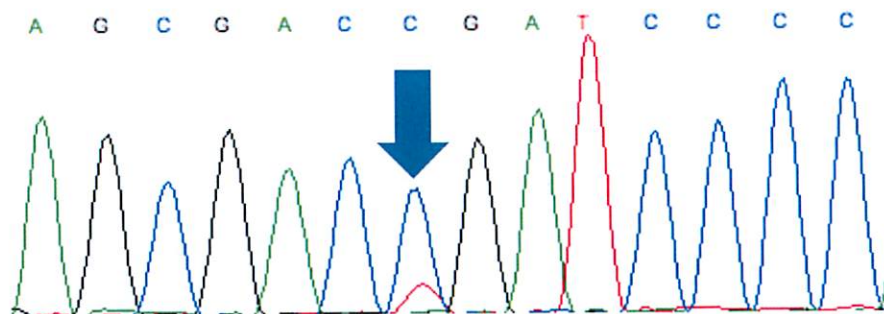
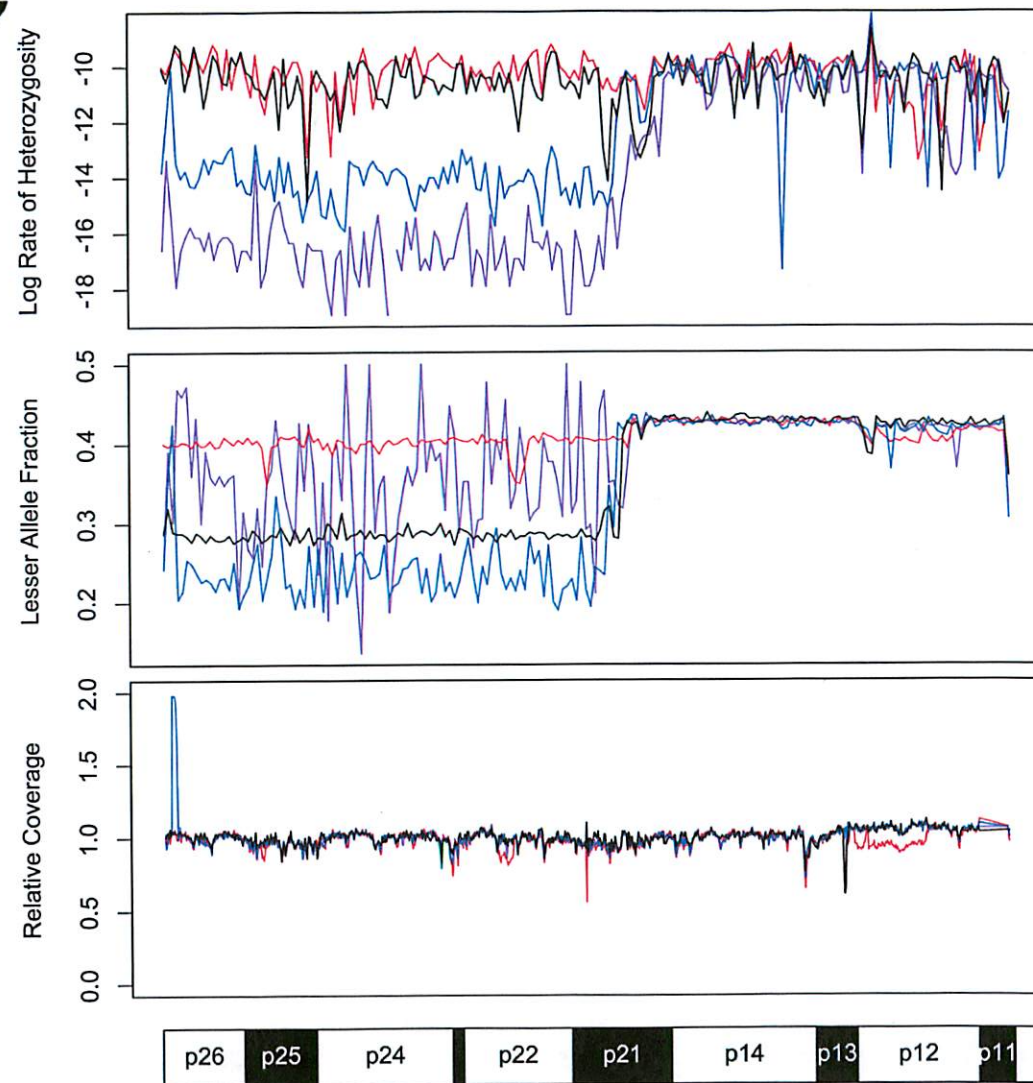
- Nonsense C/G
■ Nonsense C/A
■ Frame shift





