Relapse-specific mutations in NT5C2 in childhood acute lymphoblastic leukemia

Julia A Meyer1,2, Jinhua Wang1,3, Laura E Hogan1,4,13, Jun J Yang5, Smita Dandekar1,4, Jay P Patel6, Zuojian Tang3, Paul Zumbo7,8, Sheng Li7,8, Jiri Zavadil1,3, Ross L Levine6,9, Timothy Cardozo10, Stephen P Hunger11,12, Elizabeth A Raetz1,4, William E Evans5, Debra J Morrison1,4, Christopher E Mason7,8 & William L Carroll1,2,4

Relapsed childhood acute lymphoblastic leukemia (ALL) carries a poor prognosis, despite intensive retreatment, owing to intrinsic drug resistance1,2. The biological pathways that mediate resistance are unknown. Here, we report the transcriptome profiles of matched diagnosis and relapse bone marrow specimens from ten individuals with pediatric B-lymphoblastic leukemia using RNA sequencing. Transcriptome sequencing identified 20 newly acquired, novel nonsynonymous mutations not present at initial diagnosis, with 2 individuals harboring relapse-specific mutations in the same gene, NT5C2, encoding a 5′-nucleotidase. Full-exon sequencing of NT5C2 was completed in 61 further relapse specimens, identifying additional mutations in 5 cases. Enzymatic analysis of mutant proteins showed that base substitutions conferred increased enzymatic activity and resistance to treatment with nucleoside analog therapies. Clinically, all individuals who harbored NT5C2 mutations relapsed early, 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.

ALL is the most common pediatric malignancy, accounting for more than 25% of all childhood cancers3. Cure rates for ALL have substantially improved over the past four decades with the development of risk-stratified multiagent chemotherapy, preventive treatment to the central nervous system and the more recent introduction of augmented doses and/or schedules for the administration of conventional drugs, resulting in an overall 5-year event-free survival rate now exceeding 90% (ref. 4). In spite of these improvements, 10–20% of patients experience disease recurrence5. The prognosis for these children is dismal6, even with aggressive salvage strategies involving allogeneic stem cell transplantation7,8.

Received 31 July 2012; accepted 28 January 2013; published online 3 February 2013; doi:10.1038/ng.2558
Table 1 Validated relapse-specific somatic mutations

