Characterization of *SF3B1* Mutation in Chronic Lymphocytic Leukemia and the Role of New Splicing Modulators

Irene López Oreja
Doctoral Thesis 2021

Department of Experimental and Health Sciences
Universitat Pompeu Fabra

Under the supervision of Dr. Dolors Colomer
Experimental therapies in lymphoid neoplasms
Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS)
Drs. Sophie Bonnal and Juan Valcárcel
Gene Regulation, Stem Cells and Cancer Program
Centre for Genomic Regulation (CRG)

A tots els que m’heu fet costat en aquest camí.

ABSTRACT

Mutations in Splicing Factor 3b Subunit 1(*SF3B1*) are observed in 15% of patients with chronic lymphocytic leukemia (CLL) and are associated with worse disease prognosis. SF3B1 is involved in pre-mRNA splicing and is the target of anti-tumor drugs. Here, we analyzed RNA-sequencing data from 298 CLL patients and $SF3B1^{WT}$ and $SF3B1^{K700E}$ MEC1 CLL isogenic cell lines and characterized the pre-mRNA sequence features associated with the selection of cryptic 3'splice site (3'ss) upon *SF3B1* mutation. In addition, we validated the effects of MAP3K7 alternative splicing events in *SF3B1*-mutated CLL, and confirmed the activation of NF-κB signaling in this subgroup. Finally, we carried out the first study of the effects of the H3B-8800 splicing modulator in CLL, *in vitro* in primary CLL samples and in MEC1 WT and *SF3B1*-mutated isogenic cell lines, as well as *in vivo* in an immunodeficient NOD-SCID interleukin-2 receptor gamma $(IL2R\gamma)^{null}$ (NSG) mice model. We also characterized the transcriptomic effects of H3B-8800 treatment of MEC1 cell lines by RNA-seq. H3B-8800 showed preferential lethality towards *SF3B1*-mutated cells and synergistic effects with BCL2 inhibitor venetoclax, supporting the use of SF3B1 inhibitors as a possible novel therapeutic strategy in CLL.

RESUMEN

El 15% de los pacientes con leucemia linfática crónica (LLC) presentan mutaciones en *SF3B1* (Splicing Factor 3b Subunit 1), que les confieren un peor pronóstico. SF3B1 es una proteína implicada en pre-mRNA splicing y es la diana de drogas antitumorales. En este trabajo analizamos resultados de RNA-seq de 298 pacientes con LLC y de líneas celulares de LLC isogénicas MEC1 *SF3B1*^{WT} *y* *SF3B1*^{K700E} caracterizando secuencias del pre-mRNA implicadas en el uso de sitios 3' de splicing crípticos activados por mutaciones en *SF3B1*. Adicionalmente, validamos el efecto de uno de estos 3'ss crípticos en el gen *MAP3K7*, confirmando la activación que produce en la vía de señalización de NF-κB en el subgrupo que presenta mutaciones en *SF3B1*. Finalmente, probamos por primera vez en LLC los efectos del modulador de splicing H3B-8800, *in vitro* en células primarias de LLC y en la línea celular isogénica *SF3B1*^{K700E} MEC1, así como *in vivo*, en el modelo de ratón inmunodeprimido NOD-SCID interleukin-2 receptor gamma (*IL2Rγ*)^{null} (NSG). H3B-8800 muestra una letalidad preferencial hacia las células mutadas en SF3B1 y un efecto sinérgico en combinación con el inhibidor de BCL2 venetoclax, sugiriendo que inhibidores de SF3B1 pueden proporcionar una nueva estrategia terapéutica en la LLC.

PREFACE

Mutations in the Splicing Factor 3 b subunit 1 (SF3B1) are associated with worse prognosis in CLL. We aimed to characterize the molecular effects of these mutations on alternative splicing as a way to understand their pathogenic impact in CLL and, eventually design therapeutic approaches to overcome these adverse prognosis.

In addition, in the last years SF3B1 has emerged as the target of several anti- tumor splicing modulator drugs, including the new orally available modulator H3B-8800, which has been shown to induce preferential lethality in cancer cells displaying splicing factor mutations. In this Thesis we had the opportunity to test its efficacy for the first time in CLL, contributing to the search of therapeutic alternatives for this disease.

TABLE OF CONTENTS

ABSTRACT ..... *vii*

RESUMEN ..... *viii*

PREFACE ..... *ix*

TABLE OF CONTENTS ..... *x*

INTRODUCTION *3*

1. Pre-mRNA Splicing **3**

2. Alternative Splicing **7**

2.1. Role of alternative splicing in cancer 9

2.1.1. *SF3B1* mutations in cancer 11

2.1.2. Mechanism behind cryptic 3'ss usage in cells harboring SF3B1 mutations 13

2.1.3. Biological impact of the use of cryptic 3'ss 14

2.1.4. SF3B binding drugs for cancer therapy 15

3. Chronic Lymphocytic Leukemia **20**

3.1. Definition and epidemiology 20

3.2. IGHV Rearrangements and Somatic Hypermutations in B cell development 21

3.3. CLL subtypes 24

3.3.1. Immunogenetic subtypes 24

3.3.2. Epigenetic subtypes 24

3.4. CLL genetics 26

3.4.1. Chromosomal alterations 26

3.4.2. Mutational landscape 26

3.5. Treatment 27

3.6. *SF3B1* mutated CLL 29

AIMS *31*

RESULTS *35*

Chapter I: **37**

Transcriptomic Characterization of *SF3B1*- Mutated Chronic Lymphocytic Leukemia **37**

1. Establishment and characterization of *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 CLL isogenic cell lines 39

1. Transcriptomic analysis of *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 CLL cell lines and CLL patients' samples 43
1.1. CLL patients selection for transcriptomic analysis 43
1.2. AS events found in CLL patients and *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 CLL isogenic cell lines 49
1.3. Alt3'ss patterns related to *SF3B1*^{K700E} mutation 53
1.3.1. Validation of the Alt3'ss patterns associated with *SF3B1*^{K700E} mutation by RT-PCR 58
1.3.2. Validation of the Alt3'ss patterns associated with *SF3B1*^{K700E} mutation using minigenes 61
1.4. Gene enrichment analysis of the differentially spliced AS associated with *SF3B1* mutations 68
1.5. Possible use of ZDHHC16 AS as a biomarker to determine the functional impact of *SF3B1* mutations 69
1.6. Gene expression analysis 71
2. Functional relevance of selected AS events 74
2.1. Role of MAP3K7 alternative spliced isoforms 74
Validation of MAP3K7 Alt3'ss activation in CLL patient samples and *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 CLL isogenic cell lines 75
2.2. Effects of MAP3K7 aberrant splicing in NF-κB pathway 76
2.3. Effects of MAP3K7 aberrant splicing on p38 phosphorylation 78

Chapter II: 81

Study of the therapeutic effects of SF3B1-binding H3B-8800 in *in vitro* and *in vivo* models of Chronic Lymphocytic Leukemia 81
1. Cytotoxic effects of H3B-8800 in models of Chronic Lymphocytic Leukemia 83
1.1. Sensitivity to H3B-8800 in SF3B1-mutated CLL primary samples and cell lines 84
1.1.1. H3B-8800 effects in co-culture conditions 86
1.1.2. H3B-8800 delays leukemic infiltration in NSG *in vivo* mice model 87
2. Transcriptomic analysis of H3B-8800-treated *SF3B1* ^{WT} and *SF3B1* ^{K700E} MEC1 CLL cell lines 94
2.1. Alternative splicing analysis 94
2.2. Gene expression analyses 99
3. H3B-8800 modulation of alternative splicing events from genes related to apoptosis 103
3.1. Validation of MCL-1 AS event 103
3.2. H3B-8800 shows synergistic effects with venetoclax 105

DISCUSSION 109

Characterization of Transcriptomic Changes Associated with Mutations in the Splicing Factor *SF3B1* in Chronic Lymphocytic Leukemia 110

The SF3B1-Binding Splicing Inhibitor H3B-8800 in the Treatment of Chronic Lymphocytic Leukemia 119

CONCLUSIONS 127

MATERIALS AND METHODS 131

1. Cell lines 133
1.1. Cell culture 133
1.2. Establishment of CRISPR Cas9-edited cell lines 134
1.2.1. Design 134
1.2.2. Procedure 134
1.2.3. DNA extraction and quantification of *SF3B1* K700E mutation status using digital PCR: 136
1.2.4. Characterization of CRISPR cell lines using cell viability and MTT assay 136

2. Human samples 137

3. RNAseq analysis 137
3.1. ICGC Patients 137
3.2. H3B-8800 treated cell lines 138
3.3. Splicing analyses 138
3.3.1. Venn diagrams 139
3.3.2. Correlation analysis 139
3.3.3. Prediction of the impact on protein sequence of the alternative splicing events 139
3.4. Gene expression analysis 140
3.5. Principal Component Analysis (PCA) 140
3.6. Gene Ontology enrichment analysis 140

4. Experimental validation of RNA-seq data 141
4.1. RNA extraction and reverse transcription, semi-quantitative RT-PCR 141
4.2. Quantitative real-time PCR (RT-qPCR) 142
4.3. Fragment analysis 143
4.4. Minigenes 143
4.4.1. Cloning 143
4.4.2. Transfections 144
4.5. Western blots 146
4.6. NF-κB activity measurements 148

5. H3B-8800 treatment 148
5.1. H3B-8800-induced cytotoxicity studied by flow cytometry 148
5.2. Co-culture experiments 148

5.3. Synergy experiments with venetoclax..... 149
5.4. Stable expression of luciferase and GFP..... 149
5.5. *In vivo* model..... 149
5.5.1. Experimental design 149
5.5.2. Samples processing 150
6. Statistical analysis..... 150
ABBREVIATIONS..... 151
REFERENCES 153
ANNEX..... 181
Supplementary Tables 181
Papers published during the Thesis..... 183
ACKNOWLEDGEMENTS..... 1

INTRODUCTION

1. Pre-mRNA Splicing

The information encoded in the genes is transferred through the process of DNA transcription into messenger RNA (mRNA) molecules and then the nucleotide sequence information of mRNAs is translated into amino acid sequence information in the form of proteins[1]. In eukaryotic organisms, primary RNA polymerase II transcripts (pre-mRNA) undergo multiple processing steps, including pre-mRNA splicing, which removes non-coding parts known as introns and splices together the remaining sequences, known as exons, to generate mature mRNAs that will be exported to the cytoplasm and translated into protein[2]. Splicing takes place in the nucleus and can occur during the process of transcription (co-transcriptional splicing) or after the pre-mRNA has been transcribed (post-transcriptional splicing). Splicing is carried out by the spliceosome, an extremely complex machinery composed by five small nuclear Ribonucleoprotein particles (snRNPs) and more than 250 proteins[3]. The spliceosome recognizes specific sequences at the intron boundaries, namely the 5' splice site (5'ss) defined by a conserved GU sequence at the very 5' end of the intron, and the 3' splice site (3'ss) consisting of the branch point (BP) adenosine (located 18-35 nucleotides upstream the 3'ss), the polypyrimidine tract (Py tract) and the dinucleotide AG at the very 3' end of the intron. Both sites allow de novo spliceosome assembly at each intron[3,4] (Fig 1A).

Splicing occurs through two transesterification steps. First, the oxygen atom of the 2'hydroxyl group of the intronic BP adenosine

carries out a nucleophilic attack on the phosphate group between the 5’ exon and the intron, generating a free OH at the 3’ end of the 5’ exon and a lariat intermediate with a 2’-5’ phosphodiester bond (Fig 1B). Second, the free 3’OH of the 3’ end of the 5’ exon attacks the phosphate between the intron and the 3’ exon, releasing the intron lariat and ligating the exons[5] (Fig 1B).

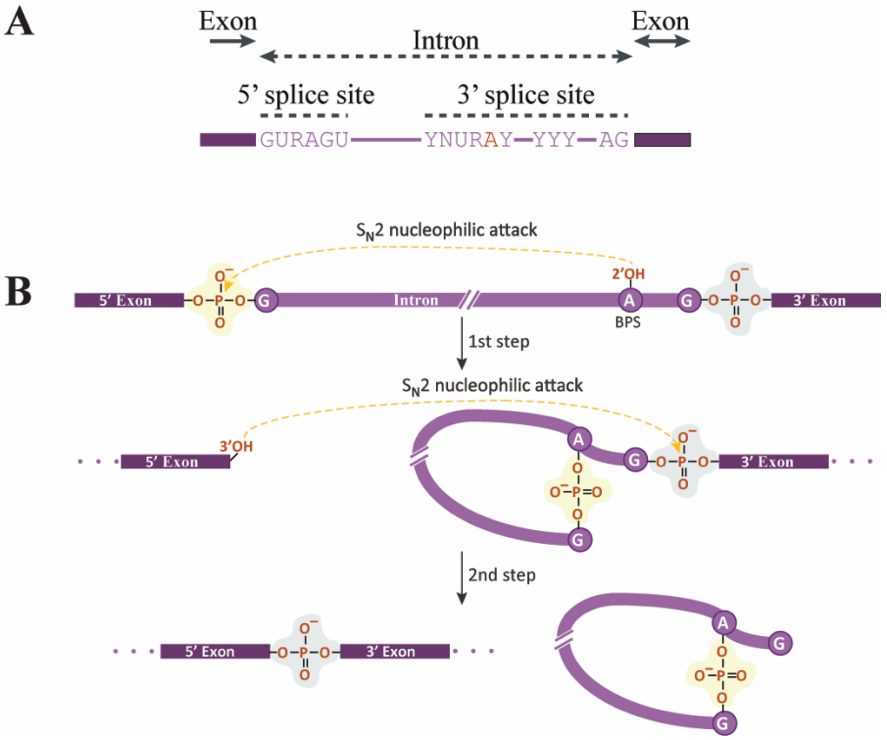


Figure 1. Intronic sequences recognized by the spliceosome and the chemical process of the splicing reaction (A) The 5’ss sequence and the 3’ss sequence are shown with the BP adenosine highlighted in red and the polypyrimidine tract (YYY)(modified from[6]). (B). In the first nucleophilic attack, the 2’OH from the BP adenosine binds to the phosphate between the 5’exon and the intron creating an intermediate intron lariat harboring an unusual 2′–5′ phosphodiester bond. In the second nucleophilic attack, the free 3’OH at the 5’ exon binds to the phosphate between the intron and the 3’ exon releasing the intron and splicing the two exons (from[4]).

The two transesterification steps are catalyzed by RNA molecules, the catalytic core being established by a structural framework made

of interactions between RNA components of the U2, U6 and U5 particles, resembling the catalytic core of group II self-catalytic RNAs[4]. During the initial step of the spliceosome assembly, namely complex E formation, U1 snRNP recognizes the 5'ss while the 3 elements of the 3'ss are recognized by the splicing factor (SF1) and the heterodimer U2AF, with U2AF2 binding to the Py tract and U2AF1 recognizing the AG dinucleotide (Fig 2, complex E) [3,7–10]. Then, U2 snRNP recruitment takes place through U2 snRNA base pairing interactions with nucleotides flanking the BP adenosine, protein-protein and RNA-protein interactions. Splicing Factor 3b subunit 1 (SF3B1), a core splicing factor which is part of the U2 snRNP, helps in the recognition of the BP by U2 snRNA, specifically recognizing the BP Adenosine base which bulges out from the U2:pre-mRNA helix [11,12](Fig 2, complex A). SF3B1 interacts with other SF3B complex subunits, and also with U2AF2, which recognizes the Py tract and also stabilizes the base pairing interaction between U2 snRNP and the BP through its positively charged RS (arginine-serine-rich) domain [12]. Structurally, SF3B1 amino-terminal region is non-structured and involved in interactions with other splicing factors, whereas the carboxy-terminal contains 22 HEAT (huntingtin elongation factor 3 protein phosphatase 2A, target of rapamycin 1) domain repeats, which fold into a shell-like structure that embraces the U2:pre-mRNA helix and holds the BP adenosine in a pocket[13].

In the next step of spliceosome assembly, U4/U6-U5 tri-snRNP associates forming complex B and several conformational rearrangements including shifts in snRNAs base pairing occur,

leading to the catalytically active spliceosome complexes (B*, B active, C and C*). After the splicing reaction, the spliceosome disassembles and is recycled for a new assembly process on a different intron, the spliced mRNA is released and the lariat is degraded[14,15] (Fig. 2).

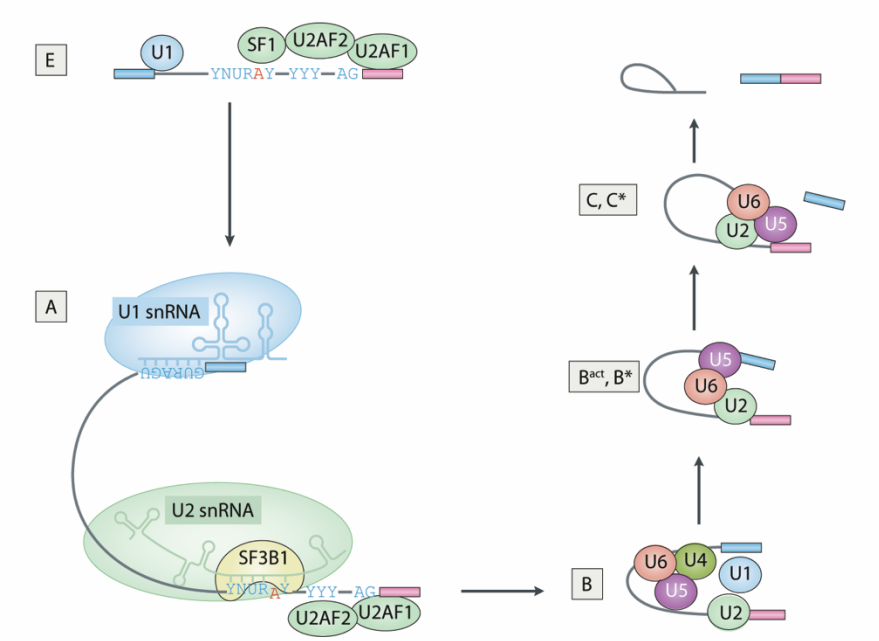


Figure 2. The Spliceosome assembly process. The spliceosome, composed of five small nuclear ribonucleoproteins (snRNPs) and numerous additional proteins, assembles anew on each pre- mRNA molecule through sequential steps involving the dynamic remodeling of its composition and conformation mediated by changes in protein–protein, protein–RNA and RNA–RNA interactions. U1 snRNP is recruited through base pairing between the 5′ splice site and U1 small nuclear RNA (snRNA), whereas SF1, U2AF2 and U2AF1 recognize the BP sequence, the polypyrimidine tract and the AG dinucleotide of the 3′ splice site, respectively, leading to the formation of complex E. U2 snRNP is then recruited to the BP through base pairing interactions with U2 snRNA, leading to complex A, in which the U2 snRNP component SF3B1 facilitates the recognition of the intron BP by U2 snRNA. The subsequent recruitment of the U4/U6–U5 tri- snRNP complex to form complex B and the remodeling of numerous interactions within complex B lead to the formation of the catalytically active conformations of the spliceosome (referred to as Bact, B*, C and C* complexes). BP adenosine is highlighted in red, the name of the complexes are indicated in the grey boxes. Bact, B active; R, purine; Y, pyrimidine; N, any nucleotide (from[6]).

Besides the major spliceosome, less than 0.05% of the introns, characterized by distinct highly conserved 5' splice site and BP sequences, are excised by the minor spliceosome. One of the differences from the major spliceosome is the presence of U11 and U12 snRNPs, which replace U1 and U2 snRNPs in recognizing the 5′ and 3′ splice sites, respectively[16,17]. Of note, SF3B1 is part of both spliceosomes.

2. Alternative Splicing

Alternative splicing is a crucial post-transcriptional mechanism of gene regulation, which allows a significant increase of protein diversity in eukaryotes by generating more than one mRNA isoform from a single gene. Different types of alternative splicing exist, namely cassette exon skipping (ES), intron retention (IR), alternative 5'ss (Alt5'ss), alternative 3'ss (Alt5'ss) and mutually exclusive exons (Fig 3). The regulation of alternative splicing site selection is mostly – but not only – attributed to cis-regulatory sequences and to trans-acting regulatory factors that bind to the cis-regulatory sequences and promote or repress the recognition of splice sites by the core splicing factors, determining the final relative abundance of different transcript isoforms. Depending on the localization and functionality of these sequences, 4 types of sequences are defined: intronic splicing enhancers (ISE), intronic splicing silencers (ISS), exonic splicing enhancers (ESE) and exonic splicing silencers (ESS). A variety of RNA binding proteins, including SR proteins, heterogeneous nuclear RNP (hnRNP) proteins, CELF, RBM and

other protein families recognize these sequences and promote or prevent the recognition of a 5'ss and/or a 3'ss[18](Fig 4).

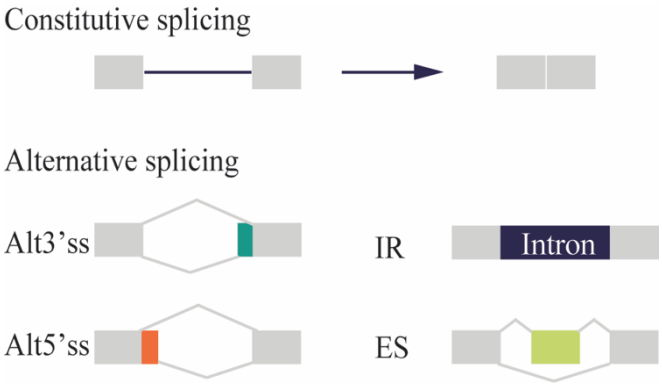


Figure 3. Alternative splicing event types. Exons are represented in rectangles, introns with a dark line. Alt3'ss, alternative 3'splice site; Alt5'ss, alternative 5'splice site; IR, intron retention; ES, exon skipping.

Alternative splicing contributes to very diverse biological processes of vital importance in humans, including neurogenesis, gametogenesis, muscle generation and contractility, B cells development or B and T cells reprogramming[19,20]. Consistent with its crucial role, splicing misregulation is the cause (or a modifier) of a wide variety of pathological disorders. Mutations in cis-acting factors, trans-acting factors or nucleotide repeat expansions (sequestering trans-acting factor regulators) contribute to the pathogenicity of muscular dystrophies, metabolic syndromes, neurodegenerative and cardiovascular diseases as well as cancer[21,22].

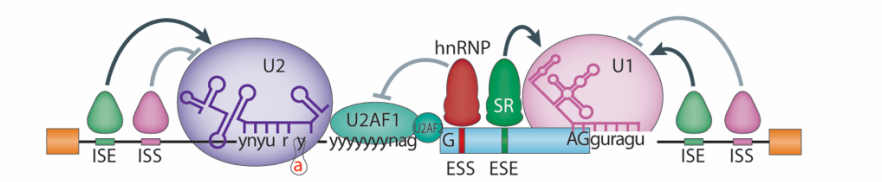


Figure 4. Consensus sequences recognition by core splicing factors U1 and U2 snRNPs and U2AF1/2 is modulated by cis-elements (ISE, intronic splicing enhancers; ISS, intronic splicing silencers; ESE, exonic splicing enhancers and ESS, exonic splicing silencers) and trans-acting factors (for example serine – arginine rich (SR) proteins and heterogeneous nuclear ribonucleoproteins (hnRNPs). The BP adenosine is highlighted in red. Modified from[22].

2.1. Role of alternative splicing in cancer

Cancer cells have AS events not present in non-malignant tissues[23]. Mutations, changes in expression and post-translational modifications of splicing factors or auxiliary proteins as well as mutations in snRNAs are different factors contributing to alterations of splicing profiles in tumors and some of these correlate with prognosis. Misregulation of splicing affects genes from all cancer's hallmarks, leading to consider alternative splicing misregulation itself as an additional hallmark[6,24,25].

Whole exome sequencing (WES) analysis across 33 tumor types from The Cancer Genome Atlas (TCGA) revealed that 119 splicing factor genes carry putative driver mutations in one or more cancer types[26] with specially high frequency in hematological malignancies[27] (Fig 5). These driver mutations are mostly present in proteins involved in the early stages of splicing recognition catalysis (complex E to A) [26]. The most recurrently mutated genes are *SF3B1*, *U2AF1*, as well at the SR protein *SRSF2* known to bind specific ESEs and stimulate recognition of the adjacent splice site by U1 or U2 snRNPs, and the splicing factor *ZRSR2* that is part of the

minor spliceosome and recognizes the 3'ss[6,27,28]. Mutations in these splicing factors are associated with specific patterns of alternative splice site selection. For example, differential inclusion of cassette exons and U12-type introns retention are observed upon mutation of *SRSF2* and *ZRSR2*, respectively[29,30].

Missense mutations in *SF3B1*, *U2AF1* and *SRSF2* occur in hotspot locations, while loss of function mutations in *ZRSR2* are distributed throughout the gene[31]. They usually appear in heterozygosis and are mutually exclusive. Two potential reasons for this mutual exclusivity could be that they have redundant downstream effects, or they end up producing synthetic lethal effects when combined in the same cell. Along this line, there is evidence that co-expression of hotspot *SF3B1*^{K700E} and *SRSF2*^{P95H} mutations display synthetic lethal interactions in hematopoietic cells in mice and they converge in promoting nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) signalling to drive myelodysplastic syndrome (MDS)[32]. Moreover, recent single cell analysis on myeloid malignancies have shown that in the rare cases where co-occurrence is detected, hotspot mutations are not present but rather involve less common mutations, which have less effects on splicing or RNA binding, reinforcing the notion of mutual exclusivity[33].

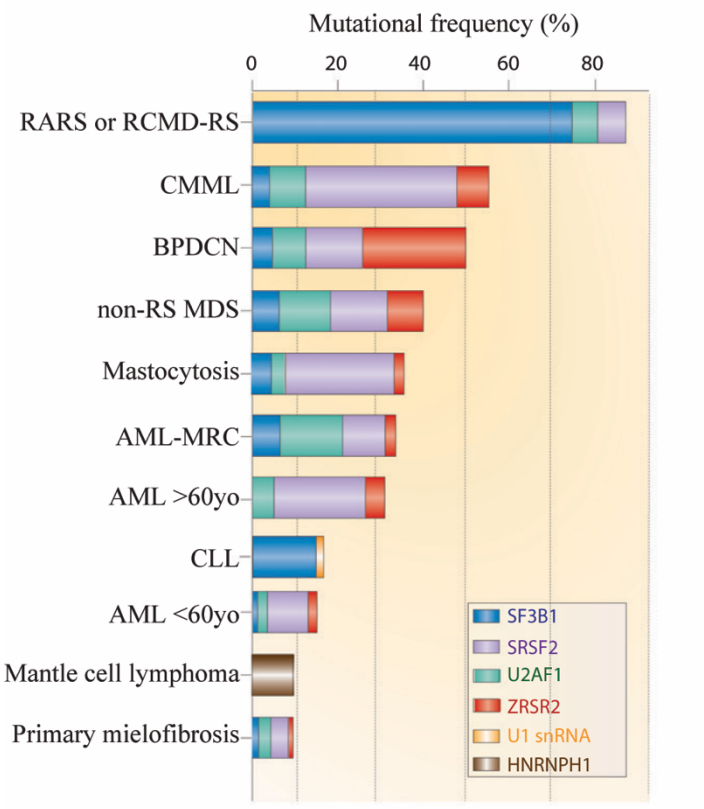


Figure 5. Frequency of mutations in the most commonly mutated RNA splicing components across hematologic malignancies. RARS, refractory anemia with ring sideroblasts; RCMD-RS, refractory cytopenia with multilineage dysplasia and ring sideroblasts; CMML, chronic myelomonocytic leukemia; BPDCN, blastic plasmacytoid dendritic cell neoplasm; non-RS MDS, non-ring syderoblast myelodysplastic syndrome; AML-MRC, acute myeloid leukemia with myelodysplasia-related changes; CLL: chronic lymphocytic leukemia; yo, years old (From [28]).

2.1.1. *SF3B1* mutations in cancer

SF3B1 mutations in cancer appear from the 4th to the 12th HEAT domains, hotspots being located between the 4th and 8th HEAT domains[26,34,35](Fig 6). The SF3B1 residues where the mutations are located contribute to the tertiary structure of the HEAT superhelix and to surface residues in the vicinity of regions of interaction with

SF3B6 (a subunit of SF3b complex) and U2AF2, suggesting they may influence SF3B1 interactions with the them[35]. However, the hotspot K700E mutation does not affect the stability of the SF3b-U2AF2 complex neither the affinity to RNA, suggesting the mechanism is more complex[35]. Recent work suggests that *SF3B1* mutations destabilize intron binding and affect the functional dynamics of SF3B1 only when binding to non-consensus or degenerated BPs[36].

SF3B1 mutations are not only present in hematological malignancies (Fig 5), but also in some solid tumors such as uveal melanoma, mucosal melanoma, pancreatic adenocarcinoma, breast cancer and bladder cancer[37–40]. Strikingly, they correlate with good prognosis in MDS and uveal melanoma but with bad prognosis in CLL[41–44].

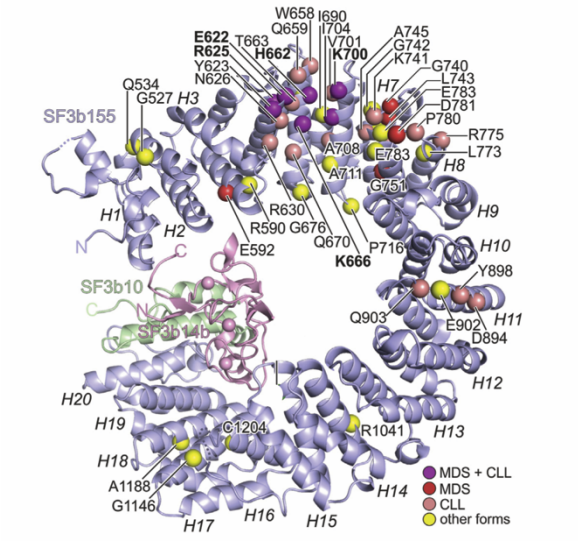


Figure 6. SF3B1 residues mutated in CLL and/or MDS and other cancers. SF3B complex is formed of 7 subunits (SF3B1-SF3B7) of which SF3B1 (also known as SF3b155 in purple), SF3B6 (also known as SF3b14a or p14 in pink) and SF3B7

(also known as SF3b10 or PHF5A in green) are shown in the figure. Mutational hot spot residues in SF3B1 are in bold, SF3B1's HEAT domains are shown with an "H" and their corresponding number. From[35]

2.1.2. Mechanism behind cryptic 3'ss usage in cells harboring SF3B1 mutations

SF3B1 mutations lead to activation of cryptic 3'ss. These sites are usually located 15 to 24 nucleotides upstream of the canonical 3'ss and have a short and weak polypyrimidine tract[45]. Furthermore, they are typically located within potential RNA secondary structures and poorly accessible compared with the canonical 3'ss they are associated with[46]. Interestingly, different experimental approaches, including the use of minigenes in several systems, have shown that different BP are used upon expression of different SF3B1 mutants (ectopic expression of mxSF3B1 in 293FT cells [47] [45](Fig 7), ectopic expression of $SF3B1^{R625G}$ in uveal melanoma cell lines or isogenic expression of $SF3B1^{K666T}$ in HEK293T cells[48]). Moreover, a synthetic minigene library in isogenic murine embryonic stem cells harboring a monoallelic K700E mutation has demonstrated that $SF3B1^{K700E}$ preferentially uses non-canonical BP which are enriched in BP that differed from the consensus YTAAY motif at position -1[49].

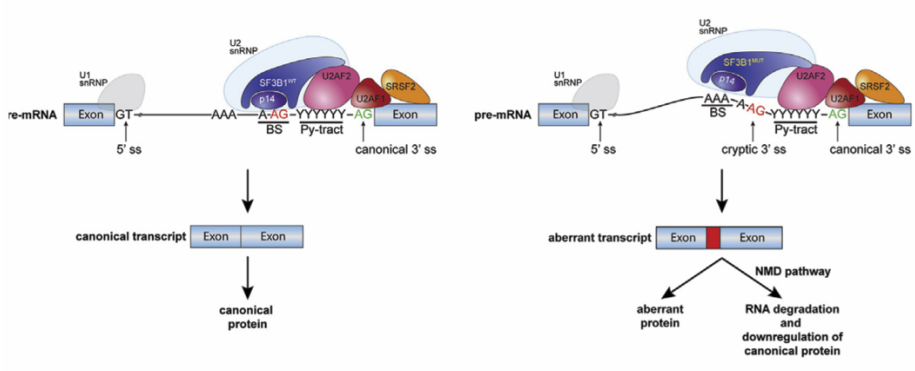


Figure 7. Schematic representation of AS in *SF3B1* WT and mutated cases. SF3B1 WT (left panel) binds to the canonical BP (BS) whereas SF3B1 mutant (right panel) binds to an upstream BP, allowing the selection of an upstream cryptic 3'ss and the formation of an aberrant transcript. Around 50% of the aberrant transcripts undergo nonsense-mediated decay (NMD) with the consequent RNA degradation and down-regulation of the canonical protein. P14 is also known as SF3B6(from[45]).

Insights into the mechanism by which *SF3B1* mutations interfere with BP recognition and alternative 3'ss selection have been provided by recent publications demonstrating that the levels of the SUGP1 spliceosomal protein decrease in the presence of *SF3B1* mutations and that this decrease, as well as mutations in SUGP1 found in a pan-cancer analysis, recapitulate splicing defects associated with *SF3B1* mutations [50,51]. In addition, these studies showed that SUGP1 interacts with U2AF2 and the HEAT domain of SF3B1 and that this interaction is disrupted upon *SF3B1* mutation. However, it is still not clear how these changes impact on BP selection.

2.1.3. Biological impact of the use of cryptic 3'ss

Approximately 50% of aberrant 3'ss activated by *SF3B1* mutations disrupt the open reading frame and generate premature stop codons (PTC), leading to nonsense-mediated decay (NMD) and

downregulation of gene and protein expression[45,52], while use of other cryptic 3'ss lead to changes in protein sequences. An interesting example of the latter is observed in the SF3B1 transcript itself. Indeed, in *SF3B1*-mutated breast cancer, uveal melanoma and MDS, a cryptic 3'ss in *SF3B1* gene has been identified which results in the insertion of eight amino acids (LLLFSLFQ) in the H3 repeat and renders *SF3B1* inactive for its function in splicing. However, its low abundance (less than 10% of the SF3B1 total transcripts) suggests a limited effect on the physiopathology of these diseases[53].

In contrast, the contribution of the usage of other cryptic 3'ss in disease is clearer. For example, aberrant 3'ss in the iron transporter ABCB7 generates a PTC which induces NMD mRNA decay and downregulation of ABCB7, responsible for iron accumulation in refractory anemia with ring sideroblasts characteristic of a group of MDS patients [54–56]. In addition, aberrant 3'ss in MAP3K7 gene, a mitogen-activated protein kinase, also produces a PTC susceptible of NMD and likely contributes to anemia in MDS by downregulation of GATA1, a master regulator of erythroid differentiation[57].

In clinical practice, *SF3B1* mutation in MDS identifies an entity characterized by ring sideroblasts, inefficient erythropoiesis and indolent clinical course, that is proposed to be considered as a distinct nosology entity[58].

2.1.4. SF3B binding drugs for cancer therapy

Different approaches are available to therapeutically target splicing at different levels of gene expression regulation. These include (i) inhibitors of epigenetic regulators, (ii) modulators of

signalling pathways regulating RNA binding proteins and splicing regulatory factors, such as inhibitors of SR protein phosphorylation, (iii) modulators which bind in the interface between SF3B1 and PHF5A components of the core SF3B complex, (iv) drugs targeting protein–RNA or protein–protein interactions that affect splice site availability, (v) antisense oligonucleotides that induce switches in splice site usage or that target specific mRNA isoforms and (vi) antibodies inhibiting aberrant protein isoforms[6,59,60].

SF3B1 binding molecules were initially identified in screens of bacterial fermentation products displaying anti-tumor properties and have evolved as attractive therapeutic possibilities for modulating splicing and proliferation in cancer cells. Three families of bacterial fermentation products and their synthetic derivatives have been studied, harboring a common pharmacophore: (i) FR901464 and synthetic spliceostatin A, meayamycin and thailanstatins; (ii) pladienolide B and its derivatives E7107, H3B-8800 and FD-895 and (iii) GEX1 and synthetic herboxidiene (Fig. 8). The design of these synthetic derivatives aimed to obtain chemically more stable products, improve the balance between efficacy and toxicity as well as to produce molecules with specificity for mutated *SF3B1*[61–63].

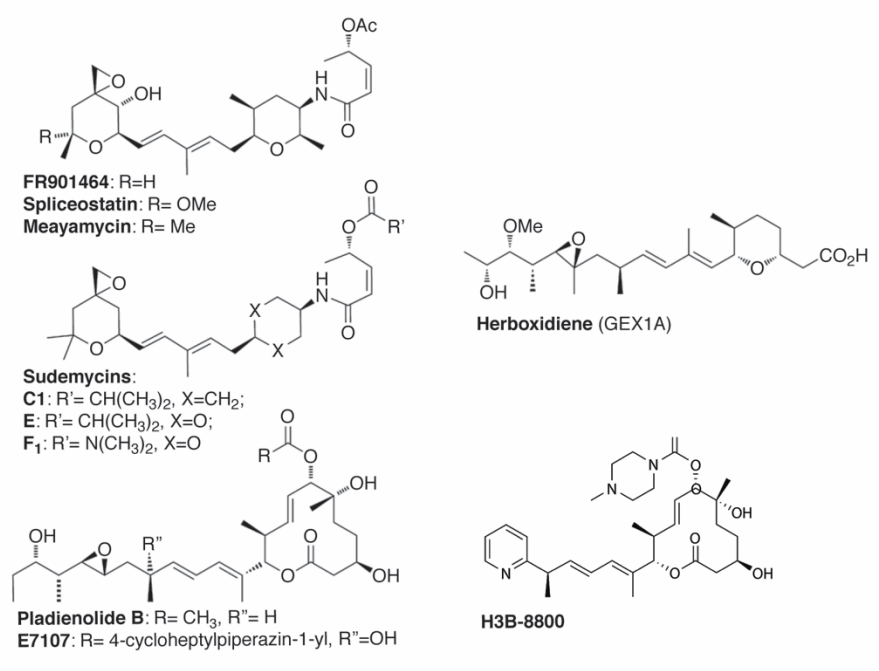


Figure 8. Molecular structures of SF3B1 binding drugs. Elements of the common pharmacophore include a conjugated diene, oxocarbonyl and peroxide groups.

Recent structural studies show how pladienolide B, E7107 and H3B-8800 interact with specific residues in the 15th to 17th HEAT domain repeats of SF3B1 and residues in the PHF5A protein (SF3B7) that jointly configure the BP adenosine-binding pocket, blocking the transition from the open to the closed conformation needed for BP adenosine recognition and subsequent 1st catalytic step of the splicing process[64–67](Fig 9) Furthermore, mutations R1074H, V1078A, V1078I in *SF3B1* and mutation Y36C in *PHF5A*, located in the BP adenosine-binding pocket, confer resistance to herboxidiene, pladienolide B, E7107 and H3B-8800[47,63,67].

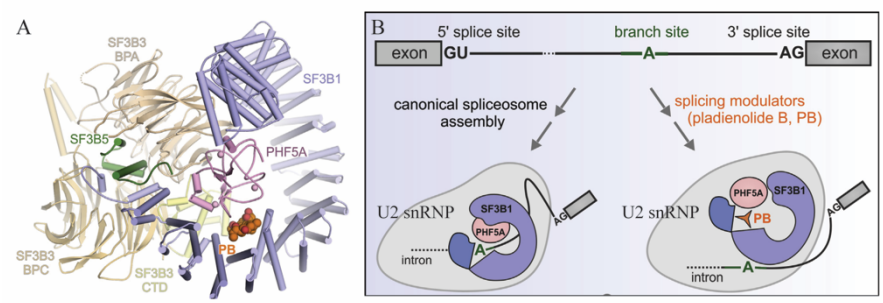


Figure 9. Structural basis of splicing modulation by SF3B binding splicing modulators. (A) Side view of the SF3B-Pladienolide B complex. SF3B subunits are shown in different colors: SF3B1 in purple, SF3B3 in yellow, SF3B5 in green and PHF5A in pink. The Pladienolide B ligand is shown as spheres and the carbon atoms are depicted in orange. (B) Schematic representation of the open and closed conformations of SF3B1. Pladienolide B binding inhibits the transition from open to closed conformation, thus inhibiting BP recognition (Modified from[64]).

SF3B1 binding molecules have cytotoxic effects in several tumor cell lines and antitumor activity in mice[68–72]. E7107 and H3B-8800 are the most advanced towards clinical practice. E7107, a pladienolide B derivative administered intravenously, entered in two phase I clinical trials in 2007 in patients with refractory and advance stage solid tumors (NCT00499499 and NCT00459823)[73,74]. Even though disease stabilization was achieved in 30% and 20% of the cases, respectively, both trials were suspended because a total of three patients presented optical neuritis.

More recently, H3B-8800, another pladienolide B derivative that is administered orally, has been engineered to target spliceosome-mutant cancer cells[63]. H3B-8800 preferentially killed *SF3B1*^{K700E} pancreatic cancer cell lines (Panc05.04) inducing caspase-3/7 cleavage after 24 hours. In NSG immunocompromised mice, subcutaneously implanted with acute myeloid leukemia K562 isogenic cell line and PDX from chronic myelomonocytic leukemia (CMML) harboring either WT or mutant *SRSF2*, H3B-8800

demonstrated modulation of splicing and a preferential killing of cells bearing mutations in spliceosome-related proteins. In addition, RNA deep sequencing analysis of H3B-8800-treated K562 cell line showed that H3B-8800 incubation inhibits splicing of short, GC-rich introns and affects splicing of mRNAs encoding spliceosome components[63]. In a phase I clinical trial conducted in patients with MDS, Acute Myeloid Leukemia (AML) and Chronic Myelomonocytic Leukemia (CMML) (NCT02841540), H3B-8800 showed a safe profile with only low grade gastrointestinal treatment-emergent adverse events. Even though no complete or partial responses were achieved, some patients experienced red blood cell transfusion independency, particularly the ones with *SF3B1* mutated MDS and presenting high levels of the aberrant transcript TMEM14C (Transmembrane Protein 14C, which is involved in heme metabolism)[75].

All-atom simulations conducted to understand the mechanism behind H3B-8800 and its preferential lethality towards cancer cells harboring mutated SF showed that this is not based on H3B-8800's higher affinity or longer residence time when bound to $SF3B1^{K700E}$ than to $SF3B1^{WT}$. This suggests that H3B-8800 selectivity could be related to its opportunistic ability to synergize and increase the splicing related-burden caused by the presence of splicing factor mutations in spliceosome-mutant cancer cells[66].

3. Chronic Lymphocytic Leukemia

3.1. Definition and epidemiology

Chronic lymphocytic leukemia /small lymphocytic lymphoma (CLL/SLL) is a B cell neoplasm characterized by monomorphic small mature B lymphocytes which coexpress CD5 and CD23[76]. The diagnosis requires $\geq 5x10^9$ /L monoclonal B lymphocytes in peripheral blood, although CLL/SLL may involve bone marrow, spleen and lymph nodes as well. Individuals with $<5x10^9$/L and without lymphadenopathy, organomegaly or other extramedullary disease are considered to have monoclonal B-cell lymphocytosis (MBL). Although CLL and SLL are the same disease, the term SLL is used for cases with circulating CLL cell count $<5x10^9$/L and documented nodal, splenic or other extramedullary involvement[76,77].

