HHS Public Access

Author manuscript

Nat Genet. Author manuscript; available in PMC 2013 September 01.

Published in final edited form as:

Nat Genet. 2013 March ; 45(3): 290–294. doi:10.1038/ng.2558.

Relapse specific mutations in *NT5C2* in childhood acute lymphoblastic leukemia

Julia A. Meyer[1,2], **Jinhua Wang**[1,3], **Laura E. Hogan**[1,4,*], **Jun J. Yang**[5], **Smita Dandekar**[1,4], **Jay P. Patel**[6], **Zuojian Tang**[3], **Paul Zumbo**[8,9], **Sheng Li**[8,9], **Jiri Zavadil**[1,3], **Ross L. Levine**[6,7], **Timothy Cardozo**[10], **Stephen P. Hunger**[11], **Elizabeth A. Raetz**[1,4], **William E. Evans**[5], **Debra J. Morrison**[1,4], **Christopher E. Mason**[8,9], and **William L. Carroll**[1,2,4,#]

[1]NYU Cancer Institute, NYU Langone Medical Center, New York, NY

[2]Department of Pathology, NYU Langone Medical Center, New York, NY

[3]NYU Center for Health Informatics and Bioinformatics, NYU Langone Medical Center, New York, NY

[4]Department of Pediatrics, NYU Langone Medical Center, New York, NY

[5]Department of Pharmaceutical Sciences, St. Jude Children's Research Hospital, New York, NY

[6]Human Oncology and Pathogenesis Program, Memorial Sloan-Kettering Cancer Center, New York, NY

[7]Leukemia Service, Memorial Sloan-Kettering Cancer Center, New York, NY

[8]Department of Physiology and Biophysics, Weill Cornell Medical College of Cornell University, New York, NY

[9]HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine, Weill Cornell Medical College of Cornell University, New York, NY

[10]Department of Pharmacology, NYU Langone Medical Center, New York, NY

[11]University of Colorado School of Medicine and Children's Hospital Colorado, Aurora CO

Abstract

Relapsed childhood acute lymphoblastic leukemia (ALL) carries a poor prognosis despite intensive retreatment, due to intrinsic drug resistance[1-2]. The biological pathways that mediate

[#]Address for Correspondence: William L. Carroll MD NYU Cancer Institute Smilow 1201 522 First Avenue New York NY 10016 Phone: 212-263-3276 Fax: 212-263-9190 william.carroll@nyumc.org.

[*]Current address: Stony Brook University, Stony Brook, NY

Author Contributions

J.A.M., L.E.H., J.J.Y, J.Z., R.L.L., T.C., W.E.E., D.J.M., C.M. and W.L.C. planned experiments. J.A.M., L.E.H., J.J.Y., S.D., J.P.P., and D.J.M. performed experiments and analyzed data. J.W., Z.T., P.Z., S.L., and C.M. performed sequencing and analyzed sequence data. T.C. performed molecular modeling. S.P.H. and E.A.R. provided patient samples and clinical data. J.A.M and W.L.C. wrote the manuscript. W.L.C. coordinated the study. All authors discussed the results and reviewed the manuscript.

The authors declare no competing financial interests.

Accession codes. Next-generation sequence data is available at NCBI Sequence Read Archive SRA048657.

resistance are unknown. Here we report the transcriptome profiles of matched diagnosis and relapse bone marrow specimens from ten pediatric B lymphoblastic leukemia patients using RNA-sequencing. Transcriptome sequencing identified 20 newly acquired novel non-synonymous mutations not present at initial diagnosis, of which two patients harbored relapse specific mutations in the same gene, *NT5C2*, a 5′-nucleotidase. Full exon sequencing of *NT5C2* was completed in 61 additional relapse specimens, identifying five additional cases. Enzymatic analysis of mutant proteins revealed that base substitutions conferred increased enzymatic activity and resistance to treatment with nucleoside analogue therapies. Clinically, all patients who harbored *NT5C2* mutations relapsed early, or within 36 months of initial diagnosis (p=0.03). These results suggest that mutations in *NT5C2* are associated with the outgrowth of drug resistant clones in ALL.

Acute lymphoblastic leukemia (ALL) is the most common pediatric malignancy, accounting for greater than 25% of all childhood cancers[3]. Cure rates for ALL have improved dramatically over the past four decades with the development of risk stratified multi-agent chemotherapy, preventive treatment to the central nervous system and the more recent introduction of augmented doses/schedules of conventional drugs, resulting in an overall five year event-free survival now exceeding 90%[4]. In spite of these improvements, 10-20% of patients experience disease recurrence[5]. The prognosis for these children is dismal[6], even with aggressive salvage strategies involving allogeneic stem cell transplant[7-8]. Relapsed ALL remains one of the leading causes of mortality for all childhood malignancies. To discover pathways that mediate B lymphoblastic leukemia (the most common subtype) relapse we sought to catalog somatic changes enriched at relapse, reasoning that such lesions may be associated with drug resistance and can provide insight into more effective treatment regimens.

