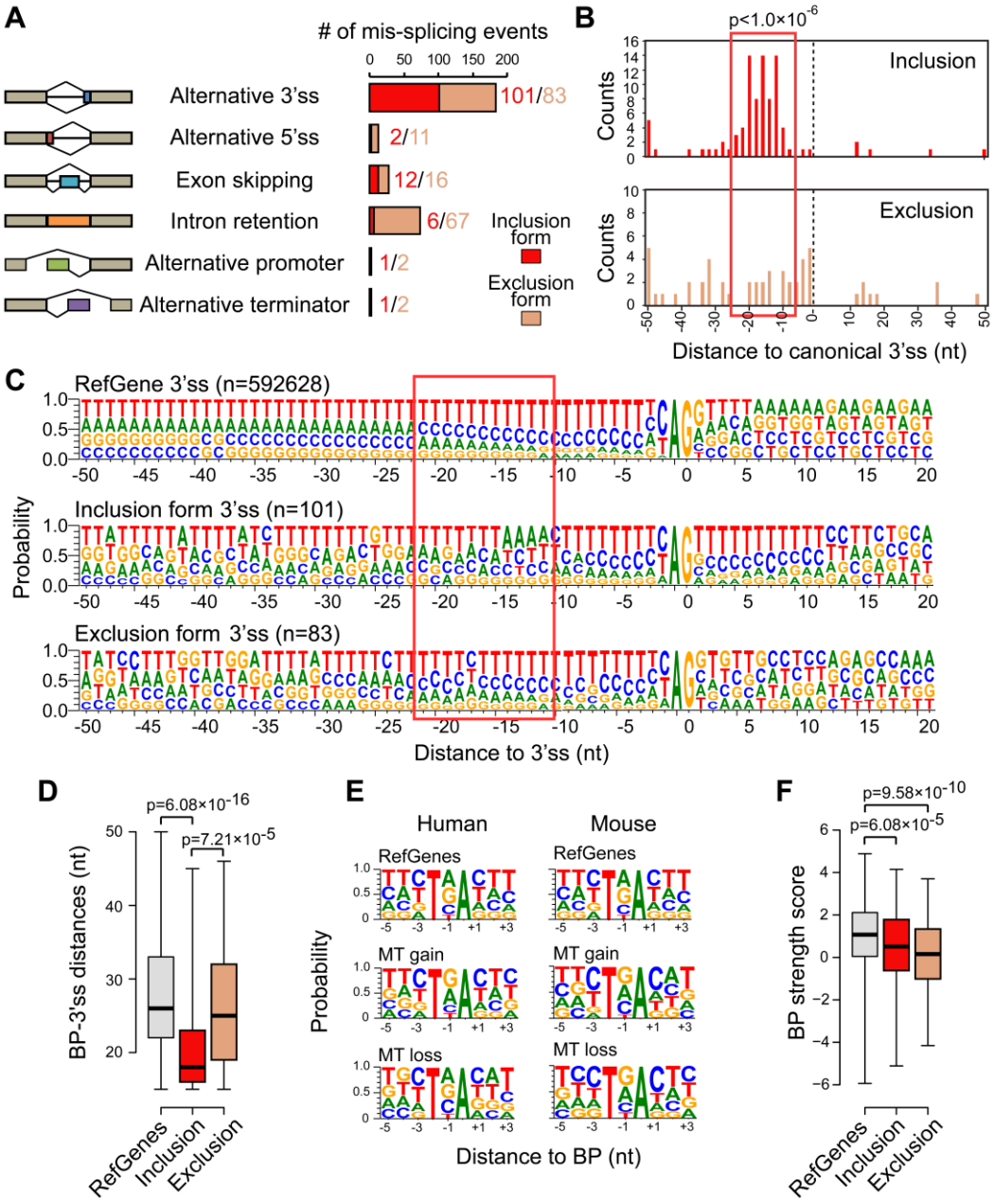

Figure S1, related to Figure 1. *SF3B1* mutation induces cryptic 3'ss selection in human CLLs. (**A**) Splice variants associated with *SF3B1* mutation in human CLLs are shown. (**B**) Histogram of distances between the cryptic and the canonical 3'ss are shown. 0 point defines the position of the canonical 3'ss. (**C**) Sequence motifs around all RefGene 3'ss, MT inclusion 3'ss and MT exclusion 3'ss are shown. The height of each letter indicates the probability that nucleotide is used at that position. (**D**) The distances between the branch point and the corresponding 3'ss are shown. 0 point defines the position of the 3'ss. (**E**) Sequence motifs associated with branch points in different groups are shown. (**F**) The BP strength scores of motifs from panel E are shown. For (D) and (F), Center lines show the means, box limits indicate the 25th and 75th percentiles and whiskers extend to minimum and maximum values.