CLL is the most common type of adult leukemia in Western countries with a 4.9 per 100,000/year incidence and around 4000 death/year (data from The Surveillance Epidemiology and End Results (SEER) Program of the National Cancer Institute, https://seer.cancer.gov/statfacts/html/clyl.html). The median age at diagnosis is 70 years and is more frequent in men (1.9:1). There are some differences between CLL incidence in ethnic subgroups with the highest incidence in white people (6.7 per 100.000/year incidence for men and 3.9 per 100.000/year for women) and the lowest in Asians (1.6 per 100.000/year incidence for and 0.8 per 100.000/year for women). In addition, relatives of patients with CLL have a higher risk of developing the disease[78]. In this regard, some risk loci have

been identified that are differentially regulated during B cell development and could explain, at least in part, this inherited predisposition[79]. CLL usually has an indolent evolution, but it may also present frequent relapses or even transform to an aggressive lymphoma (Richter syndrome)[80]. However, CLL patients have a 87.2% 5-year relative survival in USA (data from SEER Program of the National Cancer Institute).

Based on peripheral blood cell counts, hemoglobin levels and physical examination, Rai and Binet staging systems are used for patient's risk stratification[81,82] (Table 1).

Stage		Definition
Binet system		
Binet A		Hb $\geq$100 g/l (6.21 mmol/l), platelets $\geq 100 \times 10^9$/l <3 involved lymphoid sites[a]
Binet B		Hb $\geq$100 g/l (6.21 mmol/l), platelets $\geq 100 \times 10^9$/l $\geq$3 involved lymphoid sites[a]
Binet C		Hb <100 g/l (6.21 mmol/l), platelets $<100 \times 10^9$/l
Rai system		
Low-risk	Rai 0	Lymphocytosis $>5 \times 10^9$/l
Intermediate-risk	Rai I	Lymphocytosis and lymphadenopathy
	Rai II	Lymphocytosis and hepatomegaly and/or splenomegaly with/without lymphadenopathy
High-risk	Rai III	Lymphocytosis and Hb <110 g/l (6.83 mmol/l) with/without lymphadenopathy/organomegaly
	Rai IV	Lymphocytosis and platelets $<100 \times 10^9$/l with/without lymphadenopathy/organomegaly

Table 1. Staging systems for CLL. a: Binet's system takes into account five potential sites of involvement: cervical, axillary, inguinal lymphadenopathy (either uni- or bilateral), spleen and liver. From[83]

3.2. IGHV Rearrangements and Somatic Hypermutations in B cell development

Mature B cells express a B cell receptor with a uniquely rearranged immunoglobulin (IG) molecule in their surface that

allows antigen recognition. At the beginning of B cell development, in the bone marrow, the first diversification step takes place when the IG heavy chain's (IGHV) and IG light chain's variable (V), diversity (D, only in the heavy) and joining (J) gene regions are hierarchically rearranged(Fig 10). These regions are spliced to the constant regions (Cμ with the VDJ and Cκ or Cλ with VD) and assembled forming the immunoglobulin M expressed in immature B cells. After antigen encounter in the periphery, B lymphocytes proliferate in the germinal center of the lymph node where IG genes experience further diversification by the somatic hypermutation process. The enzyme activation-induced cytidine deaminase introduces mutations in the IGHV and immunoglobulin light chain V region allowing affinity maturation to the antigen. The last diversification step is the class switch recombination (CSR) when the μ constant region is replaced with a constant region for another immunoglobulin and allows the generation of BCRs that carry heavy chains of different isotypes than IgM and IgD (Fig. 10)[84,85].

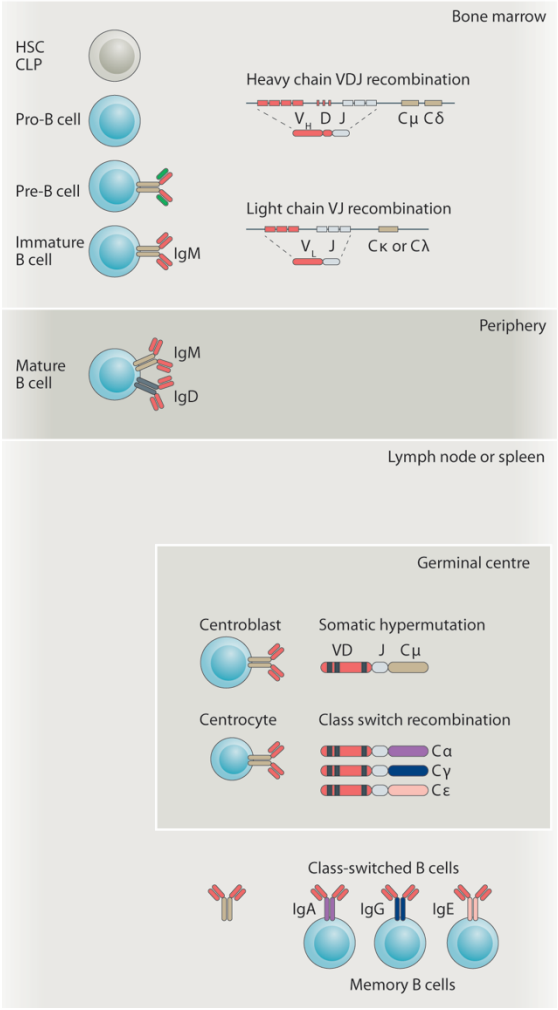


Figure 10. BCR diversification steps during B cell development. The recombination of the variable (V), diversity (D) and joining (J) regions of the heavy chain IG and V and J regions of the light chain rearrangements, somatic hypermutation and class switch recombination take place at different moments in B cell development. The constant region of the immunoglobulins (Cμ, Cδ, Cα, Cγ, Cε) determine the generation of different types of immunoglobulins (IgM, IgD, IgA, IgG or IgE). Modified from[86].

3.3. CLL subtypes

3.3.1. Immunogenetic subtypes

Two major immunogenetic CLL subtypes have been identified according to the mutational status of the IGHV genes: the ones with ≥98% identity with the germline, known as unmutated-CLL (U-CLL), originate from B lymphocytes that have not experienced the germinal center, while and the ones with <98% identity with the germline, known as mutated-CLL (M-CLL), originate from post-germinal center B lymphocytes[87]. These two subtypes are associated with distinct genetic and epigenetic alterations and they predict different clinical outcomes: the U-CLL is more aggressive than the M-CLL and has a worse response to chemotherapy treatment[88,89]. Because of its relevance, it is mandatory to analyze the IGHV mutational status during the pre-treatment evaluation of a patient[83].

In addition, about one third of CLL cases present nearly identical IG rearrangements, known as stereotypes[90]. Some of these stereotyped subsets also have particular biological properties and clinical behavior[85]. Moreover, a single point mutation in IGLV3-21^{R110} observed in all stereotype 2 cases has been associated with aggressive biological outcome [91,92].

3.3.2. Epigenetic subtypes

CLL chromatin landscape is influenced by dynamics during the normal B cell differentiation process[93]. Based on the DNA methylation signature that reflects different B cell maturation stages, CLL is divided into three clinical-biological subgroups: (i) naïve-

CLL, which has similarities to the pre-germinal center B cell; (ii) mature-CLL, which resembles the post-germinal center B cell and (iii) intermediate-CLL group with an intermediate methylation pattern between the other two groups[94]. All three groups are associated with different gene mutations, IGHV gene rearrangements, stereotypes and clinical behavior. The mature-CLL is associated with a better outcome and the naïve-CLL with a worse outcome. Recently, the discovery of IGLV3-21^{R110} mutation has enabled to distinguish between different tendencies in the intermediate-subgroup: the presence of IGLV3-21^{R110} confers an outcome comparable to the naïve-CLL and in the absence of IGLV3-21^{R110} the outcomes resembles to the mature-CLL's group[91].

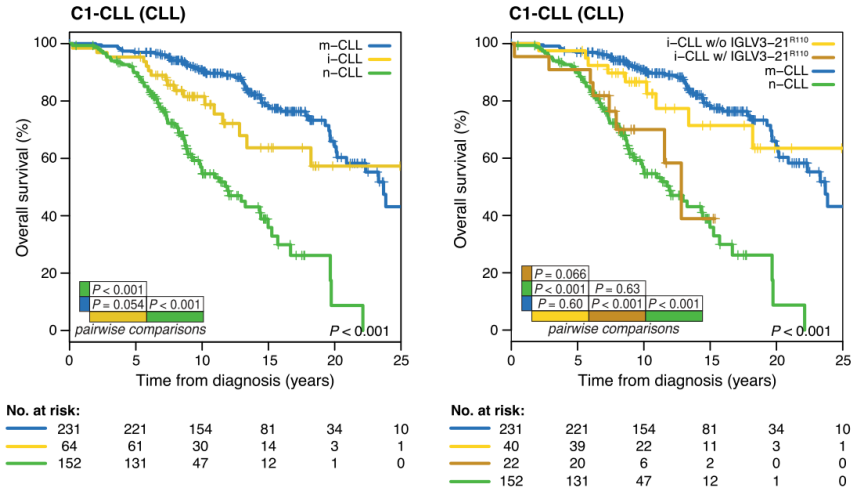


Figure 11. Comparison of overall survival among patients belonging to different epigenetic subtypes. In the first panel the IGLV3-21^{R110} presence is not taken into account, in the second panel the i-CLL cases are divided according to the presence/absence of the IGLV3-21^{R110}. n-CLL (naïve-CLL), m-CLL (mature-CLL) and i-CLL (intermediate-CLL) refer to the epigenetic subtypes of CLL based on the methylation pattern of the B cell. (from[91]).

3.4. CLL genetics

3.4.1. Chromosomal alterations

Around 80% of the CLL patients present at least one of the following chromosomal abnormalities: deletion in 13q (55% of the patients), deletion in 11q (18%), trisomy of 12q (16 %), deletion in 17p (7 %)[95]. They have a well-established prognostic value (from worse to better outcome: 17p deletion, 11q deletion, 12q trisomy, normal karyotype and deletion in 13q) and as the presence of 17p deletion has an impact in therapy election, Fluorescence In Situ Hybridization (FISH) assays must be carried out to detect it before therapy initiation[83].

3.4.2. Mutational landscape

Genome-wide sequencing has allowed the detailed characterization of the high genomic variability observed in CLL[96–99]. There are more than 60 driver genes but none of them is mutated in more than 15% of the patients at the time of diagnosis. The most recurrently mutated genes are *NOTCH1*, *SF3B1*, *TP53* and *ATM*, which are present in more than 5% of cases. Driver genes belong to several cellular pathways: B-cell signaling (e.g. *MYD88*, *BCL6*), DNA damage response (e.g. *ATM*, *TP53*), genome/chromatin structure (e.g. *ASXL1*, *MLL2*), NOTCH signaling (e.g. *NOTCH1*, *FBXW7*), RNA metabolism (e.g. *SF3B1*, *U1*), NF-κB signaling (e.g. *BIRC3*, *NFKBIE*), cell cycle (e.g. *KRAS*, *NRAS*, *BRAF*) and apoptosis (*ANKHD1*, *BAX*). Some of these mutations are associated with worse prognosis (*TP53*, *SF3B1*, *NOTCH1*)[96,100–102] or a better outcome (*MYD88*)[103]. As *TP53* mutation predicts the response to therapy, its assessment is mandatory before therapy initiation[83].

Many driver mutations are subclonal, e.g they appear in only a fraction of the tumoral cells[99,104](Fig. 12). However, some of them, such as *TP53* or *NOTCH1*, maintain their prognostic value even in subclonality[105][106].

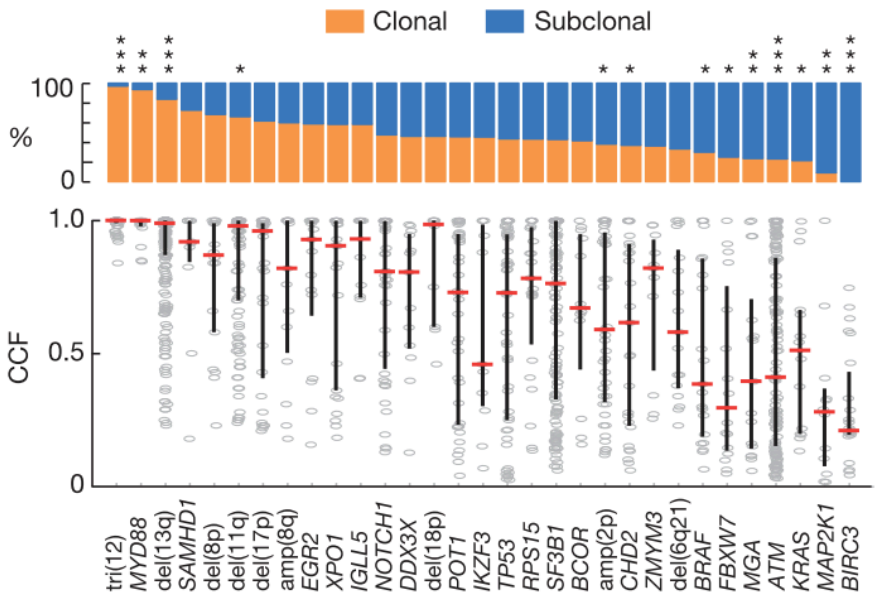


Figure 12. CLL has a subclonal architecture. The proportion in which a recurrent driver is found as clonal or subclonal is provided (top), along with the individual cancer cell fraction (CCF) values for each sample affected by a driver (tested for each driver with a Fisher's exact test, compared to the cumulative proportions of clonal and subclonal drivers excluding the driver being evaluated). Median CCF values are shown (bottom, bars represent the median and interquartile range for each driver). From[99].

3.5. Treatment

As early treatment with chemotherapeutic agents does not provide a survival benefit in patients with early stage CLL[107], the indication for treatment of CLL patients is based on the disease stage (advanced disease: Rai stage III and IV; Binet stage C) or the finding

of active disease defined as evidence of bone marrow failure; and on the presence of massive, progressive or symptomatic splenomegaly or lymphadenopathy or progressive lymphocytosis (lymphocyte doubling time of <6 months)[76].

Chemotherapy has been the cornerstone of CLL therapy. The alkylating agent chlorambucil monotherapy was the standard of care for some decades[108,109] until combinations of purine analogues (fludarabine) with cyclophosphamide improved the response in younger patients in the 90s[110–112]. Later on, the addition of the anti-CD20 monoclonal antibody, rituximab, to fludarabine and cyclophosphamide (FCR) for younger patients and the combination of anti-CD20 antibodies with alkylating agents for elderly patients with comorbidities demonstrated benefits in patients' survival[113,114]. During the last decade, the use of targeted therapy against the BCR receptor signaling pathway (Bruton's tyrosine kinase inhibitors: ibrutinib[115], acalabrutinib[116,117]; and phosphatidylinositol -3 kinase inhibitors: idelalisib[118], duvelisib[119]) and apoptosis inducing BCL2 inhibitors (venetoclax[120]) is improving patients outcomes considerably. The ESMO (European Society for Medical Oncology) recommendations for front-line treatment are summarized in figure 13.

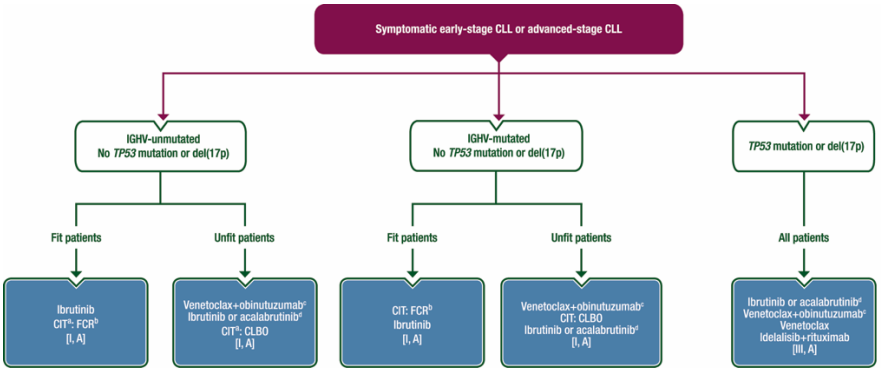


Figure 13. ESMO recommendations for front-line therapy in CLL. BR, bendamustine plus rituximab; CIT, chemoimmunotherapy; CLBO, chlorambucil plus obinutuzumab; CLL, chronic lymphocytic leukemia; FCR, fludarabine, cyclophosphamide and rituximab; IGHV, immunoglobulin heavy chain variable. a, CIT as alternative treatment, only if reasons against treatment with targeted therapies or non-availability; b, BR might be considered alternatively in patients above the age of 65 years; c, if available; d, if approved and available. From[83]

3.6. *SF3B1* mutated CLL

8–21% of CLL patients present a *SF3B1* mutation[96–99]. Most of mutations are missense and they are located within SF3B1's 4th to 8th HEAT repeats[34] (Fig. 5). K700E is the most frequent mutation and is present in half of the CLL patients with a *SF3B1* mutation. The analysis of matched samples at disease diagnosis and at relapse showed that acquisition of *SF3B1* mutations constitutes an intermediate event during disease evolution[99,106], and cannot therefore be considered as causal of disease initiation. In addition, in mice conditionally expressing a *SF3B1* mutation alone the mutation causes senescence, whereas the combination with *Atm* deletion causes a CLL-like disease[121].

In multiple studies, *SF3B1* mutations have been associated with biological and clinical characteristic of the patients, many of them

related to adverse prognosis. They are related to advanced stages, male gender, high leukocyte counts, elevated B2 microglobulin levels, high CD38 expression, U-CLL subgroup, intermediate CLL epigenetic subgroup, stereotyped BCR subset 2 and IGLV3-21^{R110} mutated subgroup, 11q deletion and fludarabine refractoriness[90,91,94,122–124]. In addition, it is an independent variable that predicts worse overall survival (OS) of CLL patients[106,123,125] and a shorter progression free survival (PFS)[44]. Patients harboring *SF3B1* mutation have a shorter PFS with fludarabine + cyclophosphamide (+ rituximab) (FCR) or lenalidomide treatments[126,127].

Transcriptomic analyses of normal B cells and CLL patients' cells have demonstrated the association between *SF3B1* mutations and alternative 3'ss usage[128,129]. Furthermore, differentially spliced and expressed genes upon *SF3B1* mutations alter multiple CLL associated pathways such as DNA damage, cell proliferation, NOTCH signaling, cytokine interactions or phosphatidylinositol signaling system[129,130].

AIMS

The main objectives of this thesis project were to characterize the impact of *SF3B1* mutations in CLL, both in terms of the effects of the mutations on alternative splicing and of the consequences for CLL cell phenotypes, and to test the effect of a novel splicing modulator (H3B-8800) in CLL. The approaches taken to address these goals will be presented in two chapters:

Chapter I: Transcriptomic Characterization of *SF3B1*-Mutated Chronic Lymphocytic Leukemia

- Generate isogenic CLL cell lines harboring *SF3B1* mutations.
- Characterize the alternative splicing changes induced by *SF3B1* mutation in patients and CLL cell lines.
- Investigate the pathogenicity associated with *SF3B1* mutations and the potential role of specific alternative splicing events in *SF3B1* mutation's pathogenicity.
- Find surrogate markers based on splicing alterations for functionally impactful *SF3B1* mutations.

Chapter II: Study of the therapeutic effects of SF3B1-binding H3B-8800 in *in vitro* and *in vivo* models of Chronic Lymphocytic Leukemia

- Assess the effect of H3B-8800 in CLL's *in vitro* and *in vivo* models.
- Determine whether H3B-8800 has a preferential effect in *SF3B1* mutated cases.
- Study the alternative splicing changes induced by H3B-8800 treatment.

RESULTS

Chapter I:

Transcriptomic Characterization of *SF3B1*- Mutated Chronic Lymphocytic Leukemia

1. Establishment and characterization of *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 CLL isogenic cell lines

In order to study the impact of *SF3B1* mutations on alternative splicing and the effect of H3B-8800 in wild type and SF3B1-mutant CLL cells, we established *SF3B1*^{K700E} MEC1 CLL isogenic cell lines using the CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)/Cas9 knock in system. Comparison between isogenic cell lines differing exclusively by the presence or absence of the K700E mutation will allow us to cleanly dissect the effects of the mutation without compound effects due to intratumoral and intertumoral heterogeneity found in CLL patient samples.

For genome editing applications, Cas9 endonuclease is targeted by a guide RNA (gRNA) harboring sequences identical or nearly identical to a given locus, near a 3 nucleotide pair sequence known as the Protospacer Adjacent Motif (PAM), where it induces a double stranded break. The cell has two mechanisms to repair this break: the most common way is non-homologous-end-joining (NEHJ) repair that will create insertions, mutations or deletions in the genome, known as indels; the other mechanism is the more precise homology directed repair (HDR): it needs a DNA template, a single-stranded donor oligonucleotide (ssODN).

A ssODN with a change in c.2098A>G position of *SF3B1* gene was specifically designed to introduce the K700E mutation.

In our first attempt, we obtained 109 clones (Fig I.1A), which upon genomic sequencing of the region, we found 53 of them to be

SF3B1^{WT}, 54 had indels, one clone corresponded to *SF3B1*^{K700N} mutation in homozygosis and one last clone harbored *SF3B1*^{K700E} mutation in heterozygosis accompanied by indels. A possible explanation for the low efficiency of correct editing is that after the Cas9 break was repaired, Cas9 continued cutting the DNA next to the PAM sequence until the Cas9 enzyme was degraded (Cas9 has a half-life of 24h).

Therefore, we carried out a new genome editing experiment introducing several changes in the protocol: we designed the ssODN complementary to the non-target strand[131], introduced a silent mutation in the PAM sequence to avoid further recognition of the PAM sequence once it was properly repaired[132], ordered phosphorothioate-modified nucleotides at the ends of the ssODN to stabilize it[133] and used 10uM of RS-1 HDR agonist to enhance the HDR[134] after cell electroporation. With these changes, from a total of 68 clones, we obtained 4 clones bearing the PAM site mutation, 3 of them additionally bearing the K700E mutation (later referred to as MUT1, MUT2, MUT3) and one bearing a K700K mutation (later referred to as WT3) (Fig I.1A). With these results, we decided to use the parental MEC1 cell line (WT1), a WT clone which underwent clonal selection after the failed CRISPR/Cas9 edition (WT2) and the *SF3B1*^{K700K} with the silent mutation in PAM (WT3) as the *SF3B1*^{WT} controls, and the three *SF3B1*^{K700E} clones (MUT1, MUT2, MUT3) with the silent mutation in PAM as the *SF3B1*^{K700E} mutated counterparts. In order to define the mutational status of the obtained clones, we used a digital PCR (dPCR) approach (Fig I.1B and C).

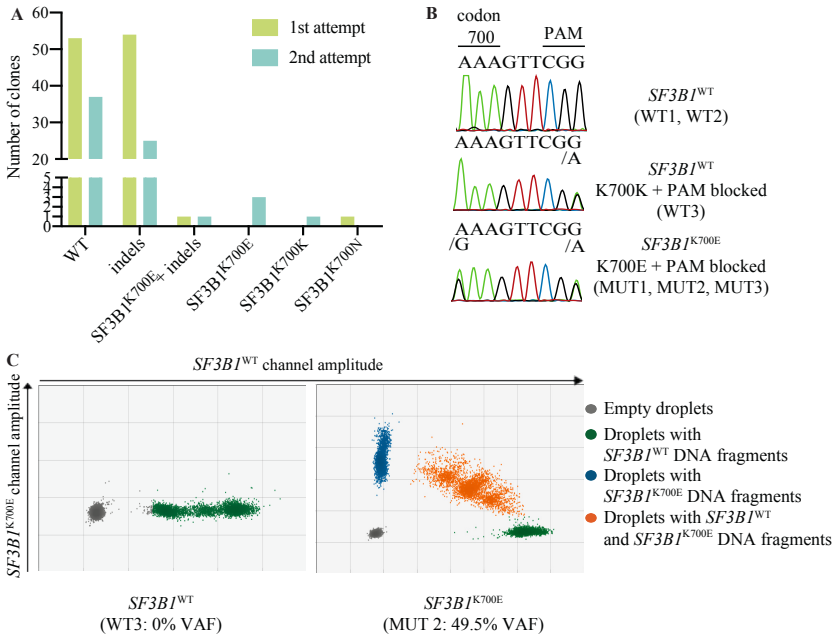


Figure I.1 Clones generated by CRISPR/Cas9 genome editing. (A) Clones generated in the two CRISPR/Cas9 editing experiments. (B) DNA sequence of CRISPR/Cas9-modified MEC1 clones. The nucleotides coding for SF3B1 codon 700 and the Cas9 targeted PAM sequence are shown. (C) Quantification by dPCR of MEC1 *SF3B1*^{WT} and *SF3B1*^{K700E} clones. Variant allele frequency (VAF) is the fraction of droplets observed matching the specific K700E variant divided by the overall number of droplets analyzed.

To further characterize the clones we generated, we first performed a western blot analysis to assess whether the levels of SF3B1 protein were affected. The results indicated that the presence of *SF3B1*^{K700E} mutation did not affect the amount of protein expressed (Fig I.2A). Second, we assessed the viability of the cell lines after 48h in culture by flow cytometry, selecting the annexin V and propidium iodide negative populations. We observed that the viability of *SF3B1*^{K700E} clones was lower than that of *SF3B1*^{WT} clones, with viability means in the range of 62% to 81% for the

SF3B1^{K700E} clones and around 90% of viability for the *SF3B1*^{WT} clones. Third, we assessed the metabolic activity (through the activity of NAD(P)H-dependent cellular oxidoreductase enzymes) of the isogenic MEC1 cell lines using colorimetric MTT assays. In this case, the variability between the experiments was higher than in viability assays, but the *SF3B1*^{K700E} clones displayed lower metabolic activity than *SF3B1*^{WT} clones.

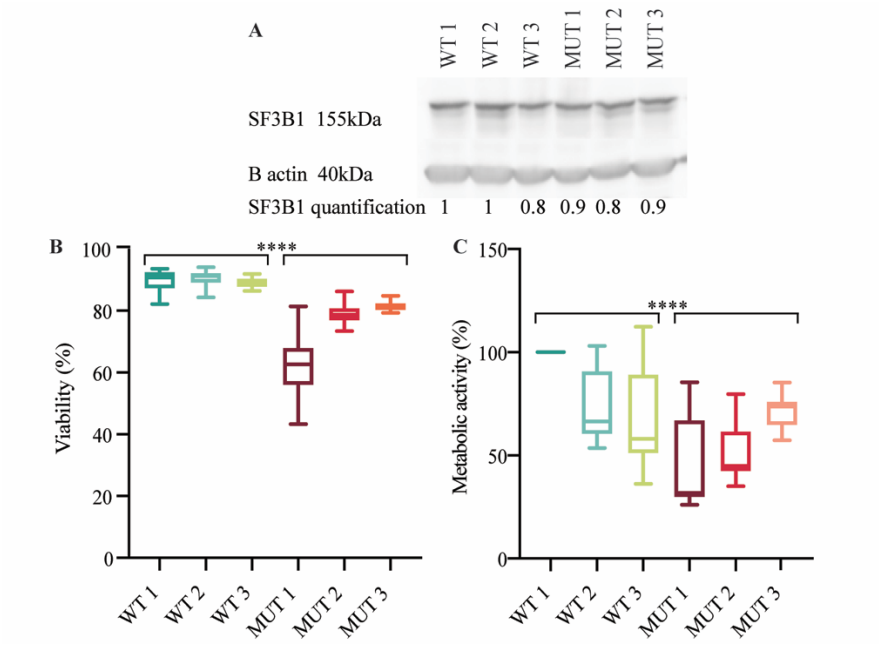


Figure I.2. Characterization of *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 cell lines. (A) SF3B1 protein expression in *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 cell lines. The quantification was calculated by normalizing the SF3B1/β actin ratio to sample WT1. (B) Viability of *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 cell lines after 48h in culture at 30.000cell/100uL concentration, assessed by flow cytometry (Annexin V-, PI-). There are 17 replicates in each group. Mann-Whitney test (****: $p<0.0001$). (C) Metabolic activity of *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 cell lines after 48h in culture at 30.000cell/100uL concentration assessed by absorbance at 570 and 670 wavelengths. There are 12 replicates in each group. Mann-Whitney test (****: $p<0.0001$).

1. Transcriptomic analysis of *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 CLL cell lines and CLL patients' samples

1.1. CLL patients selection for transcriptomic analysis

We had access to RNA deep sequencing data from 298 patients from the International Cancer Genome Consortium (ICGC) with CLL, MBL or SLL at diagnosis as well as to lymphocytes from 4 healthy donors. 27 patients presented a mutation in *SF3B1* (N626Y, R625C, H662D, T663I, K666N, K700E, V701F, G742D, M757T or K862K) and 271 were WT (patients' biological data can be found in Supplementary Table 1).

The number of reads obtained for each sample was quite variable, ranging from 30 to 200 million, most of the samples having around 50 million reads. The toolset *vast-tools* (Vertebrate Alternative Splicing and Transcription Tools) was used to quantify alternative splicing from all samples[135,136]. These tools use the Percent Spliced In (PSI) metric, i.e. the proportion of usage of a particular inclusion junction over the total (inclusion reads/(inclusion reads + exclusion reads)) to quantify AS events. Taking into consideration only events with a minimum quality score of VLOW (> 10 reads coverage supporting the PSI calculation), we quantified 10791 Alt3'ss events, 7385 Alt5'ss events, 14665 IR events and 14383 SE events. A Principal Component Analysis (PCA) was carried out with each of these 4 alternative splicing types as well as with all of the AS events together.

Standard blood collection procedures, including time until the sample is cryopreserved, can influence the transcriptional and

posttranscriptional landscape of hematopoietic cells[137]. As in the ICGC project patients' samples were collected from different institutions in Spain and were cryopreserved in our institution, we wondered if there could be differences in the transcriptome between the samples collected in our institution -which were more readily cryopreserved- and the rest. To address this issue, we did PCA considering all quantified AS events and observed that many of the samples from our institution clustered together (Fig. I.3). Even though we could not be sure that time to cryopreservation was the (only) cause behind this difference, we decided to take into account this PC1 distribution in subsequent analyses

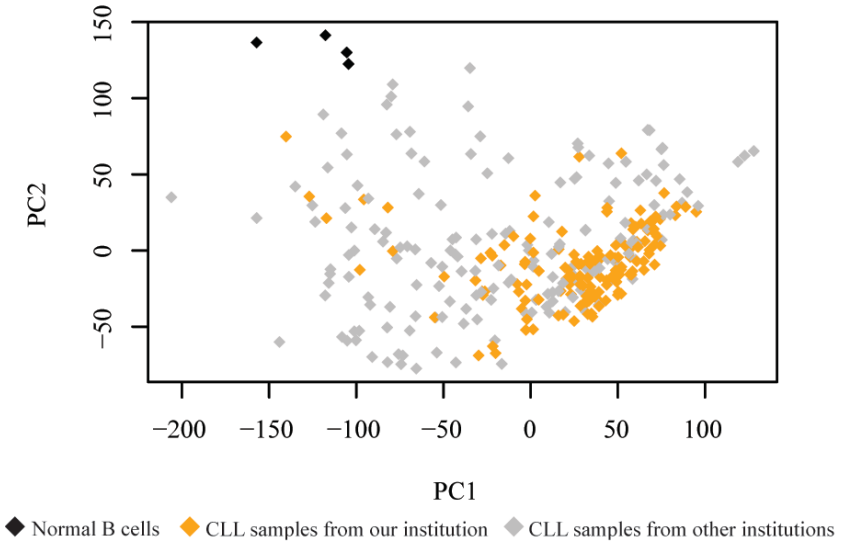


Figure I.3. Principal Component Analysis (PCA) calculated for the AS events (> 10 reads coverage supporting the PSI calculation) found in the peripheral blood mononuclear cells from 298 patients and 4 healthy donors.

When characterizing the clonal composition of a tumor, the cancer cell fraction (CCF) refers to the fraction of cancer cells presenting a mutation and is calculated using the variant allele frequency (VAF). The PCA performed using Alt3'ss data showed that PCA component 2 (variance 7%) was related to the CCF of the *SF3B1* mutation. While all the *SF3B1*^{WT} samples clustered together, the higher the CCF of a mutated sample, the bigger the difference in Alt3'ss patterns compared to *SF3B1*^{WT} samples (Fig I.4A). The samples with very low CCF (CCF<12%) clustered together with *SF3B1*^{WT} samples, except for two samples harboring *SF3B1* E862K and M757T mutations that despite having a high CCF (CCF>50%), clustered with *SF3B1*^{WT}. This finding goes in line with previous studies reporting that non-hotspot mutations, mainly the ones located from 9th to 12th HEAT domains, do not induce cryptic 3'ss selection patterns attributed to other *SF3B1* mutations located in other HEAT domains[26,138]. Additionally, PCA analysis performed with Alt5'ss, IR and SE events didn't show any clustering related to the presence and/or abundance of *SF3B1* mutations (Fig I.4B-D).

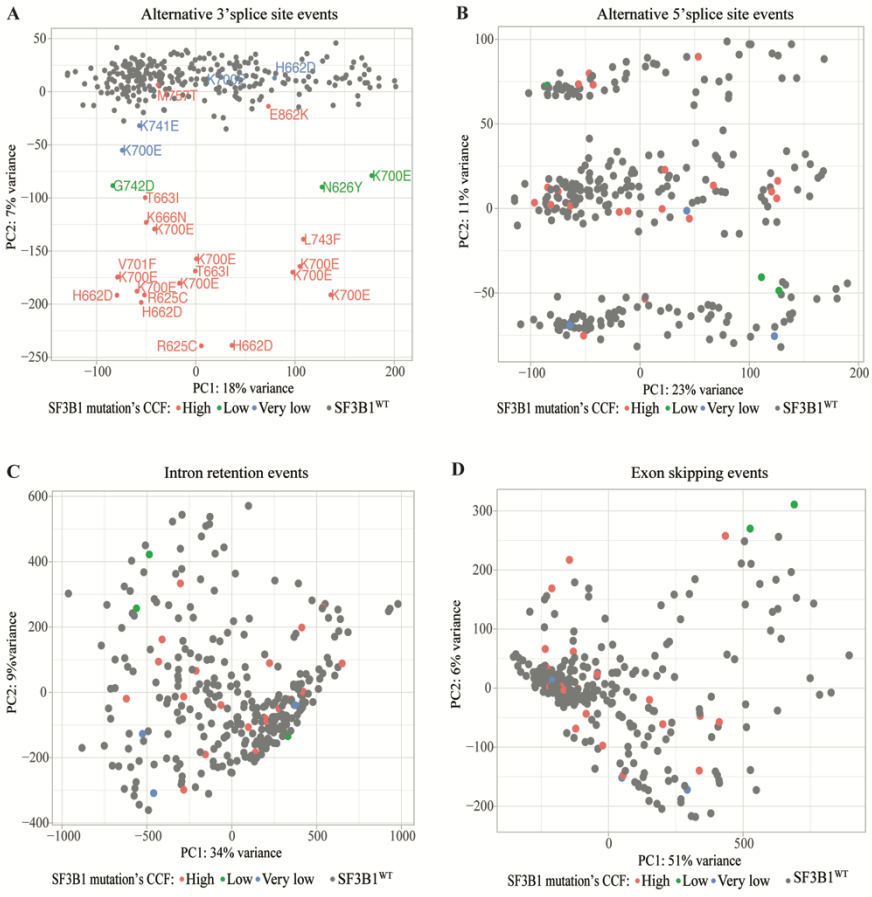


Figure I.4. Principal Component Analysis (PCA) calculated for Alt3'ss. (A), Alt5'ss events (B), IR events (C) and ES events (D) found in patients. PCA was performed on the PSI values of Alt3'ss, Alt5'ss, IR and ES events of each sample using *ggfortify and ggplot2* packages in R(v3.6.3). The mutational status is indicated in panel A. CCF is indicated by color code as defined in each panel

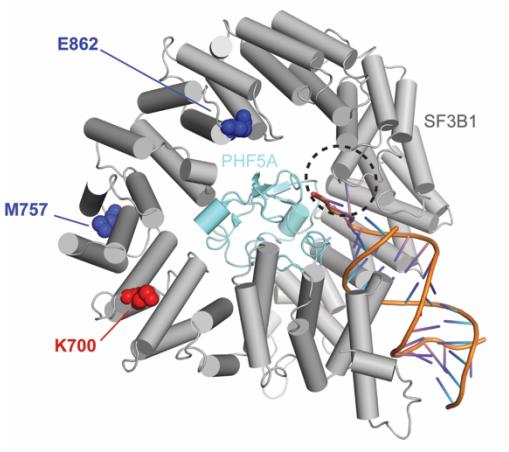


Figure I.5. Location of SF3B1 residues that induce (K700) or not (M757, E862) alterations in patterns of 3'ss usage upon mutation. The structures of SF3B1 (in grey) and PHF5A (SF3B7) (in light blue) structures are shown. Highlighted in dark blue are residues that do not lead to altered 3'ss usage upon mutation (M575T, E862K) and in red K700, a mutation hotspot (frequently K700E) which induces characteristic patterns of 3'ss usage. mRNA and U2snRNA are represented in orange (base pairing corresponds to annealing between U2 snRNA BP recognition sequence and BP nucleotides flanking the BP adenosine, which bulges out from the helix) and the binding site of drugs such as SSA, Pladienolide B or H3B-8800 is marked by the circle with dotted line. The figure was done by Suzanne Mays from Juan Valcárcel's lab using the PyMOL Molecular Graphics System, Version 1.3r1 (Schrodinger, LLC) based on PDB 6FF4 structure from the Protein Data Bank.

To further study the impact of *SF3B1* mutations, we decided to compare 15 *SF3B1* mutated patient samples presenting high CCF with 225 *SF3B1*^{WT} CLL patients samples and exclude patients with *SF3B1* E862K or M757T mutation from the analysis. None of the patients presented mutations in other splicing factors or RNA metabolism-related genes (*BUD13, CARNS1, CELF4, DDX3X, EIF4A3, MAGOH, MPHOSPH10, NXF1, RNF113A, SKIVV2L2, SF3B4, SRSF2, SRSF7, XPO1, U1, U2AF2,* nor *ZRSR2*). The biological characteristics of the patients included in this selection show that patients with *SF3B1* mutation belong more frequently to

the unmutated-IGHV subgroup and, regarding their epigenetic status, to the intermediate-CLL subgroup, in line with previous findings[90,94](Figure I.6). The most frequent mutation in our dataset was the hotspot K700E mutation, but H662D, R625C, L743F and T663I mutations were also included. The detailed biological characteristics of each patient can be found in Supplementary Table 1.

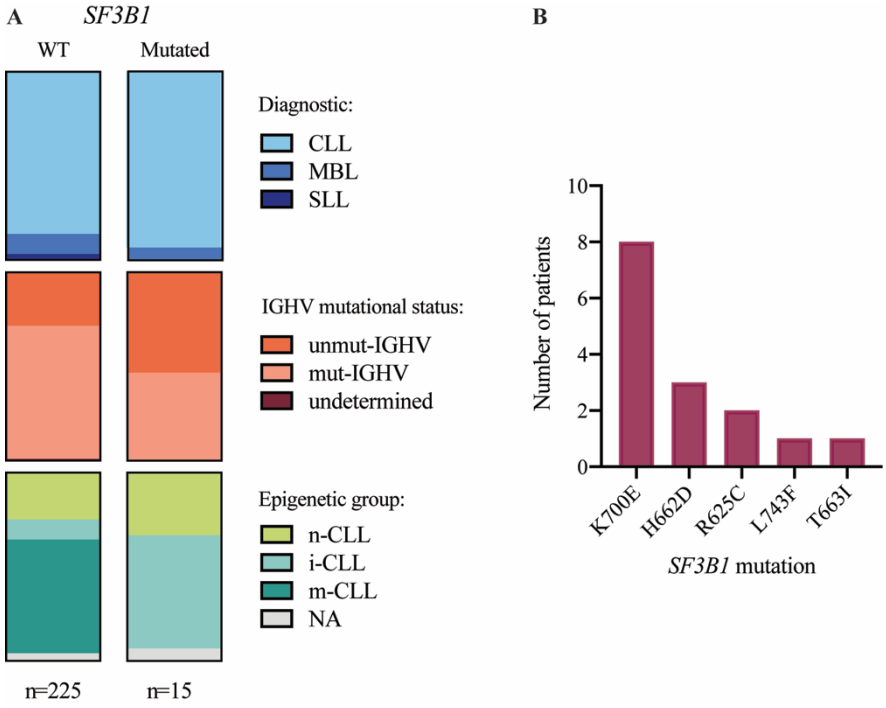


Figure I.6. Biological characteristics of the patients included in the comparison of RNA-seq data from two groups: *SF3B1* WT and *SF3B1*-mutated after discarding the patients with very-low/low CCF and *SF3B1* E862K or M757T mutations. (A) Diagnosis (CLL: chronic lymphocytic leukemia; MBL: monoclonal B lymphocytosis; SLL: small lymphocytic lymphoma), IGHV mutational status (unmut-IGHV: unmutated-IGHV; mut-IGHV: mutated-IGHV) and epigenetic (n-CLL: naïve-CLL; i-CLL: intermediate CLL; m-CLL: mature-CLL) group of the patients depending on *SF3B1* mutational status. (B) Number of patient samples with the indicated *SF3B1* mutations analyzed by RNA-seq.

1.2. AS events found in CLL patients and *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 CLL isogenic cell lines

To investigate the impact of *SF3B1* mutations on the transcriptome, we first compared RNA-seq datasets between the 15 *SF3B1* mutated patients samples presenting a high CCF, with the 225 *SF3B1*^{WT} CLL patients samples. In addition, we compared RNA-seq from MEC1 *SF3B1*^{K700E} and *SF3B1*^{WT} CLL cell lines.

All the samples included in the comparison passed the FastQC quality control from MultiQC[139]. In the case of MEC1 cell lines, each triplicate, (*SF3B1*^{WT}: WT1, WT2, WT3; *SF3B1*^{K700E}: MUT1, MUT2, MUT3) was sequenced in separate Illumina lanes and at least 200 million reads per sample were obtained.

To quantify AS and identify differentially spliced AS events associated with the *SF3B1* mutation, we used *vast-tools*. For comparing isogenic cell lines samples, we defined AS events differentially affected upon *SF3B1* mutation as events with a ΔPSI higher than 15% between the merged replicates of each condition. In the case of the patients samples, to deal with the heterogeneity and other confounding factors possibly present in the dataset, we made two comparisons considering the distribution of the patients in PC1 for all the AS events detected (> 10 reads coverage supporting the PSI calculation) (Fig I.3). We compared the samples with PC1≥ 0 on one side (8 *SF3B1* mutated samples versus 142 samples *SF3B1*^{WT} samples) and the samples with PC1< 0 (7 *SF3B1* mutated samples versus 92 samples *SF3B1*^{WT} samples). AS events needed to have at least 10 reads in 5 of *SF3B1* mutated samples and 20 of the *SF3B1*^{WT}

samples to take them into consideration for the comparison, as well as a ΔPSI higher than 15% between *SF3B1* mutated and *SF3B1*^{WT} samples. As expected, Alt3'ss were the most frequent differentially spliced AS events in the presence of *SF3B1* mutation both in patients and in isogenic MEC1 cell lines (Fig I.7). Alt3'ss were followed in frequency by IR, ES and Alt5'ss.

