An Alternatively Spliced Gain-of-Function NT5C2 Isoform Contributes to Thiopurine Resistance in Acute Lymphoblastic Leukemia

Manuel Torres-Diz[1], Clara Reglero[2], Catherine D Falkenstein[3], Annette Castro[1], Katharina E. Hayer[1,4], Caleb M. Radens[5,6], Mathieu Quesnel-Vallières[6], Zhiwei Ang[1], Priyanka Sehgal[1], Marilyn M. Li[7,8], Yoseph Barash[5,6], Sarah K. Tasian[3,9], Adolfo Ferrando[2], and Andrei Thomas-Tikhonenko[1,3,8,9 #]

[1] Division of Cancer Pathobiology, Children's Hospital of Philadelphia, Philadelphia, PA
[2] Institute for Cancer Genetics, Columbia University, New York, New York
[3] Division of Oncology, Children's Hospital of Philadelphia, Philadelphia, PA
[4] Department of Biomedical and Health Informatics, Children's Hospital of Philadelphia, Philadelphia, PA;
[5] Cell and Molecular Biology Graduate Group, University of Pennsylvania, Philadelphia, PA
[6] Department of Genetics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA
[7] Division of Genomic Diagnostic, Children's Hospital of Philadelphia, Philadelphia, PA
[8] Department of Pathology & Laboratory Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA
[9] Department of Pediatrics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA

Running title: Aberrant splicing in chemoresistant leukemia

Corresponding author: Andrei Thomas-Tikhonenko, 3056 Colket Transl Res Bldg, 3501 Civic Center Blvd, Philadelphia, PA 19104-4399 | andreit@pennmedicine.upenn.edu | Phone: 267-426-9699

Abstract

Relapse-specific mutations do not account for all chemotherapy failures in patients with B-cell acute lymphoblastic leukemia (B-ALL). By mining RNA-seq datasets of paired diagnostic/relapse pediatric B-ALL samples, we discovered pervasive alternative splicing (AS) patterns linked to relapse and affecting drivers of resistance to glucocorticoids, anti-folates, and thiopurines. Most splicing variations represent exon skipping, “poison” exon inclusion, and intron retention, phenocopying well-documented loss-of-function mutations. In contrast, relapse-associated AS of *NT5C2* mRNA yields an isoform with a cryptic 24-nt in-frame exon 6a. Inclusion of the extra 8 amino acids into this enzyme results in elevated nucleosidase activity, a known consequence of gain-of-function mutations in NT5C2 and a common determinant of 6-mercaptopurine resistance. Furthermore, in B-ALL cells NT5C2ex6a and the R238W hotspot variant confers comparable levels of resistance to 6-mercaptopurine *in vitro* and *in vivo*. These results support a role for alternative splicing as a prevalent mechanism driving chemotherapy resistance in relapsed B-ALL.

Statement of significance

Mutations in chemoresistance genes are found in relapsed/refractory acute lymphoblastic leukemia. However, in this low-mutational-burden disease, up to 30% of cases have no known relapse-specific genetic alterations. Our identification of aberrant splicing as an alternative mechanism of acquired drug resistance fills this gap and suggests new opportunities for therapeutic interventions.

Conflict of interest

The authors declare no relevant conflicts of interest.

Introduction

Leukemia is the most common cancer in children and adolescents/young adults (AYAs), accounting for almost 1/3 of all pediatric cancers, with the majority of cases classified as B-cell acute lymphoblastic leukemia (B-ALL). Only ~50% of children with first relapse survive long term, and outcomes are appreciably poorer with subsequent relapses. Early bone marrow relapses occurring within three years of diagnosis are particularly difficult to salvage (1). As a result, relapsed or refractory (r/r) high-risk B-lymphoblastic leukemia (B-ALL) cases account for a substantial number of pediatric cancer-related deaths. In addition, while most cases of B-ALL are diagnosed early in life, most deaths from B-ALL (~80%) occur in adults, largely due to r/r disease. Adults with B-ALL experience very high relapse rates and long-term event-free survival of less than 50% (2,3).

Definition of high-risk (HR) ALL is usually based on clinical factors (e.g., age at diagnosis) + cytogenetic biomarkers (e.g., Philadelphia chromosome or *KMT2A* rearrangements). Yet what molecular mechanisms are responsible for treatment failures in r/r B-ALL is only partially understood. Previous whole exome/genome sequencing efforts have identified a dozen of mutated relapse-specific targets (4-6). While some of them (e.g., *TP53*) are known to mediate responses to pan-cancer drugs such as anthracyclines, others are involved in therapeutic responses to specific leukemia therapeutics: glucocorticoids (*NR3C1*, *WHSC1*), anti-folates (*FPGS*), and thiopurines (*PRPS1*, *PRPS2*, *NT5C2*).

The most recent effort to catalog such mutations yielded up to 50 genes significantly enriched at first relapse but often as subclonal events or with low recurrence (7). Even the most recurrently involved drivers are mutated in fewer than 20% of all samples suggesting a major role for an alternative mechanism of gene dysregulation. One such mechanism is alternative spicing (AS); yet its role in acquired chemotherapy resistance remains poorly defined. To bridge this gap, we previously had generated and mined new and existing RNA-seq datasets in search for local splicing variations prevalent in B-ALL samples, but rare in normal bone marrow cell counterparts. Of note, B-ALL-specific AS was found to affect 15 out of the 20 top leukemia driver genes from the COSMIC database with frequencies far exceeding those of somatic mutations (8). This intersection suggests that AS could be an important mechanism of both leukemogenesis and drug resistance. Here we provide evidence that a subset of relapses lacking identifiable acquired mutations has an AS signature simultaneously affecting multiple r/r B-ALL drivers, including a novel NT5C2 exon 6a mRNA variant encoding a gain-of-function nucleotidase isoform with increased enzymatic activity capable of conferring resistance to 6-mercaptopurine (6-MP).

Materials and Methods

Dataset usage. We accessed RNA seq data from the TARGET ALL Phase I and II projects (https://portal.gdc.cancer.gov/projects) and from the BLUEPRINT consortium (http://dcc.blueprint-epigenome.eu). TARGET ALL data used for this analysis were accessed under Project #10088: “Alternative splicing in pediatric cancers” (request 41466-5).

RNA-seq analysis. RNA-seq reads were first trimmed to remove adapters (BBTools version 38.87) and were then aligned using STAR version 2.7.3a to the hg38 reference genome, while providing known gene isoforms through the GENCODE annotation version 32. Additionally, we used STAR flags “--quantMode GeneCounts” and “--alignSJoverhangMin 8” to quantify genes and ensure spliced reads had an overhang of at least 8 bases. Junction spanning reads were obtained from the STAR “*_SJ.out.tab” result files and each entry was normalized by dividing by the total number of junction-spanning reads and multiplying by a factor of one million. Visualization and downstream analyses were conducted in R using the ggplot2 and tidyverse packages. JPM counts for select events can be found in Supplemental Table 1.

Mutation analysis. Relapse-specific variant discovery was done using the Genome Analysis Toolkit (GATK) version 4.4.0.0 on the RNA-seq data. Shortly, aligned reads were marked for duplicated events using Picard Tools version 3.0.0. Reads spanning splicing events were split into independent reads with the SplitNCigarReads tool followed by base recalibration. To specifically call mutations that are

enriched in the relapsed samples, the Mutect2 algorithm was applied using the corresponding diagnostic sample to each relapsed file in the –normal flag. All the variants were annotated using the Funcotator tool with the hg38 somatic annotation from the FuncotatorDataSourceDownloader. All the maf files from the 48 patients were loaded into the R package maftools for variant visualization and downstream analysis.

MAJIQ splicing analysis. The Modeling Alternative Junction Inclusion Quantification (MAJIQ) algorithm version 4.4 (9) was run with each sample being quantified individually on a build comprising 219 TARGET B-ALL samples with *de novo* junction detection, but without intron retention detection. Samples were run against the gencode hg38 GFF3 annotation (v37) and required a minimum of 5 reads per sample per junction for quantification. For PSI quantification it was used the Majiq PSI and Voila tools to extract all the LSVs in all the samples with the –show-all flag. For the 48 diagnostic-relapse paired samples, the differential splicing analysis was performed with the Majiq ∆PSI followed by Voila to extract all the LSVs with the –show-all flag. Visualization and downstream analyses were conducted in R using the ggplot2, Complexheatmap and tidyverse packages. The sashimi plots were generated with the ggsashimi tool.

qRT-PCR. Total RNA was isolated using Maxwell RSC simplyRNA Cells kit with the Maxwell RSC48 Instrument (Promega) and reverse-transcribed using SuperScript IV (Invitrogen; 18090010). Primers used for mRNA isoform quantification are listed in Supplementary Tables. qRT-PCR was performed using PowerSYBR Green PCR Master Mix (Life Technologies). Reactions were performed on an Applied Biosystems Viia7 machine and analyzed with Viia7 RUO software (Life Technologies). When indicated, individual PCR products were gel-purified (QIAquick Gel Extraction Kit; Qiagen) and Nanopore sequenced.

Oxford Nanopore Technologies (ONT) targeted long-read sequencing. For each target gene we designed primers binding to exons present in all the isoforms to ensure full coverage of all the alternative splicing events. 5ng of cDNA were amplified with the long LongAmp Taq 2X Master Mix (New England Biolabs) for 25 cycles. The resulting amplicons were subjected to amplicon-seq (SQK-NBD114.24, ONT) library preparation. Subsequently, each library was loaded into a Spot-ON flow cell R10 Version (FLO-MIN114, ONT) and sequenced in a MinION Mk1C device (ONT) until it had at least 1000 reads per samples. Results were aligned using Minimap2 version 2.24-r1122 and visualized in IGV version 2.12.3.