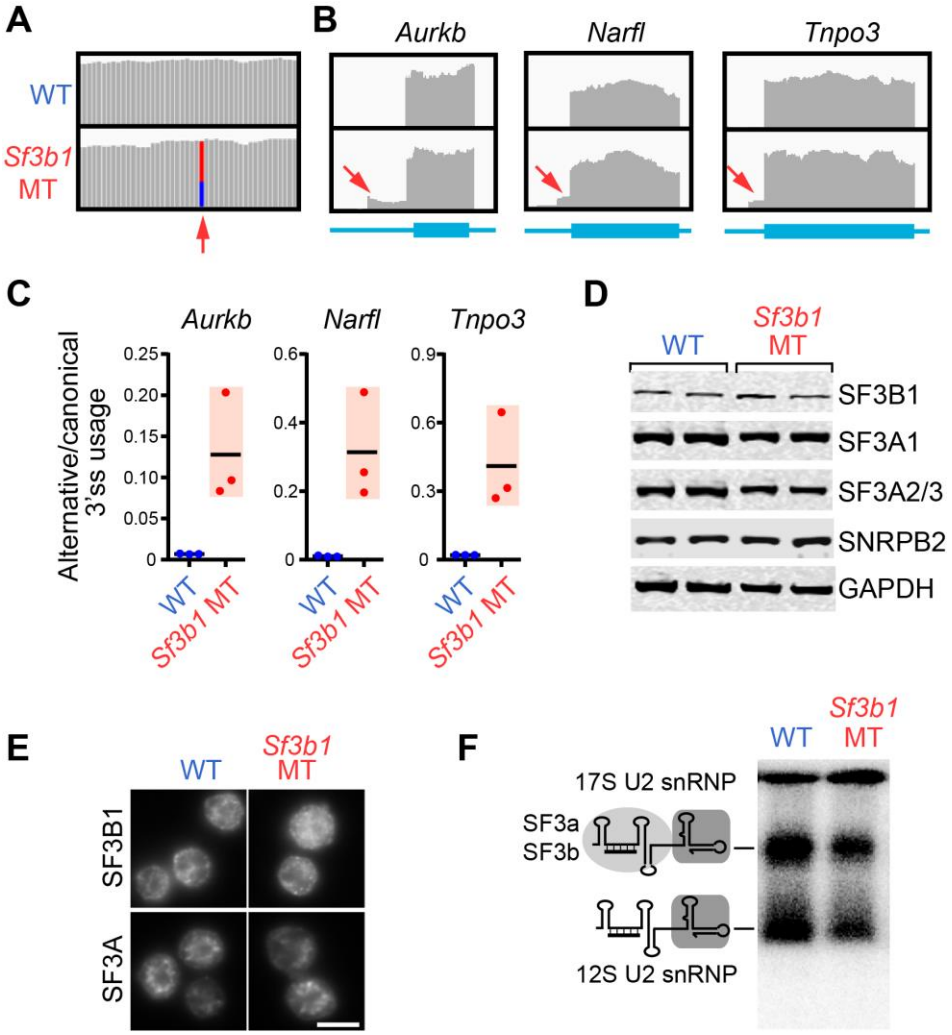

Figure S2, related to Figure 1. *Sf3b1* mutation induces cryptic 3'ss selection in B cells. (**A-B**) IGV of RNA-seq reads covering *Sf3b1-K700E* mutation (**A**) and the cryptic 3'ss of *Aurkb*, *Narfl* and *Tnpo3* genes (**B**) in the *Sf3b1* MT cells are shown. (**C**) Validation of the splice variants identified from RNA-sequencing in three replicates by quantitative RT-PCR. Center lines indicate means. (**D**) Western blot of *Sf3b1* and other components of U2 snRNP in B cells with WT or MT *Sf3b1* are shown. Two biological replicates are shown for each group. (**E**) Immunofluorescent staining of SF3B1 and SF3A proteins in the WT and *Sf3b1* MT B cells. Scale bars, 5 μm. (**F**) Functional U2 snRNP assembled in nuclear extract generated from WT or *Sf3b1* MT B cells are shown. Schematic shows the 12S U2 snRNP core and 17S U2 snRNP containing the SF3a/SF3b complex. U2 snRNP complexes were separated on a native gel and detected by a ^{32}P-labeled 2' O-methyl oligo base pairing to U2 snRNA.

