Targeting the spliceosome in chronic lymphocytic leukemia with the macrolides FD-895 and pladienolide-B

Manoj K. Kashyap,[1]* Deepak Kumar,[1]* Reymundo Villa,[2]* James J. La Clair,[2]* Chris Benner,[5] Roman Sasik,[4] Harrison Jones,[1] Emanuela M. Ghia,[1] Laura Z. Rassenti,[1,3] Thomas J. Kipps,[1,3] Michael D. Burkart,[2,] and Januario E. Castro[1,3]

[1]Moores Cancer Center; [2]Department of Chemistry and Biochemistry, [3]CLL Research Consortium, and Department of Medicine; [4]Center for Computational Biology, Institute for Genomic Medicine, University of California, San Diego, La Jolla; and [5]Integrative Genomics and Bioinformatics Core, Salk Institute for Biological Studies, La Jolla, CA, USA

*MKK, DK, RV and JJL contributed equally to this work.

Supplementary Methods and Data

Supplementary Methods

Compounds

Oligonucleotides were purchased from Integrated DNA Technologies. FD-895 was prepared by total synthesis in our laboratories.[1,2]

Cell culture: The leukemia cell lines Raji and Ramos (Burkitt's lymphoma, both mutant type *TP53*) and Jurkat (T-Cell leukemia, mutant *TP53*) were obtained from ATCC and maintained in RPMI supplemented with 10% fetal bovine serum (FBS), 2 mM L-glutamine, and 100 U/mL of penicillin and 100 μg/mL of streptomycin at 37°C in an atmosphere of 5% CO_2. A20 cell line (a BALB/c mouse B cell lymphoma line) was propagated in RPMI-1640 medium supplemented with 10% FBS and 0.05 mM 2-mercaptoethanol (ß-ME).

Chronic Lymphocytic Leukemia (CLL) sampling and cell culture: Peripheral blood mononuclear cells (PBMC) from patients with CLL were obtained from the CLL Research Consortium tissue bank. After CLL diagnosis was confirmed,[3] patients provided written informed consent for blood sample collection on a protocol approved by the Institutional Review Board of UCSD, in accordance with the Declaration of Helsinki.[4]

CLL cells and PBMCs were separated from heparinized venous blood by density gradient centrifugation using Ficoll-Hypaque media (GE Healthcare).[5] Samples with > 95% double positive cells for CD5 and CD19, as assessed by flow cytometry, were selected and used. CLL cells were maintained in RPMI 1640 medium supplemented with 10% FBS, 2 mM L-glutamine, and 100 U/mL of penicillin and 100 μg/mL of streptomycin at 37 °C in an atmosphere containing 5% CO_2.

Stromal cell support was provided using a co-culture system standardized in our laboratory. For that CLL cells were incubated with stroma -NK-tert cells using a 20 (CLL):1 (stroma -NK-tert) ratio. Stroma -NK-tert cells are fibroblast derived from human bone marrow and immortalized with human telomerase reverse transcriptase (hTERT) bearing exogene MFG-tsT-IRES-neo (RCB2350- RIKEN Bioresource Center, Japan).

Normal B cell isolation: Normal B cells were purified from buffy coats of healthy volunteer donors, obtained from San Diego blood bank. 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.

PI/DiOC$_6$ apoptosis assay *via* flow cytometry: Apoptotic and viable cells were discriminated by flow cytometry of cells stained with 40 μM 3,3'-dihexyloxacarbocyanine iodide (DiOC$_6$) (Life Technologies) and 15 μM propidium iodine (PI) (Sigma Aldrich).[6] Using this PI/DiOC$_6$ assay, viable cells excluded PI and stained brightly positive for DiOC$_6$ indicating viable mitochondria.

Calculation of % specific induced apoptosis (SIA): In order to discriminate the compound specific induced apoptosis vs. background spontaneous cell death from *in vitro* culture conditions, we calculated the percentage of specific induced apoptosis (% SIA) using the following formula: *% SIA = [(compound induced apoptosis – media only spontaneous apoptosis) / (100- media only spontaneous apoptosis)] × 100*.

Caspase colorimetric protease assay: CLL and normal B cells were treated with 100 nM FD-895 or 100 nM PLAD-B for 6 h. The whole cell protein was isolated and quantified according to the Bradford method.[7] Caspase activity was determined by using the ApoTarget caspase colorimetric protease assay kit (Life Technologies). An aliquot containing 100 μg of protein was used for each experiment.

Z-VAD assay for caspase inhibition: Z-VAD assay was performed based on the method reportedly previously.[8]

Western blot analysis: CLL cells were treated with 100 nM FD-895, 100 nM PLAD-B or 10 μM F-ara-A for 6 & 24 h for Mcl-1 and PARP, respectively. Untreated CLL cells were used as a control. A sample containing 30 μg of total protein was 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 polyclonal anti-Mcl-1 at dilution 1:200 (Catalog # sc-819, Santa Cruz Biotechnology), mouse anti-PARP at dilution 1:1000 (Catalog # 556362, BD Pharmingen), and mouse anti-β-actin at dilution 1:1000 (Catalog # 4967, BD Pharmingen). 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. Protein-antibody complexes signals were detected by exposing the X-ray films after treatment with enhanced chemiluminescence (ECL) kit (Pierce Thermo Scientific).

Reverse transcriptase PCR (RT-PCR) analysis: CLL or normal B cells (10^6 cells/well) were treated with FD-895 (0.01-1.0 µM), PLAD-B (0.01-1.0 µM), bendamustine (33-100 µM), or F-ara-A (3-33 µM) over a specified period of time (15 min, 30 min, 60 min, 120 min, or 240 min). RNA isolation and cDNA preparation, PCR reaction was performed in 25 µl of reaction volume as previously described.[1] PCR conditions were 94 °C for 3 min; 35 cycles of 94 °C for 30 s, 55 - 60 °C (60 °C for *RIOK3*, 58 °C for *DNAJB1,* 58 °C for *GAPDH,* 58 °C for *MCL-1* or 55 °C for *BCL-X*) for 30 s, and 72 °C for 1 min; followed by 72 °C for 5 min. PCR products were separated on a 2% agarose gel and stained with ethidium bromide.

Quantitative RT-PCR (qRT-PCR) analysis: CLL or normal B cells (5×10^6 cells/well) were treated with FD-895 (0.01-1.0 µM), PLAD-B (0.01-1.0 µM), or F-ara-A (10 µM) for 15 min, 60 min, or 240 min. Total RNA was extracted using mirVana miRNA isolation kit (Life Technologies) as described previously.[1] The amount of unspliced mRNA for *DNAJB1*, and *RIOK3* was determined using Power SYBR Green PCR master mix (Applied Biosystems) real-time qRT-PCR using specific primers designed for detection of the intron of each gene (***Supplementary Table 1***). PCR using 0.2 µM of each primer was performed on 20 ng of the obtained cDNA. PCR conditions were 50 °C for 2 min; 95 °C for 10 min; 35 cycles of 94 °C for 20 s, 55 °C for 20 s, followed by 72 °C for 30 s. The mRNA levels were calculated using $2^{-\Delta\Delta CT}$ method.[9] *GAPDH* was used as a control for normalization. Details of the primers used for qRT-PCR are provided in ***Supplementary Table 1***.

