LYMPHOID NEOPLASIA

Transcriptome sequencing reveals a profile that corresponds to genomic variants in Waldenström macroglobulinemia

Zachary R. Hunter,1,2 Lian Xu,1 Guang Yang,1,2 Nicholas Tsakmaklis,1 Josephine M. Vos,1 Xia Liu,1 Jie Chen,1 Robert J. Manning,1 Ji-Ji G. Chen,1 Philip Brodsky,1 Christopher J. Patterson,1 Joshua Gustine,1 Toni Dubéau,1 Jorge J. Castillo,1,2 Kenneth C. Anderson,2,3 Nikhil M. Munshi,2,4 and Steven P. Treon 1,2

1Bing Center for Waldenström’s Macroglobulinemia, Dana-Farber Cancer Institute, Boston, MA; 2Department of Medicine, Harvard Medical School, Boston, MA; 3Jerome Lipper Multiple Myeloma Center, Dana-Farber Cancer Institute, Boston, MA; and 4Department of Medicine, Boston VA Healthcare System, West Roxbury, MA

Key Points

- Transcription profiles associated with mutated MYD88, CXCR4, ARID1A, abnormal cytogenetics including 6q−, and familial WM are described.
- Mutated CXCR4 profiles show impaired expression of the tumor suppressor response induced by MYD88L265P and also G-protein/MAPK inhibitors.
- Whole-genome sequencing has identified highly prevalent somatic mutations including MYD88, CXCR4, and ARID1A in Waldenström macroglobulinemia (WM). The impact of these and other somatic mutations on transcriptional regulation in WM remains to be clarified. We performed next-generation transcriptional profiling in 57 WM patients and compared findings to healthy donor B cells. Compared with healthy donors, WM patient samples showed greatly enhanced expression of the VDJ recombination genes DNTT, RAG1, and RAG2, but not AICDA. Genes related to CXCR4 signaling were also upregulated and included CXCR4, CXCL12, and VCAM1 regardless of CXCR4 mutation status, indicating a potential role for CXCR4 signaling in all WM patients. The WM transcriptional profile was equally dissimilar to healthy memory B cells and circulating B cells likely due increased differentiation rather than cellular origin. The profile for CXCR4 mutations corresponded to diminished B-cell differentiation and suppression of tumor suppressors upregulated by MYD88 mutations in a manner associated with the suppression of TLR4 signaling relative to those mutated for MYD88 alone. Promoter methylation studies of top findings failed to explain this suppressive effect but identified aberrant methylation patterns in MYD88 wild-type patients. CXCR4 and MYD88 transcription were negatively correlated, demonstrated allele-specific transcription bias, and, along with CXCL13, were associated with bone marrow disease involvement. Distinct gene expression profiles for patients with wild-type MYD88, mutated ARID1A, familial predisposition to WM, chr6q deletions, chr3q amplifications, and trisomy 4 are also described. The findings provide novel insights into the molecular pathogenesis and opportunities for targeted therapeutic strategies for WM. (Blood. 2016;128(6):827-838)

Introduction

Whole-genome sequencing identified several highly recurring somatic mutations in patients with Waldenström macroglobulinemia (WM).1,2 In over 90% of WM patients, a single point mutation at NM_002468:c.978T>C (rs387907272) in MYD88 is found, resulting in a p.Leu265Pro (L265P) amino acid change.1,3 MYD88 is an adaptor for Toll-like (TLR) and interleukin 1 (IL1) receptors, and the MYD88L265P mutation triggers constitutive activation of NF-kB through IRAK and BTK.1,4,5 Mutated MYD88 WM patients show greater overall survival and clinical responses to the BTK inhibitoribrutinib.6,7

Activating CXCR4 frameshift or nonsense mutations in the C-terminal tail are found in 30% to 40% of WM patients, are primarily subclonal, and almost always associated with MYD88L265P.2,3,8 These somatic mutations are similar to the causal germ line variants that underlie WHIM (warts, hypogammaglobulinemia, infection, and myelokathexis) syndrome.2,9 In WM, somatic CXCR4 mutations (CXCR4<sup>WHIM</sup>) are dominants of disease presentation, as well as resistance to ibrutinib.10,11 Somatic mutations in ARID1A, a member of the SW1/Sucrose nonfermentable family and epigenetic regulator, are also found in 20% of WM patients. Gene losses affecting NF-kB signaling (ie, HIVEP2, TNFAIP3), as well as the ARID1A homolog ARID1B are present in most WM patients.2,11 Deletions in chromosome 6q with and without concurrent 6p amplifications, trisomy 3, and amplifications of 3q, as well as trisomy 4, are also commonly found in WM.12-14 Previous array-based gene expression studies of WM were largely conducted prior to these genomic discoveries and therefore the effects of recurrent somatic events on transcriptional regulation remain to be clarified.15-17 We therefore performed next-generation RNA sequencing in 57 WM patients and compared findings to sorted healthy donor-derived nonmemory (CD19<sup>+</sup>CD27<sup>−</sup>) and memory (CD19<sup>+</sup>CD27<sup>+</sup>) B cells. The latter
represent the B-cell population from where most cases of WM are thought to be derived.\textsuperscript{18,19}

