Coding and non-coding drivers of mantle cell lymphoma identified through exome and genome sequencing

Prasath Pararajalingam[1], Krysta M. Coyle[1], Sarah E. Arthur[1], Nicole Thomas[1], Miguel Alcaide[1], Barbara Meissner[2], Merrill Boyle[2], Bruno M. Grande[1], Graham Slack[2], Andrew J. Mungall[3], Randy D. Gascoyne[2], Christian Steidl[2], Joseph Connors[2], Diego Villa[2], Marco A. Marra[2], Nathalie Johnson[4], David W. Scott[2], and Ryan D. Morin[1]

[1]Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia, Canada [2]BC Cancer Centre for Lymphoid Cancer and BC Cancer Research Centre, Vancouver, British Columbia, Canada [3]Michael Smith Genome Sciences Centre, Vancouver, British Columbia, Canada [4]Department of Medicine, Jewish General Hospital, Montreal, Quebec, Canada

Key points

- RNA-binding proteins with roles in regulating alternative splicing, *DAZAP1*, *EWSR1*, *HNRNPH1*, are frequently mutated in MCL

- The majority of recurrent somatic *HNRNPH1* mutations are intronic and HNRNPH1 exhibits self-regulation through alternative splicing

Abstract

Mantle cell lymphoma (MCL) is an uncommon B-cell non-Hodgkin lymphoma (NHL) that is incurable with standard therapies. The genetic drivers of this cancer have not been firmly established and the features known to contribute to differences in clinical course remain limited. To extend our understanding of the biological pathways involved in this malignancy, we performed a large-scale genomic analysis of MCL using data from 51 exomes alongside previously published exome cohorts. To confirm our findings, we re-sequenced the genes identified in the exome cohort in 212 MCL tumors, each having clinical follow-up data. We confirmed the prognostic association of *TP53* and *NOTCH1* mutations and further nominate two additional genes, *EWSR1* and *MEF2B*, whose mutation respectively associated with poor and good outcome. Our sequencing revealed novel recurrent mutations including a unique missense hot spot in *MEF2B* and a pattern of non-coding mutations surrounding a single exon of the *HNRNPH1* gene. We sequenced the whole genomes of 34 MCLs to confirm the focal nature of *HNRNPH1* mutations. Using RNA-seq data from 110 of these cases, we identified a functional role for recurrent non-coding *HNRNPH1* mutations in disrupting an auto-regulatory feedback mechanism. Overall, we identified three novel MCL-related genes with roles in RNA trafficking or splicing, namely *DAZAP1*, *EWSR1*, and *HNRNPH1*. Taken together, these data strongly implicate a role for aberrant regulation of splicing in MCL pathobiology.

Introduction

Mantle cell lymphoma (MCL) is an uncommon B-cell lymphoma representing 4-9% of non-Hodgkin lymphoma (NHL) diagnoses worldwide[1]. MCL can be broadly divided into two clinical subtypes: nodal and leukemic non-nodal disease, with each displaying distinct natural history, and clinical and genetic features[2]. MCL commonly follows an aggressive clinical course including non-sustained responses to frontline chemo-immunotherapy and frequent relapses, although a subset of cases, including the majority of patients with leukemic non-nodal variant, exhibits significantly longer survival[3]. Clinical prognostic metrics such as MCL International Prognostic Index (MIPI) have enabled patient stratification and improvements to frontline therapy such as the inclusion of active agents (rituximab, bendamustine, cytarabine) as well as consolidative strategies (autologous stem cell transplantation), have significantly improved outcomes over the past two decades[1,4,5].

The unifying genetic feature of MCL is a chromosomal translocation event that places cyclin D1 (*CCND1*) downstream of the immunoglobulin heavy chain enhancer, causing constitutive *CCND1* expression[2,6]. The translocated *CCND1* allele can also accrue secondary mutations including non-coding mutations in the 3′ untranslated region (UTR), thereby enhancing *CCND1* mRNA stability and further elevating CCND1 protein abundance[7,8]. Through exome and targeted sequencing efforts, largely focused on non-nodal leukemic subtype, several genes have been identified as commonly mutated in MCL, including those involved in DNA damage response (*ATM*, *TP53*), epigenetic regulation (*KMT2D*, *WHSC1*), Notch signaling (*NOTCH1*, *NOTCH2*), NFκB signaling (*CARD11*, *BIRC3*, *SYK*), and ubiquitin mediated proteolysis (*UBR5*)[9–11]. The genomics of MCL have proven to be heterogeneous and diverse, and therefore larger comprehensive explorations are necessary to further understand its biology.

A limited number of recurrent mutations have been associated with prognosis in MCL treated with standard therapy. The most firmly established of these include non-silent mutations affecting *TP53*, *NOTCH1*, and *CCND1*[12–15], as well as amplifications of 3q or deletions in 17p[13,16]. With the ongoing evaluation of new therapeutics for MCL, mutations associated with acquired treatment resistance are beginning to be identified[17]. Despite a broad collection of MCL-related genes and mutations, stratification of patients by proliferation, whereby patients are separated into low-, intermediate- and high-risk categories, remains more robust than any individual driver mutation[14,18]. Genetic features that underlie differences in proliferation have yet to be identified.

This study identifies novel recurrent mutations in three RNA-binding proteins *HNRNPH1*, *DAZAP1*, and *EWSR1*, including intronic mutations affecting the splicing of exon 4 in *HNRNPH1*. Our functional characterization in MCL patient samples and cell lines showed *HNRNPH1* splicing is regulated by the HNRNPH1 protein itself via a negative feedback loop leading to alternative inclusion of exon 4 in the mature transcript. By further characterizing the mutational landscape of MCL, we implicate perturbed mRNA processing as an important mechanism in MCL biology.

Methods

Study design

We assembled a discovery cohort comprising 51 fresh-frozen diagnostic biopsies from MCL cases collected in British Columbia, Quebec (Montreal), and Ontario and subjected each of these to paired tumor/normal exome sequencing. We included available paired exome data from previous publications to yield a larger discovery cohort of 87 cases. We assembled a validation cohort from FFPE material representing 212 diagnostic tumor samples from BC for targeted sequencing. Only 23 patients overlap between these two cohorts with 189 unique cases in the validation cohort. We also subjected 16 cases from the validation cohort and 18 additional fresh frozen biopsies from BC, along with constitutional DNA, to whole genome sequencing (34 total). We performed RNA-seq on cases spanning all cohorts plus an additional 12 cases from a combination of FFPE and frozen material (110 total). Details for the samples and assays applied to each are included in **Supplemental Data** and **Supplemental Figure 1**. This study was approved by the BC Cancer REB and all participants were recruited with informed consent.

Exome data analysis

For in-house exomes, we used the Agilent SureSelect Human All Exon kits for library preparation and HiSeq2000 instruments (Illumina) for sequencing. We separately obtained exome data from 29 patients described in *Bea et al.*, which we downloaded from European Genome-Phenome Archive (EGAS00001000510) in BAM format, extracted to FASTQ and validated for data integrity using BamHash (v1.1)[9,19]. FASTQ files for tumor and matched normal exomes (7 patients) described in *Wu et al.* were kindly provided by the authors[11]. External cases qualified for inclusion if a matching normal sample was available. For cases with

sequence data from more than one tumor biopsy, we only included the earliest sample in our analysis. We used default settings for all software unless otherwise stated. We used BWA (v0.7.6a) to map reads to the GRCh38 human reference lacking alternate contigs[20] and subjected these BAM files to soft-clipping of overlapping read pairs using bamUtils clipoverlap (v1.0.13)[21]. For each BAM file, we applied GATK mark duplicates (v3.4.0), and adjusted alignments for putative indels using GATK indel realigner (v3.4.0)[22]. We used Strelka (v1.0.14)[23] to detect SSMs and indels and annotated these variants using Variant Effect Predictor (Ensembl release 83)[24] and vcf2maf.