RNA-seq analysis: RNA was isolated using mirVana miRNA isolation kit (Life Technologies, Grand Island, NY USA) from CLL and normal B cell samples after incubation for 2 h with or without FD-895 (100nM), PLAD-B (100nM) or F-ara-A (10uM).[1] Samples from two CLL patients (wild type for *TP53* and *SF3B1*), one male and one female were examined. The RNA was treated with DNase on column. Purified RNA was checked for quality control (QC) and sent for RNA sequencing (RNA-seq) to Otogenetics Corporation.

RNA-seq was performed using HiSeq reading (Illumna) at Otogenetics Corporation (Norcross, GA). The reads generated, reads efficiency, % alignment, % unique match, total reads in

analysis, total positions in analysis, reads per bp, average reads/position, and GC-content have been described for each sample in the ***Supplementary Table-2***.

Specifically, the sequences from Illumina HiSeq were sequenced as 100 bp paired-end reads and aligned to the human transcriptome (hg19, GRCh37) using the *STAR* aligner,[10] and the gene-level exon and intron FPKM levels, including intron retention ratios, were found using *HOMER.*[11] RNA-Seq density plots were generated using *HOMER* and visualized using the UCSC Genome Browser. Gene Ontology enrichment analysis was performed using *DAVID.*[12] Clustering and Heatmaps were produced using *Cluster 3.0* and *TreeView.*[13] A partial list derived from this analysis showing top 50 genes with intron retention is mentioned in ***Supplementary Table 3.***

SF3B1 mutational status:

Mutation for *SF3B1* was assessed by sequencing exons 14, 15, 16, and 18 of *SF3B1* in the genomic DNA as reported previously.[14] The details of the CLL *SF3B1* wild and *SF3B1* mutant are mentioned in ***Supplementary Table 4***.

***In vivo* study using A20 lymphoma bearing BALB/c mice:**

Six to eight week-old female BALB/c mice were obtained from Jackson Laboratory, Bar Harbor, Maine. All animal experiments were performed under an animal protocol approved by the Institutional Animal Care and Use Committee of the Moores Cancer Center, University of California, San Diego. A20 cells were injected subcutaneously (SC) in the right flank. The tumors were allowed to grow approximately a diameter of 0.5cm. In all cases 6 animals per treatment group were used. Mice were monitored daily for clinical signs of toxicity, and tumor growth was measured 3 times per week with a caliper using the following formula: volume = length × width**²** × 0.5. Animals were sacrificed when the total tumor volume of the tumors reached $\geq$ 4,000 mm3 (4 cm³). The treatment phase initiated when the tumors reached 5 to 7 mm in diameter, and it was administered by intraperitoneal (IP) injection daily for 5 consecutive days. Treatment cohorts included the following: (1) DMSO vehicle, (2) Dexamethasone (5mg/Kg/Day), (3) PLAD-B, low dose (3mg/Kg/day) or high dose (10mg/Kg/day).

Statistical analysis: Data was analyzed using Prism (GraphPad Software Inc.). IC50 values were calculated by fitting sigmoid dose-response curves with 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 $p<0.05$; $p<0.01$; and $p<0.0001$, respectively.