Methods

Sample selection and characterization

Bone marrow (BM) aspirates were collected from 57 patients with the WM consensus diagnosis.\textsuperscript{20} Participants provided informed consent for sample collection per the Dana-Farber/Harvard Cancer Center Institutional Review Board. WM cells were isolated by CD19$^+$ magnetic-activated cell sorting (MACS) microbead selection (Miltenyi Biotec, Auburn, CA) from Ficoll-Paque (Amersham-Pharmacia Biotec, Piscataway, NJ) separated BM mononuclear cells. Peripheral blood mononuclear cells from nine healthy donors (HDs) were sorted for nonmemory (CD19$^+$ CD27$^-$) and memory B cells (CD19$^+$ CD27$^+$) using a memory B-cell isolation kit (Miltenyi Biotec). RNA and DNA were purified using the AllPrep mini kit (Qiagen, Valencia, CA). Most samples were previously characterized by whole-genome sequencing and all samples were screened for MYD88 and CXCR4 gene mutations by Sanger sequencing.\textsuperscript{21} The MYD88$^{265P}$ and CXCR4 c.1013C$\rightarrow$G and c.1013C$\rightarrow$A mutations were analyzed by allele-specific polymerase chain reaction (PCR) as previously described.\textsuperscript{2,21}

Next-generation sequencing and analysis

Transcriptome profiling was conducted for the Center for Computational Biology at the Dana-Farber Cancer Institute (Boston, MA) using the NEBNext Ultra RNA library prep kit (New England BioLabs, Ipswich, MA). The paired-end samples were run 2 per lane for 50 cycles on an Illumina HiSeq (Illumina, San Diego, CA). Read-level data are available through dbGAP accession (Applied). Reads were aligned to KnownGene HG19/GRCh37 reference using STAR (Spliced Transcripts Alignment to a Reference).\textsuperscript{22} Genes with mean raw read counts of $< 10$ were not analyzed, leaving 16,888 expressed genes for analysis. Read counts per gene were obtained using featureCounts from Rsubread, and read-level data are archived in the Gene Expression Omnibus database (accession number GSE83150). Array data (Affymetrix, Santa Clara, CA) were analyzed using the Oligo and limma R bioconductor packages.\textsuperscript{23-27} Differential expression Omnibus database (accession number GSE83150). Array data were analyzed using voom from the edgeR/limma Bioconductor packages in R.\textsuperscript{23-27} Differential gene expression analysis was conducted using Ingenuity Pathway Analysis (Qiagen). Clustering and correlation analysis was conducted using the variance stabilization transformation of the count data from the Bioconductor DESeq2 package.\textsuperscript{28} In all other cases, estimates of gene expression levels are represented in transcripts per million (TPM). Log-fold change (LFC) listed in text and tables are derived from the limma analysis.

CXCR4 transduced cell lines and gene expression analysis

Previously described BCWM.1 and MWC-L1 cell lines transduced to express CXCR4$^+$ with or without activating mutations observed in WM patients were used to model CXCR4-stimulated gene expression changes in WM.\textsuperscript{18} Briefly, CXCR4 complementary DNA (cDNA) transcripts were subcloned into plenti-internal ribosomal entry site (IRES)–green fluorescent protein (GFP) vectors, and stably transduced using a lentiviral system.\textsuperscript{29} In addition to wild-type (WT) CXCR4, vectors with the mutations c.1013C$\rightarrow$G (p.Ser344fs), c.932_933insT (p.Thr311fs), and c.1030_1041delinsGT (p.Ser344fs) were generated. Cell lines were stimulated with 50 nM of the CXCR4 ligand CXCL12 (SDF1A) (R&D Systems, Minneapolis MN) for 2 hours. RNA from each sample was extracted at baseline and at 2 hours. Gene expression was assessed using Affymetrix Human Gene ST 2.0 arrays (Affymetrix, Santa Clara, CA). Array data are archived in the Gene Expression Omnibus database (accession number GSE83150). Array data were analyzed using the Oligo and limma R bioconductor packages.\textsuperscript{24,29}

Statistical analysis

Random forest analysis was implemented using the cforest function in the party package in R.\textsuperscript{30} Models were constructed using the cforest_unbiased function with 20 predictors per tree over 10,000 trees. As variable importance scores of nonsignificant genes are distributed around zero, genes were further filtered for those with permutation importance scores greater than the absolute value of the lowest scoring predictor. For reproducibility, the random seed for the analysis was set to 265. Principal component analysis (PCA) was conducted using R’s precomp, Pearson’s product-moment coefficient was used for correlation studies, and cohort mean values were compared using the Wilcoxon rank sum test.