Recurrence analysis

As genes important for lymphomagenesis can exhibit a variety of mutational patterns, we employed a voting strategy involving four separate algorithms to identify recurrently mutated genes and hot spots. Several algorithms to infer genes that are recurrently mutated in cancer cohorts have been described and each of these relies on a variety of features such as predicted functional impact[25,26], spatial clustering[27], and mutation rates[28]. We identified significantly mutated genes using a combination of MutSigCV[28], OncodriveFM[25], OncodriveFML[26] and OncodriveCLUST[27] using a false discovery rate threshold of 0.1 for each algorithm. We promoted genes identified by two or more methods for further sequencing in the validation cohort. Additionally, *NOTCH1*, *CARD11*, *NFKBIE* were included due to their importance in MCL and other B-cell NHL[11,12,29].

Targeted sequencing, whole genome sequencing and data consolidation

DNA extracted from 212 formalin-fixed paraffin-embedded (FFPE) diagnostic MCL tumor biopsies were used to generate libraries that were enriched for exons corresponding to a set of putative MCL-related genes using a hybridization-based capture approach involving

7

complementary DNA oligonucleotides[30]. These were pooled and sequenced using a MiSeq (Illumina). Paired 150-nt reads were generated, demultiplexed, and mapped to the GRCh38 human reference using Geneious (v9.1.5). Variants with minor allele frequencies greater than 0.0001 in any gnomAD population were considered germline variants and removed[31]. We also performed whole genome sequencing on FF tumor and matched normal samples from 34 MCL cases. Variants from the genome data were identified using Strelka2[32], which we found to be more robust for identifying variants with low read support or low-level contamination of blood DNA with tumor cells (a feature of some MCLs). We consolidated variants from cases sequenced by more than one method using a tiered approach. Variants found in exomes and genomes were combined per patient. In targeted sequenced tumours for which a normal exome or genome data was available, we considered variants with more than one read support in the normal to be germline variants and removed from analysis. Variants from targeted sequencing cases that did not overlap with exome or genome cases were included in the final variant set. We compared the mutation patterns to diffuse large B-cell lymphoma (DLBCL) using targeted and exome sequencing data from 1616 unique patients[30,33,34] and removed variants with minor allele frequencies greater than 0.0001 in any gnomAD population[31].

RNA-seq analysis of MCL

We performed RNA-sequencing on 110 MCL cases (**Supplemental Methods**). The majority of cases had previously undergone DNA-sequencing, while 12 cases did not. We aligned RNA-seq reads to GRCh38 human reference genome using STAR (version 2.5.3a) followed by Picard MarkDuplicates (version 2.14.1)[22,35]. We used FeatureCounts (version 1.6.0) to separately quantify the reads mapping to the relevant *HNRNPH1* splice junctions using the following parameters: juncCounts, countSplit, ignoreDup, requireBothEndsMapped, and minimum

8

mapping quality $\geq$ 10. We defined the *HNRNPH1* exon skipping ratio as the ratio of reads spanning the exon 4-6 junction to the sum of reads spanning the exon 4-5 junction and 5-6 junction. RNA-seq samples were designated HNRNPH1 exon 4 mutated or unmutated based on variants discovered in the matching DNA-sequencing sample. In cases where DNA-sequencing data was not available, we determined *HNRNPH1* exon 4 mutation status by examining the exon 4 region in the RNA-sequencing data in IGV.

Protein-RNA interactions of HNRNPH1

RNA-seq and iCLIP (individual nucleotide resolution cross-linking and immunoprecipitation) data from HeLa cells was provided by authors[36]. Using STAR, we aligned RNA-seq reads from untreated cells transfected with control siRNA (seq8 in Uren data) as described above. Here, we assigned counts using featureCounts with a flattened GTF file containing collapsed transcripts. For HNRNPH iCLIP reads, we pooled reads from two replicates and aligned each pool to GRCh38 with STAR. Peaks in iCLIP data were called by Piranha[37] in 50 bp bins, accounting for RNA abundance with log-converted RNA-seq counts used as a covariate.

Cell-based experiments

REC-1 cells were grown as previously described[12,30]. We reconstituted cycloheximide (Sigma) in DMSO and added either cycloheximide or DMSO alone to cells for up to 6 hours. We then extracted RNA or protein, as appropriate, from pellets of 2 x 10^6 cells using the RNeasy mini kit (Qiagen) or RIPA buffer, respectively, and quantified protein using the Pierce BCA kit (ThermoFisher). Equal amounts of protein were used for SDS-PAGE, and we used Western blotting to observe HNRNPH1 protein levels (abcam ab10374, 1/10 000) and histone H3 protein levels (Cell Signaling Technologies #9715, 1/1 000).

Digital PCR

For droplet digital PCR (ddPCR), we extracted RNA from FFPE-preserved samples as previously described[18] and performed reverse transcription with random hexamers (iScript, BioRad) according to manufacturer's instructions. We pre-amplified cDNA using SsoAdvanced PreAmp SuperMix (BioRad) for 15 cycles with primers targeting *HNRNPH1* (all), *HNRNPH1* (canonical only), *HNRNPH1* (alternative only), *TBP*, *YWHAZ,* and *UBC* (**Supplemental Methods**) as directed. Following a 1 in 5 dilution of pre-amplified cDNA, we performed ddPCR on each on the QX200 system (BioRad) using the primers and probes described in **Supplemental Table 1**. In each sample, we calculated the normalized expression relative to the geometric mean of the expression of three reference genes (*TBP*, *YWHAZ*, and *UBC*). Equal amounts of cell-line RNA were reverse transcribed with random hexamers as above. With diluted cDNA (1 in 10), we performed ddPCR (**Supplemental Table 1**). We normalized expression relative to the geometric mean of three reference genes (*ACTB*, *YWHAZ*, and *UBC*).

Statistical analysis

All associations between gene mutation status and binary clinical characteristics were assessed using Fisher's exact test. Association between mutations and overall survival (OS) was determined in the subset of MCL cases that were nodal and treated (n = 175) using the Kaplan-Meier method and log-rank test. We annotated MCL tumours as low-, intermediate- and high-risk based on proliferation gene expression signature using the established Nanostring nCounter-based MCL35 assay[18]. Because the high and intermediate risk cases did not exhibit significantly different overall survival, the intermediate and high proliferation groups were combined into a single group that was defined here as high risk.

10

Results

Somatic Mutation Landscape and Recurrently Mutated Genes

Several genes have previously been implicated as recurrent targets of SSM in MCL[9–11,38] though many of these candidate drivers have been inconsistent between individual studies[39]. This can be attributed to a combination of genetic heterogeneity and limited cohort sizes. To address this, we sequenced paired tumor/normal exomes from 51 MCLs diagnosed in Canada and analyzed these data alongscide published paired exome data. Three samples exhibited significantly higher mutation burden (median 1761; range 602-8780) and were excluded due to the effect of hypermutation on the detection of drivers, leading to a discovery cohort comprising of 84 cases. In total, 2122 genes were mutated in at least one tumor and tumors, on average, harbored non-silent SSMs in 38 genes (range 11-102).