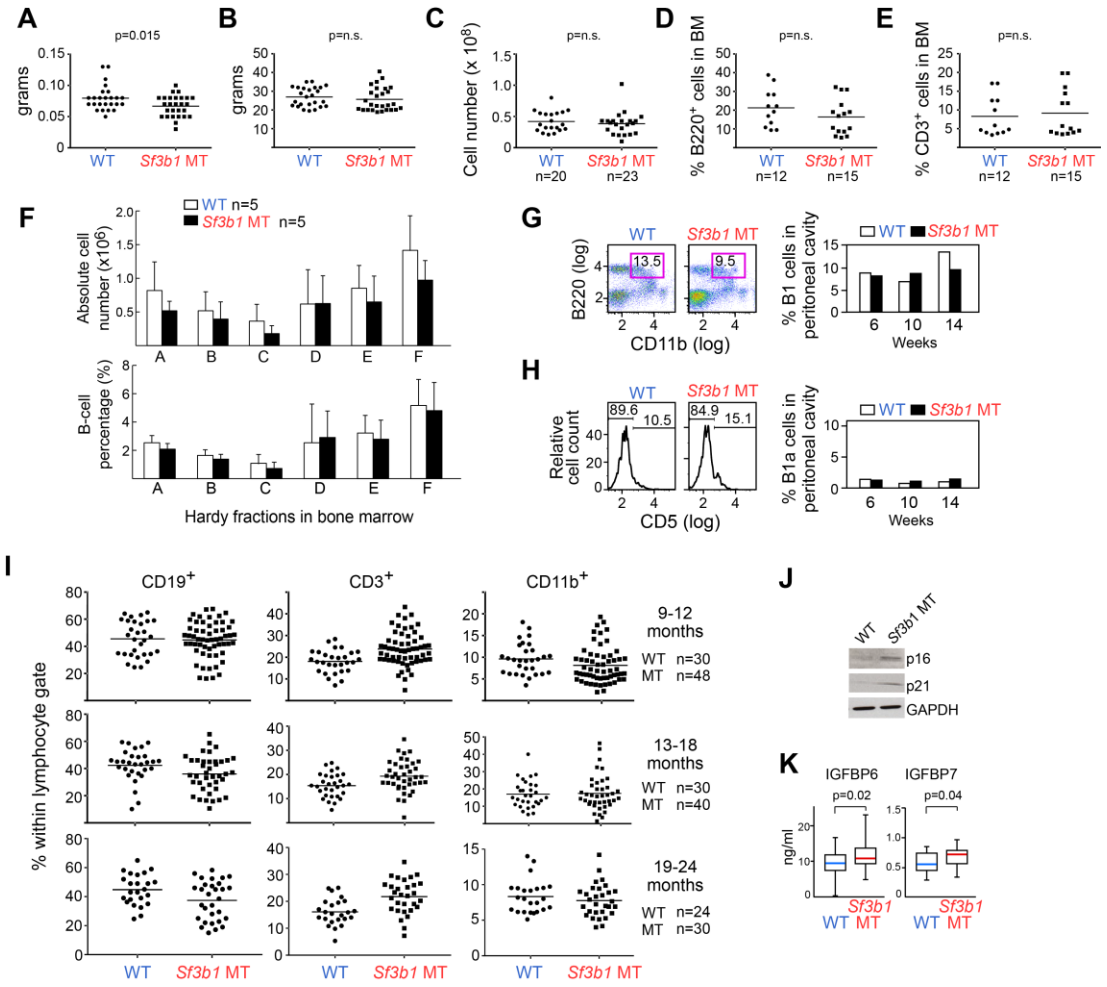

Figure S3, related to Figures 2. Expression of *Sf3b1*-K700E affects B cell development. (**A-E**) Spleen weight (**A**), body weight (**B**), total bone marrow mononuclear cells (BMC) (**C**), percentages of B220$^+$ B cells (**D**) and CD3$^+$ T cells (**E**) in bone marrows of WT and *Sf3b1* MT mice are shown. Center lines indicate the means; n.s., nonsignificant. (**F**) Absolute cell number and percentage of B cells in different hardy fractions of bone marrow mononuclear cell (BMC) are shown. Data represent Mean±SD of results derived from 5 WT and 5 MT mice. (**G** and **H**) The percentage of B220$^+$CD11b$^+$ B1 cells (**G**) and CD5$^+$ B1a cells (**H**) in the peritoneal cavity from 3 WT and 3 *Sf3b1* MT mice. (**I**) Percentages of CD19$^+$ B