Gene expression, sequencing, and methylation studies

Gene expression results were validated by using TagMan gene expression assays Hs01027785_m1 (DUSP4), Hs00244839_m1 (DUSP5), Hs00243182_m1 (RGSL1), Hs00892674_m1 (RGSL6), Hs00698249_m1 (RASSF6), Hs00396602_m1 (WNK2), Hs00924602_m1 (PRDM5) (Thermo Fisher Scientific, Waltham MA). Methylation-specific PCR assays for RASSF6, WNK2, and PRDM5 were conducted on bisulfite-converted DNA using previously established protocols.\textsuperscript{31-33}

Results

The clinical characteristics for the 57 WM patients are presented in Table 1. The distribution of somatic mutations and cytogenetic findings in these patients was similar to those previously described.\textsuperscript{34} The top 500 differentially expressed genes for each comparison discussed below are available as a part of the supplemental Data (available on the Blood Web site).

Comparing WM to healthy donor B cells

Comparison of WM and HD samples generated a list of 13,571 differentially expressed genes. To gain further insight into these gene expression differences, this gene list was sorted by absolute log$_2$ fold change, revealing a marked upregulation of DNTT, RAG1, RAG2, and IGLLi in WM samples. These genes have important functional roles in VDJ recombination and B-cell receptor (BCR) signaling. Other upregulated genes relevant to WM biology and pathogenesis included CXCL12, VCAM1, CD5L, BMP3, and IGF1. Among the top 40 differentially expressed genes sorted by absolute expression in WM as measured in TpM were CXCR4, the WM prognostic marker B2M, and 2 genes related to BCR signaling: CD79A and CD79B. To examine transcriptional similarities between the samples in an unbiased manner, multidimensional scaling of the 500 genes with the highest variance was used to best represent the data in 2 dimensions (Figure 1A). Although peripheral B cells (PBs) and HD memory B cells (MBs) clustered together, separation between HD PB/MB cell type using PCA of the 596 genes differentially expressed\textsuperscript{35,36} was not seen. In contrast, HD B cells and WM\textsuperscript{35,36} were analyzed using the Oligo and limma R bioconductor packages.\textsuperscript{24,29}
WM patients based on MYD88 and CXCR4 mutation status suggesting reduced differentiation in both MYD88L265P CXCR4WT and, in particular, MYD88WT CXCR4WT samples. Separation of patient samples by genotype can be seen in the PCA of these 19 genes in Figure 1E. BCL2 and BCL2L1 were upregulated in all WM samples, however, analysis of the larger BCL2 family revealed stratification of samples by HD cell type and MYD88/CXCR4 mutation status (supplemental Figure 2). The proapoptotic BAX gene was downregulated in WM samples and PMAIP1 was uniquely upregulated in MYD88L265P CXCR4WT WM.

Transcriptional effects of CXCR4WT

Restricting the analysis to WM patients, PCA of the top 500 high variance genes accounted for 41% of the overall variability within the first 2 components (Figure 2A). Tumor suppressors including WNK2, TP53I11, PRDM5, and CABLES1 influence the second principal component that separates MYD88L265P CXCR4WT from MYD88L265P CXCR4WHIM and MYD88WT CXCR4WT samples (supplemental Figure 1). Supervised clustering of the 3103 genes differentially expressed between MYD88L265P CXCR4WT and MYD88L265P CXCR4WHIM revealed that most expression differences in MYD88L265P CXCR4WT patients followed a pattern resembling HD samples, despite carrying the MYD88L265P mutation (Figure 2B). The top predicted upstream regulator for this gene signature was the inhibition of lipopolysaccharide signaling (Figure 2C). To further explore how CXCR4 mutations modulate MYD88 signaling in WM, the gene signature was filtered for genes involved in TLR signaling. MYD88L265P CXCR4WHIM patients showed downregulation of the lipopolysaccharide receptors TLR4 and NOD2, and upregulation of TLR7 and IRAK3 (supplemental Figure 3A). The most significant results associated with CXCR4 mutation status were silencing of genes in patients with MYD88L265P CXCR4WHIM that were elevated in MYD88L265P CXCR4WT patients (supplemental Figure 3B).