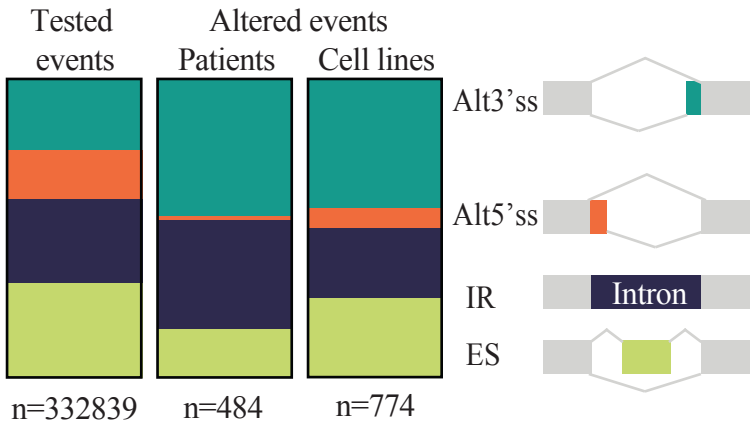


Figure I. 7. Number of differentially spliced (|ΔPSI| ≥ 15) AS event types upon *SF3B1* mutations in patients and *SF3B1*^{K700E} MEC1 cell lines comparing to all assessed AS events using *vast-tools*.

In order to determine whether the cell lines we engineered mimic the effect of SF3B1 mutations on splicing regulation observed in patient samples, we checked the overlap between the differentially spliced AS events displaying a |ΔPSI| ≥ 15 in both datasets. We found the biggest overlap in the Alt3'ss in which 56% of the Alt3'ss found in patients could be detected in MEC1 cell lines (Fig I.8A), the overlaps for the other categories were lower, with 25% of the Alt5'ss found in patient samples being detected also in MEC1 cell lines, 20% for ES events and 18% for IR events (Fig I.8A).

Next we evaluated the correlation between the effects of *SF3B1* mutation in cell lines and patient samples, taking into consideration the direction of the change, i.e. more inclusion or more skipping upon *SF3B1* mutation. We considered a $|\Delta PSI| \geq 15$ in at least in one in the patients or MEC1 cell lines to evaluate the correlation between them. This broader selection allowed us to find more overlapping events and again the highest values were found comparing Alt3'ss, with a R=0.9 correlation coefficient between patient samples and cell lines (Fig I.8B). We conclude that *SF3B1* mutations have a stronger impact on Alt3'ss but also display effects on IR, SE and Alt5'ss. Importantly, we also concluded that MEC1 *SF3B1*^{K700E} and *SF3B1*^{WT} CLL isogenic cell lines are a good model to study AS in CLL, recapitulating the changes in Alt3'ss observed in patient samples.

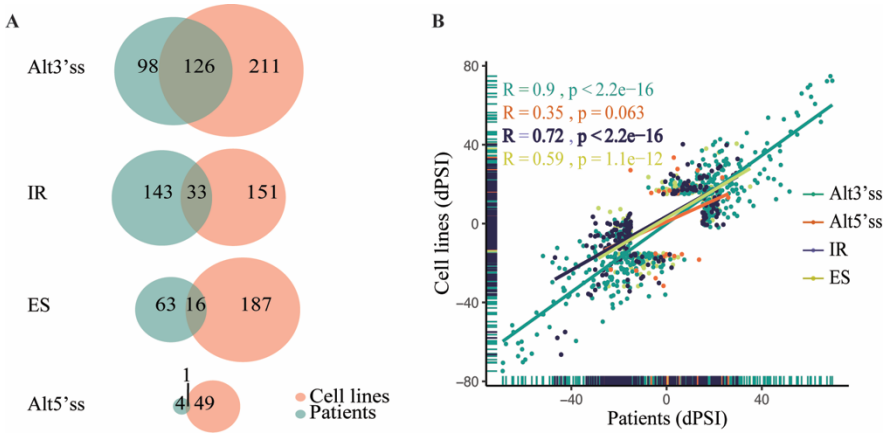


Figure I.8. (A) Overlap of the AS event types found in patients and cell lines comparing SF3B1 WT and mutated samples. (B) Correlation between AS events ΔPSI values comparing SF3B1 mutant vs WT samples found in MEC1 isogenic cell lines and in patient samples, with a $|\Delta PSI| \geq 15$ at least in one of the datasets. Pearson correlation coefficients were estimated for each AS type as indicated.

To predict the impact of AS changes induced by *SF3B1* mutations, we took the differentially spliced AS events in the patients and MEC1

cell lines (|ΔPSI| ≥ 15 in both datasets), and predicted their effects on the coding sequence (CDS) using a tool available in VastDB (Hs2_ONTO_all-v2_2). We observed an increase of the AS events predicted to undergo NMD upon *SF3B1* mutation (Fig I.9A). This finding goes in line with previous work were cryptic 3'ss have been validated to lead to NMD-mediated mRNA degradation[52]. We then checked the effect of CDS disruption on gene expression levels, analyzed by Deseq2. We observed that gene expression levels decreased significantly in those events disrupting CDS ($p<0.0001$) (Fig I.9B).

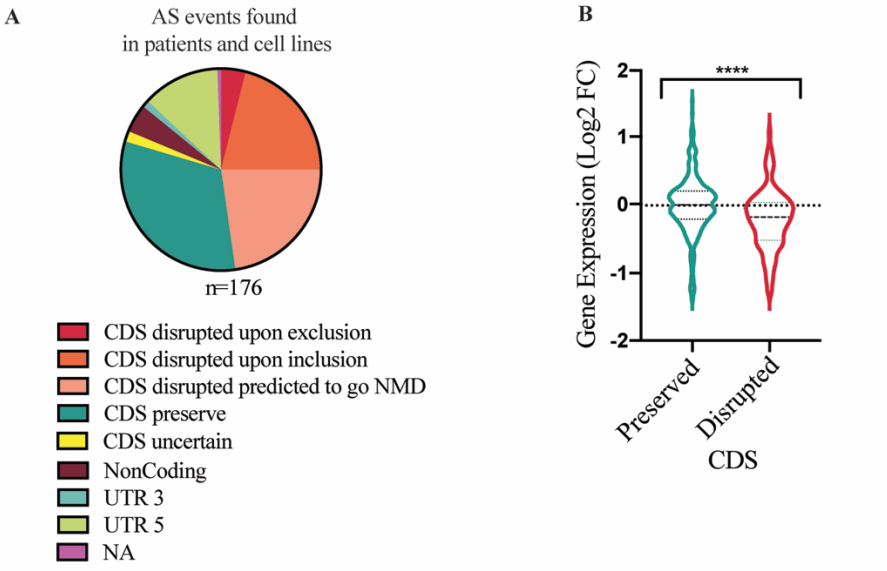


Figure I.9. Effects of AS changes observed upon SF3B1 mutation. (A) Effects on the coding sequence (CDS) of differentially spliced (|ΔPSI| ≥ 15) isoforms found in patients and MEC1 cell lines using vast-tools. The pie charts represent the fraction of events corresponding to each of the categories of predicted impact on CDS. (B) Comparison of gene expression changes associated with the events predicted to preserve or disrupt CDS. The statistical significance between both groups was evaluated using a Mann-Whitney test. ***: p-value <0.0001.

1.3. Alt3'ss patterns related to *SF3B1*^{K700E} mutation

SF3B1 mutations were previously associated with cryptic 3'ss selection. Therefor, in order to characterize all possible novel splicing events, we developed, in collaboration with André Gohr (Manuel Irimia's laboratory in System Biology department in the Centre for Genomic Regulation), a tool to extract annotated and non-annotated (de novo) splice junctions from mapped RNA-Sequencing reads and find a wider variety of Alt'3ss (refer to Material & Methods and https://gitlab.com/aghr/juncexplorer/-/wikis/home). We focused on AG type 3'ss using a common GT type 5'ss. We calculated the Percent Spliced site Usage (PSU), a metric similar to PSI, converting each junction's read numbers into relative frequencies. We considered a differentially spliced Alt3'ss when there was at least 15% change in PSU between *SF3B1*^{WT} and *SF3B1*^{K700E} conditions. With this approach, we found 160 differentially spliced Alt'3ss between *SF3B1*^{WT} and *SF3B1*^{K700E} in patients and 1633 in MEC1 cell lines. This difference in the number of events found in the patients and cell lines could be related to the fact that we had different coverages in each datasets[138] (approximately a median of 50 millions of reads in patients and 200 millions of reads in cell lines).

This tool allowed us to identify and count several alternative 3'ss associated with a given 5´ss. We observed 1 or 2 Alt3'ss as the most common occurrence in the presence of *SF3B1* mutation in CLL patient samples and MEC1 cell lines (Fig. I.10). We defined a cryptic 3'ss an alternative 3'ss with a PSU<5 in *SF3B1*^{WT} samples. In other words, a cryptic 3'ss would be an alternative 3'ss only found in

SF3B1 mutated samples (and with very low usage in SF3B1 WT samples).

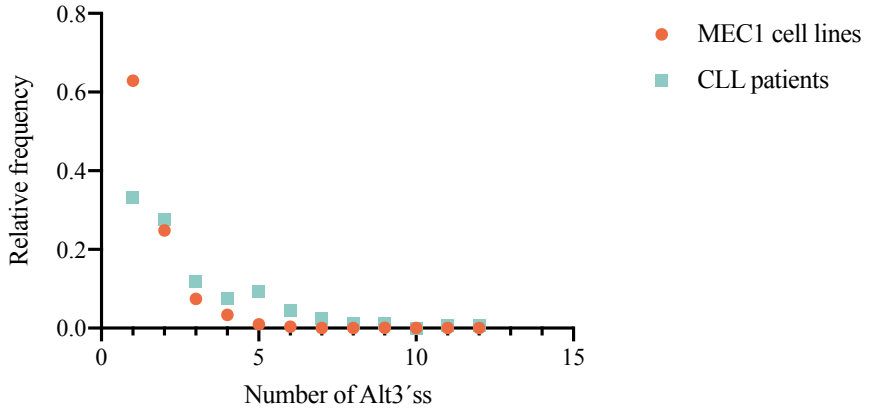


Fig I.10. Relative frequency of the number of Alt3'ss in the presence of *SF3B1* mutation in MEC1 cell lines and CLL patient samples.

Our next question was related to the distance between the Alt3'ss and the canonical 3'ss. To simplify the analysis, we decided to analyze the distribution of distances between cryptic and consensus 3'ss focusing on instances of only 2 AGs. We observed that most of the alternative or cryptic 3'ss are located upstream of the canonical 3'ss (the most frequent value was -17 nucleotides for cryptic and Alt3'ss displaying PSU values ≥30)(Fig I.11A), consistent with previous reports[45]. But we could also detect some cryptic or alternative 3'ss at much longer distances and also downstream of the canonical 3'ss. For the systematic characterization of these findings, we decided to make 4 groups of cryptic or Alt3'ss depending on the distance and position relative to the corresponding canonical 3'ss. First, we took into account if the cryptic or Alt3'ss was located

upstream or downstream of the canonical 3'ss. Then, we set a 50 nucleotides cut off to consider that a cryptic or Alt3'ss was "near" if it was at a distance shorter than 50 nucleotides, or "far" if this distance was longer than 50 nucleotides. A schematic representation of these groups is shown in figure I.11B, including also "upstream and downstream NAGNAG" patterns where the distance between the cryptic or Alt3'ss and the canonical 3'ss is only of 3 nucleotides. In the case of *SF3B1*-mutated CLL patient samples, the most frequent pattern was "upstream near" (76%) followed by "downstream near" (12%). In the case of the *SF3B1*^{K700E} MEC1 cell lines, the "upstream near" pattern was also the most frequent (38%), followed by "downstream far", "upstream far" and "downstream near" patterns, with a frequency around 20% each.

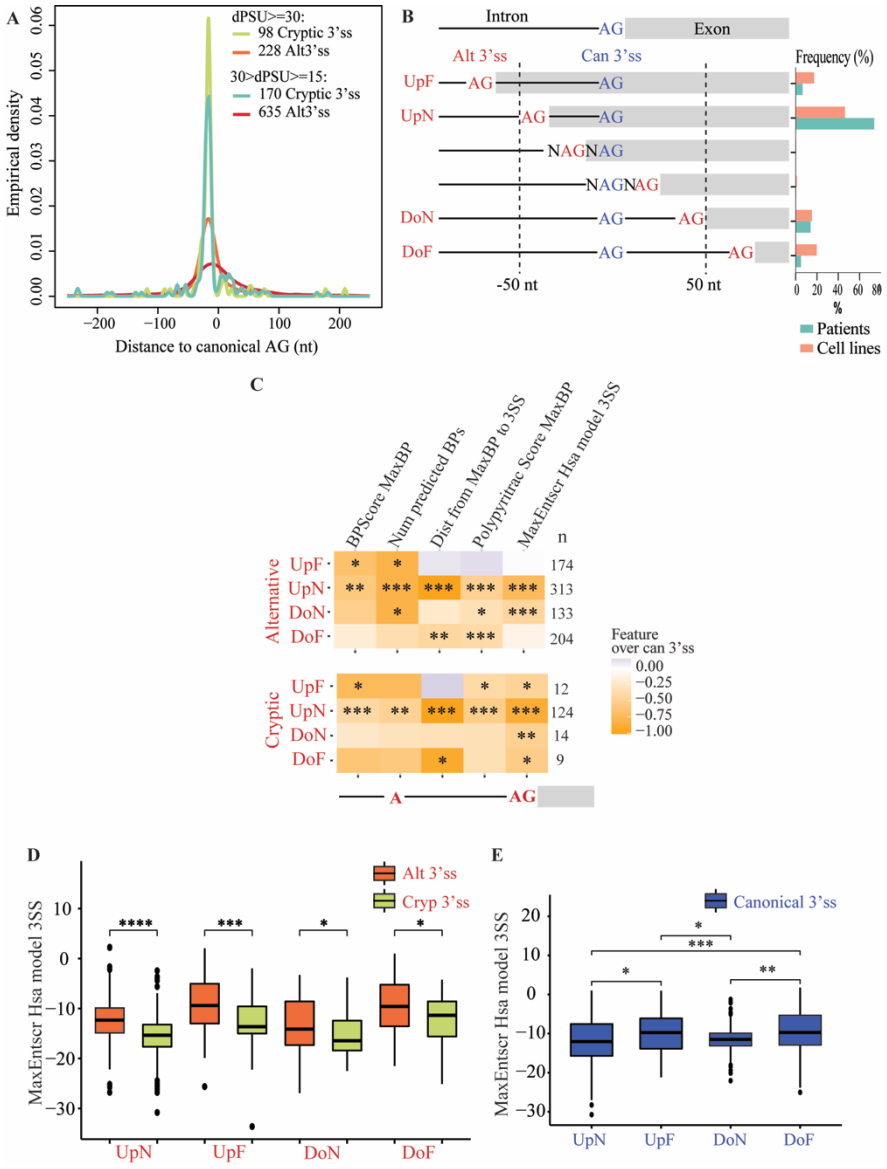


Fig I.11. Location and characteristics of cryptic and Alt3'ss activated by *SF3B1* mutations. (A) Empirical density of distances between cryptic or Alt3'ss and their canonical 3'ss. A negative distance indicates that the cryptic 3'ss or Alt3'ss is located upstream of the canonical 3'ss. Distribution of distances is shown for cryptic and for Alt3'ss with dPSU values above and below 30%, as indicated. (B) Schematic representation of subtypes of cryptic and Alt3'ss patterns relative to the canonical 3'ss (upstream near, UpN: <50 nucleotides upstream; upstream far, UpF: >50 nucleotides upstream; downstream near, DoN: <50 nucleotides downstream; downstream far, DoF: >50 nucleotides downstream) and their relative frequency in

SF3B1-mutated CLL patient samples and SF3B1^{K700E} MEC1 cell lines. (C-E) Comparison between the normalized means of a variety of intronic features calculated using Matt software[140] for each type of cryptic or Alt3'ss. (C) The mean of the features of the predicted BP, cryptic and Alt3'ss compared to those of canonical 3'ss and predicted associated BP: BPScore MaxBP (BP score for the best predicted BP), num of predicted BPs (number of predicted BPs), Dist from maxBP to 3SS (distance from the best predicted BP to the 3'ss), Polypyritrac score MaxBP (polypyrimidine tract score for the best predicted BP), MaxEntscr Hsa model 3SS (maximum entropy score of 3'ss using a model trained with human splice sites). Orange rectangles in the heatmap indicate that cryptic /Alt3'ss display lower values compared to canonical 3'ss for each of the features compared. The schematic representation below the heatmap indicates the position of the features examined within the architecture of 3' ss regions. (D) Comparison between the Maximum Entropy Score Hsa model of Alt 3'ss and cryptic 3'ss depending on the distances they are from the canonical 3'ss (UpN, UpF, DoN, DoF). (E) Comparison between the Maximum Entropy Score Hsa model of canonical 3'ss which have their Alt3'ss located at different distances (UpN, UpF, DoN, DoF). The comparisons that are statistically significant with Mann-Whitney U test are represented as follows: *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$, ****: $p<0.0001$.

To analyze characteristics of the splice sites involved that could shed light on the underlying regulatory mechanisms, we used Matt[140] (see Material & Methods) to analyze the features of these cryptic or Alt3'ss and their predicted BP compared to their canonical partners (Fig. I 11C). We calculated the maximum entropy score of the Alt3'ss that is an approximation to the overall strength of the 3' splice sites[141] and also features regarding the location and strength of the BP and polypyrimidine tract[142]. We observed that the BP score for the best predicted BP, the sequence with maximum complementarity to U2 snRNA in the nucleotides flanking the BP adenosine, is weaker in these cryptic and Alt3'ss than in the canonical 3'ss, the number of predicted BPs is lower, the distance from the best predicted BP to the 3'ss is shorter, the polypyrimidine-tract score for the best predicted BP is weaker and the Maximum Entropy Score Hsa (Homo sapiens) model (maximum entropy score of 3'ss using a model trained with human splice sites) for these cryptic and Alt3'ss

is also weaker (Fig I.11C). All these characteristics are generally shared across the 4 patterns (UpF, UpN, DoN, DoF). These analyses therefore suggest that *SF3B1* mutation leads to the activation of alternative/cryptic 3'ss that display weaker features, including weaker predicted BPs and weaker polypyrimidine-tracts. We also compared the strength of the cryptic vs Alt3'ss (Fig. I.11D). Cryptic 3'ss, which can be detected only in the presence of *SF3B1* mutations, are weaker than Alt3'ss. These results suggest that *SF3B1* mutations allow the more efficient recognition of a range of weak BP/3'splice sites.

In addition, we also compared the canonical 3'ss corresponding to each pattern (UpF, UpN, DoN, DoF). We observed that canonical 3'ss with Alt3'ss located more than 50 nucleotides away, regardless of their upstream or downstream location, have a higher entropy score than the canonical 3'ss with Alt3'ss nearer than 50 nucleotides (Fig.I.11E). These results suggest that activation of distant Alt3'ss require strong interactions of their canonical 3'ss with the splicing machinery, possibly as a first step required for Alt3'ss recognition.

1.3.1. Validation of the Alt3'ss patterns associated with *SF3B1*^{K700E} mutation by RT-PCR

We used our isogenic wild type and *SF3B1*^{K700E} MEC1 CLL cell lines to validate a selected subset of differentially spliced AS events associated with *SF3B1* mutations detected both in CLL patient samples and in MEC1 cell lines. For that purpose, we used three *SF3B1*^{WT} and three *SF3B1*^{K700E} replicates (described in Chapter I.1) and performed RT-PCR assays using oligonucleotide primers that

allow to detect two amplification products of distinct sizes corresponding to the use of the canonical and cryptic/Alt3'ss activated by the presence of the K700E mutation (Fig I.12A). Examples of each of the 4 3'ss patterns described in the previous section ("upstream near", "upstream far", "downstream near" and "downstream far") were validated, demonstrating that the cell lines recapitulate well at least some of the changes seen in patients (Fig I.12A).

Although the NAGNAG patterns found by RNA-seq were very infrequent in CLL patients and MEC1 cell lines, we tried to validate some of these events because *SF3B1* mutation had not yet been associated with NAGNAG splicing. Since the difference in length between the aberrant and canonical transcripts is only three nucleotides, we analyzed the RT-PCR product by fragment analysis. Out of the 4 cases we picked for validation, only the event in *TARBP1* gene was validated. (Fig I.12B-C).

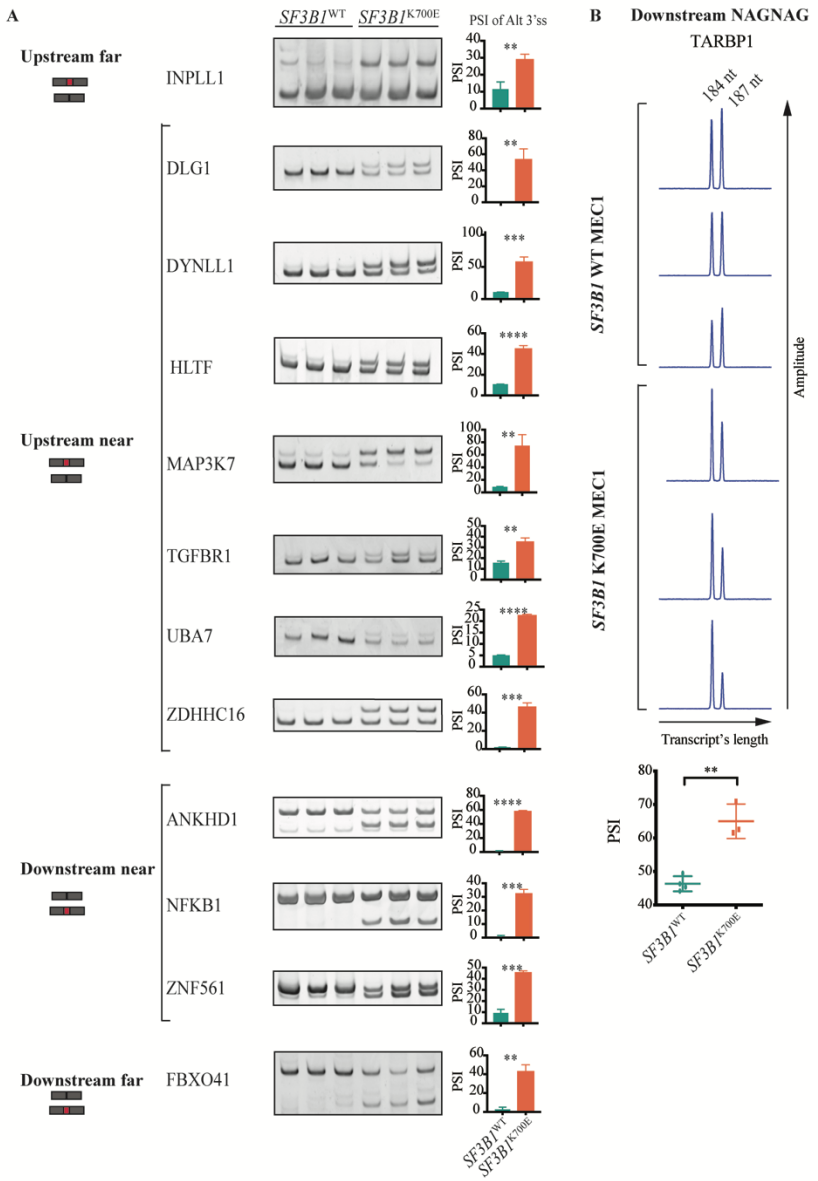


Figure I.12. Validation of Alt3'ss patterns associated with SF3B1 K700E in MEC1 cell lines. (A) RT-PCR validation of selected AS events belonging to the "upstream near", "upstream far", "downstream near" and "downstream far" 3'ss categories. Amplification products using oligonucleotides external to the pair of 3'ss analyzed were resolved by electrophoresis in native polyacrylamide gels. Results are shown for each of the wild type and K700E cell lines. Alt3'ss correspond to amplification products present only in K700E samples and match the expected sizes. The barplots next to the images represent the PSI values of the aberrant isoforms, quantified using ImageJ. Statistical significance of the changes was evaluated using unpaired

T-test (**: p value <0.01, ***: p value < 0.001, ****: p value < 0.0001). (B) Peaks of the aberrant (184 nucleotides) and canonical (187 nucleotides) isoforms of TARBP1 gene NAGNAG pattern obtained by fragment analysis of the RT-PCR amplification products using ABI3130 Genetic Analyzer (Applied Biosystems). PSI quantification of the area under the curve of the peaks (aberrant peak/(aberrant peak + canonical peak)) is shown in the lower panel. Statistical significance of the changes was evaluated using multiple t-test (**: p value <0.01).

1.3.2. Validation of the Alt3'ss patterns associated with *SF3B1*^{K700E} mutation using minigenes

To investigate the contribution of specific sequences to the activation of alternative 3'ss, e.g. the importance of predicted BP sequences, distance between canonical and Alt3'ss, etc. we designed several minigenes. For that purpose, we used an expression vector bearing a reporter derived from an Adenovirus Major Late transcript substrate, composed of 2 exons and an efficiently spliced intron [143]) under the control of a *Cytomegalovirus* promoter (pCMV 57 Δi Minx). We replaced the natural 3'ss region by different 3'ss regions comprising from 20 nucleotides upstream of the predicted BP (to ensure BP recognition and recruitment of SF3B1 on the anchoring site located 5' of the BP) to 5 nucleotides 3' of the more distal 3'ss for selected genes (Fig. I.13A). These included examples belonging to the 3'ss patterns previously described (MAP3K7 for "upstream near", INPPL1 for "upstream far", ZNF561 for "downstream near", FXBO41 for "downstream far" and TARPB1 for "downstream NAGNAG"). We transfected these minigenes by electroporation in *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 CLL cell lines, isolated RNA after 24h and analyzed the patterns of 3'ss utilization by RT-PCR and polyacrylamide gel electrophoresis electrophoresis or fragment analysis.

In the first proof of concept experiment, we observed 1) intron removal with an efficiency similar to that of the original AdML construct (Fig I.13B in MAP3K7, ZNF561 and TARBP1 constructs), 2) detectable use of Alt3'ss in the cases where the canonical and Alt3'ss were closer than 50 nucleotides (Fig I.13B in MAP3K7 and ZNF561 constructs); however, this approach did not allow the investigation of splicing profile for cases where the distance between canonical and Alt3'ss was longer than 50 nucleotides (Fig I.13B in FBXO41 and INPPL1 constructs); and 3) significant activation of Alt3'ss (upstream or downstream from the canonical 3'ss) in cells expressing the K700E mutant (Fig. I.13B, MAP3K7 and ZNF561). These results argue that the chimeric minigenes constructs are suitable to dissect the mechanisms behind the choice of alternative 3'ss upon mutation. With these results, we then focused on the 3 patterns where we could induce the splicing (MAP3K7 construct "upstream near", ZNF651 "downstream near" and TARBP1 "downstream NAGNAG").

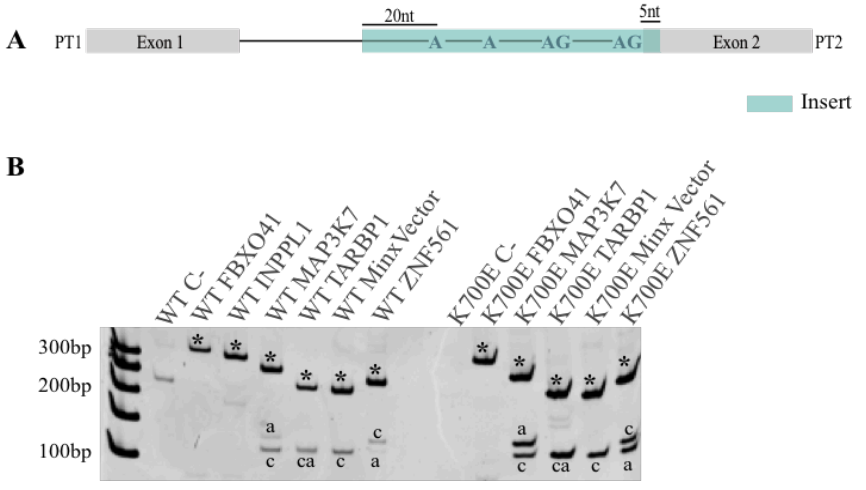


Figure I.13. Minigenes recapitulate some of the switches in 3'ss induced by *SF3B1* mutation. (A) Schematic representation of the AdML transcript where different Alt3'ss were introduced to replace its natural 3' ss region. The insert in green includes the regions with alternative 3'ss AGs, from 20 nucleotides upstream of the 5' A to 5 nucleotides downstream of the 3' AG. PT1 and PT2 correspond to exogenous sequences used for detection of the transcripts generated from the minigene, and also correspond to the sequences of the primers used for RT-PCR amplification (B) RT-PCR amplification of transcripts from the different indicated minigenes transfected in *SF3B1*^{WT} or *SF3B1*^{K700E} MEC1 cell lines. The upper band corresponds to the product of intron retention (*) and the lower bands correspond to the amplification products corresponding to the use of the canonical (c) and alternative (a) AGs. In the case of the NAGNAG pattern (ca) refers to the band were canonical and Alt3'ss are expected to migrate. RT-PCR products were separated by electrophoresis on a polyacrylamide gel, stained with Gel Red and detected with GelDocTM. Positions of molecular weight markers are shown. C indicates transfection control (no minigene).

To investigate the involvement of predicted BPs in 3'ss selection, we performed various mutations in three different minigenes.

First, in the case of ZNF561 which has an Alt3'ss 13 nts downstream of the canonical 3'ss, we observed preferential usage of the canonical site in *SF3B1* WT cells and an activation of the distal AG in *SF3B1* K700E cells (Figs. I.13B and I.14A-B). A to C mutation of the predicted branch adenosine of the BP proximal to the canonical AG (BP1 mutant) prevented the use of the distal AG in *SF3B1* K700E mutated cells, suggesting that this BP is necessary for the activation of the alternative 3'ss AG. In contrast, A to G mutation of the more distal predicted BP (BP2) led to the activation of the distal AG also in SF3B1 WT cells, arguing that recognition of this BP is required for preferential selection of the canonical 3'ss under normal conditions. Taken together, these results are compatible with a model in which 1) binding of U2 snRNP to the distal BP results in preferential use of the canonical 3' ss, while binding of U2 snRNP to the proximal BP allows use of both the canonical and the Alt3'ss, and

2) the presence of the K700E mutation allows occupancy of the proximal BP, while a wild type SF3B1 enforces preferential binding of U2 snRNP to the distal BP.

Second, in the case of MAP3K7 which has an Alt3'ss 20 nts upstream of the canonical AG, we observed a 70% usage of the canonical AG in WT cells and a 60% usage of the alternative 3'ss in K700E mutated cells (Fig.14D). A triple A mutation in the proximal predicted BP (BP1) led to a strong decrease in the use of the Alt3'ss in both cell lines, supporting a role for this BP in splice site selection. A to G mutation of the distal BP (BP2) adenosine did not affect the pattern of alternative splicing compared to the wild type minigene. This result suggests that only one of the two predicted BPs (the proximal one) effectively supports the activation of the upstream 3'ss observed upon mutation of *SF3B1*, which is consistent with the flexibility in 3'ss usage afforded by the proximal BP in ZNF561 discussed above. It is also possible that a different BP (even a non-consensus non-A BP) [144,145] plays a role in AS of MAP3K7.

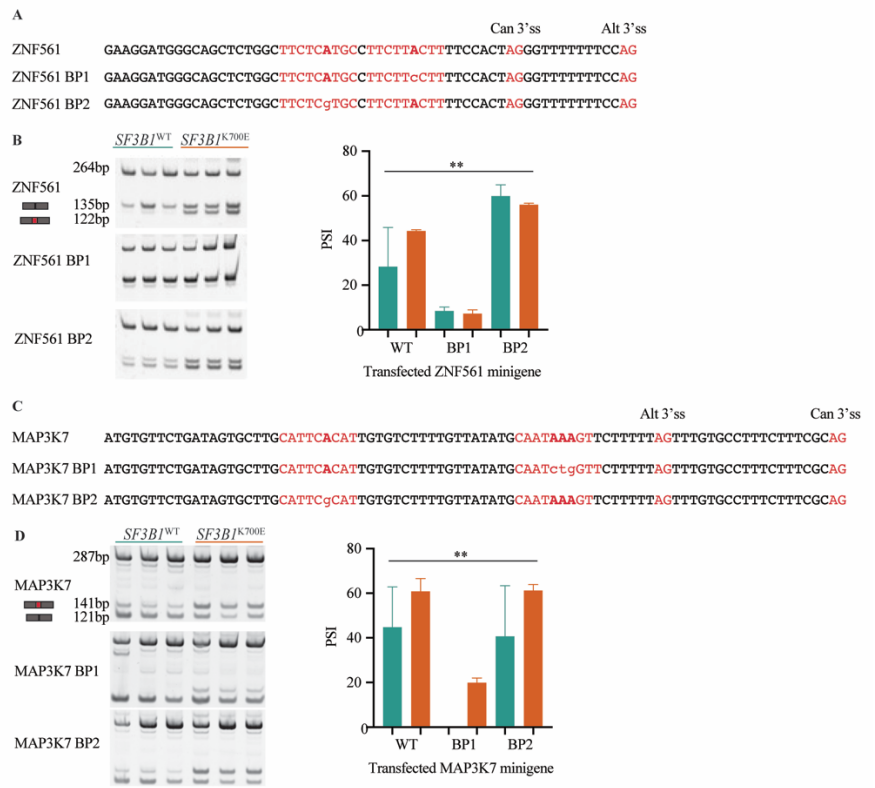


Figure I.14. Minigene mutagenesis reveals that specific predicted BPs are required for Alt3'ss selection. (A, C) Sequences of wild type and mutant alternative 3'ss regions corresponding to ZNF561(A) or MAP3K7 (C) included in the minigene constructs with the predicted BP (the adenosine in bold) and 3'ss in red. Single point mutations in BP1 or in BP2 are indicated in lower case characters. (B,D) Results of RT-PCR analysis of RNAs from *SF3B1*^{WT} or *SF3B1*^{K700E} MEC1 cell lines transfected with the indicated ZNF561(B) or MAP3K7 (D) minigenes detecting alternative and canonical spliced isoforms as well as unspliced RNAs. Amplification products were separated by polyacrylamide gel electrophoresis. Sizes of expected spliced products are indicated in the upper panels. Right panels: quantification of Alt3'ss PSI values using ImageJ. Green bars: *SF3B1*^{WT} cells. Orange bars: *SF3B1*^{K700E} cells. Kruskal-Wallis test (**: p <0.01).

Third, we used a minigene bearing the 3'ss of TARBP1 in order to validate and dissect the mechanism of regulation of a NAGNAG event. Upon transfection in MEC1 *SF3B1*^{WT} cells, both AGs were used, with a PSI of 60 for the downstream AG. In *SF3B1* mutant-cells, we observed an increase of 10% in its usage and a

corresponding decrease in the use of the proximal AG (Fig. 15). Interestingly, while we saw effects of BP mutations on the pattern of alternative splicing, we could not observe anymore a difference between MEC1 *SF3B1* WT and K700E cells lines. More specifically, upon mutation of the proximal BP, we observed a significant decrease in the use of the distal AG (aberrant) and an increase in the use of the proximal AG (canonical). Mutation of the distal BP led to roughly equal use of both AGs. These results suggest that the interplay between two alternative BPs is important to define 1) the relative use of the two 3'ss in the NAGNAG tandem, and 2) the effect of K700E mutation on AG selection.

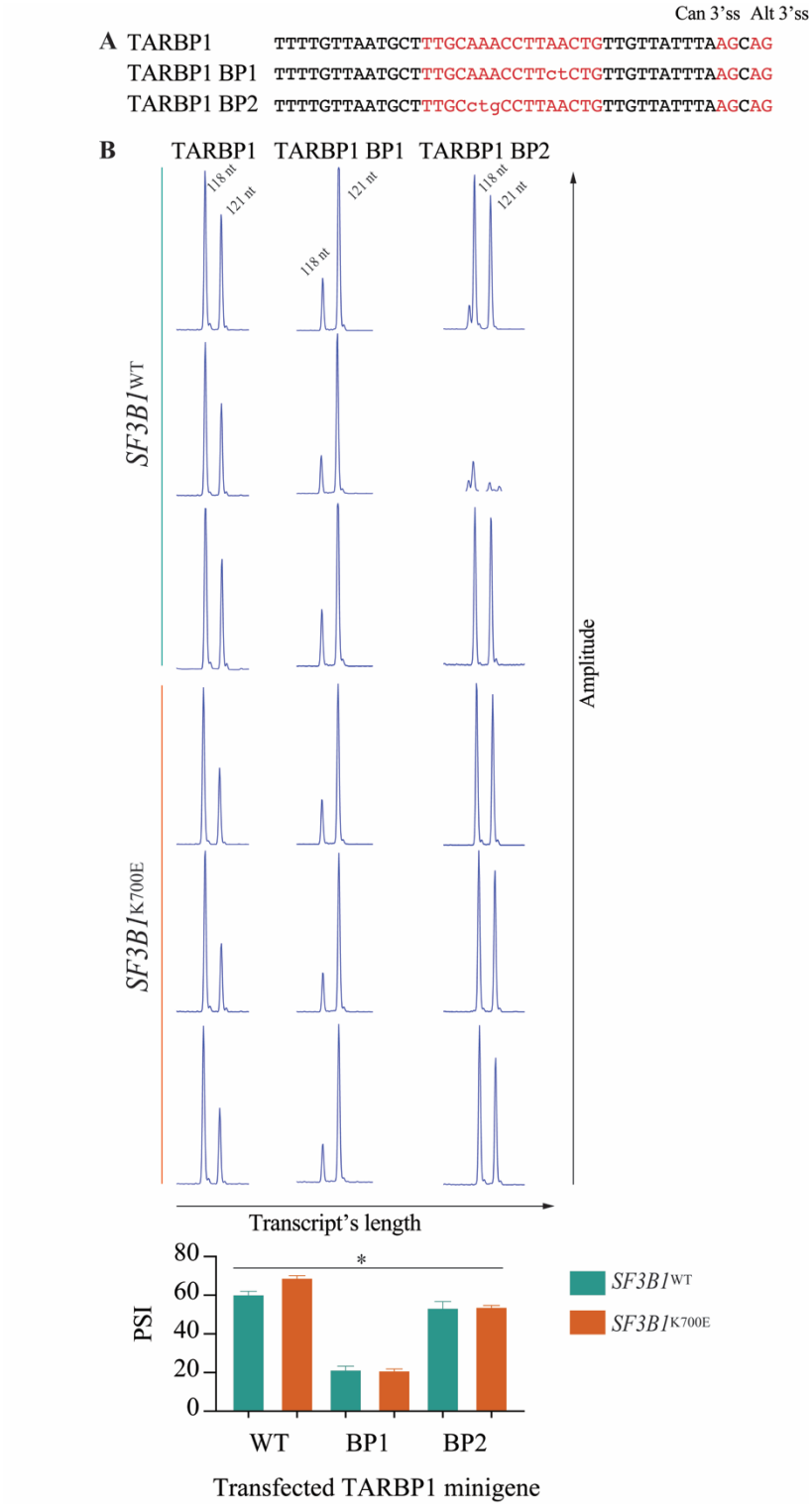

Can 3'ss Alt 3'ss
A TARBP1 TTTTGTTAATGCTTTGCAAACCTTAACTGTTGTTATTTAAGCAG
TARBP1 BP1 TTTTGTTAATGCTTTGCAAACCTTctCTGTTGTTATTTAAGCAG
TARBP1 BP2 TTTTGTTAATGCTTTGCctgCCTTAACTGTTGTTATTTAAGCAG
B
TARBP1
TARBP1 BP1
TARBP1 BP2
118 nt
121 nt
SF3B1WT
SF3B1K700E
Amplitude
Transcript's length
PSI
80
60
40
20
0
*
WT
BP1
BP2
Transfected TARBP1 minigene
SF3B1WT
SF3B1K700E

Figure I.15 TARBP1 minigene mutagenesis reveals that an interplay between two predicted BPs determines the choice between 3'ss AGs in a tandem NAGNAG sequence and the effects of *SF3B1* mutation. (A) Sequences of 3'ss region or TARBP1 minigene with the predicted BPs and 3'ss in red. Single point mutations in BP1 or in BP2 are indicated in lower case characters. (B) Peaks of the aberrant (118 nucleotides) and canonical (121 nucleotides) isoforms of TARBP1 gene's NAGNAG pattern obtained by fragment analysis of RT-PCR products from RNAs of *SF3B1*^{WT} or *SF3B1*^{K700E} MEC1 cell lines transfected with the indicated TARBP1 minigenes. PSI quantification of the area under the curve of the peaks (aberrant peak/(aberrant peak + canonical peak)) is shown in the lower panel. Statistical significance was assessed using a Kruskal-Wallis test (*: p value <0.05).

1.4. Gene enrichment analysis of the differentially spliced AS associated with *SF3B1* mutations

To assess whether the transcriptomic differences seen in the presence of *SF3B1* mutations are enriched in particular biological processes, we started by selecting all the AS events with a $|\Delta PSI| \geq 15$ detected with *vast-tools* or JuncExplorer in *SF3B1*-mutated CLL patient samples or *SF3B1*^{K700E} MEC1 cell lines. In total, we analyzed AS events from 324 genes in the patient samples and 1227 genes in the MEC1 cell lines.

Enrichment analyses performed using enrichGO package R package revealed that the genes affected are involved in diverse biological processes (Fig. I.16), including (i) immunological synapse which is essential for the interaction of CLL cells with the microenvironment; (ii) phosphatidylinositol biosynthetic processes, key components of the BCR signaling pathway; (iii) activation of MAPKK activity, involved in cell proliferation, transformation and apoptosis; (iv) chromatin organization, in line with the recent finding of hypomethylated regions in *SF3B1* mutated patient samples [146]; (v) cell cycle checkpoint and negative regulation of apoptotic signaling pathways, which could be crucial for cancer cell survival.

The variety of processes involved offers potential hints to elucidate the pathogenicity of *SF3B1* mutations in CLL, which most probably will be contributed by many of them.

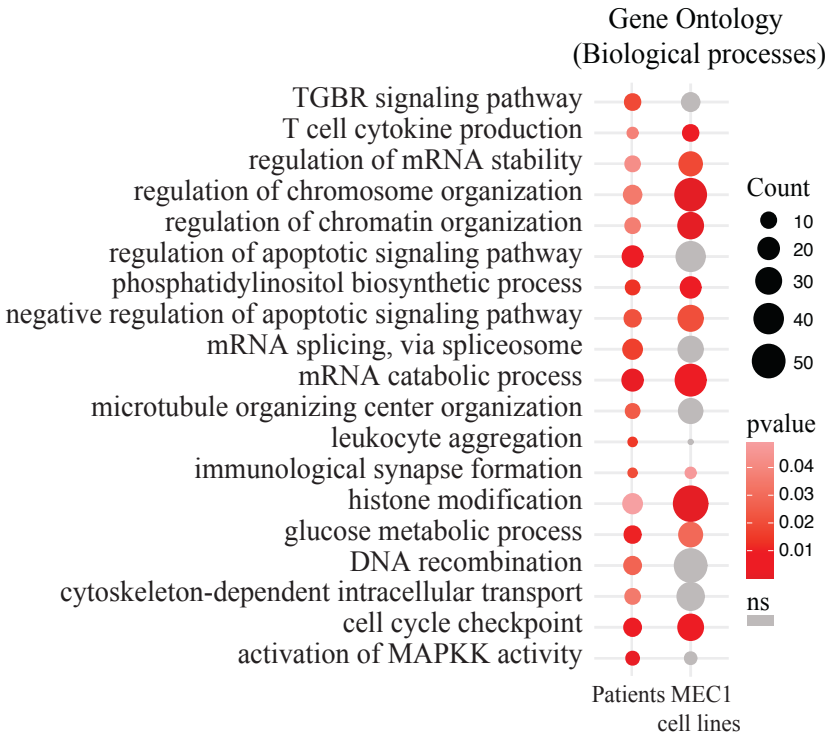


Figure I.16. *SF3B1* mutation induces AS changes in genes involved in diverse biological processes. Enriched Gene Ontology terms (enrichGO package) for biological processes in genes displaying differentially spliced events in *SF3B1* mutated CLL patient samples and *SF3B1*^{K700E} MEC1 cell lines. GO terms are indicated along with the statistical significance (p value, red circles, ns: not significant) and gene counts (black circles).