| Subject | Gene   | Chromosome | Position | Function       | Nucleotide change | Protein change | PolyPhen-2 prediction | SIFT prediction | In COSMIC database? | Encoded protein |
|---------|--------|------------|----------|----------------|-------------------|----------------|-----------------------|-----------------|---------------------|-----------------|
| 1       | RGS12  | 4          | 3287853  | Missense       | c.158C>T          | p.Ala53Val     | Damaging              | Damaging        | Yes                 | Regulator of G protein signaling 12 |
| 1       | LPHN1  | 19         | 14134808 | Missense       | c.822G>A          | p.Glu274Gln    | Damaging              | Damaging        | Yes                 | Latrophilin 1 |
| 2       | CAND1  | 12         | 6598593  | Missense       | c.1878A>C         | p.Leu626Phe    | Damaging              | Damaging        | Yes                 | Cullin-associated and neddylation-dissociated 1 |
| 2       | PRMT2  | 21         | 46903160 | Missense       | c.730A>C          | p.Met244Leu    | Benign                | Tolerated       | Yes                 | Protein arginine methyltransferase 2 |
| 2       | NIPSNAP1 | 22      | 28287562 | Missense       | c.512G>T          | p.Ser171Ile    | Damaging              | Damaging        | Yes                 | Nipsnap homolog 1 |
| 3       | USP7   | 16         | 8902368  | Missense       | c.2188A>T         | p.Thr730Ser    | Damaging              | Tolerated       | Yes                 | Ubiquitin-specific peptidase 7 |
| 4       | TULP4  | 6          | 158844705| Missense       | c.4022G>A         | p.Leu1341Arg   | Damaging              | Tolerated       | Yes                 | Tubby-like protein 4 |
| 4       | CBX3   | 7          | 26214576 | Missense       | c.206G>A          | p.Cys69Tyr     | Damaging              | Damaging        | Yes                 | Chromobox homolog 3 |
| 4       | COBRA1 | 9          | 139270653| Missense       | c.318G>A          | p.Met106Ile    | Benign                | Tolerated       | Yes                 | Cofactor of BRCA1 |
| 4       | SDF2   | 17         | 24006562 | Missense       | c.218G>A          | p.Arg73Gln     | Damaging              | Tolerated       | Noa                | Stromal cell–derived factor 2 |
| 5       | FBXO3  | 11         | 33725250 | Missense       | c.1241T>A         | p.Val414Glu    | Damaging              | Tolerated       | Yes                 | F-box protein 3 |
| 5       | SCARF1 | 17         | 1490488 | Non sense      | c.1014A>T        | p.Cys338*      | Isoform               | Change          | Yes                 | Scavenger receptor class F, member 1 |
| 6       | NEGR1  | 1          | 71849375 | Missense       | c.710C>T          | p.Pro237Leu    | Benign                | Tolerated       | Yes                 | Neuronal growth regulator 1 |
| 7       | NT5C2  | 10         | 104847097| Missense       | c.712C>T          | p.Arg238Trp    | Damaging              | Tolerated       | Noa                | 5'-nucleotidase, cytosolic II |
| 8       | DPH5   | 1          | 101233272| Missense       | c.512C>T          | p.Ser171Phe    | Damaging              | Noa             | DPH5 homolog |
| 8       | SMEK2  | 2          | 55648866 | Missense       | c.1628G>A         | p.Arg543Gln    | Damaging              | Yes             | SMEK homolog 2, suppressor of mek1 |
| 8       | MIER3  | 5          | 56262281 | Missense       | c.796G>A          | p.Glu266Lys    | Benign                | Tolerated       | Noa                | Mesoderm induction early response 1, family member 3 |
| 8       | DOPEY1 | 6          | 83912011 | Missense       | c.5591G>A         | p.Arg1864His   | Damaging              | Tolerated       | Yes                 | Dopey family member 1 |
| 8       | ZNF192 | 6          | 28229455 | Missense       | c.1418G>C         | p.Arg473Pro    | Damaging              | Tolerated       | Noa                | Zinc-finger protein 192 |
| 8       | NT5C2  | 10         | 104840473| Missense       | c.1334C>T         | p.Ser445Phe    | Damaging              | Tolerated       | Noa                | 5'-nucleotidase, cytosolic II |

Mutations were validated using remission, diagnosis and relapse genomic DNA. Chromosome positions are in reference to hg18 alignment. Nucleotide changes are in reference to the start of the coding sequences. Prediction of the structural and functional consequences of the mutations was completed using PolyPhen-2 and SIFT.

*Present in the Catalogue of Somatic Mutations in Cancer (COSMIC) database after the original submission date of this manuscript.

*Missing letters

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 further mutations in the affected exon in 9 of the 14 genes associated with cancer genomes (CAND1, CBX3, COBRA1, FBXO3, PRMT2, RGS12, SMEK2, TULP4 and USP7), as well as in 1 newly associated gene, SDF2. However, no additional tumor-specific mutations were found in these exon regions (including shared diagnosis-relapse mutations or relapse-specific mutations). Our inability to detect recurrent relapse-specific mutations in these genes indicates that some of our observed variants might be peripheral to drug resistance (so called passengers) and/or that escape mechanisms might be unique for individual patients, a finding similar to what is observed for metastasis in breast cancer19. The remaining genes (DOPEY1, DPH5, LPHN1, MIER3, NEGR1, NIPSNAP1, SCARF1 and ZNF192) were not sequenced further.