Through our analysis of this cohort, we found that 16 genes were recurrently mutated as identified by two or more algorithms used to identify driver genes. Three of the algorthims deemed each of *ATM*, *BIRC3*, *TP53*, *S1PR1*, and *B2M* to be significantly mutated, and each of *MEF2B* and *WHSC1* were identified by two methods (**Supplemental Table 2**). Notably, *CCND1*, known to be affected by somatic hypermutation, was identified by OncodriveCLUST, which relies on spatial clustering of mutations. Of the candidate MCL genes, those frequently mutated were *ATM* (44%; n = 37), *CCND1* (18%; n = 15), *TP53* (11%; n = 9), *WHSC1* (15%; n = 13), and *KMT2D* (12%; n = 10), each gene having been previously nominated by other studies (**Supplemental Figure 2**). Three genes not previously attributed to MCL were also identified by at least two methods, including *HNRNPH1* (3.6%; n = 3), *DAZAP1* (3.6%; n = 3), and *EWSR1* (3.6%; n = 3) (**Supplemental Figure 2**). These each encode RNA-binding proteins that play a role in regulating RNA metabolism and alternative splicing of pre-mRNA[40,41].

11

Confirmation of mutation pattern and prevalence

Based on the results and prior studies, we performed targeted sequencing of 18 genes in additional MCLs and used WGS to broadly resolve the exonic and intronic mutation patterns. We applied a combination of targeted sequencing (n = 212) and paired WGS (n = 34) to comprehensively explore the genetic and gene expression landscape of MCL. We consolidated variants across all samples and used the resulting non-duplicated variants for all analyses. Mutation patterns and prevalence in established MCL genes were largely consistent with prior reports (**Figure 1A**; **Supplemental Table 3**). Each of *NOTCH1*, *MEF2B* and *CCND1* have been shown to have mutation hot spots in MCL and other cancers[7,12,42,43]. In MEF2B, K23R was the predominant mutation observed in MCL (**Supplemental Figure 3**) whereas this mutation was only observed in a single diffuse large B-cell lymphoma (DLBCL) case (n = 1616; P = 4.3 x 10^{-17}, Fisher's exact test). The K23R mutation has been shown to reduce the DNA-binding ability of MEF2B, thus affecting regulation of its transcriptional targets[44]. The majority of *MEF2B* mutations found in DLBCL and follicular lymphomas (FL) affect three hot spots which are absent from MCL, namely K4, D69, and D83[43,45]. These mutations have been described to alter DNA or co-repressor binding by MEF2B, resulting in constitutive expression of MEF2B target genes[46–48]. The distinct hot spot at K23 and paucity of DLBCL/FL-related mutations implies a different role of *MEF2B* mutations in MCL.

Unsurprisingly, the incidence of non-silent mutations in newly identified genes was generally lower than those of established MCL genes. *EWSR1* was mutated in 9 cases (3.1%) and DAZAP1, in 13 cases (4.5%) (**Figure 1A**). *EWSR1* predominantly harbored frameshift or nonsense mutations in MCL (n = 8; 2.7%) and exhibited a similar pattern at a lower prevalance in a larger compendium of DLBCLs (n = 5; 0.3%) (**Figure 1B**). This suggests that *EWSR1* has

an unappreciated tumor suppressor function in both MCL and DLBCL. *DAZAP1* had a distinctive mutation pattern where mutations were clustered near the C-terminus in a region containing a nuclear localization signal (p.G383-R407)[49] and proline-rich protein-binding domain[50,51] (**Figure 1C**). Nine cases harbored putative truncating mutations with each predicted to remove or disrupt the nuclear localization signal while leaving most of the open reading frame intact. Non-synonymous mutations in this region mainly affected highly conserved residues (i.e., p.F402, p.R406, p.R407). Substitution of these residues has been shown to cause cytoplasmic accumulation of DAZAP1 in human kidney epithelial (293T) cells and simian (COS7) cells[49].

HNRNPH1 splice site mutations disrupt HNRNPH1 binding motifs

HNRNPH1 was mutated in 10% when we considered both coding and non-coding mutations, placing it as the eighth most commonly mutated gene in the cohort (**Supplemental Figure 4**). Despite limited coverage of introns by our sequencing assay, intronic variants were the most common SSM type in this gene, particularly in the regions surrounding exon 4 (**Figure 2A**). Data from paired sequencing confirmed that these were somatic and the WGS data showed they wer largely restricted to this exon and the immediate flanking regions (**Figures 2A and 2B**), HNRNP proteins are widely involved in regulating splicing by binding to pre-mRNA at specific motifs and either promoting or inhibiting usage of nearby splice sites. Distinct from other hnRNPs, HNRNPH1 (and its paralog HNRNPH2) preferentially binds RNA at poly-G motifs[36]. Strikingly, 65% (20/31) of patients with *HNRNPH1* mutations had mutations affecting a poly-G motif within or near this exon. Among the available WGS data from Burkitt lymphoma[52] (n = 106), DLBCL (n = 153), chronic lymphocytic leukemia (CLL; n = 144) and follicular lymphoma (n = 110), we only identified two DLBCL patients with *HNRNPH1* mutations in this region (1.3%)[30].

The highly specific mutation pattern led us to speculate that HNRNPH1 protein regulates its expression by modulating the splicing of the *HNRNPH1* mRNA transcript. We re-analyzed HNRNPH1 iCLIP-seq data from *Uren et al.*[36] and confirmed multiple sites of interaction between HNRNPH1 and its pre-mRNA including exon 4 (**Figure 2C**), supporting a model of direct association at these poly-G motifs. *HNRNPH1* has multiple alternative splicing events annotated including the skipping of exon 4. As annotated, isoforms with this event would be subjected to nonsense-mediated decay (NMD) due to disruption of the reading frame. Using a custom ddPCR assay which separately quantifies canonical and alternative *HNRNPH1* transcripts, we demonstrate that inhibition of the NMD process by cycloheximide causes a significant increase in alternative transcript compared to the canonical transcript (**Figure 3A**). This was consistent with a model wherein this gene regulates its own splicing by a negative feedback mechanism to limit HNRNPH1 protein abundance (**Supplemental Figure 5**). Although they do not directly affect canonical splice signals, we hypothesized that the mutations in the poly-G motifs impact the splicing of exon 4.

We analyzed RNA-seq data from 110 cases with known *HNRNPH1* mutation status to evaluate splicing differences between mutated (n = 13) and unmutated (n = 97) tumors. The mutated cases exhibited significantly greater inclusion of exon 4 than the unmutated cases (P = 2.3 x 10^{-4}; **Figure 3B**), consistent with our model (**Supplemental Figure 5**). These findings were corroborated by ddPCR of selected cases (P < 0.001; **Figure 3C**), which strongly correlated (R = 0.66, P < 0.01) with splicing ratios identified from RNA-seq data (**Supplemental Figure 5**). These results support the notion that *HNRNPH1* mutations promote the inclusion of exon 4 and the production of mature transcripts by reducing feedback inhibition, most likely by disrupting HNRNPH1 binding motifs surrounding exon 4.