Promoter methylation studies

To investigate the role of methylation in genes differentially expressed based on CXCR4 and MYD88 mutation status, 3 such genes were selected for screening based on their regulation by promoter methylation in other malignancies: PRDM5, WNK2, and RASSF6. Gene expression was validated by PCR in HD B-cell samples as well as in MYD88 and CXCR4 genotyped WM patient samples (supplemental Figure 4A). Mirroring the transcriptome data, RASSF6 expression was elevated in MYD88L265P CXCR4WT and MYD88L265P CXCR4WHIM samples but not in HD or MYD88WT samples whereas WNK2 and PRDM5 were elevated only in the MYD88L265P CXCR4WT patients. To determine relevance of these methylation sites in hematologic malignancies, WM, B-cell lymphoma, and myeloma cell lines were assessed for promoter methylation using methylation-specific PCR assays (supplemental Figure 4B). Sixteen MYD88 and CXCR4 genotyped primary WM patient samples and 4 healthy donor B-cell samples were then tested (supplemental Figure 4C). Robust promoter methylation was observed for WNK2 and PRDM5 in the MYD88WT CXCR4WT patients but was absent in HD samples. Partial promoter methylation of WNK2 and PRDM5 was also observed in some MYD88L265P samples. This correlated to a relative reduction in gene expression only for WNK2 in the MYD88L265P CXCR4WHIM genotyped patient sample with the strongest methylation signal.

Cell line models of CXCR4WHIM signaling

To better understand interactions between CXCR4WHIM and MYD88L265P in WM, lentiviral transduction was performed to overexpress CXCR4WT or 1 of 3 documented CXCR4WHIM transcripts (c.1013G>G; c.932_933insT, or c.1030_1041delinsGT) in MYD88L265P expressing BCWM.1 WM cells. Transduced cells were stimulated in triplicate with CXCL12 for 2 hours and gene expression profiling performed. The relative change in gene expression from baseline was then compared between CXCR4WT and CXCR4WHIM transduced cells revealing 61 differentially expressed genes. Of these genes, 16 of 61 (26.2%) were encompassed in the gene expression signature for MYD88L265P CXCR4WHIM patient samples (supplemental Table 1). In 8 of 16 (50%) of these genes, the direction of the change relative to MYD88L265P CXCR4WT was same as that observed in the patient samples. However, it was the genes whose direction of change was opposed that revealed the most about the underlying biology. These discordant genes were enriched for suppressors of MAPK and G-protein signaling downstream of CXCR4. They were preferentially induced in CXCR4WHIM cell lines by CXCL12 stimulation and suppressed in MYD88L265P CXCR4WT patient samples.10,38 Interrogation of these dual specificity phosphatases (DUSP) and regulator of G-protein signaling (RGS) genes revealed RGS1, RGS2, RGS13, DUSP1, DUSP2, DUSP4, DUSP5, DUSP10, DUSP16, and DUSP22 were all significantly suppressed in MYD88L265P CXCR4WHIM vs MYD88L265P CXCR4WT patients. Expression in patients of DUSP4, DUSP5, RGS13, and RGS16, which were selected for validation based on the cell line studies, can be seen in supplemental Figure 5A. The increased expression of these genes in CXCR4WHIM compared with CXCR4WT transduced lines was validated by real-time PCR (supplemental Figure 5B). Dose-response curves for these genes following CXCL12 stimulation was conducted in transduced BCWM.1 and MWC.1 (supplemental Figure 5C).

Gene signatures associated with other clinical and genomic features

For the 5 patients with ARID1A mutations, 16 differentially expressed genes were observed. The ARID1A-mutated population also demonstrated elevated BM disease infiltration (median, 90%; range, 70%-95%) when compared with non-ARID1A-mutated patients (median, 58%; range, 5%-95%; P = .0259). No statistically significant gene expression changes were observed in the 4 patients with CD79B mutations. Analysis based on somatic cytogenetic
Figure 1. Comparisons of transcriptional activity for genotyped WM patient samples and healthy donor B cells. (A) Multidimensional scaling of the top 500 genes for all HD and WM samples using Voom transformed expression values in Limma’s plotMDS function.24 (B) Pearson correlation coefficients from the correlation matrix of gene expression were averaged over all the samples for each major group. Correlation values with the HD samples were higher for those WM patients with CXCR4 mutations ($P < .025$ and $P < .036$ for HD PB and HD MB cells, respectively). (C) The 596 genes differentially expressed between the HD PB and HD MB cells were analyzed across all samples using PCA. No preferential relationship between WM samples and HD B-cell type was observed. The first 3 principal components were used to plot the samples representing 54% of the overall variability. (D) Median expression levels for 19 genes known to regulate B-cell to plasma cell transition for HD B-cell and genotyped WM samples. Error bars represent full data range per for each group. (E) PCA of the 19 genes related to B-cell to plasma cell transition shown in panel D depicting the first 3 principal components which covered over 65% of the overall gene expression variability.
abnormalities included deletion of chromosome 6q (131 genes), amplification of chromosome 6p (65 genes), amplification of chr3q including trisomy 3 (11 genes), and trisomy 4 (776 genes). A distinct gene expression signature for 263 genes was also observed in the 3 patients who had first-degree relatives with WM. No statistically significant differences in gene expression were observed based on presence of extramedullary disease, elevated B2M levels (>3 mg/dL), or International Waldenström Macroglobulinemia Prognostic Scoring System score.