1.5. Possible use of ZDHHC16 AS as a biomarker to determine the functional impact of *SF3B1* mutations

As we have described in section 2.1, the CCF of specific *SF3B1* mutations impacts on the transcriptomic profile of a patient sample, while other specific mutations such as *SF3B1* M757T and E862K do

not end up altering the 3'ss selection profile in a detectable way. In this scenario, we thought it could be useful to find an aberrant isoform that could be used as a surrogate marker for the functional impact of *SF3B1* mutations. With this aim, we selected the events detected by JuncExplorer with dPSU ≥ 50 in *SF3B1*-mutated patient samples and selected the one with the lowest average PSU in *SF3B1* WT patient samples. With this, we aimed to select an event which could maximally discriminate between *SF3B1* WT and mutated samples. Using those criteria we identified an Alt3'ss (junction ID: chr10,+,97454799,97455639) in ZDHHC16 gene with a dPSU=50.93 and an average PSU=2.87 in *SF3B1* WT patient samples (Wilcoxon test p=6.71E-10). In fact, this event has been previously described in the literature and has been used to understand the impact of *SF3B1* mutations in 3'ss usage using minigene assays[45] as well as to validate the functional impact of particular in-frame deletions in the *SF3B1* gene[147].

We validated the use of both the canonical and cryptic ZDHHC16 isoforms in CLL samples by RT-qPCR assays. While, the detection of the canonical isoform was possible in all tested samples (39 *SF3B1* WT samples, 29 *SF3B1* mutated samples with high CCF (CCF≥50), 11 with low CCF (50>CCF≥12) and 9 with very low CCF (CCF<12)), the cryptic ZDHHC16 isoform was only detected in mutated samples with CCF > 12 (except in *SF3B1* E862K or M757T cases) and 3 samples with CCF <12. (Fig. I.17). These results confirmed that cryptic ZDHHC16 isoform detection can be used to detect *SF3B1* mutations that alter the 3'ss selection profile in patient samples.

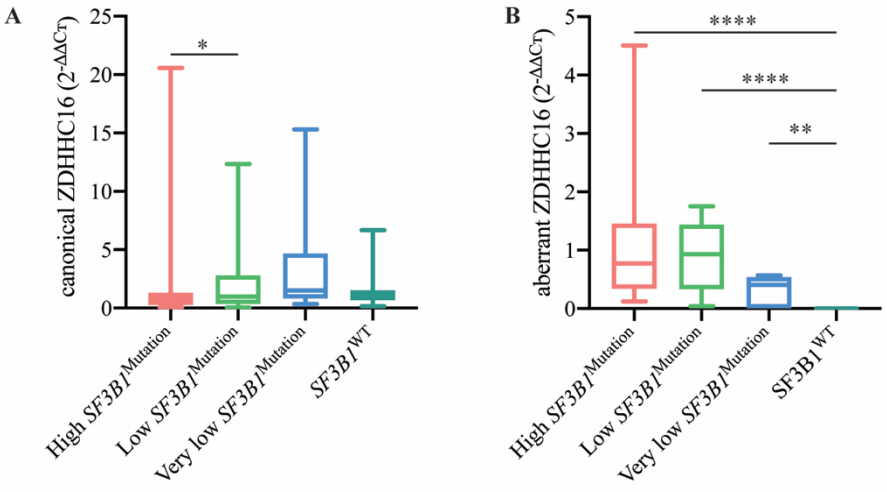


Figure I.17. Activation of cryptic 3'ss in the gene *ZDCHH16* is a surrogate marker of functional alternative splicing activity induced by *SF3B1* mutation. Canonical (A) and cryptic (B) 3'ss usage in the *ZDCHH16* gene measured by qPCR in *SF3B1* WT and mutated samples. The relative expression of each isoform was quantified by the comparative cycle threshold (Ct) method (ΔΔCt) using *RPLPO* as an endogenous control. mRNA expression levels are presented as arbitrary quantitative PCR units, normalized to the median of the WT samples. In the case of ZDHHC16 cryptic isoform, which was not detected in the WT samples, the median of the *SF3B1* mutated samples was used in the normalization. Statistical significance was assessed using a Mann-Whitney test (*: p value <0.05, **: p value < 0.001, ***: p value < 0.0001

1.6. Gene expression analysis

We also analyzed gene expression differences between SF3B1WT and SF3B1 mutated samples using Deseq2(see Material and Methods)[148]. In total, 405 genes were differentially expressed in patient samples and 940 in MEC1 cell lines (|FoldChange| > 1.5 and adjusted p value < 0.05). As an exploratory analysis, we looked for the enrichment in gene ontology categories for biological processes across these genes using enrichGO R package (see Material and Methods). The most enriched categories in patient samples were related to nuclear division and chromosome segregation. In MEC1 cell lines, the differentially expressed genes between *SF3B1*^{WT} and

SF3B1^{K700E} were related to cell adhesion, T cell activation and cytokine production and secretion (Fig I.18A). This enrichment in the cytokines category was also observed in patient samples when using KEGG terms ("Cytokine-cytokine receptor interaction", adjusted p value<0.05 and q value=0.02).

To visualize the relatedness between categories and whether they were upregulated or downregulated in MEC1 *SF3B1*^{K700E} samples, we plotted networks for the top upregulated and downregulated categories (Fig I.18B). We found interesting relationships between the categories and intriguingly, many of them (for example, cell adhesion and cytokine response) appeared both in the clusters of upregulated and of downregulated genes. In contrast, the genes involved in cell proliferation only appeared in the cluster of downregulated genes.

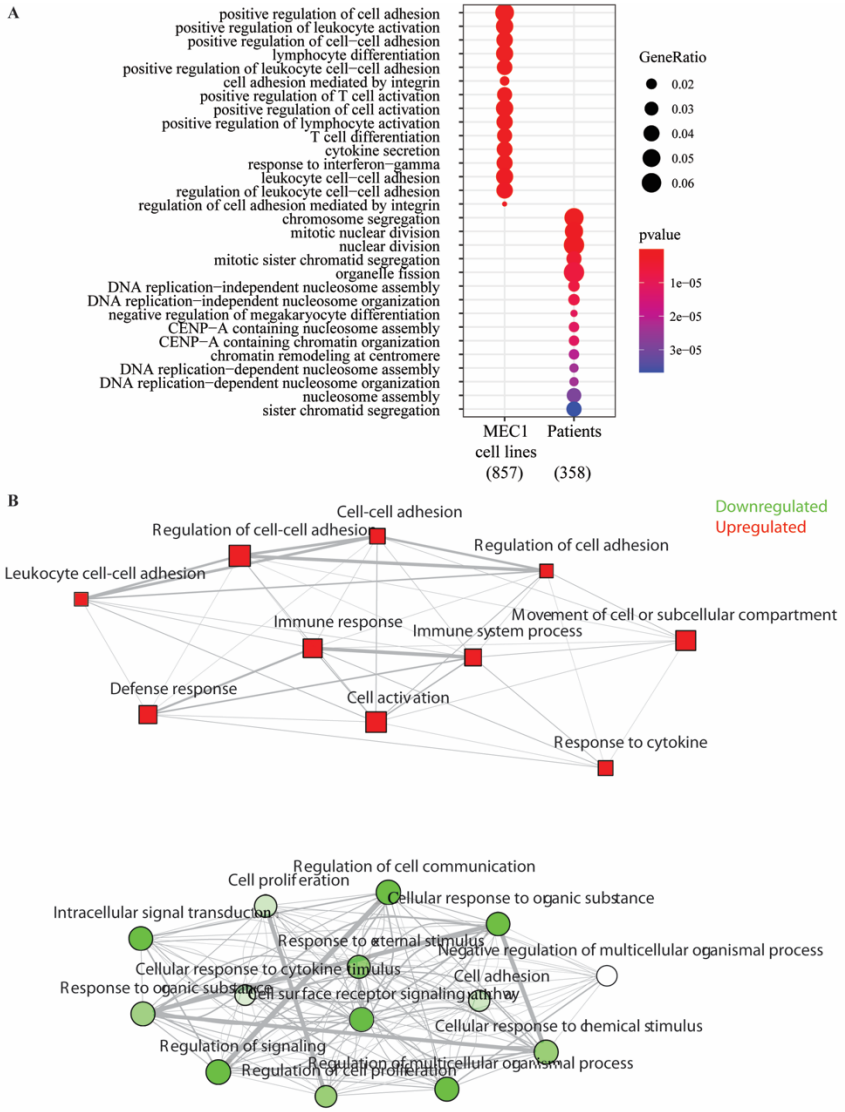


Figure I.18. Gene Ontology enrichment analysis in differentially expressed genes between *SF3B1*^{WT} and *SF3B1*^{K700E} CLL samples. (A) The 15 most enriched gene ontology terms (enrichGO package) for biological processes in differentially expressed genes in *SF3B1* mutated CLL patient samples and *SF3B1*^{K700E} MEC1 cell lines. GO terms are indicated along with the statistical significance (p value, red and blue circles) and gene counts (black circles). (B) Networks between the upregulated and downregulated genes found in the GO terms in *SF3B1*^{K700E} MEC1 cell lines plotted with iDEP94[149].

2. Functional relevance of selected AS events

2.1. Role of MAP3K7 alternative spliced isoforms

We selected MAP3K7 Alt3'ss (HsaALTA1027057-2/4) as a top candidate to study its functional role in CLL for various reasons, (i) our RNA-seq analyses showed high ΔPSI in *SF3B1*^{WT} vs *SF3B1*^{K700E} CLL patient samples (60%) and in *SF3B1*^{WT} vs *SF3B1*^{K700E} MEC1 CLL cell lines (70%), (ii) the Alt3'ss identified in the MAP3K7 gene is located upstream of exon 5, which encodes part of the kinase domain and its usage is predicted to result in the disruption of the coding sequence and degradation of the corresponding transcript by NMD (iii) we found that many genes involved in the activation of MAPKK activity activity are affected by differentially spliced events in *SF3B1*-mutated samples (Fig I.15) suggesting that these events might dysregulate this pathway and (iv) recent evidence links MAP3K7 AS with different aspects of tumorigenesis: on one hand, MAP3K7 has known roles in regulating NF-κB signaling[150] and reduced MAP3K7 protein levels in *SF3B1*-mutated NALM-6 and K562 cell lines have been related to increased nuclear p-p65 levels and enhanced NF-κB activation[151], which in the case of *SF3B1* mutated breast cancer accelerates cell migration and tumor development together with AKT activation[152]; on the other hand, it accelerates erythroid differentiation and apoptosis, producing anemia in MDS by deactivation of its downstream target P38 and premature down-regulation of P38's downstream target GATA1 transcription factor during erythroid differentiation[57,153].

Validation of MAP3K7 Alt3'ss activation in CLL patient samples and *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 CLL isogenic cell lines

We validated the activation of MAP3K7 Alt3'ss (HsaALTA1027057-2/4) in *SF3B1*^{WT} vs *SF3B1*^{K700E} MEC1 CLL isogenic cell lines by RT-PCR and quantified the canonical and Alt3'ss splicing isoforms by RT-qPCR in CLL patient samples (Fig 1.19). In addition, we also validated an Alt3'ss event in NFKB1 (HsaALTA1015253-2/3), a gene from the NF- κ B signaling pathway (Fig I.18).

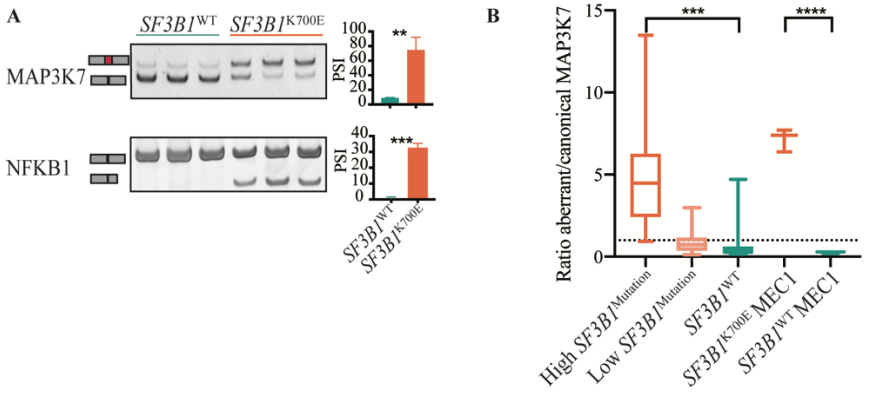


Figure I.19. Validation of AS changes in *MAP3K7* and *NFKB1* genes, associated with SF3B1 mutations. (A) Analysis of *MAP3K7* and *NFKB1* canonical and Alt3'ss usage by RT-PCR from RNA samples of *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 CLL cell lines. The positions of the different isoforms are indicated. Bar plots (right part of panel A, indicate the PSI values of the Alt3'ss isoforms, quantified using ImageJ. Statistical significane was estimated using an unpaired t-test (**: p value <0.01; ***: p value < 0.001). (B) Ratio between Alt3'ss and canonical MAP3K7 isoforms measured by RT qPCR in patient samples and in *SF3B1*^{WT} vs *SF3B1*^{K700E} MEC1 CLL cell lines. Statistical significance was assessed using unpaired t- tests (***: p value <0.001; ****:p value < 0.0001). High *SF3B1* mutation corresponds to CCF > 50%. Low *SF3B1* mutation corresponds to CCF < 50%.

Experiments using cycloheximide, an inhibitor of NMD, have shown that the Alt3'ss identified in the MAP3K7 is targeted by NMD elimination[144]. Consistent with this finding, we observed

decreased expression of the MAP3K7 protein in SF3B1^{K700E} MEC1 CLL cell lines compared to *SF3B1*^{WT} cells (Fig I.20).

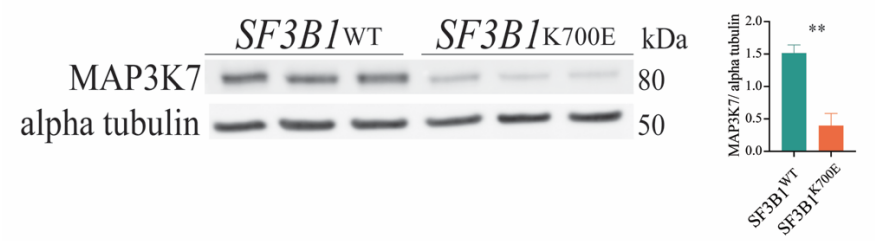


Figure I.20. MAP3K7 protein expression is decreased in *SF3B1*^{K700E} MEC1 cell lines. Western blot analysis of MAP3K7 expression, using alpha tubulin as a loading control protein. Quantification was carried out using Multi Gauge Software and normalized by calculating the MAP3K7/alpha tubulin ratio. Statistical significance was assessed using unpaired t- tests (**: p value <0.01)

2.2. Effects of MAP3K7 aberrant splicing in NF-κB pathway

To confirm the NF-κB hyperactivation reported by Lee et al in *SF3B1*-mutated cancers, including CLL[32], we measured the basal NF-κB activity in *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 CLL cell lines. Nuclear protein extracts were obtained and used to calculate the phosphorylated p65/totalp65 ratio by western blot, as well as an ELISA-based TransAM NF-κB Family Transcription Factor Assay. A trend towards higher basal level of phosphorylated p65 was detected in *SF3B1*^{K700E} samples using either of these assays (Fig I.21), although the differences did not reach statistical significance, possibly due to the low number of samples analyzed.

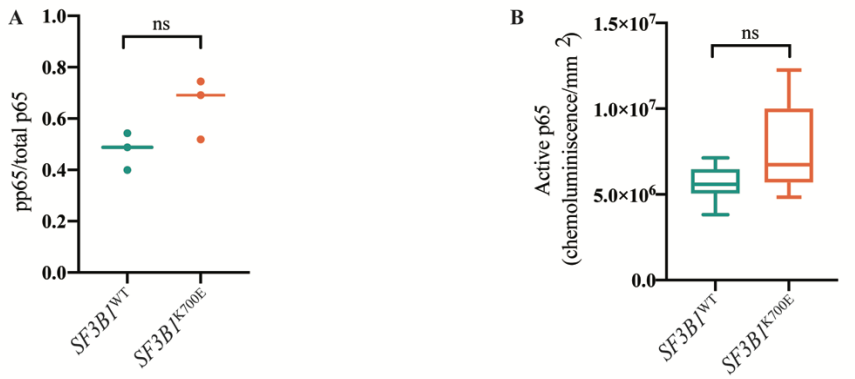


Figure I.21. Activation of the NFKB in *SF3B1*-mutated samples. (A)Western blot quantification of phosphorylated p65 (pp65) levels relative to total p65 amounts in nuclear extracts of *SF3B1*-mutated and wild type MEC1 cell lines. Statistical significance was assessed using unpaired t test. (B) Basal activated p65 levels of the same samples as in (A) measured using an ELISA-based chemiluminescence kit. Statistical significance was assessed using unpaired t test.

Differential gene expression analysis of *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 CLL cell lines revealed that "regulation of I-kappaB kinase/NF-κB signaling" (GO:0043122) gene ontology term was among the enriched biological processes (adjusted p value = 0.02 and q value = 0.01) (Fig. I.22). In addition, in the analysis of genes having differentially spliced events between *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 cell lines, "I-κB kinase/NF-κB signaling" (GO:0007249) and "regulation of I-κB kinase/NF-κB signaling" (GO:0043122) gene ontology terms were enriched with an adjusted p value = 0.0039 and q value = 0.0033. Alternative events with a ΔPSI≥15 were found in the following genes from the I-κB kinase/NF-κB signaling pathway: *NLRX1*, *MAP3K7*, *PIDD1*, *ECT2*, *CC2D1A*, *TNIP1*, *MIER1*, *EEF1D*, *CARD19*, *TRAF5*, *MAVS*, *TRIM25*, *NFKBIA*, *TRIM38*, *CTNNB1*, *GPR89A*, *INAVA*, *TRIM32*, *CARD9*, *OPTN*, *MALT1*, *UBE2V1*, *CPNE1*, *RHOH*, *ERC1* and *IKBKG*. The overlap between the genes with AS changes and differential expression in I-κB

kinase/NF-κB signaling in MEC1 *SF3B1*^{K700E} cell was of 3/26, with *MAP3K7*, *TRAF5* and *ERC1* being present in both of them.

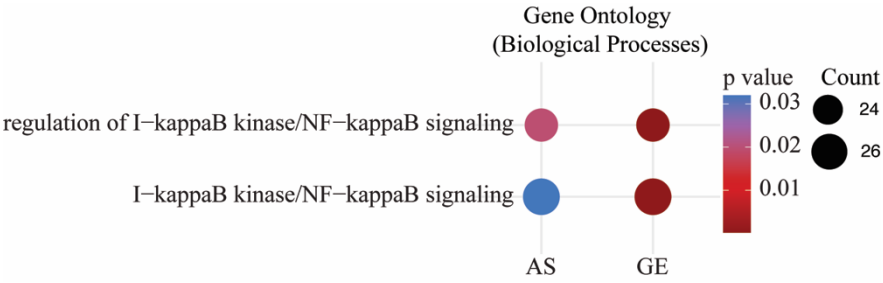


Figure I.22. *SF3B1* mutation induces AS and GE changes in genes involved in NF-κB signaling pathway. Enriched Gene Ontology terms (enrichGO R package) for genes displaying differentially spliced events or differences in gene expression in *SF3B1*^{K700E} MEC1 cell lines. GO terms are indicated along with the statistical significance (p value, red and blue circles) and gene counts (black circles).

2.3. Effects of MAP3K7 aberrant splicing on p38 phosphorylation

To validate previous findings of MAP3K7 aberrant splicing on p38 deactivation, we performed a western blot analysis of total extracts from *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 CLL cell lines. The first time we did this experiment, we obtained a very clear result where the phosphorylated form of p38 was absent in the *SF3B1*^{K700E} cell lines (Fig I.23). However, this result was not reproducible, and therefore the analyses of the effects of MAP3K7 aberrant splicing / low protein levels on p38 phosphorylation in CLL were not conclusive.

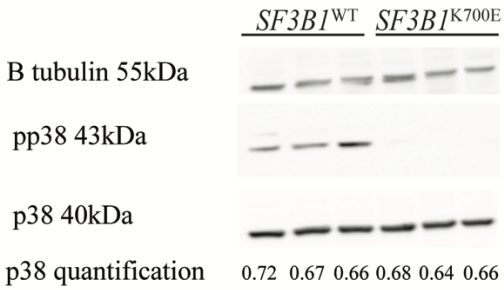


Figure I.23. Western blot analyses of phosphorylated p38(pp38) levels in total cell extracts from *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 CLL cell lines. The decreased phosphorylation of p38 (pp38) was not reproducible in other experiments. p38 was quantified using Multi Gauge Software and normalized to β tubulin amount.

Chapter II:

Study of the therapeutic effects of SF3B1-binding H3B-8800 in *in vitro* and *in vivo* models of Chronic Lymphocytic Leukemia

1. Cytotoxic effects of H3B-8800 in models of Chronic Lymphocytic Leukemia

H3B-8800 H3B-8800 is a splicing inhibitor of reported cytotoxic activity in myeloid malignancies and preferential lethality toward cancer cells bearing splicing factor mutations[63]. Interestingly, it has already been tested in a clinical trial (NCT02841540) in MDS, AML and CMML, displaying an acceptable safety profile[154]. In this context, we wanted to (i) assess the effects of H3B-8800 in *in vitro* and *in vivo* models of CLL, (ii) determine whether H3B-8800 displays preferential effects in the presence of *SF3B1* mutations, (iii) study the alternative splicing changes induced by H3B-8800 treatment in CLL cells and (iv) investigate potential combinatorial therapies for H3B-8800 in CLL.

In order to assess the cytotoxic effect of H3B-8800 splicing modulator, we used Peripheral Blood Mononuclear Cells (PBMCs) from healthy donors and CLL patients. After 48 hours incubation with H3B-8800 at 75nM we observed that B lymphocytes (CD19+) from CLL patients were more sensitive to H3B-8800 than B lymphocytes from healthy donors (Fig II.1). The cytotoxicity in CD4+ T lymphocytes and CD8+ T lymphocytes was rather minor, with higher than 90% relative viabilities (Fig. II.1), but the effect in healthy donor's T CD8+ lymphocytes was higher than in CLL patients' T CD8+ lymphocytes.

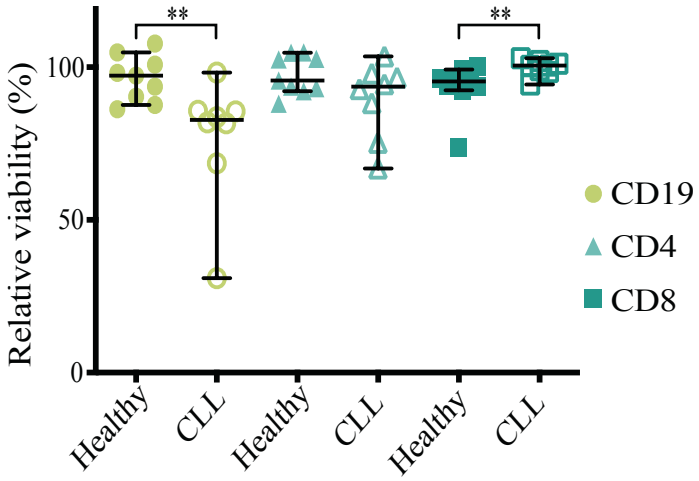


Figure II.1 Relative viability of CD19+ cells, CD4+ cells and CD8+ cells from healthy or CLL donors after 48h treatment with 75nM H3B-8800. B cells (CD19+) are represented with circles, CD4 + T lymphocytes with triangles and CD8+ T lymphocytes with squares. Assessment of statistical significance was carried out using Mann-Whitney U test (**: p value <0.01).

1.1. Sensitivity to H3B-8800 in SF3B1-mutated CLL primary samples and cell lines

Previous work reported higher efficacy of H3B-8800 in various cell lines overexpressing different *SF3B1* mutations and in *SF3B1*^{K700E} K562 cells compared to WT cells[63]. To investigate whether this is also the case in CLL, we tested the effects of H3B-8800 in *SF3B1* WT and mutated CLL primary cells (Fig II.2A). As the mutational percentage of tumor cells is known to impact on the prognosis[105] and we have also shown that in samples with low *SF3B1* mutant CCF we cannot distinguish changes in AS associated with *SF3B1* mutations (Chapter I.2.1), we took into account the CCF of the primary samples in this analysis. Along the lines of what was reported in other systems, we found that in patients with a CCF>80%, the median viability upon incubation with 75nM H3B-8800 for 48

hours was 70% compared to 79% in patient samples with CCF<80%, or 81% viability in the *SF3B1* WT patient samples. The biological characteristics of the CLL patient samples analyzed are shown in Supplementary Table 2.

To avoid potential biases due to intratumoral and intertumoral heterogeneities in CLL patient samples, we also tested the effect of H3B-8800 in isogenic *SF3B1*^{K700E} and *SF3B1*^{WT} MEC1 CLL cell lines (Fig II.2B). H3B-8800 had an impact in both *SF3B1*^{K700E} and *SF3B1*^{WT} cell lines but the cytotoxicity was more pronounced in *SF3B1*^{K700E} cells in a dose-dependent manner. Given that 75nM H3B-8800 killed nearly 50% of the MEC1 *SF3B1*^{K700E} cells after 48 hours, we defined these conditions to study the effects H3B-8800 *in vitro* (Fig. II.B).

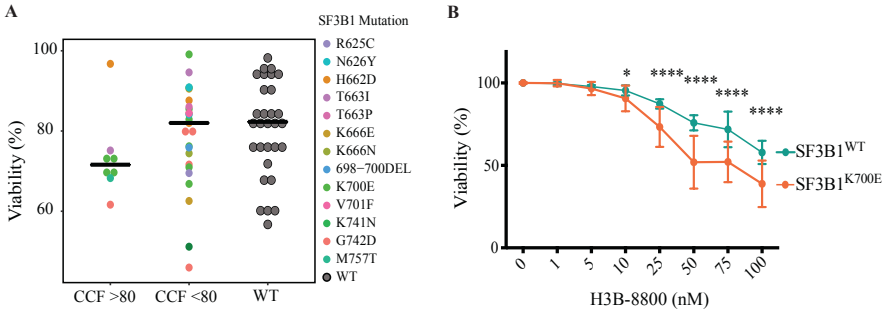


Figure II.2. Stronger effects of H3B-8800 on *SF3B1*-mutated CLL primary samples and MEC1 cell lines. (A) Analysis of viability of primary CLL cells after 48h treatment with 75 nM H3B-8800 measured by Annexin V and propidium iodide negativity by flow cytometry. *SF3B1* mutated samples are divided in two groups according to their cancer cell fraction (CCF). The identity of the mutations is indicated by the color code. (B) Analysis of viability of CLL cell lines after 48h treatment with the indicated concentrations of H3B-8800, measured by Annexin V and propidium iodide negativity by flow cytometry. Three biological replicates are shown for each group (*SF3B1*^{WT} refers to WT1, WT2, WT3 and *SF3B1*^{K700} refers to MUT1, MUT2, MUT3).. The statistical significance of the comparison between the viability of *SF3B1*^{WT} and *SF3B1*^{K700E} cells was assessed using Mann-Whitney U test (*: p value <0.05, ****: p value <0.0001).

1.1.1. H3B-8800 effects in co-culture conditions

In the proliferation centers of the lymph nodes, CLL cells are in close contact with accessory cells such as T cells[155] and nurse-like cells (lymphoma-associated macrophages)[156], which favor CLL growth and BCR activation[157]. The tumor microenvironment is also known to have an impact on the response to therapies and resistance of malignant B cells to therapeutic drugs[158]. Therefore, we tried to assess the effect of the microenvironment by measuring the cytotoxicity of H3B-8800 in co-culture conditions with bone marrow-derived mesenchymal (HS5) and human follicular dendritic cell-like (HK) cell lines, in order to mimic the germinal center of lymphoid follicles and support B cell proliferation. Under these conditions, we observed that H3B-8800-induced cytotoxicity and the preferential killing for *SF3B1* mutated samples persisted, both for primary CLL cells and for MEC1 cell lines (Fig II.3), arguing that the microenvironment does not play a fundamental role in the response or resistance to H3B-8800 treatment.

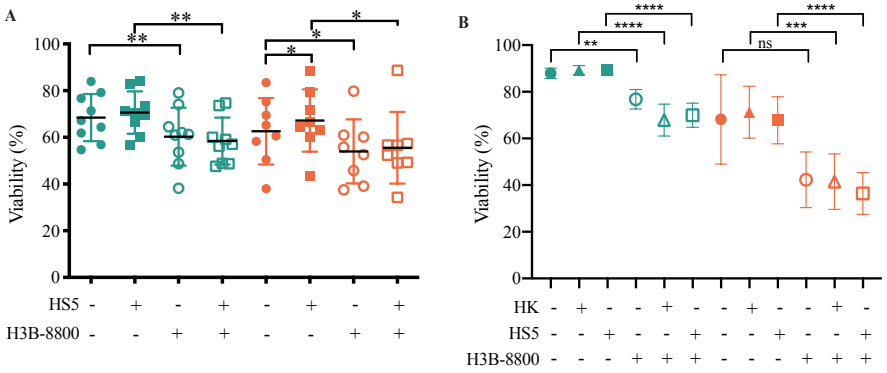


Figure II.3. H3B-8800 effects persist in coculture conditions. (A) Viability of primary cells after treatment with 50nM H3B-8800 for 48h in co-culture conditions with HS5 (bone marrow stromal cell line) cells. Green symbols correspond to *SF3B1*^{WT} samples, orange symbols to *SF3B1* mutated samples across different

conditions specified in the x axis. Statistical significance assessment using Wilcoxon tests: *: p value <0.05, **: p value <0.01. (B) Viability of cell lines viability after treatment with 75 nM H3B-8800 for 48h in coculture conditions. HK: lymph node stromal cells. HS5: bone marrow stromal cell line. Green symbols correspond to *SF3B1*^{WT} samples, orange symbols to *SF3B1* mutated samples across different conditions specified in the x axis.Statistical significance assessment using Mann-Whitney test: **: p value < 0.01, ***: p value <0.001, ****: p value <0.0001, ns: non-significant.

1.1.2. H3B-8800 delays leukemic infiltration in NSG *in vivo* mice model

To test the efficacy of H3B-8800 *in vivo*, we carried out a pilot study in 24 NSG mice in which we injected *SF3B1*^{WT} or *SF3B1*^{K700E} MEC1 isogenic cell lines subcutaneously (12 mice each group). Once the subcutaneous tumor's volume reached 150mm^3, we started the administration of either vehicle or H3B-8800 at 8mg/kg daily for 10 days. However, a drastic body weight reduction was detected in the mice receiving H3B-8800 at 8mg/kg forcing us to sacrifice them (Fig II.8A). Post mortem, we observed a reduction on spleen size as well as increased gastric size where the stomachs were brimmed with food (Fig II.8B-C). This experiment led us to reduce the dose to 6mg/kg in the next experiment.

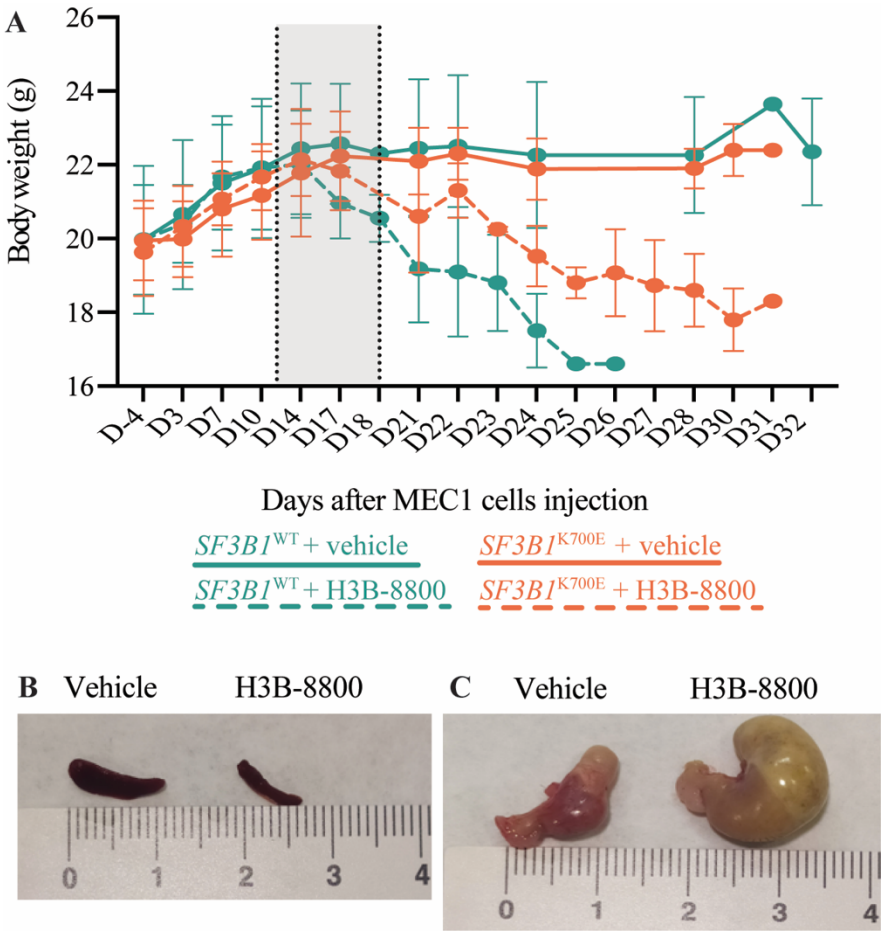


Figure II.8. Pilot study testing the effects of administration of H3B-8800 at 8mg/kg in NSG mice. (A) Body weight of mice harboring CLL subcutaneous xenografts before and after administration of H3B-8800. 6 mice were included in each group. The different treatments and xenografts are indicated by color and line codes. Highlighted in grey are the days in which the treatment was initiated once the tumor volume reached 150mm^3, which lasted 10 days. (B) Examples of spleens and (C) stomachs from mice treated with vehicle or H3B-8800.

In order to visualize the CLL cells injected in mice, we produced and injected luciferase-expressing MEC1 *SF3B1*^{K700E} or *SF3B1*^{WT} cell lines in 17 and 16 NSG mice, respectively, and followed leukemic cell infiltration by bioluminescence. Ten days after the injection, when the leukemic infiltration around the iliac crests in

mice injected with MEC1 cell lines was detectable, we started the administration of H3B-8800 at a dose of 6mg/kg (or vehicle) daily and orally for 10 days and the mice were sacrificed at day 21. One mouse bearing *SF3B1*^{WT} and treated with vehicle died at day 19.

Consistent with previous studies reporting *SF3B1* mutation-related *in vivo* growth disadvantage of Nalm-6 cell lines[129], the results of luciferase activity indicated lower leukemic infiltration capacity of *SF3B1*-mutated MEC1 cell lines (Figs. II.4 and II.5). H3B-8800 administration delayed MEC1 CLL cells infiltration in both *SF3B1*^{K700E} and *SF3B1*^{WT} groups, but the lowest infiltration was observed in the H3B-8800 treated *SF3B1*^{K700E} group (Fig II.4).

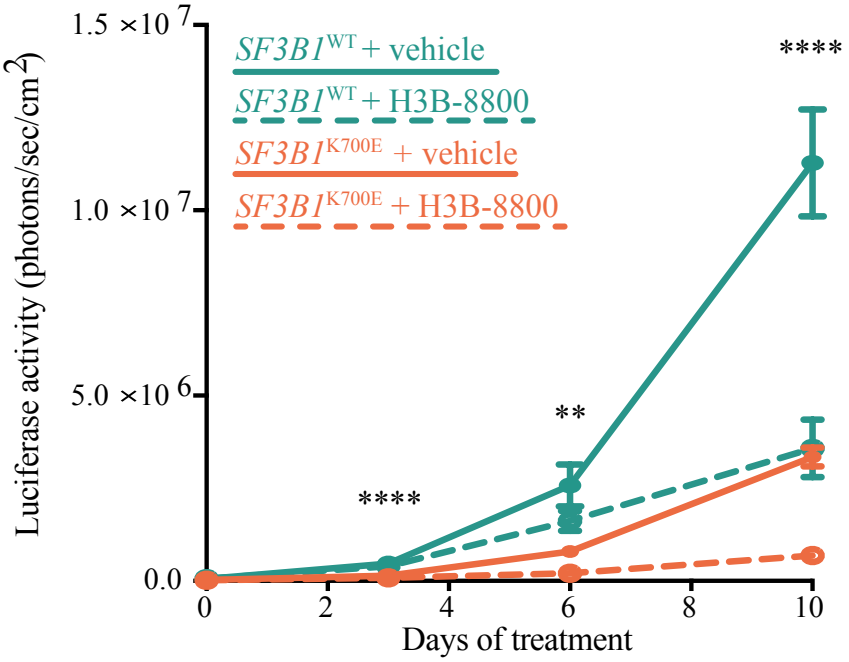


Figure II.4 H3B-8800 delays CLL leukemic infiltration. Quantification of luciferase activity of *SF3B1* WT and K700E CLL isogenic lines upon xenograft implantation in NSG mice and treatment with vehicle or H3B-8800 at 6mg/kg. The color and line codes indicate the different experimental conditions and treatments.

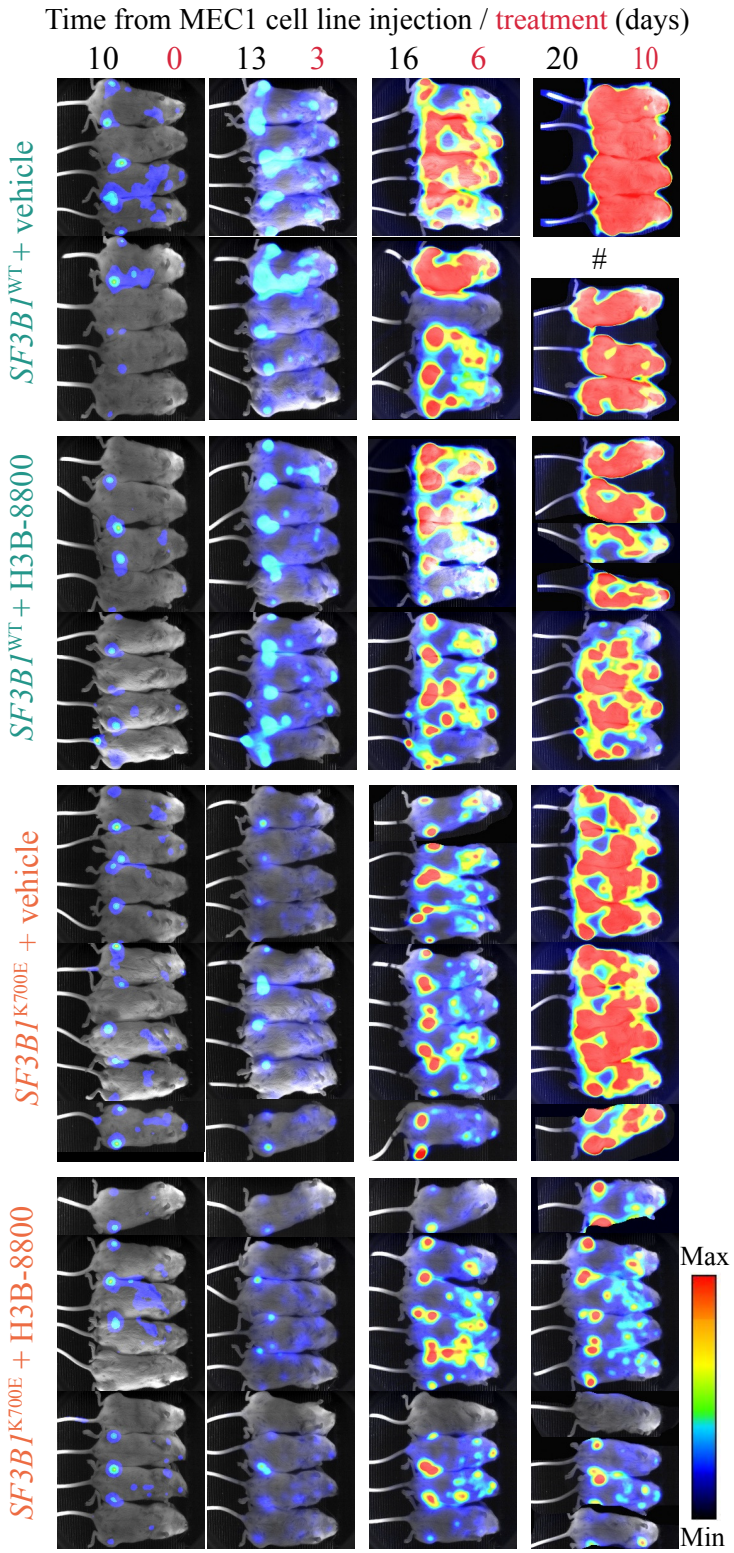

Time from MEC1 cell line injection / treatment (days)
10
0
13
3
16
6
20
10
#
SF3B1WT + vehicle
SF3B1WT + H3B-8800
SF3B1K700E + vehicle
SF3B1K700E + H3B-8800
Max
Min

Figure II.5. Analysis of leukemic infiltration of WT or K700E SF3B1 isogenic luciferase-expressing MEC1 cell lines and effects of H3B-8800 administration. Luciferase activity was captured using AEQUORIA fluorescence/luminescence imaging system. # symbol replaces the mouse which died before the experimental end point.

After sacrificing the mice, we measured leukemic infiltration in bone marrow, peripheral blood, spleen and liver. The mean values of bone marrow leukemic infiltration (measured by the presence of CD19+ B lymphocytes by flow cytometry) were 54% and 41% for the *SF3B1*^{K700E} and *SF3B1*^{WT} groups treated with vehicle, respectively. H3B-8800 treatment led to a reduction in the *SF3B1*^{K700E} group only (from 54% to 7%, Fig II.6A), indicating a preferential effect of the drug on SF3B1 mutant cells. Leukemic infiltration in peripheral blood was rather low (8% in *SF3B1*^{WT} groups and 2.6% in *SF3B1*^{K700E} groups) and reduction of infiltration (from 8% to 3%) after H3B-8800 treatment was only observed in the *SF3B1*^{WT} group (Fig II.6B).