C



A**B****C**

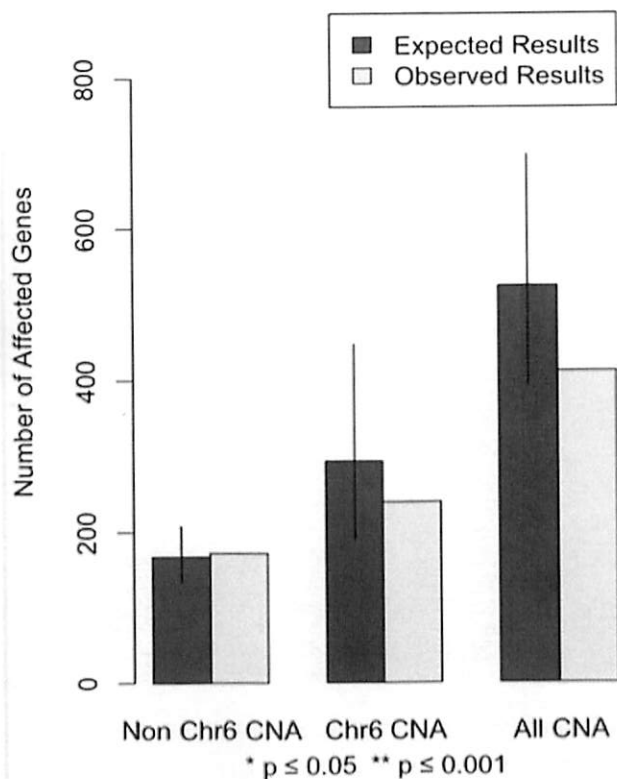
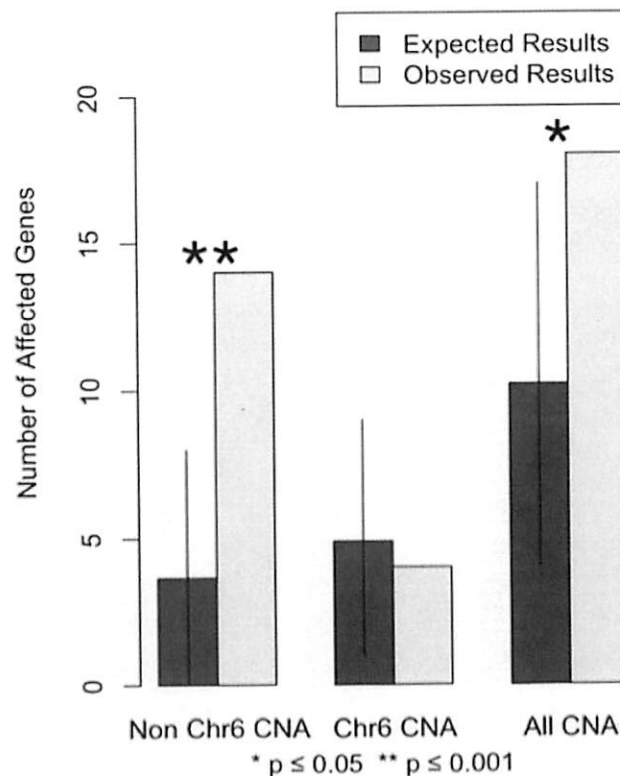
Chromosome 3p

Legend

— WM1 — WM2 — WM3 — WM4

A

Category	Functions Annotation	p-Value
Cancer	lymphohematopoietic cancer	6.07E-13
Hematological System Development and Function	quantity of B lymphocytes	1.65E-12
Cellular Growth and Proliferation	proliferation of immune cells	2.90E-12
Hematopoiesis	arrest in differentiation of lymphocytes	6.39E-12
Cellular Growth and Proliferation	proliferation of lymphocytes	8.96E-12
Hematological System Development and Function	quantity of blood cells	1.23E-11
Hematological System Development and Function	quantity of leukocytes	3.49E-11
Hematological Disease	hematological neoplasia	3.64E-11

B**Genes Affected by Somatic CNA****Number of Affected COSMIC Genes****COSMIC Genes as Percentage of Affected**