Dysregulated Spliceosome Activity Involves Increased Intron Retention Plus Upregulation and Phosphorylation of SF3B1 in Chronic Lymphocytic Leukemia

Manoj Kumar Kashyap (mkkashyap@ggn.amity.edu)
University of California
Hiren Karathia
University of Maryland
Deepak Kumar
University of California
Roberto Vera Alvarez
National Library of Medicine, National Institutes of Health
Paula A Lengerke-Diaz
Mayo Clinic
Jose Vicente Forero-Forero
Mayo Clinic
Eider Moreno
Mayo Clinic
Juliana Velez Lujan
University of California
Carlos Ivan Amaya-Chanaga
University of California
Newton Medeiros Vidal
National Library of Medicine, National Institutes of Health
Zhe Yu
University of California
Daniel Achinko
National Library of Medicine, National Institutes of Health
Emanuela M. Ghia
University of California
Michael Y. Choi
University of California
Laura Z. Rassenti
University of California
Leonardo Mariño-Ramírez
National Library of Medicine, National Institutes of Health
Stephen M. Mount
University of Maryland
Sridhar Hannenhalli
National Cancer Institute, National Institutes of Health
Thomas J. Kipps
University of California
Januario E. Castro
University of California

Research Article

Keywords: RNA splicing, CLL, intron retention, RNA-Seq, macrolide, pladienolide-B, spliceosome, Alternative RNA splicing

Posted Date: May 11th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1622375/v1

Abstract

Background

Alternative splicing (AS) is a fundamental process in eukaryotes contributing to the diversity of mRNA isoforms with variable ratios of intron/exon. SF3B1 is a pivotal protein of the spliceosome machinery. Mutations in the *SF3B1* gene have prognostic value in Chronic Lymphocytic Leukemia (CLL). Our previous studies have shown that SF3B1 inhibition with macrolides induces apoptosis specifically in CLL compared with normal cells. SF3B1 inhibition is associated with a widespread increase in intron retention (IR) on most transcripts, suggesting that IR can be used as a marker of spliceosome inhibition in CLL cells. However, it is unknown how the activity of SF3B1 contributes to the spliceosome activity and the poor clinical prognosis in CLL.

Methods

To better understand this process, we performed a comprehensive analysis to quantify the abundance of individual exonic and intronic mapped reads on annotated RNA-Seq transcripts derived from the B cells of 98 CLL patients and nine healthy volunteers (Normal B cells – NBC). We calculated ratios for intron and exon abundance for each transcript and use this as a measure of intron retention (IR) and a surrogate for alternative splicing (AS).

Results

We found that 66% of CLL B-cells transcripts had significant IR elevation compared to NBC and that transcripts with high IR were associated with low expression levels and mRNA downregulation. The IR increase in CLL B-cells was independent of prognostic factors such as *IgVH* or *SF3B1* mutations. Transcripts with the highest IR levels belonged to biological pathways associated with gene expression and RNA splicing. Also, we observed a >2-fold increase of active pSF3B1 in CLL B-cells compared to NBC. Additionally, when the CLL-B cells were treated with macrolides (pladienolide-B), a significant decrease in pSF3B1, but not total SF3B1 protein was observed.

Conclusions

Our findings suggest that IR/ARS is increased in CLL and that this process is associated with SF3B1 phosphorylation and susceptibility to SF3B1 inhibitors. These data provide additional support to the relevance of AS in carcinogenesis and evidence of pSF3B1 participation in this process.

Background

Chronic lymphocytic leukemia (CLL) is the most common B-ell malignancy among adults, characterized by apoptosis defects that provide a survival advantage to neoplastic B-lymphocytes. High-risk patients such as those with a deletion of chromosome 17 (del17p) or TP53/ATM mutation generally fails to respond to conventional chemotherapy. Thus, understanding the biology of CLL cells may enlighten the path to novel therapeutic development(1, 2).

Alternative RNA Splicing (ARS) is a widespread process in eukaryotes that contributes to post-transcription mRNA isoform diversity. These isoforms are essential for several physiological functions such as cell division, differentiation, stress response, and linage specification (3-8). Many RNAseq based approaches are used to characterize the basic pattern of the ARS events like exon skipping (ES), alternative 5’ or 3’ splice sites, mutually exclusive exons and intron retention (IR) (9, 10). Generally, these characterizations are based on capturing the abundance of RNAseq reads that map to the annotated composition of exons and introns, also known as exon and IR of a transcript (11). The exonic reads reflect mature cytoplasmic mRNAs, and recent studies suggest that mapping the RNA-Seq reads to the defined intronic regions of a gene essentially reflects nascent transcript (pre-RNA) (9, 12, 13).

EISA (exon-intron split analysis) is a practical tool that analyses separately the effect of pre-mRNA and mRNA forms a transcript. EISA quantifies changes in the exonic and intronic reads in different conditions and effectively characterizes post-transcriptional changes in gene expression (13). In recent years, IR has been the focus of several studies for its functional importance in cancer studies, its association with alternative RNA splicing, and carcinogenesis (14-17). Our previous studies support this concept, as we demonstrated *in vivo* and *in vitro* evidence that IR is significantly elevated in CLL cells compared to normal B cells (NBC). However, it is unknown the regulatory pathways that contribute to the increase in IR/AS in CLL and what fraction of the intron-using transcript, a proxy of pre-mRNA expression, contributes to the total RNA expression activity in CLL and other malignancies(18).

The spliceosome machinery is a protein complex that regulates RNA splicing in eukaryotic cells. It comprises many splicing factors, and the best characterized include splicing factor 3B subunit 1 (SF3B1), U2AF1, and SRSF2. Evidence suggests that mutations of the splicing factor genes are associated with the aberrant splicing process during tumor development and metastases(19, 20). In CLL, mutations of *SF3B1* are found in ~10-15% of cases and constitute an independent prognostic factor associated with rapid progression, short survival, and lack of response to conventional treatments; hence, SF3B1 appears to be a highly relevant target for CLL therapy(21, 22). Nevertheless, how the mutated SF3B1 (Mut-SF3B1) or imbalanced expression of the SF3B1 in CLL cells affects the accumulation of IR on transcripts is mostly unknown. Our previous *in vitro* study suggests that hyper-activation of *SF3B1* might be related to transcripts with elevated IR in CLL cells. It might explain phenotypes associated with survival, cell proliferation, and possibly treatment resistance(18).

PLAD-B and FD-895 are two polyketides known to modulate the ARS by targeting SF3B1. This protein’s specific inhibition is postulated to be responsible for their mechanism(s) of antitumor activity (18, 23-25). Our previous work demonstrated that these polyketides modulate IR level on a few selected genes in the treated CLL cells, but not in untreated CLL or NBC samples (18). Interestingly, we also found no significant difference in expression of Mut-SF3B1 *vs.* wild type *SF3B1* (Wt-*SF3B1*) in CLL cells, speculating that the contributions of *SF3B1* to changes of the IR may not be mediated through its transcriptional activity, but rather its activity at the protein level (18). However, it is not yet clear yet, how the activity of SF3B1 contributes to the regulation of IR/ARS in CLL cells.

We carried out this study to achieve following specific aims: 1) Obtain a profile of the sets of transcripts associated with IR/ARS and their correlation with prognostic factors in CLL. For this, we use bioinformatics tools to compare in- depth RNA-Seq data (available from the European Genome-phenotype Archive [EGA]) between CLL and NBC(26), 2) Validate changes in abundance of transcripts previously found altered, using RNA isolation followed by RT-PCR of another cohort of CLL vs. normal B-cells. 3) Attempt to elucidate the mechanistic pathway associated with IR/ARS alterations in CLL patients and the relevance of the function of SF3B1 at the protein level (**Figure 1**).

Methods

Strategy to identify the IR events in CLL vs NBC samples

The RNA-Seq data were obtained from the European Genome-Phenome Archive (EGA, http://www.ebi.ac.uk/ega) (27). This data comes from 98 CLL patients and

9 NBCs samples, and can be found in the following link on EGA, https://www.ebi.ac.uk/ega/studies/EGAS00001000374 (Study ID: EGAS00001000374)(26). In this study we used these datasets to study the significance of IR in CLL cells and NBC. The workflow in for the study shows different steps taken for executing the RNA-Seq analysis on this dataset (**Figure- 1**).

The details of the samples are provided in the ***Supplementary Table-1***. The details of the patients with only *SF3B1* mutations are provided in the ***Supplementary Table-2***.

The samples were classified into two risk groups: high risk and low risk based on IGHV mutational status. CLL patients with leukemic cells expressing unmutated IGHV (U-CLL) were classified as high-risk (HR) and patients with CLL cells expressing mutated IGHV (M-CLL) as low-risk (LR). Illumina platform was used to sequence RNA collected from these CLL samples. The obtained reads were mapped to the Ensembl GRCh37.62 B (hg19) reference genome using RNA sequence aligner Tophat that aligns the reads across splice junctions independently of gene annotations (28). The transcriptome-wide data was analyzed using Transcript Per Million (TPM) analysis. The transcripts that were expressed in at least 60% of both NBCs and CLL cells samples were subsequently analyzed for IR analysis.

Quantification of IR and Isoform expression

For studying this problem, we downloaded paired-end reads fastq files of RNA- Seq from the European Genome-phenome Archive (https://www.ebi.ac.uk/ega/home). To perform the intron/exon quantification, we follow the guidelines published previously (13, 29), and detailed methods used in this study are provided in **supplementary material.**

The reads from each sample were mapped onto the hg19 human genome using Tophat, version 2.0.10,(30) with default parameters. The output of that was a SAM file, which further converted to BAM file using samtools (Version 1.3) (28), *view* command and sorted the BAM file using and the samtools *sort* command. Each sorted BAM files for NBC and CLL cells were used to calculate TPM values for each genomic feature (isoform, exonic and intronic regions) for every transcript. For the TPM calculation, we downloaded tablemaker and Ballgown Bioconductor packages (31). Using the tablemaker package, we called first inbuilt Cufflinks command to reconstruct the individual transcripts and isoforms to estimate multi-mapped read count value for each of the genomic features from every BAM file and converted the output into Ballgown readable formatted file.

Later, using the Ballgown Bioconductor command, we loaded all readable formatted files of multi-mapped-corrected read values in two separate matrices in each case, for NBC or CLL cells, corresponding to overlapped exonic and intronic regions for every transcript defined in hg19-GTF. We also quantify TPM expression values for every transcript and constructed two matrices for NBC and CLL cells using the Ballgown command's "transcript-gene-table" option applying equation-1. We retained only those transcripts whose TPM expression values were >0 in more than 60% of NBC and CLL samples to assure that the analyzed transcript expresses in the large majority of sample both in NBC and CLL cells conditions. The intron-less transcripts were removed from the hg19 GFF file. With this, we could capture 25276 transcripts, which we used for the downstream analyses (***Supplementary Table-3***).

We created an *in-house* program using python to calculate TPM values(32) for every exonic and intronic region of the expressed transcripts by applying the formulas 2 & 3 shown in **Supplementary material.**

A 2D volcano plot was created to schematize the differential IR between CLL cells and NBC, CLL cells harboring Mut-*SF3B1* and NBC, CLL cells harboring Wt-*SF3B1* and NBC, CLL cells harboring Mut-*SF3B1* and CLL cells harboring Wt-*SF3B1* and finally M-CLL cells and U-CLL cells. The x-axis displays the log2 fold change ratios between two different cell types (TPM (intron/exon) / TPM(intron/exon)) and the y-axis corresponds to the $\log_{10}$ of the mean expression value (p-value) (**CLL-B (transcript expression) / NBC (transcript expression**)..