B lymphoblastic leukemia patient specimens (Supplementary Table 1) profiled using transcriptome sequencing generated an average of 84 million reads per specimen (Supplementary Table 2 and 3) and showed very strong correlation (>90% genotype concordance for >8x coverage) to previously analyzed heterozygous SNP calls from Affymetrix SNP 6.0 arrays of the same specimens (Supplemental Fig. 1)[9]. Relapse-specific single nucleotide variants (SNVs) were predicted using GATK[10-11] (Supplementary Fig. 2 and 3) and subjected to PCR validation by Sanger sequencing from corresponding remission, diagnosis, and relapse genomic DNA specimens. We identified 20 missense mutations (Table 1) that were specifically found in the relapse specimens, but absent from both remission (e.g. germline) and diagnosis (leukemia) DNA. Patient samples harbored between one and six relapse-specific variants each. All of the mutations were hemizygous, with expression of the wild type allele as well. In most cases the mutations were predicted to have a deleterious effect on protein structure that would indicate a dominant negative property or a state of haploinsufficiency. Predominant nucleotide changes were those causing C:G>T:A transitions, resulting in a transition-to-transversion ratio of 1.22 (Supplementary Fig. 4) similar to other studies[12]. While more than half of the mutations were found in genes recently identified to be mutated in cancer genome sequencing projects from head/neck, melanoma, and ovarian carcinomas[13-16], none of the relapse specific mutations were observed in previous targeted sequencing projects focused on pediatric

ALL[17-18]. Genomic DNA sequencing was completed in an additional 62 B lymphoblastic leukemia diagnosis-relapse specimen pairs to look for additional mutations within the effected exon in nine of the 14 genes associated with cancer genomes (CAND1, CBX3, COBRA1, FBXO3, PRMT2, RGS12, SMEK2, TULP4, and USP7) as well as for one novel gene, *SDF2*. However, no additional tumor specific mutations were found in these regions (shared diagnosis-relapse mutations or relapse specific mutations). Our failure to detect shared relapse specific mutations in these genes indicates that some of our observed variants may be peripheral to drug resistance (so called passengers) and/or that escape mechanisms may be unique for individual patients, a finding similar to what is observed for metastasis in breast cancer[19]. The remaining genes (DOPEY1, DPH5, LPHN1, MIER3, NEGR1, NIPSNAP1, SCARF1, ZNF192) were not additionally sequenced.

Two different mutations were observed and validated in *NT5C2*, which encodes the protein cN-II, a 5′-nucleotidase enzyme active in the cell cytoplasm[20], in two of the relapse specimens profiled by RNA-sequencing. *NT5C2*, a member of a family of seven enzymes that regulate nucleotide levels, has been shown to be responsible for the hydrolysis of nucleotides such as 5’-inosine monophosphate and 5’-guanosine monophosphate, converting them to inosine and guanosine nucleosides respectively[21] but the enzyme can also display phosphotransferase activity as well[20,22]. Both mutations were confirmed at the DNA level and were specific to the relapse specimens (Supplementary Fig. 5). To determine the frequency of mutations in *NT5C2* in ALL patients, full exon resequencing was completed in an additional 61 relapse specimens. Among the 61 patients, five additional *NT5C2* somatic mutations were discovered and also validated as relapse specific (Supplementary Fig. 5). Thus, seven out of 71 patients harbored *NT5C2* relapse specific mutations, for an overall occurrence rate of 10% (Fig. 1a-b).

Coverage at diagnosis at the *NT5C2* mutated sites discovered by RNA-sequencing was at 96X and 112X respectively. Based on this depth of sequencing, a subclone at diagnosis would have to be present in less than 1% of the bulk leukemia to be missed by this sequencing technique. To assess whether mutations in *NT5C2* were present at diagnosis as a rare subclone, backtracking using ultra-deep sequencing was performed. Amplicon resequencing of diagnosis and relapse specimen DNA identified two cases where indeed a rare clone existed at diagnosis in 0.01% and 0.02% of the total reads (25,000x and 32,000x coverage respectively) (Table 2). In the remaining five cases no mutation could be detected at diagnosis. These data suggest that the emergence of clones containing mutations in *NT5C2* is driven by powerful selective pressures presumably due to drug resistance.

Mutations in *NT5C2* were mapped onto the previously published crystal structure[23]. All five mutations found in this study mapped to a single functional unit clustered in a region thought to be involved in subunit association/dissociation through the acidic C-terminal tail of the enzyme (Fig. 1 and Supplementary Fig. 6)[24]. In addition, the focal nature of the observed mutations suggested the acquisition of novel biological properties rather than disruption of enzymatic activity. Therefore, to test the functional impact of the mutations on enzyme activity, we expressed wild-type (WT), p.Arg238Trp, p.Arg367Gln and p.Ser445Phe mutant *NT5C2* cDNA in BL21 *E. coli* cells. Protein expression was induced by isopropyl β-D-thiogalactoside and extracts were analyzed for expression by western blot

(Fig. 1c). Equal volumes of fresh protein extracts were then assayed for 5'-nucleotidase activity by monitoring the hydrolysis of inosine monophosphate and compared against a standard curve. Significantly higher activity was observed for all mutants: p.Arg238Trp, p.Arg367Gln and p.Ser445Phe, compared to WT activity ($p \leq 0.01$) (Fig. 1d). No activity above background was observed with matched uninduced samples.