Cell lines and cell culture. Human REH, Nalm-6, 697, SUP-B15, and MHH-CALL-4 B-ALL cell lines originally obtained from commercial repositories were cultured and maintained in conditions recommended by ATCC or DSMZ. Cell line authentication for REH and MHH-CALL-4 was performed by short tandem repeat profiling. All cell lines were routinely confirmed to be *Mycoplasma*-free via Sartorius EZ-PCR™ Mycoplasma detection kit testing.

Genome editing. gRNAs targeting CREBBP exons 25 and 26 (listed in Supplementary Tables) and CAS9 protein were obtained from IDT and ribonucleoprotein (RNP) complexes were assembled following manufacturing recommendations. RNP were transfected into 697 and REH cells via electroporation using the Neon transfection system (Invitrogen Neon Transfection System Model MPK5000) and the Neon 10uL Transfection kit (MPK1096 Invitrogen) with the following conditions: 1700V, 20ms and 2 pulses. The effects on exons removal were measured by cDNA PCR with primers amplifying between exons 20 and 27 (listed in Supplementary Tables) and Western blotting.

Plasmid constructs and viral infections. For NT5C2 constructs, the coding sequence of the canonical NT5C2 isoform (NM_001351169.2) was amplified using specific primers (listed in Supplementary Tables) and cloned into the pLVX backbone(Clonetech) modified with an EF1a promoter and a Basticidin resistance gene using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs). Specific primers and NEBuilder HiFi DNA Assembly Master Mix were used to introduce the exon E6a or the R238W mutation into the canonical sequence of NT5C2.

For viral particles production, 293T (ATCC) cells at a 90% confluence in a 10cm petri dish were transfected with 10µg of the NT5C2 constructs, 2.5µg of the pMD2.G (#12259 Addgene), and 7.5µg of

the psPAX2 (#12260 Addgene) using 60µL of 25µM PEI (24765 Polysciences). Supernatants were collected 48h after transfection, filtered with a 0.45µm PVDF filter and incubated with $4x10^6$ cells for 48h containing 10µg/mL plolybrene (HY-112735 MedChemExpress). Afterwards, culture media were replaced and cells were incubated with 10ug/mL Blasticidin for 96h.

NT5C2 recombinant protein production and purification. For 5′-nucleotidase assays in the absence and presence of allosteric activators, we cloned, expressed, and purified recombinant wild-type and mutant NT5C2 proteins as previously described (10). Briefly, we cloned full-length NT5C2 cDNA constructs with an N-terminal hexahistidine (His6) tag in the pET28a-LIC expression vector. We expressed recombinant proteins from Rosetta 2 (DE3) Escherichia coli cells by induction with 0.5 mmol/L isopropyl-β-D-thiogalactopyranoside overnight at 16°C and purified recombinant proteins using an ÄKTA fast protein liquid chromatography system (GE Healthcare). This was followed by affinity and size exclusion chromatography.

5'-Nucleotidase assays. We assessed 5′-nucleotidase activity of purified recombinant wild-type and mutant NT5C2 proteins using the 5′-NT enzymatic test kit (Diazyme) according to the manufacturer's instructions, as described previously (11). We calculated 5′-nucleotidase activity levels using a calibrator of known 5′-nucleotidase activity as standard. All assays were performed in triplicate in a Glomax multi-detection system plate reader (Promega).

Western blotting. For the Western blotting, cells were lysed in RIPA buffer plus protease and phospatase inhibitor and benzonase. 30ug of protein were loaded on NuPAGE gels. For detection it was used primary anti-NT5C2 antibody (3C1 Clone, WH0022978M2 Sigma-Aldrich), anti-FPGS antibody (Clone EPR11064, ab184564 Abcam), anti-CREBBP (Clone D6C5, #7389 Cell Signaling), Anti H3K9AcK14AcK18AcK23AcK27Ac (ab47915 Abcam), Anti H3K27Ac (Clone D5E4, #8173 Cell Signaling), Anti-H3 total (Clone 1B1-B2, 61475 Active Motif) in combination with anti-rabbit or anti-mouse HRP-linked secondary antibodies (Cell Signaling Technology) and Amersham's enhanced chemiluminescence Western blotting detection reagent (GE Life Sciences). A horseradish peroxidase-conjugated β-actin antibody was used to detect this protein loading control (clone 8H10D10, #12262 Cell Signaling).

Morpholino treatment. For FPGS splice switching a specific Vivo-Morpholino targeting the canonical 3' splice site of Exon 8 was designed (AAATTCCACTGGTCCGTCTGACCCC). As a control, we used the reverse sequence, as recommended by the manufacturer, in order to maintain sequence composition. Cells were incubated with the 2.5 µM morpholinos for 24h before treatment with MTX.

In vitro killing assay. For methotrexate (S1210 Selleck Chemicals), 6-mercaptopurine (852678-5G-A Sigma Aldrich), mizoribine (S1384 Selleck Chemicals), doxorubicin (HY-15142 MedChem Express), dexamethasone (D4902 Sigma Aldrich), and vincristine (HY-N0488 MedChem Express) treatments, cells were incubated with increasing concentrations of each drug for 72 hours. Vehicle control dilution corresponding to the highest drug concentration was used as a baseline measure for 100% survival. Cell viability was measured by CellTiter-Glo assays (G7572 Promega) following manufacturer's instructions. IC_{50} values were calculated using Prism software v9.3.1 (GraphPad software) with the log (inhibitor) vs. normalized response – variable slope.

In vivo cell line xenograft animal studies. Luciferase-transduced REH cells were infected with NT5C2 empty (wild-type), E6a, or R238W constructs and selected with blasticidin for 96h. NOD/SCID/Il2rgtm1wjl/SzJ (NSG) mice were intravenously injected with 1x106 native or modified REH cells and assessed for leukemia engraftment using an In Vivo Imaging System (Xenogen). Once ALL engraftment was detected ten days later, cohorts of mice (n=5/group) of each REH cell type were randomized for treatment with vehicle control or 6-mercaptopurine 100 mg/kg/day of 6-MP administered intraperitoneally daily x 5 days as described previously (12). Radiance as a surrogate for total leukemia burden was measured by in vivo bioluminescent imaging (BLI) of mice following intraperitoneal injection of luciferin substrate. BLI assesment was performed with Living Image software (Perkin-Elmer) with statistical analysis and data display in Prism (GraphPad). Pilot BLI studies of mice engrafted with each REH cell type were initially conducted to establish in vivo leukemia proliferation kinetics.

Results

A subset of r/r B-ALL has an AS signature affecting multiple chemoresistance genes. To gain new insights into molecular mechanisms that render B-ALL chemotherapy ineffective, we analyzed 48 TARGET patients with available paired diagnosis-relapse (D-R) RNA-seq datasets. We called mutations in known r/r driver genes from the RNA-seq data using Genome Analysis Toolkit (GATK) (13) and, in agreement with prior analyses, observed that in 12 out of 48 D-R pairs (25%) there were no identifiable relapse-specific mutations in 12 common r/r genes (Figure 1a). To determine whether differential gene expression (DGE) could account for the chemoresistant phenotype, we performed the Principal Component Analysis on all 96 samples. We observed no separation of diagnostic and relapse samples (Supplemental Figure 1a, red and cyan dots, respectively), suggesting that DGE as a whole is not a primary driver of B-ALL relapses. However, pervasive AS reported in our earlier B-ALL study (8) could contribute to chemoresistance by altering expression levels of distinct mRNA isoforms. Thus, we asked whether any members of the HNRNP and SRSF/TRA2 superfamilies of splicing factors (SF) are affected by mutations or copy number alterations (CNA). Using cBioPortal (14), we identified such CNAs in 14 out of 48 patients (30%). For example, deep deletions of both SRSF4 and SRSF10 genes was found in patient PAPZST and the SRSF3 gene - in samples PARAKF and PARJSR (Figure 1b). Of