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, cN-II, a member of a family of seven enzymes that regulate nucleotide amounts, has been shown to be responsible for the hydrolysis of nucleotides such as 5′-inosine monophosphate and 5′-guanosine monophosphate, converting them into inosine and guanosine nucleosides, respectively.21 However, this enzyme can also have phosphotransferase activity20,22. 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 individuals with ALL, full-exon resequencing was completed in an additional 61 relapse specimens. In the 61 affected individuals, 5 further NT5C2 somatic mutations were identified and also validated as relapse specific (Supplementary Fig. 5). Thus, 7 out of 71 affected individuals harbored relapse-specific mutations in NT5C2, for an overall occurrence rate of 10% (Fig. 1a,b).

Coverage at diagnosis at the two NT5C2 mutated sites identified by RNA sequencing was 96× and 112×. Taking into consideration this depth of sequencing, a subclone at diagnosis would have to be present in less than 1% of the bulk leukemia cells 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. Amplion resequencing of DNA from diagnosis and relapse specimens identified two cases where a rare clone indeed existed at diagnosis in 0.01% and 0.02% of the total reads (with 25,000× and 32,000× 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 affecting cN-II were mapped onto the previously published crystal structure23. 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 the NT5C2 cDNA for wild-type protein and the Arg238Trp, Arg367Gln and Ser445Phe mutants in BL21 *Escherichia coli* cells. Protein expression was induced by isopropyl-β-D-thiogalactoside (IPTG), and extracts were analyzed for expression by immunoblot (Fig. 1c). Equal volumes of fresh protein extracts were then assayed for 5′-nucleotidase activity by monitoring the hydrolysis of inosine monophosphate compared against a standard curve. Significantly higher enzymatic activity was observed for all mutants—Arg238Trp, Arg367Gln and Ser445Phe—compared to wild-type protein (*P* ≤ 0.01; Fig. 1d). No activity above background was observed with matched non-induced samples.

We hypothesized that mutations in NT5C2 allow for resistance to chemotherapy treatment, in particular, nucleoside analogs, given their effects on enzymatic function. In addition, the early emergence of NT5C2 mutations correlates with the introduction of the maintenance phase of ALL therapy in which nucleoside analogs assume a predominant role in treatment. Therefore, we investigated whether mutant forms

---

**Table 2 Deep amplicon sequencing of NT5C2 mutations**

| NT5C2 exon | Nucleotide change | Protein change | Mutant allele frequency (coverage) |
|------------|------------------|----------------|-----------------------------------|
| 9          | c.712C>T         | p.Arg238Trp    | 0.01% (25,000x) 27% (17,000x)     |
| 9          | c.712C>T         | p.Arg238Trp    | 0 (22,000x) 18% (16,000x)        |
| 9          | c.712C>T         | p.Arg238Trp    | 0 (49,000x) 31% (18,000x)        |
| 13         | c.1100G>A        | p.Arg367Gln    | 0.02% (32,000x) 25% (28,000x)     |
| 15         | c.1212insAGAC    | p.Lys404ins    | 0 (26,000x) 55% (29,000x)        |
| 15         | c.1224C>A        | p.Ser408Arg    | 0 (31,000x) 50% (22,000x)        |
| 16         | c.1334C>T        | p.Ser445Phe    | 0 (42,000x) 25% (45,000x)        |

---

**Table 3 Characteristics of subjects according to NT5C2 mutation status**

| Variable | Mutated NT5C2 (*n* = 7) | Non-mutated NT5C2 (*n* = 64) | *P* value |
|----------|-------------------------|------------------------------|-----------|
| Age at diagnosis | Less than 10 years | 4 | 39 | 0.57 |
| Ancestry | European | 3 | 47 | 0.11 |
| | African | 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 | 0 | 15 | |
| | E2aPBX1 | 0 | 1 | |
| Time to relapse | Early | 7 | 37 | 0.03 |
| | Late | 0 | 27 | |
| Risk group | Standard | 2 | 25 | 0.46 |
| | High | 5 | 39 | |