References

1. Villa R, Kashyap MK, Kumar D, et al. Stabilized Cyclopropane Analogs of the Splicing Inhibitor FD-895. *J Med Chem.* 2013;56(17):6576-6582.
2. Villa R, Mandel AL, Jones BD, La Clair JJ, Burkart MD. Structure of FD-895 revealed through total synthesis. *Org Lett.* 2012;14(21):5396-5399.
3. Matutes E, Owusu-Ankomah K, Morilla R, et al. The immunological profile of B-cell disorders and proposal of a scoring system for the diagnosis of CLL. *Leukemia.* 1994;8(10):1640-1645.
4. WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects. [http://www.wma.net/en/30publications/10policies/b3/index.html%5D].
5. Kato K, Cantwell MJ, Sharma S, Kipps TJ. Gene transfer of CD40-ligand induces autologous immune recognition of chronic lymphocytic leukemia B cells. *J Clin Invest.* 1998;101(5):1133-1141.
6. Castro JE, Prada CE, Aguillon RA, et al. Thymidine-phosphorothioate oligonucleotides induce activation and apoptosis of CLL cells independently of CpG motifs or BCL-2 gene interference. *Leukemia.* 2006;20(4):680-688.
7. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem.* 1976;72:248-254.
8. Zhang S, Wu CC, Fecteau JF, et al. Targeting chronic lymphocytic leukemia cells with a humanized monoclonal antibody specific for CD44. *Proc Natl Acad Sci U S A.* 2013;110(15):6127-6132.
9. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods.* 2001;25(4):402-408.
10. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29(1):15-21.
11. Heinz S, Benner C, Spann N, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell.* 2010;38(4):576-589.
12. Huang dW, Sherman BT, Tan Q, et al. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res.* 2007;35(Web Server issue):W169-175.
13. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A.* 1998;95(25):14863-14868.
14. Schwaederlé M, Ghia E, Rassenti LZ, et al. Subclonal evolution involving SF3B1 mutations in chronic lymphocytic leukemia. *Leukemia.* 2013;27(5):1214-1217.
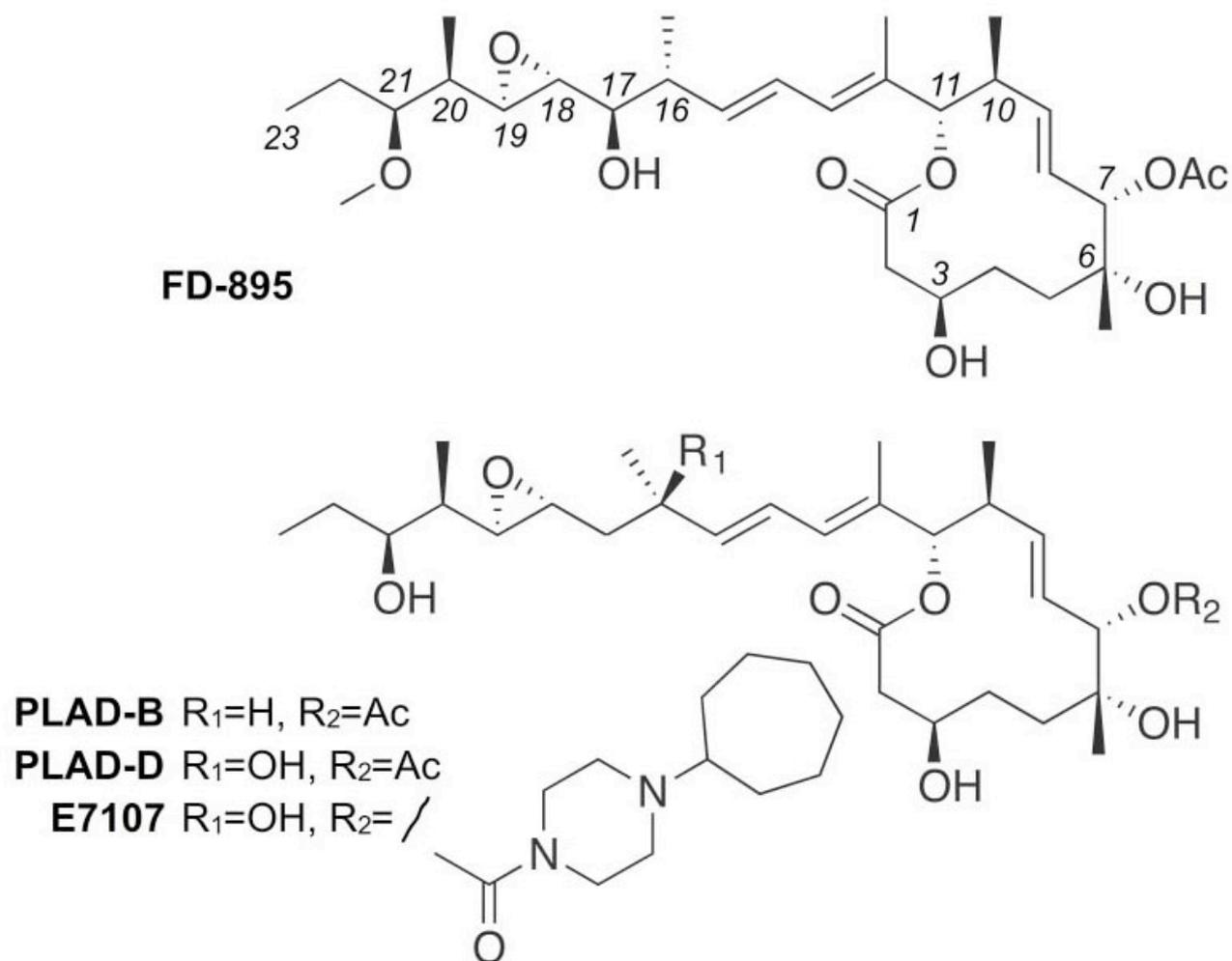
Supplementary Figure S1. Structures of FD-895, pladienolide B, pladienolide D, and E7107
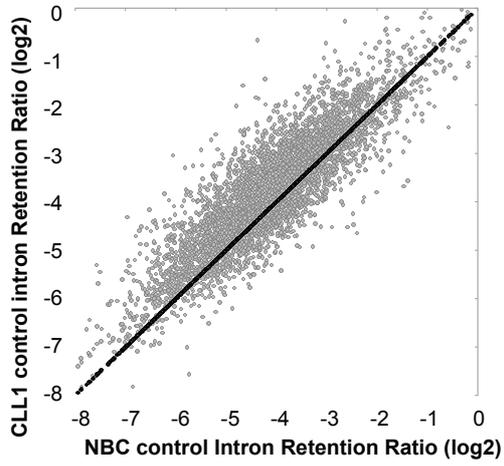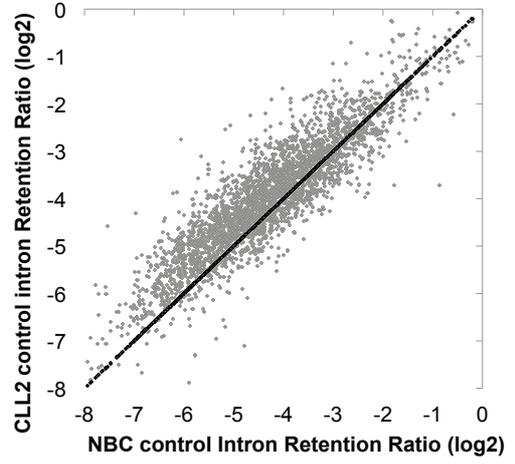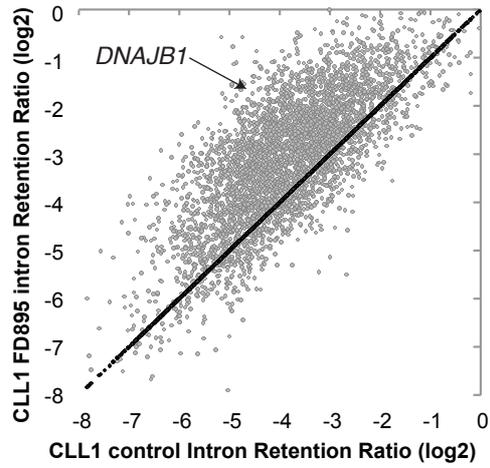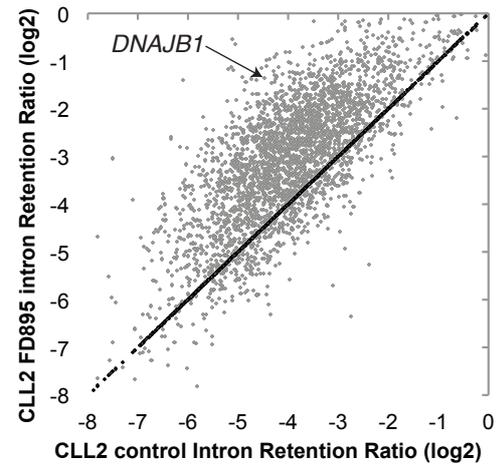
Supplementary Figure S2: RNA Sequencing Analysis in samples treated with FD-895.