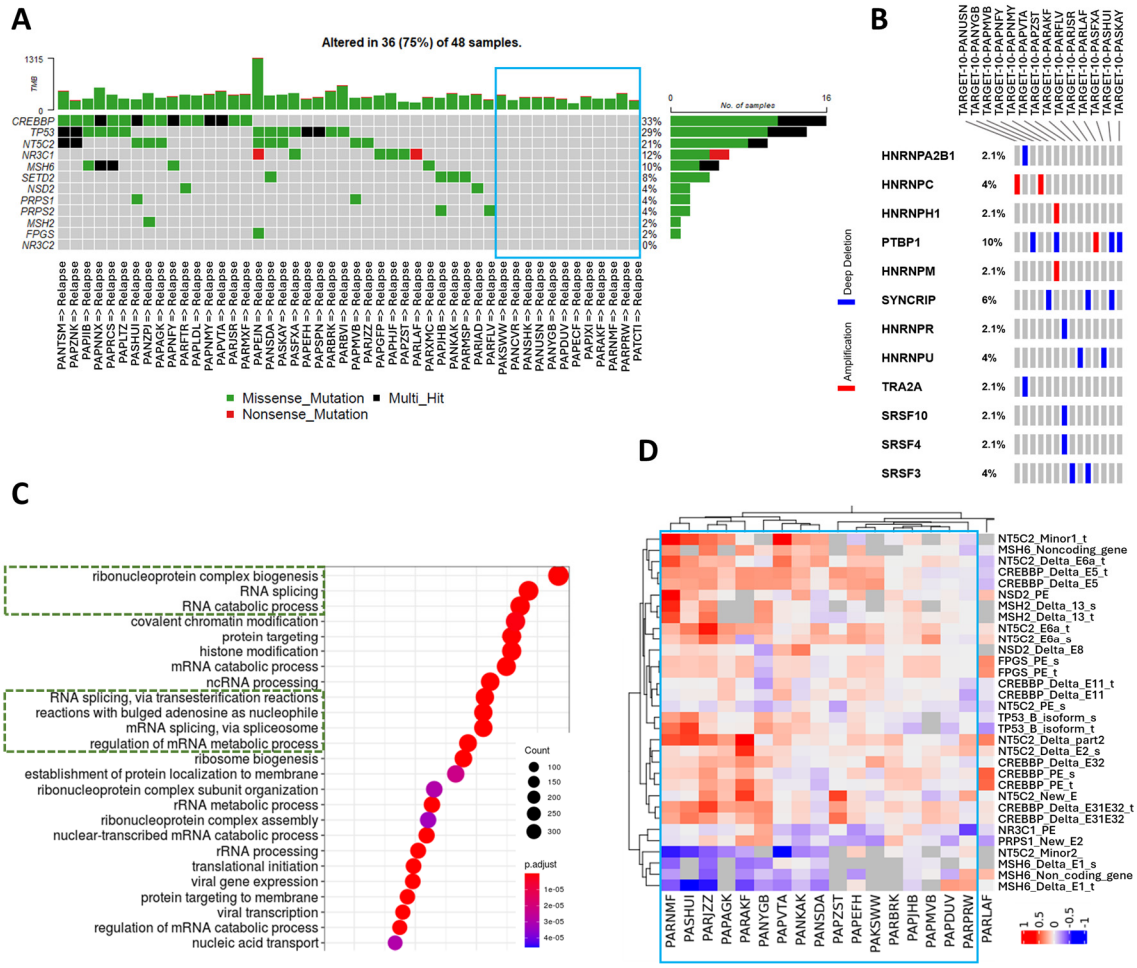


Figure 1. Splicing signature of relapsed leukemias. a. Oncoprint showing distribution of acquired mutations in known relapse/resistance genes in 48 paired TARGET samples. The red rectangle on the right side marks samples with no identifiable relapse-specific mutations. **b,** Oncoprint showing distribution of mutations in select splicing factor genes in 48 TARGET patients under investigation. **c.** Gene Set Enrichment Analysis of transcripts comprising one of the unsupervised clusters. Green rectangles denote GO categories involving RNA splicing. **d. H**eatmap showing changes in splicing in a cluster of 17 relapse vs. diagnostic B-ALL samples (light blue box). The dendrograms on top represent the results of unsupervised clustering without excluding samples with known relapse-specific mutations. The adjacent sample PARLAF with several profoundly mis-spliced transcripts is shown for comparison. The red and blue colors represent increases/decreases in inclusion of exonic segments, respectively.

note, these deletions affect only the latest among longitudinal samples, attesting to their possible role in B-ALL relapses.

To explore the extent and functional consequences of splicing dysregulation, we ran the MAJIQ 2.1 splicing algorithm on paired diagnostic/relapse samples while filtering for local splicing variations (LSV) with ≥20% difference in junction inclusion with 95% confidence (Supplemental Figure 1b). We further filtered for samples without known driver mutations and for LSVs affecting these wild-type drivers. After performing hierarchical clustering, we discovered a cluster of 5 relapses with concurrent mis-splicing of multiple chemoresistance genes (Supplemental Figure 1c, light blue rectangle). To elucidate the underlying molecular mechanisms, we performed gene set enrichment analysis (15) on all transcripts differentially expressed in D-R pairs in that cluster of patients. Top 3 and 7 out of top 12 categories pertained to RNA splicing (Figure 1c). To determine whether AS is only in play in samples with unmutated drivers of chemoresistance or has broader implications, we extended MAJIQ splicing analysis to all 48 D-R pairs. Unsupervised clustering re-discovered the same 5 patient samples and assigned additional 12 r/r leukemias to the same cluster for the total of 17 (Figure 1d, light blue rectangle), potentially implicating AS in ~1/3 of B-ALL relapses. We obtained further evidence to SF dysregulation at the level of poison exon inclusion (16) and ensuing nonsense-mediated decay (NMD) (17). For example, in all relapse samples, almost 100% of transcripts encoding the aforementioned SRSF10 carried stop-codon-containing exon 3a, rendering them non-functional (Supplemental Figure 1d).

Table 1. r/r B-ALL drivers affected by AS. Events **in bold** are under investigation in this study.

Gene Name	Junctions coordinates (hg38)	Affected exons (MANE v1.0 nomenclature)	Effect on protein
CREBBP	**3736815-3739578**	**Δex26**	**In-frame truncation**
CREBBP	**3736815-3740399**	**Δex25-26**	**In-frame truncation**
CREBBP	3745595-3749627 3745354-3745440	"poison" exon 21a inclusion	Nonsense-mediated decay
CREBBP	3774693-3778011	Δex11	In-frame truncation
CREBBP	3782926-3793386	Δex5	In-frame truncation
FPGS	**127807775-127808234**	**ex8 alt 5' site**	**Frameshift**
MSH2	47475270-47478272	Δex13	Frameshift
MSH6	47696097-47708183 47696097-47698818	Alt 5'UTR	Non-coding isoforms
MSH6	47737506-47790927	Δex1	Loss of AUG
NR3C1	143396525-143399656	Novel Δex2a	Early termination
NSD2	1935262-1939654	Δex8	Frameshift
NSD2	1952231-1952752	int11 retention	Non-coding transcript
NT5C2	103100943-103101045	int8 retention	Frameshift
NT5C2	**103105381-103105706**	**Novel exon 6a** (previouslueferred to as exon 4a in *Black et al*)	**In frame insertion**
NT5C2	103143168-103174858	Novel exon 3a	Unknown
NT5C2	103174982-103193236	Δex2	Shorter 5'UTR
NT5C2	103106706-103174858	Δex4	Frameshift
NT5C2	103111805-103174858	alt ex3 5' site/alt exon 4a	Minor isoform
NT5C2	103111805-103139406	alt exon 4a	Minor isoform
PRPS1	107628750-107630017	Novel exon 1a	Early termination
TP53	7673339-7673535	Novel exon 9a	Hypomorphic β isoform

We then considered more carefully the nature of AS events contributing to this cluster (Table 1) as they related to resistance to commonly used B-ALL therapeutics: glucocorticoids (*NR3C1* (18)), thiopurines (*PRPS1* (19) & *NT5C2* (11,20), and anti-folates (*FPGS* (21,22)). In additional, they affected relapse genes associated with more general, not exclusively cancer cell-intrinsic mechanisms. Examples included the master epigenetic regulator *CREBBP* (23-25) and its functional interactor *SETD2* (26) and the pan-cancer suppressor *TP53* recently shown to drive resistance to most B-ALL therapeutics (27,28) . Overall, the majority of the splicing variants in Table 1 are putative loss-of-function events, as they either directly disrupt open reading frames (mimicking nonsense mutations) or render the transcript subject to NMD (17). For instance, *TP53* exon 9i contains a premature stop codon, and its inclusion results in the previously validated p53β protein isoform with multiple functional defects (29,30). However, the effects of *CREBBP* exon 25/26 skipping and the alternative 5' splice site in *FPGS* exon 8 [previously reported in the literature (31-33)] could not be inferred *a priori* in the context of the loss-of-function model and required experimental investigation.

AS events of *CREBBP* and *FPGS* in r/r B-ALL are loss-of-function. While *CREBBP* mRNA was affected by several AS events, in-frame skipping of exons 25 and 26 was of potential functional significance, as it removes 261bp (87aa) from the HAT domain. This splicing event was also among the most robustly detected in our analysis of B-ALL relapses: on average, 50% of all *CREBBP* reads

