Novel *SF3B1* in-frame deletions result in aberrant RNA splicing in CLL patients

Anant A. Agrawal[1*], Michael Seiler[1*], Lindsey T. Brinton[2], Rose Mantel[2], Rosa Lapalombella[2], Jennifer A. Woyach[2], Amy J. Johnson[2], Ping Zhu[1], Markus Warmuth[1], Lihua Yu[1], John C. Byrd[2], Peter G. Smith[1], James S. Blachly[2#], Silvia Buonamici[1#]

[1]H3 Biomedicine, Cambridge, MA, USA

[2]Ohio State University Comprehensive Cancer Center, Columbus, OH, USA

[*]Co-first author, [#]Co-senior author

Supplemental Material

Table S1. CLL patient information and cytogenetics

Sample name	SF3B1 K700del	SF3B1 Q698del 1	SF3B1 Q698del 2
Patient ID	CLL1	CLL2	CLL3
SF3B1 mutation	p.K700_V701delinsN	p.Q698_K700delinsQ	p.Q698_K700delinsQ
Age (yrs)	55	54	42
Sex	Male	Female	Male
Rai stage	IV	IV	II
IGHV mutation	Unmutated	Unmutated	Unmutated
Cytogenetics	t(8;17) resulting in gain of 8q and loss of 17p. Deletion 17p and 3 copies of 8q24 by FISH	Complex karyotype with t(13;17). Deletion of 17p and 1 copy of 13q14.3 by FISH	Complex karyotype with isochromosome 17; deletion 17p by FISH
SF3B1 variant allele fraction (%)	27.3	42.3	8.8

Material and Methods

Patient samples

Patients with unknown *SF3B1* mutation status were selected from a cohort of previously-treated CLL patients enrolled in IRB approved trials at The Ohio State University in accordance with the Declaration of Helsinki. Interphase fluorescent *in-situ* hybridization (FISH) and metaphase cytogenetics were studied on fresh samples, while remaining cells were enriched with a negative selection cocktail and separated with density gradient centrifugation and then viably cryopreserved.

Cytogenetics

CLL patient cells were cultured and stimulated with pokeweed mitogen, PMA and CpG oligonucleotide prior to microscopy of metaphases for determination of cytogenetics. Interphase FISH for 8q24, 11q22.3, 12cen, 13q14.3, and 17p13.1 was performed with probes from Abbot Molecular.

RNAseq sample preparation, quantification and analysis

Viably cryopreserved CLL primary tumors were thawed and total RNA was extracted with the Qiagen miRNeasy kit. RNA samples were treated with ambion turbo DNAse. Sequencing libraries were prepared using the Illumina TruSeq stranded mRNA library preparation kit according to the manufacturer's directions, and then sequenced on a HiSeq 2500 instrument targeting approximately 35-40x10^6 read pairs per sample. All raw RNA sequencing reads are available in GEO under accession GSE95352. Raw RNA sequence reads were aligned to hg19 using STAR.[1] Splice junctions sharing a common splice site were identified and measured as percent spliced in (PSI) as done previously,[2] and RNASeq mutation frequency was determined using samtools.[3] *SF3B1*-mutant specific alternative 3'ss (n = 194) were obtained and hierarchical clustering of z-normalized junction PSI data was used to identify samples with *SF3B1*-mutant associated aberrant splicing.[2]

DNAseq sample preparation and analysis

DNA was prepared with the Qiagen QIAamp DNA mini kit according to manufacturer's instructions. Sequencing libraries were prepared using custom adapters and enriched for *SF3B1* with hybrid capture probes from Integrated DNA Technologies. Libraries were sequenced on a Miseq instrument (Illumina).

Raw DNA reads from samples with in-frame deletions were aligned using BWA[4] to genome indices based on the sequence of *SF3B1*, *SF3B1* K700del, and *SF3B1* Q698del. The coverage ratio in the deletion region between the deletion genome index and the intact *SF3B1* index were used to determine allele frequency. DNA sequencing reads are available in the sequence read archive (SRA) under accession SRP100874.

Structural structure modeling

Homology model of SF3B1 in-frame deletions were generated by the SWISS-MODEL web server based on the cryo-EM and crystal structures of Hsh155 and SF3B1 (PDB ID: 5GM6 and 5IFE), respectively. Figures were generated in PyMOL.

