Genomic and transcriptomic profiling reveals distinct molecular subsets associated with outcomes in mantle cell lymphoma

Shuhua Yi, … , Lugui Qiu, Lili Wang


Graphical abstract

Find the latest version:

https://jci.me/153283/pdf
Genomic and Transcriptomic Profiling Reveals Distinct Molecular Subsets Associated with Outcomes in Mantle Cell Lymphoma

Shuhua Yi1*, Yuting Yan1,2*, Meiling Jin2*, Supriyo Bhattacharya3, Yi Wang1, Yiming Wu2, Lu Yang2, Eva Gine4, Guillem Clot4, Lu Chen5, Ying Yu1, Dehui Zou1, Jun Wang1, An T. Phan2, Rui Cui1,6, Fei Li7, Qi Sun1, Qiongli Zhal8, Tingyu Wang1, Zhen Yu1, Lanting Liu1, Wei Liu1, Rui Lyv1, Weiwei Sui1, Wenyang Huang1, Wenjie Xiong1, Huijun Wang1, Chengwen Li1, Zhijian Xiao1, Mu Hao1, Jianxiang Wang1, Tao Cheng1, Silvia Beà4, Alex F. Herrera5,9, Alexey Danilov5,9, Elias Campo4, Vu N. Ngo2, Lugui Qiu1,7#, Lili Wang2,5#

Affiliations
1State Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Institute of Hematology & Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Tianjin, China
2Department of Systems Biology, Beckman Research Institute, City of Hope Comprehensive Cancer Center, Monrovia, CA
3Division of Translational Bioinformatics, Beckman Research Institute, City of Hope Comprehensive Cancer Center, Irwindale, CA
4Lymphoid Neoplasm Program, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Hematology Department, Hospital Clínic; Departament d'Anatomia Patològica, Universitat de Barcelona, Barcelona, Spain
5Toni Stephenson Lymphoma Center, Beckman Research Institute, City of Hope Comprehensive Cancer Center, Duarte, CA
6Department of Hematology, Tianjin First Center Hospital, Tianjin, China
7Department of Hematology, The First Affiliated Hospital of Nanchang University, Institute of Hematology, Academy of Clinical Medicine of Jiangxi Province, Nanchang, Jiangxi Province, China
8Department of Pathology, National Clinical Research Center for Cancer, Key Laboratory of Cancer Prevention and Therapy of Tianjin, Tianjin Medical University Cancer Institute and Hospital, Tianjin, China
9Department of Hematology & Hematopoietic Cell Transplantation, City of Hope Comprehensive Cancer Center, Duarte, CA
☆, # indicate these authors contributed equally to this study.

Correspondence:
Lili Wang, MD, PhD, 1215 S 5th Ave, Monrovia, CA 91016. Phone: 626.218.7016, lilwang@coh.org
Lugui Qiu, MD, 288 Nanjing Road, Tianjin 30020, China. Phone: 086.22.23909172, Email: qiulg@ihcams.ac.cn
Key words: MCL, genetic lesions, clonal evolution
Abstract

Mantle cell lymphoma (MCL) is a phenotypically and genetically heterogeneous malignancy in which the genetic alterations determining clinical behavior are not fully understood. Here we performed a comprehensive whole-exome sequencing analysis of 152 primary samples derived from 134 MCL patients, including longitudinal samples from 16 patients and matched RNA sequencing data from 48 samples. We classified MCL into four robust clusters (C). C1 featured mutated IGHV, CCND1 mutation, amp(11q13) and active BCR signaling. C2 was enriched with del(11q)/ATM mutations and upregulation of NF-κB and DNA repair pathways. C3 was characterized by mutations in SP140, NOTCH1 and NSD2 with downregulation of BCR signaling and MYC targets. C4 harbored del(17p)/TP53 mutations, del(13q) and del(9p), and active MYC pathway and hyperproliferation signatures. Patients in these four clusters had distinct outcomes (5-year OS rates for C1-C4 were 100%, 56.7%, 48.7%, and 14.2%, respectively, p<0.001). We also inferred the temporal order of genetic events and studied clonal evolution of 16 patients before treatment and at progression/relapse. Eleven of these samples showed drastic clonal evolution that was associated with inferior survival while the other samples showed modest or no evolution. Our study thus identifies genetic subsets that clinically define this malignancy and delineates clonal evolution patterns and their impact on clinical outcomes.
Introduction

Mantle cell lymphoma (MCL) is an aggressive subtype of non-Hodgkin’s B cell lymphoma that has a median overall survival of approximately 5 years (1-7). MCL patients can be generally grouped into two types based on their clinical behavior: aggressive conventional MCL (cMCL) and indolent leukemic non-nodal MCL (nnMCL) (2, 5). Several clinical and molecular features are used to distinguish these two types of MCL, including involvement of lymph nodes, expression of SOX11 (2, 5), and somatic hypermutation on the B-cell receptor (BCR) immunoglobulin heavy variable (IGHV) genes (1, 5, 8, 9). Due to significant heterogeneity in the clinical outcome of patients with MCL (3, 4, 10), differentiating patients who will have poor clinical outcomes from patients who will achieve durable response with standard therapies remains a challenge. The MCL international prognostic index (MIPI) (7, 11-14) and tumor Ki67 (11) expression are used to stratify newly diagnosed MCL patients. However, these traditional prognostic markers have not enabled tailored therapeutic strategies for MCL. In the era of novel therapies for MCL (1, 15, 16), better prognostic and predictive models that account for the biologic heterogeneity of the disease are needed to stratify patients.

In the last decade, unbiased massively parallel sequencing of whole exomes (WES) and RNA sequencing (RNA-seq) of MCL have identified recurrent mutations (TP53, ATM, NOTCH1/2, CCND1, HNRNPH1, KMT2D) associated with MCL (6, 17-22) and genetic lesions (del(9p), ARID1A, SMARCA4) that contribute to resistance to chemoimmunotherapy or targeted therapies (21, 23-25). However, several constraints have limited previous analyses. First, use of relatively small cohorts in studies that utilized an unbiased discovery approach (17, 21, 26) has curtailed the ability to define patterns of genetic lesions and their associations with clinical outcomes. In a larger study (25), only 8 genes were assessed, allowing limited evaluation of the prognostic importance of co-occurring genetic alterations. Second, limited availability of matched RNA-seq and WES data impeded connecting the genotype with phenotype. Lastly, lack of longitudinal samples restricted the ability to evaluate clonal evolution of MCL in relation to disease progression to chemoimmunotherapy.

To address these limitations, we performed WES on 152 MCL tumor samples from 134 patients (Table 1). Longitudinally collected samples were available for 16 patients and 48 samples had matched RNA-seq data (Supplemental Figure 1).