Association between IR and transcript expression in CLL *vs.* NBC

To assess the differences in IR and transcript expression association between CLL cells and NBC, first we classified all the transcripts into two sets: (i) transcript with significantly high IR values in CLL cells and (ii) transcript with significantly high IR values in NBC. For this, we compared the transcript-specific distribution of TPM in NBC-Intronic vs CLL cells -Intronic values. We performed two separate one-sided Wilcoxon test, each for (CLL cells > NBC) and (NBC > CLL cells) conditions. The transcripts satisfying (p <0.05; False Discovery Ratio FDR=5%) for CLL cells > NBC were separated as set-i and the same for the condition NBC > CLL cells were separated for set-ii transcripts.

Within each set, we further classified three sets of transcripts: (A) over-expressed, (B) under- expressed and (C) non-differentially expressed. For this, we used TPM values from the transcripts expression and performed one-sided Wilcoxon test each for (CLL cells > NBC) and (NBC > CLL cells) at $p < 0.05$ and FDR=5% for set-A and set-B. For set-C we separately checked for expression difference for CLL cells ≠ NBC and NBC ≠ CLL cells at $p>0.05$ and considered respectively as set-C for CLL cells and NBC.

After capturing several transcripts fulfilling the six sets, we first constructed 2x3 contingency table and performed a Chi-square test. We also performed Fisher's of 2x2 contingency table for high/low intron-used transcript vs. over/under expression and recorded ODDs score.

To demonstrate sample-wise association study between high intron-retaining transcript in CLL cells with expression, we selected top 200 transcripts from the set-IA based on ranked P-value and captured the TPM values for CLL-intron and CLL-transcript from every CLL samples and also the TPM values for NBC-intron and NBC-transcript for each of the 200 NBC transcripts. Later we converted the TPM values into Z-scores and plotted them into heatmaps using an in-house built R-script. The heatmap column was sorted: first, NBC samples, second, M-CLL samples and last U-CLL samples.

Analysis of Biological Pathways and Molecular Functions affected due to IR in CLL *vs.* NBC

We used the top 25% transcripts with high IR in CLL cells as compared with NBC. The biological pathway enrichment was done with PANTHER (Protein Analysis Through Evolutionary Relationships) classification system. Top 25% transcripts used with significantly high IR in CLL cells compared with NBC ($p \leq 0.05$; FDR=0.1). We selected the transcript gene symbols used those as input, chose "Panther Overrepresentation Test" and picked Reactome Pathway as the annotation set. The *p*-value was adjusted using a Bonferroni correction (33, 34). Molecular function enrichment was done using WEB-based Gene Set Analysis Toolkit (WebGestalt, http://www.webgestalt.org) (35).

Network analysis

We used Reactome Functional Interaction (FI) network database a highly reliable, manually curated pathway-based protein functional interaction network(36), to analyze RNA-Seq data obtained from 98 CLL patients and 9 NBCs. Reactome FI network database identified protein interaction data for 9821 genes. Protein interaction networks were used to assign sets of genes to
discrete subnetworks(37). A subnetwork is defined as a graphical representation of genes, represented as nodes, and functional interactions represented as co lines between nodes. The color of the nodes represents the log2 ratio of IR in CLL versus in NBCs. A version V3.4 of Cytoscape was used to
identify and visualize the first neighbors of SF3B1 in the subnetwork (38).

B-cell isolation: Peripheral blood mononuclear cells (PBMC) derived from CLL cells were obtained from the CLL tissue bank of the Moores Cancer Center, UC San Diego, La Jolla. Upon confirmation of the diagnosis of CLL(39),
patients provided written informed consent for blood sample collection on a protocol approved by the Institutional Review Board of UCSD, following with the D of Helsinki (40). Normal B-cells (NBC) were purified from buffy coats of healthy volunteer donors. Positive isolation with Dynabeads CD19 pan B
(Life Technologies) and DETACHaBEAD CD19 (Life Technologies) were used to achieve more than 95% purity by flow cytometry analysis.

CLL cells: Chronic lymphocytic leukemia B-cells (CLL cells) were either used
as pure (purified for profiling using negative selection kit (23), or studied by gating on double-positive cells for CD19/CD5 surface markers.

SF3B1 sequencing for characterization of CLL samples to classify as Mut- SF3B1 or Wt-SF3B1

For profiling of the CLL cells to know the *SF3B1* mutational status, we took $5e10^6$ CLL cells and processed them for DNA isolation as described previously (41). The details of the CLL samples used for SF3B1 and pSF3B1 profiling along with their SF3B1 mutational status are provided in ***Table-1***.

Table-1

SF3B1 mutational status in CLL patients used for expression profiling of SF3B1 and pSF3B1

Sample ID	Type of *SF3B1* Mutation	Type of *SF3B1* Mutation
CLL101	Heterozygous	K700E
CLL102	Heterozygous	K700E
CLL103	Heterozygous	E622D
CLL104	Heterozygous	N626Y
CLL105	Heterozygous	K700E
CLL106	Heterozygous	K700E
CLL107	Heterozygous	G740E
CLL108	Heterozygous	K700E
CLL109	Heterozygous	K700E
CLL110	Heterozygous	R625C

The DNA amplification for different *SF3B1* mutations focused on exons 14 and 15 was carried out using primers listed in ***Table #*** **2**.

Table # 2

List of primers for characterization of mutation in CLL B- cells for *SF3B1* gene

SF3B1	Primer Direction	Primer Sequence	PCR Product (BPs)
Exon 14	Forward	5' TCTGTTTATGGAATTGATTATGGAA 3'	424
Exon 14	Reverse	3' GGGCAACATAGTAAGACCCTGT 5'	
Exon 15	Forward	5' TTGGGGCATAGTTAAAACCTG 3'	209
Exon 15	Reverse	3' AAATCAAAAGGTAATTGGTGGA 5'	

Reverse transcription PCR (RT-PCR)

For validation of candidates for IR, CLL or normal B cells (NBCs) ($5e10^6$ cells/well) were subjected to RNA isolation. Total RNA was isolated using total the RNA isolation kit from Life Technologies (Grand Island, NY USA). For IR validation studies, the potential source of false positives could be the traces of DNA present in CLL or NBC samples after RNA isolation; isolated RNA samples were treated with DNase I to remove DNA contamination and then cDNA was prepared as described previously(23). The list of primers used in the study for validation of IR for transcripts of different genes is shown in ***Table #*** **3.**

PCR conditions were as follows: 95°C for 3 min; 35 cycles of 95°C for 30 s; 55- 58°C (dependent on transcript specific primers) for 30 s, and 72°C for 45 s; followed by 72°C for 5 min using a PTC-100 thermocycler (MJ Research). PCR products were separated on a 2% agarose gel and stained with ethidium bromide (23).

Expression of SF3B1 and phospho-SF3B1 in CLL samples

The CLL cells and normal B cells were seeded the day before the experiment and cultured overnight. Before plating, the DMSO content was removed by making the total volume to 10 ml by adding HBSS and the cells were centrifuged at 300g for 5 minutes at RT. The supernatant was discarded, and the cells were re-suspended in appropriate volume to have at least 1e6 cells/50 µl/well.

The cells were plated and incubated with 10 µg/ml of anti-IgM (catalog # 2022- 08, Southern Biotech) for 15 minutes at 37°C (42, 43). After incubation, the cells were centrifuged immediately at 4°C to stop the reaction and washed twice
with FACS buffer. The cells were labeled with CD19/CD5 (for CLL cells) and CD19/CD20 (for NBCs) and incubated for 30 minutes at 4°C. The cells were incubated in pre-warmed cytofix buffer at 37°C for 10 minutes before use and centrifuged at 300g for 5 minutes; the supernatant was removed. The cells were disrupted by pipetting back and forth and permeabilized using Phosphoflow PERM buffer by incubating for 30 minutes on ice. The cells were washed twice with FACS buffer and centrifuged at 300g for 10-15 minutes to remove the supernatant.

The PBMCs isolated from human blood were processed for CD19/CD20 staining at 4^0C. The cells were washed twice with FACS buffer and fixed immediately using BD™ Phosphoflow Perm Buffer (Catalog # 558052) for 10 minutes at 37°C.

Table # 3

List of primers used for validation of intron retention for different genes in CLL and normal B-cells

Gene Symbol	HGNC Approved Name of the molecule (Genomic region probed)	HGNC ID	HPRD ID	Refseq ID	Coordinates on UCSC Genome Browser
ADTRP	Androgen dependent TFPI regulating protein (chr6:11716510-11716630)	21214	12836	NM_032744	chr6:11716510- 11716529
					chr6:11716611- 11716630
RPL39L	Ribosomal protein L39-like (chr3:186,847,868- 186,848039)	17094	09610	NM_052969	chr3:186,847,868- 186,847,887
					chr3:186,848,019- 186,848,039
HS3ST1	Heparan sulfate (glucosamine) 3-O- sulfotransferase 1 (chr4:11,424,163-11,424,329)	5194	04455	NM_005114	chr4:11,424,163- 11,424,182
					chr4:11,424,310- 11,424,329
CTLA4	Cytotoxic T lymphocyte- associated antigen 4 (chr2:203871925-203872008)	2505	00474	NM_005214	chr2:203871925-203871944
					chr2:203871989- 203872008
FMOD	Fibromodulin (chr1:203343741-203343920)	3774	*02591*	NR_103757	chr1:203343741- 203343760
					chr1:203343896- 203343920
GUCY2C	Guanylate Cyclase 2C (chr12:14625022-14625136)	4688	*07529*	NM_004963	chr12:14625022- 14625041
					chr12:14625117- 14625136

FP: forward primer, **RP**: reverse primer, **Tm**: Melting temperature

Post-incubation, cells were centrifuged and after decanting the supernatant the cells were fixed with BD PhosphoflowTM Perm buffer II (catalog # 558052) for 30 minutes on ice. The cells were centrifuged at 300g for 5 minutes and the supernatant was discarded. Further, cells were washed twice with FACS buffer. The Fc receptor (FcR) was blocked by using FcR blocking agent (catalog # 130- 059-901, Miltenyi Biotec) in a ratio of 1:25 at room temperature for 10 minutes. The cells were treated using 1:200 dilutions of rabbit anti-Human SF3B1 (D7L5T, catalog #14434, Cell Signaling Technology), rabbit anti-human phosphor-SF3B1 (Catalog # 25009S, Cell Signaling Technology), or normal rabbit IgG (catalog # 2729, Cell Signaling Technology) as a non-specific isotype control for 30 minutes at room temperature. The cells were washed twice with FACS buffer and incubated with Alexa Fluor® 488 Conjugated goat anti-rabbit IgG (H+L), F(ab') fragment (catalog #4412, Cell Signaling Technology) with a dilution of 1:50 for 30 minutes at RT. The cells were washed twice with FACS buffer and subjected to run on the flow cytometer. The data analysis was carried out using flowjo software. All the samples were run in duplicate.