cells, CD3$^+$ T cells or CD11b$^+$ macrophages in WT or *Sf3b1* MT mice aged between 9-12, 13-18 and 19-24 months are shown. Center lines indicate the means. (**J**) Western blot of senescence markers p16 and p21 in B cells from WT and *Sf3b1* MT mice are shown. (**K**) The levels of the secreted proteins IGFBP6 and IGFBP7 in the serum of the WT and *Sf3b1* MT mice (n=30 per group) are shown. Center lines indicate the means; box limits indicate the 25th and 75th percentiles; whiskers extend to minimum and maximum values.

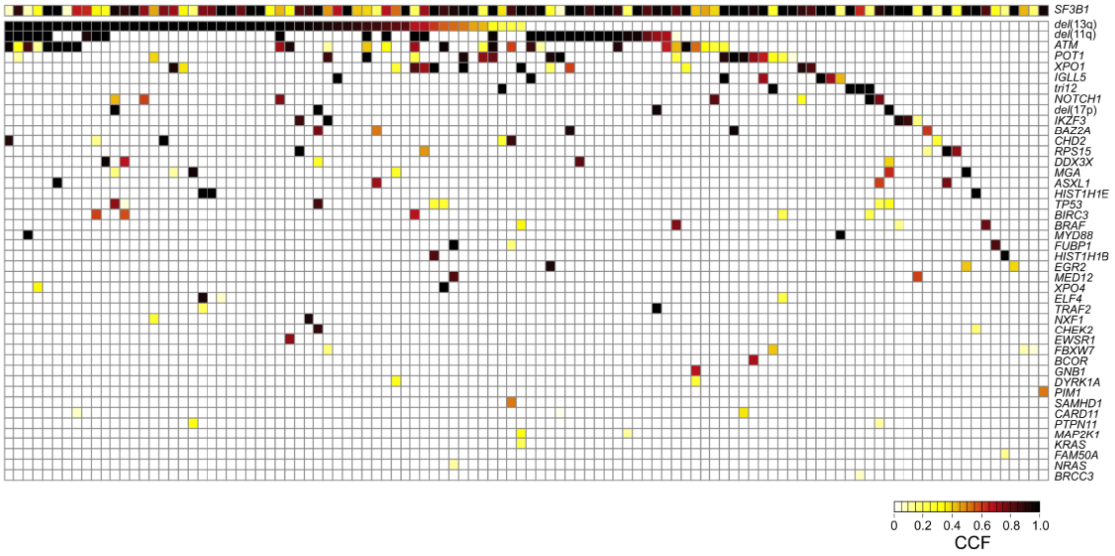

Figure S4, related to Figure 3. Co-occurrence of somatic mutations in 116 CLL samples with mutant *SF3B1*. The map summarizes results of whole exome sequencing analysis of 538 CLL samples in *Landau et al. Nature 2015*. CCF, cancer cell fraction.