*Fisher’s exact test.*

---

**Figure 1** Relapse-specific mutations in NT5C2 alter enzymatic activity. (a) Dimer of human cytosolic 5′-nucleotidase II (cN-II) subunits. Two such dimers, linked by a different interface, form the tetrameric active form of this enzyme. The backbone traces of the structures are shown as ribbons. 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 by an asterisk. Note that the C terminus of one monomer extends into a groove in the other monomer 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. The p.Arg238Trp, p.Arg367Gln and p.Ser445Phe alterations are shown as space-filling spheres colored red for oxygen, blue for nitrogen and white for carbon. The projected locations of the insertion (p.Lys404ins) and point alteration (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 with relapse-specific mutations and the encoded protein alterations. Three mutations were found at the same site in exon 9 encoding amino acid 238. (c) Immunoblot analysis of cN-II protein induction by IPTG in BL21 cells. Protein lysates (10 μg per lane) were blotted with antibody against cN-II WT, wild type. (d) Equivalent volumes of BL21 protein lysate were subjected to a 5′-nucleotidase assay (Diazyme). Mean activity levels were normalized by protein concentration for each sample. Columns show the mean of three independent experiments ± s.d. *P* values were calculated using two-sided unpaired Student’s t tests (*P* ≤ 0.01).
acquired mutations relapsed early, within 36 months of initial diagnosis.

Figure 2 NT5C2 mutations confer chemoresistance to purine nucleoside analog treatment. (a–f) Reh cells infected with control GFP lentivirus or with virus expressing wild-type (WT) or mutant NT5C2 were treated with increasing concentrations of 6-thioguanine (a), 6-mercaptopurine (b), cytarabine (c), gemcitabine (d), doxorubicin (e) or prednisolone (f) and assayed for apoptosis. Columns show a mean of three independent determinations ± s.d. from a representative experiment repeated three times with similar results. P values were calculated using two-sided unpaired Student’s t tests (*P < 0.001). (g) Immunoblot of infected Reh cells showing the presence of Flag-tagged cN-II proteins compared to GFP control and Reh cells alone. Actin is shown as a loading control.

of cN-II could provide protection from the apoptosis induced by treatment with various chemotherapeutic agents used clinically for childhood ALL. The B-lymphoblastic leukemia cell line Reh was transduced with lentiviruses encoding wild-type or mutant (Arg238Trp, Arg367Gln or Ser445Phe) cN-II and assayed for apoptosis after incubation with various chemotherapeutic agents for 24–72 h (Fig. 2). Compared to cells expressing wild-type protein, cells expressing mutant forms of cN-II were significantly more resistant to apoptosis after treatment with the purine analogs 6-mercaptopurine and 6-thioguanine (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 cN-II–mediated chemoresistance, we also examined the effects of the NT5C2 mutations on the intracellular accumulation of thiopurine nucleotides, which are active metabolites of 6-mercaptopurine. After treatment with 6-mercaptopurine, Reh cells transduced with lentiviruses expressing mutant forms of cN-II showed reduction in the level of thioguanine nucleotides compared to control cells expressing wild-type protein or GFP (Supplementary Fig. 7), consistent with the thiopurine resistance resulting from the NT5C2 mutations noted at relapse.

The characteristics of affected individuals with and without NT5C2 mutations were analyzed (Table 3), and, notably, all individuals who acquired mutations relapsed early, within 36 months of initial diagnosis (P = 0.03). Median time to relapse for those with a NT5C2 mutation was 516 d compared to 930 d for those without a NT5C2 mutation (Supplementary Fig. 8). Among all individuals that relapsed early, 16% of cases harbored mutations. This finding is consistent with previous data indicating potential differences in the biological pathways that mediate early versus late relapse.