To identify genes associated with relevant WM clinical parameters, all genes were tested for correlation with serum immunoglobulin M (IgM) levels, hemoglobin, and BM disease involvement. After filtering the results by significance ($P < .05$) and the absolute value of the correlation estimate $>0.3$, 312 genes were found to be associated with IgM, 965 with hemoglobin, and 3738 with BM involvement. These genes were combined with $MYD88/CXCR4$ genotype, age, sex, and treatment status in a random forest regression analysis for the first 37 study patients, leaving the remaining 20 patients for
Figure 3.
cross-validation. Filtering genes/predictors by variable importance score resulted in 40, 236, and 143 associated genes for IgM, BM, and hemoglobin, respectively (supplemental Figure 6). Functional analysis was conducted on each gene list and top genes from each group were selected based on biological significance. The 40 genes associated with IgM were predicted to be downstream of IL6 signaling ($P = .0018$) and included many genes of WM pathogenic interest including TNFRSF17, BMP3, IRS1, and CEPZOS (Figure 3A). Many genes relevant to WM biology, including CXCL13, TP53, CXCR4, MYD88, CDC23, and AKAP1, were associated with BM disease involvement (Figure 3B). Although MYD88 and CXCR4 were both correlated to BM disease involvement ($\rho = -0.50; \rho < .0001$ and $\rho = 0.46; \rho = .0003$, respectively), they were negatively correlated with each other ($\rho = -0.46; P = .0004$). As shown in supplemental Figure 7, this relationship may be influenced by MD88/CXCR4 mutation status. The gene list associated with BM disease involvement included MYD88 and CXCR4, as well as PIK3AP1, PIK3CD, AJUBA, and OS9 (Figure 3C). Based on these findings, CXCL13 was included in an independent serum cytokine profiling study of 86 WM patients that validated the CXCL13 correlation between BM (Spearman $\rho = 0.562; P = 2.192 \times 10^{-8}$) and hemoglobin (Spearman $\rho = -0.567; P = 1.53 \times 10^{-8}$).39

**Analysis of mutant allele burden**

The percentage of reads supporting the mutant allele relative to WT allele was calculated for MYD88, CXCR4, ARID1A, and CD79B in the cross-validation. Filtering genes/predictors by variable importance score resulted in 40, 236, and 143 associated genes for IgM, BM, and hemoglobin, respectively (supplemental Figure 6). Functional analysis was conducted on each gene list and top genes from each group were selected based on biological significance. The 40 genes associated with IgM were predicted to be downstream of IL6 signaling ($P = .0018$) and included many genes of WM pathogenic interest including TNFRSF17, BMP3, IRS1, and CEPZOS (Figure 3A). Many genes relevant to WM biology, including CXCL13, TP53, CXCR4, MYD88, CDC23, and AKAP1, were associated with BM disease involvement (Figure 3B). Although MYD88 and CXCR4 were both correlated to BM disease involvement ($\rho = -0.50; \rho < .0001$ and $\rho = 0.46; \rho = .0003$, respectively), they were negatively correlated with each other ($\rho = -0.46; P = .0004$). As shown in supplemental Figure 7, this relationship may be influenced by MD88/CXCR4 mutation status. The gene list associated with BM disease involvement included MYD88 and CXCR4, as well as PIK3AP1, PIK3CD, AJUBA, and OS9 (Figure 3C). Based on these findings, CXCL13 was included in an independent serum cytokine profiling study of 86 WM patients that validated the CXCL13 correlation between BM (Spearman $\rho = 0.562; P = 2.192 \times 10^{-8}$) and hemoglobin (Spearman $\rho = -0.567; P = 1.53 \times 10^{-8}$).39