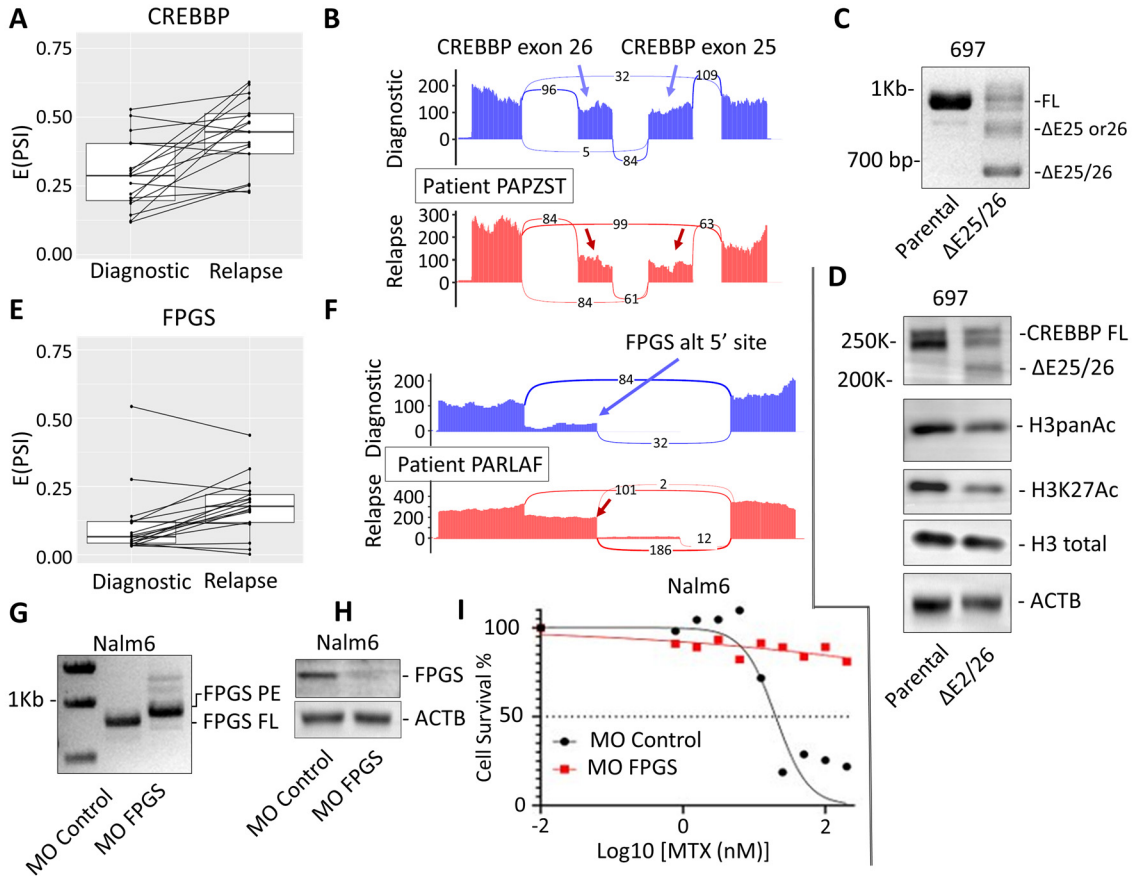


Figure 2. LOF AS events affecting CREBBP and FPGS. a. Box plot showing percentage of read connecting CREBBP exon 24 to exon 27 in diagnosis/relapse pairs depicted in panel 1d. PSI values refer to percent-spliced-in. **b,** Sashimi plots visualizing the same CREBBP events in a representative PAPZST sample. **c.** RT-PCR analysis of CREBBP transcripts in 697 cells transfected with exon 25 and 26-specific sgRNAs packaged into Cas9 particles. **d.** Immunoblotting analysis of the same samples probed with antibodies recognizing indicated proteins and most-translational modifications on histone H3. Beta actin was used as a loading control. **e.** Box plot showing percentage of read connecting 5' splice sites in FPGS exon 8 to the downstream exon 9 in diagnosis/relapse pairs depicted in panel 1d. PSI values refer to percent-spliced-in. **f,** Sashimi plots visualizing the same FPGS event in a representative PARLAF sample. **g.** RT-PCR analysis of FPGS transcripts in Nalm6 cells transfected with control or FPGS exon 8 5' splice site-specific Morpholino. **h.** Immunoblotting analysis of the same samples probed with an anti-FPGS-antibody. **i.** IC50 plot representing survival of these cells after exposure to increasing concentrations of methotrexate (MTX).

corresponded to the resultant ex24-ex27 junction (Figure 2a). In the representative example in Figure 2b, fewer than 30% of reads in that region (32) corresponded to the Δex25/26 isoform in the diagnostic sample (blue sashimi plots), whereas in the pair relapse sample this event predominated (99 reads in the red sashimi plots). To determine the effects of this event on protein function, we designed two sets of short guide RNAs mapping to the exons in question and used them to remove the exon 25-26 segment from 697 B-ALL cells via CRISPR-mediated genome editing. Using semi-quantitative RT-PCR, we detected decreased expression of the *CREBBP* full-length isoform and robust skipping of its exons 25 and/or 26, with combined skipping of both exons being the most readily detectable event (Figure 2c). This was also apparent when protein levels were analyzed using immunoblotting (Figure 2d, top panel). Since the splicing event affects the HAT domain, we chose, as a functional readout, to test its effects on histone modifications. In these experiments, *CREBBP* exon 25/26 skipping resulted in a commensurate reduction in H3K27Ac while preserving total H3 levels (Figure 2d). These results support the notion that *CREBBP* exon 25/26 skipping yields a protein isoform deficient in histone acetyl transferase activity.

The AS event affecting *FPGS* exon 8 was also pervasive: only in one pair of samples was its direction reversed (Figure 2e). In a representative relapse sample depicted in Figure 2f, 65% of reads connected to the downstream exon 9 (186 in the red sashimi plot) originated at the alternative 5' splice site, whereas in the paired diagnostic sample this percentage ratio was reversed (32 reads in the blue sashimi plot). To model this event *in vitro*, we designed a morpholino (MO) antisense splice blocker targeting the canonical 5' splice site and forcing the use of the alternative downstream site. In transfected Nalm6 B-ALL cells this MO almost completely redirected splicing towards the expected downstream site, resulting in partial intron inclusion (Figure 2g). Since this intronic region contains a stop codon, the use of *FPGS* MO also resulted in a sharp decrease in FPGS protein levels, as evidenced by immunoblotting data (Figure 2h). As FPGS is a known driver of sensitivity to methotrexate (21,22), we tested the effects of MO-induced FPGS AS on cell survival in response to this antifolate drug. In these experiments, redirecting splicing of *FPGS* exon 8 towards the downstream splice site resulted in almost complete resistance to methotrexate and unattainable IC50 (Figure 2i). These data solidified our conclusion that AS phenocopies loss-of-function mutations in this gene previously linked to chemoresistance in r/r B-ALL (5).

AS of *NT5C2* yields an mRNA isoform conferring resistance to thiopurines. Across all relapse-associated AS events identified in our analysis (Table 1), the alternative isoform of *NT5C2* with the cryptic in-frame exon 6a [previously referred to by us as exon 4a (8)] stood out. This microexon is annotated in RefSeq, but not included in the The Matched Annotation from the NCBI and EMBL-EBI (MANE), which is a universal standard for clinical reporting (34). Additionally, in GENCODE V44 the sole microexon-containing transcript ENST00000674860.1 is classified as non-coding. Of note, increased frequencies of this event were observed in most relapsed B-ALL samples compared with the matched diagnostic specimens (Figure 3a). In a representative relapse sample depicted in Figure 3b, >80% of reads originating in the upstream canonical exon 6 (352 in the red sashimi plot) connected to exon 6a, whereas in the paired diagnostic sample <25% did (80 reads in the blue sashimi plot). To determine whether *NT5C2* exon 6a can be found in translatable cap-to-poly(A) protein-coding transcripts we amplified its partial cDNA and subjected the resultant 1.7-kb fragment (Supplemental Figure 2a) to targeted resequencing using the Oxford Nanopore Technologies approach (35). Using this assay, we readily detected exon *NT5C2* 6a inclusion in the majority of reads corresponding to the Philadelphia chromosome-like MHH-CALL4 cell line with the IGH::CRLF2 rearrangement (36) (Supplemental Figure 2b). Further support for the translatability of this isoform comes from its inclusion in the Uniprot (37) database under accession code A0A6Q8PHP0 and the inclusion of the exon 6a-encoded tryptic peptide SQVAVQKR in The Peptide Atlas (38) under the accession code PAp02559851.

Inclusion of exon 6a in the *NT5C2* translated sequence results in the addition of 8 extra amino acids near the ATP-binding effector site 2. To determine the effect of this alteration on NT5C2 function, we produced the corresponding recombinant protein (Supplemental Figure 2c,d) and tested its enzymatic activity in a nucleotidase assays with IMP as a substrate under basal conditions and following allosteric