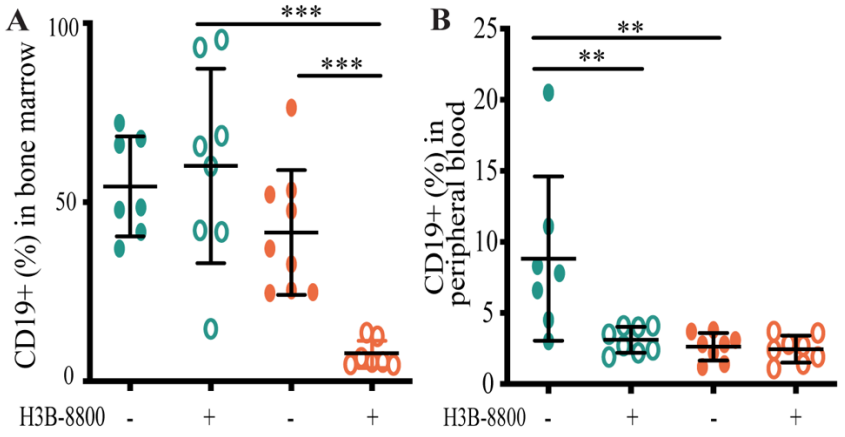


Figure II.6. Analysis of leukemic infiltration in the bone marrow (A) and peripheral blood (B), measured as the percentage of CD19+ cells in flow cytometry assays, for *SF3B1*^{WT} and *SF3B1*^{K700E} xenograft-implanted mice, treated with vehicle of H3B-8800, as indicated. *SF3B1*^{WT} samples are indicated in green, *SF3B1*^{K700E}

samples in orange. Statistical significance was assessed using Mann-Whitney test (**: p value <0.01 ; ***: p value <0.001).

In addition, liver and spleen weight were higher in the vehicle-treated *SF3B1*^{WT} group compared to the *SF3B1*^{K700E} group. H3B-8800 treatment decreases liver and spleen weights in both groups. (Fig II.7A, II.7B, II.7C). CD79a immunohistochemical staining to detect B lymphocyte infiltration in spleen showed (i) massive tumor infiltration in the SF3B1^{WT} group, (ii) lower infiltration in the SF3B1^{K700E} group and (iii) a reduction of the infiltration in both groups after H3B-8800 treatment (Fig II.7D).

We conclude that *SF3B1*^{K700E} CLL MEC1 cells display lower infiltration capacity and higher sensitivity to H3B-8800, at least regarding bone marrow infiltration.

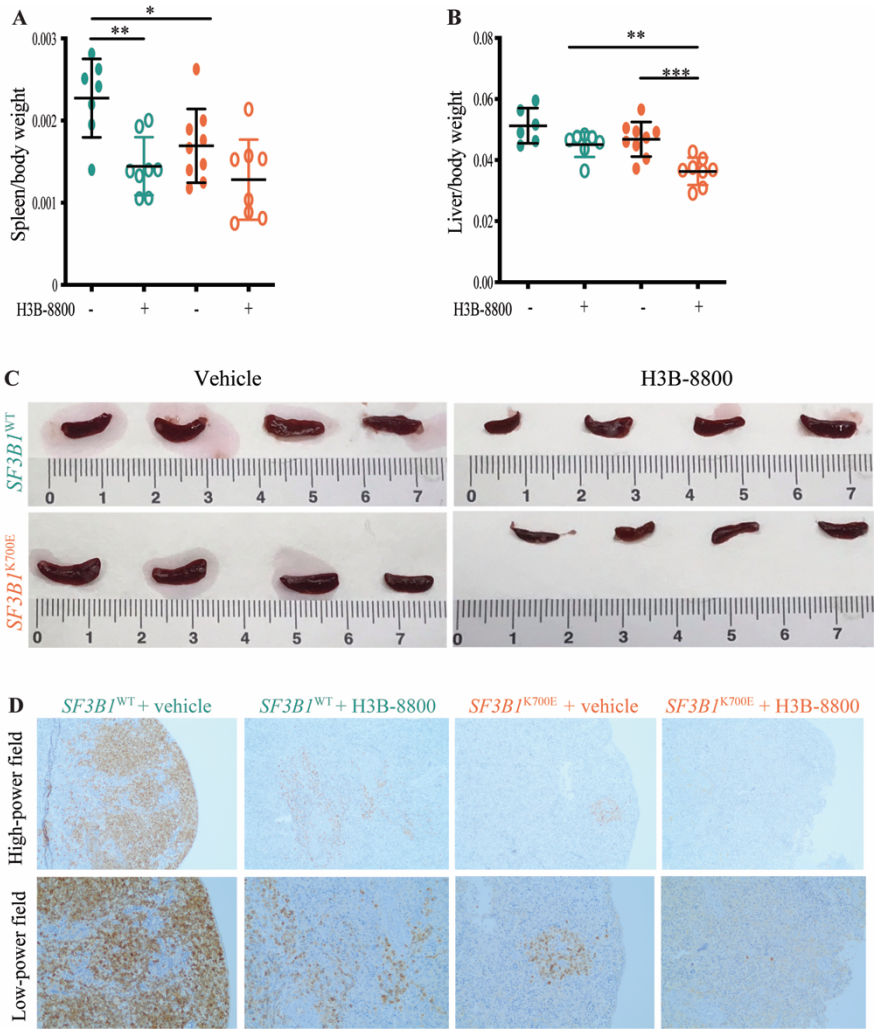


Figure II.7. H3B-8800 decreases leukemic infiltration in NSG mice's spleens. Normalized quantification of spleen (A) and liver (B) weights from mice with WT (green) or SF3B1 mutant (orange) xenografts treated with vehicle or H3B-8800 as indicated. Statistical significance was assessed using Mann-Whitney test (**: p value < 0.01; ***: p value < 0.001). (C) Representative spleens from NSG mice implanted with *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 cell lines after H3B-8800 or vehicle administration, as indicated. (D) CD79a immunohistochemical staining in spleen from H3B-8800-treated versus vehicle-treated mice implanted with *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 cell lines as indicated.

2. Transcriptomic analysis of H3B-8800-treated *SF3B1* ^{WT} and *SF3B1* ^{K700E} MEC1 CLL cell lines

2.1. Alternative splicing analysis

In order to identify the impact of H3B-8800 on alternative splicing and gene expression in *SF3B1* ^{K700E} CLL cells we treated 4 million *SF3B1* ^{WT} and *SF3B1* ^{K700E} MEC1 cells (in triplicate) with 75nM H3B-8800 or DMSO (control) for 6 hours before extracting the RNA and carrying out RNA deep sequencing using Illumina pair-end sequencing technology and analyzing the datasets using the toolset *vast-tools* as in Chapter I. The results obtained from the comparison between untreated *SF3B1*^{WT} and *SF3B1*^{K700E} groups have been discussed in Chapter I, sections 2.2 to 2.4.

We observed that H3B-8800 induced massive intron retention and exon skipping in both cell lines, with nearly 8000 introns detected as retained and more than 10000 exons detected as skipped when using a $|\Delta PSI| \geq 15$ and PSI range < 5 cut offs (Fig II.9A-C). Furthermore, a large overlap between the AS events regulated upon H3B-8800 treatment in *SF3B1*^{WT} and *SF3B1*^{K700E} cells was observed for both AS categories (Fig II.9B-C), indicating that *SF3B1* mutational status does not influence the response to H3B-8800 in any major way.

Intriguingly, a larger fraction of distinct effects of the drug on *SF3B1*^{WT} and *SF3B1*^{K700E} cells was observed when analyzing alternative 5'ss or 3'ss events (Fig II.9D-E).

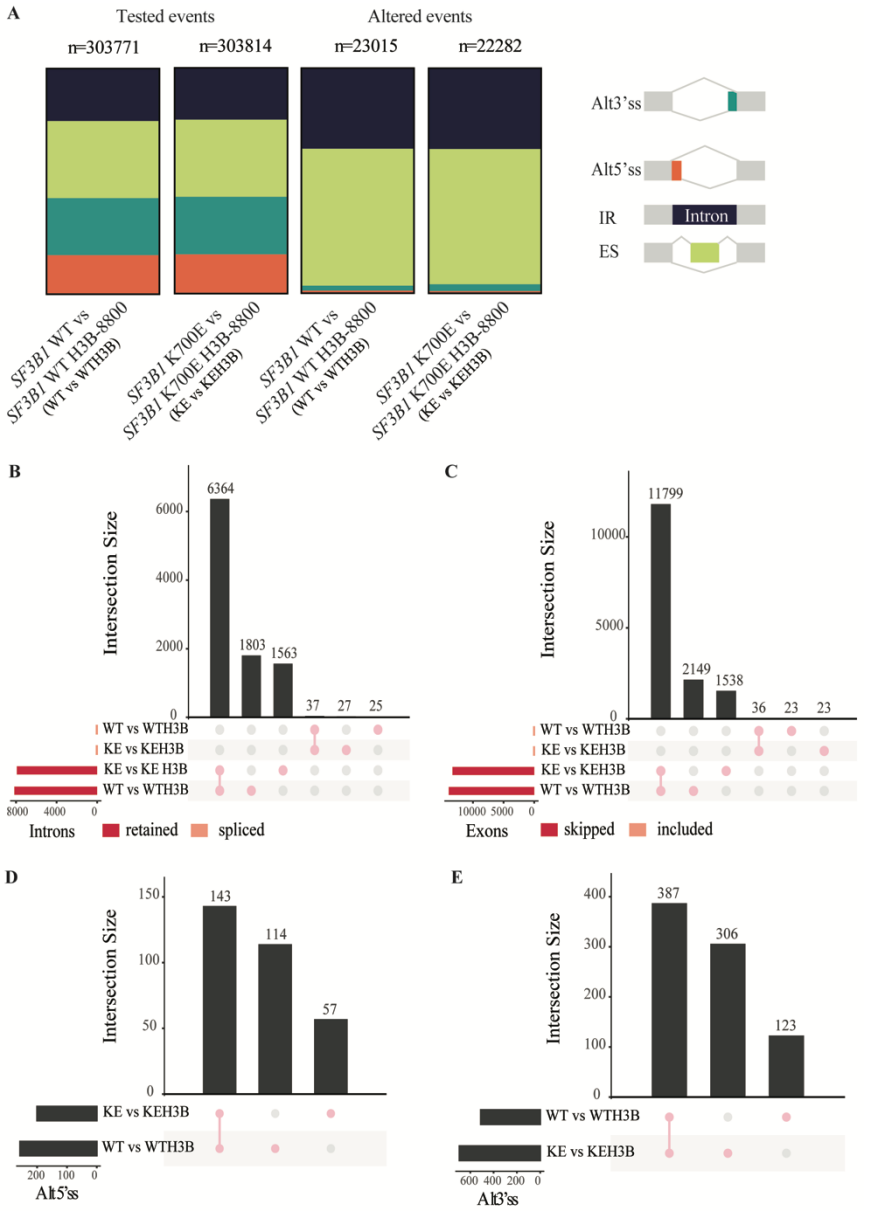


Figure II.9. H3B-8800 induces overlapping intron retention and exon skipping events in H3B-8800-treated *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 cell lines. (A) Relative frequency of differentially spliced (|ΔPSI| ≥ 15) AS event types in *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 cell lines upon H3B-8800 treatment compared to all assessed AS events using VAST-TOOLS for each comparison. (B-E) Upset plots displaying the numbers and overlaps between Intron Retention (B), Skipped Exons (C), Alt5'ss (D) and Alt3'ss (E) events misregulated upon H3B-8800 treatment in *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 cell lines for events mapped in all

conditions. WT vs WTH3B: *SF3B1*^{WT} vs *SF3B1*^{WT} treated with H3B-8800; KE vs KEH3B: *SF3B1*^{K700E} vs *SF3B1*^{K700E} treated with H3B-8800.

Next we focused on the analysis of splicing features of AS changes in intron retention and exon skipping in control versus H3B-8800-treated *SF3B1*^{WT} cells. We analyzed features using the Matt package[140](see Material and Methods), which allows the investigation of different parameters including the length of exons / introns, the strength of 5'ss / 3'ss and the comparison of such parameters between different groups.

For exons, we compared our group of regulated exons (|ΔPSI (control- H3B-8800 treated)| ≥ 15 and PSI range < 5 upon H3B-8800 treatment) to different control groups (constitutive exons: PSI > 95 in both groups (n=4922); alternative non-changing exons in at least one group: 10<PSI<90 and |ΔPSI| ≤ 5 (n=3288); and cryptic exons: PSI < 5 in both groups(n=4976)). Interestingly, we observed that skipped exons upon H3B-8800 treatment are shorter than the exons in all the control groups (constitutive, alternative and cryptic) (Fig II.10). Then, depending on the exon groups used as control, the features of the sequences surrounding the skipped exons vary. Compared to alternative non-changing and cryptic exons, skipped exons display stronger 5'ss and 3'ss according to the maximum entropy score Hsa (Homo sapiens) model (maximum entropy score of 3'ss using a model trained with human splice sites), and the surrounding introns and exons are shorter. On the other hand, compared to constitutive exons, skipped exons display weaker 5'ss and 3'ss according to the maximum entropy score Hsa model, and the surrounding introns are longer (Fig II. 10).

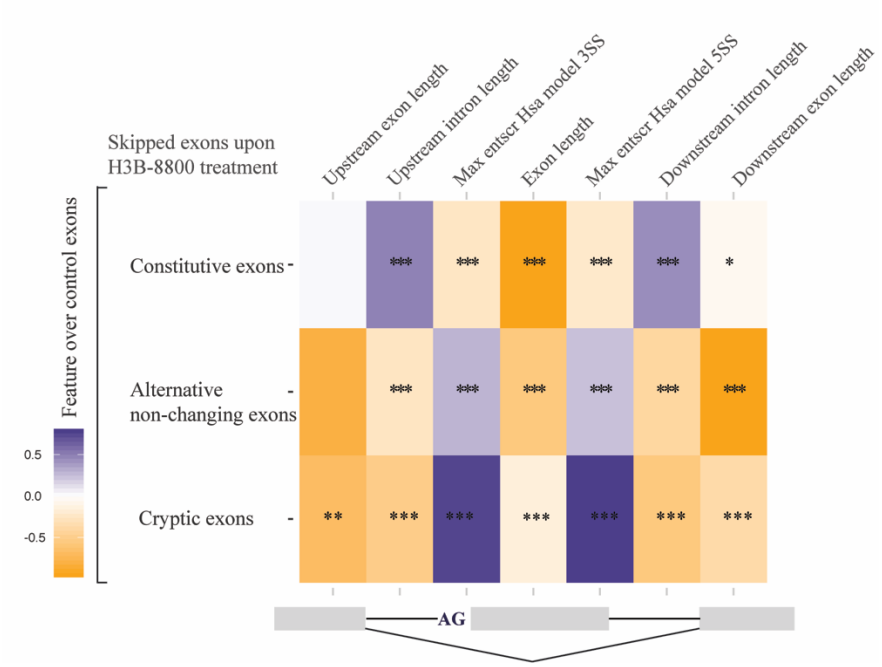


Figure II.10. Sequence feature analysis of exons skipped upon H3B-8800 treatment of *SF3B1*^{WT} MEC1 cells using Matt[140]. Significant features are indicated at the top of the figure and the degree and statistical significance level for the comparison with constitutive, alternative non-changing and cryptic exons indicated through the heat map and Mann-Whitney U test (*: $p<0.05$, **: $p<0.01$, ***: $p<0.001$). Exons skipped upon H3B-8800 treatment were defined as ($|\Delta PSI| \geq 15$ and PSI range < 5 upon H3B-8800 treatment. Constitutive exons were defined as displaying PSI > 95 in both groups; alternative non-changing exons as 10<PSI<90 in at least one group and $|\Delta PSI| \leq 5$ and cryptic exons as PSI < 5 in both groups.

In order to analyze the features of the retained introns upon H3B-8800 treatment (n=8171), we compared them with introns that were spliced more efficiently (n=62) and with introns that did not change (-1 <ΔPSI <1, 5 < PSI <95 in *SF3B1*^{WT} control group, n=675) under the same conditions. When comparing those three groups we observed that (i) retained introns and their flanking exons are shorter than the ones from non-changing introns (ii) their 5'ss and 3'ss are weaker (Fig II.11A) and (iii) display higher GC content (GCC), while

better spliced introns display lower GCC (Fig II.11A). An RNA map of GCC showed that retained introns are located in a more general GC-rich genomic context, as seen by the increased GCC in adjacent exons and surrounding introns as compared to control introns and better spliced ones (Fig II. 11B).

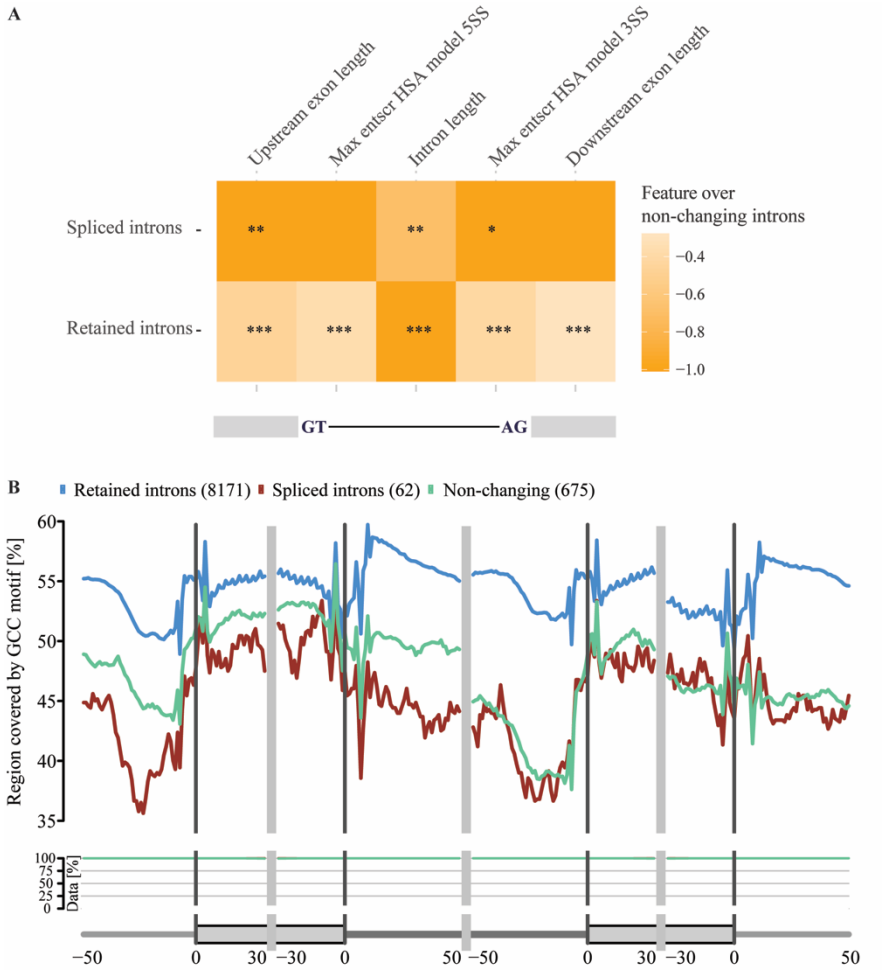


Figure II.11. Introns retained after H3B-8800 treatment are shorter and have higher GC content. (A) Sequence feature analysis of retained and more spliced introns upon H3B-8800 treatment of *SF3B1*^{WT} MEC1 cells using Matt[140]. Retained introns were defined as |ΔPSI (control – treated)| ≥ 15 and PSI range < 5, spliced introns as ΔPSI ≤ 15 and PSI range < 5 and non-changing introns as -1 <ΔPSI <1, 5 < PSI <95 in *SF3B1*^{WT} control group. Significant features are indicated at the top

of the figure and the degree and statistical significance level for the comparison of retained and more spliced introns with non-changing introns indicated through the heat map and Mann-Whitney U test (*: $p<0.05$, **: $p<0.01$, ***: $p<0.001$). (B) Distribution of GC content along the genomic exon/intron region (RNA map) for retained, more spliced and non-changing introns using Matt[140]. The upper plot shows the frequency of GCC motif; the middle plot shows the coverage of the region and the lower scheme, the number of nucleotides in exons and introns taken into account for the analysis.

2.2. Gene expression analyses

To analyze the gene expression changes induced after 6 hours of H3B-8800 treatment, we analyzed RNA-seq data from the experiment described before (II.5.1) using Deseq2. A heatmap of the normalized expression (divided by standard deviation) of the 1000 most variable genes across all samples performed with iDEP94[149] showed that the clustering was primarily driven by the treatment and secondarily by *SF3B1* mutational status (Fig. II.12).

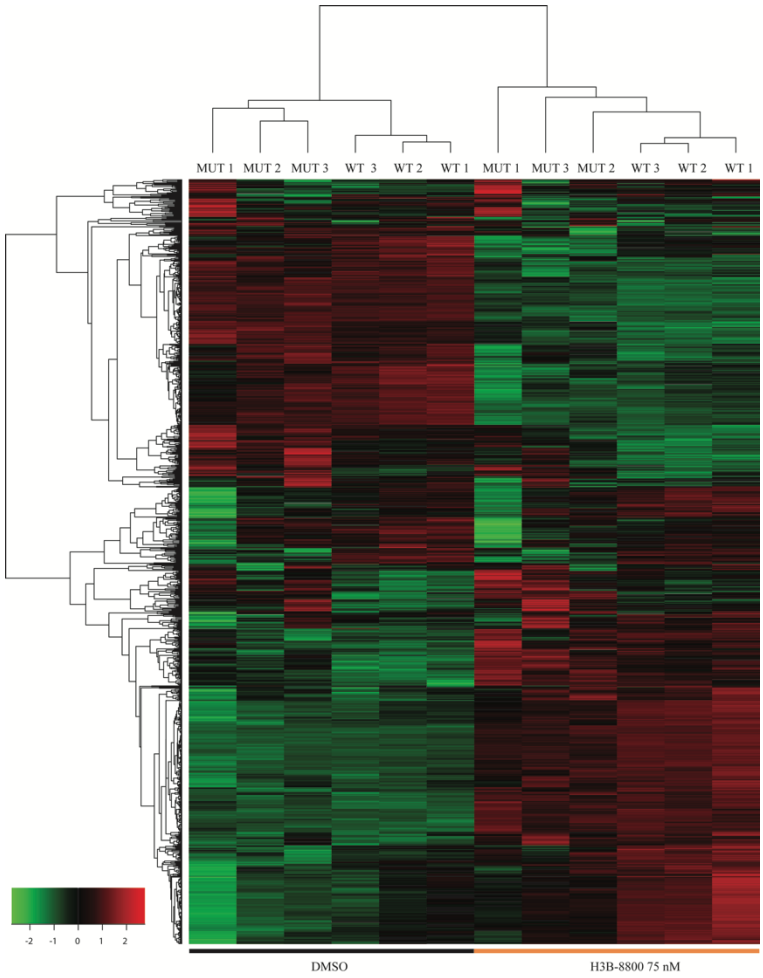


Figure II.12. H3B-8800 treatment induces similar gene expression changes in *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 cell lines. Heatmap of normalized expression values (divided by SD) for the 1000 most variable genes across the samples using iDEP94[149].

To study differentially expressed genes upon H3B-8800 treatment in *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 cells (differences between DMSO control and H3B-8800 treatment in each of the groups), we defined a cut off of |log2 Fold Change > 1| and an adjusted p value <0.05. We obtained 6719 dysregulated genes in the H3B-8800 treated *SF3B1*^{WT} group and 6077 genes in the H3B-8800 treated

SF3B1^{K700E} group (5590 genes were dysregulated in both datasets). Those sets of genes were used to identify enriched gene ontology terms using enrichGO R package. The most enriched categories were related to RNA metabolism and splicing in H3B-8800 treated *SF3B1*^{WT} and *SF3B1*^{K700E} groups. SF1, U2AF1 and U2AF2, SF3b complex components (PHF5A, SF3B1-SF3B6), several RBM proteins, SRSFs and hnRNPs were among the splicing factors and auxiliary proteins affected, urging caution regarding the interpretation of sequence features as direct effect of the drug on the function of SF3b alone (Supplementary Table 3).

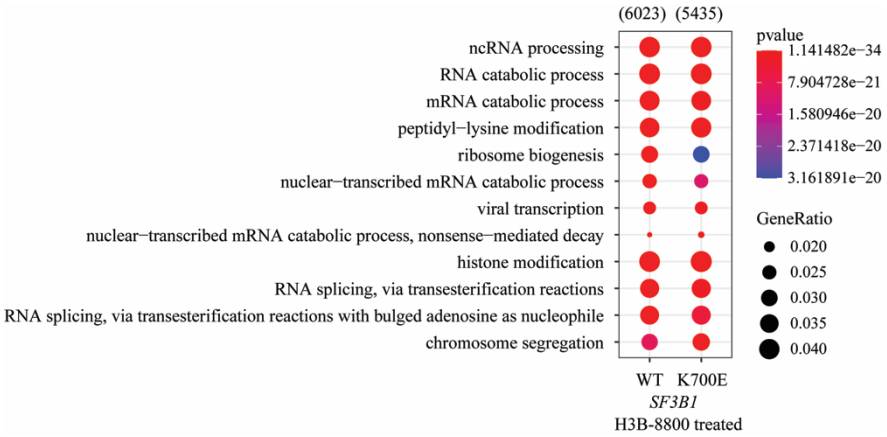


Figure II.12. Gene Ontology enrichment analysis of biological processes for all the genes differentially expressed (|log2 Fold Change| > 1 and p value <0.05) in *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1cell lines after treatment with 75nM H3B-8800 for 6 hours calculated using enrichGO. GO terms are indicated on the left, and the enrichment properties (pvalue and GeneRatio, the ratio of the affected genes in each pathway) indicated by the heat map color and size of the circles, respectively.

Interestingly, and potentially relevant to explain effects in viability reported in Chapter II section 1.1, a GO category related to apoptosis ("intrinsic apoptotic signaling pathway" (GO:0097193))

was enriched in H3B-8800 treated *SF3B1*^{WT} and *SF3B1*^{K700E} groups (with adjusted p value < 0.05 and q value < 0.05).

To look for an explanation for *SF3B1* mutated cells' higher sensitivity to H3B-8800 we studied the differentially expressed genes after H3B-8800 treatment between the MEC1 *SF3B1*^{WT} and *SF3B1*^{K700E} cells. We defined a cut off of |log2 Fold Change > 0.58| and an adjusted p value <0.05 and found 993 genes were differentially expressed. Those sets of genes were used to identify enriched gene ontology terms using enrichGO R package . The most enriched categories were related to T cell activation, cytokine production and leukocyte cell adhesion (Fig. II.13). Interestingly, these biological processes were also enriched when we compared differentially expressed genes in MEC1 *SF3B1*^{WT} and *SF3B1*^{K700E} cells without H3B-8800 treatment (Fig. I.18 in chapter I), revealing the gene expression differences between MEC1 *SF3B1*^{WT} and *SF3B1*^{K700E} cells are maintained despite H3B-8800 treatment.

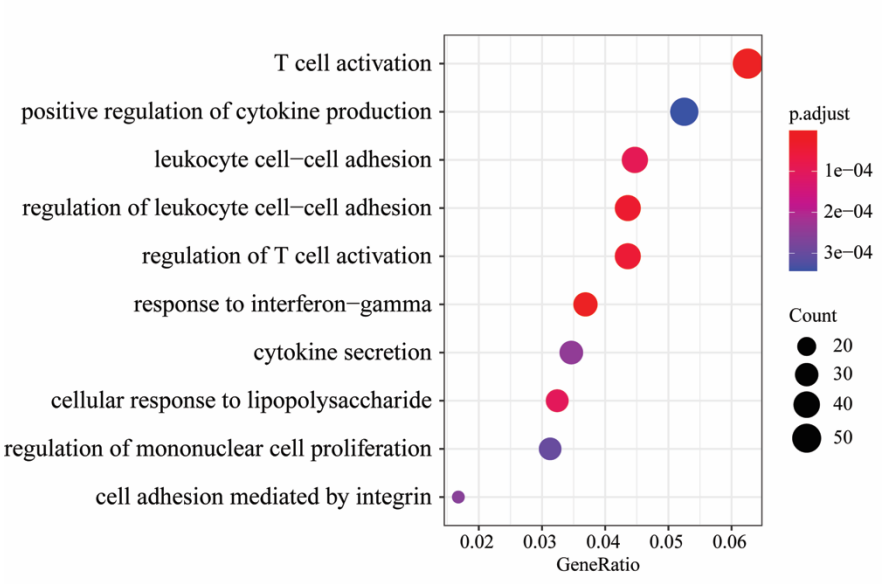


Figure II.13. Gene Ontology enrichment analysis of biological processes for all the genes differentially expressed between (|log2 Fold Change| > 0.58 and adjusted p value <0.05) in *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1cell lines after treatment with 75nM H3B-8800 for 6 hours calculated using enrichGO. GO terms are indicated on the left, and the enrichment properties (pvalue and the gene counts used to calculate the Gene Ratio (the ratio of the affected genes in each pathway)) indicated by the heat map color and size of the circles, respectively.

3. H3B-8800 modulation of alternative splicing events from genes related to apoptosis

3.1. Validation of MCL-1 AS event

Our RNA-seq analysis identified the well-described cassette exon 2 in MCL1gene[159] (HsaEX0038237) as changing in H3B-8800 treated *SF3B1*^{WT} and *SF3B1*^{K700E} ΔPSI=-20.3 and ΔPSI=-25.7 (ΔPSI= PSI upon H3B-8800 treatment - PSI DMSO condition), respectively. Exon 2 skipping leads to an mRNA that encodes a proapoptotic isoform while exon inclusion generates an mRNA encoding an antiapoptotic isoform (Fig II.14A). This AS event is known to be modulated by *SF3B1*-binding splicing modulators such as sudemycins or E7107[71,160]. We also observed increased skipping of MCL1 exon 2 upon treatment with 75nM H3B-8800 in both patient primary cells (Fig II.14B) and in MEC1 cell lines (Fig II.14C-D). However we have not yet confirmed increased proportion of the pro-apoptotic form at the protein level nor investigated the involvement of this AS change on the viability changes observed in CLL samples and MEC1 cell lines upon H3B-8800 treatment (Fig. II.2).

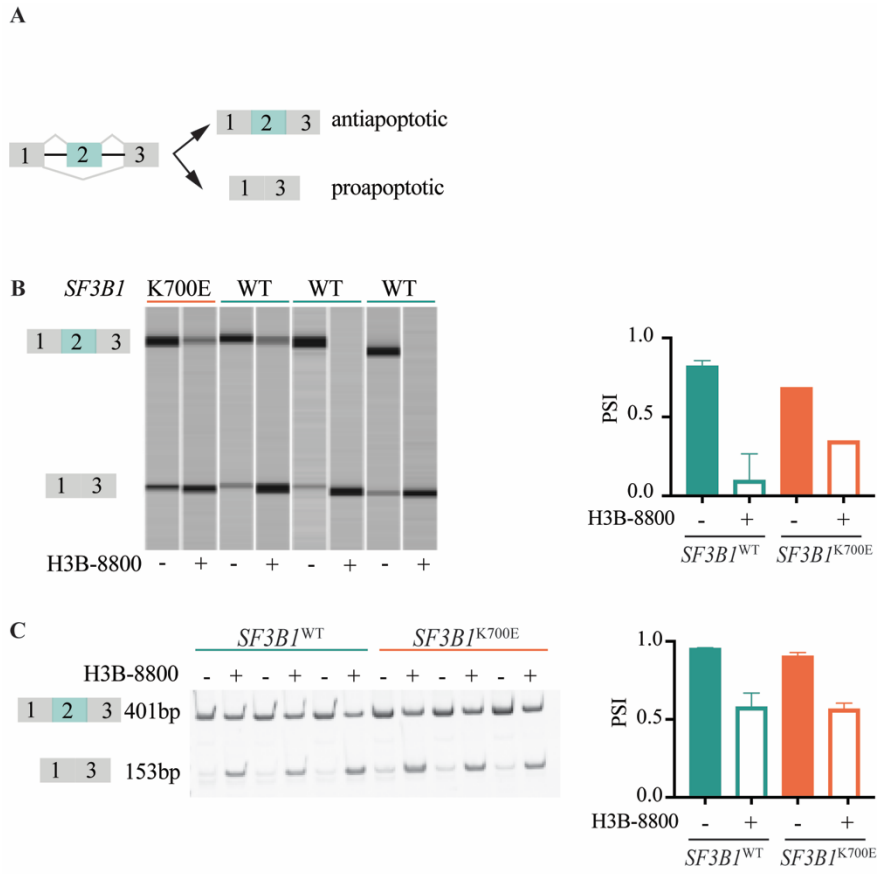


Figure II.14. H3B-8800 induces the proapoptotic isoform of MCL1. (A) Schematic representation of the AS event involving MCL-1 exon 2. (B) Validation of MCL1 exon 2 skipping in 4 patient samples after 3h treatment with 75nM H3B-8800. RT-PCR amplification products corresponding to WT and K700E SF3B1 patient samples treated with vehicle or H3B-8800, using primers complementary to exons 1 and 3, were analyzed by capillary electrophoresis. The panel on the right shows mean and standard deviation quantification of exon 2 PSI values under the different conditions. (C) Validation of MCL1 exon 2 skipping in WT and SF3B1 K700E isogenic MEC1 cell lines after 6h of treatment with 75nM H3B-8800, analyzed by RT-PCR as in (B) and polyacrylamide gel electrophoresis. The panel on the right shows mean and SD quantification of exon 2 PSI values under the different conditions.

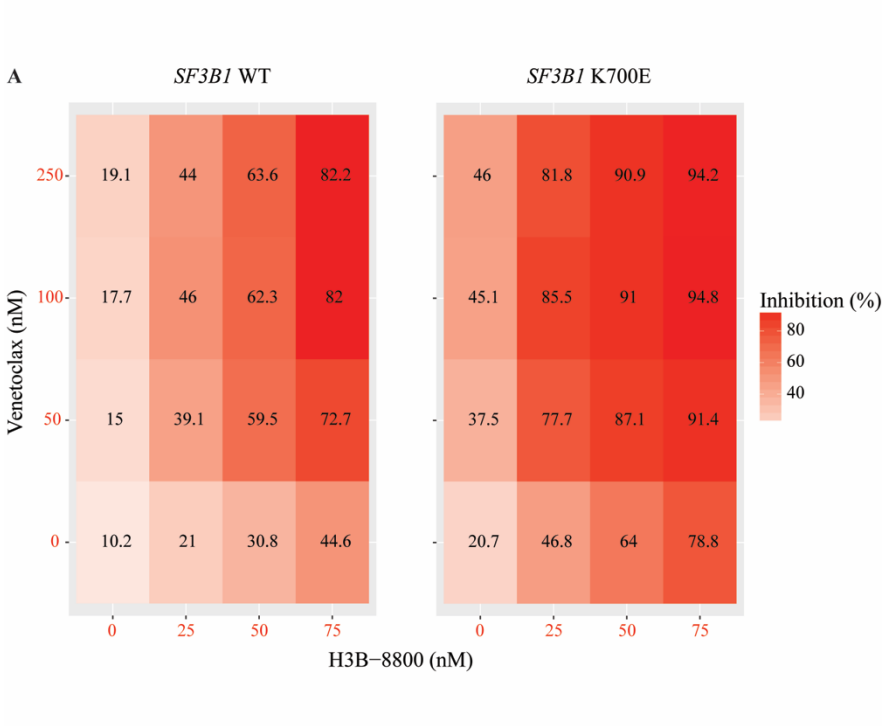
3.2. H3B-8800 shows synergistic effects with venetoclax

Venetoclax (ABT-199), a selective inhibitor of the antiapoptotic protein BCL2, is used in first line therapy for high-risk CLL patients and in patients with relapsed or refractory CLL[83].

It was previously reported that BCL2 inhibitors display a synergistic effect with E7107 or herboxidiene splicing modulators[160][161]. In fact, E7107 sensitizes venetoclax-resistant CLL cells by decreasing their MCL1 dependence and increasing their BCL2 dependence[161]. In this context, we postulated that the combination of venetoclax with H3B-8800 splicing modulator might enhance the cytotoxic effect of each of these drugs when used individually.

To test this hypothesis, we treated *SF3B1*-WT and -mutated MEC1 cell lines with venetoclax (0-250nM) and H3B-8800 (0-75nM) for 48 hours and assessed cytotoxicity by Annexin V and propidium iodide positivity by flow cytometry (Fig. II.15A and C). This data was used to calculate a ZIP synergy score using SynergyFinder[162] (Fig. II.15B). The synergy score is the difference between the observed effect and the expected effect of a drug. If the score is positive, the drug combination could be classified as synergistic; if it is negative, as antagonistic. The synergy score can be interpreted as the proportion of cellular responses that can be attributed to the drug interactions[162]. In the case of MEC1 cell lines, we obtained ZIP synergy scores around 30 for *SF3B1*^{WT} and *SF3B1*^{K700E} cells; i.e. 30% of the cytotoxicity observed in the combination therapy with H3B-8800 and venetoclax could be attributed to the interaction between the effects of these drugs. The

concentrations with highest synergy scores were 100-250nM venetoclax with 75nM H3B-8800 in the MEC1 *SF3B1*^{WT} cells (Fig. II.15C) and 100-250nM venetoclax with 25nM H3B-8800 in the *SF3B1*^{K700E} cells. Altogether, these results demonstrated that the combination of H3B-8800 with venetoclax resulted in enhanced cytotoxic effects.



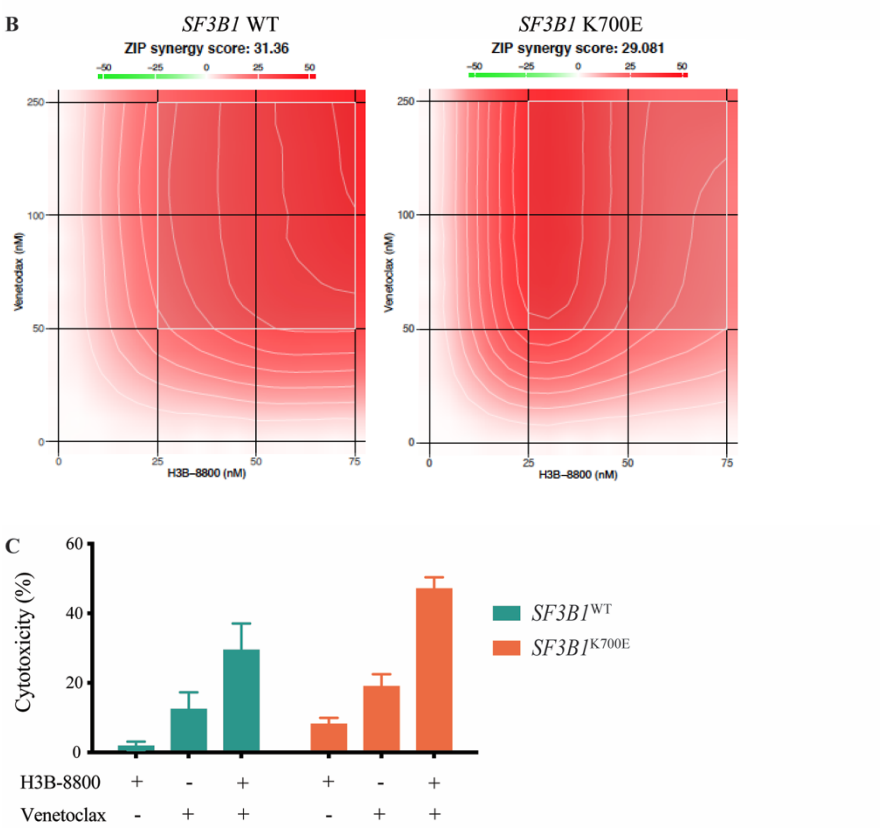


Figure II.15. H3B-8800 displays synergistic effects with venetoclax in MEC1 CLL cell lines. (A) Dose-response inhibition matrices for the indicated doses of H3B-8800 and venetoclax combination therapy at 48h in *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 cell lines. Inhibition values correspond to the cytotoxicity assessed by annexin V and propidium iodide positivity by flow cytometry. (B) ZIP synergy score representation based on the dose-response inhibition matrices from panel (A) for the indicated H3B-8800 and venetoclax combination doses in *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 cell lines. Synergy scores were calculated using SynergyFinder[162]. (C) Cytotoxicity with 100nM venetoclax and/or 75nM H3B-8800 synergistic doses in *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1 cell lines. Cytotoxicity was assessed by annexin V and propidium iodide positivity by flow cytometry.

DISCUSSION

Characterization of Transcriptomic Changes Associated with Mutations in the Splicing Factor *SF3B1* in Chronic Lymphocytic Leukemia

Genome-wide studies highlight the genomic complexity of CLL[98,99], a disease characterized by a relatively high number of cancer driver genes, each of them mutated at relatively low-frequency. *SF3B1* is one of the most recurrently mutated genes in CLL and its mutational status correlates with worse disease prognosis[106,124,125], which makes its study particularly interesting in order to understand pathogenic mechanisms and develop future therapies. In this work, we mainly focused on deciphering the effects of *SF3B1* mutation on alternative splicing regulation in CLL, and explored the molecular and cellular effects of a SF3B1 inhibitor, H3B-8800, of potential therapeutic use.

For the study of the role of *SF3B1* mutations in AS, we analyzed the RNA-seq data of one of the largest CLL patient cohorts. Working with patient samples is inherently difficult due to interpatient heterogeneity. Challenges we had to face include: first, CLL patients can harbor more than one driver alteration[99,106]; second, *SF3B1* mutations are associated with specific biological features of the disease, such as the U-CLL IGHV rearrangement status, the intermediate epigenetic subgroup or IGHV3-21^{R110} mutation[90,91,94]. These features make the comparison between the *SF3B1* WT and mutant samples challenging not only because each group is heterogeneous, but also because the two groups are not

equally homogeneous. To partially overcome these difficulties and work with a homogeneous model, we have established the first CLL isogenic cell lines harboring or not a *SF3B1* mutation. Although other *SF3B1* wild type and mutant isogenic cell lines had been previously developed, such as Nalm-6 B cell precursor leukemia cell line, murine embryonic stem cells or induced pluripotent stem cells [45,163,164], these were not available for CLL. Our work therefore provides a *SF3B1* mutated CLL model which can be easily handled in the lab and can be used for further research on the molecular and cellular consequences of SF3B1 mutations. In addition, the gene editing strategy that we have developed may help to engineer additional CLL or other cancer cell lines as well as additional *SF3B1* mutations.

Diving into the heterogeneity of SF3B1 mutations

PCA analysis of samples from patients bearing or not mutations in *SF3B1* showed specific clustering of mutations and alternative 3'ss usage, which depended not only on the mutational status of *SF3B1* but also on its subclonality. The samples with the lowest CCF clustered next to *SF3B1* WT samples, while samples with higher mutational burden are characterized by their patterns of alternative 3'ss usage, suggesting that this molecular phenotype is relevant for cellular phenotypes leading to clonal selection and possibly disease progression (see below). In addition, *SF3B1* E862K and M757T showed an alternative 3'ss profile that was similar to the SF3B1 WT samples, suggesting that these mutations do not have an impact in alternative 3'ss activation.

The subclonality of *SF3B1* mutations is known to have a direct link with patients' prognosis. Indeed, *SF3B1* clonal mutations with a variant allele frequency (VAF) > 12% are predictive of shorter time to first treatment (TTFT), while lower percentages do not[105]. However, in the recent analysis of the COMPLEMENT I clinical trial where chlorambucil or chlorambucil and ofatumumab treatments were compared, patients affected by low VAF clones showed shorter Progression-Free Survival (PFS)[165]. These discrepancies highlight the need to take into account and reporting the VAF when analysing the results of clinical trials in the era of new targeted therapies.