Results

Mutated cancer driver genes and mutational signatures in MCL
Our samples were collected from 134 MCL patients (123 newly diagnosed and 11 with relapsed disease, 15 nnMCL and 119 cMCL) with a broad range of clinical characteristics including different MIPI risk as well as IGHV unmutated and mutated (Table 1, Supplemental Table 1-2). The median follow-up time was 31.0 (range 4.5-107.3) months. The 3-year OS rate in the cohort was 69.6% (95 CI 60.1-78.5%).

We detected a median of 29 nonsynonymous mutations in protein-coding sequences per sample (range, 8-72), and a median mutation burden of 1.35 mutations per megabase (Mb), which is similar to that previously reported in MCL and other hematologic cancers (Supplemental Figure 3A-B)(17, 19, 21, 26). We identified 33 recurrently mutated genes (mutated in >5 samples, mutation frequency >3%, Figure 1A), which include known and novel recurrent mutations (LRP1B, PCLO, RYR2, PCDH10, OBSCN, TACC2, FAT3, LRP2, SVEP1, ZFHX4, MPDZ, DCDC1, IKBKB, ARID1A) (Figure 1A, bold).

To determine which of the mutated genes are likely to contribute to lymphomagenesis, we used the CLUMPS algorithm (27) to identify clustering of mutations predicted to have significant impacts on three-dimensional protein structures or interference with its binding partners. Mutations in TP53 and CCND1 were found with significant clustering (P < 0.05) whereas mutations in ATM, SP140 and SMARCA4 showed moderate clustering (Supplemental Figure 4A, P < 0.1). To determine how individual mutation affects the clustering P value, we systematically removed each mutation and recalculated the weighted average proximity (WAP) score and the resulting change in -log10(P value), ∆lpvalue (Figure 1B, Supplemental Figure 4C). A positive ∆lpvalue indicates that the mutation clusters with other mutations in the protein, and that the removal of this mutation adversely affects the CLUMPS score significance. Conversely, a negative ∆lpvalue indicates that the mutation does not cluster significantly with other mutations. In TP53, we found both categories of mutations, with positive and negative ∆lpvalues (Figure 1C). For example, mutations at R248 and I195 show negative ∆lpvalues, while mutations at R273 have positive ∆lpvalues, indicating that these two groups of TP53 mutations may exert their effects in different ways. Two of the mutations with negative ∆lpvalue, R248 and S241 are at the DNA binding interface, suggesting that these mutations may interfere with the DNA recognition by TP53. In contrast, mutations with positive ∆lpvalues, such as R158, V156 and Y205 are clustered within structural domains of TP53, that are distant from the DNA binding interface. These mutations may affect the function of TP53 through different mechanisms compared to the ones that are un-clustered. We further observed that the ∆lpvalue of WAP score for TP53 mutation was significantly different between SOX11 negative and positive patients (Figure 1B, Supplemental Figure 4B). This implies that the SOX11 positive patients tend to have mutations that cluster together in the TP53 structure compared to SOX11 negative patients (Figure 1C).

The hotspot mutations at the C47 and Y44 of CCND1 impacting the WAP score were not in direct contact
with its binding partner CDK4, but in a loop region that packs between two helices and maintains a half helical turn in the loop (Supplemental Figure 4E). The significance was supported by the observation that both mutations can increase CCND1 protein stability and promote ibrutinib resistance in MCL (28).

To delineate the roles of recurrent mutations in MCL biology, we examined genome-wide CRISPR/Cas9 perturbation screen results from DepMap (https://depmap.org/portal/) for leukemia and lymphoma as well as our own CRISPR/Cas9 perturbation screen results in the MCL cell line JeKo-1. Silencing of SP140, SMARCA4, PLCO, TP53, and TRAF2 in JeKo-1 cells conferred a cell growth advantage while knockdown of these genes in other cell lines had modest or no impact on cell growth (Figure 1D, Supplemental Figure 3C), indicating that these genes may act as tumor suppressors in MCL.

Leveraging our WES dataset, we identified four mutation signatures prevalent in MCL using the MutationalPatterns pipeline (Supplemental Figure 5A-B) (29). This includes an age-related signature involving C-to-T transitions at CpG sites; a c-AID signature characterized by increased C>T/G mutations at a known activation-induced cytidine deaminase (AID) hotspot (SBS84), an enzyme essential for somatic hypermutation of germinal center B cells; and signatures 5 and 40, common signatures that were prevalent in most cancers and leukemia/lymphoma, respectively. The c-AID signature mainly comprised clustered mutations (Supplemental Figure 5C). Of note, most of the signatures are contributed by aging-related signature and signature 40 (range: 36.2-100%, median 68.8%, Supplemental Figure 5D).

Copy number alterations in MCL
With this dataset, we identified 20 recurrent SCNAs (Figure 1A, Figure 2, q value ≤ 0.1, frequency ≥ 10%). Of note, tumor-only pipeline generated highly correlated SCNA calls in the 89 paired samples, which were well correlated with fluorescence in situ hybridization results (ROC, p<0.001 for del(17p), del(13q), del(11q)) (Supplemental Figures 6A-B and 7). In addition to previously reported SCNAs linked to OS (del(9p), del(17p), del(13q) and del(8p23.3)) (18, 30), we also identified driver SCNAs including (del(15q11-13) and amp(11q13.3) (Figure 1A).

To determine how SCNAs affect gene expression, we performed an integrative analysis in samples with WES and RNA-seq data (n=48). We focused on identifying genes that showed significant changes within the deleted or amplified regions by comparing samples with or without the lesions. Pathways that were significantly affected by amplification included RNA catabolic and translation pathway (EIF4G1, RPL4, DDX6, PRL15) and MYC pathway (MYC, NME1). Pathways that were perturbed by deletion included DNA repair and cell cycle (ATM, CDKN1B, POT1) and RNA splicing (HNRNPK, NCBP1, SRSF1) (Supplemental Figure 6C-D).
Our WES dataset revealed significant relationships between several genetic alterations (Figure 2B). In addition to known co-occurrence between TP53 mutation and del(17p), ATM mutation and del(11q), del(9p21.3) and del(17p) (Supplemental Figure 8A-C), we also observed a high co-occurrence of del(9p21.3) with del(8p) ($q<0.001$) and del(13q) ($q=0.004$) (Figure 2C). Moreover, we found low co-occurrence of genetic alterations such as mutations in TP53 and ATM or del(11q) (Supplemental Figure 8B, $q<0.05$), indicating tumor cells harboring these events may originate from a different genetic trajectory.

**Association of genetic features with clinical outcomes**

We examined associations between genetic alterations and key MCL features. Overall, we observed a high number of SCNAs associated with unmutated IGHV status and SOX11 expression (Supplemental Figure 9A-B). The c-AID mutation signature was strongly associated with mutated IGHV status while the aging signature correlated with unmutated IGHV and SOX11 expression (Supplemental Figure 9C-D). Moreover, the number of SCNAs could predict clinical outcomes (Supplemental Figure 9E-F).