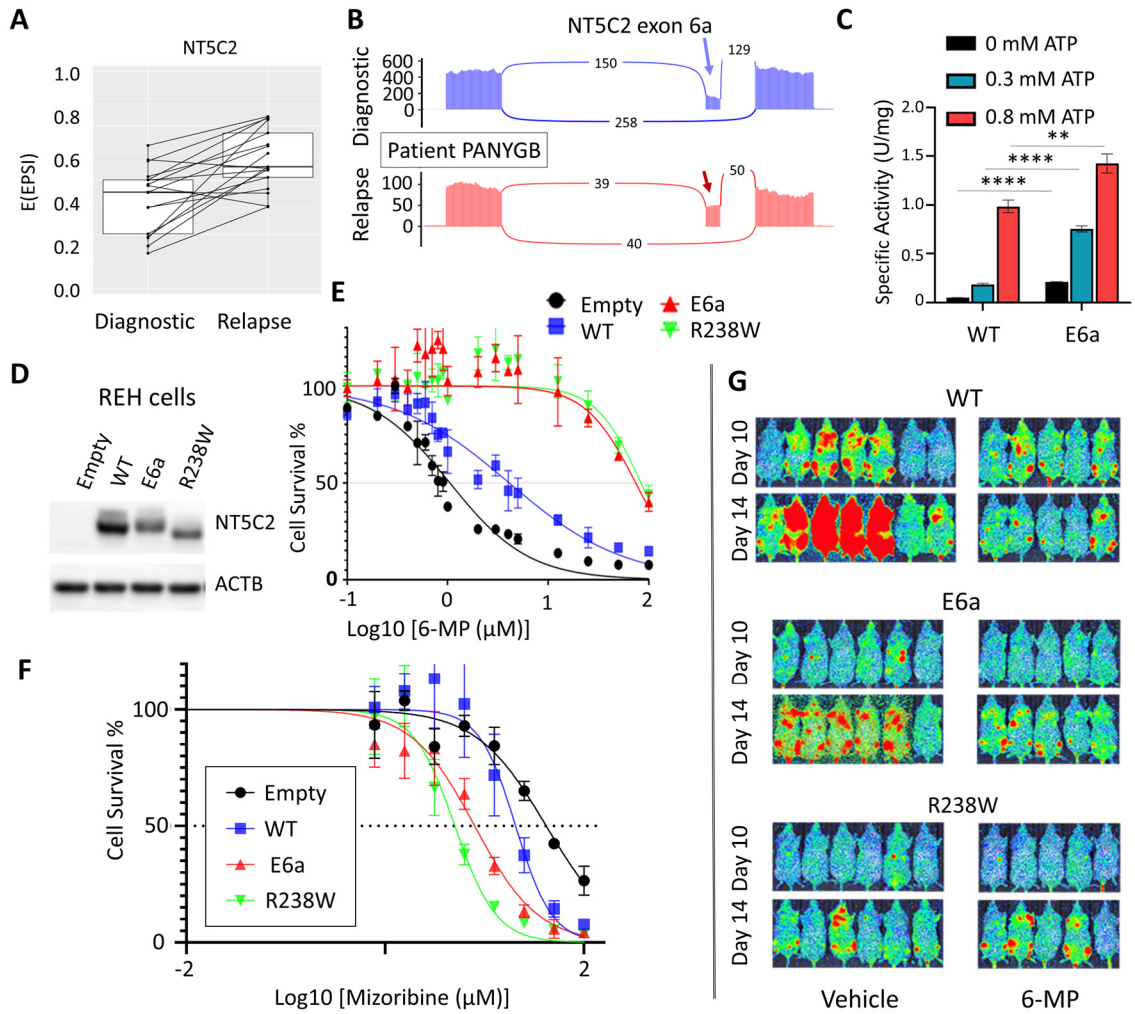


Figure 3. The NT5C2ex6a isoform as a driver of resistance to thiopurines. **a.** Box plot showing percentage of read connecting NT5C2 exon 6 to exon 6a in 6 diagnosis/relapse pairs depicted in panel 1c. PSI values refer to percent-spliced-in. **b.** Sashimi plots visualizing the same NT5C2 events in a representative PARJZZ sample. **c.** In vitro nucleotidase assays assessing the enzymatic activity of the canonical (WT) and E6a NT5C2 isoforms in the presence of increasing concentrations of ATP. Data are shown as mean ± SD. Asterisks indicate statistical significance per Student's t-test. **d.** Expression levels of transduced NT5C2 isoforms in REH cells. "WT", "E6a", and "R238W" denote, respectively, the canonical isoform, the NT5ex6a splice isoform, and the R238W hotspot mutant. "Empty" denotes vector-only cells. **e.** The IC_{50} plot representing survival of these cells exposed to increasing concentrations of 6-MP. **f.** The IC_{50} plot representing survival of these cells exposed to increasing concentrations of mizoribine. **g.** Bioluminescent detection of the same cells additionally expressing the firefly luciferase gene. Cells were xenografted into NSG mice and imaged on days 10 and 14.

activation with ATP (10). In the cell-free assays, bacterially produced NT5C2ex6a exhibited elevated enzymatic activity compared with canonical NT5C2 protein in response to low concentrations of ATP (Figure 3c). This biochemical response was consistent with a gain of function mechanism, similar to that observed for most relapsed-associated *NT5C2* mutations characterized to date (10).

To follow on these results and to formally test the potential role of *NT5C2* ex6a in 6-MP resistance, we generated REH B-ALL cell models expressing canonical *NT5C2*, *NT5C2* ex6a, and *NT5C2* R238W, a validated gain-of-function resistance driver with the most common hotspot mutation (10) (Figure 3d). Cell viability assays carried out with increasing concentrations of 6-MP demonstrated that expression of the *NT5C2* ex6a isoform triggers 6-MP resistance comparable to that induced by the *NT5C2* R238W mutation and well above that conferred by the canonical isoform of wild-type *NT5C2* (Figure 3e and Supplemental Figure 2e), although the effects were somewhat less potent in Nalm-6 cells (Supplemental Figure 2f). Metabolically, increased enzymatic activity of NT5C2 mutants results in increased clearance of thiopurine mononucleotide monophosphate metabolites, but also in depletion of

intracellular nucleotide pools, which in turn renders leukemia cells more sensitive to mizoribine, an inosine monophosphate dehydrogenase (IMPDH) inhibitor (12). Therefore, we evaluated the relationship between *NT5C2* ex6a and sensitivity to this drug. Once again, both *NT5C2* R238W- and *NT5C2* ex6a-expressing cells showed convergent phenotypes with respect to increased sensitivity to mizoribine treatment compared with cells expressing the canonical isoform of wild-type *NT5C2* (Figure 3f and Supplemental Figure 2g).

We further noted that the NT5C2 exon 6a-encoded peptide SQVAVQKR contains a characteristic SQ motif, which constitutes the recognition site for the DNA-PK/ATR/ATM subfamily of kinases of the PIKK family (39), also reflected in the Kinase Library (40). This observation suggested a potential role for protein phosphorylation at this site in the regulation of NT5C2 activity. To test this possibility, we generated mutant forms of NT5C2 ex6a, in which the ex6a-encoded serine residue was replaced by the common phosphomimetic aspartic acid (S131D) or by alanine (S131A) as phosphorylation-deficient control. Evaluation of responses to 6-MP treatment of REH cells expressing each of these constructs revealed that while the parental NT5C2 ex6a isoform and its S131A derivative conferred equivalent levels of 6-MP resistance, the S131D substitution resulted in the markedly increased resistance phenotype, with one log further increase in IC50 (Supplemental Figure 2h). These results suggest that phosphorylation events at the S131 site augment *NT5C2* exon 6a-driven 6-MP resistance.

To directly assess the impact of *NT5C2* ex6a inclusion on leukemia cells growth and therapy responses *in vivo*, we implanted luciferase-expressing REH cells expressing the canonical wild type *NT5C2* isoform, *NT5C2* R238R and *NT5C2* ex6a into NSG mice and imaged animals on days 5, 8, 12, and 14. Previous reports linked hyperactive NT5C2 with a “fitness cost” phenotype and decreased cell proliferation resulting from the depletion of intracellular nucleotide pools (12). Consistent with these observations, expression of *NT5C2* R238W and *NT5C2* ex6a induced clear growth retardation when compared with expression of the canonical *NT5C2* isoform (Supplemental Figure 3a). Nevertheless, we observed robust leukemia engraftment, allowing us to test the effect of the three different *NT5C2* variants on chemotherapy response by treating each experimental group with either vehicle or 6-MP (Figure 3g). When compared to the initial treatment timepoint (day 10 post-engraftment), the WT model shows a negative fold change (plotted on the Y-axis in Supplemental Figure 3b), which was indictive of cell death. In contrast, the E6a model demonstrates slower cell division under treatment, but the fold change remains non-negative, signifying that the treatment does not induce cell death. Notably, the E6a model exhibits an identical fold change to the R238W mutant model. These findings further support the equivalent roles of these two NT5C2 variants in promoting chemoresistance.

Discussion

The overarching goal of precision medicine is matching each patient to the best available treatment. Towards this end, modern cancer diagnostics labs employ multi-plexed panels capable of identifying mutations in >100 genes known or suspected to be cancer drivers and/or predictive/prognostic biomarkers. This clinical effort is designed to identify original driver mutations as well as mutations acquired under therapeutic pressure. In the best-case scenario, somatic mutations acquired in these patients might be druggable. Salient examples include BCR::ABL1 and ABL class fusions as targetable by imatinib or dasatinib, CRLF2, JAK2, and EPOR rearrangements as targetable by ruxolitinib, and KMT2A rearrangements as targetable by menin inhibitors [reviewed in (41)]. However, most pediatric leukemias, especially B-ALL, are mutationally “quiet”, limiting the utility of gene panels as currently implemented. A recent clinical study performed at our institution showed that only 12% patients with leukemia (mostly acute myelogenous leukemia) treated at our institution have identifiable mutations that qualify them for a specific targeted therapy (42). While epigenetic mechanisms, such as DNA methylation and histone modifications, have also emerged as important players in many types of cancer, their diagnostic and especially therapeutic utilities in B-ALL remain limited (43), especially in B-ALL (44).

Yet more recently, RNA-centric mechanisms such as formation of fusion transcripts and circular RNAs, RNA editing, and in particular aberrant mRNA splicing have been recognized as key drivers of neoplastic growth (45,46). Indeed mutations in genes encoding splicing factors (SF3B1, SRSF2, etc.)

are very common in adult leukemias [reviewed in (47)], raising the possibility that targeting aberrant splicing programs could be a viable therapeutic option (48). Admittedly, such mutations are exceedingly rare in pediatric AML and ALL; however, a growing body of work by our group and others has revealed that inclusion of “poison” and other non-canonical exons in SF-encoding transcripts is among the most consistent splicing aberration in both B-ALL and T-ALL (8,49) as well as in many solid tumors without “hardwired” changes in the splicing machinery (16,50).