The study of the impact that different *SF3B1* mutations have in the molecular phenotype has shown that rare or novel (not hotspot) mutations in *SF3B1* can have similar effects to those of hotspot mutations[147,166]. Similarly, many rare mutations in *U2AF1* and *SRSF2* also completely or partially phenocopy the effects of hotspot mutations in these genes[167]. Interestingly, this opens the possibility of using common surrogate markers for mutations in each of these splicing factors, allowing not only the identification of functional mutations but also of functional VAFs. With this aim in mind, we looked for an AS isoform that would be specific for *SF3B1* mutations and we identified a cryptic isoform in the gene ZDHHC16 which is detected in *SF3B1* mutations with a VAF>13% and not in *SF3B1* WT cases nor in cases with E862K or M757T mutations, which as alluded to earlier, cluster together with *SF3B1* WT samples in terms of the patterns of 3'ss usage. Therefore, we propose that the presence of the ZDHHC16 cryptic isoform could be used as a surrogate marker of *SF3B1* mutations and VAFs with an impact in

Alt3'ss selection. Interestingly, in a phase I clinical trial testing the SF3B1-targeting splicing inhibitor H3B-8800 (see below) in AML, CMML and MDS (NCT02841540), aberrant expression of an AS isoform in the gene *TMEM14C* identified a group of patients who particularly benefited from this therapy, evidencing the applicability of this type of approach[75].

However, it is important to point out that mutations such as *SF3B1* E862K or M757T, which do not reproduce the AS alterations of the hotspot mutations, might remain to be pathogenic through other mechanisms and therefore cannot be dismissed as harmless mutations. Along these lines, *SF3B1* is known to have other functions in addition to its role as a component of U2snRNP. For instance, it contributes to sister chromatid cohesion and mitotic chromosome segregation[168] and it also binds to nucleosomes[169] so there could be other mechanism by which *SF3B1* mutations contribute to disease progression.

Transcriptomic changes associated with SF3B1 mutations

Since the first reports of *SF3B1* mutations and aberrant splicing in cancer[41,104,128,170], important efforts have been made to characterize their alternative splicing profiles. The usage of cryptic 3'ss selection has been reported in previous publications highlighting the activation of cryptic 3'ss typically located ≈15 to 24 nucleotides upstream of the canonical 3'ss[45,46,48,129,138], with additional less frequent cases of Alt'3ss located at longer distances[46] or even closely downstream of the canonical 3'ss[45,48,129] Our analyses

confirm these observations and in addition explore intronic features that can contribute to mediate the effects of SF3B1 mutations.

We classified the Alt3'ss in 4 groups depending on the location (upstream/downstream) and the distance (>/<50 nucleotides) to the canonical 3'ss. Interestingly we observed, both in CLL primary cells and CLL cell lines, that cryptic 3'ss can be located not only upstream but also downstream of the canonical 3'ss and that the distance between the two 3'ss, both in upstream and downstream cases, can be longer than 50 nucleotides and therefore is not as constrained as previously thought.

Regarding the sequence features involved, we observed that Alt3'ss and cryptic 3'ss typically have weaker BP, Py tract and 3'ss than the canonical 3'ss, consistent with previous reports[48]. These properties are common regardless of Alt3'ss or cryptic 3'ss location. Moreover, cryptic 3'ss features are even weaker than Alt'3ss features, suggesting that very weak 3'ss cannot be recognized by WT SF3B1. We also confirm weaker polypyrimidine tracts in the Alt3'ss compared to the canonical 3'ss, as reported by Darman et al [45]. Regarding the BP usage, mutation of predicted BPs for Alt3'ss and canonical 3'ss in minigene assays provided results consistent with the use of different BP by the Alt3'ss and the canonical 3'ss[45,48,171,172]. However, our results are also consistent with more elaborate scenarios, e.g. involving a complex interplay between various BPs and U2 snRNP particles harboring WT or mutant SF3B1 present in the same cell. For example, in the case of ZNF561 downstream cryptic 3'ss, our results are consistent with a model in which WT U2 snRNP acts through a distal BP while K700E U2

snRNP activates the cryptic downstream 3'ss through a downstream BP, thus expanding the range of mechanisms involved. In the case of MAP3K7, in contrast, our results suggest that the same BP is used by WT and K700E U2 snRNP with different consequences for the relative use of the canonical and the upstream Alt3'ss. In this sense, we have set up a collaboration with the group of Jonathan Staley (University of Chicago) to directly determine the location of functional BPs by sequencing of lariat intermediates. This group has developed protocols for co-transcriptional lariat sequencing that analyse nascent intron lariats and enables genome-wide BP annotation with unprecedented efficiency[173]. In addition, they have been able to identify genetic elements such as the polypyrimidine tract, intron length, and regional GC content that predict the timing of splicing relative to transcription and the order or intron removal. Such a powerful tool would be particularly interesting in our study, as it may allow the identification of the BP used under normal conditions, upon drug treatment or upon *SF3B1* mutation. The data obtained will be confronted with the predictions using various BP predictor tools such as SVM and used to fine tune our feature analysis and evaluate our mechanistic hypotheses.

Additionally, it is important to mention that nanopore sequencing, which yields reads as long as 2 megabases[174], can be very useful to broaden our knowledge of the effects of *SF3B1* mutations on AS. In fact, in a recent study carried out in normal B cell and CLL samples, the authors report strong down-regulation of intron retention events (i.e. increased efficiency of intron removal) in *SF3B1* mutated CLL samples[175]. In this case, *SF3B1*^{WT} samples

present an increase in the intron retention events that is not seen in normal B cells nor in *SF3B1* mutated CLL samples. In addition, these studies also reveal additional splicing complexity, e.g. identifying multiple co-occurring events at long distances of the Alt3'ss usage associated with *SF3B1* mutation.

The impact of AS events induced by SF3B1 mutations on CLL pathogenesis

In recent years, RNA-seq analysis of CLL samples have provided a wealth of information about changes in AS events associated with *SF3B1* mutations. Many efforts have been made to elucidate the role of certain AS events in CLL pathogenesis. We have focused our work in AS of the MAP3K7 gene. MAP3K7 is a kinase that mediates tumor necrosis factor α, interleukin-1β, and Toll-like receptor signaling through the NF-κB, JNK, and MAPK pathways. It has been reported that *SF3B1* mutation induces cryptic splicing of MAP3K7, resulting in hyperactivation of NF-κB pathway [32], a pathway that is activated in CLL cells in the lymph node[176]. We observed that MAP3K7 protein expression was decreased in the *SF3B1*^{K700E} lines compared to *SF3B1*^{WT} lines and this decrease correlated with an increase of NF-κB activation in *SF3B1*^{K700E} cases. Also, downregulation of p38 phosphorylation was observed, as previously reported in MDS, related to aberrant MAP3K7 AS and NMD[57]. The fact that p38 MAPK can act as a tumor suppressor[177] and its phosphorylation is decreased upon *SF3B1* mutation in CLL cells, make this an interesting pathway for further research, as it could shed light on the reasons why *SF3B1* mutation confers selective advantages to CLL

cells. Even though we did not perform rescue experiments to validate the relationship between MAP3K7 aberrant splicing and its downstream effects in NF-κB pathway and p38 phosphorylation, these links are well characterized in MDS and breast cancer[32,57,178]. These results point out commonalities in Alt3'ss activation upon *SF3B1* mutation in different tumor types.

Other AS events that may confer advantages to *SF3B1* mutated CLL cells have been described. For example, an Alt3'ss in DVL2, a gene that negatively regulates Notch pathway[179], has been linked to Notch1 pathway activation in *SF3B1* mutated CLL[129,180]. The Notch pathway is involved in B cell development[181] and is activated in CLL patients with *NOTCH1* mutations[182], conferring them specific clinical and biological features as well as a poor prognosis[125,183]. However, DVL2 AS may not be the only factor contributing to NOTCH1 activation because genome-wide DNA methylation studies have shown that *SF3B1* mutated patients display decreased methylation levels in 67 genomic regions, mainly nearby telomeric regions, which are enriched in gene bodies of cancer related genes such as NOTCH1[146].

Another interesting example is related to activation of an Alt3'ss in the gene encoding PPP2R5A, a protein phosphatase 2A (PP2A) subunit. NMD of this isoform promotes MYC stability and impairs apoptosis by increasing MYC S^{62} and BCL2 S^{70} phosphorylation in *SF3B1* mutated tumors, including CLL[178].

One of the difficulties we faced at the time of selecting an aberrant splicing isoform with a potential biological impact in CLL for further

study was selecting among a long list of appealing candidates. The use of systematic approaches such as CRISPR screening may be instrumental to identify key events that mediate the impact of the mutation. A precedent is the work by Inoue et al[166], where they designed a single-guide RNA (sgRNA) library that covered specific targets of mutant *SF3B1* predicted to undergo NMD and using this approach they identified the inclusion of a poison exon in BRD9, a component of the BAX chromatin remodeling complex, whose miss-splicing correction by antisense oligonucleotides or CRISPR-directed mutagenesis suppressed tumor growth in uveal melanoma[166]. This AS isoform in BRD9 was also found in CLL[166]. Even though this approach was beyond the scope of this Thesis, we are collaborating with Miquel Anglada from the lab of Luis Serrano at Systems Biology Program at the Centre of Genomic Regulation to apply a systematic method with our dataset. Based on the dependency on specific genes for cancer cell survival, predicted from large-scale RNAi screens in hundreds of cancer cell lines[184], he is developing a method to compute the contribution of a particular AS event to the gene dependency, taking into account AS events and GE data. Ideally, this approach could help us to compare the essentiality of a gene in *SF3B1* WT and mutated conditions and to identify AS events that are essential for the survival of CLL cells.

The SF3B1-Binding Splicing Inhibitor H3B-8800 in the Treatment of Chronic Lymphocytic Leukemia

SF3B binding molecules may have a place in CLL therapy

CLL therapy has experienced important transformations in the last decade. Since the BTK inhibitor ibrutinib demonstrated its efficacy on refractory high-risk CLL patients in 2013[185], targeted therapy has been gradually incorporated in the clinical practice and is now part of the front-line therapy for CLL patients. The targeted drugs currently approved by both the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA) are the Bruton's tyrosine kinase (BTK) inhibitors ibrutinib and acalabrutinib, the BCL-2 inhibitor venetoclax and the phosphatidylinositol 3-kinase (PI3K) inhibitors idelalisib and duvelisib. While this has resulted in significant improvement of the patients' prognosis, CLL remains to be considered an incurable disease. For instance, ibrutinib monotherapy mainly induces partial remissions, which are usually preserved with long-term therapy, but in turn can be compromised by tolerability issues and the development of resistance[86]. In the case of venetoclax, the antagonist of the anti-apoptotic BCL2 protein, it demonstrated the induction of important responses with negative minimal residual disease from Phase 1 trials, and improved results in combination with anti-CD20 antibodies[120,186,187]. However, the emergence of resistance to venetoclax therapy, including BCL2 mutations or TP53 mutations or deletions [188–191], is leading to the search for new strategies[192]. To overcome these therapeutic

limitations, current efforts are focused on finding a combinatory therapy that induces strong responses within a fixed period, i.e. not requiring long-term treatments that may pave the way for the emergence of resistance.

In this new context, there is limited data about the responses of patients carrying recurrent mutations to the newly approved target therapies. In patients with *SF3B1* mutations, (i) they show a shorter progression free survival with chlorambucil (+ ofatumumab) as well as with obinutuzumab + chlorambucil combinations[165,193]; and (ii) they present shorter duration of response to venetoclax [194]. Regarding therapy with BTK inhibitors, *SF3B1* mutations are enriched in ibrutinib treated samples[195] and in the patients developing a Richter transformation after ibrutinib therapy, *SF3B1* mutations are more frequent than mutations in BTK[196]. In this scenario where *SF3B1* mutations are linked to an aggressive phenotype even when new targeted therapies are used, there could be a place for *SF3B1*-targeting therapies in CLL.

SF3B1 binding splicing modulators have already shown promising effects in CLL. Spliceostatin A, sudemycin and pladienolide B and its derivative FD-895 cause more cytotoxicity on primary CLL samples than in healthy B lymphocytes, inducing apoptosis at least in part through the switch of MCL1 splicing toward its proapoptotic isoform [71,197,198]. In relation to conditions mimicking the CLL microenvironment, in co-culture conditions with stroma-NK-tert cells for example, pladienolide B and FD-895 overcome the stromal cell support; however in the presence of CD40L plus IL4, which mimic T-cell help, spliceostatin A-

associated apoptosis is reduced[197,198]. In the case of H3B-8800, we show that it also causes more cytotoxicity in primary CLL cells than in healthy B lymphocytes and the effect was maintained in coculture conditions with HK and HS5 cells, suggesting that the compound has sufficient efficacy to overcome the effect of the microenvironment. Furthermore, H3B-8800 also modulates MCL-1 splicing towards the generation of the pro-apoptotic isoform.

Regarding *in vivo* mice models, in immunocompromised NSG mice engrafted with primary cells from CLL patients, sudemycin D6 decreases CLL cell counts in peripheral blood and spleen[71]. Even if it is not strictly a CLL model, in BALB/c mice bearing subcutaneous A20 lymphoma tumors, pladienolide B causes tumor regression and improves survival[198].

Our data further supports the observation that drugs targeting the spliceosome and SF3B can impact on cancer development in mouse models. Indeed, we showed that H3B-8800 in the context of CLL also reduces leukemic infiltration in a NSG mouse model engrafted with SF3B1$^{WT/K700E}$ MEC1 cell lines, and decreases CLL cell counts in peripheral blood and bone marrow.

In addition, combined therapies involving SF3B1-binding splicing modulators and drugs used for CLL treatment have been tested. For instance, the combination of sudemycin D1 with ibrutinib produces enhanced *in vitro* cytotoxicity involving IBTK's (BTK physiological inhibitor) splicing modulation[71]. Also, combinations of SF3B1 binding compounds and BCL-2 inhibitors have been further explored. Combinations of spliceostatin A and venetoclax or

spliceostatin A and the BCL2 inhibitor ABT-263 overcome the pro-survival effect of the microenvironment (CD40L and IL4) that could not be overcome otherwise in monotherapy[197]. Moreover, Ten Hacken et al[161] showed that splicing modulation with the pladienolide B derivative E7107 sensitizes CLL cells to venetoclax therapy by reducing their dependency of MCL1 and increasing the dependence on BCL2. Furthermore, E7107 overcomes venetoclax resistance *in vivo* in the Eμ-TCL1–based adoptive transfer murine model of CLL and increases mice overall survival alone or in combination with venetoclax (Fig D1). Our preliminary data showed that H3B-8800 modulates MCL1 splicing and that H3B-8800 has a synergistic effect with venetoclax. A model has been proposed whereby SF3B-targeting splicing modulators preferentially perturb RNA splicing of MCL1 and BCL2A1, but not BCL2L1 (BCLxL), leading to selective cytotoxicity in MCL1- or BCL2A1-dependent cancer cells. In combination with BCLxL/BCL2 inhibitors, splicing modulators can enhance the cytotoxicity through broader inhibition of the BCL2 family genes that act cooperatively in antiapoptosis/pro-survival in cancer cells [160]. Further studies are needed to investigate whether H3B-8800 acts through the same mechanism as E7107 to sensitize cells to venetoclax, i.e. by reducing MCL1 and increasing BCL2 dependence. For that purpose, it will be necessary to examine MCL1 and BCL2 protein expression upon H3B-8800 and venetoclax monotherapy, as well as in combination.

SF3b binding molecules induce preferential lethality in cells bearing mutated splicing factors

H3B-8800 was designed to exert preferential cytotoxicity in splicing factor-mutant cells. It showed stringer effect in NSG mice models harboring *SF3B1*^{K700E} leukemic cells and *SRSF2*^{P95H} CMML cells compared to their WT counterparts[63]. The *SF3B1*^{K700E} MEC1 cell lines we engineered are more sensitive to H3B-8800 than *SF3B1*^{WT} cells at concentrations of 10nM and higher. CLL primary samples with diverse mutations in the *SF3B1* gene and high CCF (>80%) also show a tendency towards higher sensitivity, although it does not reach statistical significance. However, this difference is not observed in patient samples with lower percentages of *SF3B1* mutations. In addition, in NSG mice bearing *SF3B1*^{K700E} MEC1 isogenic cell lines, the H3B-8800-treated *SF3B1*^{K700E} group shows the lowest CLL infiltration, as measured by bioluminescence and also displays the lowest bone marrow infiltration and spleen infiltration, arguing that H3B-8800 also has a preferential lethality in *SF3B*-mutant CLL cells.

These findings would align with the idea that cancer cells with heterozygous mutations in SF are dependent on their WT allele for splicing function. Several lines of evidence support this idea.

First, SF mutations appear in a mutually exclusive manner in myeloid malignancies[31]. In CLL, mutations in U1 snRNP that are present in 10% of the patients and *SF3B1* mutations also seem to be mutually exclusive[102]. This might indicate that the splicing burden of splicing factor-mutant cells is unsteady and accumulating two

perturbations makes the cancer cells non-viable. Along these lines, the simultaneous expression of *SF3B1* and *SRSF2* mutations cause synthetic lethality[32] and single-cell analysis has revealed that in the rare cases in which mutations in splicing factors appear concomitantly, they do not display hotspot mutations and display more limited effects on RNA splicing or RNA binding[33]. In addition, different studies show that in the presence of *SF3B1*, *SRSF2* or *U2AF1* mutations, the WT allele is required for cell survival[199–201]. Taking this into consideration, we could think that as CLL primary cells with lower percentages of *SF3B1* mutation have higher proportion of the WT allele, this makes them less sensitive to H3B-8800 treatment than the samples with high CCF. Rather than the presence or absence of a mutation, the allelic load may be the key factor for the cell survival.

Second, I would like to mention that H3B-8800 is not the only drug showing a preferential effect towards splicing factor-mutant cancer cells, although it is perhaps the best characterized one. For instance, sudemycin D6 has demonstrated a more cytotoxic effect in *U2AF1*^{S34F} cell lines compared to WT cells[202] and *SRSF2*-mutated myeloid leukemia models and *SF3B1*-mutated hematopoietic stem and progenitor cells are more sensitive to E7107 treatment[200,203]. In CLL cells, sudemycin D1 and D6 were also more cytotoxic in cases bearing *SF3B1* mutation or other RNA splicing related genes[71].

H3B-8800 induced transcriptomic alterations

Different SF3b binding molecules have some common and differential effects on splicing regulation[204]. In general, they cause the retention of short introns with high GC content and the skipping of short exons harboring weak splice sites, as it has been observed with spliceostatin A, sudemycin, E7107 and H3B-8800 among others[63,200,202,204]. Along these lines, our results with H3B-8800-treated MEC1 cells go in the same direction. Our findings are similar to observations in H3B-8800-treated K562 cell line[63]: while they detect the presence of a weaker BP in short retained introns with high GC content, we observe that these introns have a weaker 3'ss. In addition, we also report that those introns have a weaker 5'ss and shorter flanking exons. Regarding the exons skipped upon H3B-8800 treatment, we identified for the first time specific features that have not been looked at in other datasets. These skipped exons are shorter, with shorter flanking introns and exons and stronger 5'ss and 3'ss compared to alternative non-changing exons. These finding suggest that H3B-8800 selectively modulates introns and exons with specific features.

Finally, the general dysregulation found in spliceosome encoding genes upon H3B-8800 treatment can be of great interest. Several splicing factors and regulators are differentially spliced or expressed according to previous work and our findings[63,205]. Even if this dysregulation appears to be independent of the mutational status of splicing factors, we hypothesize that its impact is stronger on cancer cells bearing splicing factor mutations, as these cells depend on the

WT allele and would be more vulnerable to splicing perturbations. However, most probably, the differences in lethality between cancer cells bearing SF mutants can also be related to the fact that differentially spliced AS events and GE changes can be found in the *SF3B1*^{K700E} cells compared to *SF3B1*^{WT} cells after H3B-8800 exposure. The systematic study of these AS events and GE changes could shed light on the preferential lethality induced by H3B-8800 in *SF3B1* mutant CLL cells.

In this Thesis, with alternative splicing and SF3B1 as my fellow travelers, I have investigated the pathogenic mechanisms that are associated with worse prognosis in CLL, as well as the therapeutic opportunities that these insights can offer in the treatment of this disease. I have had the opportunity to serve as the bridge between two labs with different background to face the challenge from different perspectives. I hope that the conclusions of this work will help to better understand CLL and one day improve the management and outcome of patients with CLL.

CONCLUSIONS

Conclusions from chapter I: Transcriptomic Characterization of *SF3B1*-Mutated Chronic Lymphocytic Leukemia

- We used CRISPR-mediated genome editing to generate isogenic MEC1 CLL lines harboring *SF3B1*^{WT} or (heterozygous) *SF3B1*^{K700E} alleles and demonstrated their utility as models to study alternative splicing changes associated with *SF3B1* mutations in CLL patient samples.
- *SF3B1* mutation induces the selection of weaker Alt 3'ss located upstream or downstream of the canonical 3'ss in CLL cells.
- *SF3B1* mutation induces alternative splicing changes in genes involved in biological processes relevant in CLL.
- Aberrant MAP3K7 alternative 3' splice usage upon *SF3B1* mutation in isogenic MEC1 cell lines decreases MAP3K7 protein levels and results in NF-κB hyperactivation.
- An aberrant isoform in the gene ZDHHC16 might be used as a surrogate marker for the presence of functional *SF3B1* mutations in CLL and possibly other tumors.

Conclusions from chapter II: Study of the therapeutic effects of SF3B1-binding H3B-8800 in in vitro and in vivo models of Chronic Lymphocytic Leukemia

- H3B-8800 displays cytotoxic effects in CLL primary cells, MEC1 cell lines and CLL cells in NSG mice.
- H3B-8800 induces preferential lethality in *SF3B1*-mutated CLL cells.
- H3B-8800 treatment induces the skipping of short exons flanked by stronger splice sites compared to alternative non-changing exons and retention of short introns with high GC content.
- The combination of H3B-8800 with venetoclax displays synergistic cytotoxic effects in CLL MEC1 cell lines, suggesting convergent effects of targeting different pro-apoptotic Bcl2 genes and isoforms.

MATERIALS AND METHODS

1. Cell lines

1.1. Cell culture

MEC1 (DSMZ: ACC497) CLL cell lines were subcultured at 5 × 10^5 cells/ml in Iscove's modified Dulbecco's medium (IMDM) GlutaMAX™ medium (GIBCO 31980-022) supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS), 1% (v/v) penicillin-streptomycin 10.000U/mL (GIBCO, 15140122) and 1%(v/v) GlutaMax 100X (GIBCO, 25030-24) in a humidified atmosphere at 37°C containing 5% carbon dioxide. The follicular dendritic cell line HK, kindly provided by Dr. Yong Sung Choi (Alton Ochsner Medical Foundation, New Orleans, LA)[206], was cultured in RPMI1640 (61870036, GIBCO) supplemented with 20% FBS, 1% (v/v) penicillin-streptomycin 10.000U/mL and 1%(v/v) GlutaMax 100X. HS5 (ATCC: CRL-11882) cell line was cultured in Dulbecco's modified Eagle Medium (DMEM) (GIBCO, 10565018) supplemented with 10% (v/v) heat-inactivated FBS, 1% (v/v) penicillin-streptomycin 10.000 U/ml and 1% (v/v) GlutaMax 100X (GIBCO, 25030-24).

Every batch of cell lines was tested for Mycoplasma with EZ-PCR™ Mycoplasma Detection Kit (20-700-20, Biological Industries) periodically when they were in culture.

1.2. Establishment of CRISPR Cas9-edited cell lines

1.2.1. Design

crRNAs were design using DESKGEN AI library design software. We designed the single strand ODN complementary to the non-target strand[131] with 2 nucleotide changes compared to the reference sequence: first, we replace c.2098A>G to induce a lysine to glutamic acid change and introduced a PAM silent mutation changing XGG PAM sequence to XGA (c.2106G>A) to avoid further recognition of the PAM sequence once it was properly repaired[132](Fig. M.1A). Phosphorothioate-modified nucleotides at the ends of the ssODN were introduced to stabilize the oligos, which were ordered to IDT[133]. Oligonucleotide sequences used for the generation and characterization of *SF3B1*^{K700E} MEC1 cell line are shown in Table 1.

1.2.2. Procedure

crRNA and Alt-R® CRISPR-Cas9 tracrRNA, ATTO™ 550 (IDT Ref. 1075927) annealing was done at 95C for 5 min. Assembly of the crRNA:tracRNA duplex (10uM) and the Cas9 protein (10uM) was done and mixed with ssODN (10uM) and Alt-R® Cas9 Electroporation Enhancer (IDT Ref. 1075915). Neon Transfection System (Termo Fisher Ref. MPK 5000) with a 20ms pulse at 1600 volts was used to electroporate MEC1 cell lines. Cells were seeded in 300uL of IMDM medium prepared as previously described plus 10uM of RS-1(Sigma Aldrich, R9782) to enhance homology directed repair (HDR)[134]. 72 hours later, cells were seeded by serial dilution in single cell cultures in round botton 96 wells plates (Corning Incorporated COSTAR®, 3799) for 3 weeks and then

screening of the clones was carried out using Taqα1 restriction enzyme, which cut the sequence containing the silent mutation we introduced in the PAM sequence (c.2106G>A)(Fig. M.1B).

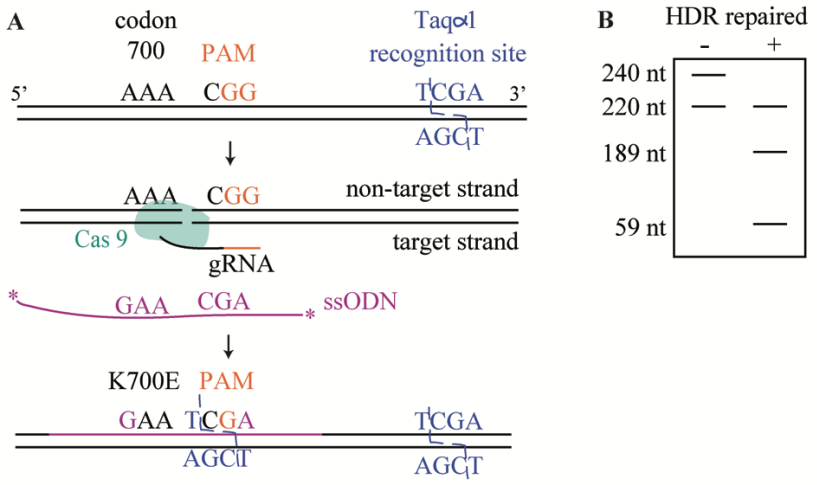


Figure M.1. Design of the CRISPR/Cas9 experiments. (A) Representation of the CRISPR Cas9 *SF3B1* K700E knock-in strategy, indicating the position of codon 700, PAM sequence and Taqα1 recognition site, position of sgRNA and donor ssODN with nucleotide changes to induce the K700E mutation and PAM sequence editing. (B) Expected size of PCR amplified bands corresponding to HDR clones after digestion with Taqα1 restriction enzyme, depending on absence or presence of edited sites.

Primer's name	**Sequence**
Guide RNA	5'-/AltR1/rUrGrGrArUrGrArGrCrArGrCrArGrArArArGrUrUrGrUrUrUrUrUrArGrArGrCrUrArUrGrCrU/AltR2/-3'
CRISPR template used in the 1st experiment	TAACTTAGGTAATGTTGGGGCATAGTTAAAACCTGTGTTTGGTTTTGTAGGTCTTGTGGATGAGCAGCAGGAAGTTCGGACCATCAGTGCTTTGGCCATTGCT
CRISPR template with K700E and PAM-blocking mutation	C*A*TAAAGGCTTTAACACAGAATCAAAAGATTCGATACCATAAGGAGTTGCTGCTTCAGCCAAGGCAGCAATGGCCAAAGCACTGATGGTTCGAACTTCCTGCTGCTCATCCACAAGACCTACAAAACC*A*A
SF3B1 gRNA detection F	CTGGATGATATTGTGTAACTTAGG
SF3B1 gRNA detection R	ATACCATAAGGAGTTGCTGC

SF3B1 amplicon F	CTGCTGACAGGCTATGGTTCATG
SF3B1 amplicon R	GGCATATTCTGCATCCATAAGAGG

Table 1. Oligonucleotide sequences used for the generation and characterization of *SF3B1*^{K700E} MEC1 cell line.

1.2.3. DNA extraction and quantification of *SF3B1* K700E mutation status using digital PCR:

DNA extraction was carried out using QIAamp® DNA Micro Kit (<1x10^6.) or QIAamp® DNA Mini Kit (>1x10^6) (Qiagen), depending of the number of cells. The amount of extracted DNA was quantified using a NanoDrop 1000 spectrophotometer (Thermo scientific). Then, nanoliter-sized droplets were generated by QX200 Droplet Generator (Biorad, #1864002) and amplified by PCR using commercial labeled primers to quantify the presence of K700E *SF3B1* mutation. Fluorescence assays were read using a QX200 Droplet Reader (Biorad, #1864003) and data analysis was carried out using Quanta Soft software (v1.7).

1.2.4. Characterization of CRISPR cell lines using cell viability and MTT assay

30.000 MEC1 cells/well were seeded in a 96-well flat bottom polystyrene culture plate and incubated for 48 hours. Cell viability was measured by flow cytometry using Annexin V (FITC) and propidium iodide (Invitrogen, BMS500FI-300). For the MTT assay 10.000 cells/well were seeded and incubated for 48 hours. Then, Thiazolyl Blue Tetrazolium Bromide 98% at a concentration of 0,5 mg/ml in PBS was added to the cells and were incubated for 2 hours until the presence of violet crystal was observed at the optical microscope (Leica DMI1). The stop solution (isopropanol and HCl 24:1) was added and the cells were carefully resuspended before

reading the plates at the TECAN spectrophotometer at 570 and 670 wavelengths.

2. Human samples

Peripheral blood mononuclear cells (PBMCs) from CLL patients (with ≥90% tumor B cells) and healthy donors were isolated, cryopreserved and stored within the Hematopathology collection registered at the Biobank (Hospital Clinic-IDIBAPS; R121004-094). Clinical and biological data of each patient are detailed in Supplementary Table 1. *SF3B1* mutations were analyzed in previous sequencing studies[98,100,207].

3. RNAseq analysis

3.1. ICGC Patients

We analyzed 298 patients from the Internationa Cancer Genome Consortium (ICGC) presenting CLL, MBL or SLL at diagnosis. We then compared 15 *SF3B1* mutated patient samples and 225 *SF3B1* WT CLL patient samples, none of them presenting mutations in other splicing factors or RNA metabolism-related genes. *SF3B1* mutated patients presented high cancer cell fraction (>50%). Patients with *SF3B1* E862K or M757T were not included in the comparison groups. The biological and clinical characteristics are shown in Supplementary Table 1. The samples were collected before any treatment administration and were processed as described in[98].

3.2. H3B-8800 treated cell lines

4 million cells were treated with 75nM H3B-8800 or DMSO for 6 hours. RNA extraction was carried out using Quiagen RNeasy Mini kit. RNA sequencing was performed by the CRG Genomics Core Facility (Illumina HiSeq 2500 v4). Stranded mRNA-seq libraries were prepared and 2 samples/lane were sequenced following 2x125 nt paired-ended protocol. Triplicates were sequenced for each condition in separate lines.

3.3. Splicing analyses

Splicing analyses of RNA-Seq data were performed using the toolset Vertebrate Alternative Splicing and Transcription Tools[135,136] (*vast-tools*, version 2.1.3 for the patients and 2.5.1 for MEC1 cell lines) (https://github.com/vastgroup/vast-tools). For each event, a minimum 10 reads coverage was required. To detect differentially spliced events between two conditions (*SF3B1* mutation or H3B-8800 treatment), we required a ΔPSI of at least 15% between the merged replicates of each condition and a minimum range of 5 between the PSI values of the two conditions.

For a wider search of annotated and not annotated Alt 3'ss we developed JuncExplorer (https://gitlab.com/aghr/juncexplorer). JuncExplorer allows the identification and quantification of annotated and not annotated (de-novo) splice junctions from mapped RNA-Seq reads. Percent Splice Site Usage (PSU) values are calculated by converting the read numbers into relative frequencies per case for each of its splice sites for each of the 5' and 3'ss. Alt 3'ss cases were considered when one donor site (5') had several acceptor

sites (3') within a single intron. To detect differentially spliced events between two conditions (*SF3B1* mutation or H3B-8800 treatment), we required a ΔPSU of at least 15% between the merged replicates of each condition. Cryptic 3'ss correspond to alternative 3'ss not detected in *SF3B1* WT samples and detected in *SF3B1* mutated samples. Downstream analysis of alternative splicing events was performed using Matt (http://matt.crg.eu) with special focus on exon and intron features, splice sites strengths[141] and BP predictions[142].

3.3.1. Venn diagrams

Venn diagrams were generated with *VennDiagram* package in R (v3.6.3).

3.3.2. Correlation analysis

Pearson coefficient correlation plots were obtained with *ggplot* and *ggpubr* packages in R (v3.6.3).

3.3.3. Prediction of the impact on protein sequence of the alternative splicing events

Prediction of impact on protein sequence was obtained from VastDB (https://vastdb.crg.eu/wiki/Downloads). First, AS sequences were mapped to non-coding (UTRs) or coding (CDS) sequences. Then, the events were predicted to disrupt the CDS if their inclusion or skipping would induce a frameshift in their open reading frame (ORF), they would induce a premature stop codon (PTC) predicted to trigger NMD or an in-frame stop codon would generate a protein 100 amino acids shorter than the reference protein. The rest of the CDS mapping events were considered to preserve the transcript's coding-potential.

3.4. Gene expression analysis

Illumina universal adapters sequences were removed from the RNA-Seq reads with Cutadapt version 2.4 discarding all reads with length shorter than 15. The adapter-trimmed reads were mapped with STAR 2.7.1.a[208] to the human genome version GRCh38. Read counts per gene where generated with STAR (--quantMode GeneCounts) using uniquely mapping reads only. DESeq2 version 1.14.1[148] was used to determine fold changes and p values for all annotated genes comparing these conditions: 15 *SF3B1* mutated CLL patients vs 225 *SF3B1* WT CLL patients, *SF3B1*$^{WT/K700E}$ MEC1 cell lines treated with 75nM H3B-8800 after 6 hours. In case of patient data, we corrected for batch effects caused by the different provenance (hospital) of the different data sets included in this study. Therefore, we define a batch variable and added it to the model definition of the generalized linear model used by DESeq2. iDEP v.0.94 [149]was used to do the heatmaps.

3.5. Principal Component Analysis (PCA)

PCA was performed on the PSI values of Alt3'ss, Alt5'ss, IR and ES of each sample using *ggfortify and ggplot2* packages in R(v3.6.3).

3.6. Gene Ontology enrichment analysis

Differentially expressed genes or genes bearing differentially spliced (DPSU ≥15) AS events (Alt3'ss, Alt5'ss, IR and ES) between *SF3B1* wild-type and mutated patients' and MEC1 cell line's groups were analyzed using *enrichGO* and *clusterProfiler* packages in

R(v3.6.3). Statistical significance was defined with adjusted p value < 0.05 and the positive false discovery rate with q value <0.05.

4. Experimental validation of RNA-seq data

4.1. RNA extraction and reverse transcription, semi-quantitative RT-PCR

RNA extraction and DNAse treatment were performed using Maxwell simplyRNA kit (Promega, AS1340), illustra^{TM} RNAspin Mini RNA Isolation Kit (GE Healthcare, 25-0500-72) or RNeasy Mini kit (Qiagen, 74104), following the manufacturers' instructions. In total, 50 ng of total RNA were reverse transcribed with Superscript III (Invitrogen, 12087539) following the manufacturer's recommendations. PCR reactions were carried out using GoTaq enzyme (Promega, M3001). The PCR products were separated by electrophoresis on 6% acrylamide gels run in TBE 1X, stained with gel red (Biotium, 41003) and analyzed with GelDocTM XR+ (BIO RAD). The bands were quantified with ImageJ v1.53a. The primers used are listed in Table 2.

Primer's name	**Sequence**
INPLL1_F	CCTGTAGAGGGTGAGCGAGA
INPLL1_R	GAGCTGTGGGAGTCTCAGGA
DLG1_R	TCCTGACAAATTTGGATCCTG
DLG1_F	TTATACTGGCCAGCTTCAATG
DYNLL1_R	TCGGTAGCGACGGTATCTCT
DYNLL1_F	CTGTTGCATCTCTTCCGACA
HLTF_R	ATGCAGATGTCCTGGGTCTT
HLTF_F	AATCCAGGCAAATTGCACA

MAP3K7_F	AATATGCTGAAGGGGGCTCT
MAP3K7_R	TCCCTGTGAATTAGCGCTTT
TGFBR1_F	GCACCCTCTTCAAAAACTGG
TGFBR1_R	CACACTGGTCCAGCAATGAC
UBA7_F	CACTGCACAAGTTCCAGCAC
UBA7_R	CACTAGGGCCTCATCCAGTG
ZDHHC16_F	ACCTCTGGTTCCTGTGCATC
ZDHHC16_R	CCCTACTCACTCTGCCCTTG
ANKHD1_F	GCAGACTTACGCACTGTGGA
ANKHD1_R	TCTGCTTTCATCCGTGTCAG
NFKB1_F	GCAGATGACATCCAGATTCG
NFKB1_R	CGAAGCTGGACAAACACAGA
ZNF561_F	TTCATAAGCTGGCCTGAGGT
ZNF561_R	CTTGCCAGGTAGTCCTCCAC
FBXO41_F	GGGGAAGAGGGTGATGTCTC
FBXO41_R	GCGTGTCCAGGTAGGTGAAG
TARBP1_F	TTTAAGCCATCCCCTACGTG
TARBP1_R	ATTCCTTCCACATCCACAGC

Table 2. Primers used for the validation of the different Alt3'ss patterns.

4.2. Quantitative real-time PCR (RT-qPCR)

Total RNA from CLL primary samples was extracted using the TRIZOL reagent (Invitrogen, 15596026) according to the manufacturer's instructions. One microgram of total RNA was then reverse transcribed to cDNA using random primers and the MMLV reverse transcriptase (Invitrogen, 28025013). MAP3K7 and ZDHHC16 expression levels were determined in StepOnePlus™ Real-Time PCR System (Applied Biosystems) using specific primers for aberrant and canonical isoforms (Table 3). The relative expression of each gene was quantified by the comparative cycle threshold (Ct) method (DDCt) using RPLPO as an endogenous

housekeeping gene. mRNA expression levels were provided as arbitrary quantitative PCR units, taking as a calibrator the median of the WT samples. In the case of ZDHHC16 aberrant isoform, which was not detected in the WT samples, the median of the *SF3B1* mutated samples was taken as reference.

Primer's name	**Sequence**
MAP3K7 F	GGAACACTGTAAACACCAACT
MAP3K7 aberrant junction R	GGGCTCTTTATATAATGTTTGTGC
MAP3K7 canonical junction R	CTCTTTATATAATGTGCTGCATGG
MAP3K7 probe	TGCGTGGGCAGCAGTATAATATGGC
ZDHHC16 F	CGAGAAAGGATGACTCACAAGAG
ZDHHC16 R	CCTTCTTGTTGATGTGCCTTTC
ZDHHC16 aberrant junction probe	CAGTCTTCGCCCCTCTTTTCTTAG
ZDHHC16 canonical junction probe	TCTGGTTCCTGTGCAGTTCTGTGG

Table 3. Primers used for the qPCR validations.

4.3. Fragment analysis

For the amplification of the RT-PCR products, one of the primers was labeled with FAM to obtain a fluorochrome-labeled single-strand (denatured) PCR product. These products were size separated in a capillary sequencing polymer and detected via automated scanning with a laser using ABI3130 Genetic Analyzer (Applied Biosystems ®). Results were analysed with geneMapper v6.0.

4.4. Minigenes

4.4.1. Cloning

pCMV 57 Δi Minx plasmid harboring 2 exons and an intron from the Adenovirus Major Late Promoter transcript, flanked by PT1 and

PT2 sequences that allow detection of transcripts from the expression vector [209] by RT PCR analysis, was used as a scaffold in which different 3'ss (from 20 nucleotides upstream of the predicted BP to 5 nucleotides downstream of the 3'ss) were introduced replacing equivalent sequences of the AdML intron (Table 4). For this, pCMV 57 Δi Minx plasmid was amplified by PCR using TaqPlus Precision (Agilent, 600212), deleting the AdML intron 2 sequences to be deleted, and DNA oligonucleotide primers corresponding to the sequences of interest were ordered to IDT and amplified with GoTag (Promega, M3001). The PCR products were purified by QIAquick PCR Purification Kit (Qiagen, Ref. 28106) and the methylated DNA template was digested with the enzyme DpnI (New England BioLabs, R0176L). The amplified vector was purified by agarose gel electrophoresis and purified using QIAquick Gel Extraction Kit (Qiagen, Ref. 28704) following the manufacturer's instructions. 20ng of the vector were used for Gibson cloning at a 1:8 vector-insert ratio with a mix of reagents from the Protein Technologies Unit core facility of the Centre for Genomic Regulation. For transformation, Stellar bacteria (Takara Bio, 636763) were used. Single colonies from LB Ampi plates were grown overnight in LB supplemented with ampicillin, DNA was purified using QIAprep Spin Miniprep Kit (Qiagen, Ref. 27106) and minigenes were sequenced (Eurofins) to verify the products of mutagenesis.

4.4.2. Transfections

MEC1 *SF3B1*^{WT} and *SF3B1*^{K700E} cell lines were transfected with 50ng of minigene DNA plasmids and Alt-R® Cas9 Electroporation Enhancer (IDT Ref. 1075915) using Neon Transfection System

(TermoFisher Ref. MPK 5000) with a 20ms pulse at 1600 volts. 24h later, cells were collected for RNA extraction and validation by RT-PCR. The PCR products from the TARBP1 gene were analyzed by fragment analysis as described above. PCR products of other minigenes were separated by electrophoresis on standard native polyacrylamide gels as described above.