These findings provide a detailed look at relapse mechanisms and how sequence alterations can directly result in the chemoresistant phenotypes observed in patients who relapse. In particular, we discovered multiple relapse-specific mutations in NT5C2, a gene not previously associated with somatic mutations in cancer. Our data show a direct relationship between acquired somatic mutations and chemoresistance to a specific class of drugs used in treatment, purine analogs, as opposed to defects in pathways shared across classes of cytotoxic agents. A previous study did not correlate cytosolic 5’-nucleotidase activity with in vitro resistance to 6-thioguanine in blasts from children at diagnosis with ALL, although a weak correlation was seen with the total amount of enzyme. However these studies focused on cases at diagnosis, and, presumably, these cases all contained wild-type NT5C2. In addition, previous studies have correlated high NT5C2 mRNA levels with resistance to cytarabine in patients with acute myeloid leukemia, whereas other studies showed that the purified enzyme does not hydrolyze araC monophosphate. 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 might have a role in resistance to purine analogs.\(^{29,30}\) However, the discovery of acquired mutations in NT5C2 in individuals 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 the emergence of resistant clones during maintenance therapy and/or to treat relapsed ALL. Inhibitors of 5'-nucleotidase have already been developed, given their potential usefulness in cancer therapy and the prevention of drug resistance to anti-retroviral treatment.\(^{31,32}\) Taken together, our data demonstrate that discovery-based approaches can identify recurrent mutations in individuals with cancer who relapse after cytotoxic chemotherapy.

**METHODS**

Methods and any associated references are available in the online version of the paper.

**Accession codes.** Next-generation sequence data are available at the NCBI Sequence Read Archive (SRA) under accession SRA048657.

**ACKNOWLEDGMENTS**

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. We gratefully acknowledge the Children’s Oncology Group (COG) for patient specimens; the New York University Genome Technology Center for expert assistance with Illumina (B. Bayas) and Roche 454 (E. Venturini) deep-sequencing experiments (supported by grants from the National Cancer Institute to COG, including U10 CA98543 (COG Specimen Banking). J.A.M. is supported by NIH grant T32 CA009161. L.E.H. was supported by grants from the National Cancer Institute to COG, including U10 CA98543 (COG Specimen Banking). J.A.M. is supported by NIH grant T32 CA009161. L.E.H. was supported by the American Society of Hematology and St. Baldrick’s Foundation. S.P.H. is the Ergen Family Chair in Pediatric Cancer.

**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.E.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.E.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.