We hypothesized that mutant forms of *NT5C2* allow for resistance to chemotherapy treatment, in particular nucleoside analogues, given its enzymatic function. In addition the early emergence of *NT5C2* mutants correlates with the introduction of the maintenance phase of ALL therapy where nucleoside analogues assume a predominant role in treatment. Therefore we investigated whether or not mutant forms of *NT5C2* could provide protection from apoptosis after treatment with various chemotherapeutic agents used in the treatment of childhood ALL. The B lymphoblastic leukemia cell line, Reh, was transduced with lentiviruses encoding *NT5C2* WT or mutants (p.Arg238Trp, p.Arg367Gln, or p.Ser445Phe) and assayed for apoptosis after incubation with various chemotherapy agents used to treat childhood ALL for 24-72hrs. Compared to cells expressing WT protein, cells expressing mutant forms of *NT5C2* were significantly more resistant to apoptosis after treatment with purine analogues 6-MP and 6-TG (Fig. 2a-b). As expected, no resistance was seen when the experiment was repeated with cytarabine, doxorubicin, gemcitabine, or prednisolone (Fig. 2c-f). To further understand the mechanistic basis of the *NT5C2*-mediated chemoresistance, we also examined the effects of *NT5C2* mutations on the intracellular accumulation of thiopurine nucleotides, active metabolites of 6-MP. Following 6-MP treatment, Reh cells transduced with mutant forms of *NT5C2* showed drastic reduction of thio-guanoside nucleotides compared to WT or GFP infected control cells (Supplementary Fig. 7), consistent with the thiopurine resistance resulting from *NT5C2* mutations noted at relapse.

The characteristics of patients with and without *NT5C2* mutations were analyzed (Table 3) and interestingly, all patients who acquired mutations relapsed early, or within 36 months of initial diagnosis (p=0.03). Median time to relapse for those with a *NT5C2* mutation was 516 days compared to 930 for those without a *NT5C2* mutation (Supplementary Fig. 8). Among all patients that relapsed early, 19% of cases harbored mutations. This finding is consistent with previous data indicating potential differences in biological pathways that mediate early vs. late relapse[9].

These findings provide a detailed look at relapse mechanisms and how sequence alterations can directly result in the chemoresistant phenotypes observed in relapse patients. In particular, we discovered multiple relapse specific mutations in *NT5C2*, a gene not previously associated with somatic mutations in cancer. Our data demonstrates a direct relationship between acquired somatic mutations and chemoresistance to a specific class of drugs used in treatment, purine analogues, as opposed to defects in pathways shared across classes of cytotoxic agents. A previous study failed to correlate cytosolic 5′-nucleotidase activity with *in vitro* resistance to 6-TG in blasts from children at diagnosis with ALL, although a weak correlation was seen with the total amount of enzyme[25]. However these studies focused on patients at diagnosis and presumably these cases all contained WT *NT5C2*. In addition, previous studies have correlated high *NT5C2* levels with resistance to cytarabine in patients with AML[26-27], while other studies show the purified enzyme does not

hydrolyze araC monophosphate[28]. Our results in ALL are in agreement with the later finding. We hypothesize that the emergence of clones containing *NT5C2* mutations early in maintenance after completing phases of rotational multiagent chemotherapy correlates with a greater reliance on these agents. We and others have identified additional genes whose expression may play a role in resistance to purine analogs[29-30]. However the discovery of acquired mutations in *NT5C2* among patients with early relapse, a group with a uniformly poor outcome, provides a focal point to develop insight into major biological pathways that mediate drug resistance in vivo and potentially to develop new therapies targeting *NT5C2* to prevent emergence of resistant clones during maintenance therapy and/or to treat relapsed ALL. Inhibitors of 5′-nucleotidase have already been developed given their potential in cancer therapy and the prevention of drug resistance to anti-retroviral treatment[31-32]. Taken together our data demonstrates that discovery based approaches can identify recurrent mutations in cancer patients which relapse after cytotoxic chemotherapy.

Online Methods

Patient Specimens

Cryopreserved matched pairs of pediatric B lymphoblastic leukemia marrow specimens from diagnosis and relapse and when available, remission, were obtained from the Children's Oncology Group (COG) ALL cell bank from ten patients (Supplementary Table 1) from trials: AALL0232, AALL0331, and COG 9906 (ClinicalTrials.gov: NCT00075725, NCT00103285, NCT00005603 respectively). All specimens were Ficoll-enriched prior to cryopreservation and contained >80% blasts as measured by flow cytometry prior to enrichment. All patients (or parents) provided written consent for the banking and future research use of these specimens in concordance with the regulations of the institutional review boards of all participating institutions.

Patient timing to relapse calculated from initial diagnosis date. Samples were chosen based on bone marrow blast percentage at the time of banking submission, as well as by Affymetrix SNP6.0 chip. All samples with less than 20% disparity between the two methods and with >80% blasts in both diagnosis and relapse samples were considered for sequencing. RNA was extracted from diagnosis and relapse bone marrow samples using RNEasy Mini Kits (Qiagen) and quality verified by an Agilent Bioanalyzer 2100 (Agilent Technologies).

RNA Sequencing and Analysis

Libraries were prepared according to standard protocol for sequencing using the Illumina Genome Analyzer IIx (Illumina). Final cDNA libraries were evaluated for fragment size distribution by 2100 Agilent Bioanalyzer (DNA 1000 chip) and quantified by Quanti-iT Picogreen dsDNA Assay kit (Invitrogen). All libraries were sequenced using 54 base pair reads on the Illumina Genome Analyzer GAIIx. Image collection and analysis was completed using the Illumina CASAVA pipeline. Reads in raw FASTQ files were aligned to the human reference genome (hg18) using the Burroughs-Wheeler Aligner (v0.5.8a)[33] allowing up to two mismatches. Data is deposited at the NCBI Sequence Read Archive (SRA048657). Mapped reads in the raw BAM files were then recalibrated and locally realigned to call single nucleotide variants (SNVs) and insertion/deletions (indels) using the

Genome Analysis Toolkit (GATK).[10] After removing duplicate reads, only those reads with mapping qualities Q≥30 were used to predict SNVs and indels, again using GATK.[11]