**Analysis of mutant allele burden**

The percentage of reads supporting the mutant allele relative to WT allele was calculated for MYD88, CXCR4, ARID1A, and CD79B in the cross-validation. Filtering genes/predictors by variable importance score resulted in 40, 236, and 143 associated genes for IgM, BM, and hemoglobin, respectively (supplemental Figure 6). Functional analysis was conducted on each gene list and top genes from each group were selected based on biological significance. The 40 genes associated with IgM were predicted to be downstream of IL6 signaling ($P = .0018$) and included many genes of WM pathogenic interest including TNFRSF17, BMP3, IRS1, and CEPZOS (Figure 3A). Many genes relevant to WM biology, including CXCL13, TP53, CXCR4, MYD88, CDC23, and AKAP1, were associated with BM disease involvement (Figure 3B). Although MYD88 and CXCR4 were both correlated to BM disease involvement ($\rho = -0.50; \rho < .0001$ and $\rho = 0.46; \rho = .0003$, respectively), they were negatively correlated with each other ($\rho = -0.46; P = .0004$). As shown in supplemental Figure 7, this relationship may be influenced by MD88/CXCR4 mutation status. The gene list associated with BM disease involvement included MYD88 and CXCR4, as well as PIK3AP1, PIK3CD, AJUBA, and OS9 (Figure 3C). Based on these findings, CXCL13 was included in an independent serum cytokine profiling study of 86 WM patients that validated the CXCL13 correlation between BM (Spearman $\rho = 0.562; P = 2.192 \times 10^{-8}$) and hemoglobin (Spearman $\rho = -0.567; P = 1.53 \times 10^{-8}$).39

**Analysis of mutant allele burden**

The percentage of reads supporting the mutant allele relative to WT allele was calculated for MYD88, CXCR4, ARID1A, and CD79B in the
had mutant allele burdens in excess of 50%, an unexpected mapping efficiency between indels and single-nucleotide variants, MYD88L265P allele burden at the RNA level for the 11 patients with nonsense mutations (Figure 4A). The median percentage of mutant reads was 37.4% (range, 18.4%-61.4%) and 54.5% (range, 18.2%-60.9%) for MYD88 and CXCR4, respectively (P = 0.053) with the mutant allele burden of CXCR4 being greater than MYD88 in 9 of 11 patients (81.8%). Sanger sequencing of paired DNA and cDNA confirmed the mutant allele was underrepresented in the cDNA relative to the DNA for MYD88L265P in many patients whereas the opposite was true for CXCR4.

RNA sequencing data. Although the relevant mutations were observed in the RNA in all cases, many samples with CXCR4WHIM mutations had mutant allele burdens in excess of 50%, an unexpected finding for a typically subclonal variant. As there can be differences in mapping efficiencies between indels and single-nucleotide variants, MYD88L265P allele burden was then compared with the CXCR4WHIM allele burden at the RNA level for the 11 patients with nonsense CXCR4

Figure 5. Summary of key gene expression differences. Gene expression levels for all study samples are shown depicting key differences for WM vs normal B-cell samples as well as WM patient intrapatient differences based on MYD88 and CXCR4 genotype. The heat map represents log_{10} transformed Tpm values with median and range Tpm values listed for each relevant comparator. P values have been adjusted for multiple hypotheses testing using the Benjamini-Hochberg FDR.

VOLUME 128, NUMBER 6 834 HUNTER et al BLOOD, 11 AUGUST 2016
true for CXCR4 mutations (Figure 4B). This reduced expression of MYD88<sup>265P</sup> at the RNA level was also observed in some CXCR4<sup>WT</sup> patients.

**Discussion**

This study represents the first profile of the transcriptional landscape of WM by next-generation sequencing. Among the most striking findings was a >7.8 log-fold upregulation of VDJ recombination related genes including RAG1, RAG2, and DNTT in WM patients samples (Figures 5 and 6). The class switch recombination gene AICDA was not observed at meaningful levels consistent with the lack of immunoglobulin class switching in WM. The overexpression of RAG1 and RAG2 is notable because specific somatic mutation patterns such as deletions in BTG1 and ETV6 are associated with their aberrant expression, and are commonly observed in WM patients. The universal expression of CXCR4 and the upregulation of its ligand CXCL12 are suggestive of autocrine/paracrine signaling. CXCR4 activation has been reported to increase cell adhesion to VCAM1 via VLA4, and VCAM1 was the eighth most upregulated gene relative vs HD samples. Taken together, these findings suggest that VCAM1, CXCR4, and CXCL12 may facilitate the homotypic WM cell clustering observed in the BM of many WM patients. Likewise, IGF1, an inducer of AKT1 survival signaling in WM, was a top hit among the most upregulated genes. IGF1 and its receptor IGF1R may therefore be potential therapeutic targets in WM and warrant further investigation.