Western blot analysis

CLL cells were treated with 100 nM of PLAD-B for 15, 60, and 180 minutes. The lysates were prepared using a modified RIPA buffer. Untreated CLL cells were used as a control. The samples containing at least 20 µg of total protein were subjected to 4-20% Criterion Precast Gel (Bio-Rad) SDS-PAGE, followed by transfer using polyvinylidene difluoride (PVDF) membrane (Millipore). After blocking with 5% bovine serum albumin (BSA) in TBST (20 mM Tris•HCl, 137 mM NaCl, 0.1% Tween-20 pH 7.6), the membrane was incubated with primary antibody overnight at 4 °C. The following primary antibodies were used: rabbit monoclonal anti-SF3B1 at dilution 1:1000 (14434, clone D7L5T, Cell Signaling Technology), mouse anti-phospho-SF3B1 (Anti-pSF3B1) at dilution 1:1000 (clone, D8D8V, site: Thr313, catalog # 25009S, Cell Signaling Technology), and rabbit anti-β-actin at dilution 1:5000 (Catalog # 4967, Cell Signaling Technology).(24) After washing twice with TBST, the membranes were incubated with horseradish peroxidase (HRP)-labeled anti-rabbit (Catalog # sc-2030, Santa Cruz Biotechnology) or HRP-labeled anti-mouse (Catalog # sc-2031, Santa Cruz Biotechnology) secondary antibodies with a dilution of 1:5000 dilution for 1 h at room temperature in TBST with 5% skimmed milk. Protein-antibody complexes signals were detected by exposing the X-ray films (Catalog # E3018, HyBlot CL) after treatment with Enhanced chemiluminescence (ECL) kit (catalog # 32106, Pierce Thermo Scientific).

Statistical analyses

Data was analyzed using GraphPad Prism 5.0 (GraphPad Software, Inc.). The error bars represent standard deviation (SD). Statistical differences for the mean values are indicated as follows: *, **, *** and **** denote $p<0.05$; $p<0.01$; $p<0.001$, $p<0.0001$ respectively.

Results

Our initial study found that the intron/exon's ratio was significantly higher in CLL than NBCs **(Supplementary Figure-1)**. The observations were quite intriguing, so we decided to test our observation in a large cohort of CLL and NBC samples. For the same, we analyzed the RNA-Seq data obtained from EGA(26).

Classification of RNA-Seq data

The immunoglobulin heavy chain variable region genes (IGHV) were available for 97 (99%) CLL samples. Forty-one had CLL cells expressing unmutated-IGHV (HR) with 98% homology to the germline IGHV, whereas fifty-six had CLL
cells expressing mutated IGHV genes (LR). One sample without known IGHV mutational status was excluded from the analysis (**Supplementary table-**

1). Eighty-five patients had CLL cells harboring Wt-*SF3B1*, whereas nine patients had CLL cells carrying the Mut-*SF3B1* gene. The *SF3B1* mutational status was not available for four samples, so these samples were excluded from the analysis (**Supplementary table-2**).

Identification of IR events in CLL *vs.* NBC samples

The transcriptome-wide data was analyzed using TPM analysis. The ranscripts were expressed in at least 60% of both NBCs and CLL samples were subsequently analyzed for IR analysis. The global distribution of IR for 14811 transcripts is shown in the volcano plot of **Figure 2**. The lists of transcripts are provided in ***Supplementary Table-3***. Also, an overall pattern for intron/exon ratio for all CLL cells (regardless of their subtypes) *vs.* NBC has been shown in **Supplementary Figure-2**A. The scatter plot of CLL cells vs NBCs presents overall distribution of IR (**Supplementary Figure 2A)** where we observed that there were 20906 transcripts (61%) above the regression line and 13894 (39%) below the regression line, indicating that a large majority of transcripts had >1 intron/exon ratio in CLL cells as compared with NBC, consistent with our previously published study on CLL(18).

The volcano plot created to schematize the differential IR between two cell types gives a global view of expression level and IR. The upper right quadrant in **Supplementary Figure 2B** represents values positive for both quantities.

The isoforms plotted in this quadrant were interpreted as more expression
and more IR in CLL cells. The lower right quadrant was for positive values of the "log2 of CLL cells (intron/exon) / NBC (intron/exon)" and low values of the "log10 (CLL cells **(transcript expression) / NBC (transcript expression)**)". This quadrant showed isoforms with more IR in CLL cells but with lower expression. The upper left quadrant shows negative values of "$\log_2$ of CLL cells (intron/exon)

/ NBC (intron/exon)" and positive values of the "$\log_{10}$ (CLL cells **(transcript expression) / NBC (transcript expression)**)". It showed more IR in NBC but more expression in CLL cells. Finally, the lower left quadrant was opposite of the right upper quadrant. It shows more expression and more IR in NBC than in CLL cells (**Supplementary Figure 2B)**. The volcano plots making a comparison between CLL cells harboring Mut-SF3B1 and NBC (**Figure 2B**), CLL cells harboring Wt-SF3B1 and NBC (**Figure 2C**), CLL cells harboring Mut-SF3B1 and CLL cells harboring Wt-SF3B1 (**Figure 2D**) and finally M-CLL cells and U-CLL cells (**Figure 2E**) have been interpreted similarly.

Further, the scatter plot of **Figure 2F** shows the comparison between IR ratio on the x-axis (Log_2 CLL intron / NBC intron) and exon ratio on the y-axis (Log_2 CLL exon / NBC exon), for all transcripts between CLL cells *vs* NBC.

Association between intron retention and transcript expression in CLL

cells To assess the transcripts' global association of IR and expression of the transcripts, we first took the complete 25276 reference transcripts. We classified them based on levels of IR and expression of the transcripts in CLL cells and NBC using a one-sided Wilcoxon test (see Methods). Using this, we found six sets of distinct transcripts: set-IA, set-IB, and set-IC, corresponding to transcripts contributing to high IR with upregulation, high IR with downregulation and high IR with non-differential expression in CLL. Similarly, set-IIA, set-IIB, and set-IIC correspond to transcripts contributing similarly to three sets in NBC. We found a total of 16725 transcripts contributing to differentially intron-retaining transcripts in CLL cells and NBC, among those, 11969 (set-I=71%) were significantly high intron-retained in CLL cells, and 4756 (set- II=29%) were significantly high used in NBC with p-value < 0.05 (FDR=5%). Within the 11969 transcripts, the set-IA

contains 10436 (87%), set-IB contains 188 (2%) and set-IC contains 1345 (11%). We constructed a 2x3 contingency table (Figure 2A) and performed a Chi-square test. The Chi-square test value was highly significant with a p-value < 1.0e-300. The data presented in the bar graph (**Figure 3A**), show overexpression of high and low intron-retaining transcripts in CLL and NBCs. We plotted set-IA, set-IB, set-IIA and set-IIB, performed Fisher's exact test and observed the result with p- value< 1.0e-300 and ODDs=863. These results suggest that in CLL cells the IR contributes positively to the up-regulation of corresponding transcripts.

To study the sample-wise association between the IR and transcript expression, we sorted the set-I (11969 transcripts), ranked them by p-value from lowest to the highest order (**Figure 3A**), and selected the top 200 transcripts (see **Methods and Supplementary Information**). We recorded values for these 200 transcripts; for IR and transcript expression observed in all CLL cells and NBC samples. We segregated the CLL samples based on IGHV mutational status in order to understand the association with disease prognosis. An overall sample- wise IR in CLL cells was positively associated with the corresponding transcripts sample-wise expression as compared to IR and transcript expression in NBCs
(**Figure 3B & 3C**). However, the association between IR and transcript expression observed, comparing unmutated-IGHV CLL cases (HR) with NBC's had a statistically higher significance (p<0.05) than the association kept comparing mutated-IGHV CLL cases (LR) with NBCs. Further, the degree of association was increased case of IGHV unmutated (poor prognosis) cases.

Analysis of Biological Pathways and Molecular Functions Altered Due to IR between CLL *vs.* NBC in transcripts with maximum IR

Next, we analyzed the top 25% transcripts from Figure-1A with high IR in CLL compared with normal B-cells (**Supplemental table 4**). We observed everal different pathways affected by IR F**igure 4A**), including extracellular matrix including collagen organization, inflammation mediated by chemokine and cytokine, Hemostasis, antigen processing (T-cell activation), immune-regulatory interactions between a lymphoid and non-lymphoid cell, and cellular response
to stress. However, we found RNA processing, splicing, and gene expression pathway as the major ones where 30% of the transcript input belongs to this pathway. In summary, the IR containing transcripts were enriched for biologically important pathways in CLL cells.

We analyzed the top 25% transcripts for molecular function (MFs) and assessed the effect of IR on molecular function in CLL cells. We observed several different MFs affected due to high IR in CLL/NBC (**Figure-4B**). Among the top MFs

were protein binding, ion binding, nucleic acid binding, hydrolase activity, and transferase activity.

Network Analysis

We analyzed RNA-Seq data obtained from 98 CLL patients and 9 NBCs sing Reactome Functional Interaction (FI) network database, a highly reliable, manually curated pathway-based protein functional interaction network; this analysis allowed us to obtain protein interaction data for 9821 genes. Protein interaction networks were used to assign sets of genes to discrete subnetworks.

Cytoscape V3.4 software enables us to identify the first neighbors of SF3B1 and visualize the subnetwork centered on SF3B1. In this subnetwork, the nodes' color of the nodes represents the log2 ratio of IR in CLL cells versus in NBCs. (**Figure-4C**).

Furthermore, the network analysis provided insight into the pathways that emerged due to high IR in NBC: the immune system, metabolism of RNA, DNA replication, cell cycle, and metabolism of proteins **(Supplementary Figure-3)**. When analyzed in CLL cells, the network analysis provided insight into the pathways that emerged due to high IR in CLL cells: immune system, homeostasis **(Supplementary Figure-4)**.

Validation of candidate transcripts for IR in CLL cells and NBC

To validate our findings from the high-throughput RNA-Seq data, we chose six transcripts with high IR in CLL cells and selected a region showing the event; the area, including the intron, was amplified, and we found that in all CLL cells but no in NBC the intron was present.

The six transcripts included novel as well as known transcripts that have been described previously in CLL cells. One of the transcripts for the *CTLA4* gene has been shown for the intronic region retained in CLL cells but not in NBC (**Figure 5A**). The six transcripts that were validated using RT-PCR were: cytotoxic T-Lymphocyte associated protein 4 (*CTLA4*), androgen-dependent TFPI regulating protein (*ADTRP*), fibromodulin (*FMOD*), heparin sulfate-glucosamine 3- sulfotransferase 1 (*HS3ST1*), guanylate cyclase 2C (GUCY2C), ribosomal protein L39 like (RPL39L) (**Figure 5B**). The IR ratio for *CTLA4* was 29.5 in CLL cells as compared to NBC. We were able to amplify the intronic region in all CLL cells but not detected in NBC.

We validated C6orf105, which was later designated as the *ADTRP* gene **(Figure 5B)**. The IR ratio between CLL cells and NBC was 25.6, indicating a significantly higher IR in CLL cells compared with NBCs. Another known gene was (*FMOD*), for which we found a significant increase in IR in CLL cells compared with NBC. Previously, *FMOD* was overexpressed in CLL compared with NBC (44, 45). The ratio of IR in CLL cells /NBC was 23.4, and it was also validated in CLL cells **(Figure 5B)**. Overall, we observed aberrant splicing and in particular high IR in CLL cells compared to NBC in ~73% of the transcripts.

Up-regulation of total SF3B1 and pSF3B1 in CLL Cells

A large number of studies in solid tumors and hematological malignancies reported aberrant regulation of spliceosome complex. Since, SF3B1 is one of the significant spliceosome complex proteins associated with leukemia, we reasoned to study SF3B1, one of the spliceosome complex's significant proteins(46). We found that the expression of SF3B1 and pSF3B1 was significantly higher in CLL cells compared with NBC's (p<0.0001 and p<0.01, respectively) **(Figure 6A**). However, no significant difference was observed between SF3B1 and pSF3B1 expression levels in NBCs.