Data was subjected to a set of post processing filters: i) a minimum of ≥8x coverage per variant site; ii) reads supporting the variant in > 20% of the total reads per site; iii) bidirectional sequence support of variant reads; iv) no more than 1 variant within 5bp distance v) minimum of 8x wild type (WT) coverage at the corresponding site in the paired diagnosis sample. Variants were filtered for known SNPs from the most current dbSNP database, dbSNP 135, and 1000 Genomes Project.[34] Finally, only those variants present in genes with the most conservative annotation by RefSeq were considered (removal of all XM_ annotations). All predicted variants were then manually inspected on the paired BAM files using the Integrative Genomics Viewer (IGV).[35] Predicted SNVs were compared to COSMIC v55 database[14] and processed using PolyPhen-2 prediction program and SIFT[36-37]. SNVs Schematic of filtering process for the SNV detection is outlined in Supplementary Figure 2. Schematic for indel detection is outlined in Supplementary Figure 3.

Fusion Detection

Paired end data (n=8) was processed using an in-house pipeline BEGAT (unpublished, CEM). Results were filtered for candidates using following criteria: i) coverage less than 8x reads; ii) region less than 10Kb away; iii) homologous gene filter to filter out mapping errors between gene isoforms and paralogs and iv) removal of those fusions mapped to repetitive regions (Supplementary Note).

Targeted Validation

Variant validation was completed in eight out of ten discovery specimens, for which matched remission, diagnosis, and relapse genomic DNA were available (two samples not validated based on missing remission gDNA). Primers were designed within 400 base pairs of the variant site and amplified by PCR. PCR products were sequenced using Sanger sequencing and trace files were manually inspected for variation from the reference genome using the Mutation Surveyor program (Softgenetics). All validated mutations were reconfirmed with a second PCR and Sanger reaction. Full exon sequencing of *NT5C2* (Ensembl transcript ID ENST00000404739) was completed by Sanger sequencing using exon specific primers (Genewiz, Inc) (Supplementary Information).

454 Amplicon Sequencing

Targeted amplicon sequencing was performed using the Roche 454 Genome Sequencer FLX + deep sequencing platform. PCR amplicons spanning the mutated sites were tagged using Roche 454 adaptor-MIDs primer sets and added to PCR primers designed for the bi-directional sequencing. Amplicons were then purified by AMpure XP beads (Beckman) to remove excess primer and quantified by fluorometry again using the Quant-iT PicoGreen dsDNA Assay Kit. A titration test was performed on the amplicon libraries using a low-volume emulsion PCR amplicon kit according to the Roche 454 protocol and followed by the emulsion-based clonal amplification (emPCR amplification, Lib-A). The libraries were sequenced on the Roche 454 Genome Sequencer FLX + sequencing system (454 Life

Sciences) at ultra-deep coverage (17,000-50,000x) using a two region 70 × 75-mm Titanium PicoTiterPlate and the mutation analysis was performed using the Roche 454 Amplicon Variant Analyzer package.

Mutation Modeling

Molecular graphics of *NT5C2* were rendered with ICM-Pro (Molsoft, LLC). Molecular surface rendering and exact-boundary electrostatic mapping onto that surface were calculated as previously described[38-39].

NT5C2 Protein Expression and 5'-Nucleotidase Assay

Full length *NT5C2* wild-type and mutant (p.Arg238Trp, p.Arg367Gln and p.Ser445Phe) cDNA constructs (purchased from Genewiz, Inc) were cloned into the pET30a expression vector (gift from J.D. Ernst) using NdeI and HindIII restriction sites. pET30a expression vectors were transformed into BL21 DE3 pLysS chemically competent *Escherichia coli* (Invitrogen). *NT5C2* expression was induced using 1mM isopropyl-β-D-thiogalactoside (IPTG) for 5hrs incubated at 37° C. Cells were pelleted at 8,000 × g for 2min at 4°C and resuspended in lysis buffer (50mM NaH2PO4, 300mM NaCl, 10mM imidazole) with 1x protease inhibitors (GE Healthcare), 1mg/ml lysozyme was added and incubated on ice for 30min. Lysate was then centrifuged at 15,000 × g for 10min at 4°C. Protein was subject to electrophoresis on 9% SDS-Tris acrylamide gel and transferred to PDVF membranes. Membranes were incubated with 1:5000 rabbit polyclonal anti-*NT5C2* antibody (Abcam, Inc.) followed by secondary anti-rabbit horseradish peroxidase-conjugated antibody, 1:10,000 (GE Healthcare) and developed using ECL (GE Healthcare). Ten microliters of purified protein extract was used to assess enzymatic activity of WT and mutant proteins using the 5'-Nucleotidase Enzymatic Test Kit (Diazyme, Inc.) according to protocol. Mean ± s.d. is presented of three independent experiments.

Cell Culture and Drug Treatment

Reh cells obtained from ATCC were grown in RPMI1640 medium supplemented with 10% FBS, 10mM HEPES, 1%Pen/Strep under 5% CO_2 at 37°C. 293T cells (ATCC) were grown in DMEM medium supplemented with 10% FBS and 1% Pen/Strep under 5% CO_2 at 37°C. 6-mercaptopurine (6MP), 6-thioguanine (6TG), cytarabine, doxorubicin, gemcitabine and prednisolone (Sigma) were serially diluted in RPMI before use at indicated concentrations.