Association of recurrent mutations with OS

We next explored the relationship between mutation the individual genes and OS in nodal and treated MCL cases that also underwent DNA sequencing. The clinical characteristics of the patients used for survival analysis were largely similar to previous MCL cohorts (**Table 1**). Consistent with previous reports, mutations in *TP53* (**Figure 4A**) and *NOTCH1* (**Figure 4B**) were associated with shorter OS (P = 5.2×10^{-9} and P = 4.0×10^{-3}, respectively). *MEF2B* K23R mutated cases (n = 9) exhibited significantly longer overall median OS (**Figure 4C**, median not reached; P = 0.04). Among the RNA-binding proteins, we found a novel association between *EWSR1* mutations and worse OS (P = 1.8×10^{-4}). Given the robust association between proliferation and prognosis, we next examined whether these prognostic mutations were unequally distributed between MCLs with high or low proliferation after combining intermediate risk cases with high risk cases (**Supplemental Figure 6A**). Mutations in either *TP53* (P = 4.3×10^{-5}) or *MEF2B* (P = 1.4×10^{-2}) were associated with high- or low-proliferation tumors, respectively (**Supplemental Figure 6B**), while mutations in RNA-processing genes were equally distributed between high- and low-proliferation tumors. Through multivariate analysis we found that only MCL35-derived risk status (P = 1.1×10^{-2}) and *TP53* mutations (P = 9.6×10^{-5}) were independently prognostic of OS (**Supplemental Table 5**).

Discussion

This study represents the largest profiling of mutations in MCL to date. Using 291 MCL cases, we validated the recurrence of mutations in genes with known relevance to MCL, including *ATM*, *KMT2D*, *TP53*, *CCND1*, and *NOTCH1*. Using clinical data available for the bulk of these cases, we confirmed the prognostic association of mutations in both *TP53* and *NOTCH1*. *NOTCH1* was not independently prognostic when considering *TP53* mutations and MCL35-

15

derived risk categories. We also found that *WHSC1* mutations were not associated with OS as was previously discovered[53], which may be due to the limited sample size of that study. We observed prognostic associations for *EWSR1* mutations and *MEF2B* K23R mutations although neither were significant in the multivariate analysis. Confirmation of prognostic association of these genes will require substantially larger cohort sizes. Although the role of *MEF2B* mutations in prognosis remains to be clarified, the high frequency and specificity of K23R mutations in MCL highlights the need for functional characterization of this hot spot as has been done for hot spots common in DLBCL[46,48].

Althoughly less frequently mutated in MCL, *EWSR1* is an established cancer gene that is typically discussed in the context of the EWSR1-FLI1 fusion oncoprotein that drives Ewing sarcoma[54]. The pattern of mutations observed here implies a separate tumor suppressor role of this gene in MCL and DLBCL. FET family proteins, including EWSR1, each comprise N-terminal transcriptional activating domains and RNA-binding domains near the C-terminus. One of the physiologic functions of EWSR1 entails the coupling of transcription and splicing through its interactions with RNA Polymerase II and recruitment of splicing factors through its C-terminal domain[55–57]. Depletion can induce alternative splicing of EWSR1 targets including genes involved in DNA repair and genotoxic stress signaling[58]. Notably, EWSR1 has been implicated in regulating CCND1 by promoting formation of the less oncogenic CCND1a isoform relative to the shorter CCND1b isoform[59]. Although the targets of EWSR1 have not been established in MCL, our data are consistent with the notion that loss of EWSR1 activity alters RNA metabolism and splicing of genes relevant to MCL.

The DAZAP1 protein contains two N-terminal, highly conserved RNA recognition motif (RRM) domains[60], which allow it to interact with the pre-mRNA transcripts of its targets and regulate

16

alternative splicing[51,61,62]. Its C-terminal proline-rich domain[63] is responsible for interactions with other splicing factors[61]. The protein primarily localizes in the nucleus but can also be found in the cytoplasm and within mitochondria[49]. There have been conflicting reports of nuclear localization signals (NLS) in either the N- and C-terminus of this gene[49,64]. A truncated protein lacking the C-terminus or with mutations at highly conserved residues, R393, F402, and F406, has been shown to accumulate in the cytoplasm[49]. In addition to mutations in MCL affecting each of F402 and F406, we observed multiple cases with C-terminal truncating mutations, which would each remove the C-terminal NLS while leaving most of the protein intact. Furthermore, we and others have demonstrated a similar pattern of mutations in a subset of DLBCLs[30,34]. The existence of recurrent *EWSR1* and *DAZAP1* mutations in both malignancies add to the limited genetic features shared between DLBCL and MCL, along with inactivating mutations in *KMT2D* and *TP53*. Based on the effect of mutagenesis experiments[49], we hypothesize that the commonly observed mutations cause reduced nuclear occupancy of DAZAP1 and affect interactions with other proteins. This could disrupt several processes, including transcription, alternative splicing, mRNA transport and translation[40,51,61].

HNRNPH1 is a member of the HNRNPH/F family of heterogeneous nuclear ribonucleoproteins[65]. HNRNPH1 binds to various cis-regulatory elements, and depending on the sequence context and interacting proteins, can promote or suppress the use of nearby splice sites[66]. Our data support a model wherein HNRNPH1 protein normally limits its own accumulation by suppressing the inclusion of exon 4 and directing its mRNA for degradation by NMD. In this model, the observed intronic mutations would destroy motifs surrounding exon 4 thereby favoring exon inclusion and enhancing the production of the mature transcript, an observation that is supported by RNA-seq (110 cases) and ddPCR (34 cases) analysis. The

17

possibility that exon 4 inclusion is correlated with HNRNPH1 levels, and the paucity of mutations in this region in other B-cell NHL may indicate a differential role for HNRNPH1 in B-cell development or lymphomagenesis and are consistent with a more important role of HNRNPH1 in MCL biology. Similar to the predicted effects of other RNA binding proteins with a multiplicity of targets, enhanced activity of HNRNPH1 is expected to have widespread effects on the splicing landscape in MCL[36].

Mutations in these novel MCL-related genes (*EWSR1*, *DAZAP1*, and *HNRNPH1*) are compelling as they implicate mRNA maturation, splicing and/or trafficking in lymphomagenesis. Evidence is accumulating which relates alterations in RNA-binding proteins and splice factors in numerous cancers to various aspects of cancer cell biology. Specifically, small changes in RNA-binding proteins can have large downstream effects on gene expression and can thus impact multiple hallmarks of cancer[67]. For example, the splicing factor SF3B1 was identified as recurrently mutated in CLL[68–70] and further detailed investigations have identified widespread alternative splicing affecting multiple cellular pathways[71,72]. The identification of pleiotropic downstream effects including DNA damage response, apoptosis, and Notch signaling[71,73], indicate that widespread disruptions to RNA processing can enhance cancer cell survival by multiple pathways. The evidence for alterations affecting multiple splicing factors in the literature[74,75], including those observed in this study, suggest an emerging role of RNA metabolism and splicing in B-cell lymphomas.

In summary, through genomic analysis of 303 MCL tumors, we identified novel recurrently mutated genes with a range of mutation incidences. We implicate an important role for RNA-binding proteins and RNA processing in MCL as compared to other B-cell lymphomas, suggesting that RNA metabolism and splicing have a specific role in MCL pathology, although

the downstream targets of these genes in MCL have yet to be characterized. Further work which links these mutations to dysregulation of specific RNA molecules will highlight the relevance of RNA processing in MCL.