Effect of Anti-IgM on SF3B1 and pSF3B1 expression in CLL cells and NBC

We tested whether SF3B1 and pSF3B1 expression could be modulated by anti- IgM stimulation. After *in vitro* stimulation of CLL cells with anti-IgM for 15 minutes, no difference in SF3B1 expression was observed between IgM stimulated and un-stimulated; the same effect was observed between stimulated and un-stimulated NBCs. Also, no difference in SF3B1 expression was observed we compared CLL cells; we made a comparison between Mut-*SF3B1* and Wt-*SF3B1* (**Figure 6B**). But, when we compare both groups of CLL cells (Mut-SF3B1 and Wt-SF3B1) with NBC, higher expression of total SF3B1 was evident in CLL (p-value<0.05).

When the expression of pSF3B1 levels was observed between Wt-*SF3B1 vs.* Mut-*SF3B1* CLL samples, no differences were observed; however, in both cases, there was a significant rise in protein expression after IgM stimulation. The pSF3B1 levels in both Wt-SF3B1 and Mut-SF3B1 had a marked overrepresentation compared with NBC in both treatments, with IgM stimulation or without stimulation.

Modulation of Expression of SF3B1 and effect on post-translational modification upon macrolide probe treatment in CLL cells

Since we observed IR in more than >70% of the transcripts, we hypothesized that the CLL spliceosome machinery could have an operation defect that makes it does not complete the intron removal function as it should. Due to the lack of an ideal method to measure spliceosome activity in cells, we reasoned to use a previously reported macrolide molecule PLAD-B to modulate the splicing factor subunit SF3B1 (a driver protein for spliceosome activity) and to determine its phosphorylation. In particular, we interrogated the threonine amino acid site, which gets phosphorylated at site Thr313, reported to indicate active spliceosome(47).

We treated CLL cells with 100 nM of PLAD-B for 15, 60, 180, 360, or 960 minutes and interrogated for expression of total SF3B1 and pSF3B1. We found no significant change in total SF3B1 expression between untreated and treated CLL cells over time. However, in CLL cells, the level of pSF3B1 significantly decreased after 15 minutes of PLAD-B treatment. In contrast, the protein level of total-SF3B1 does not change, suggest that the PLAD-B could control pSF3B1 form, and in turn, modulate the IR program of the transcriptome of CLL cells (**Figure 6C)**. In contrast, no change in the expression of SF3B1

and pSF3B1 was noticed upon treatment with 10 μM of Fludarabine (F-ara-A), a conventional chemotherapy agent used in the clinic for CLL patient's treatment.

Discussion

Many studies focused on transcriptome profiling of CLL cells in recent years. They have reported aberrant splicing events and mutations in genes such as *SF3B1*, *SRSF2*, *U2AF1*, and *U2AF2*, which belong to splicing machinery, *i.e.*, spliceosome, have been reported (10, 19, 20, 26, 41, 46, 48). Furthermore, mutations in *SF3B1* are associated with poor prognosis in CLL patients(49, 50).

Alternative splicing is the process, which leads to diversity in the genome *via* the generation of novel isoforms for different genes controlled by the spliceosome complex. Several splicing events have been reported to be associated with other diseases, including cancers (14, 18)·(51-53), and other diseases like diabetes(54), cardiac hypertrophy(55), and amyotrophic lateral sclerosis(56). Many factors, the length of the exons and introns play a critical role in recognizing the splice site, and when small, it is most efficient (57). The splicing events are based on the combination of intron and exons are altered. In our study we investigated the degree to which IR was observed in the transcriptome of CLL cells compared to NBC.

Evaluation of the total events that are part of AS is very challenging, and because most AS takes place with partial or complete involvement of the intron, we decided to align the reads, quantify the retained-intron or exon across the transcriptome and calculate for each transcript the intron/exon ratio and use this value as an IR measure and a mathematical surrogate for AS.

In an earlier RNA-Seq study, we used fragments per kilobase of exon model per million mapped reads (FPKM) as the quantitative measure. Since then, there has been advancement in RNA-Seq, and FPKM can be converted to transcript per million (TPM). Due to the inconsistences amount samples introduced while quantifying the transcript abundance with reads per kilobase of exon model per million reads (RPKM) or FPKM, Wagner *et al.* introduced TPM and demonstrated that TPM respects the average invariance and eliminate statistical biases inherent in the RPKM or FPKM measures (58), and overall the effect of current study, we used TPM for RNA-Seq analysis.

In recent years, among different types of splicing events, IR came into highlight because of its association with the inactivation of tumor suppressor genes nd reports in other malignancies including breast cancer, lung adenocarcinoma, CLL, hepatocellular carcinoma(61), and even in normal tissues/cells like T cell(62), normal granulocyte differentiation(4), and neuronal development(8). IR has been reported as a common mechanism for the inactivation of tumor suppressor genes(53).

In this study, we observed that 66% of the total transcripts have high IR, and an increased level of transcript expression. Further, we assessed change in ratio of an abundance of IR in pre-mRNA *vs.* exon-using mRNA transcripts in CLL cells and found that high ratio of the pre-mRNA to mRNA is more prominent than low ratio, which globally suggests that highly accumulated pre-mRNA transcripts may down-regulate expression of mRNA in CLL cells.

Further, among top IR harboring transcripts, RNA processing and splicing were the significant pathways represented in our findings. There are reports of genes with retained introns in concordance with our data, where an increase in transcript expression was observed for *SLC25A37*, *SPTA1,* and *SF3B1* genes(63)

For all the six gene validated and analyzed, we found that none of the NBC samples showed intron region presence, whereas all the CLL samples showed presence of the intronic fragment indicating that IR occurs pre-dominantly in CLL cells.

Among known genes, we found cortactin gene for which there was high IR as well as an increase in the transcript expression. *CTTN* has been reported in association with CLL with significant increased mRNA expression in CLL cells compared with NBC (64). We validated cytotoxic T-lymphocyte-associated protein 4 (*CTLA4*) and Fibromodulin (*FMOD*) transcripts, which have been reported to be expressed in CLL. Another example of a known gene was *FMOD*. We found high IR and over-expression for *FMOD* CLL specimens has been described previously (17). In a recent study, *FMOD* was one of the seven genes included in an expression panel that was reliably able to distinguish clonal expansions of CLL cells from normal B-lymphocytes. *FMOD* has been reported to be upregulated in CLL and to be exclusively expressed in CLL and mantle cell lymphoma. In a recent study, *FMOD* was one of the seven genes included in an expression panel that was reliably able to distinguish CLL cells' clonal expansions from normal B-lymphocytes. *FMOD* has been reported to be upregulated in CLL and exclusively expressed in CLL and mantle cell lymphoma(65).

Another gene *CTLA4* that code for cytotoxic T-lymphocyte associated protein 4 (*CTLA4*) the ratio of IR ratio was 29.5 in CLL as compared to normal B-cells. *CTLA4* is also known as CD152 and has been reported in $CD4^+$ and $CD8^+$ T cells. It is a surface protein, which bears one transmembrane (TM) domain and a signal peptide (SP). We were able to amplify *CTLA4* intronic region in all CLL but not in normal B-cells. *CTLA4* has been reported to be downregulated in CLL samples(66). CTLA4 consists of two ligands present on B-cells: CD80 (B7-1) and CD86 (B7-2), and has robust affinity towards these, and hence by binding to CD80, it can lead to inhibition of the T-cell activation through AKT phosphorylation(67, 68). We were able to amplify the intronic region in all CLL but not detected in normal B-cells. We found that none of the NBC samples showed intron region presence. In contrast, all the CLL samples showed the intronic fragment's presence, indicating that IR occurs pre-dominantly in CLL B- cells.

A widespread IR occurrence suggested an aberrant splicing pattern of splicing in CLL cells, further suggesting abnormal spliceosome activity. There has been an association between poor prognosis of CLL disease and Mut-*SF3B1*. Still, in this work, we found no significant IR difference between Mut-*SF3B1* and Wt-*SF3B1* harboring CLL cells, suggesting that the Mut-SF3B1 does not significantly influence the change of IR in CLL cells transcriptome. This may indicate that

the K700E and other mutations assessed here are not related to the overall IR effect in CLL cells. In agreement with previous findings (49), mutations in SF3B1 associated with poor outcome induced subtle but broad changes in gene expression across multiple pathways with no relation with spliceosome activity.

In this study and in agreement with previous findings,(50, 69) we reported an increased expression of *SF3B1* transcript and SF3B1 protein in CLL cells compared to NBCs. Still, there is no clear evidence whether it is because of chromatin hypomethylation of *SF3B1*(70), or hyper-phosphorylation of SF3B1 at 313th position threonine amino acid residue. The *SF3B1* mutant cases have been observed to bear detectable levels of *SF3B1* transcript, which indicates the possibility that SF3B1 has a gain of function(49, 50, 71), which explain that even Mut-SF3B1 also have overexpression of SF3B1 protein as compared to NBC. The transcript expression of SF3B1 has been reported previously as upregulated in CLL cells compared to NBC(50, 69).

Since we observed overall IR events in CLL B-cells as compared to NBCs, our study indicates that these aberrant IR events could be associated with abnormalities in the spliceosome machinery in CLL B-cells. Part of the evidence for that is the pSF3B1 activity due to 313 threonine residue, which is found only in a catalytically active spliceosome form associated with chromatin and where ~80% of the pre-mRNA splicing occurs(72). Moreover, our results suggest that there are significant abnormalities in IR/AS in cancer cells compared with their normal counterparts and highlight this pathway's role in carcinogenesis and potentially unveiling future therapeutic targets based on transcriptome analysis and gene expression.

After targeting the spliceosome with the SF3B1 macrolide inhibitor PLAD-B, we observed downregulation of pSF3B1 but no changes in total SF3B1 expression. The effect on pSF3B1 expression caused by PLAD-B appears to be highly specific as cytotoxic agents like F-ara-A did not induce pSF3B1 downregulation despite of triggering cell death and even when this compound was used at supra- physiological concentrations. These data provide further support of the relevance of pSF3B1 and its role in splicing activity, cancer cell survival, and the mechanism of action of spliceosome inhibitors such as PLAD-B.

PLAD-B effectively induces cell death effectively in Mut-*SF3B1* or Wt-SF3B1 harboring CLL B-cells, which is intriguing. One possible reason could be the location of *SF3B1* mutations as in MDS, these are located in a close cluster, but in CLL, SF3B1 mutations are spread in the heat-repeat region(73).

Our study reported many non-coding RNAs (ncRNAs) with high intron-usage, but their relevance concerning CLL disease is unknown. Interestingly, many RNAs that have been identified belonging to ncRNAs provide evidence on the translation of products in the form of peptides from the antisense transcripts and introns(74). For future directions, we plan to study intron-usage for transcripts, which retain the introns, harbor the premature termination codon, and examine how it targets the mRNA for nonsense-mediated decay (NMD). These studies could be critical as post-transcriptional mRNA processing. The transcripts' stability is emerging as significant modes of gene regulation in cancer, even though little is known about how leukemic/cancerous cells control mRNA processing or target mRNAs for degradation.

Overall, our results suggest a crucial role of IR in CLL, IR/AS patterns and pathways affected by this transcriptome alterations, an the potential impact of this in leukemia's pathophysiology. More importantly, we hypothesize that by extension, there are broader implications of our findings implicating IR/AS in the process of carcinogenesis in general. Furthermore, we found that pSF3B1 de- phosphorylation mediated by SF3B1 inhibitor (PLAD-B) is critical in the mechanism of these compounds' action and provides evidence that increased signaling through the spliceosome machinery components may be responsible for the IR/AS transcriptome alterations present in CLL.

