JPET Fast Forward. Published on June 28, 2011 as DOI: 10.1124/jpet.111.182873

JPET #182873

Genetic variants in NT5C2 are associated with its expression and cytarabine sensitivity in HapMap cell lines and in AML patients.

Amit K Mitra, Kristine R Crews, Stanley Pounds, Xueyuan Cao, Tanya Feldberg, Yogita Ghodke, Varsha Gandhi, William Plunkett, M. Eileen Dolan, Christine Hartford, Susana Raimondi, Dario Campana, James Downing, Jeffrey E Rubnitz, Raul C Ribeiro, and Jatinder K. Lamba.

Department of Experimental and Clinical Pharmacology (A.K.M, T.F, Y.G and J.K.L), University of Minnesota, MN; Department of Experimental Therapeutics (V.G., W.P.), MD Anderson Cancer Center, Houston, TX; Department of Pathology, Department of Medicine, University of Chicago, IL (MED); Department of Pharmaceutical Sciences (K.R.C), Biostatistics (S.P and X.C), Pathology, (J.D. and S.C) and Oncology (D.C, J.E.R and R.C.R), St Jude Children's Research Hospital, Memphis, TN, USA

Running title: Pharmacogenetics of 5'nucleotidase NT5C2.

Corresponding Author: Jatinder K Lamba, Department of Experimental and Clinical Pharmacology, University of Minnesota, Minneapolis, 55455. Tel: 612-624-8651
Email: lamba004@umn.edu

Corresponding author:
Jatinder Lamba, PhD
Associate Professor,
Department of Experimental and Clinical Pharmacology,
College of Pharmacy,
Institute of Human Genetics,
University of Minnesota,
Minneapolis, MN-55455
Phone: (612) 624 8651
Fax: (612) 625 4294
Email: lamba004@umn.edu

Text pages: 32

Tables: 4

Figs: 8

References: 41

Abstract: 249 words

Introduction: 490 words

Discussion: 1021 words

Nonstandard abbreviations:

5'UTR: 5' untranslated region; Ara-C: 1-β-D-arabinofuranosyl-cytosine (Cytarabine); CEU: Centre d' Etude du Polymorphisme Humain; NT5C: cytoplasmic 5' nucleotidases; ESE: Exonic splicing enhancer; IVS: Intervening sequence (intron); HapMap: Haplotype Map (of human genome); mt: mutant; LD: Linkage disequilibrium SNPs: Single Nucleotide Polymorphisms; WT: wild-type; YRI: Yoruba people in Ibadan, Nigeria; BMGC: Biomedical Genomics Center; EFS: event free survival; OS: overall survival; MSP: Methylation specific PCR; MRD: Minimal residual disease.

Recommended section assignment: Metabolism, Transport, and Pharmacogenomics

Abstract:

Cytosolic 5'-nucleotidase II (NT5C2) is involved in the development of ara-C resistance and has been associated with clinical outcome in patients receiving ara-C- based chemotherapy. NT5C2 inactivates ara-C by dephosphorylating ara-CMP to ara-C. In this study, we sequenced NT5C2 in genomic DNA samples from International HapMap project panels with European (CEU; n=90) or African (YRI; n=90) ancestry. We identified 41 genetic variants (one insertion-deletion and 40 SNPs), including three non-synonymous SNPs (Thr3Ala, Lys47Arg, Gln136Arg). Twenty-five SNPs were novel and 16 overlapped with the HapMap data. Subjects with African ancestry had NT5C2 mRNA expression levels that was significantly higher than those with European ancestry ($p = 0.005$). Further, there was a correlation between NT5C2 mRNA expression and ara-C sensitivity in CEU but not in YRI cell lines. None of the non-synonymous SNPs demonstrated any effect on NT5C2 activity. The genotypes of several SNPs were significantly associated with NT5C2 mRNA expression and/or ara-C sensitivity in CEU cell lines but very few were significant in YRI cell lines. Most interestingly, SNPs (LD group CEU.12) in the 5'UTR were associated with NT5C2 expression and ara-C sensitivity in HapMap cell lines, and with NT5C2 mRNA expression and ara-C sensitivity in diagnostic leukemic blasts from pediatric patients with acute myeloid leukemia. Functional genomics analysis demonstrated that the promoter SNP rs11191612 was associated with altered luciferase activation in reporter assays and altered DNA-protein binding in gel shift assays. These results suggest that genetic variations in NT5C2 influence its expression and, potentially, cellular responses to nucleoside analogs.

JPET Fast Forward. Published on June 28, 2011 as DOI: 10.1124/jpet.111.182873
This article has not been copyedited and formatted. The final version may differ from this version.