Acknowledgments

Author Contributions

PP analyzed and interpreted the data; SEA, MA performed library preparation; KMC and NT analyzed CLIP-seq and RNA-sequencing data; KMC performed cell-based and FFPE-based assays; MA performed sequencing and analysis targeted sequencing; BMG provided bioinformatic support and interpreted the data; BM, MB, GS, AM performed nucleic acid extractions and sample quality control; DV and DWS provided clinical data and reviewed the cases; DWS, MAM, NJ, RG, CS, JC and RDM interpreted data, designed the study, and with PP, wrote the manuscript.

Disclosure of Conflicts of Interest

RDG, JMC, DV and DWS are inventors of Nanostring nCounter-based MCL35 assay.

19

Tables

Table 1. Characteristics of patient samples gathered in British Columbia. MCL tumor samples from 232 patients were obtained from patients living in British Columbia. Of the 232 tumors, 175 were nodal and initially treated and were used in survival analysis.

	Nodal, treated (n=175)	Others (n=57)*	Total (n=232)
Median age (Range)	63 (31-85)	70 (39-90)	64 (31-90)
Male (%)	153 (77)	38 (67)	173 (75)
Performance status >1 (%)	37/166 (22)	8/53 (15)	45/219 (20)
Elevated LDH (%)	54/169 (32)	4/53 (15)	58/222 (26)
Blastoid (%)	21 (12)	1 (2)	22 (10)
Bulky (mass >= 10cm) (%)	21 (18)	5 (9)	36 (16)
Ann Arbor Stage (%)			
1	8/174 (5)	5/55 (7)	12/229 (5)
2	9/174 (5)	3/55 (6)	12/229 (5)
3	14/174 (9)	7/55 (13)	22/229 (10)
4	142/174 (82)	41/55 (74)	183/229 (80)
Bone marrow involvement (%)	122 (70)	33 (58)	155 (67)
GI tract involvement (%)	15 (9)	12 (21)	27 (12)
B symptoms (%)	52/174 (30)	5/55 (9)	57/229 (25)
MIPI (%)			
High	28/174 (16)	2/56 (4)	30/230 (13)
Intermediate	39/174 (22)	20/56 (36)	59/230 (26)
Low	107/174 (62)	34/56 (61)	141/230 (61)
Treatment			
Splenectomy	2	1	3
Palliative chemotherapy	2	0	2
CVPR	3	1	4
R-CHOP	153	4	157
BR	15	1	15
Observation	0	51	51

* Samples not used in survival analysis: Nodal observed, 47; Non-nodal, 10
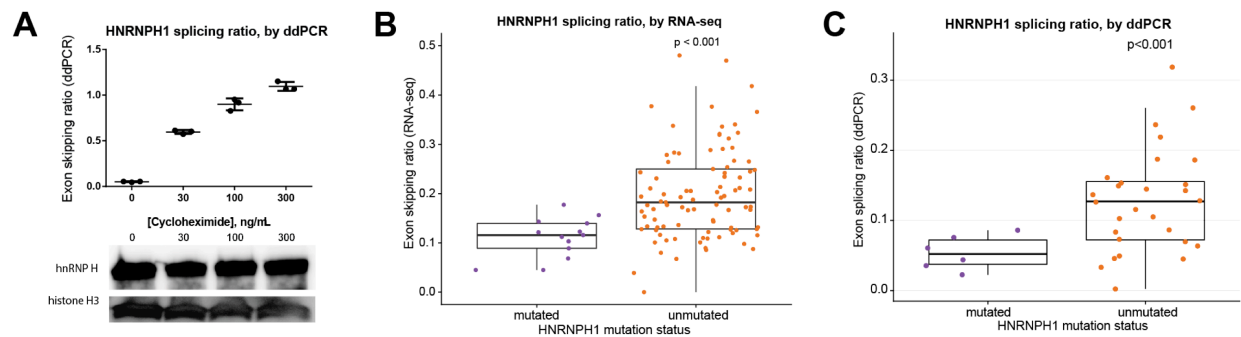
Figures

Figure 1. MCL mutation pattern. (A) Mutations observed across 291 MCL samples in 18 candidate MCL genes. Mutations shown here are limited to non-silent mutations for all genes with the exception of *HNRNPH1*, for which intronic and silent mutations were also included. Spatial distribution of mutations observed in (B) *EWSR1*, and (C) *DAZAP1* in MCL (top) compared to DLBCL (bottom).
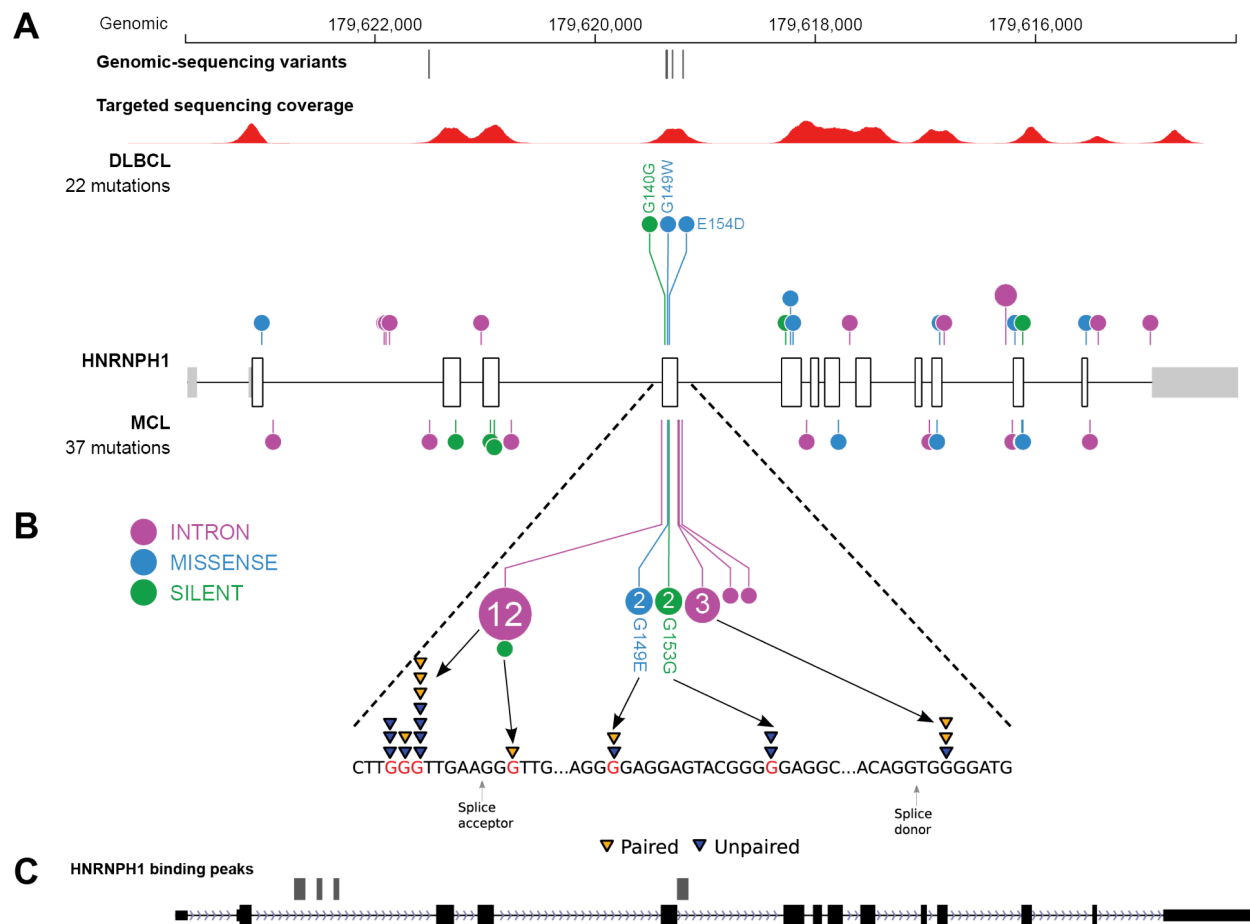
Figure 2. *HNRNPH1* mutations in MCL cluster near exon 4 in poly-G motifs. (A) Somatic mutations found in genomic sequencing cases and targeted sequencing coverage of a representative sample. The prevalence and pattern of mutations in *HNRNPH1* is compared between DLBCL (above) and MCL (below). (B) Splice site and intronic mutations were observed both upstream and downstream of exon 4 affecting poly-G motifs. Paired mutations (orange triangles) are mutations found to be somatic by sequencing matched constitutional DNA ($n = 7$). Unpaired mutations (blue triangles) are mutations found in tumor-only DNA sequencing ($n = 12$). (C) HNRNPH1 iCLIP binding peaks show that HNRNPH1 binds near exon 4 of the transcript (shown is Refseq isoform NM_001257293).