Cell culture and expression plasmids

HEK293FT (Life technologies) cells were used for overexpressing mxSF3B1 and ZDHHC16 minigene plasmids. These cells were cultured in DMEM+10%FBS according to the manufacturer's instruction. Both mxSF3B1 and ZDHHC16 minigene expression plasmids have been described in detail previously. In-frame deletions p.K700del and p.Q698del were introduced using site-directed mutagenesis kit (QuikChange II XL, Agilent).

Real time quantitative PCR (qPCR)

RNA was extracted from cultured HEK293FT cells using RNeasy spin columns (Qiagen). cDNA was prepared using SuperScript Vilo, and diluted to 20ng input for qPCR. mxSF3B1 transcript levels were detected by Viia7 using the following TaqMan primers-probe sets: FWD 5'-TTG ATG TCG CCA ACA CTT GAA-3', REV 5'-GCT AAA CCA GCA GCC TTG G-3' and Probe 5'-TTG GAT GAC TTG GTC CGC CCA T-3'.

Co-immunoprecipitation and western blotting

Transfected HEK293FT cell lines were lysed in RIPA buffer plus protease-inhibitor cocktail (Mini-complete, EDTA-free, Roche). α-HA magnetic beads (Thermo Scientific, 88838) were used for co-immunoprecipitation of HA-SF3B1 according to manufacturer's instructions. Input lysates and elution samples were diluted in RIPA buffer with 4X LDS Sample Buffer (Nupage, Life Technologies) and 10X Reducing Reagent (Nupage, Life Technologies). 30ug of total protein were loaded per well in 4-12% Bis-Tris SDS Page gels (Novex, Life Technologies). Gels were transferred to nitrocellulose membranes using iBlot system (Life Technologies). Membranes were blocked in pre-filtered LI-COR Blocking Buffer (Odyssey) for 1 hour and then cut into sections. Each section was probed separately with antibodies to each of the following proteins: HA-tag, SF3B1, SF3B2, SF3B3, SF3B4, SF3B14a/p14, SF3B14b/PHF5A. The following primary antibodies were used at 1:1000 dilution for western blot analysis in LI-COR buffer: α-SF3B1 mouse monoclonal antibody (MBL, D221-3), α-SF3B2 rabbit polyclonal antibody (Protein Tech, 10919-1-AP), α-SF3B3 rabbit polyclonal antibody (Protein Tech, 14577-1-AP), α-SF3B4 goat polyclonal antibody (Santa Cruz, 14276), α-p14 rabbit polyclonal antibody (Protein Tech, 12379-1-AP), α-PHF5A rabbit polyclonal antibody (Protein Tech, 15554-1-AP), α-HA mouse monoclonal antibody (Abcam, ab18181). α-rabbit and α-goat IRDye-800CW secondary antibody (LI-COR) was used at 1:5000 dilution and α-mouse IRDye-680LT secondary antibody (LI-COR) was used at 1:20,000 dilution. Western blot was imaged using Odyssey V3.0 imager.

Minigene assay

HEK293FT cells were seeded on 6 wells/plate and transfected using Fugene (Roche). 1 μg of DNA per pcDNA-DEST40 HA-mxSF3B1 and *ZDHHC16* minigene constructs were used for each transient transfection, generated in triplicates. 48 hours later, cells were collected to isolate RNA for PCR analysis. RNA was isolated as described above and the cDNA was used for PCR. For all minigene transcripts, the following forward and reverse primers were used: FWD 5'-TCT TGT CTA CCT CTG GTT CCT-3'; REV 5'-TAG AAG GCA CAG TCG AGG-3'. Products were

amplified for 29 cycles using OneTaq 2XMasterMix (NEB), loaded onto 2.5% agarose gels, and run at 80V for 4 hours.

Cell viability assay

Primary CLL patient cells were treated in triplicate with the spliceosome inhibitor E7107 at concentrations ranging from 1.6 to 100 nM, and cell viability was assessed by MTS colorimetric assay.

References

1. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* Jan 01 2013;29(1):15-21.

2. Darman RB, Seiler M, Agrawal AA, et al. Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3' Splice Site Selection through Use of a Different Branch Point. *Cell reports.* Nov 03 2015;13(5):1033-1045.

3. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* Aug 15 2009;25(16):2078-2079.

4. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* Jul 15 2009;25(14):1754-1760.