Comparison of IR log2 ratios (FPKM intron / FPKM exon) in untreated CLL1 and untreated NBC (A) and CLL2 and untreated NBC (B). The black line represents the diagonal where IR ratios are equal in both samples. The comparison of IR log2 ratios in CLL1 control and FD-895 treated samples (C) and CLL2 control and FD-895 treated samples (D). The black line represents the diagonal where IR ratios are equal in both samples.
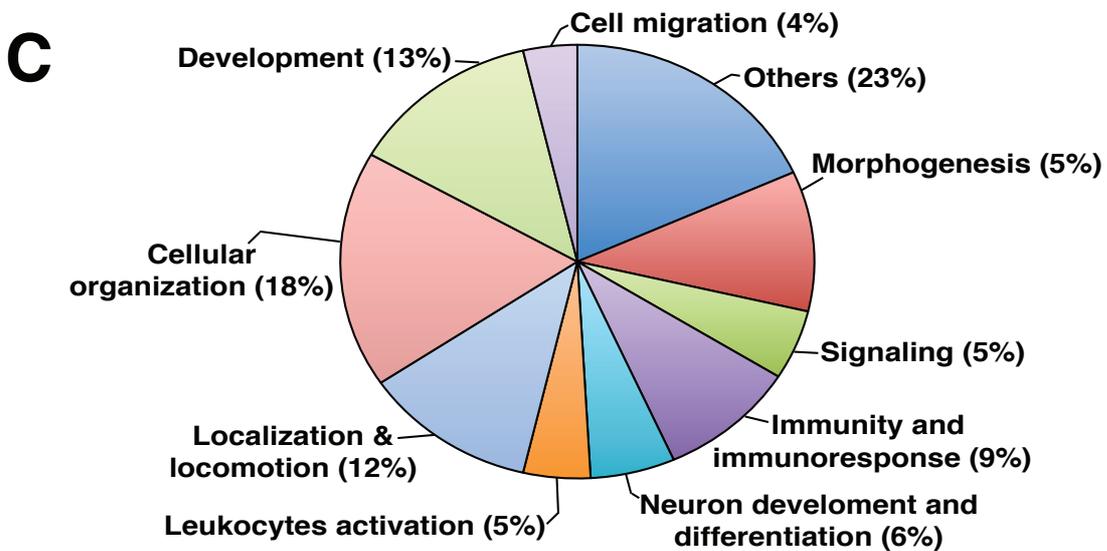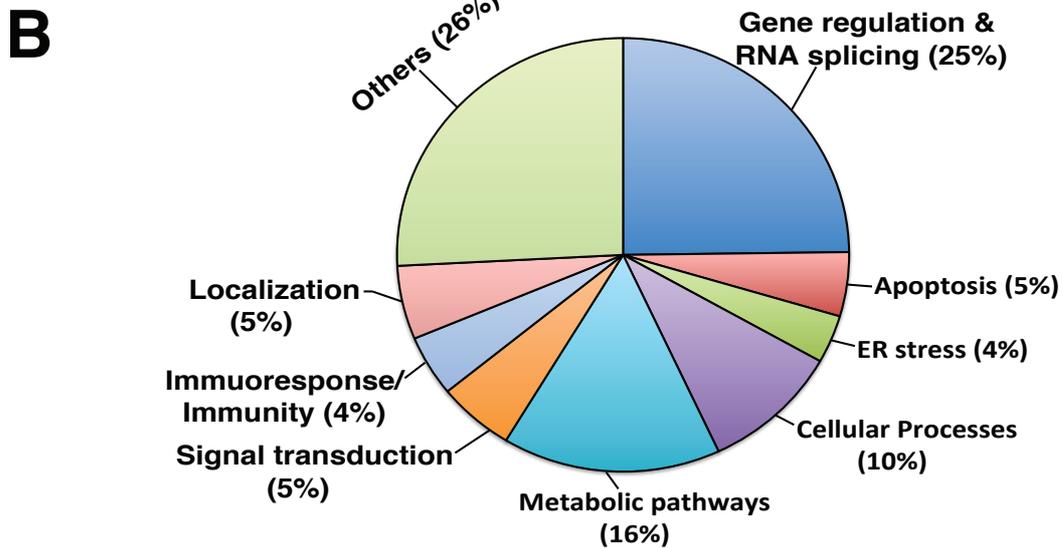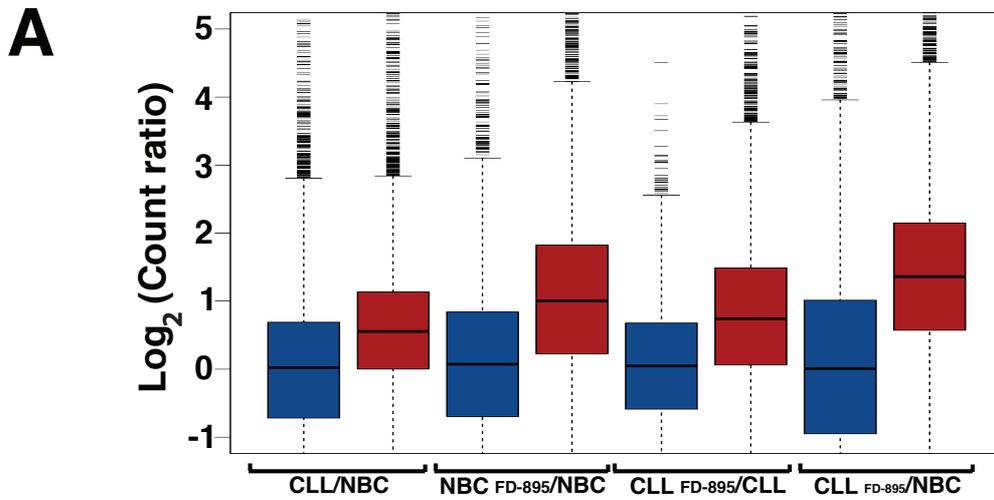
Supplementary Figure S3. Induction of intron retention (IR) by FD-895 in NBC And CLL cells and Gene Ontology analysis in affected biological pathways.