24

Figure 3. *HNRNPH1* **mutations impact RNA splicing.** (A) We separately quantified canonical and alternative *HNRNPH1* transcripts by digital PCR in REC-1 cells cultured with cycloheximide (an inhibitor of NMD). This revealed an increasing proportion of the alternative transcript (skipped exon 4) with high concentrations. This increase was not associated with increased HNRNPH protein as determined by immunoblot. (B) Mutated *HNRNPH1* cases showed significantly lower exon skipping ratios compared to unmutated cases, as measured by RNA-seq. (C) Digital PCR was used to separately quantify alternative and canonical *HNRNPH1* transcripts in mutant (n = 6) and wildtype (n = 30) cases. Mutant cases exhibit lower rate of exon skipping and higher overall abundance of HNRNPH1 mRNA.
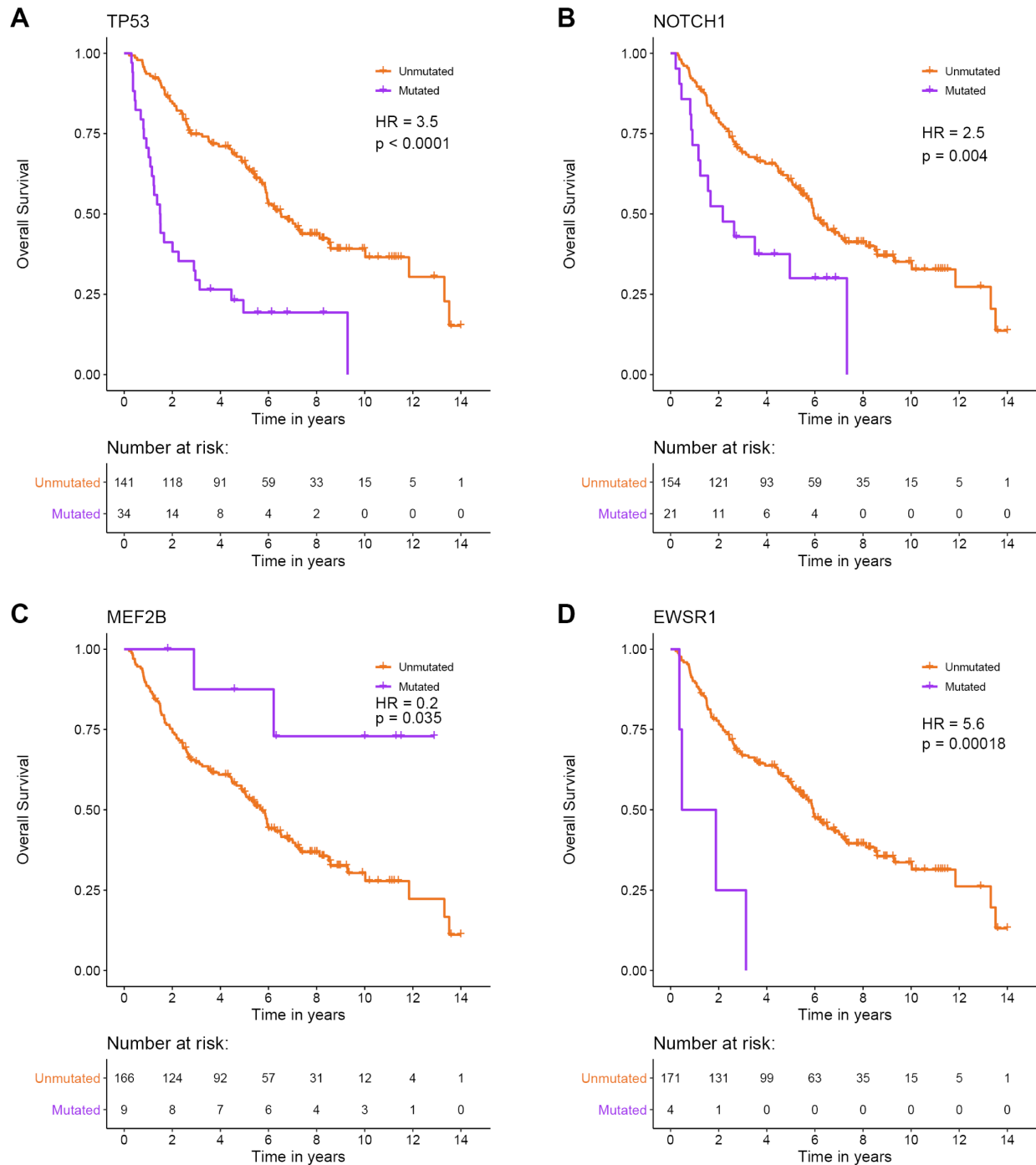
Figure 4. *TP53*, *NOTCH1*, *MEF2B*, and *EWSR1* mutations are associated with overall survival in nodal MCL. Survival analysis on the nodal MCL cases revealed four genes with mutations significantly associated with outcome. (A) *TP53* mutated cases (n = 34) had shorter overall survival (P = 5.2×10^{-9}). (B) *NOTCH1* mutated cases (n = 21) had shorter overall

survival (P = 4.0 x 10^{-3}). (C) In contrast, *MEF2B* K23R cases (n = 9) exhibited longer overall

survival (P = 3.5 x 10^{-2}). (D) *EWSR1* mutated cases (n = 4) exhibited shorted overall survival (P

= 1.8 x 10^{-4}). *TP53*, *NOTCH1*, and *EWSR1* remained significant after correcting for multiple

hypothesis testing (FDR < 0.1).

References

1. Vose JM. Mantle cell lymphoma: 2017 update on diagnosis, risk-stratification, and clinical management. *American Journal of Hematology*. 2017;92(8):806–813.