We examined the prognostic significance of somatic mutations. Mutations in SP140, SMARCA4, TRAF2, and PCDH10 were predictive of poor PFS (Figure 3A-B). SP140 mutations occurred at 8% frequency in our cohort and 9 out of 11 mutations were frameshift and nonsense mutations that resulted in a truncated form of SP140 (Figure 3A), highly suggestive of loss-of-function mutations. We further identified 10 samples (7.5%) harboring SP140 deletion (loss of 2q36.3-37.1), all of which showed downregulation of SP140 expression compared with samples lacking the deletion (Figure 3C-D). Mutation or deletion of SP140 was predictive for shorter PFS and OS and associated with SOX11 expression, suggesting this gene may be a potential tumor suppressor in MCL (Supplemental Figures 9G and 10A). Consistent with previous publication (25, 31), the presence of TP53 or NOTCH1 aberrations was associated with shorter PFS (Supplemental Figure 10B-C, Figure 3E). Of note, TP53, NOTCH1, PCDH10 mutations as well as SP140 mutation/deletion retained significance for PFS and OS when MIPI risk and IGHV mutation status were added (Figure 3E).

We then assessed the contribution of recurrent SCNAs to MCL progression. Consistent with previous observations (30, 32-35), loss of 17p13.3 (35%) and 9p21.3 (40%) predicted inferior PFS and OS (Supplemental Figure 11) and it remained significant in the multivariate analysis (Figure 3E). Recurrent SCNAs in this cohort including del(12p13.31), del(13q14.2), del(15q11-13), del(8p23.3), amp(13q31) were also associated with shortened PFS and OS (Supplemental Figure 11), but not significant in the multivariable analysis (Figure 3E). Del(9p21.3), del(1p21.1), del(11q22.3), del(13q14.2), and del(6q25.3) were associated with unmutated IGHV and SOX11 expression ($P<0.05$, Supplemental Figure 9G). Most
of the genetic alterations also remained significant among patients who received the cytarabine-based regimen (Supplemental Figure 10D-E).

We further gained insights into the contribution of the most frequent deletion in MCL, chromosome 9 deletion. We first examined genes that may render cell growth advantage through analysis of our CRISPR/Cas9 perturbation screen in JeKo-1 cells and found many critical tumor suppressors were located on chr9, including CDKN2A, SMARCA2, FBXO10, and TOR1B (Figure 4A, Z score ≥ 1). We next classified the WES samples with del(9) into three groups based on the deleted region: 9p-, 9q-, or large region (both 9p/9q) (Figure 4A). Del(9p) was more frequent (23/54) than del(9q) (14/54) or both (17/54, Figure 4A). These deletions also influenced gene expression as reflected by our unsupervised RNA-seq analysis of MCL samples containing and lacking these deleted regions (Figure 4B). Consistent with previous study(36), we found 24 downregulated genes on chr9 were significantly associated with PFS and OS (Figure 4C, HR<1, P<0.05), and all three types of deleted regions were predictive of inferior clinical outcomes (Figure 4D, log rank paired comparison, P<0.05), irrespective of the size and location of the deleted region.

Coordinate genetic signatures classify MCL into four subsets that have unique gene expression patterns and distinct clinical behavior

To identify genetic subtypes based on shared genetic features in MCL, we applied an NMF consensus clustering algorithm (37, 38) to 35 recurrent genetic alterations and discovered 4 robust subsets of tumors characterized by distinct genetic signatures (Supplemental Figure 12A, Figure 5A). The four subtypes differed significantly in PFS and OS (Figure 5B, P<0.001). Patients with the C1 subtype had a more favorable outcome than those with C2, C3, and C4 subtypes. Median PFS was not reached for C1, 41.2 months for C2, 30.7 months for C3, and 16.1 months for C4 (log rank, P<0.001). 5-year OS rates for C1-C4 were 100%, 56.7%, 48.7%, and 14.2%, respectively. Differences in survival of patients of the 4 subtypes also remained significant among patients who received the cytarabine-based regimen (Supplemental Figure 12B). Moreover, molecular cluster was an independent risk factor when MIPI risk and IGHV mutation status were included in the multivariate analysis, however, this is mainly driven by C4 and C1 (Supplemental Figure 12C, C4 vs. C1, P=0.017).

To determine the robustness of these genetic clusters, we assessed whether these genetic alterations can stratify MCL patients using a published genetically well-annotated MCL cohort for validation(26) (Barcelona cohort, Figure 6). Projection of cluster features classified patients into 4 distinct clusters with C1 has favorable clinical course and C2/C3 fell in between C1 and C4. There is significant statistical
difference among clusters \((P=0.001)\), in which C1 vs. C4 \((P<0.001)\), C2 vs. C4 \((P<0.035)\), and C3 vs. C4 \((P=0.014)\) reached significance in pairwise test.

To explore phenotypic differences among the MCL genetic subtypes, we performed an integrative analysis using matched RNA-seq data \((n=48)\) across the 4 subsets (Figure 7, C1-C4, n=12, 11, 16, and 9, respectively). We first assessed whether the recurrent mutated genes identified from WES were expressed at the RNA level and discovered most of these mutations were highly expressed (Supplemental Figure 13A-B). Likewise, the frequent SCNAs also resulted in significant dysregulated gene expression (Figure 7A), which we further validated by RT-PCR analysis of MCL samples containing and lacking the SCNAs (Supplemental Figure 13C). Our analysis revealed that each genetic subset has a unique gene expression pattern (Figure 7B, Supplemental Figure 14A). Consistent with the differing cellular origins for the two types of MCL \((1, 5)\), C1 was enriched for gene expression signatures of memory B cells and C2, C3, and C4 appeared to have a signature of CCR6 negative light zone B cells or naïve B cells (Figure 7B-C, Supplemental Figure 14A). We further tested a previous reported 16-gene signature that unbiased distinguished cMCL and nnMCL \((39)\) and found 35 out of 36 C2-C4 patients was classified as cMCL while 9 out of 12 C1 patients was in nnMCL (Supplemental Figure 14B).

**Cluster 1.** (16% of samples). Most of the C1 samples were IGHV mutated, and featured with mutant CCND1, TP53 and amp(11q13). Most of the TP53 mutations in C1 were with negative WAP score (Supplemental Figure 4B). Patients with and without TP53 mutations had similar overall survival \((P=0.470, \text{Supplemental Figure 15A-B})\). C1 had the lowest SOX11 expression (Supplemental Figure 15C). Phenotypically, C1 was enriched with a memory B cell phenotype and active BCR signaling (Figure 5A, Figure 7B-C, Supplemental Figure 14A). We observed enrichment of BCR signaling in the Barcelona cohort although it is insignificant due to the small number of available microarray gene expression data (Figure 6D).

**Cluster 2.** (23% of samples). Of 31 samples in C2, 28 harbored del(11q) (minimal deleted region contains ATM) while 19 of these 28 samples had a co-occurring ATM mutation. Consistent with these genetic lesions, genes involved in DNA replication, DNA repair, and hyper proliferation were all upregulated (Figure 7B). Expression of genes involved in TNFα signaling via the NFκb pathway, interferon α and γ response was significantly enriched in both discovery and validation cohorts (Figure 6D, 7B).