Primer's name	**Sequence**
AdML EcoRVPT1Ex1	ATGCAGATATCgtcgacgacacttgctcaacACTCTTGGA TCGGAAACCCGT
AdML NotIPT2Ex2	ACGTAGCGGCCGCaagcttgcatcgaatcagtagGGATCCC CACTGGAAAGAC
VectorS (PCRATW)	GGTTGAGGACAAACTCTTCG
VectorAS (PCRATW)	AAGGAAACCCTGGACTACTG
PT1	GTCGACGACACTTGCTCAAC
PT2	AAGCTTGCATCGAATCAGTAG
SinglestrandF	CAGTAGTCCAGGGTTTCCTT
SinglestrandR	CGAAGAGTTTGTCCTCAACC
TARBP1 insert	CAGTAGTCCAGGGTTTCCTTTTTTGTTAATGCTT TGCAAACCTTAACTGTTGTTATTTAAGCAGGAG AAGGTTGAGGACAAACTCTTCG
TARBP1 insert BP1	CAGTAGTCCAGGGTTTCCTTTTTTGTTAATGCTT TGCAAACCTTCTCTGTTGTTATTTAAGCAGGGA GAAGGTTGAGGACAAACTCTTCG
TARBP1 insert BP2	CAGTAGTCCAGGGTTTCCTTTTTTGTTAATGCTT TGCCTGCCTTAACTGTTGTTATTTAAGCAGGGA GAAGGTTGAGGACAAACTCTTCG
MAP3K7 insert	CAGTAGTCCAGGGTTTCCTTATGTGTTCTGATA GTGCTTGCATTCACATTGTGTCTTTTGTTATATG CAATAAAGTTCTTTTTAGTTTGTGCCTTTCTTTC GCAGTGCTGGGTTGAGGACAAACTCTTCG
MAP3K7 insert BP2	CAGTAGTCCAGGGTTTCCTTATGTGTTCTGATA GTGCTTGCATTCGCATTGTGTCTTTTGTTATATG CAATAAAGTTCTTTTTAGTTTGTGCCTTTCTTTC GCAGTTGCTGGGTTGAGGACAAACTCTTCG
MAP3K7 insert BP1	CAGTAGTCCAGGGTTTCCTTATGTGTTCTGATA GTGCTTGCATTCACATTGTGTCTTTTGTTATATG CAATCTGGTTCTTTTTAGTTTGTGCCTTTCTTTC GCAGTTGCTGGGTTGAGGACAAACTCTTCG

ZNF561 insert	CAGTAGTCCAGGGTTTCCTTGAAGGATGGGCA GCTCTGGCTTCTCATGCCTTCTTACTTTTCCACT AGGGTTTTTTTCCAGGGGAACGGTTGAGGACA AACTCTTCG
ZNF561 insert BP2	CAGTAGTCCAGGGTTTCCTTGAAGGATGGGCA GCTCTGGCTTCTCGTGCCTTCTTACTTTTCCACT AGGGTTTTTTTCCAGGGGAACGGTTGAGGACA AACTCTTCG
ZNF561 insert BP1	CAGTAGTCCAGGGTTTCCTTGAAGGATGGGCA GCTCTGGCTTCTCATGCCTTCTTCCTTTTCCACT AGGGTTTTTTTCCAGGGGAACGGTTGAGGACA AACTCTTCG
INPPL1 insert	CAGTAGTCCAGGGTTTCCTTTTGGCCATGATGC CGGGGCCCTTTAACCCTCTTTCCATGGAAGTCA CTTTACAGCTGCATCGTGCCTCCTACTCCACTG AGTGTGGGAGGCCCAAACGGCTGCCCACTGAC CCCTGCCCACAGAATGGGGGTTGAGGACAAAC TCTTCG
FBXO41 insert	CAGTAGTCCAGGGTTTCCTTCACACACATCCCC ACACCAGCCTTGACCCTCCACGCTCACTTTTGA ACCCCTGGTATTCTGCAGCCCCGCCCCGAGGGA GGCACTGGGCGGGGTCGGCGAGCAGAGAGGGT CAGCCCCTCACGCTCCAATGAGGTCATCAGCCC CAGGGTTGAGGACAAACTCTTCG
Minigene_exon_F	CTCCGAACGGTAAGAGCCTA
Minigene_exon_R	AAAAAGGGACAGGATAAGTATGACA

Table 4. Oligonucleotide sequences used to build and analyze the minigenes designed to recapitulate changes in 3' splice site selection induced by SF3B1 mutations and dissect relevant sequence elements.

4.5. Western blots

5 x 10^6 cells/condition were used for protein extraction with Tris-TritonX Lysis buffer (20mM Tris-HCl pH 7.6, 150mM NaCl, 1% TritonX-100), protease inhibitor cocktail (Sigma-Aldrich, 11697498001) and phosphatase inhibitor cocktail (Sigma-Aldrich, 4906845001) for 30 min at 4ºC. Lysates were centrifuged at 13200 rpm for 15 min at 4ºC. Protein concentration was assessed by Bradford assay (Bio-Rad, #5000006). Total extracts were subjected to 10% SDS-PAGE separation and transferred to an Immobilon-P

PVDF Membrane (Merk Millipore, IPVH00010). Primary antibodies (Table 5) were incubated overnight at 4ºC and secondary antibodies (Table 5) for 1 hour at room temperature. Antibodies were incubated in BSA and washed with TBST. Pierce ECL western blotting substrate (Termofisher, #32106) was used to develop the membrane and Amersham ECL Select Western Blotting Detection Reagent (Cytiva, RPN2235) for the detection of low protein levels. Immunodetection was carried out using a Image Quant™ LAS 4000 mini and quantified by Multi Gauge Software (FUJIFILM).

Antibody's name	**Ref.**	**Provider**
Anti-SAP155 antibody (SF3B1)	AB37578	Abcam
TAK1 Antibody	#4505	Cell Signaling Technology
Phospho-p38 MAPK (Thr180/Tyr182) (D3F9) XP® Rabbit mAb	#4511T	Cell Signaling Technology
p38 MAPK (D13E1) XP® Rabbit mAb	#8690T	Cell Signaling Technology
Phospho-NF-κB p65 (Ser536) (93H1) Rabbit mAb	#3033	Cell Signaling Technology
NFκB p65 Anticuerpo	sc-8008	Santa Cruz Biotechnology
NF-κB1 p105/p50 (D4P4D) Rabbit mAb	#13586	Cell Signaling Technology
Monoclonal Anti-α-Tubulin antibody produced in mouse	T6074	Sigma-Aldrich
Monoclonal Anti-β-Actin antibody produced in mouse, clone AC-15	A5441	Sigma-Aldrich
Monoclonal Anti-β-Tubulin antibody produced in mouse	T4026	Sigma-Aldrich
Lamin B1 (D9V6H) Rabbit mAb	#13435S	Cell Signaling Technology
Anti-rabbit IgG, HRP-linked Antibody	#7074	Cell Signaling Technology
Anti-mouse IgG, HRP-linked Antibody	#7076	Cell Signaling Technology

Table. 5 Primary and secondary antibodies used for Western blot analyses.

4.6. NF-κB activity measurements

NF-κB activity was measured by western blot and using ELISA-based TransAM NF-κB Family Transcription Factor Assay Kit (Active motif, 43296) in nuclear extracts obtained using Nuclear extract kit (Ref. 40010).

5. H3B-8800 treatment

5.1. H3B-8800-induced cytotoxicity studied by flow cytometry

30.000 MEC1 cells/well or 200.000 primary cells/well were seeded in a 96-well flat bottom polystyrene culture plates (Termo Scientific, 167008) and incubated for 48 hours with H3B-8800 (kindly provided by H3B Biomedicine) at different concentrations. Cytotoxicity was measured by Annexin V (Pacific blue/FITC) and PI-mediated (Invitrogen, BMS500FI-300 and A35122) flow cytometry (Attune acoustic focusing cytometer from Applied Biosystems for primary cells and BD LSRFortessa^{TM} Flow Cytometer for the cell lines).

5.2. Co-culture experiments

HS5 and HK cells were seeded on day 1 at a densities of 20.000 cells/well and 10.000 cells/well, respectively. On day 2, 20.000 MEC1 cells/well or 200.000 primary cells/well were added and incubated with 75nM H3B-8800 for 48h. Cytotoxicity was assessed by flow cytometry as described above, after selection of MEC1 cells by CD19+ (BD, 345789) positivity.

5.3. Synergy experiments with venetoclax

Experiments were carried out as above, except that H3B-8800 and venetoclax (ABT-199, Selleckchem, S8048) at various concentrations were incubated for 48 hours and the synergism between them was studied using SynergyFinder 2.0[162].

5.4. Stable expression of luciferase and GFP

293T cells were transfected using jetPEI® (Ref. 101-10N) with pLV430G-oFL-T2A-eGFP, a lentivirus expressing the luciferase gene cloned in an EGFP-expressing vector (gift from Amer Najjar, M. D. Anderson Cancer Center, Houston, TX) and viral stocks were collected from the culture supernatant 48 h post-transfection. MEC1 cell lines were infected and GFP-positive cells were sorted by flow cytometry 7 days afterwards.

5.5. *In vivo* model

5.5.1. Experimental design

6-8 weeks old NSG mice were used which lack B, T and NK lymphocytes. 10 millions of MEC1^{K700E} (17 mice) or MEC1^{K700K} (16 mice) were injected intravenously. Leukemic infiltration was followed by bioluminescence using a Fluorescence/luminescence imaging system AEQUORIA (Hamamatsu Photonics K.K.). 10 days after the injection, once leukemic infiltration was detected by bioluminiscence, H3B-8800 (6mg/kg) or vehicle control (95% methylcellulose 0.5% and 5% ethanol 100%) was administrated orally and daily for 10 days until the mice were sacrificed 24h after receiving the last dose. The experimental design was approved by the

Animal Testing Ethic committee of the University of Barcelona and Generalitat de Catalunya (Protocol #9680).

5.5.2. Samples processing

Peripheral blood was extracted from the submaxillary vein. For bone marrow extraction, we first cut both femurs of each mouse at the neck and next to the condyles, and then injected IMDM medium with a syringe and a 25G needle to facilitate bone marrow extraction. Peripheral blood, bone marrow, spleen and liver samples were first processed with ACK lysing buffer from Lonza (BP10-548E) then, leukemic infiltration was measured using CD19+ kit from Invitrogen (63-0198-42) to detect the tumoral B lymphocytes in a BD LSRFortessa^{TM} Flow Cytometer. Mouse spleens were fixed in 10% formaldehyde, dehydrated by ethanol after 24 hours and embedded in dissolved paraffin. The specimens were finally stained by hematoxylin-eosin (HE) (Pan Reac Appli Chem, 256991.1212 and Química Clínica Aplicada, 99 14 15) and CD79a (Roche, 790-4432) immunohistochemical staining and then analyzed in an Olympus BX53 optical microscope.

6. Statistical analysis

Statistical tests were performed as indicated in figure legends with R (v3.6.3) or GraphPad Prism (v8.4.0).

ABBREVIATIONS

Abbreviation	Explanation
3'ss	3'splice site
5'ss	5'splice site
AS	Alternative splicing
Alt3'ss	Alternative 3'ss
Alt5'ss	Alternative 5'ss
BCR	B cell receptor
BP	Branch point
CMML	Chronic Myelomonocytic Leukemia
CRISPR/Cas9	Clustered Regularly Interspaced Short Palindromic Repeats/Cas 9
CLL	Chronic lymphocytic leukemia
dPCR	Digital PCR
ΔPSI	Delta PSI
dPSU	Delta Percent Spliced site Usage
ES	Exon skipping
ESMO	European Society for Medical Oncology
ESE	Exonic splicing enhancer
ESS	Exonic splicing silencer
FBS	Fetal bobine serum
FCR	Fludarabine, cyclophosphamide and rituximab combination
Fig.	Figure
HDR	Homology directed repair
HEAT	Huntingtin elongation factor 3 protein phosphatase 2A, target of rapamycin 1
IG	Immunoglobulin
IGHV	Immunoglobulin heavy chain variable
IR	Intron retention
ISE	Intronic splicing enhancer

ISS	Intronic splicing silencer
M-CLL	IGHV mutated chronic lymphocytic leukemia
MBL	Monoclonal B lymphocytosis
MDS	Myelodysplastic syndrome
mRNA	Messenger RNA
NEHJ	non-homologous-end-joining
NMD	Nonsense-mediated decay
NSG	NOD-SCID interleukin-2 receptor gamma (*IL2Rγ*)^{null}
nt	Nucleotide(s)
OS	Overall survival
PBMC	Peripheral blood mononuclear cells
PPT	Polypyrimidine tract
Pre-mRNA	Precursor mRNA
PFS	Progression-free survival
PSI	Percent spliced in
PSU	Percent spliced usage
PTC	Premature stop codon
Py tract	Polypyrimidine tract
RBM	RNA binding motif
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
SF1	Splicing Factor 1
SF3B1	Splicing Factor B subunit 1
SLL	Small lymphocytic lymphoma
snRNP	Small nuclear ribonucleoproteins
SR protein	Serine-arginine-rich protein
ss	Splice site
U-CLL	IGHV unmutated chronic lymphocytic leukemia
U2AF1	U2 auxiliary factor 1
U2AF2	U2 auxiliary factor 2
UTR	Untranslated region
WT	Wild-type

REFERENCES

1. Crick, F. Central dogma of molecular biology. Nature, 227(5258). *Nat. Biotechnol.* **1970**, *227*, 561.
2. Francis, C. Split Genes and RNA Splicing. *Science (80-.).* **1979**, *204*, 264–271, doi:10.1126/science.373120.
3. Wahl, M.C.; Will, C.L.; Lührmann, R. The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell* **2009**, *136*, 701–718, doi:10.1016/j.cell.2009.02.009.
4. Papasaikas, P.; Valcárcel, J. The Spliceosome: The Ultimate RNA Chaperone and Sculptor. *Trends Biochem. Sci.* **2016**, *41*, 33–45, doi:10.1016/j.tibs.2015.11.003.
5. Ruskin, B.; Krainer, A.R.; Maniatis, T.; Green, M.R. Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. *Cell* **1984**, *38*, 317–331, doi:10.1016/0092-8674(84)90553-1.
6. Bonnal, S.C.; López-Oreja, I.; Valcárcel, J. Roles and mechanisms of alternative splicing in cancer — implications for care. *Nat. Rev. Clin. Oncol.* **2020**, *17*, 457–474, doi:10.1038/s41571-020-0350-x.
7. Wu, S.; Romfo, C.M.; Nilsen, T.W.; Green, M.R. Functional recognition of the 3′ splice site AG by the splicing factor U2AF35. *Nature* **1999**, *402*, 832–835, doi:10.1038/45590.
8. Valcárcel, J.; Gaur, R.K.; Singh, R.; Green, M.R. Interaction of U2AF65 RS region with pre-mRNA branch point and promotion of base pairing with U2 snRNA [corrected]. *Science* **1996**, *273*, 1706–1709, doi:10.1126/science.273.5282.1706.
9. Abovich, N.; Rosbash, M. Cross-Intron Bridging Interactions in the Yeast Commitment Complex Are Conserved in Mammals. *Cell* **1997**, *89*, 403–412, doi:10.1016/S0092-8674(00)80221-4.
10. Taggart, A.J.; Lin, C.L.; Shrestha, B.; Heintzelman, C.; Kim, S.; Fairbrother, W.G. Large-scale analysis of branchpoint usage across species and cell lines. *Genome Res.* **2017**, *27*, 639–649, doi:10.1101/gr.202820.115.
11. Query, C.C.; Moore, M.J.; Sharp, P.A. Branch nucleophile selection in pre-mRNA splicing: evidence for the bulged duplex model. *Genes Dev.* **1994**, *8*, 587–597, doi:10.1101/gad.8.5.587.
12. Spadaccini, R.; Reidt, U.; Dybkov, O.; Will, C.; Frank, R.; Stier, G.; Corsini, L.; Wahl, M.C.; Lührmann, R.; Sattler, M. Biochemical and NMR analyses of an SF3b155-p14-U2AF-

RNA interaction network involved in branch point definition during pre-mRNA splicing. *RNA* **2006**, *12*, 410–425, doi:10.1261/rna.2271406.
13. Wang, C.; Chua, K.; Seghezzi, W.; Lees, E.; Gozani, O.; Reed, R. Phosphorylation of spliceosomal protein SAP 155 coupled with splicing catalysis. *Genes Dev.* **1998**, *12*, 1409–1414, doi:10.1101/gad.12.10.1409.
14. Konarska, M.M.; Vilardell, J.; Query, C.C. Repositioning of the reaction intermediate within the catalytic center of the spliceosome. *Mol. Cell* **2006**, *21*, 543–553, doi:10.1016/j.molcel.2006.01.017.
15. Will, C.L.; Luhrmann, R. Spliceosome Structure and Function. *Cold Spring Harb. Perspect. Biol.* **2011**, *3*, a003707–a003707, doi:10.1101/cshperspect.a003707.
16. Turunen, J.J.; Niemelä, E.H.; Verma, B.; Frilander, M.J. The significant other: splicing by the minor spliceosome. *Wiley Interdiscip. Rev. RNA* **2013**, *4*, 61–76, doi:10.1002/wrna.1141.
17. Verma, B.; Akinyi, M. V; Norppa, A.J.; Frilander, M.J. Minor spliceosome and disease. *Semin. Cell Dev. Biol.* **2017**, 1–10, doi:10.1016/j.semcdb.2017.09.036.
18. Chen, M.; Manley, J.L. Mechanisms of alternative splicing regulation: Insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 741–754, doi:10.1038/nrm2777.
19. Gallego-Paez, L.M.; Bordone, M.C.; Leote, A.C.; Saraiva-Agostinho, N.; Ascensão-Ferreira, M.; Barbosa-Morais, N.L. Alternative splicing: the pledge, the turn, and the prestige: The key role of alternative splicing in human biological systems. *Hum. Genet.* **2017**, *136*, 1015–1042, doi:10.1007/s00439-017-1790-y.
20. Monzón-Casanova, E.; Matheson, L.S.; Tabbada, K.; Zarnack, K.; Smith, C.W.J.; Turner, M. Polypyrimidine tract-binding proteins are essential for B cell development. *Elife* **2020**, *9*, e53557, doi:10.7554/eLife.53557.
21. Daguenet, E.; Dujardin, G.; Valcárcel, J. The pathogenicity of splicing defects: mechanistic insights into pre- mRNA processing inform novel therapeutic approaches . *EMBO Rep.* **2015**, *16*, 1640–1655, doi:10.15252/embr.201541116.
22. Scotti, M.M.; Swanson, M.S. RNA mis-splicing in disease. *Nat. Rev. Genet.* **2016**, *17*, 19–32, doi:10.1038/nrg.2015.3.

23. Kahles, A.; Lehmann, K.-V.; Toussaint, N.C.; Hüser, M.; Stark, S.G.; Sachsenberg, T.; Stegle, O.; Kohlbacher, O.; Sander, C.; Caesar-Johnson, S.J.; et al. Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients. *Cancer Cell* **2018**, *0*, 1–14, doi:10.1016/J.CCELL.2018.07.001.
24. Rahman, M.A.; Krainer, A.R.; Abdel-Wahab, O. SnapShot: Splicing Alterations in Cancer. *Cell* **2020**, *180*, 208-208.e1, doi:10.1016/j.cell.2019.12.011.
25. Oltean, S.; Bates, D.O. Hallmarks of alternative splicing in cancer. *Oncogene* **2014**, *33*, 5311–5318, doi:10.1038/onc.2013.533.
26. Seiler, M.; Peng, S.; Agrawal, A.A.; Palacino, J.; Teng, T.; Zhu, P.; Smith, P.G.; Buonamici, S.; Yu, L. Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types. *Cell Rep.* **2018**, *23*, 282–296, doi:10.1016/j.celrep.2018.01.088.
27. Anczukow, O.; Krainer, A.R. Splicing-factor alterations in cancers. *Rna* **2016**, *22*, 1285–1301, doi:10.1261/rna.057919.116.
28. Chen, S.; Benbarche, S.; Abdel-Wahab, O. Splicing factor mutations in hematologic malignancies. *Blood* **2021**, *138*, 599–612, doi:10.1182/blood.2019004260.
29. Komeno, Y.; Huang, Y.-J.; Qiu, J.; Lin, L.; Xu, Y.; Zhou, Y.; Chen, L.; Monterroza, D.D.; Li, H.; DeKelver, R.C.; et al. SRSF2 Is Essential for Hematopoiesis, and Its Myelodysplastic Syndrome-Related Mutations Dysregulate Alternative Pre-mRNA Splicing. *Mol. Cell. Biol.* **2015**, *35*, 3071–3082, doi:10.1128/mcb.00202-15.
30. Madan, V.; Kanojia, D.; Li, J.; Okamoto, R.; Sato-Otsubo, A.; Kohlmann, A.; Sanada, M.; Grossmann, V.; Sundaresan, J.; Shiraishi, Y.; et al. Aberrant splicing of U12-type introns is the hallmark of ZRSR2 mutant myelodysplastic syndrome. *Nat. Commun.* **2015**, *6*, 6042, doi:10.1038/ncomms7042.
31. Yoshida, K.; Sanada, M.; Shiraishi, Y.; Nowak, D.; Nagata, Y.; Yamamoto, R.; Sato, Y.; Sato-Otsubo, A.; Kon, A.; Nagasaki, M.; et al. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **2011**, *478*, 64–69, doi:10.1038/nature10496.
32. Lee, S.C.W.; North, K.; Kim, E.; Jang, E.; Obeng, E.; Lu, S.X.; Liu, B.; Inoue, D.; Yoshimi, A.; Ki, M.; et al. Synthetic

Lethal and Convergent Biological Effects of Cancer-Associated Spliceosomal Gene Mutations. *Cancer Cell* **2018**, *34*, 225-241.e8, doi:10.1016/j.ccell.2018.07.003.
33. Taylor, J.; Mi, X.; North, K.; Binder, M.; Penson, A.; Lasho, T.; Knorr, K.; Haddadin, M.; Liu, B.; Pangallo, J.; et al. Single-cell genomics reveals the genetic and molecular bases for escape from mutational epistasis in myeloid neoplasms. *Blood* **2020**, *136*, 1477–1486, doi:10.1182/blood.2020006868.
34. Wan, Y.; Wu, C.J. SF3B1 mutations in chronic lymphocytic leukemia. *Blood* **2013**, *121*, 4627–4634, doi:10.1182/blood-2013-02-427641.
35. Cretu, C.; Schmitzová, J.; Ponce-Salvatierra, A.; Dybkov, O.; De Laurentiis, E.I.; Sharma, K.; Will, C.L.; Urlaub, H.; Lührmann, R.; Pena, V. Molecular Architecture of SF3b and Structural Consequences of Its Cancer-Related Mutations. *Mol. Cell* **2016**, *64*, 307–319, doi:10.1016/j.molcel.2016.08.036.
36. Borišek, J.; Saltalamacchia, A.; Gallì, A.; Palermo, G.; Molteni, E.; Malcovati, L.; Magistrato, A. Disclosing the impact of carcinogenic SF3b mutations on pre-mRNA recognition via all-atom simulations. *Biomolecules* **2019**, *9*, 1–17, doi:10.3390/biom9100633.
37. Biankin, A. V; Waddell, N.; Kassahn, K.S.; Gingras, M.; Muthuswamy, L.B.; Johns, A.L.; Miller, D.K.; Wilson, P.J.; Wu, J.; Chang, D.K.; et al. Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. **2012**, *491*, 399–405, doi:10.1038/nature11547.Pancreatic.
38. Koboldt, D.C.; Fulton, R.S.; McLellan, M.D.; Schmidt, H.; Kalicki-Veizer, J.; McMichael, J.F.; Fulton, L.L.; Dooling, D.J.; Ding, L.; Mardis, E.R.; et al. Comprehensive molecular portraits of human breast tumours. *Nature* **2012**, *490*, 61–70, doi:10.1038/nature11412.
39. Martin, M.; Maßhöfer, L.; Temming, P.; Rahmann, S.; Metz, C.; Bornfeld, N.; Van De Nes, J.; Hitpass, L.K.; Hinnebusch, A.G.; Horsthemke, B.; et al. Exome sequencing identifies recurrent somatic mutations in EIF1AX and SF3B1 in uveal melanoma with disomy 3. *Nat. Genet.* **2013**, *45*, 933–936, doi:10.1038/ng.2674.
40. Martelotto, L.G.; De Filippo, M.R.; Ng, C.K.Y.; Natrajan, R.; Fuhrmann, L.; Cyrta, J.; Piscuoglio, S.; Wen, H.-C.; Lim,

R.S.; Shen, R.; et al. Genomic landscape of adenoid cystic carcinoma of the breast. *J. Pathol.* **2015**, *237*, 179–189, doi:https://doi.org/10.1002/path.4573.
41. Furney, S.J.; Pedersen, M.; Gentien, D.; Dumont, A.G.; Rapinat, A.; Desjardins, L.; Turajlic, S.; Piperno-Neumann, S.; de la Grange, P.; Roman-Roman, S.; et al. SF3B1 mutations are associated with alternative splicing in uveal melanoma. *Cancer Discov.* **2013**, *3*, 1122–1129, doi:10.1158/2159-8290.CD-13-0330.
42. Malcovati, L.; Karimi, M.; Papaemmanuil, E.; Ambaglio, I.; Jädersten, M.; Jansson, M.; Elena, C.; Gallì, A.; Walldin, G.; Porta, M.G.D.; et al. SF3B1 mutation identifies a distinct subset of myelodysplastic syndrome with ring sideroblasts. *Blood* **2015**, *126*, 233–241, doi:10.1182/blood-2015-03-633537.
43. Gangat, N.; Mudireddy, M.; Lasho, T.L.; Finke, C.M.; Nicolosi, M.; Szuber, N.; Patnaik, M.M.; Pardanani, A.; Hanson, C.A.; Ketterling, R.P.; et al. Mutations and prognosis in myelodysplastic syndromes: karyotype-adjusted analysis of targeted sequencing in 300 consecutive cases and development of a genetic risk model. *Am. J. Hematol.* **2018**, *93*, 691–697, doi:10.1002/ajh.25064.
44. Zhang, Z.; Chen, S.; Chen, S.; Chen, G.; Zhang, R.; Li, J.; Qu, J. SF3B1 mutation is a prognostic factor in chronic lymphocytic leukemia: A meta-analysis. *Oncotarget* **2017**, *8*, 69916–69923, doi:10.18632/oncotarget.19455.
45. Darman, R.B.; Seiler, M.; Agrawal, A.A.; Lim, K.H.; Peng, S.; Aird, D.; Bailey, S.L.; Bhavsar, E.B.; Chan, B.; Colla, S.; et al. Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3' Splice Site Selection through Use of a Different Branch Point. *Cell Rep.* **2015**, *13*, 1033–45, doi:10.1016/j.celrep.2015.09.053.
46. Kesarwani, A.K.; Ramirez, O.; Gupta, A.K.; Yang, X.; Murthy, T.; Minella, A.C.; Pillai, M.M. Cancer-associated SF3B1 mutants recognize otherwise inaccessible cryptic 3' splice sites within RNA secondary structures. *Oncogene* **2017**, *36*, 1123–1133, doi:10.1038/onc.2016.279.
47. Yokoi, A.; Kotake, Y.; Takahashi, K.; Kadowaki, T.; Matsumoto, Y.; Minoshima, Y.; Sugi, N.H.; Sagane, K.; Hamaguchi, M.; Iwata, M.; et al. Biological validation that SF3b is a target of the antitumor macrolide pladienolide.

FEBS J. **2011**, *278*, 4870–4880, doi:10.1111/j.1742-4658.2011.08387.x.
48. Alsafadi, S.; Houy, A.; Battistella, A.; Popova, T.; Wassef, M.; Henry, E.; Tirode, F.; Constantinou, A.; Piperno-Neumann, S.; Roman-Roman, S.; et al. Cancer-associated SF3B1 mutations affect alternative splicing by promoting alternative branchpoint usage. *Nat. Commun.* **2016**, *7*, 10615, doi:10.1038/ncomms10615.
49. Gupta, A.K.; Murthy, T.; Paul, K. V; Ramirez, O.; Fisher, J.B.; Rao, S.; Rosenberg, A.B.; Seelig, G.; Minella, A.C.; Pillai, M.M. Degenerate minigene library analysis enables identification of altered branch point utilization by mutant splicing factor 3B1 (SF3B1). *Nucleic Acids Res.* **2018**, *1*, 1–11, doi:10.1093/nar/gky1161.
50. Zhang, J.; Ali, A.M.; Lieu, Y.K.; Liu, Z.; Gao, J.; Rabadan, R.; Raza, A.; Mukherjee, S.; Manley, J.L. Disease-Causing Mutations in SF3B1 Alter Splicing by Disrupting Interaction with SUGP1. *Mol. Cell* **2019**, *76*, 82-95.e7, doi:10.1016/j.molcel.2019.07.017.
51. Liu, Z.; Zhang, J.; Sun, Y.; Perea-Chamblee, T.E.; Manley, J.L.; Rabadan, R. Pan-cancer analysis identifies mutations in SUGP1 that recapitulate mutant SF3B1 splicing dysregulation. *Proc. Natl. Acad. Sci.* **2020**, 201922622, doi:10.1073/pnas.1922622117.
52. Leeksma, A.C.; Derks, I.A.M.; Kasem, M.H.; Kilic, E.; de Klein, A.; Jager, M.J.; van de Loosdrecht, A.A.; Jansen, J.H.; Navrkalova, V.; Faber, L.M.; et al. The Effect of SF3B1 Mutation on the DNA Damage Response and Nonsense-Mediated mRNA Decay in Cancer. *Front. Oncol.* **2021**, *10*, 1–7, doi:10.3389/fonc.2020.609409.
53. Bergot, T.; Lippert, E.; Douet-Guilbert, N.; Commet, S.; Corcos, L.; Bernard, D. Human cancer-associated mutations of SF3B1 lead to a splicing modification of its own RNA. *Cancers (Basel)*. **2020**, *12*, 1–16, doi:10.3390/cancers12030652.
54. Boultwood, J.; Pellagatti, A.; Nikpour, M.; Pushkaran, B.; Fidler, C.; Cattan, H.; Littlewood, T.J.; Malcovati, L.; Della Porta, M.G.; Jädersten, M.; et al. The Role of the Iron Transporter ABCB7 in Refractory Anemia with Ring Sideroblasts. *PLoS One* **2008**, *3*, e1970.
55. Shiozawa, Y.; Malcovati, L.; Gallì, A.; Sato-Otsubo, A.;

Kataoka, K.; Sato, Y.; Watatani, Y.; Suzuki, H.; Yoshizato, T.; Yoshida, K.; et al. Aberrant splicing and defective mRNA production induced by somatic spliceosome mutations in myelodysplasia. *Nat. Commun.* **2018**, *9*, 3649, doi:10.1038/s41467-018-06063-x.

56. Dolatshad, H.; Pellagatti, A.; Liberante, F.G.; Llorian, M.; Repapi, E.; Steeples, V.; Roy, S.; Scifo, L.; Armstrong, R.N.; Shaw, J.; et al. Cryptic splicing events in the iron transporter ABCB7 and other key target genes in SF3B1-mutant myelodysplastic syndromes. *Leukemia* **2016**, *30*, 2322–2331, doi:10.1038/leu.2016.149.
57. Lieu, Y.K.; Liu, Z.; Ali, A.M.; Wei, X.; Penson, A.; Zhang, J. SF3B1 mutant-induced missplicing of MAP3K7 causes anemia in myelodysplastic syndromes. **2020**.
58. Malcovati, L.; Stevenson, K.; Papaemmanuil, E.; Neuberg, D.; Bejar, R.; Boultwood, J.; Bowen, D.T.; Campbell, P.J.; Ebert, B.L.; Fenaux, P.; et al. SF3B1-mutant MDS as a distinct disease subtype: a proposal from the International Working Group for the Prognosis of MDS. *Blood* **2020**, *136*, 157–170, doi:10.1182/blood.2020004850.
59. Lee, S.C.-W.; Abdel-Wahab, O. Therapeutic targeting of splicing in cancer. *Nat. Med.* **2016**, *22*, 976–986, doi:10.1038/nm.4165.
60. Desterro, J.; Bak-Gordon, P.; Carmo-Fonseca, M. Targeting mRNA processing as an anticancer strategy. *Nat. Rev. Drug Discov.* **2019**, doi:10.1038/s41573-019-0042-3.
61. Webb, T.R.; Joyner, A.S.; Potter, P.M. The development and application of small molecule modulators of SF3b as therapeutic agents for cancer. *Drug Discov. Today* **2013**, *18*, 43–9, doi:10.1016/j.drudis.2012.07.013.
62. Bonnal, S.; Vigevani, L.; Valcárcel, J. The spliceosome as a target of novel antitumour drugs. *Nat. Rev. Drug Discov.* **2012**, *11*, 847–859, doi:10.1038/nrd3823.
63. Seiler, M.; Yoshimi, A.; Darman, R.; Chan, B.; Keaney, G.; Thomas, M.; Agrawal, A.A.; Caleb, B.; Csibi, A.; Sean, E.; et al. H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. *Nat. Med.* **2018**, *24*, 497–504, doi:10.1038/nm.4493.
64. Cretu, C.; Agrawal, A.A.; Cook, A.; Will, C.L.; Fekkes, P.; Peter, G.; Lührmann, R.; Larsen, N.; Buonamici, S.; Pena, V. Structural basis of splicing modulation by antitumor

macrolide compounds. *Mol. Cell* **2018**, 1–29, doi:10.1016/j.molcel.2018.03.011.
65. Finci, L.I.; Zhang, X.; Huang, X.; Zhou, Q.; Tsai, J.; Teng, T.; Agrawal, A.; Chan, B.; Irwin, S.; Karr, C.; et al. The cryo-EM structure of the SF3b spliceosome complex bound to a splicing modulator reveals a pre-mRNA substrate competitive mechanism of action. *Genes Dev.* **2018**, *32*, 309–320, doi:10.1101/gad.311043.117.
66. Spinello, A.; Borišek, J.; Malcovati, L. Investigating the Molecular Mechanism of H3B-8800 : A Splicing Modulator Inducing Preferential Lethality in Spliceosome-Mutant Cancers. **2021**.
67. Teng, T.; Tsai, J.H.; Puyang, X.; Seiler, M.; Peng, S.; Prajapati, S.; Aird, D.; Buonamici, S.; Caleb, B.; Chan, B.; et al. Splicing modulators act at the branch point adenosine binding pocket defined by the PHF5A–SF3b complex. *Nat. Commun.* **2017**, *8*, 15522, doi:10.1038/ncomms15522.
68. Sakai, Y.; Tsujita, T.; Akiyama, T.; Yoshida, T.; Mizukami, T.; Akinaga, S.; Horinouchi, S.; Yoshida, M.; Yoshida, T. GEX1 compounds, novel antitumor antibiotics related to herboxidiene, produced by Streptomyces sp. II. The effects on cell cycle progression and gene expression. *J. Antibiot. (Tokyo)*. **2002**, *55*, 863–872, doi:10.7164/antibiotics.55.863.
69. Sakai, T.; Sameshima, T.; Matsufuji, M.; Kawamura, N.; Dobashi, K.; Mizui, Y. Pladienolides, new substances from culture of Streptomyces platensis Mer-11107. I. Taxonomy, fermentation, isolation and screening. *J. Antibiot. (Tokyo)*. **2004**, *57*, 173–179, doi:10.7164/antibiotics.57.173.
70. Nakajima, H.; Hori, Y.; Terano, H.; Okuhara, M.; Manda, T.; Matsumoto, S.; Shimomura, K. New antitumor substances, FR901463, FR901464 and FR901465: II. Activities against experimental tumors in mice and mechanism of action. *J. Antibiot. (Tokyo).* **1996**, *49*, 1204–1211, doi:10.7164/antibiotics.49.1204.
71. Xargay-Torrent, S.; López-Guerra, M.; Rosich, L.; Montraveta, A.; Roldán, J.; Rodríguez, V.; Villamor, N.; Aymerich, M.; Lagisetti, C.; Webb, T.R.; et al. The splicing modulator sudemycin induces a specific antitumor response and cooperates with ibrutinib in chronic lymphocytic leukemia. *Oncotarget* **2015**, *6*, 22734–22749, doi:10.18632/oncotarget.4212.

72. Iwata, M.; Ozawa, Y.; Uenaka, T.; Shimizu, H.; Niijima, J.; Kanada, R.M.; Fukuda, Y.; Nagai, M.; Kotake, Y.; Yoshida, M.; et al. E7107, a new 7-urethane derivative of pladienolide D, displays curative effect against several human tumor xenografts. *Cancer Res.* **2004**, *64*, 691 LP – 691.
73. Hong, D.S.; Kurzrock, R.; Naing, A.; Wheler, J.J.; Falchook, G.S.; Schiffman, J.S.; Faulkner, N.; Pilat, M.J.; O'Brien, J.; LoRusso, P. A phase I, open-label, single-arm, dose-escalation study of E7107, a precursor messenger ribonucleic acid (pre-mRNA) splicesome inhibitor administered intravenously on days 1 and 8 every 21 days to patients with solid tumors. *Invest. New Drugs* **2014**, *32*, 436–444, doi:10.1007/s10637-013-0046-5.
74. Eskens, F.A.L.M.; Ramos, F.J.; Burger, H.; O'Brien, J.P.; Piera, A.; De Jonge, M.J.A.; Mizui, Y.; Wiemer, E.A.C.; Carreras, M.J.; Baselga, J.; et al. Phase I pharmacokinetic and pharmacodynamic study of the first-in-class spliceosome inhibitor E7107 in patients with advanced solid tumors. *Clin. Cancer Res.* **2013**, *19*, 6296–6304.
75. Steensma, D.P.; Wermke, M.; Klimek, V.M.; Greenberg, P.L.; Font, P.; Komrokji, R.S.; Yang, J.; Brunner, A.M.; Carraway, H.E.; Ades, L.; et al. Phase I First-in-Human Dose Escalation Study of the oral SF3B1 modulator H3B-8800 in myeloid neoplasms. *Leukemia* **2021**, doi:10.1038/s41375-021-01328-9.
76. Hallek, M.; Cheson, B.D.; Catovsky, D.; Caligaris-Cappio, F.; Dighiero, G.; Döhner, H.; Hillmen, P.; Keating, M.; Montserrat, E.; Chiorazzi, N.; et al. Guidelines for diagnosis, indications for treatment, response assessment and supportive management of chronic lymphocytic leukemia. *Blood* **2018**, *131*, blood-2017-09-806398, doi:10.1182/blood-2017-09-806398.
77. Swerdlow, S.H.; Campo, E.; Pileri, S.A.; Harris, N.L.; Stein, H.; Siebert, R.; Advani, R.; Ghielmini, M.; Salles, G.A.; Zelenetz, A.D.; et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms., doi:10.1182/blood-2016.
78. Berndt, S.I.; Camp, N.J.; Skibola, C.F.; Vijai, J.; Wang, Z.; Gu, J.; Nieters, A.; Kelly, R.S.; Smedby, K.E.; Monnereau, A.; et al. Meta-analysis of genome-wide association studies discovers multiple loci for chronic lymphocytic leukemia.

Nat. Commun. **2016**, *7*, 10933, doi:10.1038/ncomms10933.
79. Speedy, H.E.; Beekman, R.; Chapaprieta, V.; Orlando, G.; Law, P.J.; Martín-García, D.; Gutiérrez-Abril, J.; Catovsky, D.; Beà, S.; Clot, G.; et al. Insight into genetic predisposition to chronic lymphocytic leukemia from integrative epigenomics. *Nat. Commun.* **2019**, *10*, doi:10.1038/s41467-019-11582-2.
80. Rossi, D.; Spina, V.; Gaidano, G. Biology and treatment of Richter syndrome. *Blood* **2018**, *131*, blood-2018-01-791376, doi:10.1182/blood-2018-01-791376.
81. Rai, K.R.; Sawitsky, A.; Cronkite, E.P.; Chanana, A.D.; Levy, R.N.; Pasternack, B.S. Clinical staging of chronic lymphocytic leukemia. *Blood* **1975**, *46*, 219–234.
82. Binet, J.L.; Auquier, A.; Dighiero, G.; Chastang, C.; Piguet, H.; Goasguen, J.; Vaugier, G.; Potron, G.; Colona, P.; Oberling, F.; et al. A new prognostic classification of chronic lymphocytic leukemia derived from a multivariate survival analysis. *Cancer* **1981**, *48*, 198–206, doi:10.1002/1097-0142(19810701)48:1<198::aid-cncr2820480131>3.0.co;2-v.
83. Eichhorst, B.; Robak, T.; Montserrat, E.; Ghia, P.; Niemann, C.U.; Kater, A.P.; Gregor, M.; Cymbalista, F.; Buske, C.; Hillmen, P.; et al. Chronic lymphocytic leukaemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **2021**, *32*, 23–33, doi:10.1016/j.annonc.2020.09.019.
84. Schroeder, H.W.; Cavacini, L. Structure and function of immunoglobulins. *J. Allergy Clin. Immunol.* **2010**, *125*, S41–S52, doi:10.1016/j.jaci.2009.09.046.
85. ten Hacken, E.; Gounari, M.; Ghia, P.; Burger, J.A. The importance of B cell receptor isotypes and stereotypes in chronic lymphocytic leukemia. *Leukemia* **2019**, *33*, 287–298, doi:10.1038/s41375-018-0303-x.
86. Burger, J.A.; O'Brien, S. Evolution of CLL treatment — from chemoimmunotherapy to targeted and individualized therapy. *Nat. Rev. Clin. Oncol.* **2018**, *15*, 510–527, doi:10.1038/s41571-018-0037-8.
87. Seifert, M.; Sellmann, L.; Bloehdorn, J.; Wein, F.; Stilgenbauer, S.; Dürig, J.; Küppers, R. Cellular origin and pathophysiology of chronic lymphocytic leukemia. *J. Exp. Med.* **2012**, *209*, 2183–2198.
88. Damle, R.N.; Wasil, T.; Fais, F.; Ghiotto, F.; Valetto, A.;

Allen, S.L.; Buchbinder, A.; Budman, D.; Dittmar, K.; Kolitz, J.; et al. Ig V Gene Mutation Status and CD38 Expression As Novel Prognostic Indicators in Chronic Lymphocytic Leukemia. *Blood* **1999**, *94*, 1840–1847.
89. Hamblin, T.J.; Davis, Z.; Gardiner, A.; Oscier, D.G.; Stevenson, F.K. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood* **1999**, *94*, 1848–1854.
90. Stamatopoulos, K.; Agathangelidis, A.; Rosenquist, R.; Ghia, P. Antigen receptor stereotypy in chronic lymphocytic leukemia. *Leukemia* **2017**, *31*, 282–291, doi:10.1038/leu.2016.322.
91. Nadeu, F.; Royo, R.; Clot, G.; Duran-Ferrer, M.; Navarro, A.; Martin, S.; Lu, J.; Zenz, T.; Baumann, T.S.; Jares, P.; et al. IGLV3-21R110 identifies an aggressive biological subtype of chronic lymphocytic leukemia with intermediate epigenetics. *Blood* **2020**, doi:10.1182/blood.2020008311.
92. Maity, P.C.; Bilal, M.; Koning, M.T.; Young, M.; Bergen, C.A.M. Van; Renna, V.; Imkeller, K.; Busse, C.E.; Degano, M.; Bakchoul, T.; et al. IGLV3-21 * 01 is an inherited risk factor for CLL through the acquisition of a single-point mutation enabling autonomous BCR signaling. **2020**, *117*, doi:10.1073/pnas.1913810117.
93. Beekman, R.; Chapaprieta, V.; Russiñol, N.; Vilarrasa-Blasi, R.; Verdaguer-Dot, N.; Martens, J.H.A.; Duran-Ferrer, M.; Kulis, M.; Serra, F.; Javierre, B.M.; et al. The reference epigenome and regulatory chromatin landscape of chronic lymphocytic leukemia. *Nat. Med.* **2018**, doi:10.1038/s41591-018-0028-4.
94. Queirós, A.C.; Villamor, N.; Clot, G.; Martinez-Trillos, A.; Kulis, M.; Navarro, A.; Penas, E.M.M.; Jayne, S.; Majid, A.; Richter, J.; et al. A B-cell epigenetic signature defines three biologic subgroups of chronic lymphocytic leukemia with clinical impact. *Leukemia* **2015**, *29*, 598–605, doi:10.1038/leu.2014.252.
95. Döhner, H.; Stilgenbauer, S.; Benner, A.; Leupolt, E.; Kröber, A.; Bullinger, L.; Döhner, K.; Bentz, M.; Lichter, P. Genomic Aberrations and Survival in Chronic Lymphocytic Leukemia. *N. Engl. J. Med.* **2000**, *343*, 1910–1916, doi:10.1056/nejm200012283432602.
96. Wang, L.; Lawrence, M.S.; Wan, Y.; Stojanov, P.; Sougnez,

C.; Stevenson, K.; Werner, L.; Sivachenko, A.; DeLuca, D.S.; Zhang, L.; et al. SF3B1 and Other Novel Cancer Genes in Chronic Lymphocytic Leukemia. *N. Engl. J. Med.* **2011**, *365*, 2497–2506, doi:10.1056/NEJMoa1109016.