2. Chen D, Viswanatha DS, Zent CS, et al. Indolent Mantle Cell Lymphoma: A Distinct Subgroup Characterized by Leukemic Phase Disease without Lymphadenopathy. *Blood*. 2009;114(22):3937–3937.

3. Abrisqueta P, Scott DW, Slack GW, et al. Observation as the initial management strategy in patients with mantle cell lymphoma. *Annals of Oncology: Official Journal of the European Society for Medical Oncology*. 2017;28(10):2489–2495.

4. LIM S, HORSMAN J, HANCOCK B. The Mantle Cell Lymphoma International Prognostic Index: Does it work in routine practice? *Oncology Letters*. 2010;1(1):187–188.

5. Maddocks K. Update on Mantle Cell Lymphoma. *Blood*. 2018;blood–2018–03–791392.

6. Cheah CY, Seymour JF, Wang ML. Mantle Cell Lymphoma. *Journal of Clinical Oncology*. 2016;34(11):1256–1269.

7. Mohanty A, Sandoval N, Das M, et al. CCND1 mutations increase protein stability and promote ibrutinib resistance in mantle cell lymphoma. *Oncotarget*. 2016;7(45):73558–73572.

8. Wiestner A, Tehrani M, Chiorazzi M, et al. Point mutations and genomic deletions in CCND1 create stable truncated cyclin D1 mRNAs that are associated with increased proliferation rate and shorter survival. *Blood*. 2007;109(11):4599–4606.

9. Bea S, Valdes-Mas R, Navarro A, et al. Landscape of somatic mutations and clonal evolution in mantle cell lymphoma. *Proceedings of the National Academy of Sciences*. 2013;110(45):18250–18255.

10. Zhang J, Jima D, Moffitt AB, et al. The genomic landscape of mantle cell lymphoma is related to the epigenetically determined chromatin state of normal B cells. *Blood*. 2014;123(19):2988–2996.

11. Wu C, Miranda NF de, Chen L, et al. Genetic heterogeneity in primary and relapsed mantle cell lymphomas: Impact of recurrent CARD11 mutations. *Oncotarget*. 2016;7(25):38180–38190.

12. Kridel R, Meissner B, Rogic S, et al. Whole transcriptome sequencing reveals recurrent NOTCH1 mutations in mantle cell lymphoma. *Blood*. 2012;119(9):1963–1971.

13. Halldórsdóttir AM, Lundin A, Murray F, et al. Impact of *TP53* mutation and 17p deletion in mantle cell lymphoma. *Leukemia*. 2011;25(12):1904–1908.

14. Rosenwald A, Wright G, Wiestner A, et al. The proliferation gene expression signature is a quantitative integrator of oncogenic events that predicts survival in mantle cell lymphoma. *Cancer Cell*. 2003;3(2):185–197.

15. Tam CS, Anderson MA, Pott C, et al. Ibrutinib plus Venetoclax for the Treatment of Mantle-Cell Lymphoma. *New England Journal of Medicine*. 2018;378(13):1211–1223.

16. Espinet B, Salaverria I, Beà S, et al. Incidence and prognostic impact of secondary cytogenetic aberrations in a series of 145 patients with mantle cell lymphoma. *Genes, Chromosomes and Cancer*. 2010;49(5):439–451.

17. Agarwal R, Chan Y-C, Tam CS, et al. Dynamic molecular monitoring reveals that SWI–SNF mutations mediate resistance to ibrutinib plus venetoclax in mantle cell lymphoma. *Nature Medicine*. 2019;25(1):119.

18. Scott DW, Abrisqueta P, Wright GW, et al. New Molecular Assay for the Proliferation Signature in Mantle Cell Lymphoma Applicable to Formalin-Fixed Paraffin-Embedded Biopsies. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*. 2017;35(15):1668–1677.

19. Óskarsdóttir A, Másson G, Melsted P. BamHash: A checksum program for verifying the integrity of sequence data. *Bioinformatics*. 2016;32(1):140–141.

20. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:1303.3997 [q-bio]*. 2013;

21. Jun G, Wing MK, Abecasis GR, Kang HM. An efficient and scalable analysis framework for variant extraction and refinement from population scale DNA sequence data. *Genome Research*. 2015;gr.176552.114.

22. Van der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics*. 2013;43:11.10.1–33.

23. Saunders CT, Wong WSW, Swamy S, et al. Strelka: Accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics*. 2012;28(14):1811–1817.

24. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biology*. 2016;17:122.

25. Gonzalez-Perez A, Lopez-Bigas N. Functional impact bias reveals cancer drivers. *Nucleic Acids Research*. 2012;40(21):e169–e169.

26. Mularoni L, Sabarinathan R, Deu-Pons J, Gonzalez-Perez A, López-Bigas N. OncodriveFML: A general framework to identify coding and non-coding regions with cancer driver mutations. *Genome Biology*. 2016;17:128.

27. Tamborero D, Gonzalez-Perez A, Lopez-Bigas N. OncodriveCLUST: Exploiting the positional clustering of somatic mutations to identify cancer genes. *Bioinformatics (Oxford, England)*. 2013;29(18):2238–2244.

28. Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*. 2013;499(7457):214–218.

29. Mansouri L, Noerenberg D, Young E, et al. Frequent NFKBIE deletions are associated with poor outcome in primary mediastinal B-cell lymphoma. *Blood*. 2016;128(23):2666–2670.

30. Arthur SE, Jiang A, Grande BM, et al. Genome-wide discovery of somatic regulatory variants in diffuse large B-cell lymphoma. *Nature Communications*. 2018;9(1):4001.

31. Karczewski KJ, Francioli LC, Tiao G, et al. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *bioRxiv*. 2019;531210.

32. Kim S, Scheffler K, Halpern AL, et al. Strelka2: Fast and accurate calling of germline and somatic variants. *Nature Methods*. 2018;15(8):591.

33. Reddy A, Zhang J, Davis NS, et al. Genetic and Functional Drivers of Diffuse Large B Cell Lymphoma. *Cell*. 2017;171(2):481–494.e15.

34. Schmitz R, Wright GW, Huang DW, et al. Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. *New England Journal of Medicine*. 2018;378(15):1396–1407.

35. Dobin A, Davis CA, Schlesinger F, et al. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)*. 2013;29(1):15–21.

36. Uren PJ, Bahrami-Samani E, Araujo PR de, et al. High-throughput analyses of hnRNP H1 dissects its multi-functional aspect. *RNA Biology*. 2016;13(4):400–411.

37. Uren PJ, Bahrami-Samani E, Burns SC, et al. Site identification in high-throughput RNA–protein interaction data. *Bioinformatics*. 2012;28(23):3013–3020.

38. Meissner B, Kridel R, Lim RS, et al. The E3 ubiquitin ligase UBR5 is recurrently mutated in mantle cell lymphoma. *Blood*. 2013;121(16):3161–3164.

39. Ahmed M, Zhang L, Nomie K, Lam L, Wang M. Gene mutations and actionable genetic lesions in mantle cell lymphoma. *Oncotarget*. 2016;7(36):58638–58648.

40. Smith RWP, Anderson RC, Smith JWS, et al. DAZAP1, an RNA-binding protein required for development and spermatogenesis, can regulate mRNA translation. *RNA*. 2011;17(7):1282–1295.

41. Paronetto MP, Miñana B, Valcárcel J. The Ewing sarcoma protein regulates DNA damage-induced alternative splicing. *Molecular Cell*. 2011;43(3):353–368.