Transient Transfection and Lentivirus Gene Transfer

WT and mutant (p.Arg238Trp, p.Arg367Gln, and p.Ser445Phe) *NT5C2* cDNA were cloned into the lentiviral vector pLenti (gift from M.R. Phillips) using restriction sites SalI and XbaI. All plasmids were sequence verified. cDNA constructs were transfected along with helper plasmids into 293T cells using the Calcium Phosphate method to produce replication-defective virus. Supernatant was harvested 48hrs later and used to transduce Reh cells (*NT5C2* sequence verified as WT) supplemented with 8ug/ml Polybrene (Sigma). Viral media was replaced after 24hrs post infection. Cells were monitored 72hrs post infection for infection rate by GFP positive cells using the FACScan (BD). 200,000 infected cells were plated per well in 200ul of media in triplicate for drug treatment with 6-mercaptopurine (6-

MP), 6-thioguanine (6-TG), cytarabine, doxorubicin, gemcitabine and prednisolone (Sigma). Cells were incubated for 24-72hrs and then assayed for apoptosis by Annexin V-PE and 7AAD staining (Annexin V-PE Apoptosis Detection Kit, BD Pharmingen) followed by flow cytometric analysis using the FACScan (BD). The percent Annexin V/7AAD positive and negative cells were analyzed by FlowJo software (version 7.6.1, Tree Star Inc). Data was plotted relative to no chemotherapy treatment and error bars represent the s.d. of three independent determinations. 1 ×10^6 cells were harvested for protein at time of plating. Briefly, cells were pelleted at 1200rpm for 5mins and resuspended in 100ul RIPA buffer with 1x protease inhibitors (GE Healthcare), incubated on ice 15min, and centrifuged at 14,000rpm for 10min at 4°C. Protein was subject to electrophoresis on 9% SDS-Tris acrylamide gel and transferred to PDVF membranes. Membranes were incubated with 1:5000 anti-Flag antibody (Sigma) followed by secondary anti-mouse horseradish peroxidase-conjugated antibody, 1:10,000 (GE Healthcare) and developed using ECL (GE Healthcare).

HPLC Determination of Nucleotides

Reh cells were transiently infected as described above. Post infection cells were treated with 10uM 6-MP for 24hrs in duplicate. After 24hrs, 5 × 10^6 cells were washed twice with PBS and cells pellets frozen at -80C. Intraceullar accumulation of thioguanine nucleotides (6-MP active metabolites) was determined by a reversed-phase liquid chromatographic assay, as described previously[40].

Statistical Analysis

Statistical analysis of enzymatic assays and chemoresistance assays were performed using the two-sided unpaired Student's *t*-test. Statistical analysis of the clinical and biological characteristics of patients with *NT5C2* mutations was performed using Fisher's Exact test. $P < 0.05$ was considered statistically significant.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We would like to thank the members of the Carroll laboratory as well as L.B. Gardner, M. Karajannis, and I. Osman for their critical review of the manuscript. This work was supported by NIH R01 CA140729, R21 CA152838-02 grants to W.L. Carroll and New York University Cancer Center Support Grant 5 P30 CA16087-30 in collaboration with the NYU Genome Technology Center; additional support by grants from the NCI to the Children's Oncology Group including U10 CA98543 (COG Chair's grant), U10 CA98413 (COG Statistical Center), and U24 CA114766 (COG Specimen Banking). JAM is supported by NIH T32 CA009161; LEH was supported by the American Society of Hematology and St. Baldrick's Foundation; SPH is the Ergen Family Chair in Pediatric Cancer. We gratefully acknowledge the Children's Oncology Group for patient specimens; NYU Genome Technology Center for expert assistance with Illumina (Berrin Baysa) and Roche 454 (Elisa Venturini) deep sequencing experiments (supported in part by NIH/NCRR S10 RR026950-01 grant); P. Grace and J.D. Ernst for expression vectors; and F. Tsai, M.R. Phillips, and S.M. Brown for technical guidance.

References

1. Raetz EA, et al. Reinduction platform for children with first marrow relapse in acute lymphoblastic lymphoma. J Clin Oncol. 2008; 26:3971–3978. [PubMed: 18711187]