Conclusion

We showed that clearly there is significantly higher abundance of retained introns in CLL cells. The pathway analysis as well as macrolide treatment provided evidence in favor of the dysregulated RNA splicing machinery that could be targeted in CLL. These findings have the potential to be extrapolated in other malignancies as well. The role of not only intron retention, but also of other AS events could of clinical significance in CLL, therefore that could be an area for further investigation to decipher the role of AS in leukemogenesis

Abbreviations

AS: Alternative Splicing

CLL: ChronicLymphocyticLeukemiaDEG: Differentially Expressed GeneEISA: Exon–intronsplitanalysis

ES: Exon Skipping

FDR: FalseDiscoveryRatio

FPKM: FragmentsPer KilobaseperMillion

IGHV: Immunoglobulin heavy-chain Variable-region geneIR: IntronRetention

M-CLL: IGHV Mutated CLLMAPQ: Mappingquality

MATS: Multivariate Analysis of Transcript SplicingMut-SF3B1: Mutated Splicing Factor 3B subunit 1NGS: Next-Generation Sequence

RPKM: Reads Per Kilobase per MillionSF3B1: Splicing Factor 3B subunit 1TPM: Transcripts Per Million

U-CLL: IGHVUnmutatedCLL

Wt-SF3B1: Unmutated Splicing Factor 3B subunit 1

Declarations

Ethics approval and consent to participate

For all of the patients who participated in this study, written informed consent was obtained. It was approved by the Ethical Committee of Moores Cancer Center, University of California San Diego, La Jolla ethical Regulations and conducted according to the guidelines of the Declaration of Helsinki. Animals were not used in the study.

Consent for Publication

All authors consent to the publication of the manuscript in Genome Medicine. Further, figures or tables are original, so there was no requirement of taking permission or consent from anyone.

Availability of Data and Materials

All data generated and analyzed during our study are included either in the published article or supplementary data associated with the manuscript.

Competing Interests

The authors declare that they have no competing or conflict of interests.

Funding

The authors would like to thank grant support: Lymphoma Research Foundation (LRF, grant #285871) to J.E.C., the National Institutes of Health (PO1-CA081534)-CLL Research Consortium Grant to T.J.K, J.E.C and UC San Diego Foundation Blood Cancer Research Fund (T.J.K.) and the Bennett Family Foundation (J.E.C.). MKK is the recipient of the TARE fellowship (Grant # TAR/2018/001054) from the Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India, and sanction #: 5/13/55/2020/NCD-III from ICMR, New Delhi, India.

Authors' Contributions

J.E.C., T.J.K., and M. K. K. conceived and guided the research. M.K.K., H.K., D.K., and J.V.L. were involved in designing of the experiments, analyzing the data, and interpreting the results. M.K.K., D.K., J.V.L., C.A., E.G., R.V.A., C.I.A.,

N.M.V., D.A., S.H., S.M.M., V.F.F.J, and Z.Y. were involved in conducting the experiments, data analysis, and interpretation of the results. J.E.C., M.K.K., and H.K., wrote the manuscript. J.E.C, MYC, L.R., E.G., and T.J.K. provided patient samples, clinical and laboratory data. S.H., L.M.R., and S.M.M., P.A.L.D, J.V.F.F, and E.M. participated in the critical discussion of the results and editing of the manuscript.

Acknowledgements

We thank the EGA for providing us extensive large RNA-Seq data set on CLL and normal B-cells samples.

References

1. Hewamana S, Dearden C. Treatment options for high-risk chronic lymphocytic leukaemia. Ther Adv Hematol. 2011 Jun;2(3):147-59.
2. Stilgenbauer S, Zenz T. Understanding and managing ultra high-risk chronic lymphocytic leukemia. Hematology Am Soc Hematol Educ Program. 2010;2010:481- 8.
3. Wong JJ, Gao D, Nguyen TV, Kwok CT, van Geldermalsen M, Middleton R, et al. Intron retention is regulated by altered MeCP2-mediated splicing factor recruitment. Nat Commun. 2017 May 8;8:15134.
4. Wong JJ, Ritchie W, Ebner OA, Selbach M, Wong JW, Huang Y, et al. Orchestrated intron retention regulates normal granulocyte differentiation. Cell. 2013 A 95.
5. Ergun A, Doran G, Costello JC, Paik HH, Collins JJ, Mathis D, et al. Differential splicing across immune system lineages. Proc Natl Acad Sci U S A. 2013 Aug;110(35):14324-9.
6. Moore MJ, Wang Q, Kennedy CJ, Silver PA. An alternative splicing network links cell-cycle control to apoptosis. Cell. 2010 Aug;142(4):625-36.
7. Shaked I, Zimmerman G, Soreq H. Stress-induced alternative splicing modulations in brain and periphery: acetylcholinesterase as a case study. Ann N Y Acad Sci. 2008 Dec;1148:269-81.
8. Yap K, Lim ZQ, Khandelia P, Friedman B, Makeyev EV. Coordinated regulation of neuronal mRNA steady-state levels through developmentally controlled intron retention. Genes Dev. 2012 Jun 1;26(11):1209-23.
9. Park E, Pan Z, Zhang Z, Lin L, Xing Y. The Expanding Landscape of Alternative Splicing Variation in Human Populations. Am J Hum Genet. 2018 Jan;102(1):11-26.
10. DeBoever C GE, Shepard PJ, Rassenti L, Barrett CL, Jepsen K, Jamieson, CH CD, Kipps TJ, Frazer KA. Transcriptome sequencing reveals potential mechanism of cryptic 3' splice site selection in SF3B1-mutated cancers. PLoS Comput Biol. 2015;11(3):e1004105.

11. Bjørklund SS, Panda A, Kumar S, Seiler M, Robinson D, Gheeya J, et al. Widespread alternative exon usage in clinically distinct subtypes of Invasive Ductal Carcinoma. Sci Rep. 2017 Jul;7(1):5568.
12. Zhao S, Zhang Y, Gamini R, Zhang B, von Schack D. Evaluation of two main RNA-seq approaches for gene quantification in clinical RNA sequencing: polyA+ selection versus rRNA depletion. Sci Rep. 2018 Mar;8(1):4781.
13. Gaidatzis D, Burger L, Florescu M, Stadler MB. Erratum: Analysis of intronic and exonic reads in RNA-seq data characterizes transcriptional and post- transcriptional regulation. Nat Biotechnol. 2016 Feb;34(2):210.
14. Eswaran J, Horvath A, Godbole S, Reddy SD, Mudvari P, Ohshiro K, et al. RNA sequencing of cancer reveals novel splicing alterations. Sci Rep. 2013;3:1689.
15. Zhang Q, Li H, Jin H, Tan H, Zhang J, Sheng S. The global landscape of intron retentions in lung adenocarcinoma. BMC Med Genomics. 2014;7:15.
16. Dvinge H, Bradley RK. Widespread intron retention diversifies most cancer transcriptomes. Genome medicine. 2015;7(1):45.
17. Liao W, Jordaan G, Nham P, Phan RT, Pelegrini M, Sharma S. Gene expression and splicing alterations analyzed by high throughput RNA sequencing of chronic lymphocytic leukemia specimens. BMC Cancer. 2015;15:714.
18. Kashyap MK, Kumar D, Villa R, La Clair JJ, Benner C, Sasik R, et al. Targeting the spliceosome in chronic lymphocytic leukemia with the macrolides FD-895 and pladienolide-B. Haematologica. 2015 Jul;100(7):945-54.
19. Singh B, Eyras E. The role of alternative splicing in cancer. Transcription. 2017 03;8(2):91-8.
20. Urbanski LM, Leclair N, Anczuków O. Alternative-splicing defects in cancer: Splicing regulators and their downstream targets, guiding the way to novel cancer therapeutics. Wiley Interdiscip Rev RNA. 2018 07;9(4):e1476.
21. Zhang Z, 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 Sep;8(41):69916-23.
22. Te Raa GD, Derks IA, Navrkalova V, Skowronska A, Moerland PD, van Laar J, et al. The impact of SF3B1 mutations in CLL on the DNA-damage response. Leukemia. 2015 May;29(5):1133-42.
23. Dhar S, La Clair JJ, Leon B, Hammons JC, Yu Z, Kashyap MK, et al. A Carbohydrate-Derived Splice Modulator. Journal of the American Chemical Society. 2016 Apr 20;138(15):5063-8.
24. Kumar D, Kashyap MK, La Clair JJ, Villa R, Spaanderman I, Chien S, et al. Selectivity in Small Molecule Splicing Modulation. ACS chemical biology. 2016 Oct 21;11(10):2716-23.
25. Burkart MD LCJ, Jones BD, Mandel A, Villa R, Castro JE, Kashyap MK, Kumar D, inventor University of California, assignee. Anti-cancer polyketide compounds. USA. 2017.
26. Ferreira PG JP, Rico D, Gómez-López G, Martínez-Trillos A, Villamor N,, Ecker S G-PA, Knowles DG, Monlong J, Johnson R, Quesada V, Djebali, S PP, López-Guerra M, Colomer D, Royo C, Cazorla M, Pinyol M, Clot G,, Aymerich M RM, Kulis M, Tamborero D, Gouin A, Blanc J, Gut M, Gut I,, Puente XS PD, Martin-Subero JI, López- Bigas N, López-Guillermo A,, Valencia A L-OC, Campo E, Guigó R.
Transcriptome characterization by RNA sequencing identifies a major molecular and clinical subdivision in chronic lymphocytic leukemia. Genome Res. 20 26.
27. Lappalainen I, Almeida-King J, Kumanduri V, Senf A, Spalding JD, Ur-Rehman S, et al. The European Genome-phenome Archive of human data consented for biomedical research. Nat Genet. 2015 Jul;47(7):692-5.
28. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Aug;25(16):207 9.
29. Anders S, Reyes A, Huber W. Detecting differential usage of exons from RNA- seq data. Genome Res. 2012 Oct;22(10):2008-17.
30. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 2009 May;25(9):1105-11.
31. Frazee AC, Pertea G, Jaffe AE, Langmead B, Salzberg SL, Leek JT.
Ballgown bridges the gap between transcriptome assembly and expression analysis. Nat Biotechnol. 2015 Mar;33(3):243-6.
32. Vera Alvarez R PL, Mariño-Ramírez L, Landsman D. TPMCalculator: one-step software to quantify mRNA abundance of genomic features. Bioinformatics. 2019;35(11):1960-2.
33. Mi H, Lazareva-Ulitsky B, Loo R, Kejariwal A, Vandergriff J, Rabkin S, et al. The PANTHER database of protein families, subfamilies, functions and pathways. Nucleic Acids Res. 2005 Jan;33(Database issue):D284-8.
34. Mi H, Poudel S, Muruganujan A, Casagrande JT, Thomas PD. PANTHER version 10: expanded protein families and functions, and analysis tools. Nucleic A Jan;44(D1):D336-42.
35. Wang J, Duncan D, Shi Z, Zhang B. WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): update 2013. Nucleic Acids Res. 2013 Jul;41(Web Server issue):W77- 83.
36. Fabregat A, Sidiropoulos K, Garapati P, Gillespie M, Hausmann K, Haw R, et al. The Reactome pathway Knowledgebase. Nucleic Acids Res. 2016 Jan;44(D1):D481- 7.
37. Chuang HY, Rassenti L, Salcedo M, Licon K, Kohlmann A, Haferlach T, et al. Subnetwork-based analysis of chronic lymphocytic leukemia identifies pathways that associate with disease progression. Blood. 2012 Sep 27;120(13):2639-49.
38. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, et al. Integration of biological networks and gene expression data using Cytoscape. Nat Protoc. 2007;2(10):2366-82.