Such AS SF events are richly represented in the TARGET cohort of pediatric patients with B-ALL, many with relapsed/refractory disease (8). Here we report that among their important consequences is the pervasive dysregulation, at the level of mRNA splicing, of key chemosensitivity genes. In most cases (CREBBP, FPGS, and other examples in Table 1) the AS events are of the loss-of-function nature, making it challenging to devise pharmacological interventions. Still, an RNA-based diagnostic test detecting, for example, the heavy usage of FPGS ex8 alt 5’ site could guide clinical decisions regarding the use of methotrexate. Thus, we envision that targeted RNA-seq panels, either short- or long-read, will become valuable companion diagnostics, despite known limitations of underlying technologies (51). In fact, their utility might further extend to immunotherapy cases, where the presence of abnormally processed mRNAs encoding surface proteoforms have been shown to interfere with the use of chimeric antigen receptors and bispecific T cell engagers (52-57).

The case of NT5C2 is particularly interesting from the therapeutic standpoint, since NT5C2 splicing aberrations leading to the inclusion of cryptic exon 6a ultimately create a proteoform with gain-of-function properties. This was apparent in both cell-free nucleotidase assays (Figure 3c) and in vitro and in vivo 6-MP resistance assays performed with engineered B-ALL cell lines (Figure 3e,g, Supplemental Figure 2f,h, Supplemental Figure 2b). Not only could robust expression of the NT5C2ex6a isoform predict resistance to thiopurines, but may suggest additional therapeutic options such as treatment with the emerging class of direct NT5C2 small-molecule inhibitors such as CRCD2 (58) and also with the already FDA-approved drug mizoribine (59). Indeed, NT5C2ex6a-expressing REH and NALM6 cells were found to exhibit heightened sensitivity to mizoribine compared to the isogenic cell lines reconstituted with the canonical NT5C2 isform (Figure 3f, Supplemental Figure 2g). Of interest, at least in one leukemia (PARJZZ) NT5C2 exon 6a inclusion (Figure 1D) co-exists with the hotspot mutation R367Q reported earlier (4), raising interesting questions about tumor heterogeneity or possible cooperation between genetic and transcriptional mechanisms of resistance.

Other ways to therapeutically target NT5C2 dysregulation are likely to exist. Our recently published work elucidated the role of NT5C2 post-transcriptional modification in chemoresistance, with phosphorylation of the Ser-502 residue by a yet-to-be-identified protein kinase increasing its activity against 6-MP (58). Our new data showing that substituting Ser at position 131 with a phosphomimetic Asp augments NT5C2ex6a-induced resistance to 6-MP (Supplemental Figure 2h) lend further credence to the model linking NT5C2 phosphorylation and chemoresistance function. If further research confirms the involvement of ATR/ATM family kinases in phosphorylating the NT5C2 exon 6a-encoded SQ motif, a variety of small-molecule inhibitors of these enzymes (60) will be available for preclinical studies and clinical trials aiming to re-sensitize r/r B-ALL to chemotherapy and in doing so improve outcomes in this aggressive childhood cancer.

Keywords: acute lymphoblastic leukemia, chemoresistance, splicing, thiopurines, NT5C2

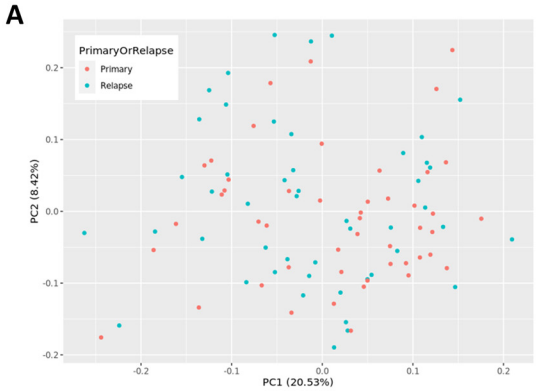
Financial support: This work was supported by grants from the National Institutes of Health (U01 CA232563 and U01 CA232563-S3 to ATT and YB, and U01 CA232486 and U01 CA243072 to SKT), United States Department of Defense (CA180683P1 to SKT), St. Baldrick's-Stand Up to Cancer (SU2C-AACR-DT-27-17 to ATT and SKT), The V Foundation for Cancer Research (T2018-014 to ATT), The Emerson Collective (886246066 to ATT). AC acknowledges support from CHOP SPROUT Program. SKT is a Scholar of the Leukemia & Lymphoma Society. SKT is Joshua Kahan Endowed Chair in Pediatric Leukemia Research and ATT is Mildred L. Roeckle Endowed Chair in Pathology at Children’s Hospital of Philadelphia.

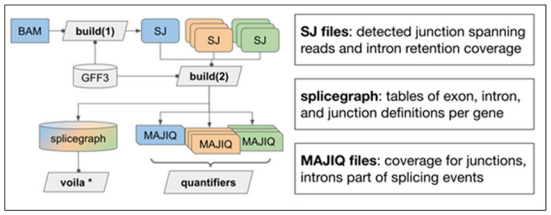
References

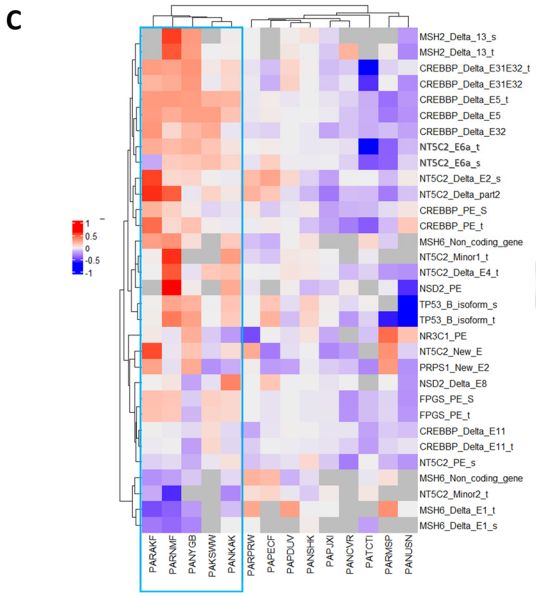
1. Hunger SP, Raetz EA. How I treat relapsed acute lymphoblastic leukemia in the pediatric population. Blood 2020;136(16):1803-12.
2. Roberts KG, Mullighan CG. Genomics in acute lymphoblastic leukaemia: insights and treatment implications. Nat Rev Clin Oncol 2015;12(6):344-57.
3. Jeha S, Choi J, Roberts KG, Pei D, Coustan-Smith E, Inaba H*, et al.* Clinical significance of novel subtypes of acute lymphoblastic leukemia in the context of minimal residual disease-directed therapy. Blood Cancer Discov 2021;2(4):326-37.
4. Ma X, Edmonson M, Yergeau D, Muzny DM, Hampton OA, Rusch M*, et al.* Rise and fall of subclones from diagnosis to relapse in pediatric B-acute lymphoblastic leukaemia. Nat Commun 2015;6:6604.
5. Li B, Brady SW, Ma X, Shen S, Zhang Y, Li Y*, et al.* Therapy-induced mutations drive the genomic landscape of relapsed acute lymphoblastic leukemia. Blood 2020;135(1):41-55.
6. Oshima K, Khiabanian H, da Silva-Almeida AC, Tzoneva G, Abate F, Ambesi-Impiombato A*, et al.* Mutational landscape, clonal evolution patterns, and role of RAS mutations in relapsed acute lymphoblastic leukemia. Proc Natl Acad Sci U S A 2016;113(40):11306-11.
7. Waanders E, Gu Z, Dobson SM, Antic Z, Crawford JC, Ma X*, et al.* Mutational landscape and patterns of clonal evolution in relapsed pediatric acute lymphoblastic leukemia. Blood Cancer Discov 2020;1(1):96-111.
8. Black KL, Naqvi AS, Asnani M, Hayer KE, Yang SY, Gillespie E*, et al.* Aberrant splicing in B-cell acute lymphoblastic leukemia. Nucleic Acids Res 2018;46(21):11357-69.
9. Vaquero-Garcia J, Barrera A, Gazzara MR, Gonzalez-Vallinas J, Lahens NF, Hogenesch JB*, et al.* A new view of transcriptome complexity and regulation through the lens of local splicing variations. eLife 2016;5:e11752.
10. Dieck CL, Tzoneva G, Forouhar F, Carpenter Z, Ambesi-Impiombato A, Sanchez-Martin M*, et al.* Structure and Mechanisms of NT5C2 Mutations Driving Thiopurine Resistance in Relapsed Lymphoblastic Leukemia. Cancer Cell 2018;34(1):136-47.
11. Tzoneva G, Perez-Garcia A, Carpenter Z, Khiabanian H, Tosello V, Allegretta M*, et al.* Activating mutations in the NT5C2 nucleotidase gene drive chemotherapy resistance in relapsed ALL. Nat Med 2013;19(3):368-71.
12. Tzoneva G, Dieck CL, Oshima K, Ambesi-Impiombato A, Sanchez-Martin M, Madubata CJ*, et al.* Clonal evolution mechanisms in NT5C2 mutant-relapsed acute lymphoblastic leukaemia. Nature 2018;553(7689):511-4.
13. Brouard JS, Bissonnette N. Variant Calling from RNA-seq Data Using the GATK Joint Genotyping Workflow. Methods Mol Biol 2022;2493:205-33.
14. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA*, et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012;2(5):401-4.
15. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA*, et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102(43):15545-50.
16. Thomas JD, Polaski JT, Feng Q, De Neef EJ, Hoppe ER, McSharry MV*, et al.* RNA isoform screens uncover the essentiality and tumor-suppressor activity of ultraconserved poison exons. Nat Genet 2020;52(1):84-94.
17. Popp MW, Maquat LE. Organizing principles of mammalian nonsense-mediated mRNA decay. Annu Rev Genet 2013;47(1):139-65.
18. Kuster L, Grausenburger R, Fuka G, Kaindl U, Krapf G, Inthal A*, et al.* ETV6/RUNX1-positive relapses evolve from an ancestral clone and frequently acquire deletions of genes implicated in glucocorticoid signaling. Blood 2011;117(9):2658-67.
19. Li B, Li H, Bai Y, Kirschner-Schwabe R, Yang JJ, Chen Y*, et al.* Negative feedback-defective PRPS1 mutants drive thiopurine resistance in relapsed childhood ALL. Nat Med 2015;21(6):563-71.
20. Meyer JA, Wang J, Hogan LE, Yang JJ, Dandekar S, Patel JP*, et al.* Relapse-specific mutations in NT5C2 in childhood acute lymphoblastic leukemia. Nat Genet 2013;45(3):290-4.
21. Rots MG, Pieters R, Peters GJ, Noordhuis P, van Zantwijk CH, Kaspers GJ*, et al.* Role of folylpolyglutamate synthetase and folylpolyglutamate hydrolase in methotrexate accumulation and polyglutamylation in childhood leukemia. Blood 1999;93(5):1677-83.