**Cluster 3.** (32% of samples). Besides enriched NOTCH1 mutations, C3 also harbored mutations in NSD2 (WHSC1), KMT2D, SP140 as well as amp(13q) and del(6q). In contrast to C2, we observed significant downregulation of genes implicated in TNFα signaling via the NFκb pathway and interferon γ response.
but with activated NOTCH signaling. Additionally, BCR signaling, MYC targets, and IL2 STAT-5 signaling were all downregulated in C3 in both discovery and validation cohorts (Figure 6D, 7B-C).

Cluster 4. (28% of samples). This subtype harbored the most SCNAs, including deletions (del(17p), del(13q), del(9p)) and mutations (TP53 and TRAF2) (Figure 5A). Mutations in TP53 was enriched for positive WAP score and predict for poor survival in C4 (Supplemental Figure 15A). Phenotypically, C4 had gene signatures of active MYC pathway, hyperproliferation, and light zone CCR6 negative B cells in both discovery and validation cohorts (Figure 6D, 7B-C). C4 was associated with the highest incidence of blastoid or pleomorphic MCL (25.0%, \( P=0.016 \)) and SCNAs \( (P<0.001, \text{Supplemental Figure 15D-E}) \) but had the lowest contribution to the clustered cAID mutation signature \( (P<0.001, \text{Supplemental Figure 15F}) \). Consistent with this, C4 had the worst clinical outcome with median PFS and OS of 16.1 and 30.0 months, respectively.

Temporal ordering of genetic events and clonal evolution during progression of MCL

To understand Intratumoral heterogeneity and identity the relationship of clonal and subclonal genetic events, we used the ABSOLUTE algorithm\( (40) \) to determine CCF for each of the genetic lesions from our 134 patients. We classified a mutation or SCNA as clonal when the CCF was \( >0.9 \) and subclonal otherwise\( (41-43) \). In total, we identified 516 clonal and 173 subclonal events. Del(11q22.3), del(9p21.3), and ATM mutations tended to be clonal events while mutations in NSD2, PCLO, KMT2C, LRP1B were more likely to be subclonal events (Figure 8A, \( P<0.05 \)).

We further inferred temporal relationships between pairs of genetic events. We first identified instances in which a clonal event was found together with a subclonal event within the same sample, as these pairs reflected the acquisition of one lesion (clonal) followed by another (subclonal). We obtained 22 clonal and subclonal pairs and constructed a temporal map of the evolutionary trajectories of MCL based on the connections (Figure 8B). Both mutations and SCNAs can be the early events (all started with IGH-CCND1 translocation) with six points of departure involving mutated ATM, CCND1, del(1p) del(11q), amp(8q), and del(9q) (Figure 8B). The number of clonal events, but not of subclonal events, was associated with the clinical outcomes (PFS and OS, \( P<0.001 \), Figure 8C, Supplemental Figure 16C-D), highlighting the initiating genetic events and complex genetics driving the clinical outcomes.

To assess clonal evolution in relation to disease progression, we analyzed CCFs for each alteration in 33 longitudinally collected samples from 16 patients (Supplemental Figure 16A) and used PhylogicNDT to cluster dynamic changes and construct phylogeny tree over the timepoints (Figure 9A-C, Supplemental Figure 17)(44). We observed three patterns of tumor evolution upon treatment: 1) no clonal evolution, no change in number of clones, CCF change\(<0.2 \) (n=1 pair); 2) modest clonal evolution, 0.2\( \leq \text{CCF} \)
change ≤ 0.5 (n=4 pairs); 3) drastic clonal evolution, CCF change > 0.5 (n=11 pairs) (Figure 9D). Although the time intervals between collection of first and second samples were essentially identical between drastic evolution and modest or no evolution (Figure 9E, 30.0 vs. 28.1 months, P=0.861), drastic evolution had higher number of clusters and was significantly associated with poor survival (Figure 9E-F, median survival from second sampling 17.1 months vs. not reached, P=0.023), revealing a strong association between clonal evolution and increased disease aggressiveness. Five out of 11 patients whose samples showed drastic evolution (4 cases from C1 or C3 to C4, one case from C1 to C2) also had a cluster change, while all 5 patients whose samples showed modest or no evolution retained the same cluster status (Supplemental Figure 17). Patients whose samples with cluster changes had poor survival after relapse even though their relapse interval appears to be longer (Supplemental Figure 16E-F), highlighting genetic heterogeneity drives the progression of disease.

Discussion

In the past decade, numerous studies have profiled genome-wide genetic alterations, gene expression, and epigenomic changes in MCL (17, 19, 21, 23, 24, 26, 31, 45-48). These studies not only generated insights on the molecular features (2, 30) and mechanisms of pathogenesis (21, 26), drug resistance (23, 24), cellular origin of subsets of MCL (26, 48) but also revealed vast genetic complexity and phenotypic heterogeneity present within MCL, which has become a barrier in connecting genotype with disease phenotype in MCL. Here, starting with a large WES dataset along with matched transcriptome data, we classified MCL into four clusters based on shared genetic lesions and determined their gene expression signatures as well as associations with clinical outcomes. We further studied clonal evolution patterns prevalent in MCL and inferred the order of genetic lesions along the development of MCL.

Our cluster analyses have a few implications. First, C1 is highly enriched for nnMCL and C2-4 are mostly cMCL. Whereas cMCL may present with similar clinical symptoms, they fell into three distinct genetic subsets, which all accompanied by coordinated dysregulated cellular pathways. This analysis provides clues for future biomarker-driven clinical trial design that align particular treatments (e.g. Bruton’s tyrosine kinase inhibitors) with patients most likely to benefit (e.g. downregulation of BCR pathway signaling as observed in C3 or C4). Based on “goldilocks” model of BCR singling and B cell survival, which states that B cell survival is dependent on the tuning of BCR signaling such that it is neither overly strong nor overly weak (49). Within this context, cells with inherent increased BCR signaling would be anticipated to be less sensitive to a BTK inhibitor like ibrutinib, as these cells are away from the minimum threshold of BCR signaling needed for cell survival. In fact, MCL cell lines including JeKo and Mino, which carry complex copy number variations and mutations (likely to be C3-4) (https://depmap.org/portal/), are sensitive to BTK inhibitor while JVM2 (genetic feature similar to C1) is insensitive to BTK treatment (50). Second, our
results emphasize the influence of distinct genetic features on the clinical outcomes. Despite the different treatment regimens and patient population (Chinese and European descent) between our discovery and validation cohorts, all MCLs fell into 4 distinct clusters. These molecular clusters open a door to precision medicine as they can serve as steppingstone between genetic discovery and its application to clinical practice. Last, selection of treatment for MCL based on individual genetic alteration may not optimal because clinical response is determined by a cluster of genetic factors. A good example illustrated this scenario is TP53 mutations as mutant TP53 was associated with inferior clinical courses from both previous reports (51, 52) and our analysis (Supplemental Figure 10B). In particular, C1 and C4 all harbored TP53 mutations (36% and 63%), however, the association of mutant TP53 and clinical outcome depends on the co-occurring genetic events (Figure 5A, Supplemental Figure 15A) and mutation sites (Figure 1B-C, Supplemental Figure 4B). We observed that TP53 mutations lose its adverse prognostic significance in C1 patients, which is validated by Barcelona cohort (Supplemental Figure 15B). This led us to consider whether the mutation sites may influence the function of TP53 protein. Based on our CLUMP analysis, some of the mutations (e.g. R273, positive \( \Delta lpvalue \), enriched in SOX11 positive samples, Figure 1B) cluster together in the protein structure while others (e.g. R248, negative \( \Delta lpvalue \), more enriched in SOX11 negative samples, Supplemental Figure 4B) do not cluster that well. In supporting of our notion from CLUMP analysis, two recent publications reported that both R248 and R273 act in a dominant negative manner but have different level of impact on the function of TP53 in myeloid leukemia (53, 54), adding an extra layer of complexity for the genetic cluster. Altogether, it is anticipated that integrated characterization of changes in MCL genetic clusters and gene expression following treatment with differing therapeutic interventions will further improve the design for precision medicine in MCL.