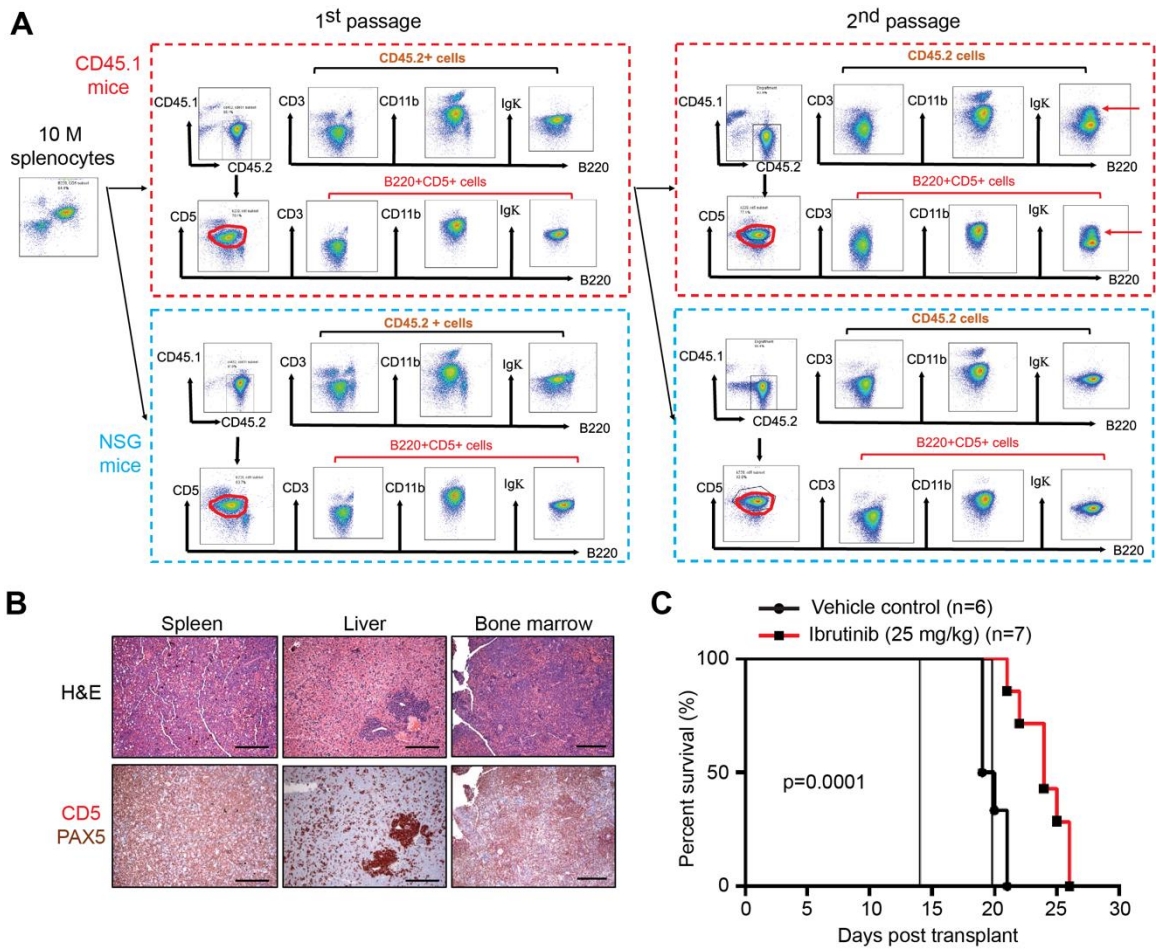

Figure S5, related to Figures 4. Transplantation of DM-CLL cells. (**A**) Serial transplantation of DM-CLL cells in immunocompetent recipient mice (CD45.1) and immunodeficient recipient mice (NSG) are shown. 1st and 2nd engraftments were evaluated 6-weeks and 4-weeks post tail-vein injection into CD45.1 and NSG mice, respectively. In the first round of engraftment, DM-CLL cells were engrafted into CD45.1 and NSG mice and multiple lineage-specific markers from flow cytometry analysis are shown. In the second round of engraftment, we further engrafted spleen CLL cells from CD45.1 recipient mice into CD45.1 and NSG recipient mice. Red arrows indicate a subpopulation of CLL cells with higher IgK expression arise in CD45.1 mice. (**B**) H&E staining and IHC staining of CD5 and PAX5 proteins show infiltration of CLL-like cells in multiple tissues in NSG mice. Scale bar: 500 μm. (**C**) Kaplan–Meier survival curve of DM-CLL transplanted NSG mice treated with 25 mg/kg Ibrutinib (n=7) or vehicle control (n=6). Drugs were administered daily between 14-20 days post-transplantation.