39. Rassenti LZ, Huynh L, Toy TL, Chen L, Keating MJ, Gribben JG, et al. ZAP-70 compared with immunoglobulin heavy-chain gene mutation status as a predictor of disease progression in chronic lymphocytic leukemia. N Engl J Med. 2004 Aug;351(9):893-901.
40. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA. 2000 Dec;284(23):3043-5.
41. Schwaederle M, Ghia E, Rassenti LZ, Obara M, Dell' Aquila ML, Fecteau JF, et al. Subclonal evolution involving SF3B1 mutations in chronic lymphocytic leukemia. Leukemia. 2013 Apr;27(5):1214-7.
42. Lumb S, Fleischer SJ, Wiedemann A, Daridon C, Maloney A, Shock A, et al. Engagement of CD22 on B cells with the monoclonal antibody epratuzumab stimulates the phosphorylation of upstream inhibitory signals of the B cell Commun Signal. 2016 Jun;10(2):143-51.
43. Talab F, Allen JC, Thompson V, Lin K, Slupsky JR. LCK is an important mediator of B-cell receptor signaling in chronic lymphocytic leukemia cells. Mol Cancer Res. 2013 May;11(5):541-54.
44. Mayr C, Bund D, Schlee M, Moosmann A, Kofler DM, Hallek M, et al. Fibromodulin as a novel tumor-associated antigen (TAA) in chronic lymphocytic leukemia (CLL), which allows expansion of specific CD8+ autologous T lymphocytes. Blood. 2005 Feb;105(4):1566-73.
45. Mikaelsson E, Danesh-Manesh AH, Lüppert A, Jeddi-Tehrani M, Rezvany MR, Sharifian RA, et al. Fibromodulin, an extracellular matrix protein: characterization of its unique gene and protein expression in B-cell chronic lymphocytic leukemia and mantle cell lymphoma. Blood. 2005 Jun;105(12):4828-35.
46. Wang L, Lawrence MS, Wan Y, Stojanov P, Sougnez C, Stevenson K, et al. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. The New England journal of medicine. 2011 Dec 29;365(26):2497-506.
47. Girard C, Will CL, Peng J, Makarov EM, Kastner B, Lemm I, et al. Post- transcriptional spliceosomes are retained in nuclear speckles until splicing completio Commun. 2012;3:994.
48. Quesada V, Conde L, Villamor N, Ordóñez GR, Jares P, Bassaganyas L, et al. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. Nat Genet. 2012 Jan;44(1):47-52.
49. Wang L BA, Fan J, Wan Y, Gambe R, Li S, Hergert S, Yin S, Freeman SS,, Levin JZ FL, Seiler M, Buonamici S, Smith PG, Chau KF, Cibulskis CL, Zhang, W RL, Ghia EM, Kipps TJ, Fernandes S, Bloch DB, Kotliar D, Landau DA,, Shukla SA AJ, Reed R, DeLuca DS, Brown JR, Neuberg D, Getz G, Livak KJ,, Meyerson MM KP, Wu CJ. Transcriptomic Characterization of SF3B1 Mutation Reveals Its Pleiotropic Effects in Chronic Lymphocytic Leukemia. Cancer Cell. 2016;30(5):750-63.
50. Wan Y WC. SF3B1 mutations in chronic lymphocytic leukemia. Blood. 2013;23(23):4627-34.
51. Hsu TY, Simon LM, Neill NJ, Marcotte R, Sayad A, Bland CS, et al. The spliceosome is a therapeutic vulnerability in MYC-driven cancer. Nature. 2015 Sep;525(7569):384-8.
52. Wong JJ AA, Ritchie W, Rasko JE. Intron retention in mRNA: No longer nonsense: Known and putative roles of intron retention in normal and disease biolo 2016;38(1):41-9.
53. Jung H LD, Lee J, Park D, Kim YJ, Park WY, Hong D, Park PJ, Lee E. Intron retention is a widespread mechanism of tumor-suppressor inactivation. Nat Genet. 2015 2015 Nov;47(11):1242-8.
54. Newman JRB CA, Mika M, New FN, Onengut-Gumuscu S, Atkinson MA, Rich SS, McIntyre LM, Concannon P. Disease-specific biases in alternative splicing and tissue-specific dysregulation revealed by multitissue profiling of lymphocyte gene expression in type 1 diabetes. Genome Res 2017;27(11):1807-15.
55. Kim T KJ, Oh JG, Hong SE, Kim DH. Pressure-overload cardiac hypertrophy is associated with distinct alternative splicing due to altered expression of splicing factors. Mol Cells. 2014;37(1):81-7.
56. Wang Q CE, Manley JL, Rio DC. Widespread intron retention impairs protein homeostasis in C9orf72 ALS brains. Genome Res. 2020;30(12):1705-15.
57. Amit M, Donyo M, Hollander D, Goren A, Kim E, Gelfman S, et al. Differential GC content between exons and introns establishes distinct strategies of splice-site recognition. Cell Rep. 2012 May;1(5):543-56.
58. Wagner GP, Kin K, Lynch VJ. Measurement of mRNA abundance using RNA- seq data: RPKM measure is inconsistent among samples. Theory Biosci. 2012 Dec;131(4):281-5.
59. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011 Aug;12:323.
60. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, et al. A survey of best practices for RNA-seq data analysis. Genome Biol. 2016 Jan;17:13.
61. Zhou J, Lai PB, Tsui SK. Identification of a non-coding KLF4 transcript generated from intron retention and downregulated in human hepatocellular carcinoma. Int J Oncol. 2015 Oct;47(4):1554-62.
62. Cho V, Mei Y, Sanny A, Chan S, Enders A, Bertram EM, et al. The RNA-binding protein hnRNPLL induces a T cell alternative splicing program delineated by differential intron retention in polyadenylated RNA. Genome Biol. 201
63. Pimentel H PM, Gee SL, Mohandas N, Pachter L, Conboy JG. A dynamic intron retention program enriched in RNA processing genes regulates gene expression during terminal erythropoiesis. Nucleic Acids Res. 2016;44(2):838-51.
64. Gattazzo C, Martini V, Frezzato F, Trimarco V, Tibaldi E, Castelli M, et al. Cortactin, another player in the Lyn signaling pathway, is over-expressed and alternatively spliced in leukemic cells from patients with B-cell chronic lymphocytic leukemia. Haematologica. 2014 Jun;99(6):1069-77.
65. McCarthy BA YS, Tipping M, Yan XJ, Wang XP, Bennett F, Li W, Lesser M, Paul S, Boyle E, Moreno C, Catera R, Messmer BT, Cutrona G, Ferrarini M, Kolitz JE, Allen SL, Rai KR, Rawstron AC, Chiorazzi N. A seven-gene expression panel distinguishing clonal expansions of pre-leukemic and chronic lymphocytic leukemia B cells from normal B lymphocytes. Immunol Res. 2015;63(1- 3):90-100.

66. Mittal AK CN, Rohlfsen RA, Gupta P, Joshi AD, Hegde GV, Bociek RG, Joshi SS. Role of CTLA4 in the proliferation and survival of chronic lymphocytic leukemia. PLoS One. 2013;8(8):e70352.
67. Collins AV BD, Gilbert RJ, Iaboni A, Manso-Sancho R, Walse B, Stuart, DI vdMP, Davis SJ. The interaction properties of costimulatory molecules revisited. Immunity. 2002;,17(2):201-10.
68. Parry RV CJ, Frauwirth KA, Lanfranco AR, Braunstein I, Kobayashi SV, Linsley PS, Thompson CB, Riley JL. CTLA-4 and PD-1 receptors inhibit T-cell activation by distinct mechanisms. Mol Cell Biol. 2005;25(21):9543-53.
69. Rossi D, Bruscaggin A, Spina V, Rasi S, Khiabanian H, Messina M, et al. Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. Blood. 2011 Dec 22;118(26):6904- 8.
70. Kulis M HS, Bibikova M, Queirós AC, Navarro A, Clot G, Martínez-Trillos, A CG, Brun-Heath I, Pinyol M, Barberán-Soler S, Papasaikas P, Jares, P BS, Rico D, Ecker S, Rubio M, Royo R, Ho V, Klotzle B, Hernández L, Conde, L L-GM, Colomer D, Villamor N, Aymerich M, Rozman M, Bayes M, Gut M,, Gelpí JL OM, Fan JB, Quesada V, Puente XS, Pisano DG, Valencia A, López-, Guillermo A GI, López-Otín C, Campo E, Martín- Subero JI. Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. Nat Genet. 2012;44(11):1236-42.
71. Yang J QJ, Yao DM, Qian SX, Qian W, Lin J, Xiao GF, Wang CZ, Deng ZQ, Ma JC, Chen XX. SF3B1 mutation is a rare event in Chinese patients with acute and chronic myeloid leukemia. Clin Biochem. 2013 May;46(7-8):701-3.
72. Girard C WC, Peng J, Makarov EM, Kastner B, Lemm I, Urlaub H, Hartmuth K, Lührmann R. Post-transcriptional spliceosomes are retained in nuclear speckles until splicing completion. Nat Commun. 2012;3:994.
73. Cretu C SJ, Ponce-Salvatierra A, Dybkov O, De Laurentiis EI, Sharma K, Will CL, Urlaub H, Lührmann R, Pena V. Molecular Architecture of SF3b and Structural Consequences of Its Cancer-Related Mutations. Mol Cell. 2016;64(2):307-19.
74. Prabakaran S HM, Chauhan R, Winter D, Tweedie-Cullen RY, Dittrich C, Hong E, Gunawardena J, Steen H, Kreiman G, Steen JA. Quantitative profiling of peptides from RNAs classified as noncoding. Nat Commun. 2014;5:5429.