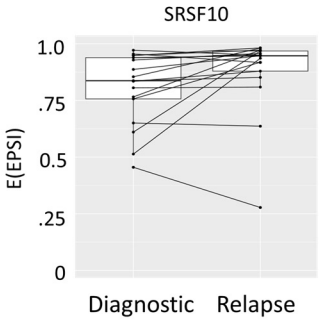
22. Lopez-Lopez E, Autry RJ, Smith C, Yang W, Paugh SW, Panetta JC*, et al.* Pharmacogenomics of intracellular methotrexate polyglutamates in patients' leukemia cells in vivo. J Clin Invest 2020;130(12):6600-15.
23. Mullighan CG, Zhang J, Kasper LH, Lerach S, Payne-Turner D, Phillips LA*, et al.* CREBBP mutations in relapsed acute lymphoblastic leukaemia. Nature 2011;471(7337):235-9.
24. Gang EJ, Hsieh YT, Pham J, Zhao Y, Nguyen C, Huantes S*, et al.* Small-molecule inhibition of CBP/catenin interactions eliminates drug-resistant clones in acute lymphoblastic leukemia. Oncogene 2014;33(17):2169-78.
25. Zhang J, Mullighan CG, Harvey RC, Wu G, Chen X, Edmonson M*, et al.* Key pathways are frequently mutated in high-risk childhood acute lymphoblastic leukemia: a report from the Children's Oncology Group. Blood 2011;118(11):3080-7.
26. Mar BG, Bullinger LB, McLean KM, Grauman PV, Harris MH, Stevenson K*, et al.* Mutations in epigenetic regulators including SETD2 are gained during relapse in paediatric acute lymphoblastic leukaemia. Nat Commun 2014;5:3469.
27. Oshima K, Zhao J, Perez-Duran P, Brown JA, Patino-Galindo JA, Chu T*, et al.* Mutational and functional genetics mapping of chemotherapy resistance mechanisms in relapsed acute lymphoblastic leukemia. Nat Cancer 2020;1(11):1113-27.
28. Yang F, Brady SW, Tang C, Sun H, Du L, Barz MJ*, et al.* Chemotherapy and mismatch repair deficiency cooperate to fuel TP53 mutagenesis and ALL relapse. Nat Cancer 2021;2(8):819-34.
29. Bourdon JC, Fernandes K, Murray-Zmijewski F, Liu G, Diot A, Xirodimas DP*, et al.* p53 isoforms can regulate p53 transcriptional activity. Genes Dev 2005;19(18):2122-37.
30. Tang Y, Horikawa I, Ajiro M, Robles AI, Fujita K, Mondal AM*, et al.* Downregulation of splicing factor SRSF3 induces p53beta, an alternatively spliced isoform of p53 that promotes cellular senescence. Oncogene 2013;32(22):2792-8.
31. Stark M, Wichman C, Avivi I, Assaraf YG. Aberrant splicing of folylpolyglutamate synthetase as a novel mechanism of antifolate resistance in leukemia. Blood 2009;113(18):4362-9.
32. Wojtuszkiewicz A, Raz S, Stark M, Assaraf YG, Jansen G, Peters GJ*, et al.* Folylpolyglutamate synthetase splicing alterations in acute lymphoblastic leukemia are provoked by methotrexate and other chemotherapeutics and mediate chemoresistance. Int J Cancer 2016;138(7):1645-56.
33. Zhou X, Wang J, Patel J, Valentine M, Shao Y, Newman S*, et al.* Exploration of Coding and Non-coding Variants in Cancer Using GenomePaint. Cancer Cell 2021;39(1):83-95.e4.
34. Morales J, Pujar S, Loveland JE, Astashyn A, Bennett R, Berry A*, et al.* A joint NCBI and EMBL-EBI transcript set for clinical genomics and research. Nature 2022;604(7905):310-5.
35. Schulz L, Torres-Diz M, Cortes-Lopez M, Hayer KE, Asnani M, Tasian SK*, et al.* Direct long-read RNA sequencing identifies a subset of questionable exitrons likely arising from reverse transcription artifacts. Genome Biol 2021;22(1):190.
36. Ghandi M, Huang FW, Jane-Valbuena J, Kryukov GV, Lo CC, McDonald ER, 3rd*, et al.* Next-generation characterization of the Cancer Cell Line Encyclopedia. Nature 2019;569(7757):503-8.
37. TheUniProtConsortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res 2019;47(D1):D506-D15.
38. Desiere F, Deutsch EW, King NL, Nesvizhskii AI, Mallick P, Eng J*, et al.* The PeptideAtlas project. Nucleic Acids Res 2006;34(suppl_1):D655-D8.
39. Traven A, Heierhorst J. SQ/TQ cluster domains: concentrated ATM/ATR kinase phosphorylation site regions in DNA-damage-response proteins. Bioessays 2005;27(4):397-407.
40. Johnson JL, Yaron TM, Huntsman EM, Kerelsky A, Song J, Regev A*, et al.* An atlas of substrate specificities for the human serine/threonine kinome. Nature 2023;613(7945):759-66.
41. Raetz EA, Bhojwani D, Devidas M, Gore L, Rabin KR, Tasian SK*, et al.* Children's Oncology Group blueprint for research: Acute lymphoblastic leukemia. Pediatr Blood Cancer 2023;70 Suppl 6(S6):e30585.
42. Surrey LF, MacFarland SP, Chang F, Cao K, Rathi KS, Akgumus GT*, et al.* Clinical utility of custom-designed NGS panel testing in pediatric tumors. Genome Med 2019;11(1):32.
43. Jing D, Huang Y, Liu X, Sia KCS, Zhang JC, Tai X*, et al.* Lymphocyte-Specific Chromatin Accessibility Pre-determines Glucocorticoid Resistance in Acute Lymphoblastic Leukemia. Cancer Cell 2018;34(6):906-21 e8.
44. Nordlund J, Syvänen AC. Epigenetics in pediatric acute lymphoblastic leukemia. Semin Cancer Biol 2018;51:129-38.
45. Dvinge H, Kim E, Abdel-Wahab O, Bradley RK. RNA splicing factors as oncoproteins and tumour suppressors. Nat Rev Cancer 2016;16(7):413-30.
46. Stanley RF, Abdel-Wahab O. Dysregulation and therapeutic targeting of RNA splicing in cancer. Nature Cancer 2022;3(5):536-46.