42. Fabbri G, Rasi S, Rossi D, et al. Analysis of the chronic lymphocytic leukemia coding genome: Role of NOTCH1 mutational activation. *Journal of Experimental Medicine*. 2011;208(7):1389–1401.

43. Morin RD, Mendez-Lago M, Mungall AJ, et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature*. 2011;476(7360):298–303.

44. West AG, Shore P, Sharrocks AD. DNA binding by MADS-box transcription factors: A molecular mechanism for differential DNA bending. *Molecular and Cellular Biology*. 1997;17(5):2876–2887.

45. Ying CY, Dominguez-Sola D, Fabi M, et al. MEF2B mutations lead to deregulated expression of the BCL6 oncogene in Diffuse Large B cell Lymphoma. *Nature immunology*. 2013;14(10):1084–1092.

46. Brescia P, Schneider C, Holmes AB, et al. MEF2B Instructs Germinal Center Development and Acts as an Oncogene in B Cell Lymphomagenesis. *Cancer Cell*. 2018;34(3):453–465.e9.

47. Lei X, Kou Y, Fu Y, et al. The Cancer Mutation D83V Induces an α-Helix to β-Strand Conformation Switch in MEF2B. *Journal of Molecular Biology*. 2018;430(8):1157–1172.

48. Pon JR, Wong J, Saberi S, et al. *MEF2B* mutations in non-Hodgkin lymphoma dysregulate cell migration by decreasing MEF2B target gene activation. *Nature Communications*. 2015;6:7953.

49. Lin Y-T, Yen PH. A novel nucleocytoplasmic shuttling sequence of DAZAP1, a testis-abundant RNA-binding protein. *RNA*. 2006;12(8):1486–1493.

50. Dai T, Vera Y, Salido EC, Yen PH. Characterization of the mouse Dazap1 gene encoding an RNA-binding protein that interacts with infertility factors DAZ and DAZL. *BMC Genomics*. 2001;2:6.

51. Choudhury R, Roy SG, Tsai YS, et al. The splicing activator DAZAP1 integrates splicing control into MEK/Erk-regulated cell proliferation and migration. *Nature Communications*. 2014;5:3078.

52. Grande BM, Gerhard DS, Jiang A, et al. Genome-wide discovery of somatic coding and non-coding mutations in pediatric endemic and sporadic Burkitt lymphoma. *Blood*. 2019;blood–2018–09–871418.

53. Yang P, Zhang W, Wang J, et al. Genomic landscape and prognostic analysis of mantle cell lymphoma. *Cancer Gene Therapy*. 2018;25(5):129–140.

54. Grünewald TGP, Cidre-Aranaz F, Surdez D, et al. Ewing sarcoma. *Nature Reviews Disease Primers*. 2018;4(1):5.

55. Bertolotti A, Melot T, Acker J, et al. EWS, but Not EWS-FLI-1, Is Associated with Both TFIID and RNA Polymerase II: Interactions between Two Members of the TET Family, EWS and hTAFII68, and Subunits of TFIID and RNA Polymerase II Complexes. *Molecular and Cellular Biology*. 1998;18(3):1489–1497.

56. Knoop LL, Baker SJ. The Splicing Factor U1C Represses EWS/FLI-mediated Transactivation. *Journal of Biological Chemistry*. 2000;275(32):24865–24871.

57. Dutertre M, Sanchez G, De Cian M-C, et al. Cotranscriptional exon skipping in the genotoxic stress response. *Nature Structural & Molecular Biology*. 2010;17(11):1358–1366.

58. Paronetto MP. Ewing Sarcoma Protein: A Key Player in Human Cancer. *International Journal of Cell Biology*. 2013;

59. Sanchez G, Delattre O, Auboeuf D, Dutertre M. Coupled alteration of transcription and splicing by a single oncogene: Boosting the effect on cyclin D1 activity. *Cell Cycle (Georgetown, Tex.)*. 2008;7(15):2299–2305.

60. Akindahunsi AA, Bandiera A, Manzini G. Vertebrate 2xRBD hnRNP proteins: A comparative analysis of genome, mRNA and protein sequences. *Computational Biology and Chemistry*. 2005;29(1):13–23.

61. Chen H-Y, Yu Y-H, Yen PH. DAZAP1 regulates the splicing of Crem, Crisp2 and Pot1a transcripts. *Nucleic Acids Research*. 2013;41(21):9858–9869.

62. Pastor T, Pagani F. Interaction of hnRNPA1/A2 and DAZAP1 with an Alu-Derived Intronic Splicing Enhancer Regulates ATM Aberrant Splicing. *PLoS ONE*. 2011;6(8):e23349.

63. Kurihara Y, Watanabe H, Kawaguchi A, et al. Dynamic changes in intranuclear and subcellular localizations of mouse Prrp/DAZAP1 during spermatogenesis: The necessity of the C-terminal proline-rich region for nuclear import and localization. *Archives of Histology and Cytology*. 2004;67(4):325–333.

64. Lin Y-T, Wen W-C, Yen PH. Transcription-dependent nuclear localization of DAZAP1 requires an N-terminal signal. *Biochemical and Biophysical Research Communications*. 2012;428(3):422–426.

65. Geuens T, Bouhy D, Timmerman V. The hnRNP family: Insights into their role in health and disease. *Human Genetics*. 2016;135:851–867.

66. Wang E, Aslanzadeh V, Papa F, et al. Global profiling of alternative splicing events and gene expression regulated by hnRNPH/F. *PloS One*. 2012;7(12):e51266.

67. Hong S. RNA Binding Protein as an Emerging Therapeutic Target for Cancer Prevention and Treatment. *Journal of Cancer Prevention*. 2017;22(4):203–210.

68. Quesada V, Conde L, Villamor N, et al. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nature Genetics*. 2011;44(1):47–52.

69. Rossi D, Bruscaggin A, Spina V, et al. Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: Association with progression and fludarabine-refractoriness. *Blood*. 2011;118(26):6904–6908.

70. Wang L, Lawrence MS, Wan Y, et al. SF3B1 and Other Novel Cancer Genes in Chronic Lymphocytic Leukemia. *New England Journal of Medicine*. 2011;365(26):2497–2506.

71. Wang L, Brooks AN, Fan J, et al. Transcriptomic Characterization of SF3B1 Mutation Reveals Its Pleiotropic Effects in Chronic Lymphocytic Leukemia. *Cancer Cell*. 2016;30(5):750–763.

72. Hacken E ten, Valentin R, Regis FFD, et al. Splicing modulation sensitizes chronic lymphocytic leukemia cells to venetoclax by remodeling mitochondrial apoptotic dependencies. *JCI Insight*. 3(19):

73. Raa GD te, Derks IaM, Navrkalova V, et al. The impact of *SF3B1* mutations in CLL on the DNA-damage response. *Leukemia*. 2015;29(5):1133–1142.

74. LeFave CV, Squatrito M, Vorlova S, et al. Splicing factor hnRNPH drives an oncogenic splicing switch in gliomas. *The EMBO Journal*. 2011;30(19):4084–4097.

75. Golan-Gerstl R, Cohen M, Shilo A, et al. Splicing factor hnRNP A2/B1 regulates tumor suppressor gene splicing and is an oncogenic driver in glioblastoma. *Cancer Research*. 2011;71(13):4464–4472.