Figures

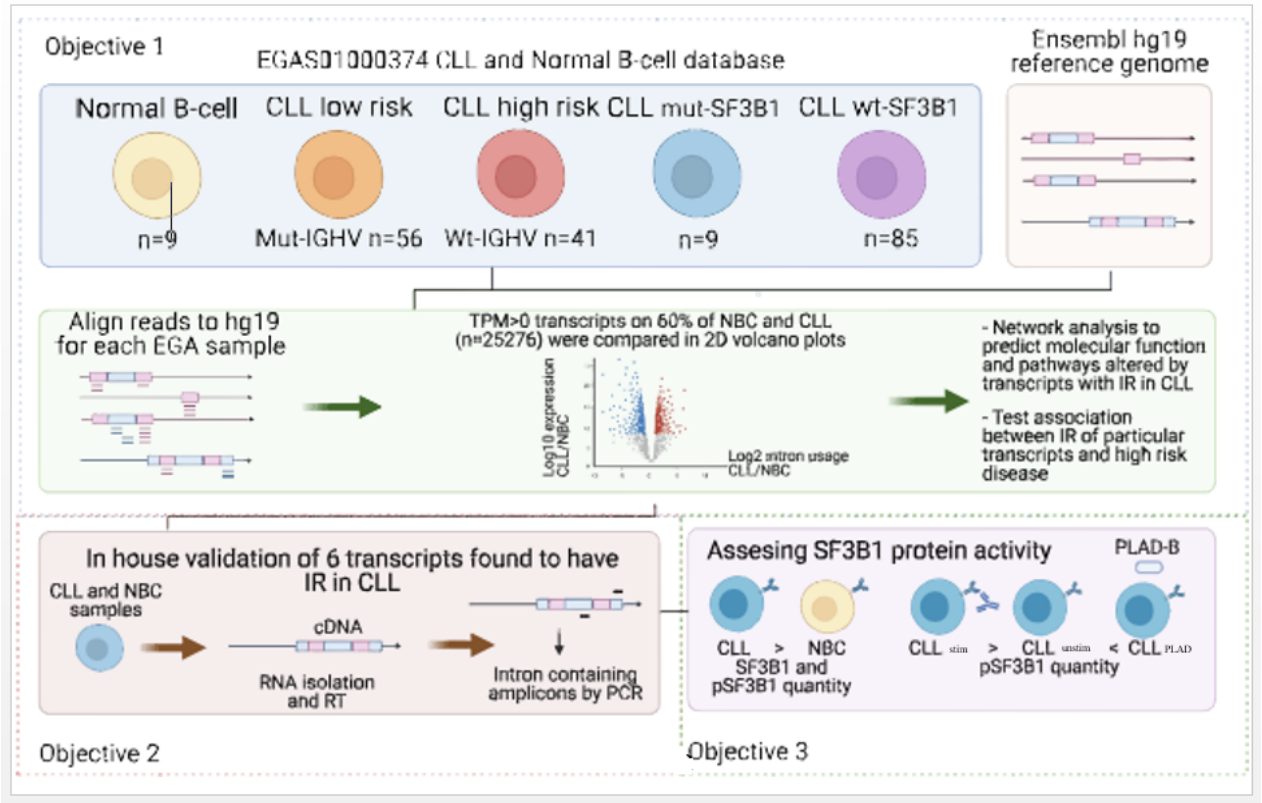

Figure 1

The workflow for RNA-Seq Analysis for IR Study in CLL vs NBC For studying the global IR pattern between CLL vs NBC, the RNA-Seq data obtained from EGA was processed for analysis after alignment with the reference genome hg19. The transcripts with >0 TPM values were selected if those were detected in at least 60% of the population. A comparative analysis of IR between different CLL Cells subtypes vs NBC was done. Further biological pathways and molecular function were studies using PANTHER. A SF3B1 centric sub-network was constructed to study if there are involvements of genes/transcripts involve in RNA splicing. A selected number of transcripts to show IR was validated using cDNA derived from CLL cells and NBC (synthesized from total RNA treated with DNaseI to remove the DNA contamination) after subjecting to RT-PCR followed by agarose gel electrophoresis. SF3B1 and pSF3B1 profiling was done after IgM stimulation of CLL and NBC. Further, the effect of macrolide PLAD-B and fludarabine (F-ara-A) was studied *in vitro* on CLL cells after treatment for

different time-points. The protein was isolated after cell lysis and after SDS-PAGE, the protein was transferred on the membrane for detection of SF3B1, pSF3B1, and loading control (β-actin).

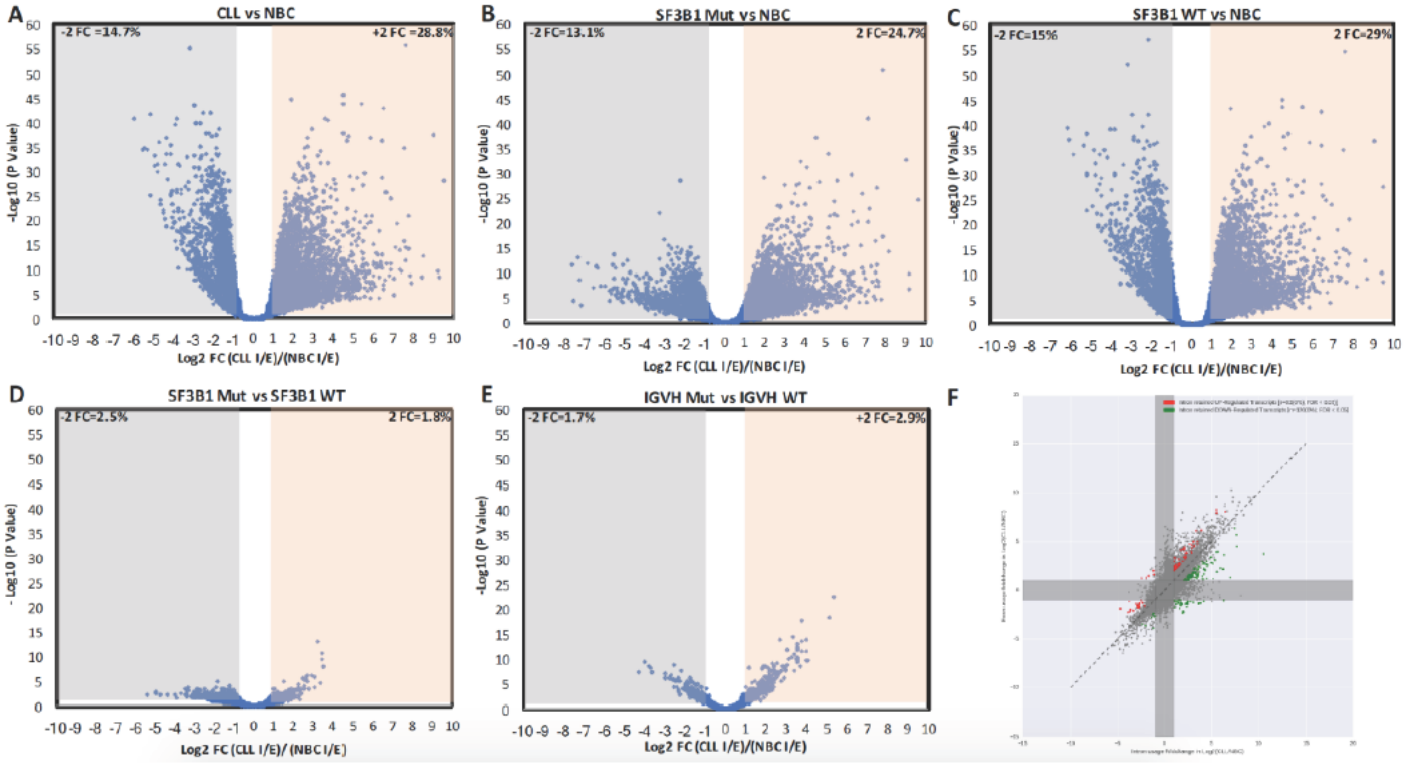

Figure 2

RNA sequencing analysis for global intron retention in CLL versus normal B cells and correlation with transcript expression

RNA-Seq analysis was conducted on sequences obtained from the EGA data derived from CLL samples (n=97 including high risk=41, low risk=56) and normal B cells (n=9 from healthy controls in triplicate). **(A-E)** Volcano plot displaying the differential intron retention log2 ratios TPM (intron/exon) between **(A)** CLL-B cells and NBC (CLL-B vs NBC). The scatter plot shows 14811 transcripts for intron retention between CLL-B cells versus NBC. **(B)** CLL-B cells carrying Mut-SF3B1 and NBC (Mut-SF3B1 vs NBC). **(C)** CLL-B cells carrying Wt-SF3B1 and NBC (Wt-SF3B1 vs NBC). **(D)** CLL-B cells carrying Mut-SF3B1 and CLL-B cells carrying Wt-SF3B1 (Mut-SF3B1 vs Wt-SF3B1). **(E)** CLL-B cells carrying Mut-IgV$_H$ and CLL-B cells carrying Wt-IgV$_H$ (Mut-IgV$_H$ vs Wt-IgV$_H$). The bottom black line

represents the diagonal regression line where intron retention ratios are equal both samples compared. **(F)** Further, the data was presented in form of a quadrant between CLL-B cells versus NBC to show the distribution of ratio of intronic and exonic TPM values for 14811 transcripts.

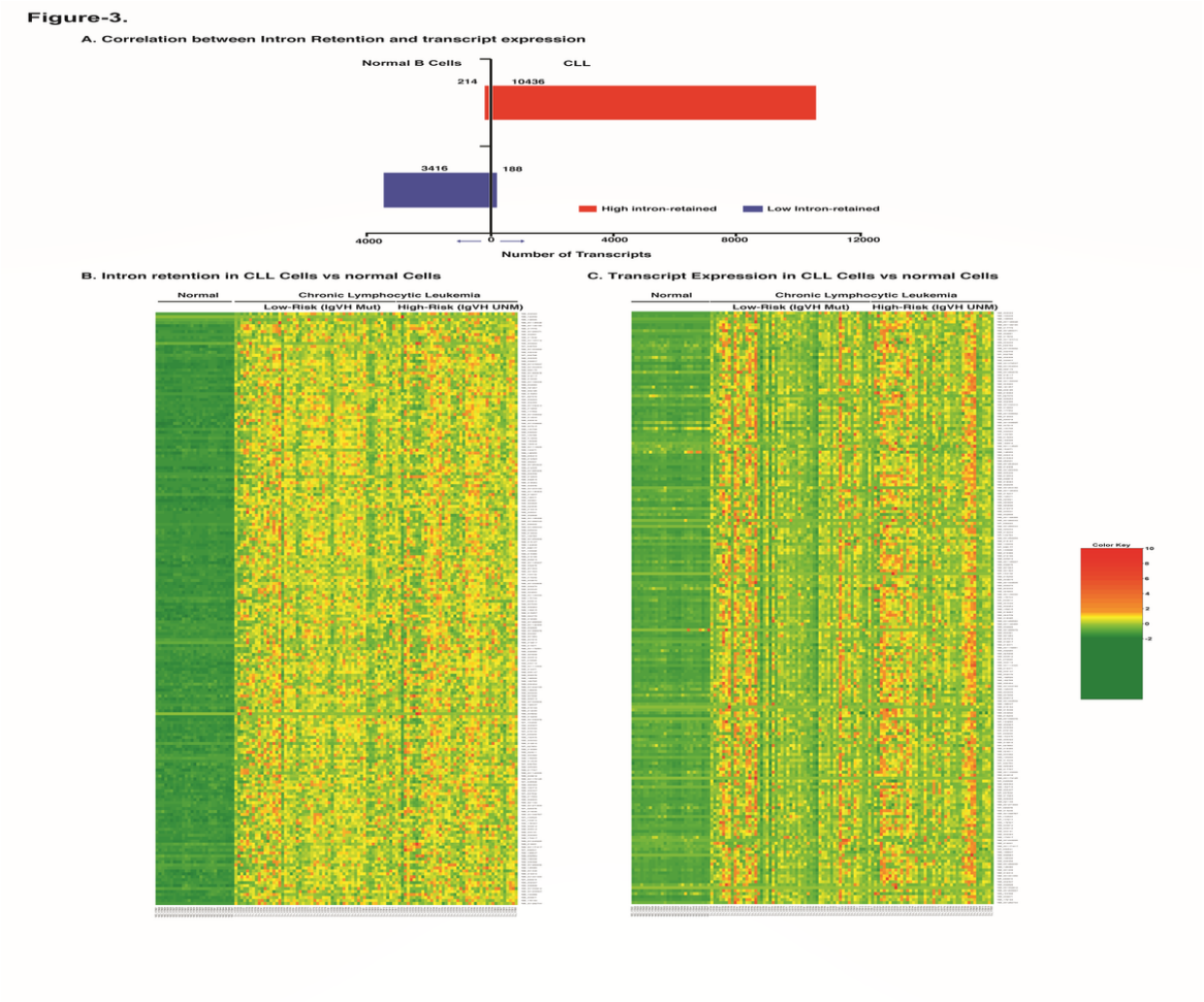

Figure 3

The correlation between intron retention and transcript expression in various cell types

(A) 16725 transcripts that were accumulated differentially used-intron respectively in CLL (set-I) and NBC (set-II) at p<=0.05 (FDR=5%) and among those the three subsets each of the set corresponding to over-expressed (set-IA, set-IIA), under-expressed (set-IB, set-IIB) and non-differentially-expressed (set- IC, set-IIC) transcripts consequently found in CLL and NBC at p<0.05 (FDR=5%). The Chi-square p-value for the 2x3 table is <=1.0e-300. The bar graph represents two sets of bars, each for a fraction of high intron-used in CLL (set-I) and NBC (set-II) cases. It is associated with upregulated (set-IA, set-IIA) and downregulated (set-IB, set-IIB) at transcription level. The 2x2 contingency table shows significant p-value < 1.0e-300 and ODDs=886. **(C)** The Heatmap Z-scores of TPM values for sample-wise intron retention in NBC and CLL from each of the top 200 transcripts selected based on high intron retention in CLL with fulfilling condition CLL > NBC at p<0.05; FDR=5%. **(D)** The Heatmap Z-scores of TPM values for sample-wise transcripts expression in NBC and CLL from each of the top 200 transcripts selected based on high intron retention in CLL with fulfilling condition CLL > NBC at p<0.05; FDR=5%.