97. Quesada, V.; Conde, L.; Villamor, N.; Ordóñez, G.R.; Jares, P.; Bassaganyas, L.; Ramsay, A.J.; Beà, S.; Pinyol, M.; Martínez-Trillos, A.; et al. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat. Genet.* **2011**, *44*, 47–52, doi:10.1038/ng.1032.
98. Puente, X.S.; Beà, S.; Valdés-Mas, R.; Villamor, N.; Gutiérrez-Abril, J.; Martín-Subero, J.I.; Munar, M.; Rubio-Pérez, C.; Jares, P.; Aymerich, M.; et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature* **2015**, *526*, 519–524, doi:10.1038/nature14666.
99. Landau, D.A.; Tausch, E.; Taylor-Weiner, A.N.; Stewart, C.; Reiter, J.G.; Bahlo, J.; Kluth, S.; Bozic, I.; Lawrence, M.; Böttcher, S.; et al. Mutations driving CLL and their evolution in progression and relapse. *Nature* **2015**, *526*, 525–530, doi:10.1038/nature15395.
100. Puente, X.S.; Pinyol, M.; Quesada, V.; Conde, L.; Ordóñez, G.R.; Villamor, N.; Escaramis, G.; Jares, P.; Beà, S.; González-Díaz, M.; et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature* **2011**, *475*, 101–105, doi:10.1038/nature10113.
101. Zenz, T.; Kröber, A.; Scherer, K.; Häbe, S.; Bühler, A.; Benner, A.; Denzel, T.; Winkler, D.; Edelmann, J.; Schwänen, C.; et al. Monoallelic TP53 inactivation is associated with poor prognosis in chronic lymphocytic leukemia: results from a detailed genetic characterization with long-term follow-up. *Blood* **2008**, *112*, 3322–3329, doi:10.1182/blood-2008-04-154070.
102. Shuai, S.; Suzuki, H.; Diaz-Navarro, A.; Nadeu, F.; Kumar, S.A.; Gutierrez-Fernandez, A.; Delgado, J.; Pinyol, M.; López-Otín, C.; Puente, X.S.; et al. The U1 spliceosomal RNA is recurrently mutated in multiple cancers. *Nature* **2019**, doi:10.1038/s41586-019-1651-z.
103. Martínez-Trillos, A.; Navarro, A.; Aymerich, M.; Delgado, J.; López-Guillermo, A.; Campo, E.; Villamor, N. Clinical impact of MYD88 mutations in chronic lymphocytic

leukemia. *Blood* 2016, *127*, 1611–1613.
104. Landau, D.A.; Carter, S.L.; Stojanov, P.; Mckenna, A.; Stevenson, K.; Lawrence, M.S.; Sougnez, C.; Stewart, C.; Sivachenko, A.; Wang, L.; et al. Evolution and Impact of Subclonal Mutations in Chronic Lymphocytic Leukemia. **2013**, doi:10.1016/j.cell.2013.01.019.
105. Nadeu, F.; Delgado, J.; Royo, C.; Baumann, T.; Stankovic, T.; Pinyol, M.; Jares, P.; Navarro, A.; Martín-García, D.; Bè, S.; et al. Clinical impact of clonal and subclonal TP53, SF3B1, BIRC3, NOTCH1, and ATM mutations in chronic lymphocytic leukemia. *Blood* **2016**, *127*, 2122–2130, doi:10.1182/blood-2015-07.
106. Nadeu, F.; Clot, G.; Delgado, J.; Martín-García, D.; Baumann, T.; Salaverria, I.; Beà, S.; Pinyol, M.; Jares, P.; Navarro, A.; et al. Clinical impact of the subclonal architecture and mutational complexity in chronic lymphocytic leukemia. *Leukemia* **2017**, *32*, 645–653, doi:10.1038/leu.2017.291.
107. Hoechstetter, M.A.; Busch, R.; Eichhorst, B.; Bühler, A.; Winkler, D.; Eckart, M.J.; Vehling-Kaiser, U.; Schimke, H.; Jäger, U.; Hurtz, H.J.; et al. Early, risk-adapted treatment with fludarabine in Binet stage A chronic lymphocytic leukemia patients: results of the CLL1 trial of the German CLL study group. *Leukemia* **2017**, *31*, 2833–2837, doi:10.1038/leu.2017.246.
108. Han, T.; Ezdinli, E.Z.; Shimaoka, K.; Desai, D. V Chlorambucil vs. combined chlorambucil-corticosteroid therapy in chronic lymphocytic leukemia. *Cancer* **1973**, *31*, 502–508, doi:10.1002/1097-0142(197303)31:3<502::aid-cncr2820310303>3.0.co;2-7.
109. Knospe, W.H.; Jr, V.L.; Huguley, C.M. Bi-weekly chlorambucil treatment of chronic lymphocytic leukemia. *Cancer* **1974**, *33*, 555–562, doi:10.1002/1097-0142(197402)33:2<555::AID-CNCR2820330234>3.0.CO;2-I.
110. Eichhorst, B.F.; Busch, R.; Hopfinger, G.; Pasold, R.; Hensel, M.; Steinbrecher, C.; Siehl, S.; Jäger, U.; Bergmann, M.; Stilgenbauer, S.; et al. Fludarabine plus cyclophosphamide versus fludarabine alone in first-line therapy of younger patients with chronic lymphocytic leukemia. *Blood* **2006**, *107*, 885–891, doi:10.1182/blood-

2005-06-2395.
111. Catovsky, D.; Richards, S.; Matutes, E.; Oscier, D.; Dyer, M.; Bezares, R.F.; Pettitt, A.R.; Hamblin, T.; Milligan, D.W.; Child, J.A.; et al. Assessment of fludarabine plus cyclophosphamide for patients with chronic lymphocytic leukaemia (the LRF CLL4 Trial): a randomised controlled trial. *Lancet (London, England)* **2007**, *370*, 230–239, doi:10.1016/S0140-6736(07)61125-8.
112. Flinn, I.W.; Neuberg, D.S.; Grever, M.R.; Dewald, G.W.; Bennett, J.M.; Paietta, E.M.; Hussein, M.A.; Appelbaum, F.R.; Larson, R.A.; Moore, D.F.J.; et al. Phase III trial of fludarabine plus cyclophosphamide compared with fludarabine for patients with previously untreated chronic lymphocytic leukemia: US Intergroup Trial E2997. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **2007**, *25*, 793–798, doi:10.1200/JCO.2006.08.0762.
113. Hallek, M.; Fischer, K.; Fingerle-Rowson, G.; Fink, A.M.; Busch, R.; Mayer, J.; Hensel, M.; Hopfinger, G.; Hess, G.; von Grünhagen, U.; et al. Addition of rituximab to fludarabine and cyclophosphamide in patients with chronic lymphocytic leukaemia: a randomised, open-label, phase 3 trial. *Lancet (London, England)* **2010**, *376*, 1164–1174, doi:10.1016/S0140-6736(10)61381-5.
114. Goede, V.; Fischer, K.; Busch, R.; Engelke, A.; Eichhorst, B.; Wendtner, C.M.; Chagorova, T.; de la Serna, J.; Dilhuydy, M.-S.; Illmer, T.; et al. Obinutuzumab plus Chlorambucil in Patients with CLL and Coexisting Conditions. *N. Engl. J. Med.* **2014**, *370*, 1101–1110, doi:10.1056/NEJMoa1313984.
115. Byrd, J.C.; Hillmen, P.; O'Brien, S.; Barrientos, J.C.; Reddy, N.M.; Coutre, S.; Tam, C.S.; Mulligan, S.P.; Jaeger, U.; Barr, P.M.; et al. Long-term follow-up of the RESONATE phase 3 trial of ibrutinib vs ofatumumab. *Blood* **2019**, *133*, 2031–2042, doi:10.1182/blood-2018-08-870238.
116. Byrd, J.C.; Wierda, W.G.; Schuh, A.; Devereux, S.; Chaves, J.M.; Brown, J.R.; Hillmen, P.; Martin, P.; Awan, F.T.; Stephens, D.M.; et al. Acalabrutinib monotherapy in patients with relapsed/refractory chronic lymphocytic leukemia: updated phase 2 results. *Blood* **2020**, *135*, 1204–1213, doi:10.1182/blood.2018884940.
117. Byrd, J.C.; Hillmen, P.; Ghia, P.; Kater, A.P.; Chanan-Khan,

A.; Furman, R.R.; O'Brien, S.; Yenerel, M.N.; Illés, A.; Kay, N.; et al. Acalabrutinib Versus Ibrutinib in Previously Treated Chronic Lymphocytic Leukemia: Results of the First Randomized Phase III Trial. *J. Clin. Oncol.* **2021**, *39*, 3441–3452, doi:10.1200/JCO.21.01210.

118. Sharman, J.P.; Coutre, S.E.; Furman, R.R.; Cheson, B.D.; Pagel, J.M.; Hillmen, P.; Barrientos, J.C.; Zelenetz, A.D.; Kipps, T.J.; Flinn, I.W.; et al. Second Interim Analysis of a Phase 3 Study of Idelalisib (ZYDELIG®) Plus Rituximab (R) for Relapsed Chronic Lymphocytic Leukemia (CLL): Efficacy Analysis in Patient Subpopulations with Del(17p) and Other Adverse Prognostic Factors. *Blood* **2014**, *124*, 330, doi:10.1182/blood.V124.21.330.330.
119. Flinn, I.W.; Hillmen, P.; Montillo, M.; Nagy, Z.; Illés, Á.; Etienne, G.; Delgado, J.; Kuss, B.J.; Tam, C.S.; Gasztonyi, Z.; et al. The phase 3 DUO trial: duvelisib versus ofatumumab in relapsed and refractory CLL/SLL. *Blood* **2018**, *132*, blood-2018-05-850461, doi:10.1182/blood-2018-05-850461.
120. Kater, A.P.; Seymour, J.F.; Hillmen, P.; Eichhorst, B.; Langerak, A.W.; Owen, C.; Verdugo, M.; Wu, J.; Punnoose, E.A.; Jiang, Y.; et al. Fixed Duration of Venetoclax-Rituximab in Relapsed/Refractory Chronic Lymphocytic Leukemia Eradicates Minimal Residual Disease and Prolongs Survival: Post-Treatment Follow-Up of the MURANO Phase III Study. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **2019**, *37*, 269–277, doi:10.1200/JCO.18.01580.
121. Yin, S.; Gambe, R.G.; Sun, J.; Carrasco, R.D.; Wu, C.J.; Wang, L.; Yin, S.; Gambe, R.G.; Sun, J.; Martinez, A.Z.; et al. A Murine Model of Chronic Lymphocytic Leukemia Based on B Cell-Restricted Expression of Sf3b1 Mutation and Atm Deletion. *Cancer Cell* **2019**, 1–14, doi:10.1016/j.ccell.2018.12.013.
122. Rossi, D.; Bruscaggin, A.; Spina, V.; Rasi, S.; Khiabanian, H.; Messina, M.; Fangazio, M.; Vaisitti, T.; Monti, S.; Chiaretti, S.; et al. Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. *Blood* **2011**, *118*, 6904–6908, doi:10.1182/blood-2011-08-373159.
123. Jeromin, S.; Weissmann, S.; Haferlach, C.; Dicker, F.; Bayer, K.; Grossmann, V.; Alpermann, T.; Roller, A.; Kohlmann,

A.; Haferlach, T.; et al. SF3B1 mutations correlated to cytogenetics and mutations in NOTCH1, FBXW7, MYD88, XPO1 and TP53 in 1160 untreated CLL patients. *Leukemia* **2014**, *28*, 108–117, doi:10.1038/leu.2013.263.

124. Baliakas, P.; Hadzidimitriou, A.; Sutton, L.A.; Rossi, D.; Minga, E.; Villamor, N.; Larrayoz, M.; Kminkova, J.; Agathangelidis, A.; Davis, Z.; et al. Recurrent mutations refine prognosis in chronic lymphocytic leukemia. *Leukemia* **2015**, *29*, 329–336, doi:10.1038/leu.2014.196.
125. Rossi, D.; Rasi, S.; Spina, V.; Bruscaggin, A.; Monti, S.; Ciardullo, C.; Deambrogi, C.; Khiabanian, H.; Serra, R.; Bertoni, F.; et al. Integrated mutational and cytogenetic analysis identifies new prognostic subgroups in chronic lymphocytic leukemia. *Blood* **2013**, *21*, doi:10.1182/blood.
126. Stilgenbauer, S.; Schnaiter, A.; Paschka, P.; Zenz, T.; Rossi, M.; Döhner, K.; Bühler, A.; Böttcher, S.; Ritgen, M.; Kneba, M.; et al. Gene mutations and treatment outcome in chronic lymphocytic leukemia: Results from the CLL8 trial. *Blood* **2014**, *123*, 3247–3254.
127. Takahashi, K.; Hu, B.; Wang, F.; Yan, Y.; Kim, E.; Vitale, C.; Patel, K.P.; Strati, P.; Gumbs, C.; Little, L.; et al. Clinical implications of cancer gene mutations in patients with chronic lymphocytic leukemia treated with lenalidomide. **2018**, *131*, doi:10.1182/blood-2017-11-817296.
128. Ferreira, P.G.; Jares, P.; Rico, D.; Gómez-López, G.; Martínez-Trillos, A.; Villamor, N.; Ecker, S.; González-Pérez, A.; Knowles, D.G.; Monlong, J.; et al. Transcriptome characterization by RNA sequencing identifies a major molecular and clinical subdivision in chronic lymphocytic leukemia. *Genome Res.* **2014**, *24*, 212–226, doi:10.1101/gr.152132.112.
129. Wang, L.; Brooks, A.N.; Fan, J.; Wan, Y.; Gambe, R.; Li, S.; Hergert, S.; Yin, S.; Freeman, S.S.; Levin, J.Z.; et al. Transcriptomic Characterization of SF3B1 Mutation Reveals Its Pleiotropic Effects in Chronic Lymphocytic Leukemia. *Cancer Cell* **2016**, *30*, 750–763, doi:10.1016/j.ccell.2016.10.005.
130. Lütge, A.; Lu, J.; Hüllein, J.; Walther, T.; Sellner, L.; Wu, B.; Oakes, C.C.; Dietrich, S.; Huber, W.; Zenz, T. Subgroup-specific gene expression profiles and mixed epistasis in chronic lymphocytic leukemia. **2021**, 0–21.

131. Richardson, C.D.; Ray, G.J.; Dewitt, M.A.; Curie, G.L.; Corn, J.E. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. **2016**, *34*, doi:10.1038/nbt.3481.
132. Paquet, D.; Kwart, D.; Chen, A.; Sproul, A.; Jacob, S.; Teo, S.; Olsen, K.M.; Gregg, A.; Noggle, S.; Tessier-lavigne, M.; et al. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* **2016**, *533*, 125–129, doi:10.1038/nature17664.
133. Renaud, J.; Boix, C.; Charpentier, M.; De Cian, A.; Cochennec, J.; Duvernois-Berthet, E.; Perrouault, L.; Tesson, L.; Edouard, J.; Thinard, R.; et al. Improved Genome Editing Efficiency and Flexibility Using Modified Oligonucleotides with TALEN and CRISPR-Cas9 Nucleases. *Cell Rep.* **2016**, *14*, 2263–2272, doi:10.1016/j.celrep.2016.02.018.
134. Pinder, J.; Salsman, J.; Dellaire, G. Nuclear domain ' knock-in ' screen for the evaluation and identification of small molecule enhancers of CRISPR-based genome editing. **2015**, *43*, 9379–9392, doi:10.1093/nar/gkv993.
135. Irimia, M.; Weatheritt, R.J.; Ellis, J.D.; Parikshak, N.N.; Gonatopoulos-Pournatzis, T.; Babor, M.; Quesnel-Vallières, M.; Tapial, J.; Raj, B.; O'Hanlon, D.; et al. A highly conserved program of neuronal microexons is misregulated in autistic brains. *Cell* **2014**, *159*, 1511–1523, doi:10.1016/j.cell.2014.11.035.
136. Tapial, J.; Ha, K.C.H.; Sterne-Weiler, T.; Gohr, A.; Braunschweig, U.; Hermoso-Pulido, A.; Quesnel-Vallières, M.; Permanyer, J.; Sodaei, R.; Marquez, Y.; et al. An atlas of alternative splicing profiles and functional associations reveals new regulatory programs and genes that simultaneously express multiple major isoforms. *Genome Res.* **2017**, *27*, 1759–1768, doi:10.1101/gr.220962.117.
137. Dvinge, H.; Ries, R.E.; Ilagan, J.O.; Stirewalt, D.L.; Meshinchi, S.; Bradley, R.K. Sample processing obscures cancer-specific alterations in leukemic transcriptomes. *Proc. Natl. Acad. Sci.* **2014**, *111*, 16802–16807, doi:10.1073/pnas.1413374111.
138. DeBoever, C.; Ghia, E.M.; Shepard, P.J.; Rassenti, L.; Barrett, C.L.; Jepsen, K.; Jamieson, C.H.M.; Carson, D.; Kipps, T.J.; Frazer, K.A. Transcriptome Sequencing Reveals Potential Mechanism of Cryptic 3' Splice Site Selection in

SF3B1-mutated Cancers. *PLOS Comput. Biol.* **2015**, *11*, e1004105, doi:10.1371/journal.pcbi.1004105.
139. Ewels, P.; Magnusson, M.; Lundin, S.; Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **2016**, *32*, 3047–3048, doi:10.1093/bioinformatics/btw354.
140. Gohr, A.; Irimia, M. Matt: Unix tools for alternative splicing analysis. *Bioinformatics* **2019**, *35*, 130–132, doi:10.1093/bioinformatics/bty606.
141. Yeo, G.; Burge, C.B. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J. Comput. Biol.* **2004**, *11*, 377–394, doi:10.1089/1066527041410418.
142. Corvelo, A.; Hallegger, M.; Smith, C.W.J.; Eyras, E. Genome-wide association between branch point properties and alternative splicing. *PLoS Comput. Biol.* **2010**, *6*, 12–15, doi:10.1371/journal.pcbi.1001016.
143. Zillmann, M.; Zapp, M.L.; Berget, S.M. Gel electrophoretic isolation of splicing complexes containing U1 small nuclear ribonucleoprotein particles. *Mol. Cell. Biol.* **1988**, *8*, 814–821, doi:10.1128/mcb.8.2.814-821.1988.
144. Li, Z.; Zhao, B.; Shi, Y.; Liang, Y.; Qian, R.; Wan, Y. Characterization of the aberrant splicing of MAP3K7 induced by cancer-associated SF3B1 mutation. *J. Biochem.* **2021**, doi:10.1093/jb/mvab023.
145. Mercer, T.R.; Clark, M.B.; Andersen, S.B.; Brunck, M.E.; Haerty, W.; Crawford, J.; Taft, R.J.; Nielsen, L.K.; Dinger, M.E.; Mattick, J.S. Genome-wide discovery of human splicing branchpoints. *Genome Res.* **2015**, *25*, 290–303, doi:10.1101/gr.182899.114.
146. Pacholewska, A.; Grimm, C.; Herling, C.D.; Lienhard, M.; Königs, A.; Timmermann, B.; Altmüller, J.; Mücke, O.; Reinhardt, H.C.; Plass, C.; et al. Altered DNA methylation profiles in SF3B1 mutated CLL patients. *Int. J. Mol. Sci.* **2021**, *22*, doi:10.3390/ijms22179337.
147. Agrawal, A.A.; Seiler, M.; Brinton, L.T.; Mantel, R.; Lapalombella, R.; Woyach, J.A.; Johnson, A.J.; Zhu, P.; Warmuth, M.; Yu, L.; et al. Novel SF3B1 in-frame deletions result in aberrant RNA splicing in CLL patients. *Blood Adv.* **2017**, *1*, 995–1000, doi:10.1182/bloodadvances.2017007062.
148. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of

fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550, doi:10.1186/s13059-014-0550-8.
149. Ge, S.X.; Son, E.W.; Yao, R. iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinformatics* **2018**, *19*, 534, doi:10.1186/s12859-018-2486-6.
150. Sato, S.; Sanjo, H.; Takeda, K.; Ninomiya-Tsuji, J.; Yamamoto, M.; Kawai, T.; Matsumoto, K.; Takeuchi, O.; Akira, S. Essential function for the kinase TAK1 in innate and adaptive immune responses. *Nat. Immunol.* **2005**, *6*, 1087–1095, doi:10.1038/ni1255.
151. Lee, S.C.W.; North, K.; Kim, E.; Jang, E.; Obeng, E.; Lu, S.X.; Liu, B.; Inoue, D.; Yoshimi, A.; Ki, M.; et al. Synthetic Lethal and Convergent Biological Effects of Cancer-Associated Spliceosomal Gene Mutations. *Cancer Cell* **2018**, 225–241, doi:10.1016/j.ccell.2018.07.003.
152. Liu, B.; Liu, Z.; Chen, S.; Ki, M.; Erickson, C.; Reis-Filho, J.S.; Durham, B.H.; Chang, Q.; de Stanchina, E.; Sun, Y.; et al. Mutant SF3B1 promotes AKT and NF-kB driven mammary tumorigenesis. *J. Clin. Invest.* **2020**, doi:10.1172/JCI138315.
153. de Thonel, A.; Vandekerckhove, J.; Lanneau, D.; Selvakumar, S.; Courtois, G.; Hazoume, A.; Brunet, M.; Maurel, S.; Hammann, A.; Ribeil, J.A.; et al. HSP27 controls GATA-1 protein level during erythroid cell differentiation. *Blood* **2010**, *116*, 85–96, doi:10.1182/blood-2009-09-241778.
154. Steensma, D.P.; Wermke, M.; Klimek, V.M.; Greenberg, P.L.; Font, P.; Komrokji, R.S.; Yang, J.; Brunner, A.M.; Carraway, H.E.; Ades, L.; et al. Results of a Clinical Trial of H3B-8800, a Splicing Modulator, in Patients with Myelodysplastic Syndromes (MDS), Acute Myeloid Leukemia (AML) or Chronic Myelomonocytic Leukemia (CMML). *Blood* **2019**, *134*, 673, doi:10.1182/blood-2019-123854.
155. Patten, P.E.M.; Buggins, A.G.S.; Richards, J.; Wotherspoon, A.; Salisbury, J.; Mufti, G.J.; Hamblin, T.J.; Devereux, S. CD38 expression in chronic lymphocytic leukemia is regulated by the tumor microenvironment. *Blood* **2008**, *111*, 5173–5181, doi:10.1182/blood-2007-08-108605.

156. Bürkle, A.; Niedermeier, M.; Schmitt-Gräff, A.; Wierda, W.G.; Keating, M.J.; Burger, J.A. Overexpression of the CXCR5 chemokine receptor, and its ligand, CXCL13 in B-cell chronic lymphocytic leukemia. *Blood* **2007**, *110*, 3316–3325, doi:10.1182/blood-2007-05-089409.
157. Burger, J.A.; Wiestner, A. Targeting B cell receptor signalling in cancer: Preclinical and clinical advances. *Nat. Rev. Cancer* **2018**, *18*, 148–167, doi:10.1038/nrc.2017.121.
158. Shain, K.H.; Tao, J. The B-cell receptor orchestrates environment-mediated lymphoma survival and drug resistance in B-cell malignancies. *Oncogene* **2014**, *33*, 4107–4113, doi:10.1038/onc.2013.379.
159. Bae, J.; Leo, C.P.; Hsu, S.Y.; Hsueh, A.J. MCL-1S, a splicing variant of the antiapoptotic BCL-2 family member MCL-1, encodes a proapoptotic protein possessing only the BH3 domain. *J. Biol. Chem.* **2000**, *275*, 25255–25261, doi:10.1074/jbc.M909826199.
160. Aird, D.; Teng, T.; Huang, C.; Pazolli, E.; Banka, D.; Cheung-ong, K.; Eifert, C.; Furman, C.; Wu, Z.J.; Seiler, M.; et al. Sensitivity to splicing modulation of BCL2 family genes defines cancer therapeutic strategies for splicing modulators. *Nat. Commun.* **2019**, doi:10.1038/s41467-018-08150-5.
161. Ten Hacken, E.; Valentin, R.; Regis, F.F.D.; Sun, J.; Yin, S.; Werner, L.; Deng, J.; Gruber, M.; Wong, J.; Zheng, M.; et al. Splicing modulation sensitizes chronic lymphocytic leukemia cells to venetoclax by remodeling mitochondrial apoptotic dependencies. *JCI insight* **2018**, *3*, doi:10.1172/jci.insight.121438.
162. Ianevski, A.; Giri, A.K.; Aittokallio, T. SynergyFinder 2.0: Visual analytics of multi-drug combination synergies. *Nucleic Acids Res.* **2021**, *48*, W488–W493, doi:10.1093/NAR/GKAA216.
163. Tushar Murthy, Kiran V. Paul, Alexander C. Minella, and M.M.P. The Development and Use of Scalable Systems for Studying Aberrant Splicing in SF3B1-Mutant CLL. *Methods Mol. Biol.* **2018**, *1881*, 1–8, doi:10.1002/9781444300857.ch36.
164. Dalton, W.B.; Helmenstine, E.; Walsh, N.; Gondek, L.P.; Kelkar, D.S.; Read, A.; Natrajan, R.; Christenson, E.S.; Roman, B.; Das, S.; et al. Hotspot SF3B1 mutations induce

metabolic reprogramming and vulnerability to serine deprivation. *J. Clin. Invest.* **2019**, doi:10.1172/jci125022.
165. Tausch, E.; Schneider, C.; Robrecht, S.; Zhang, C.; Dolnik, A.; Bloehdorn, J.; Bahlo, J.; Al-Sawaf, O.; Ritgen, M.; Fink, A.M.; et al. Prognostic and predictive impact of genetic markers in patients with CLL treated with obinutuzumab and venetoclax. *Blood* **2020**, doi:10.1182/blood.2019004492.
166. Inoue, D.; Chew, G.-L.; Liu, B.; Michel, B.C.; Pangallo, J.; D'Avino, A.R.; Hitchman, T.; North, K.; Lee, S.C.-W.; Bitner, L.; et al. Spliceosomal disruption of the non-canonical BAF complex in cancer. *Nature* **2019**, *574*, 432–436, doi:10.1038/s41586-019-1646-9.
167. Pangallo, J.; Kiladjian, J.-J.; Cassinat, B.; Renneville, A.; Taylor, J.; Polaski, J.T.; North, K.D.; Abdel-Wahab, O.; Bradley, R.K. Rare and private spliceosomal gene mutations drive partial, complete, and dual phenocopies of hotspot alterations. *Blood* **2020**, doi:10.1182/blood.2019002894.
168. Sundaramoorthy, S.; Vázquez-novelle, M.D.; Lekomtsev, S.; Howell, M.; Petronczki, M. Functional genomics identifies a requirement of pre-mRNA splicing factors for sister chromatid cohesion. **2014**, *33*, 2623–2642.
169. Kfir, N.; Lev-Maor, G.; Glaich, O.; Alajem, A.; Datta, A.; Sze, S.K.; Meshorer, E.; Ast, G. SF3B1 Association with Chromatin Determines Splicing Outcomes. *Cell Rep.* **2015**, *11*, 618–629, doi:10.1016/J.CELREP.2015.03.048.
170. Gentien, D.; Kosmider, O.; Nguyen-Khac, F.; Albaud, B.; Rapinat, A.; Dumont, A.G.; Damm, F.; Popova, T.; Marais, R.; Fontenay, M.; et al. A common alternative splicing signature is associated with SF3B1 mutations in malignancies from different cell lineages. *Leukemia* **2014**, *28*, 1355–1357, doi:10.1038/leu.2014.28.
171. Carrocci, T.J.; Zoerner, D.M.; Paulson, J.C.; Hoskins, A.A. SF3b1 mutations associated with myelodysplastic syndromes alter the fidelity of branchsite selection in yeast. *Nucleic Acids Res.* **2017**, *45*, 4837–4852, doi:10.1093/nar/gkw1349.
172. Tang, Q.; Rodriguez-Santiago, S.; Wang, J.; Pu, J.; Yuste, A.; Gupta, V.; Moldón, A.; Xu, Y.Z.; Query, C.C. SF3B1/Hsh155 HEAT motif mutations affect interaction with the spliceosomal ATPase Prp5, resulting in altered branch site selectivity in pre-mRNA splicing. *Genes Dev.* **2016**, *30*, 2710–2723, doi:10.1101/gad.291872.116.

173. Zeng, Y.; Zeng, H.; Fair, B.J.; Krishnanomhan, A.; Hou, Y.; Ruthenburg, A.J.; Li, Y.I.; Staley, J.P. Profiling of Nascent Lariat Intermediates Reveals Key Genetic Determinants of Human Co-Transcriptional Splicing. **2021**, 1–79.
174. Payne, A.; Holmes, N.; Rakyan, V.; Loose, M. Whale watching with BulkVis: A graphical viewer for Oxford Nanopore bulk fast5 files. *bioRxiv* **2018**, 312256, doi:10.1101/312256.
175. Tang, A.D.; Soulette, C.M.; van Baren, M.J.; Hart, K.; Hrabeta-Robinson, E.; Wu, C.J.; Brooks, A.N. Full-length transcript characterization of SF3B1 mutation in chronic lymphocytic leukemia reveals downregulation of retained introns. *Nat. Commun.* **2020**, *11*, 1438, doi:10.1038/s41467-020-15171-6.
176. Herishanu, Y.; Pérez-Galán, P.; Liu, D.; Biancotto, A.; Pittaluga, S.; Vire, B.; Gibellini, F.; Njuguna, N.; Lee, E.; Stennett, L.; et al. The lymph node microenvironment promotes B-cell receptor signaling, NF-κB activation, and tumor proliferation in chronic lymphocytic leukemia. *Blood* **2011**, *117*, 563–574, doi:10.1182/blood-2010-05-284984.
177. Martínez-Limón, A.; Joaquin, M.; Caballero, M.; Posas, F.; de Nadal, E. The p38 pathway: From biology to cancer therapy. *Int. J. Mol. Sci.* **2020**, *21*, 1–18, doi:10.3390/ijms21061913.
178. Liu, Z.; Yoshimi, A.; Wang, J.; Cho, H.; Chun-Wei Lee, S.; Ki, M.; Bitner, L.; Chu, T.; Shah, H.; Liu, B.; et al. Mutations in the RNA Splicing Factor SF3B1 Promote Tumorigenesis through MYC Stabilization. *Cancer Discov.* **2020**, *10*, 806–821, doi:10.1158/2159-8290.CD-19-1330.
179. Collu, G.M.; Hidalgo-Sastre, A.; Acar, A.; Bayston, L.; Gildea, C.; Leverentz, M.K.; Mills, C.G.; Owens, T.W.; Meurette, O.; Dorey, K.; et al. Dishevelled limits Notch signalling through inhibition of CSL. *Development* **2012**, *139*, 4405–4415, doi:10.1242/dev.081885.
180. Pozzo, F.; Bittolo, T.; Tissino, E.; Vit, F.; Vendramini, E.; Laurenti, L.; D'Arena, G.; Olivieri, J.; Pozzato, G.; Zaja, F.; et al. *SF3B1*-mutated chronic lymphocytic leukemia shows evidence of NOTCH1 pathway activation including CD20 downregulation. *Haematologica* **2020**, 0–0, doi:10.3324/haematol.2020.261891.
181. Arruga, F.; Vaisitti, T.; Deaglio, S. The NOTCH Pathway

and Its Mutations in Mature B Cell Malignancies. *Front. Oncol.* **2018**, *8*, 550.
182. Arruga, F.; Gizdic, B.; Serra, S.; Vaisitti, T.; Ciardullo, C.; Coscia, M.; Laurenti, L.; D'Arena, G.; Jaksic, O.; Inghirami, G.; et al. Functional impact of NOTCH1 mutations in chronic lymphocytic leukemia. *Leukemia* **2014**, *28*, 1060–1070.
183. Villamor, N.; Conde, L.; Martínez-Trillos, A.; Cazorla, M.; Navarro, A.; Beà, S.; López, C.; Colomer, D.; Pinyol, M.; Aymerich, M.; et al. NOTCH1 mutations identify a genetic subgroup of chronic lymphocytic leukemia patients with high risk of transformation and poor outcome. *Leukemia* **2013**, *27*, 1100–1106, doi:10.1038/leu.2012.357.
184. McFarland, J.M.; Ho, Z. V.; Kugener, G.; Dempster, J.M.; Montgomery, P.G.; Bryan, J.G.; Krill-Burger, J.M.; Green, T.M.; Vazquez, F.; Boehm, J.S.; et al. Improved estimation of cancer dependencies from large-scale RNAi screens using model-based normalization and data integration. *Nat. Commun.* **2018**, *9*, 1–13, doi:10.1038/s41467-018-06916-5.
185. Byrd, J.C.; Furman, R.R.; Coutre, S.E.; Flinn, I.W.; Burger, J.A.; Blum, K.A.; Grant, B.; Sharman, J.P.; Coleman, M.; Wierda, W.G.; et al. Targeting BTK with Ibrutinib in Relapsed Chronic Lymphocytic Leukemia. *N. Engl. J. Med.* **2013**, *369*, 32–42, doi:10.1056/NEJMoa1215637.
186. Roberts, A.W.; Davids, M.S.; Pagel, J.M.; Kahl, B.S.; Puvvada, S.D.; Gerecitano, J.F.; Kipps, T.J.; Anderson, M.A.; Brown, J.R.; Gressick, L.; et al. Targeting BCL2 with Venetoclax in Relapsed Chronic Lymphocytic Leukemia. *N. Engl. J. Med.* **2015**, *374*, 311–322, doi:10.1056/NEJMoa1513257.
187. Fischer, K.; Al-Sawaf, O.; Bahlo, J.; Fink, A.-M.; Tandon, M.; Dixon, M.; Robrecht, S.; Warburton, S.; Humphrey, K.; Samoylova, O.; et al. Venetoclax and Obinutuzumab in Patients with CLL and Coexisting Conditions. *N. Engl. J. Med.* **2019**, *380*, 2225–2236, doi:10.1056/NEJMoa1815281.
188. Eugen Tausch; William Close; Anna Dolnik; Johannes Bloehdorn; Brenda Chyla; Lars Bullinger; Hartmut Döhner; Daniel Mertens; Stephan Stilgenbauer Venetoclax resistance and acquired BCL2 mutations in chronic lymphocytic leukemia. *Haematologica* **2019**, *104*, e434–e437, doi:10.3324/haematol.2019.222588.
189. Blombery, P.; Anderson, M.A.; Gong, J.N.; Thijssen, R.;

Birkinshaw, R.W.; Thompson, E.R.; Teh, C.E.; Nguyen, T.; Xu, Z.; Flensburg, C.; et al. Acquisition of the recurrent Gly101Val mutation in BCL2 confers resistance to venetoclax in patients with progressive chronic lymphocytic leukemia. *Cancer Discov.* **2019**, *9*, 342–353, doi:10.1158/2159-8290.CD-18-1119.
190. Manuscript, A. NIH Public Access. **2014**, *54*, 1364–1366, doi:10.3109/10428194.2012.742528.The.
191. Thijssen, R.; Diepstraten, S.T.; Moujalled, D.M.; Chew, E.; Flensburg, C.; Shi, M.X.; Dengler, M.A.; Litalien, V.; MacRaild, S.; Chen, M.; et al. Intact TP53 function is essential for sustaining durable responses to BH3-mimetic drugs in leukemias. *Blood* **2021**.
192. Lin, V.S.; Lew, T.E.; Handunnetti, S.M.; Blombery, P.; Nguyen, T.; Westerman, D.A.; Kuss, B.J.; Tam, C.S.; Roberts, A.W.; Seymour, J.F.; et al. BTK inhibitor therapy is effective in patients with CLL resistant to venetoclax. *Blood* **2020**, *135*, 2266–2270, doi:10.1182/blood.2020004782.
193. Tausch, E.; Beck, P.; Schlenk, R.F.; Jebaraj, B.M.C.J.; Dolnik, A.; Yosifov, D.Y.; Hillmen, P.; Offner, F.; Janssens, A.; Govind Babu, K.; et al. Prognostic and predictive role of gene mutations in chronic lymphocytic leukemia: results from the pivotal phase III study COMPLEMENT1. *Haematologica* **2020**, *105*, 2440–2447.
194. Roberts, A.W.; Ma, S.; Kipps, T.J.; Coutre, S.E.; Davids, M.S.; Eichhorst, B.; Hallek, M.; Byrd, J.C.; Humphrey, K.; Zhou, L.; et al. Efficacy of venetoclax in relapsed chronic lymphocytic leukemia is influenced by disease and response variables. *Blood* **2019**, *134*, 111–122.
195. Gángó, A.; Alpár, D.; Galik, B.; Marosvári, D.; Kiss, R.; Késüs, V.; Aczél, D.; Eyüpoglu, E.; Nagy, N.; Nagy, Á.; et al. Dissection of subclonal evolution by temporal mutation profiling in chronic lymphocytic leukemia patients treated with ibrutinib. *Int. J. Cancer* **2020**, *146*, 85–93.
196. Kanagal-Shamanna, R.; Jain, P.; Patel, K.P.; Routbort, M.; Bueso-Ramos, C.; Alhalouli, T.; Khoury, J.D.; Luthra, R.; Ferrajoli, A.; Keating, M.; et al. Targeted multigene deep sequencing of Bruton tyrosine kinase inhibitor–resistant chronic lymphocytic leukemia with disease progression and Richter transformation. *Cancer* **2019**, *125*, 559–574.
197. Larrayoz, M.; Blakemore, S.J.; Dobson, R.C.; Blunt, M.D.;

Rose-Zerilli, M.J.J.; Walewska, R.; Duncombe, A.; Oscier, D.; Koide, K.; Forconi, F.; et al. The SF3B1 inhibitor spliceostatin A (SSA) elicits apoptosis in chronic lymphocytic leukaemia cells through downregulation of Mcl-1. *Leukemia* **2016**, *30*, 351–360.
198. Kashyap, M.K.; Kumar, D.; Villa, R.; La Clair, J.J.; Benner, C.; Sasik, R.; Jones, H.; Ghia, E.M.; Rassenti, L.Z.; Kipps, T.J.; et al. Targeting the spliceosome in chronic lymphocytic leukemia with the macrolides FD-895 and pladienolide-B. *Haematologica* **2015**, *100*, 945, doi:10.3324/haematol.2014.122069.
199. Zhou, Q.; Derti, A.; Ruddy, D.; Rakiec, D.; Kao, I.; Lira, M.; Gibaja, V.; Chan, H.M.; Yang, Y.; Min, J.; et al. A chemical genetics approach for the functional assessment of novel cancer genes. *Cancer Res.* **2015**, *75*, 1949–1958, doi:10.1158/0008-5472.CAN-14-2930.
200. Lee, S.C.-W.; Dvinge, H.; Kim, E.; Cho, H.; Micol, J.-B.; Chung, Y.R.; Durham, B.H.; Yoshimi, A.; Kim, Y.J.; Thomas, M.; et al. Modulation of splicing catalysis for therapeutic targeting of leukemia with mutations in genes encoding spliceosomal proteins. *Nat. Med.* **2016**, *22*, 672–678, doi:10.1038/nm.4097.
201. Fei, D.L.; Motowski, H.; Chatrikhi, R.; Prasad, S.; Yu, J.; Gao, S.; Kielkopf, C.L.; Bradley, R.K.; Varmus, H. Wild-Type U2AF1 Antagonizes the Splicing Program Characteristic of U2AF1-Mutant Tumors and Is Required for Cell Survival. *PLoS Genet.* **2016**, *12*, e1006384, doi:10.1371/journal.pgen.1006384.
202. Shirai, C.L.; White, B.S.; Tripathi, M.; Tapia, R.; Ley, J.N.; Ndonwi, M.; Kim, S.; Shao, J.; Carver, A.; Saez, B.; et al. Mutant U2AF1-expressing cells are sensitive to pharmacological modulation of the spliceosome. *Nat. Commun.* **2017**, *8*, doi:10.1038/ncomms14060.
203. Obeng, E.A.; Chappell, R.J.; Seiler, M.; Chen, M.C.; Campagna, D.R.; Schmidt, P.J.; Schneider, R.K.; Lord, A.M.; Wang, L.; Gambe, R.G.; et al. Physiologic Expression of Sf3b1 K700E Causes Impaired Erythropoiesis, Aberrant Splicing, and Sensitivity to Therapeutic Spliceosome Modulation. *Cancer Cell* **2016**, *30*, 404–417, doi:10.1016/j.ccell.2016.08.006.
204. Vigevani, L.; Gohr, A.; Webb, T.; Irimia, M.; Valcárcel, J.

Molecular basis of differential 3′ splice site sensitivity to anti-tumor drugs targeting U2 snRNP. *Nat. Commun.* **2017**, *8*, 2100, doi:10.1038/s41467-017-02007-z.
205. Zhou, Y.; Han, C.; Wang, E.; Lorch, A.H.; Serafin, V.; Cho, B.K.; Gutierrez Diaz, B.T.; Calvo, J.; Fang, C.; Khodadadi-Jamayran, A.; et al. Posttranslational regulation of the exon skipping machinery controls aberrant splicing in leukemia. *Cancer Discov.* **2020**, *10*, 1388–1410, doi:10.1158/2159-8290.CD-19-1436.
206. Kim, H.S.; Zhang, X.; Choi, Y.S. Activation and proliferation of follicular dendritic cell-like cells by activated T lymphocytes. *J. Immunol.* **1994**, *153*, 2951–2961.
207. Quesada, V.; Ramsay, A.J.; Rodríguez, D.; Puente, X.S.; Campo, E.; López-Otín, C. The genomic landscape of chronic lymphocytic leukaemia: Biological and clinical implications. *BMC Med.* **2013**, *11*, doi:10.1111/bjh.13254.
208. Dobin, A.; Davis, C.A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T.R. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **2013**, *29*, 15–21, doi:10.1093/bioinformatics/bts635.
209. Sakamoto, H.; Inoue, K.; Higuchi, I.; Ono, Y.; Shimura, Y. Control of Drosophila Sex-lethal pre-mRNA splicing by its own female-specific product. *Nucleic Acids Res.* **1992**, *20*, 5533–5540, doi:10.1093/nar/20.21.5533.

ANNEX

Supplementary Tables

Supplementary tables are found in the following link: https://www.dropbox.com/sh/m9j5d0zytx07g9l/AABdyYW6oLEe7QbdXS1KFdzWa?dl=0

Table S.1. Biological and clinical information of the patients from the ICGC project analyzed by RNA-seq.

Table S.2. Biological and clinical information of the patients samples treated with H3B-8800.

Table S.3. Gene Ontology terms enriched. for the genes differentially expressed in *SF3B1*^{WT} and *SF3B1*^{K700E} MEC1cell lines after treatment with 75nM H3B-8800 .Relative to Figure II.12

Papers published during the Thesis

1. Bonnal, S.C.; López-Oreja, I.; Valcárcel, J. Roles and mechanisms of alternative splicing in cancer — implications for care. *Nat. Rev. Clin. Oncol.* **2020**, *17*, 457–474, doi:10.1038/s41571-020-0350-x.

2. Gimenez, N.; Tripathi, R.; Giró, A.; Rosich, L.; López-Guerra, M.; López-Oreja, I.; Playa-Albinyana, H.; Arenas, F.; Mas, J.M.; Pérez-Galán, P.; et al. Systems biology drug screening identifies statins as enhancers of current therapies in chronic lymphocytic leukemia. *Sci. Rep.* **2020**, *10*, 1–16, doi:10.1038/s41598-020-78315-0.

3. López-Oreja, I.; Playa-Albinyana, H.; Arenas, F.; López-Guerra, M.; Colomer, D. Challenges with approved targeted therapies against recurrent mutations in cll: A place for new actionable targets. *Cancers (Basel).* **2021**, *13*, doi:10.3390/cancers13133150.

ACKNOWLEDGEMENTS

Gracias

por depositar vuestra confianza en mí y guiarme

por acompañarme y enseñarme

por escucharme y arroparme

por las experiencias compartidas

y por ampliar mi horizonte

Eskerrik asko

urruti egon arren, gertu sentitzen zaituztedan guztiei,

etxekoei

eta zuri