2. Klumper E, et al. In vitro cellular drug resistance in children with relapsed/refractory acute lymphoblastic leukemia. Blood. 1995; 86:3861–3868. [PubMed: 7579354]
3. Li J, Thompson TD, Miller JW, Pollack LA, Stewart SL. Cancer incidence among children and adolescents in the United States, 2001-2003. Pediatrics. 2008; 121:e1470–1477. [PubMed: 18519450]
4. Hunger SP, et al. Improved survival for children and adolescents with acute lymphoblastic leukemia between 1990 and 2005: a report from the children's oncology group. J Clin Oncol. 2012; 30:1663–1669. [PubMed: 22412151]
5. Pui CH, Evans WE. Treatment of acute lymphoblastic leukemia. N Engl J Med. 2006; 354:166–178. [PubMed: 16407512]
6. Chessells JM, et al. Long-term follow-up of relapsed childhood acute lymphoblastic leukaemia. British journal of haematology. 2003; 123:396–405. [PubMed: 14616997]
7. Eapen M, et al. Outcomes after HLA-matched sibling transplantation or chemotherapy in children with B-precursor acute lymphoblastic leukemia in a second remission: a collaborative study of the Children's Oncology Group and the Center for International Blood and Marrow Transplant Research. Blood. 2006; 107:4961–4967. [PubMed: 16493003]
8. Gaynon PS, et al. Bone marrow transplantation versus prolonged intensive chemotherapy for children with acute lymphoblastic leukemia and an initial bone marrow relapse within 12 months of the completion of primary therapy: Children's Oncology Group study CCG-1941. J Clin Oncol. 2006; 24:3150–3156. [PubMed: 16717292]
9. Hogan LE, et al. Integrated genomic analysis of relapsed childhood acute lymphoblastic leukemia reveals therapeutic strategies. Blood. 2011
10. McKenna A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010; 20:1297–1303. [PubMed: 20644199]
11. DePristo MA, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011; 43:491–498. [PubMed: 21478889]
12. Ding L, et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. Nature. 2010; 464:999–1005. [PubMed: 20393555]
13. Stransky N, et al. The mutational landscape of head and neck squamous cell carcinoma. Science. 2011; 333:1157–1160. [PubMed: 21798893]
14. Forbes SA, et al. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. Nucleic Acids Res. 2011; 39:D945–950. [PubMed: 20952405]
15. Wei X, et al. Exome sequencing identifies GRIN2A as frequently mutated in melanoma. Nat Genet. 2011; 43:442–446. [PubMed: 21499247]
16. Network TCGAR. Integrated genomic analyses of ovarian carcinoma. Nature. 2011; 474:609–615. [PubMed: 21720365]
17. Greenman C, et al. Patterns of somatic mutation in human cancer genomes. Nature. 2007; 446:153–158. [PubMed: 17344846]
18. Mullighan CG, et al. CREBBP mutations in relapsed acute lymphoblastic leukaemia. Nature. 2011; 471:235–239. [PubMed: 21390130]
19. Shah SP, et al. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. Nature. 2009; 461:809–813. [PubMed: 19812674]
20. Bianchi V, Spychala J. Mammalian 5′-nucleotidases. J Biol Chem. 2003; 278:46195–46198. [PubMed: 12947102]
21. Tozzi MG, et al. Nucleoside phosphotransferase activity of human colon carcinoma cytosolic 5′-nucleotidase. Arch Biochem Biophys. 1991; 291:212–217. [PubMed: 1659319]
22. Tozzi MG, et al. Cytosolic 5′-nucleotidase/phosphotransferase of human colon carcinoma. Advances in experimental medicine and biology. 1991; 309B:173–176. [PubMed: 1664180]
23. Wallden K, et al. Crystal structure of human cytosolic 5′-nucleotidase II: insights into allosteric regulation and substrate recognition. The Journal of biological chemistry. 2007; 282:17828–17836. [PubMed: 17405878]
24. Spychala J, Chen V, Oka J, Mitchell BS. ATP and phosphate reciprocally affect subunit association of human recombinant High Km 5′-nucleotidase. Role for the C-terminal polyglutamic

acid tract in subunit association and catalytic activity. European journal of biochemistry / FEBS. 1999; 259:851–858. [PubMed: 10092873]

25. Pieters R, et al. Relation of 5′-nucleotidase and phosphatase activities with immunophenotype, drug resistance and clinical prognosis in childhood leukemia. Leukemia research. 1992; 16:873–880. [PubMed: 1405718]
26. Galmarini CM, et al. Expression of high Km 5′-nucleotidase in leukemic blasts is an independent prognostic factor in adults with acute myeloid leukemia. Blood. 2001; 98:1922–1926. [PubMed: 11535530]
27. Galmarini CM, et al. Deoxycytidine kinase and cN-II nucleotidase expression in blast cells predict survival in acute myeloid leukaemia patients treated with cytarabine. British journal of haematology. 2003; 122:53–60. [PubMed: 12823345]
28. Mazzon C, et al. Cytosolic and mitochondrial deoxyribonucleotidases: activity with substrate analogs, inhibitors and implications for therapy. Biochem Pharmacol. 2003; 66:471–479. [PubMed: 12907246]
29. Yang JJ, et al. Genome-wide copy number profiling reveals molecular evolution from diagnosis to relapse in childhood acute lymphoblastic leukemia. Blood. 2008
30. Diouf B, et al. Somatic deletions of genes regulating MSH2 protein stability cause DNA mismatch repair deficiency and drug resistance in human leukemia cells. Nature medicine. 2011; 17:1298–1303.
31. Gallier F, et al. Structural insights into the inhibition of cytosolic 5′-nucleotidase II (cN-II) by ribonucleoside 5′-monophosphate analogues. PLoS Comput Biol. 2011; 7:e1002295. [PubMed: 22174667]
32. Jordheim LP, et al. Identification and characterization of inhibitors of cytoplasmic 5'- nucleotidase cN-II issued from virtual screening. Biochem Pharmacol. 2013; 85:497–506. [PubMed: 23220537]
33. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25:1754–1760. [PubMed: 19451168]
34. A map of human genome variation from population-scale sequencing. Nature. 2010; 467:1061–1073. [PubMed: 20981092]
35. Robinson JT, et al. Integrative genomics viewer. Nat Biotechnol. 2011; 29:24–26. [PubMed: 21221095]
36. Adzhubei IA, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010; 7:248–249. [PubMed: 20354512]
37. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009; 4:1073–1081. [PubMed: 19561590]
38. Totrov M, Abagyan R. The contour-buildup algorithm to calculate the analytical molecular surface. Journal of structural biology. 1996; 116:138–143. [PubMed: 8742735]
39. Totrov M, Abagyan R. Rapid boundary element solvation electrostatics calculations in folding simulations: successful folding of a 23-residue peptide. Biopolymers. 2001; 60:124–133. [PubMed: 11455546]
40. Dervieux T, et al. HPLC determination of thiopurine nucleosides and nucleotides in vivo in lymphoblasts following mercaptopurine therapy. Clin Chem. 2002; 48:61–68. [PubMed: 11751539]