Panel-A: Boxplot showing the ratio of RNA-seq reads of canonical exon transcripts containing no intronic sequence (blue), and intron containing reads (red), measured in fragments per kilobase per million fragments mapped (FPKM). The ratios were established to compare samples from either NBC or CLL cells treated as indicated. **Panel-B:** Shows highly significant (Bonferroni-adjusted $p < 0.01$) biological processes affected by intron retention (IR) in FD-895 treated CLL cells compared to untreated CLL cells. The GO analysis indicates that FD-895 induces IR in genes that belong to cancer related pathways linked to mRNA splicing, protein targeting to ER, apoptosis and metabolic pathways among others. **Panel-C**: Highly significant (Bonferroni-adjusted $p < 0.01$) biological processes affected by IR in untreated CLL cells compared to untreated normal B cells. The gene ontology (GO) analysis indicates that compared to NBC, untreated CLL cells show base line IR abnormalities in pathways related to cellular development, organization, localization & locomotion among others.
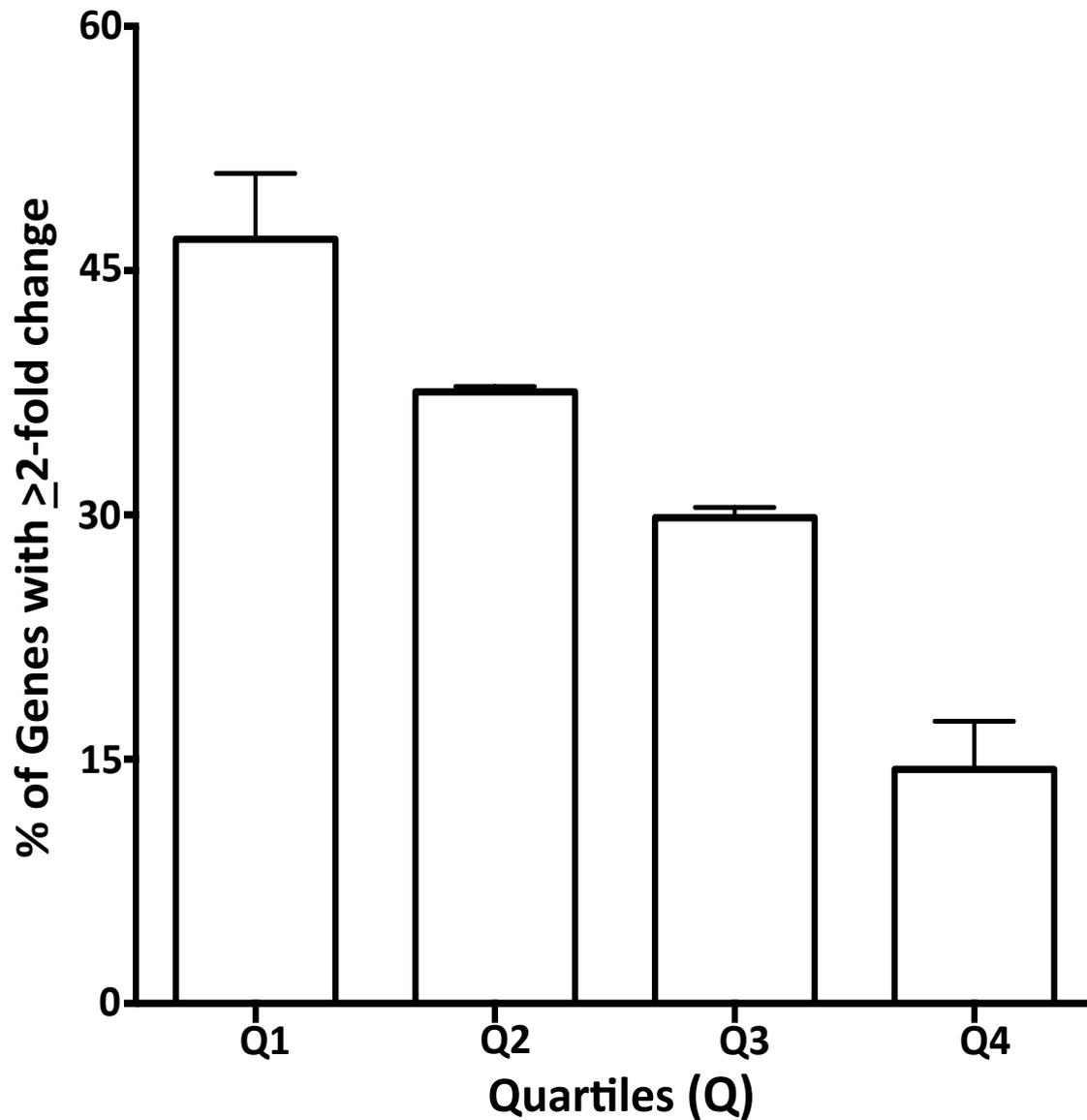
Quartiles based on IR baseline pre-treatment (CLL/NBC):

Quartile	IR fold change CLL/NBC	% of genes with >2 fold change after treatment with FD-895
Q1	0-0.957	46.94%
Q2	0.958-1.224	37.56%
Q3	1.225-1.611	29.84%
Q4	1.612-16.315	14.38%

Supplementary Figure S4. Quartile analysis for assessment of change IR after FD-895 treatment in CLL

Using the RNA sequence data presented in *Figure 1A*, we calculated a fold change of IR for genes found in CLL compared to a control sample from NBC. These fold change of IR (CLL vs. NBC) data was divided in quartiles using the ranges shown in this table. Then, we examined the percentage of genes from each quartile that showed a ≥ 2 fold increase in IR after treatment with FD-895. We found that the % of genes with ≥ 2 fold IR increase after treatment with FD-895 was inversely proportional to the base line IR ratio, with more genes that fell in this category in Q1 (lowest IR base line) compared with the other quartiles. The figure represents the Mean+/S.D. value drawn from both two independent CLL samples (CLL1 and CLL2). Y-axis represents the % genes showing ≥ 2-fold IR increase and the x-axis shows each one of the quartiles (Q) - Q1-4.
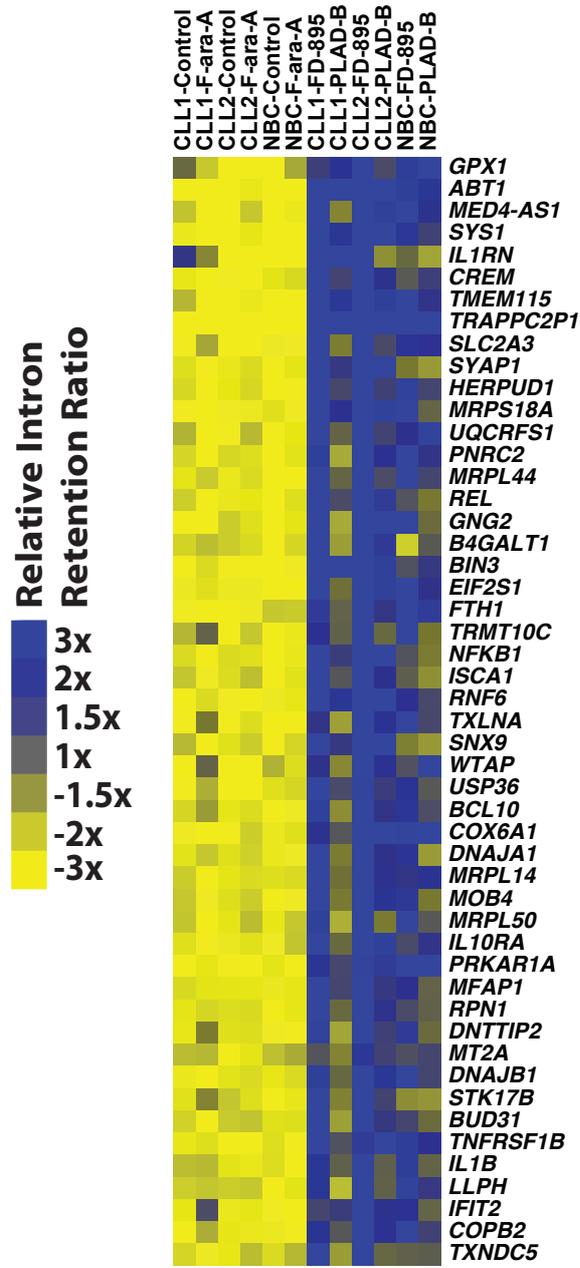
Supplementary Figure S5. Hierarchical clustering for 50 genes in the heat map depicting the relative IR log2 ratios

The heat map shows 50 genes with intron retention after FD-895 treatment. These genes belong to pathway representing RNA splicing, gene regulation, signal transduction, ER stress and apoptosis.