The genetic heterogeneity in MCL also has impact on clinical outcomes and disease trajectory. As previously reported (26), we also confirmed that the high number of CNAs is associated with inferior OS (Supplemental figure 9F). CNAs tend to be clonal events and the higher number of clonal driver events is predictive of poor survival. It appears that MCL may originate from several different genetic traits, each arising from one or a combination of genetic lesions. Each trait has different intermediate and later genetic events, suggesting a stepwise acquisition of traits (Figure 8B). In nearly all MCL, t(11;14) is a foundation event although it by itself does not lead to MCL (55). We postulate that the second hit could be a genetic trait-starting event such as mutations in \( ATN \) and \( CCND1 \), del(11q) and del(9p) (Figure 8B). An example supporting this idea is B-cell specific inactivation of \( ATN \) (one of the traits in our study) synergizes with ectopic Cyclin D1 expression to promote pre-germinal center lymphoma in mice (56).
Several studies reported the clonal evolution in MCL with the implication that heterogeneous genetic alterations associated with MCL relapse (8, 23, 47). Our study revealed branched evolution is a common feature upon chemotherapy and predictive of clinical outcomes, suggesting that intratumor heterogeneity forms the fuel for relapse and drug resistance. Although mutant TP53/del(17p) was reported to be associated with disease relapse (25), we only observed frequency of del(9p) and amp(3q)) arise in response to therapies (>20% CCF changes in 50% samples) (Supplemental Figure 16B). Large cohort of sample analysis are expected to fully understand the genetic events and role of clonal evolution in driving MCL.

In summary, this integrative analysis provides a framework to assess unappreciated genetic heterogeneity in the clinically defined subtypes of MCLs and forms the basis for designing precision therapies for aggressive MCL with genetic factors and oncogenic pathways as tractable targets.

Methods

Samples and genomics studies

Diagnostic biopsy and/or blood samples representing 152 MCL tumors were obtained from 134 MCL patients. Ninety-five patients received standard high dose cytarabine-based aggressive regimen (Supplemental Figure 1B) while others received non-aggressive treatment (Supplemental Methods). Tumor cells were collected from bone marrow, blood, and lymph nodes (125 cryopreserved and 27 FFPE), with 102 (67.1%) had matched germline tissue (Supplemental Table 1). Thirty-three were longitudinal tumor samples collected from 16 patients at diagnosis (pre-treatment), at progression, or at relapse following treatment (Supplemental Table 2).

WES libraries were prepared using Agilent SureSelect Human All ExonV6 (Agilent Technologies) and sequenced on Hiseq 4000 platform (Illumina). Raw reads were aligned to the human reference genome (GRCh37/hg19) using Burrows-Wheeler Aligner(57). Somatic single nucleotide variations (SNVs) and somatic copy number alterations (SCNAs) were called using GATK best practice somatic mutation and somatic copy number variant discovery pipelines (58, 59), respectively. A tumor-only pipeline was used by filtering a panel of normal (16196 normal samples, Supplemental Methods)(60) from the GATK4 pipeline results for samples without matched normal tissue, which yields comparable mutation calls in paired samples (Supplemental Figure 2). MutationalPattern (29) was used to determine de novo mutation signatures. CLUMPS (clustering of mutations in protein structures) method (27) was used to assess the significance of mutational clustering in a given three-dimensional structure. The details of the calculation of the weighted average proximity (WAP) score were described previously (27). The ABSOLUTE algorithm was used to calculate the tumor purity, ploidy, and cancer cell fraction (CCF) for SNV and SCNA(40). Statistical methods were adapted to infer the order of genetic alterations (43). Phylogenetic
analysis was performed on longitudinally collected samples using the PhylogicNDT package(61). Detailed methods are available in the supplementary material.

**CRISPR library screen**

The genome-wide CRISPR library screen was carried out using the Human GeCKO v2 Library, 2-plasmid system (a gift from Feng Zhang; Addgene #100000049) and following the protocol as described(62). Briefly, the library, which contains 122,417 unique single-guide (sg) RNAs targeting the human genome with 6 sgRNAs per gene. The entire library together with helper plasmids pMD2.g and psPAX2 were then transfected into HEK 293T cells and lentiviral supernatants were collected after two days followed by spin infection at 1200 xg, in two replicates, of doxycycline-inducible Cas9-expressing JeKo-1 cells for 1 h in the presence of 8 μg/ml polybrene. Transduced cells were selected by puromycin for 3 days and doxycycline (1μg/ml) was added to induce Cas9 expression followed by culturing for additional 14 days.

Genomic DNA was harvested on days 0 (day 3 in puromycin) and 14 and subjected to high-throughput sequencing to determine sgRNA abundance. MAGeCK(63) software was used as to quantify sgRNA depletion or enrichment.