**COMPETING FINANCIAL INTERESTS**

The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/ng.2558. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Raetz, E.A. et al. Reinduction platform for children with first marrow relapse in acute lymphoblastic leukaemia. J. Clin. Oncol. 26, 3971–3978 (2008).
2. Klumper, E. et al. In vitro cellular drug resistance in children with relapsed/refractory acute lymphoblastic leukaemia. Blood 86, 3861–3868 (1995).
3. Li, J., Thompson, T.D., Miller, J.W., Pollack, L.A. & Stewart, S.L. Cancer incidence among children and adolescents in the United States, 2001–2003. Pediatrics 121, e1470–e1477 (2008).
4. Hunger, S.P. et al. Improved survival for children and adolescents with acute lymphoblastic leukaemia between 1990 and 2005: a report from the children’s oncology group. J. Clin. Oncol. 33, 1663–1669 (2012).
5. Pui, C.H. & Evans, W.E. Treatment of acute lymphoblastic leukaemia. N. Engl. J. Med. 354, 166–178 (2006).
6. Chessells, J.M. et al. Long-term follow-up of relapsed childhood acute lymphoblastic leukaemia. Br. J. Haematol. 123, 396–405 (2003).
7. Eapen, M. et al. Outcomes after HLA-matched sibling transplantation or chemotherapy in children with B-precursor acute lymphoblastic leukaemia in a second remission: a collaborative study of the Children’s Oncology Group and the Center for International Blood and Marrow Transplant Research. Blood 107, 4961–4967 (2006).
8. Gaynor, P.S. et al. Bone marrow transplantation versus prolonged intensive chemotherapy for children with acute lymphoblastic leukaemia 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. 24, 3150–3156 (2006).
9. Hogan, L.E. et al. Integrated genomic analysis of relapsed childhood acute lymphoblastic leukaemia reveals therapeutic strategies. Blood 118, 5218–5226 (2011).
10. McKeever, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303 (2010).
11. Dib, N. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. 43, 491–498 (2011).
12. Ding, L. et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. Nature 464, 999–1005 (2010).
13. Stransky, N. et al. The mutational landscape of head and neck squamous cell carcinoma. Science 333, 1157–1160 (2011).
14. Forbes, S.A. et al. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. Nucleic Acids Res. 39, D945–D950 (2011).
15. Wei, S. et al. Exome sequencing identifies GRIN2A as frequently mutated in melanoma. Nat. Genet. 43, 442–446 (2011).
16. Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. Nature 474, 609–615 (2011).
17. Greenman, C. et al. Patterns of somatic mutation in human cancer genomes. Nature 446, 153–158 (2007).
18. Mullighan, C.G. et al. CREBBP mutations in relapsed acute lymphoblastic leukaemia. Nature 471, 235–239 (2011).
19. Shah, S.P. et al. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. Nature 461, 809–813 (2009).
20. Bianchi, V. & Spychala, J. Mammalian 5'-nucleotidases. J. Biol. Chem. 278, 46195–46198 (2003).
21. Tozzi, M.G. et al. Nucleoside phosphotransferase activity of human colon carcinoma cytosolic 5'-nucleotidase. Arch. Biochem. Biophys. 291, 212–217 (1991).
22. Tozzi, M.G. et al. Cytosolic 5'-nucleotidase/phosphotransferase of human colon carcinoma. Adv. Exp. Med. Bioi. 309B, 173–176 (1991).
23. Waldén, K. et al. Crystal structure of human cytosolic 5'-nucleotidase II. Insights into allosteric regulation and substrate recognition. J. Biol. Chem. 282, 17828–17836 (2007).
24. Spychala, J., Chen, V., Oka, J. & Mitchell, B.S. ATP and phosphate reciprocally affect subunit association of human recombiant high Km 5'-nucleotidase. Role of the C-terminal polyglutamic acid tract in subunit association and catalytic activity. Eur. J. Biochem. 259, 851–858 (1999).
25. Pieters, R. et al. Relation of 5'-nucleotidase and phosphatase activities with immunophenotype, drug resistance and clinical prognosis in childhood leukaemia. Leuk. Res. 16, 873–880 (1992).
26. Galmari, C.M. et al. Expression of high Km 5'-nucleotidase in leukemic blasts is an independent prognostic factor in adults with acute myeloid leukemia. Blood 98, 1922–1926 (2001).
27. Galmari, C.M. et al. Deoxycytidine kinase and cN-II nucleotidase expression in blast cells predict survival in acute myeloid leukaemia patients treated with cytarabine. Br. J. Haematol. 122, 53–60 (2003).
28. Mazzon, C. et al. Cytosolic and mitochondrial deoxynucleotidases: activity with substrate analogs, inhibitors and implications for therapy. Biochem. Pharmacol. 66, 471–479 (2003).
29. Yang, J.J. et al. Genome-wide copy number profiling reveals molecular evolution from diagnosis to relapse in childhood acute lymphoblastic leukaemia. Blood 112, 4178–4183 (2008).
30. Diouf, B. et al. Somatic deletions of genes regulating MSH2 protein stability cause DNA mismatch repair deficiency and drug resistance in human leukaemia cells. Nat. Med. 17, 1298–1303 (2011).
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. 7, e1002295 (2011).
32. Jordheim, L.P. et al. Identification and characterization of inhibitors of cytoplasmic 5'-nucleotidase cN-II issued from virtual screening. Biochem. Pharmacol. 85, 547–560 (2013).
33. Smith, M. et al. Uniform approach to risk classification and treatment assignment for children with acute lymphoblastic leukaemia. J. Clin. Oncol. 14, 18–24 (1996).
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 affected individuals (Supplementary Table 1) from trials AALL0232, AALL0331 and COG 9906 (NCT00075725, NCT00103285 and NCT0005603 at ClinicalTrials.gov, respectively). All specimens were Ficoll enriched before cryopreservation and contained >80% blasts, as measured by flow cytometry before enrichment. All subjects (or their parents) provided written consent for banking and future research use of these specimens in accordance with the regulations of the institutional review boards of all participating institutions.