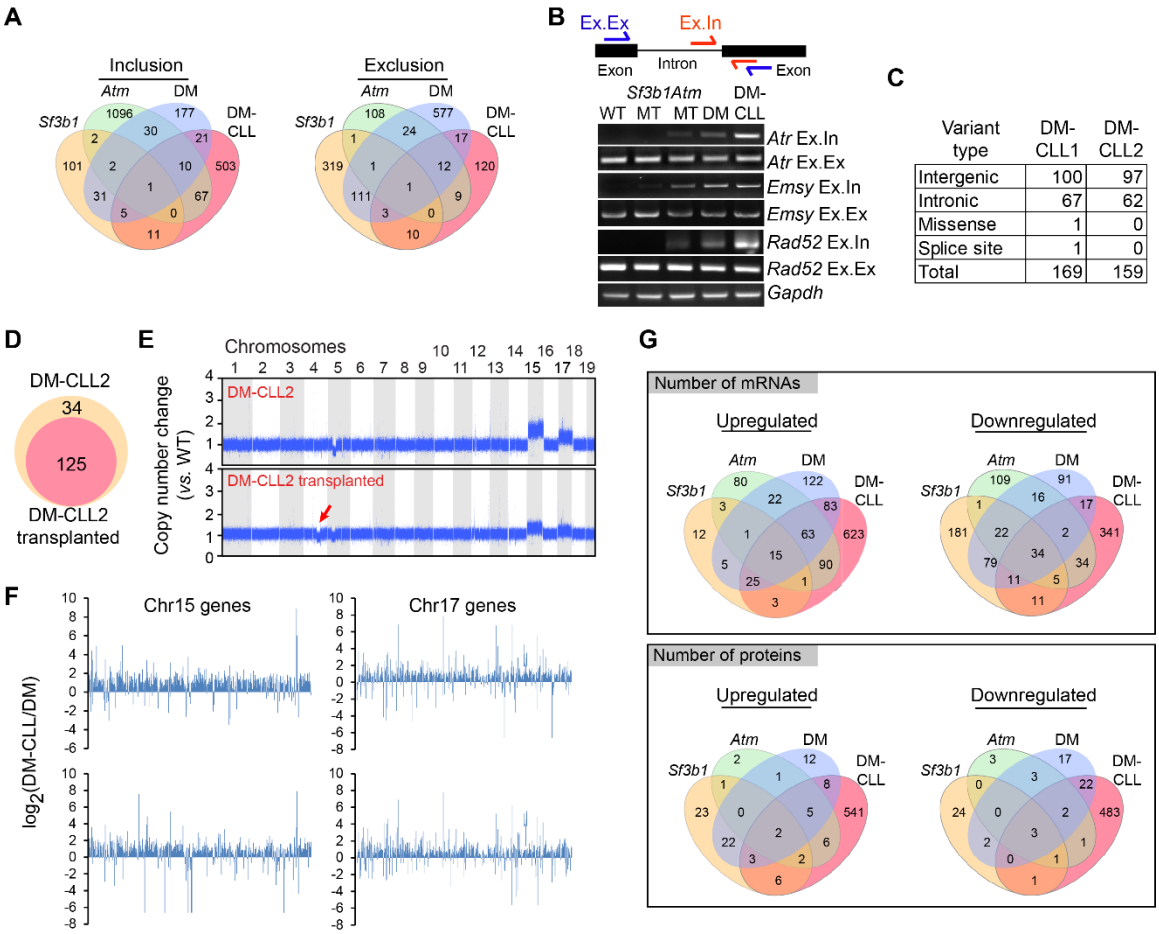

Figure S6, related to Figures 5 and 6. Comprehensive characterization of the mouse models. (**A**) Venn diagrams show the overlaps of mis-splicing events amongst different groups. (**B**) RT-PCR validation detects intron retention events in multiple DNA damage response genes. (**C**) The number of clonal somatic mutations identified in two different DM-CLLs. (**D**) The number of overlapping SNVs between the original DM-CLL2 cells and transplanted CLL cells are shown. (**E**) IGV of copy number variants in the original DM-CLL2 cells and transplanted CLL cells are shown. (**F**) DM-CLL/DM fold changes of genes localized in Chromosomes 15 and 17 are shown. Genes were arranged based on their position along the chromosomes. (**G**) Venn diagrams show the overlaps between differentially expressed genes/proteins in each MT group versus WT. See Table S6 for details.