**Consensus clustering of genetic alterations**

All recurrent mutated genes (frequency ≥ 5%), IGHV mutational status, and significant regions of SCNAs (GISTIC2.0, q value ≤ 0.1 and frequency ≥ 10%) were assembled into a gene matrix and NMF consensus clustering was used to identify genetic clusters as previously reported(64). Briefly, all genetic lesions based on the following rules: non-silent mutations, 1; IGHV mutation, 2; low-level copy number (CN) deletion (1.0 ≤ CN ≤ 1.7 copies), 1; high-level deletion (CN < 1.0 copies), 2; low-level amplification (2.3 ≤ CN ≤ 3.7 copies), 1; high-level amplification (CN > 3.7 copies), 2. The NMF consensus clustering algorithm was used to assign samples into different clusters. Both cophenetic coefficient and silhouette values for K=2 to K=10 were calculated to determine the best solution as shown in Supplemental Figure 12A (K=4). Fisher’s exact test was used to identify markers for each cluster by testing whether the frequency in a cluster to another cluster was higher than presence or absence of the variants. P values for multiple comparisons were adjusted using the Benjamini-Hochberg correction. Genetic alterations with q value < 0.1 were defined as markers. The main cluster algorithm code was downloaded from GitHub ([https://github.com/broadinstitute/DLBCL_Nat_Med_April_2018](https://github.com/broadinstitute/DLBCL_Nat_Med_April_2018))(64). The results were visualized as a heatmap using R package ComplexHeatmap 2.4.2(65). In the Barcelona cohort, we adopted single-cell projective non-negative matrix factorization (scPNMF)(66) method to project features extracted from discovery cohort. Parameter “-K 15, method = “KL”” was used and samples were assigned into a nearest cluster of from discovery cohort based on UMAP. 33 samples with matched gene expression profiling data available are used to validate our expression features in different clusters.
Integrative genomics and transcriptomics pathway analysis

RNA-seq libraries were generated with NEBNext UltraTM RNA Library Prep Kit (New England Bio) and sequenced on Hiseq platform (Illumina). Raw reads were aligned to the human reference genome (GRCh38/hg38) using STAR (67) and expression levels of mRNAs were normalized to transcript per million (TPM). To directly compare pathway expression for each cluster, the log2-transformed TPM values for all genes in the gene set were averaged to provide a signature value for each sample and then the value for samples assigned to each cluster was calculated as the cluster average expression of the signature. These values were linearly transformed and F test used to compare each cluster.

Statistics

Survival curves were estimated using the Kaplan-Meier method and log-rank test was used to assess statistical significance for progression free survival (PFS) and overall survival (OS) between cohorts. Multivariate Cox regression analysis was used to assess the independent prognostic impact from MIPI risk, IGHV mutational status, and individual genetic factors for outcomes in the MCL cohort. Student’s t test or Mann-Whitney U tests were used to evaluate differences between continuous variables.

Study approval

All samples were obtained from MCL patients. Written informed consent was obtained from all participants, in accordance with the Declaration of Helsinki and approved by the Institute of Hematology and Blood Disease Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College Ethics Committee.
Author contributions


Acknowledgements

We are grateful to Drs. Catherine Wu, Donna Neuberg (Dana-Farber Cancer Institute) and Pavan Bachireddy (MD Anderson Cancer Center) for constructive and valuable discussions. We acknowledge Drs. Yanan Cao (Ruijin Hospital, China) and Zhaohui Gu (City of Hope) for germline mutation annotations. We thank Dr. Chip Stewart (Broad Institute) for computational analysis of clonal evolution. This work was supported by grants from the National Nature Science Foundation of China (81970187, 82170193, 81920108006 and 81900203), Chinese Academy of Medical Sciences Innovation Fund for Medical Sciences (2019-I2M-2-009, 2016-I2M-3-013, 2017-I2M-3-018), National Science and Technology Major Project from China (2017ZX09304024). This work was supported by Startup package from City of Hope (L.W.), grants from the National Institutes of Health (NCI) R01CA240910 and R01CA21623 (to L.W.), R01CA244576 (to A.V.D.). EC is funded by the Spanish Ministerio de Ciencia e Innovación (Grant RTI2018-094274-B-I00, National Institutes of Health 1P01CA229100 and Generalitat de Catalunya Suport Grups de Recerca AGAUR (Grant 2017-SGR-1142), E.C. is an Academia Researcher of the “Institució Catalana de Recerca i Estudis Avançats” (ICREA) of the Generalitat de Catalunya. SB is funded by Fondo de Investigaciones Sanitarias, Instituto de Salud Carlos III and European Regional Development Fund “Una manera de hacer Europa” (grant number PI17/01061 to SB) and Generalitat de Catalunya Suport Grups de Recerca AGAUR 2017-SGR-709. A.F.H is supported by the Emmet and Toni Stephenson Leukemia and Lymphoma Society Scholar Award, and the Lymphoma Research Foundation Larry and Denise Mason Clinical Investigator Career Development Award. We thank Dr. Keely Walker from City of Hope for manuscript edits.

Conflict-of-interest disclosure

A.F.H. reports Research funding from BMS, Merck, Genentech, Inc./F. Hoffmann-La Roche Ltd, Gilead Sciences, Seattle Genetics, AstraZeneca, ADC Therapeutics. Consultancy for BMS, Merck, Genentech,
Inc./F. Hoffmann-La Roche Ltd, Kite Pharma/Gilead, Seattle Genetics, Karyopharm, Takeda, Tubulis, AstraZeneca.
Reference


Figure 1. Recurrent somatic genetic alterations and mutation signatures in MCL. A. Recurrent somatic mutations and copy number alterations (rows) identified following whole exome sequencing of 134 primary samples (columns) obtained from patients with newly diagnosed (green) and relapsed (red) MCL. Samples were annotated for prior treatment, MIPI risk, IGHV status, and Sox11 expression level when collected. Left: Blue labels, recurrent copy number deletion; red labels, recurrent copy number amplification; black labels, somatic mutations. Bold labels, novel copy number alterations/mutations. Right: percentage of samples mutated. Top: total number of genetic alterations across the cohort. B. Contributions of individual mutations to the collective WAP score of TP53. The changes in WAP score $p$ value due to removal of individual mutations are plotted as function of residue number. The radius of the circles around each point in the graphs represent the number of patients with that mutation. The color indicates SOX11 expression. C. TP53 dimer bound to DNA fragment, PDB ID: 3IGK. One of the monomers is shown in yellow, the other in grey. DNA is shown as in orange. The mutations observed in SOX11+ and SOX11- patients are colored in magenta and green, respectively. D. $\beta$ scores from genome-wide of CRISPR/Cas9 screens of JeKo-1 of genes identified as having recurrent mutations.
Figure 2. Recurrent CNAs, co-occurring genetic events, and clinical association. A. Significant copy number amplifications (left, red) and deletions (right, blue). Left sides of the mirror plots show the incidences of significant focal CNA events. Right sides of the mirror plots show q values for each region.
Genes located in the peak of relevant cytobands are listed. B. Pairwise associations between recurrent genetic alterations found in the 134 MCL samples. Low and high co-occurrence are shown in blue and red, respectively. Intensity of the color reflects the odds ratio. Statistically significant association as determined by q value is marked by asterisks. C. The number of samples with co-occurrence of the indicated genetic events in the cohort 134 MCL samples. q indicates the significance of Fisher's exact test.
Figure 3. Associations of somatic mutations with clinical outcomes. A, Lollipop diagrams of selected putative driver genes showing mutation subtype, position, and frequency. Bottom: Y axis indicates the number of identified mutations in the COSMIC database. B, Kaplan–Meier plots (with log-rank P values) of PFS and OS associated with presence and absence of selected mutations. C, Samples with SP140 mutations or deletions did not overlap in the cohort. D, Deletion of SP140 affected its gene expression. SP140 expression TPM value was extracted and plotted from MCL samples with SP140
deletion, mutation, or wild-type. E. Forest plots of the multivariate analysis of MIPI risk groups and individual genetic factors for PFS and OS in our MCL cohort.
Figure 4