Time to relapse was calculated from the initial diagnosis date. Samples were chosen on the basis of bone marrow blast percentage at the time of bank submission, as well as by Affymetrix SNP 6.0 chip data. All were chosen on the basis of bone marrow blast percentage at the time of relapse. All specimens were Ficoll enriched before cryopreservation and contained >80% blasts as measured by flow cytometry before enrichment. All subjects (or their parents) provided written consent for banking and future research use of these specimens in accordance with the regulations of the institutional review boards of all participating institutions.

Targeted validation. Variant validation was completed in eight out of ten discovery specimens for which genomic DNA from matched remission, diagnosis and relapse was available (two samples were not validated owing to unavailable remission genomic DNA). Primers were designed within 400 bp of the variant site, and sequences were 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 confirmed with a second PCR and Sanger reaction. Full-exon sequencing of NT5C2 (Ensembl transcript ENST00000404739) was completed by Sanger sequencing using exon-specific primers (Genewiz) (Supplementary Table 4).

Roche 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-multiplex identifier (MID) tags primer sets and added to PCR primers designed for bidirectional sequencing. Amplicons were then purified with AMPure XP beads (Beckman Coulter) to remove excess primer and quantified by fluorometry 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, which was followed by emulsion-based clonal amplification (emPCR amplification; Lib-A). Libraries were sequenced on the Roche 454 Genome Sequencer FLX + sequencing system (454 Life Sciences) at ultra-deep coverage (17,000–50,000×) using a two-region 70-mm × 75-mm Titanium PicoTiterPlate, and mutation analysis was performed using the Roche 454 Amplicon Variant Analyzer package.

Mutation modeling. Molecular representations of cN-II were rendered with ICM-Pro (Molecular). Molecular surface rendering and exact boundary electrostatic mapping onto that surface were carried out as previously described19,40.

cN-II protein expression and 5′-nucleotidase assay. Full-length NT5C2 cDNA for wild-type and mutant (Arg238Trp, Arg367Gln and Ser445Phe) (purchased from Genewiz) was cloned into the pET30a expression vector (a gift from J.D. Ernst) using NdeI and HindIII restriction sites. pET30a expression vectors were transformed into BL21 DE3 pLysS chemically competent E. coli (Invitrogen). NT5C2 expression was induced using 1 mM IPTG with 5 h of incubation at 37 °C. Cells were pelleted at 8,000 g for 2 min at 4 °C and resuspended in lysis buffer (50 mM NaH2PO4, 300 mM NaCl and 10 mM imidazole) with 1× protease inhibitors (GE Healthcare). Then, 1 mg/ml lysozyme was added, and samples were incubated on ice for 30 min. Lysates were centrifuged at 15,000 g for 10 min at 4 °C. Protein was subjected to electrophoresis on 9% SDS-Tris acrylamide gels and transferred to PVDF membranes. Membranes were incubated with a 1:5,000 dilution of rabbit polyclonal antibody to cN-II (ab96084, Abcam), incubated with a 1:10,000 dilution of horseradish peroxidase (HRP) -conjugated secondary antibody to rabbit (GE Healthcare) and developed using enhanced chemiluminescence (ECL; GE Healthcare). Purified protein extract (10 ml) was used to assess the enzymatic activity of wild-type and mutant proteins using the 5′-Nucleotidase Enzymatic Test kit (Diazyme) according to the provided protocol. Data are represented as the mean ± s.d. from three independent experiments.

Cell culture and drug treatment. Reh cells obtained from the American Type Culture Collection (ATCC) were grown in RPMI1640 supplemented with 10% FBS, 10 mM HEPES and 1% penicillin-streptomycin.
under 5% CO₂ at 37 °C. 293T cells (ATCC) were grown in DMEM supplemented with 10% FBS and 1% penicillin-streptomycin under 5% CO₂ at 37 °C. 6-mercaptopurine, 6-thioguanine, cytarabine, doxorubicin, gemcitabine and prednisolone (Sigma) were serially diluted in RPMI before use at the indicated concentrations.