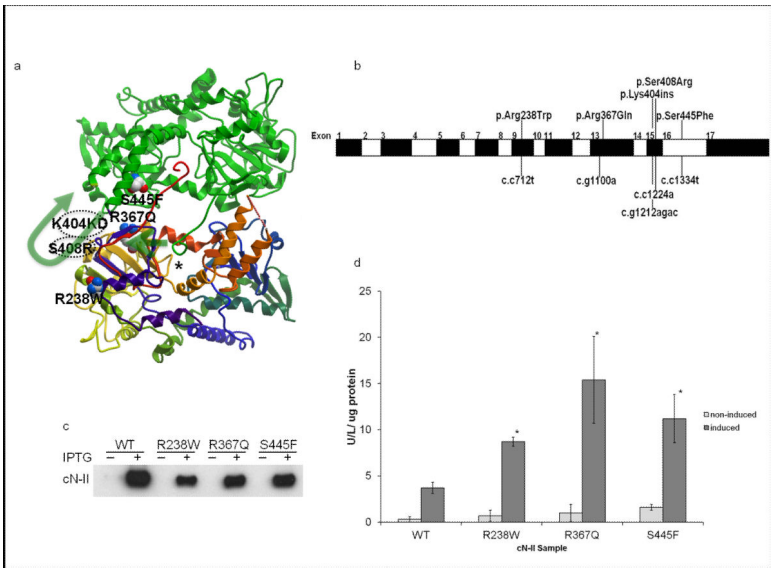

Figure 1. Relapse Specific Mutations in *NT5C2* Alter Enzymatic Activity

a, Dimer of human cytosolic 5′-nucleotidase II subunits. There are two such dimers, linked by a different interface to form the tetrameric active form of this enzyme. The backbone traces of the structures are displayed as ribbon diagrams. The bottom monomer ribbon is colored in a gradient from its N-terminus (purple) to its C-terminus (red). The location of the active site is indicated with an asterisk (*). Note that the C-terminus of one monomer extends into a groove in the other monomer in order to form the dimer. The upper monomer ribbon is colored green for contrast. The location of the disordered loop at positions 400-417 is indicated as an orange dashed line in the bottom monomer and as a transparent green U-shaped arrow in the top monomer to show its expected area of interaction. Mutations p.Arg238Trp, p.Arg367Gln and p.Ser445Phe are displayed as space-filling spheres colored red for oxygen, blue for nitrogen and white for carbon and are also labeled. The projected locations of the insertion mutation (p.Lys404LysAsp) and point mutation (p.Ser408Arg) in the disordered loop, which is not visible in the crystal structure, are indicated by dashed circles and labeled. A straight transparent green arrow indicates the expected trajectory of the acidic C-terminal tail of the upper monomer, which is not present in the crystal structure, as it lies across the bottom monomer. **b**, *NT5C2* coding region diagram with mutations. Three mutations were found at the same site in exon 9 encoding amino acid 238. **c**, Western blot analysis of cN-II protein induction by IPTG in BL21 cells. 10ug of each protein lysate run per lane and blotted with antibody against *NT5C2*. **d**, Equivalent volumes of BL21 protein lysate were subjected to the 5′-nucleotidase assay (Diazyme, Inc) according to protocol. Mean activity levels were normalized per protein concentration for each sample. Columns show the mean of three independent experiments ± s.d. *P* values calculated using two-sided unpaired Student's *t*-test (*$p \leq 0.01$).

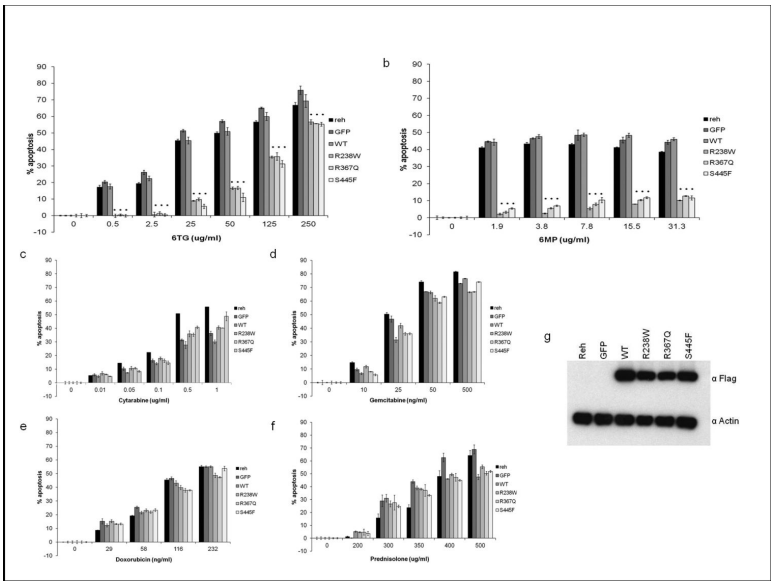

Figure 2. ***NT5C2*** **Mutants Confer Chemoresistance to Purine Nucleoside Analogue Treatment**

Reh cells transiently infected with WT, mutant or control GFP lentivirus were treated with increasing concentrations of **a**, 6-TG, **b**, 6-MP, **c**, cytarabine, **d**, gemcitabine, **e**, doxorubicin, **f**, prednisolone and assayed for apoptosis. Columns show a mean of three independent determinations ± s.d. from a representative experiment repeated 3 times with similar results. *P* values calculated using two-sided unpaired Student's *t*-test (**p* < 0.001). **g**, Western blot of infected Reh cells demonstrated presence of flag tagged *NT5C2* constructs compared to GFP control and Reh cells alone. Actin shown as loading control.