A

B

C

D

Gene expression associated to PFS and OS (HR<1, P<0.05)

D

Percent survival

Without chr9 deletion

Mainly 9q deletion

Mainly 9p deletion

Chr9 large region deletion

P<0.001

P<0.001

677
Figure 4. Deletion of chromosome 9 was associated with poor survival. A. Chromosome 9 deletion in samples from our cohort. Top: blue line indicates percentage of MCL samples with chromosome 9 deletion at the location. Known tumor suppressors and oncogenes present on chromosome 9 are color-coded based on their Z score in the CRISPR/Cas9 screen in JeKo-1 cells. Bottom: Deletions in 9p (purple), 9q (blue), or large regions (dark red) in samples from our cohort. Homozygous minimal 9p deletions are marked in red. Cancer cell fraction (Supplemental Methods) of chromosome 9 deletion is shown in gray scale. B. Unsupervised clustering analysis of gene expression in chromosome 9 distinguishes MCL samples with deletions in different region. C. Volcano plot of genes on chromosome 9 that are differentially expressed between MCL samples that have and that do not have chromosome 9 deletions. Downregulated genes that were significantly associated with shorter PFS and OS are indicated in red (Cox regression hazard ratio (HR) <1, \( P<0.05 \)). D. Kaplan-Meier plots of PFS and OS according to type of chromosome 9 deletion.
Figure 5. Coordinate genetic signatures group MCL into four clusters associated with clinical outcome. A. Non-negative matrix factorization (NMF) consensus clustering was performed using all somatic mutations and SCNAs in the 134 MCL samples (columns). Clusters 1-4 are shown with their associated landmark genetic alterations (boxed for each cluster). Left bar graph shows the correlation of genetic alterations associated with each cluster (q value, Fisher's exact test). Non-synonymous
mutations, black; low-level deletion (1.0 \leq \text{copy number} [\text{CN}] \leq 1.7 \text{ copies}), light blue; high-level deletion (\text{CN} \leq 1.0 \text{ copies}), dark blue; low-level amplification (3.7 \geq \text{CN} \geq 2.3 \text{ copies}), orange; high-level amplification (\text{CN} \geq 3.7 \text{ copies}), red. Header shows cluster association (C1, black; C2, green; C3, blue; C4, red), clinical group (conventional MCL, yellow green; indolent MCL, light green), Sox11 expression (negative, green; positive, brown), MIPI risk (high risk, dark pink; intermediate risk, median pink; low risk, light pink), pathology status (blastoid or pleomorphic, crimson; classic, bright lilac), and treatment regimen (Standard cytarabine-based aggressive regimen, dark blue; other regimen, light blue). B-C. Kaplan-Meier plots of PFS and OS of patients grouped into the four clusters. Star indicates $P<0.05$ by log-rank test.
Figure 6. Molecular cluster and gene expression signature validated in Barcelona cohort. A. Sample inclusion description in the validation cohort. B. Projective non-negative matrix factorization consensus clustering was performed using genetic alterations identified from our discovery cohort (Figure...
Clusters 1-4 are shown with their associated landmark genetic alterations (boxed for each cluster). Header shows cluster association (C1, black; C2, green; C3, blue; C4, red). C. Kaplan-Meier plots of OS of patients grouped into the four clusters. P indicates log-rank test. Number indicates samples included in each cluster. D. Integration of genetic and transcriptomic analyses identified gene expression signatures for each genetic subset. The heatmap was generated using normalized enrichment score (NES). Asterisks indicate the significance level of the enrichment.
Figure 7. Integrative analysis of genome and transcriptome reveals a unique gene expression signature in each cluster. **A.** Recurrent somatic mutations, SCNA$s$, and gene expression associated with SCNA$s$. Top panel: x-axis shows the chromosome location of recurrent somatic mutations, y-axis indicates the frequency of mutations detected in our MCL cohort (n=134). Genes shown in purple have a mutation incidence > 5%. Bottom panel: left y-axis indicates proportions of copy number deletion (DEL) and amplification (AMP). Each dot represents a gene at its chromosome location. Genes with absolute copy number < 1.7 or > 2.3 were defined as deleted or amplified, respectively. Genes with a deletion incidence > 10% are shown in blue and genes with an amplification incidence > 10% as red. **B.** Integration of genetic and transcriptomic analyses identified unique gene expression signatures for each genetic subset. The Hallmark and KEGG gene sets and Signature database were used for Gene Set Enrichment Analysis. The heatmap was generated using normalized enrichment score (NES). Red indicates an upregulated pathway in the cluster compared to other clusters, while blue indicates a downregulated pathway. Asterisks indicate the significance level of the enrichment. **C.** Proposed model for the four MCL subgroups. Clusters 1-4 were all associated with distinct genetic events and gene expression signatures. C1 had indolent disease and carried memory B cell gene signature. C2-4 had more aggressive clinical courses and expressed CCR6 negative light zone or naïve B cells gene signature.
Figure 8. Clonal driver events associated with clinical outcomes. A. Cancer cell fraction (CCF) values for each sample affected by a recurrent genetic alteration across all 134 samples. Median CCF values are shown (top, bars represent the median and interquartile range for each genetic alteration). Alterations with a CCF value > 0.9 were defined as a clonal event. The cumulative proportion of a
recurrent genetic alteration found as clonal (blue) or subclonal (red) in the cohort is shown in bottom plot.

B. Computational inference of temporal order of genetic alterations in MCL. Arrows indicate when paired clonal and subclonal genetic alterations were found in the same sample. Dashed lines indicate the temporal order was found in ≥ 3 samples; solid lines that the temporal order was found in ≥ 5 samples.