A

5 kb | ———————————————————— | hg19

CLL1-Control
CLL2-Control
NBC-Control
CLL1-FD895
CLL2-FD895
NBC-FD895

Relative RNA-Seq Read Density

GAPDH

B

20 kb | ———————————————————— | hg19

CLL1-Control
CLL2-Control
NBC-Control
CLL1-FD895
CLL2-FD895
NBC-FD895

Relative RNA-Seq Read Density

RIOK3

C

5 kb | ———————————————————— | hg19

CLL1-Control
CLL2-Control
NBC-Control
CLL1-FD895
CLL2-FD895
NBC-FD895

Relative RNA-Seq Read Density

MCL1

16
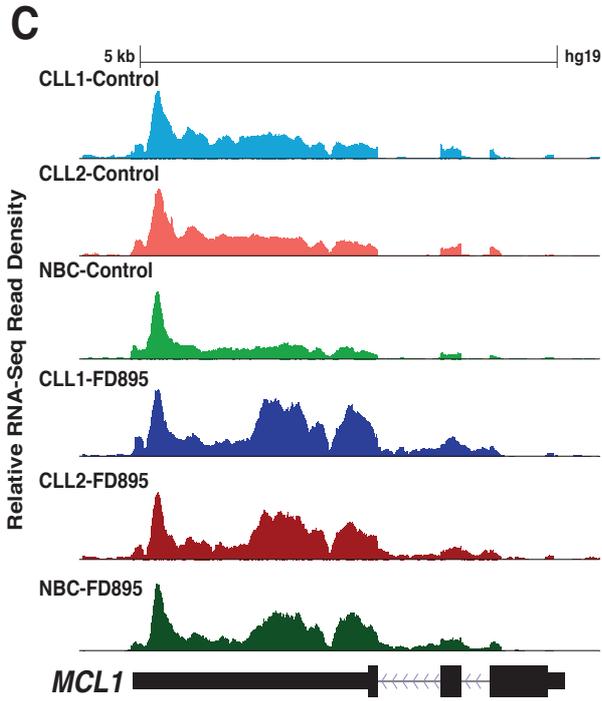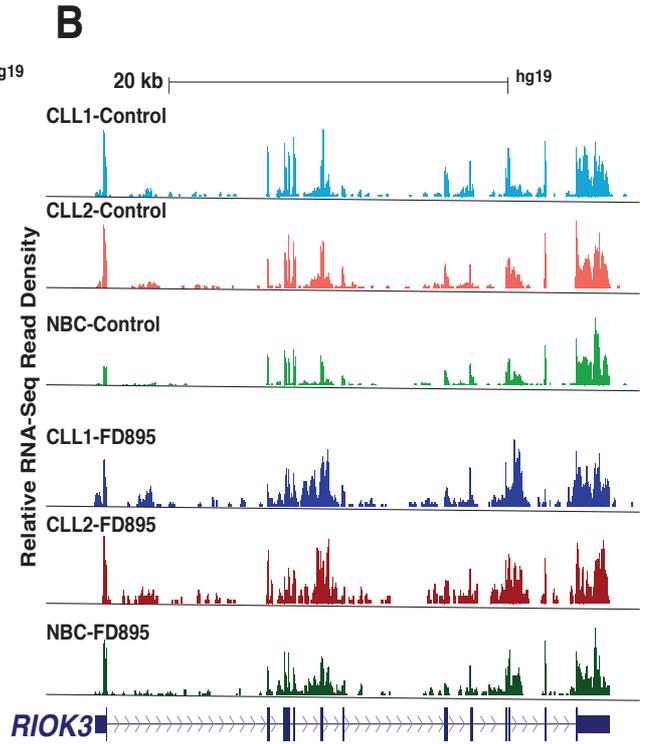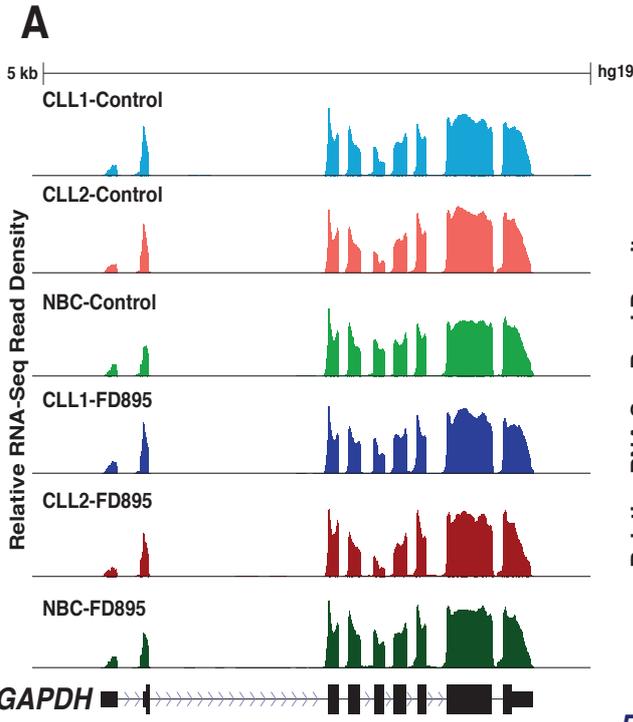
Supplementary Figure S6. Screenshot of *GAPDH*, *RIOK3*, and *MCL1* genes showing relative RNA-seq read density mapping to the UCSC reference genome (hg19) to show effect on intron retention after two hrs of treatment with FD-895 (100nM).

Panel A: The RNA-seq reads mapping of house keeping gene *GAPDH* to show effect of spliceosome modulator on intron retention between different treatments. **Panel B:** The RNA-seq reads mapping of RIOK kinase 3 encoding gene (*RIOK3*) to show effect of spliceosome modulator on intron retention between different treatments. **Panel C:** The RNA-seq reads mapping of anti-apoptotic gene *MCL1* to show effect of spliceosome modulator on intron retention between different treatments.
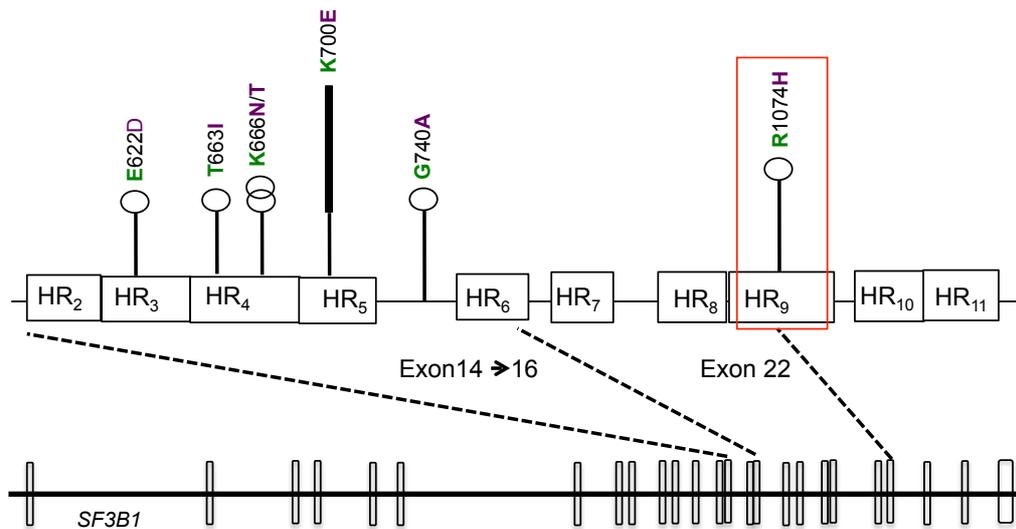
Green: Wild
Purple: Mutated
Red Box: Resistance to Pladienolide B but there is no report of presence of this mutation in CLL-B.