Table 1

Validated Relapse Specific Somatic Mutations

Patient	Gene	Chromosome	Position	Function	Nucleotide change	Protein Change	PolyPhen-2 Prediction	SIFT Prediction	In COSMIC database?	Gene Name
1	*RGS12*	4	3287853	missense	c.c158t	p.Ala53Val	Damaging	Damaging	Yes	regulator of G-protein signaling 12
1	*LPHN1*	19	14134808	missense	c.c822g	p.Glu274Gln	Damaging	Damaging	Yes	latrophilin 1
2	*CAND1*	12	65985593	missense	c.a1878c	p.Leu626Phe	Damaging	Damaging	Yes	cullin-associated and neddylation-dissociated 1
2	*PRMT2*	21	46903160	missense	c.a730c	p.Met244Leu	Benign	Tolerated	Yes	protein arginine methyltransferase 2
2	*NIPSNAP1*	22	28287562	missense	c.g512t	p.Ser171Ile	Damaging	Damaging	Yes	nipsnap homolog 1
3	*USP7*	16	8902368	missense	c.a2188t	p.Thr730Ser	Damaging	Tolerated	Yes	ubiquitin specific peptidase 7
4	*TULP4*	6	158844705	missense	c.t4022g	p.Leu1341Arg	Damaging	Tolerated	Yes	tubby like protein 4
4	*CBX3*	7	26214576	missense	c.g206a	p.Cys69Tyr	Damaging	Damaging	Yes	chromobox homolog 3
4	*COBRA1*	9	139270653	missense	c.g318a	p.Met106Ile	Benign	Tolerated	Yes	cofactor of BRCA1
4	*SDF2*	17	24006562	missense	c.g218a	p.Arg73Gln	Damaging	Tolerated	No^	stromal cell-derived factor 2
5	*FBXO3*	11	33725250	missense	c.t1241a	p.Val414Glu	Damaging	Tolerated	Yes	F-box protein 3
5	*SCARF1*	17	1490488	nonsense	c.a1014t	p.stop338Cys	Isoform Change	Tolerated	Yes	scavenger receptor class F, member 1
6	*NEGR1*	1	71849375	missense	c.c710t	p.Pro237Leu	Benign	Tolerated	Yes	neuronal growth regulator 1
7	*NT5C2*	10	104847097	missense	c.c712t	p.Arg238Trp	Damaging	Damaging	No^	5′-nucleotidase, cytosolic II
8	*DPH5*	1	101233272	missense	c.c512t	p.Ser171Phe	Damaging	Damaging	No^	DPH5 homolog
8	*SMEK2*	2	55648886	missense	c.g1628a	p.Arg543Gln	Damaging	Damaging	Yes	SMEK homolog 2, suppressor of mek1
8	*MIER3*	5	56262281	missense	c.g796a	p.Glu266Lys	Benign	Tolerated	No^	mesoderm induction early response 1, family member 3
8	*DOPEY1*	6	83912011	missense	c.g5591a	p.Arg1864His	Damaging	Tolerated	Yes	dopey family member 1
8	*ZNF192*	6	28229455	missense	c.g1418c	p.Arg473Pro	Damaging	Tolerated	No^	zinc finger protein 192
8	*NT5C2*	10	104840473	missense	c.c1334t	p.Ser445Phe	Damaging	Tolerated	No^	5′-nucleotidase, cytosolic II

Mutations validated using remission, diagnosis and relapse genomic DNA. Chromosome position in reference to hg18 alignment. Nucleotide change in reference to start of coding sequence. Prediction of structural/functional consequence of mutation completed using Poly Phen-2 and SIFT.

^Present in COSMIC database post original submission date of article. Gene name from Hugo Gene Nomenclature Committee.

Table 2

Deep Amplicon Sequencing of *NT5C2* Mutations

NT5C2 Exon	Nucleotide change	Protein change	Mutant Allele Frequency (coverage)	
			Diagnosis	Relapse
9	c.c712t	p.Arg238Trp	0.01% (25,000x)	27% (17,000x)
9	c.c712t	p.Arg238Trp	0 (22,000x)	18% (16,000x)
9	c.c712t	p.Arg238Trp	0 (49,000x)	31% (18,000x)
13	c.g1100a	p.Arg367Gln	0.02% (32,000x)	25% (28,000x)
15	c.g1212agac	p.Lys404LysAsp	0 (26,000x)	55% (29,000x)
15	c.c1224a	p.Ser408Arg	0 (31,000x)	50% (22,000x)
16	c.c1334t	p.Ser445Phe	0 (42,000x)	25% (45,000x)

Table 3

Characteristics of Patients According to *NT5C2* Mutation Status

Variable		NT5C2 Mutated n=7	Non-Mutated n=64	P-value
Age at Diagnosis:	Less than 10 years	4	39	0.57
	≥ to 10 years	3	25	
Race:	White	3	47	0.11^
	Black	1	6	
	Asian	1	3	
	Other	1	5	
	Unknown	1	3	
Sex:	Female	2	27	0.39
	Male	5	37	
Cytogenetics:	ETV6/RUNX1	1	13	0.12#
	Hyperdiploid	--	15	
	E2aPBX1	--	1	
	Normal	6	35	
Time to Relapse:	Early	7	37	0.03
	Late	--	27	
Risk Group*:	Standard	2	25	0.46
	High	5	39	

^Fisher's Exact test p-value of all other race groups vs. white.

#Fischer's Exact test p-value of normal vs. all other cytogenetic groups.

*NCI risk group (Smith et al. 1996).