Despite previous studies suggesting that WM was derived from MB, we did not observe increased similarity at the transcription level for MB relative PB as shown in Figure 1B, C, and E. WM represents MB, we did not observe increased similarity at the transcription level for MB relative PB as shown in Figure 1B, C, and E. WM represents
upregulation of IRAK3 in CXCR4\textsuperscript{WHIM} patients may be the most proximal explanation for this phenomenon. Although IRAK3 is a negative regulator of the IRAK4/1 signaling cascade, it has been shown that BTK activation downstream of MYD88 occurs independently of IRAK4/1 and is likely unaffected by IRAK3 upregulation.\textsuperscript{4,49} In addition, TLRI7 is upregulated in CXCR4\textsuperscript{WHIM} patients and can use the MYD88/IRAK4/IRAK3 complex to initiate regulatory signaling through MAPK3-dependent activation of NF-κB.\textsuperscript{50}

Compared with the other WM genotypes and healthy donor B cells, WM patients with the MYD88\textsuperscript{L265P}CXCR4\textsuperscript{WT} genotype strongly upregulate a number of surface receptors and signaling molecules including IL17RB, GPER1, WNT5A, and IGFI. IL17RB signaling may provide an additional source of NF-κB activation whereas GPER1 and IGFI activate AKT1 and MAPK signaling in marginal zone lymphoma and WM, respectively.\textsuperscript{61,62} However, WNT5A is a suppressor of B-cell proliferation.\textsuperscript{53} Many of the genes upregulated in the MYD88\textsuperscript{L265P}CXCR4\textsuperscript{WT} patient cohort such as PMAIP1, WNK2, PRDM5, CABLES1, CXXC4, and CDKN1C are tumor suppressors that either promote apoptosis, inhibit cell cycle, or block MAPK signaling.\textsuperscript{54-56} Normalization of MYD88\textsuperscript{L265P} associated tumor suppressor expression in CXCR4-mutated WM patients may provide the clearest explanation for the selective inhibition of MYD88 signaling observed in CXCR4\textsuperscript{WHIM} patients. This yin-and-yang relationship between MYD88 and CXCR4 is further supported by the modulation of allele specific transcription, decreasing and increasing the mutation allele burden of each gene, respectively. CXCR4 and MYD88 total transcription levels were negatively correlated and inversely predictive of BM disease involvement. Although aberrant promoter methylation of WNK2 and PRDM5 was prominent in MYD88\textsuperscript{WT}CXCR4\textsuperscript{WT} samples, it did not explain the low levels of WNK2 and PRDM5 observed in patients with CXCR4\textsuperscript{WHIM} mutations, thereby strengthening the case that these findings are related to CXCR4\textsuperscript{WHIM} signaling rather than unrelated epigenetic events.

The strongest gene markers for patients with MYD88\textsuperscript{L265P}CXCR4\textsuperscript{WHIM} was the upregulation of CXCR7, which like CXCR4 is activated by CXCL12, and TSPAN33 which is upregulated in activated B cells.\textsuperscript{57,58} IL15 was uniquely suppressed in patients with MYD88\textsuperscript{L265P}CXCR4\textsuperscript{WHIM} and warrants functional studies to assess its role in WM microenvironment. Preclinical studies have shown that CXCR4\textsuperscript{WHIM} transduced cell lines are resistant to PIK3CD inhibitors such as idelalisib.\textsuperscript{10,38} Both PIK3R5 and PIK3CG were significantly higher in MYD88\textsuperscript{WT}CXCR4\textsuperscript{WT} samples. PIK3R5 interacts with the β/γ G-protein subunits downstream of G-protein–coupled receptors such as CXCR4 to activate PI3KCG which may explain the observed resistance.\textsuperscript{61} The G-protein and MAPK inhibitors RGS1, RGS2, RGS13, DUSP1, DUSP2, DUSP4, DUSP5, DUSP10, DUSP16, and DUSP22 as well as the MAPK-inducible tumor suppressor ERFII1 were significantly downregulated in patients with MYD88\textsuperscript{L265P}CXCR4\textsuperscript{WHIM} vs MYD88\textsuperscript{L265P}CXCR4\textsuperscript{WT}.\textsuperscript{62} The fact that DUSP2, DUSP4, DUSP5, RGS13, and ERFII1 were all significantly induced by CXCL12 in CXCR4\textsuperscript{WHIM} vs CXCR4\textsuperscript{WT}-transduced lines supports a role for these genes as negative regulators of CXCR4 signaling and may have implications for MYD88 MAPK signaling as well.\textsuperscript{63,64} This is similar to a number of secondary events such as MYBBP1A mutations and deletions of TRAF3, TNFAIP3, and HIVEP2, which are thought to represent loss of negative regulators for NF-κB signaling in WM.\textsuperscript{2,11,65} The mechanism(s) by which negative regulators of CXCR4 signaling are lost in patients with CXCR4\textsuperscript{WHIM} mutations remain to be clarified warranting further investigation.