**Transient transfection and lentivirus gene transfer.** NT5C2 cDNA for wild-type and mutant (Arg238Trp, Arg367Gln and Ser445Phe) was cloned into the lentiviral vector pLenti (a gift from M.R. Phillips) using SalI and XbaI restriction sites. All plasmids were sequence verified. cDNA constructs were transfected into 293T cells along with helper plasmids using the calcium phosphate method to produce replication-defective virus. Supernatant was harvested 48 h later and used to transduce Reh cells (whose NT5C2 sequence was verified as wild type) supplemented with 8 mg/ml polybrene (Sigma). Virus-containing medium was replaced 24 h after infection. Cells were monitored 72 h after infection for infection efficiency by the detection of GFP-positive cells using a FACScan (BD). We plated 200,000 infected cells per well in 200 ml of medium in triplicate for drug treatment with 6-mercaptopurine, 6-thioguanine, cytarabine, doxorubicin, gemcitabine and prednisolone (Sigma). Cells were incubated for 24–72 h and then assayed for apoptosis by Annexin V–PE and 7-AAD staining (Annexin V–PE Apoptosis Detection kit, BD Pharmingen) followed by flow cytometry analysis using a FACScan. The percentages of cells positive and negative for Annexin V and/or 7-AAD staining were analyzed with FlowJo software (version 7.6.1, Tree Star). Data were plotted relative to results obtained with no chemotherapy treatment, and error bars represent the standard deviation from three independent determinations. We harvested 1 × 10⁶ cells for protein at the time of plating. Briefly, cells were pelleted at 200 g for 5 min and resuspended in 100 ml of RIPA buffer with 1× protease inhibitors (GE Healthcare), incubated on ice for 15 min and centrifuged at 15,000 g for 10 min at 4 °C. Protein was subjected to electrophoresis on 9% SDS-Tris acrylamide gels and transferred to PVDF membranes.

Membranes were incubated with a 1:5,000 dilution of antibody to Flag (F3165, Sigma), incubated with a 1:10,000 dilution of HRP-conjugated secondary antibody to mouse (GE Healthcare) and developed using ECL (GE Healthcare).

**HPLC determination of nucleotides.** Reh cells were transiently infected with NT5C2 constructs. After infection, cells were treated with 10 µM 6-mercaptopurine for 24 h in duplicate. After 24 h, 5 × 10⁶ cells were washed twice with PBS, and cells pellets were frozen at –80 °C. Intracellular accumulation of thioguanine nucleotides (6-mercaptopurine active metabolites) was determined by a reversed-phase liquid chromatography assay as described previously⁴¹.

**Statistical analysis.** Statistical analysis of enzymatic and chemoresistance assays was performed using the two-sided unpaired Student’s t test. Statistical analysis of the clinical and biological characteristics of study subjects with NT5C2 mutations was performed using Fisher’s exact test. P < 0.05 was considered to be statistically significant.

---

34. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
35. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).
36. Robinson, J. T. et al. Integrative genomic viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
37. Adzhubei, I. A. et al. A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
38. Kumar, P., Henikoff, S. & Ng, P.C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **4**, 1073–1081 (2009).
39. Totrov, M. & Abagyan, R. The contour-buildup algorithm to calculate the analytical molecular surface. *J. Struct. Biol.* **116**, 138–143 (1996).
40. Totrov, M. & Abagyan, R. Rapid boundary element solvation electrostatics calculations in folding simulations: successful folding of a 23-residue peptide. *Biomolecules* **60**, 124–133 (2001).
41. Dervieux, T. et al. HPLC determination of thiopurine nucleosides and nucleotides in vivo in lymphoblasts following mercaptopurine therapy. *Clin. Chem.* **48**, 61–68 (2002).