References:
1. Yokoi A et al, *FEBS J.* 2011;**278** (24):4870-80
2. Webb RT et al. *Drug Discovery Today.* 2013;**18** (1-2):43-9
3. Wan Y and Wu CJ, *Blood.* 2013;**121** (23):4627-34
4. Schwaederle M et al. *Leukemia.* 2013;**27**:1214-1217
5. Oscier DG et al. *Blood.* 2012; **121** (3): 468-75
6. Landau DA et al. *Cell.* 2013; **152** (4):714-726

Supplementary Figure S7. Schema showing distribution of mutations in different Heat repeats regions (HR) of *SF3B1* gene.

The figure showing distribution of mutations identified in our and studies by other groups in *SF3B1* gene. For the position of the mutations, the purple letter indicates the original amino acid followed by position of amino acid and the purple color indicates the amino acid after the mutation. The R1074H mutations shown in red box in HR9 has never been reported in any of the study on CLL but in colorectal cancer cell lines after maintaining those in media containing PLAD-B for prolonged period.

Supplementary Tables

Supplementary Table 1: Sequences of primers used in the RT-PCR and qRT-PCR experiments

RT PCR primers		
Primer	**Location**	**Sequence**
DNAJB1- FW	Exon 2	5' GAACCAAAATCACTTTCCCCAAGGAAGG 3'
DNAJB1 - RV	Exon 3	5' AATGAGGTCCCCACGTTTCTCGGGTGT 3'
RIOK3 - FW	Exon 3	5' CCAGTGACCTTATGCTGGCTCAGAT 3'
RIOK3 - RV	Exon 4	5' GGTCTGTAGGGATCATCACGAGTA 3
BCL-X - FW	Exon 2	5' GAGGCAGGCGACGAGTTTGAA 3'
BCL-X - RV	Exon 3	5' TGGGAGGGTAGAGTGGATGGT 3'
MCL-1 - FW	Exon 1	5' CTCGGTACCTTCGGGAGCAGGC 3'
MCL-1 - RV	Exon 3	5' CCAGCAGCACATTCCTGATGCC 3'
GAPDH - FW	Exon 3	5' TGGTCACCAGGGCTGCTT 3'
GAPDH - RV	Exon 4	5' AGCTTCCCGTTCTCAGCCTT 3'
qRT PCR primers		
DNAJB1- FW	Intron 2	5' GGCCTGATGGGTCTTATCTATGG 3'
DNAJB1- RV	Intron 2	5' TTAGATGGAAGCTGGCTCAAGAG 3'
RIOK3 - FW	Intron 3	5' TCAATGGAGATAGCAAAGGTATTATAAC 3'
RIOK3 - RV	Intron 3	5' AGATTTACTTAGGAGCACATTATGAGTG 3'
GAPDH - FW	Exon 3	5' TGGTCACCAGGGCTGCTT 3'
GAPDH - RV	Exon 4	5' AGCTTCCCGTTCTCAGCCTT 3'
FW and RV denote forward and reverse primers		

Supplementary Table 2: Details of the parameters from RNA-seq data

Directory/Experiment File	Genome	Total reads	% Aligned	% Unique Match	% Multimappers	% unmapped	Aligner	Genome	Total reads in analysis	Total positions in analysis	Est. Genome Size	Reads per bp	Avg. Reads per position	GC-content
NBC-PLAD-B	hg19	32437585	51.10%	45.30%	5.90%	48.90%	STAR	hg19	14693711	7497476	3092662209	0.004751	1.413	0.483
NBC-F-ara-A	hg19	33181747	33.80%	30.80%	3.00%	66.20%	STAR	hg19	10234242	7731582	3093307713	0.003309	0.882	0.469
NBC-FD895	hg19	23057170	90.40%	75.00%	15.40%	9.60%	STAR	hg19	17296312	4784233	3093561832	0.005591	2.772	0.496
NBC-control	hg19	23540499	84.20%	73.70%	10.50%	15.80%	STAR	hg19	17347130	7583077	3061420103	0.005666	1.748	0.478
CLL2-PLAD-B	hg19	43358539	84.30%	76.60%	7.70%	15.70%	STAR	hg19	33206561	10793915	3093242291	0.010735	2.105	0.482
CLL2-F-ara-A	hg19	49132267	18.70%	17.50%	1.10%	81.30%	STAR	hg19	8620084	6677622	3092153001	0.002788	0.823	0.472
CLL2-FD895	hg19	27607036	23.30%	21.60%	1.70%	76.70%	STAR	hg19	5965561	3940652	3092691589	0.001929	1.026	0.477
CLL2-control	hg19	43670385	23.80%	22.30%	1.40%	76.20%	STAR	hg19	9747419	7348961	3092779879	0.003152	0.856	0.474
CLL1-PLAD-B	hg19	25885854	88.60%	74.90%	13.70%	11.40%	STAR	hg19	19381741	6101538	3093181770	0.006266	2.311	0.479
CLL1-F-ara-A	hg19	27481703	75.30%	65.80%	9.60%	24.60%	STAR	hg19	18068764	7152187	3093136971	0.005842	1.948	0.474
CLL1-FD895	hg19	48559602	25.20%	22.10%	3.10%	74.80%	STAR	hg19	10739956	6610220	3093379330	0.003472	1.153	0.479
CLL1-control	hg19	36518872	34.60%	31.50%	3.10%	65.40%	STAR	hg19	11501147	8003113	3093323972	0.003718	0.977	0.468
Pladienolide-B	PLAD-B													
Fludarabine	F-ara-A													

Supplementary Table 3: A partial list showing top 50 genes with intron retention in untreated CLL as compared to control normal B cells