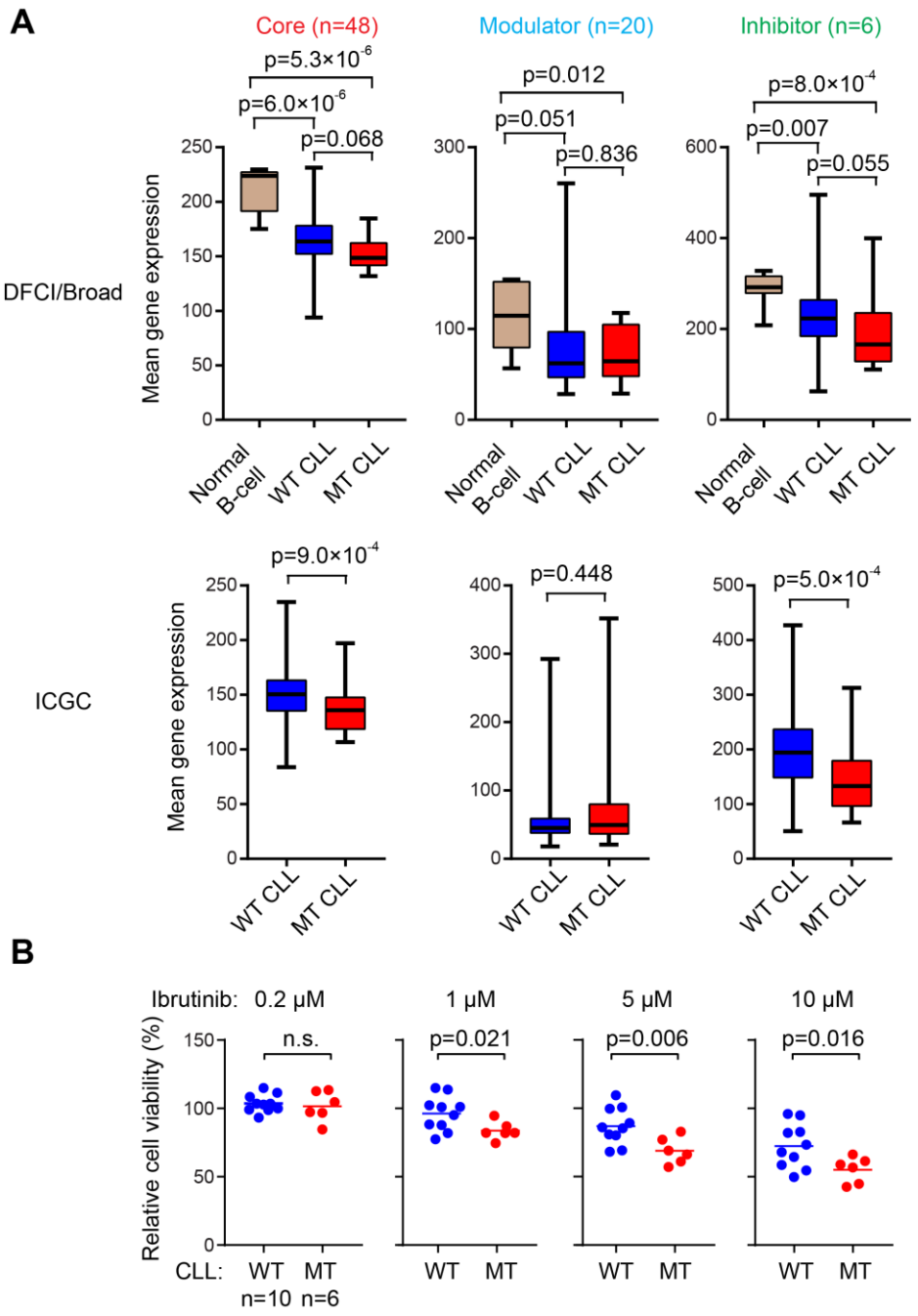

Figure S7, related to Figure 7. Altered BCR signaling genes in human CLLs with *SF3B1* mutations. (**A**) Mean expression of BCR signaling core components (n=48), modulators (n=20) and inhibitors (n=6) in different groups are shown. See Table S7 for BCR signaling genes. Center lines show the means, box limits indicate the 25th and 75th percentiles and whiskers extend to minimum and maximum values. (**B**) 10 *SF3B1* WT and 6 *SF3B1* MT CLLs were treated with different doses of ibrutinib for 48 hours and the relative cell viability relative to DMSO control group were calculated. Each dot represents mean of three replicates of each individual CLL case. Center lines show the means of the group.