47. Choi PS, Thomas-Tikhonenko A. RNA-binding proteins of COSMIC importance in cancer. J Clin Invest 2021;131(18):e151627.
48. Bradley RK, Anczukow O. RNA splicing dysregulation and the hallmarks of cancer. Nat Rev Cancer 2023.
49. Zhou Y, Han C, Wang E, Lorch AH, Serafin V, Cho BK*, et al.* Posttranslational Regulation of the Exon Skipping Machinery Controls Aberrant Splicing in Leukemia. Cancer Discov 2020;10(9):1388-409.
50. Leclair NK, Brugiolo M, Urbanski L, Lawson SC, Thakar K, Yurieva M*, et al.* Poison Exon Splicing Regulates a Coordinated Network of SR Protein Expression during Differentiation and Tumorigenesis. Mol Cell 2020;80(4):648-65 e9.
51. Byron SA, Van Keuren-Jensen KR, Engelthaler DM, Carpten JD, Craig DW. Translating RNA sequencing into clinical diagnostics: opportunities and challenges. Nat Rev Genet 2016;17(5):257-71.
52. Sotillo E, Barrett DM, Black KL, Bagashev A, Oldridge D, Wu G*, et al.* Convergence of Acquired Mutations and Alternative Splicing of CD19 Enables Resistance to CART-19 Immunotherapy. Cancer Discov 2015;5(12):1282-95.
53. Zheng S, Gillespie E, Naqvi AS, Hayer KE, Ang Z, Torres-Diz M*, et al.* Modulation of CD22 Protein Expression in Childhood Leukemia by Pervasive Splicing Aberrations: Implications for CD22-Directed Immunotherapies. Blood Cancer Discov 2022;3(2):103-15.
54. Cortes-Lopez M, Schulz L, Enculescu M, Paret C, Spiekermann B, Quesnel-Vallieres M*, et al.* High-throughput mutagenesis identifies mutations and RNA-binding proteins controlling CD19 splicing and CART-19 therapy resistance. Nat Commun 2022;13(1):5570.
55. Zhao Y, Aldoss I, Qu C, Crawford JC, Gu Z, Allen EK*, et al.* Tumor-intrinsic and -extrinsic determinants of response to blinatumomab in adults with B-ALL. Blood 2021;137(4):471-84.
56. Witkowski MT, Lee S, Wang E, Lee AK, Talbot A, Ma C*, et al.* NUDT21 limits CD19 levels through alternative mRNA polyadenylation in B cell acute lymphoblastic leukemia. Nat Immunol 2022;23(10):1424-32.
57. Ang Z, Paruzzo L, Hayer KE, Schmidt C, Torres Diz M, Xu F*, et al.* Alternative splicing of its 5'-UTR limits CD20 mRNA translation and enables resistance to CD20-directed immunotherapies. Blood 2023;142(20):1724-39.
58. Reglero C, Dieck CL, Zask A, Forouhar F, Laurent AP, Lin WW*, et al.* Pharmacologic Inhibition of NT5C2 Reverses Genetic and Nongenetic Drivers of 6-MP Resistance in Acute Lymphoblastic Leukemia. Cancer Discov 2022;12(11):2646-65.
59. Valvezan AJ, McNamara MC, Miller SK, Torrence ME, Asara JM, Henske EP*, et al.* IMPDH inhibitors for antitumor therapy in tuberous sclerosis complex. JCI Insight 2020;5(7).
60. Morgado-Palacin I, Day A, Murga M, Lafarga V, Anton ME, Tubbs A*, et al.* Targeting the kinase activities of ATR and ATM exhibits antitumoral activity in mouse models of MLL-rearranged AML. Sci Signal 2016;9(445):ra91.

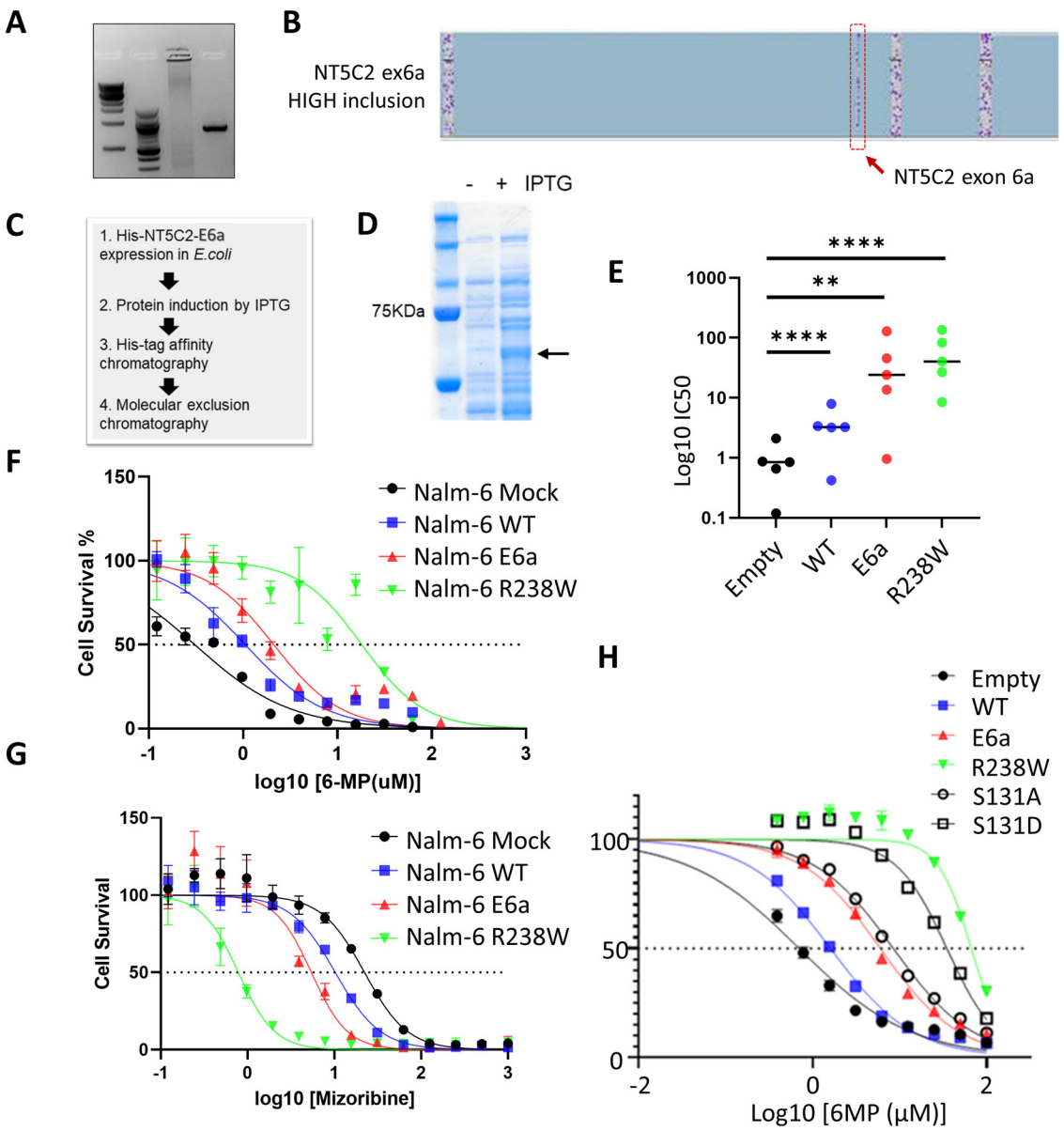




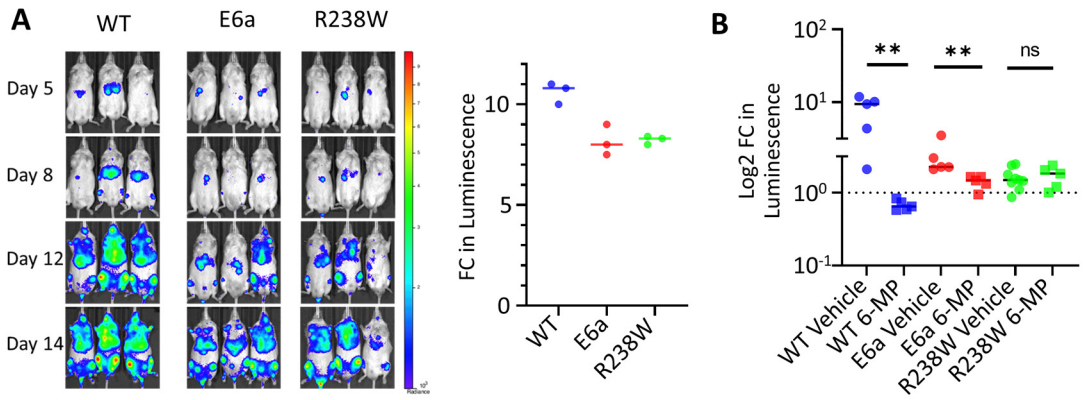




Supplemental Figure 1. **a.** Principal Component Analysis performed in 48 pairs samples (96 total) using RNA-seq data. Red and blue dots represent diagnostic and relapse samples respectively. **b.** MAJIQ-based workflow to identify relapse-specific splicing events. See text for more explanations. **c.** Heatmap showing changes in splicing in a cluster of 14 relapse vs. diagnostic B-ALL samples (light blue box) lacking known relapse-specific mutations. The dendrograms on top represent the results of unsupervised clustering. The red and blue colors represent increases/decreases in inclusion of exonic segments, respectively. **d.** Box plot showing percentage of read corresponding to SRSF10 poison exon inclusion in diagnosis/relapse pairs depicted in panel 1d.



Supplemental Figure 2. The role of NT5C2ex6a in chemoresistance in vitro. **a.** RT-PCR amplification of the full-length NT5C2ex6a cDNA from MHH-CALL4 cells. **b**. Targeted long-read ONT sequencing of the NT5C2ex6a. **c.** General schema of NT5C2 ex6a protein purification. **d.** Electrophoretic detection of the NT5C2 ex6a isoform (arrow) following induction by IPTG. **e.** Average IC50 values from 5 independent experiments depicted in Figure 3e. Asterisks indicate statistical significance per Student's t-test. **f.** The IC_{50} plot representing survival of NT5C2-expressing Nalm6 cells exposed to increasing concentrations of 6-MP. **g.** The IC_{50} plot representing survival of NT5C2-expressing Nalm6 cells exposed to increasing concentrations of mizoribine. **h.** The IC_{50} plot representing survival of additional NT5C2 variant-expressing REH cells exposed to increasing concentrations of 6-MP.



Supplemental Figure 3. The role of NT5C2ex6a in chemoresistance in vivo. a. LEFT: Ivis bioluminescence image of mice engrafted with REH Luc NT5C2 WT, E6a or R238W cells at the indicated time points. RIGHT: Quantitation of data from the previous panel, expressed as fold change in photon counts between d14 and d5. **b.** Quantitation of data from Figure 3g, expressed as fold change in photon counts between d14 and d10. Asterisks indicate statistical significance per Student's t-test.