Rank Order Based on ratios of IR ratios between CLL/NBC	Refseq ID	Chromosme Location	Gene Description	% CLL/NBC
1	NM_019037	chr8	EXOSC4|RRP41|RRP41A|Rrp41p|SKI6|Ski6p|hRrp41p|p12A|-|8q24.3|protein-coding	454
2	NM_024081	chr11	PRRG4|PRGP4|TMG4|-|11p13|protein-coding	308
3	NM_001004304	chr12	ZNF740|Zfp740|-|12q13.13|protein-coding	153
4	NR_040662	chr6	HCP5|6S2650E|D6S2650E|P5-1|-|6p21.3|ncRNA	125
5	NM_001099694	chr19	ZNF578|-|-|19q13.41|protein-coding	112
6	NM_000137	chr15	FAH|-|-|15q25.1|protein-coding	107
7	NM_001145347	chr19	ZNF576|-|-|19q13.31|protein-coding	100
8	NM_016042	chr9	EXOSC3|PCH1B|RP11-3J10.8|RRP40|Rrp40p|bA3J10.7|hRrp-40|p10|CGI-102|9p11|protein-coding	93
9	NM_018019	chr17	MED9|MED25|-|17p11.2|protein-coding	79
10	NM_005101	chr1	ISG15|G1P2|IFI15|IP17|UCRP|hUCRP|-|1p36.33|protein-coding	70
11	NM_013375	chr6	ABT1|hABT1|-|6p22.2|protein-coding	55
12	NM_002575	chr18	SERPINB2|HsT1201|PAI|PAI-2|PAI2|PLANH2|-|18q21.3|protein-coding	49
13	NM_001160154	chr2	MGAT4A|GNT-IV|GNT-IVA|GnT-4a|-|2q12|protein-coding	44
14	NM_172211	chr1	CSF1|CSF-1|MCSF|RP11-195M16.2|1p13.3|protein-coding	43
15	NM_001135191	chr2	ASAP2|AMAP2|CENTB3|DDEF2|PAG3|PAP|Pap-alpha|SHAG1|-|2p25|2p24|protein-coding	32
16	NM_000882	chr3	IL12A|CLMF|IL-12A|NFSK|NKSF1|P35|-|3q25.33|protein-coding	30
17	NM_181708	chr12	BCDIN3D|-|-|12q13.12|protein-coding	30
18	NM_001285549	chr2	ZDBF2|-|-|2q33.3|protein-coding	26
19	NM_001093771	chr12	TXNRD1|GRIM-12|TR|TR1|TRXR1|TXNR|-|12q23-q24.1|protein-coding	26
20	NM_002657	chr20	PLAGL2|ZNF900|-|20q11.21|protein-coding	25
21	NM_001134655	chr16	ZNF213|CR53|ZKSCAN21|ZSCAN53|-|16p13.3|protein-coding	25
22	NM_018132	chr6	CENPQ|C6orf139|CENP-Q|-|6p12.3|protein-coding	22
23	NM_004209	chr16	SYNGR3|-|-|16p13|protein-coding	22
24	NM_152266	chr19	C19orf40|FAAP24|-|19q13.11|protein-coding	22
25	NM_003131	chr6	SRF|MCM1|-|6p21.1|protein-coding	16
26	NM_020745	chr6	AARS2|AARSL|COXPD8|MT-ALARS|MTALARS|RP11-444E17.1|6p21.1|protein-coding	16
27	NM_001164721	chr1	PTAFR|PAFR|-|1p35-p34.3|protein-coding	14
28	NR_038943	chr10	ADD3-AS1|-|-|-|ncRNA	14
29	NM_001370	chr2	DNAH6|DNHL1|Dnahc6|HL-2|HL2|hCG_1789665|2p11.2|protein-coding	14
30	NM_145294	chr16	WDR90|C16orf15|C16orf16|C16orf17|C16orf18|C16orf19|-|16p13.3|protein-coding	13
31	NM_032043	chr17	BRIP1|BACH1|FANCJ|OF|-|17q22.2|protein-coding	12
32	NM_023007	chr1	JMJD4|-|-|1q42.13|protein-coding	11
33	NR_027131	chrX	NKAPP1|CXorf42|-|Xq24|pseudo	11
34	NM_181553	chr16	CMTM3|BNAS2|CKLFSF3|-|16q21|protein-coding	11
35	NM_003830	chr19	SIGLEC5|CD170|CD33L2|OB-BP2|OBBP2|SIGLEC-5|-|19q13.3|protein-coding	11
36	NM_005456	chr11	MAPK8IP1|IB1|JIP-1|JIP1|PRKM8IP|-|11p11.2|protein-coding	10
37	NM_006963	chr10	ZNF22|HKR-T1|KOX15|ZNF422|Zfp422|RP11-285G1.13-001|10q11|protein-coding	9
38	NM_001256508	chr11	TBC1D10C|CARABIN|EPI64C|-|11q13.2|protein-coding	9
39	NM_001127596	chr1	FCGR3A|CD16|CD16A|FCG3|FCGR3|FCGRIII|FCR-10|FCRIII|FCRIIIA|IGFR3|RP11-5K23.1|1q23|protein-coding	9
40	NM_006779	chr11	CDC42EP2|BORG1|CEP2|-|11q13|protein-coding	8
41	NM_014818	chr11	TRIM66|C11orf29|TIF1D|TIF1DELTA|-|11p15.4|protein-coding	8
42	NM_032429	chr10	LZTS2|LAPSER1|RP11-108L7.8|10q24|protein-coding	8
43	NM_144706	chr2	C2orf15|-|-|2q11.2|protein-coding	8
44	NM_004673	chr1	ANGPTL1|ANG3|ANGPT3|ARP1|AngY|UNQ162|dJ595C2.2|PSEC0154|1q25.2|protein-coding	8
45	NR_037861	chr6	PPT2-EGFL8|-|-|6p|ncRNA	7
46	NM_032387	chr17	WNK4|PHA2B|PRKWNK4|-|17q21-q22|protein-coding	7
47	NM_001097579	chrX	GPR34|-|RP11-204C16.6|Xp11.4|protein-coding	6
48	NM_001136482	chr19	C19orf38|HIDE1|-|19p13.2|protein-coding	6
49	NR_109990	chr20	RP5-1103G7.4|-|-|-|ncRNA	4
50	NM_001135768	chr19	PVR|CD155|HVED|NECL5|Necl-5|PVS|TAGE4|-|19q13.2|protein-coding	3

Supplementary Table 4: SF3B1 mutations observed within the examined patient-derived CLL cells

Patient	Mutation	Exon	Nucleotide Change	COSMIC ID
CLL001	ND	N/A	N/A	N/A
CLL002	ND	N/A	N/A	N/A
CLL003	E622D	14	c.1866G>C	132938
CLL004	K700E	15	c.2098A>G	84677
CLL005	K700E	15	c.2098A>G	84677
CLL006	T663I	14	c.1988C>T	145921
CLL007	E622D	14	c.1866G>C	132938
CLL008	G740E	15	c.2219G>A	133120
CLL009	K700E	15	c.2098A>G	84677
CLL010	K700E	15	c.2098A>G	84677
CLL011	K666T	14	c.1997A>C	131556
CLL012	K666N	14	c.1996G>C	132937
CLL001 and CLL002 were used for RNA-seq analysis.				
ND: No mutation detected; N/A: Not applicable				