C. Kaplan-Meier plot of PFS according to the number of clonal driver events.
Figure 9

(A) Clusters 1-5

(B) Clusters 2 and 3

(C) MCL34_T Treatment-Naive R maintenance for 2 years MCL34_R Relapse

(D) Extreme evolution

(E) Sampling interval (months)

(F) Survival from first sampling and second sampling

746
Figure 9. Clonal evolution pattern in MCL and its association with clinical outcome. A-C. Depiction of tumor clonal evolution from diagnosis to relapse in a representative patient (MCL34). A. Dynamic changes in genetic alterations during disease progression. Representative genetic alterations for each cluster are listed in the plot. B. Clonal evolution estimated using PhylogicNDT. The mean cancer cell fraction (CCF) and 95% confidence interval of each cluster are indicated. C. Fish-plot showing the clonal evolution process. The width of each timepoint indicates the clonal fractions of each subclone population. D. Joint distributions of CCF values of genetic alterations across two (or more) time points (ND: newly diagnosed; P: progression; R: relapse; R1: first relapse; R2: second relapse) were estimated using clustering analysis. Each line corresponds to cluster of genetic alterations (range 3-33) and illustrates the dynamic changes in CCF at the different time points for clusters. We classified any CCF increase or decrease >0.5 between two time points for any cluster as extreme evolution. CCF changes between 0.2 and 0.5 or less than 0.2 were classified as moderate evolution or no evolution, respectively. E. Sample interval and number of clonal clusters in patients with either extreme evolution or with modest or no evolution. F. Kaplan-Meier plot of survival from either first sampling (left) or second sampling (right).
Table 1. Clinical characteristics of the 134 MCL patients with WES analysis.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Total cohort (n=134)</th>
<th>Patients with Standard regimen (n=95)</th>
<th>Patients with other regimen (n=39)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median age at diagnosis, years (range)</td>
<td>59 (37-80)</td>
<td>56 (37-65)</td>
<td>68 (47-80)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sex, Male/Female</td>
<td>3.5/1</td>
<td>3.3/1</td>
<td>3.9/1</td>
<td>0.739</td>
</tr>
<tr>
<td>Presence of B symptoms, n (%)</td>
<td>64/134 (48)</td>
<td>46/95 (48)</td>
<td>21/39 (54)</td>
<td>0.568</td>
</tr>
<tr>
<td>Bone marrow involvement, n (%)</td>
<td>124/134 (93)</td>
<td>86/95 (91)</td>
<td>38/39 (97)</td>
<td>0.167</td>
</tr>
<tr>
<td>Nodal involvement, n (%)</td>
<td>86/127 (68)</td>
<td>62/92 (67)</td>
<td>24/35 (69)</td>
<td>0.899</td>
</tr>
<tr>
<td>Splenomegaly, n (%)</td>
<td>92/120 (80)</td>
<td>61/84 (73)</td>
<td>31/36 (86)</td>
<td>0.109</td>
</tr>
<tr>
<td>Elevated LDH, n (%)</td>
<td>56/132 (42)</td>
<td>37/93 (40)</td>
<td>19/39 (49)</td>
<td>0.343</td>
</tr>
<tr>
<td>Elevated β2-microglobulin, n (%)</td>
<td>54/81 (67)</td>
<td>36/56 (64)</td>
<td>18/25 (72)</td>
<td>0.496</td>
</tr>
<tr>
<td>MIPI score at diagnosis, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Low risk</td>
<td>34/132 (26)</td>
<td>30/93 (32)</td>
<td>4/39 (10)</td>
<td></td>
</tr>
<tr>
<td>Intermediate risk</td>
<td>49/132 (37)</td>
<td>38/93 (41)</td>
<td>11/39 (28)</td>
<td></td>
</tr>
<tr>
<td>High risk</td>
<td>49/132 (37)</td>
<td>25/93 (27)</td>
<td>24/39 (62)</td>
<td></td>
</tr>
<tr>
<td>Nanostring L-MCL16 assay, n (%)†</td>
<td></td>
<td></td>
<td></td>
<td>0.571</td>
</tr>
<tr>
<td>cMCL</td>
<td>38/48 (79)</td>
<td>27/35 (77)</td>
<td>11/13 (85)</td>
<td></td>
</tr>
<tr>
<td>nnMCL</td>
<td>10/48 (21)</td>
<td>8/35 (23)</td>
<td>2/13 (15)</td>
<td></td>
</tr>
<tr>
<td>Ki-67 index ≥30%, n (%)</td>
<td>26/57 (46)</td>
<td>20/44 (45)</td>
<td>6/13 (46)</td>
<td>0.965</td>
</tr>
<tr>
<td>Morphology, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>0.522</td>
</tr>
<tr>
<td>Classic</td>
<td>109/126 (87)</td>
<td>75/88 (85)</td>
<td>34/38 (89)</td>
<td></td>
</tr>
<tr>
<td>Blastoid/pleomorphic</td>
<td>17/126 (13)</td>
<td>13/88 (15)</td>
<td>4/38 (11)</td>
<td></td>
</tr>
<tr>
<td>SOX11 positive, n (%)</td>
<td>102/125 (82)</td>
<td>76/89 (85)</td>
<td>26/36 (72)</td>
<td>0.085</td>
</tr>
<tr>
<td>IGHV Unmutated, n (%)</td>
<td>114/134 (85)</td>
<td>81/95 (85)</td>
<td>33/39 (85)</td>
<td>0.924</td>
</tr>
<tr>
<td>Cytogenetics by FISH, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17p deletion</td>
<td>39/120 (33)</td>
<td>23/84 (27)</td>
<td>16/36 (44)</td>
<td>0.067</td>
</tr>
<tr>
<td>13q deletion</td>
<td>20/85 (24)</td>
<td>14/58 (24)</td>
<td>6/27 (22)</td>
<td>0.846</td>
</tr>
<tr>
<td>11q deletion</td>
<td>24/82 (29)</td>
<td>20/56 (36)</td>
<td>4/26 (15)</td>
<td>0.060</td>
</tr>
</tbody>
</table>
‡ Clot et al. (39)

‡ 152 MCL samples were obtained from 134 MCL patients. Longitudinal samples collected at initial diagnosis and relapse or first and second relapse were collected from 16 patients; for one patient, two tumor samples were collected from bone marrow and lymph node at the diagnosis.

§ 3 from spleen, 2 from intestine, 1 cerebrospinal fluid and 3 from other organs.

<table>
<thead>
<tr>
<th>Karyotyping, n (%)</th>
<th>0.719</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;3 aberrations</td>
<td>74/89 (83)</td>
</tr>
<tr>
<td>≥3 aberrations</td>
<td>15/89 (17)</td>
</tr>
</tbody>
</table>

| Sequenced sample, n (%)‡ |
|--------------------------|-------|
| Bone marrow              | 89/152 (59) | 57/112 (51) | 32/40 (80) |
| Peripheral blood         | 35/152 (23) | 28/112 (25) | 7/40 (18) |
| Lymph node               | 19/152 (13) | 18/112 (16) | 1/40 (2) |
| Other tissue§            | 9/152 (6) | 9/112 (8) | 0/40 (0) |
| With matched RNA-seq, n (%) | 48/134 (36) | 35/95 (37) | 13/39 (33) |

| Status at sampling, n (%) |
|---------------------------|-------|
| With treatment-naïve samples | 123/134 (92) | 88/95 (93) | 35/39 (90) |
| With relapse samples      | 24/134 (18) | 19/95 (20) | 5/39 (13) |
| With longitudinal samples | 18/134 (13) | 17/95 (18) | 1/39 (3) |

| Follow-up data           |
|--------------------------|-------|
| 3-year PFS, % (95% CI)   | 51 (41-61) | 61 (49-72) | 25 (9-41) |
| 3-year OS, % (95% CI)    | 70 (60-79) | 79 (69-88) | 45 (26-65) |