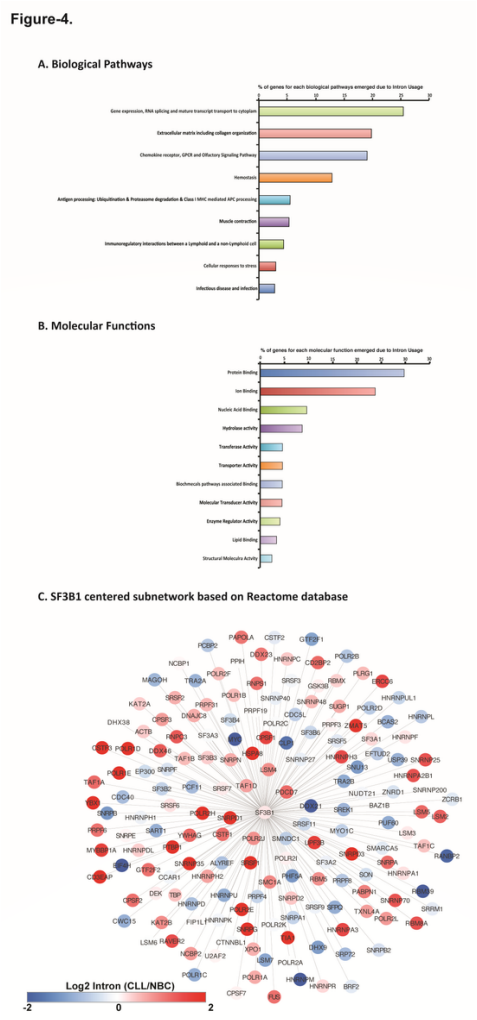

Figure 4

Biological Pathway, Molecular functions, and Network analysis based on intron retention

For biological pathway analysis, we took top 25% transcripts with high intron retention in CLL as compared with normal B cells (p<0.05, FDR=0.1). The gene symbols for corresponding transcripts were used as input and subjected for Reactome Pathway. The p-value was adjusted using a Bonferroni correction. The X-axis shows the % of the genes enriched and Y-axis shows the description of

the pathways. (**B)** Molecular function analysis was done for top 25% of transcripts with high intron retention CLL/NBC Network analysis by selecting transcripts with high intron retention in CLL as compared with normal B cells using WebGestalt. The height of the bar represents the % of genes observed in the category. **(C)** Subnetwork centered on SF3B1 and its first neighbors. The SF3B1 centered subnetwork was based on Reactome FI network database and cytoscape V3.4 was used to identify the first neighbors of SF3B1. Nodes and links represent genes and functional interactions, respectively. The color of each nodes scale with log2 intron retention ratio in CLL cells vs NBCs as indicated in the scale at the bottom.

Figure-5.

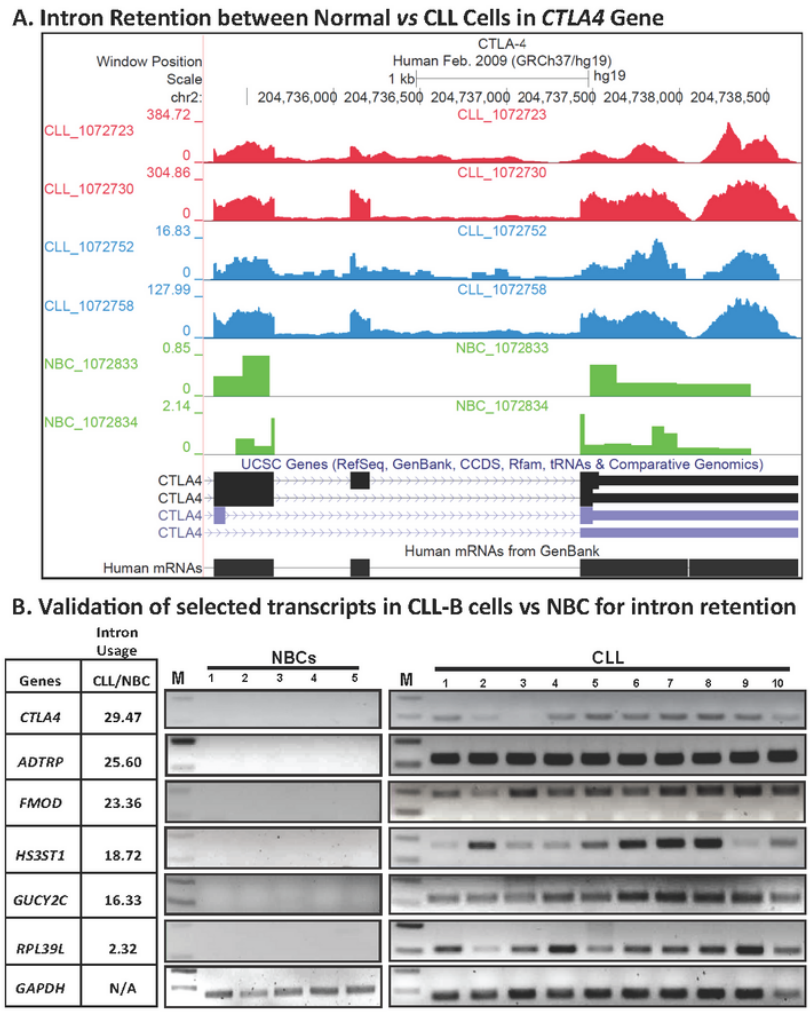

Figure 5

Selection and validation of transcripts for intron retention using RT-PCR

(A) RNA-Seq read mapping of the *CTLA4* gene from the UCSC reference genome (hg19), six representative samples belonging to high or low-risk CLL cells, or normal B cells are shown. The intron retention tracks are shown in green (for normal B cells), blue (low-risk CLL), and red (high-risk CLL) colors. It is clear that the reads map to the intronic region of *CTLA4* transcript as annotated in the UCSC database. The absolute read counts for each sample are indicated on the y-axis. **(B)** The TPM ratios of CLL/NBC for intron and exon are shown for the transcripts selected for validation. RNA was isolated from pure CLL-B cells or NBC and after DNase I digestion cDNA was prepared. Selected intronic region for assessment of intron retention were amplified using RT-PCR for *PPP2R5B*, *ADTRP*, *RPL39L*, *HS3ST1*, *CTLA4*, *FMOD*, and *GUCY2C* transcripts. *GAPDH* were used as controls for normalization/loading of RNA.

Figure-6.

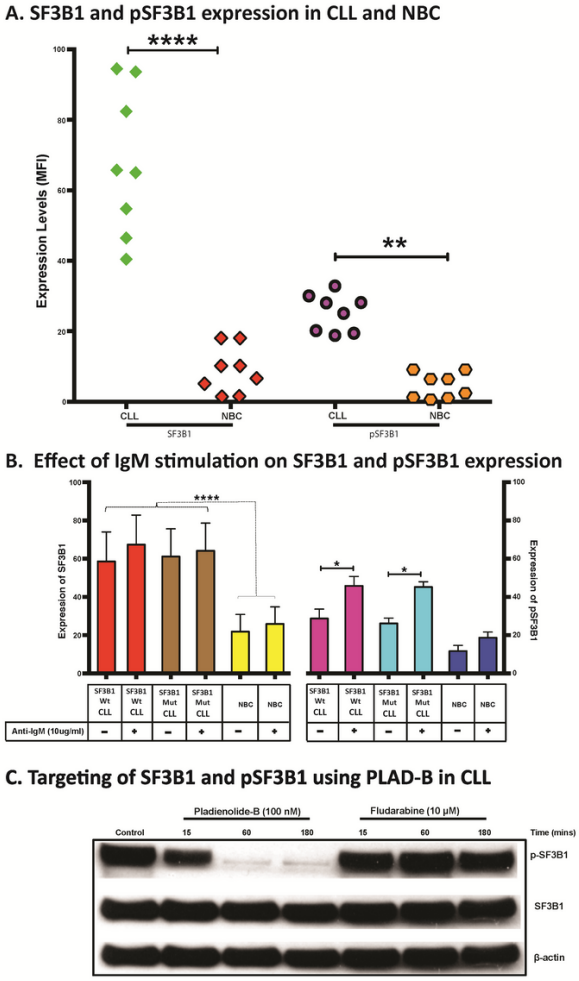

Figure 6

Effect of IgM stimulation on SF3B1 and pSF3B1 expression in normal and CLL cells

(A) Normal B cells or CLL-B cells were thawed and overnight incubated at 37^{o}C. Cells were surface stained using CD19/CD5 (CLL-B cells) or CD19/CD20 (Normal B cells) followed by intracellular staining using anti-SF3B1 and anti-

phospho-SF3B1 antibodies for total and phospho-SF3B1 protein using flow cytometry. **(B)** Normal B cells or CLL-B cells were thawed and overnight incubated at 37^{o}C. Cells were stimulated with Anti-IgM (10 µg/ml) for 15 minutes. Post-surface staining using CD19/CD5 (CLL-B cells) or CD19/CD20 (Normal B cells), cells were subjected for intracellular staining using anti-SF3B1 and anti- phospho-SF3B1 antibodies for total and phospho-SF3B1 using flow cytometry. All the samples were run in duplicate and the data is presented with the means and their respective SD. Statistical significance was determined by using Bonferroni correction test for multiple comparison test, where *, **, ***, **** represent $p < 0.05$; $p<0.01$, $p < 0.001$, and $p < 0.0001$ respectively. **(C)** A total of five million CLL-B cells were incubated overnight and treated with 100 nM of Pladienolide-B (PLAD-B) or 10 µM of Fludarabine (F-ara-A) for 15, 60 and 180 minutes. Post-incubation of splicing modulators or chemotherapy, cells were harvested and lysed using modified RIPA. A total of 200 µg of protein was run on SDS-PAGE and subjected for western blot. Antibodies against SF3B1, and phospho-SF3B1 were used for assessing total and phospho-SF3B1 levels. β- actin was used as a loading control, while cells incubated in media only were used as the negative control (-).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigure1.pdf
- SupplementaryFigure2.pdf
- SupplementaryFigures3.pdf

- SupplementaryFigures4.pdf