Of the 1155 genes that were differentially expressed between MYD88\textsuperscript{L265P} and MYD88\textsuperscript{WT} patients, 552 were not observed in the MYD88\textsuperscript{L265P}CXCR4\textsuperscript{WHIM} vs MYD88\textsuperscript{L265P}CXCR4\textsuperscript{WT} gene signature. MYD88\textsuperscript{WT} patient expression was quite heterogeneous indicating pathogenetic diversity in this population. Large-scale genomic exploration will invariably be needed to help clarify the genetic basis for MYD88\textsuperscript{WT} WM disease. Regardless, a clear downregulation of genes associated with NF-κB signaling including IL6, IRAK2, TNFAIP3, NFκB2, TIPARP, PIM1, and PIM2 was observed in these patients. Other genes of interest included the downregulation of CD40, and upregulation of PTBP3, CD86, and CXCR3. Notably, IGFI1, PIK3AP1, and AKT2 were all upregulated, suggesting a reliance on PI3K signaling. The antiapoptotic gene BCL2, known to play an important role in WM, was overexpressed in all patient samples while the proapoptotic BAX was underexpressed, suggesting a mechanism independent of the MYD88 and CXCR4 mutations.\textsuperscript{55-67}

Somatic mutations in ARID1A were associated with increased CXCL13, as well as increased BM infiltration. CXCL13 has not been previously described in WM and was also a strong predictor of BM disease involvement and hemoglobin levels. CXCL13 may also play a role in mast cell recruitment.\textsuperscript{68} Mast cells can be observed admixed with the WM B cells in patient BM histology and are thought to play an important supportive role in the tumor microenvironment of WM.\textsuperscript{69} Deletions of chromosome 6q were associated with over 131 differentially expressed genes. Levels of previously identified target genes such as the NF-κB negative regulator HIVEP2, as well as BCLAF1, FOXO3 and ARID1B were all suppressed in the presence of 6q deletions.\textsuperscript{2} Although only 3 patients had strong familial predisposition to WM, these patients demonstrated significant dysregulation of several cancer-associated genes including RB1, STAT5B, ZNF300, MAPK9, and NFκB1. These studies highlight the pivotal roles of MYD88 and CXCR4 signaling in WM. CXCR4 mutations appear to function primarily as dampeners of negative regulators that are upregulated in response to mutant MYD88 signaling. The upregulation of the ligand, adhesion targets, and CXCR4 itself in all WM patients provides evidence for uniform CXCR4 dysregulation in WM and supports the development of CXCR4 antagonists for WM therapy. The upregulation of PIK3 pathway members and the increased promoter methylation in MYD88\textsuperscript{WT} patients creates a strong rationale for preclinical studies of PIK3 inhibitors and demethylating agents in this population. The upregulation of BCL2 across all WM patient genotypes also supports the development of BCL2 antagonists. Finally, CXCL13 is highly expressed by WM cells, and its expression correlates with important clinical parameters. Further studies to delineate therapeutic targeting of this cytokine in WM are warranted.

**Acknowledgments**

The authors acknowledge the contributions of Yaoyu Wang and John Quackenbush at the Center for Cancer Computational Biology at the Dana-Farber Cancer Institute for RNA Sequencing, and Mathew Temple, Leutz Buon, Terry Haley, and the Dana-Farber Research Computing Center for their invaluable assistance. The authors acknowledge the generous support of the WM patients who provided their samples.

This work was supported by Peter S. Bing, a Translational Research grant from the Leukemia & Lymphoma Society, a grant
References


in Waldenstrom’s macroglobulinemia.


Transcriptome sequencing reveals a profile that corresponds to genomic variants in Waldenström macroglobulinemia

Zachary R. Hunter, Lian Xu, Guang Yang, Nicholas Tsakmaklis, Josephine M. Vos, Xia Liu, Jie Chen, Robert J. Manning, Jiaji G. Chen, Philip Brodsky, Christopher J. Patterson, Joshua Gustine, Toni Dubeau, Jorge J. Castillo, Kenneth C. Anderson, Nikhil M. Munshi and Steven P. Treon