Introduction:

Cytarabine (1–β-arabinofuranosylcytosine, ara-C) a deoxycytidine nucleoside analog, is one of the most effective chemotherapeutic agents used in the treatment of acute myeloid leukemia (AML) (Wang et al., 1970). Ara-C is a prodrug that requires activation through intracellular phosphorylation to ara-C-triphosphate (ara-CTP). Incorporation of ara-CTP in place of deoxycytidine triphosphate (dCTP) results in chain termination, thereby blocking DNA and RNA synthesis and causing leukemic cell death, which in turn is associated with therapeutic response to ara-C (Galmarini et al., 2001b;Kufe et al., 1980;Major et al., 1981;Raza et al., 1992). Thus, the intracellular concentration of ara-CTP is one of the determinants of the clinical efficacy of ara-C chemotherapy. Deoxycytidine kinase is the key enzyme catalyzing the first phosphorylation step resulting in the formation of ara-CMP. However, cytoplasmic 5'nucleotidases catalyze dephosphorylation of ara-CMP, thereby reducing the amount of ara-C for conversion to ara-CTP (Amici and Magni, 2002).

Mammalian 5'-nucleotidases are a family of seven known enzymes (five cytosolic, one mitochondrial and one extracellular) that catalyze the dephosphorylation of ribo- and deoxyribo-nucleoside monophosphates to their respective nucleoside and inorganic phosphates (Bianchi et al., 1986). Intracellular nucleotidases like cytosolic 5'-nucleotidase II (NT5C2 or cN-II, or high Km 5'-nucleotidase) and cytoplasmic 5'-nucleotidase-III (NT5C3) are involved in the final step of dephosphorylation prior to export of nucleosides out of the cell. While NT5C2 has a preference for 6-hydroxypurine nucleotide mono-phosphates (like IMP, dIMP, GMP, dGMP, and XMP) as substrates, NT5C3 only hydrolyzes pyrimidine monophosphates (Amici and Magni, 2002;Pesi et al., 1994).

Multiple reports have suggested involvement of cytosolic nucleotidases in drug resistance to ara-C and/or gemcitabine and their influence on clinical outcome. It has been shown that NT5C2 (cN-II) expression levels correlate with *in vitro* sensitivity of primary leukemic cells to ara-C (Galmarini et al., 2005). Development of resistance to nucleoside analogues in cancer cell lines has been associated with increased NT5C2 expression (Dumontet et al., 1999;Schirmer et al., 1998). In adult AML patients, high mRNA levels of NT5C2 as well as higher ratio of NT5C2/DCK mRNA expression along with unfavorable karyotype have been associated with shorter overall survival (Galmarini et al., 2001a;

JPET Fast Forward. Published on June 28, 2011 as DOI: 10.1124/jpet.111.182873
This article has not been copyedited and formatted. The final version may differ from this version.

Galmarini 2003). High NT5C2 expression has also been associated with disease free survival, overall survival and resistance to ara-C (Galmarini et al., 2005). Suzuki et al reported that high NT5C2 mRNA levels in patients with high-risk myelodysplastic syndrome (MDS) treated with ara-C containing chemotherapy was associated with shorter median or overall survival (Suzuki et al., 2007).

The inter-patient variability in the expression/activity of NT5C2 due to genetic variation could thus contribute to differences in treatment response. We identified genetic variations in NT5C2 by re-sequencing it in the genomic DNA from HapMap European and African ancestry panels, and determined the association of SNPs with NT5C2 mRNA expression levels and ara-C sensitivity in HapMap cell lines. Then, we functionally characterized selected SNPs *in vitro* and determined the association between potentially significant NT5C2 variants with clinical measures in pediatric AML patients undergoing treatment including ara-C.

Materials and Methods

Reagents:

The Expand Hi Fidelity PCR system was obtained from Roche (Indianapolis, IN); the TOPA TA cloning kit, pcDNA3.1 Directional TOPO Expression Kit and the pET directional cloning kit were obtained from Invitrogen (Carlsbad, CA, USA). Restriction enzymes and JM109 cells were obtained from Promega (Madison, WI). RNeasy Minikit and Plasmid Plus kit were procured from Qiagen Inc. Anti-Human antibody was obtained from Abcam (Cambridge, MA). All other reagents used were of molecular biology grade.

HapMap cell lines:

We used Epstein-Barr virus-transformed B-lymphoblastoid HapMap cell lines derived from 30 Centre d' Etude du Polymorphisme Humain (CEU) trios (2 parents and a child) (n = 90, European descent) and 30 Yoruba trios (n = 90, African descent, referred to as YRI) to identify genetic variants in NT5C2. The purpose of using the same cell lines that have been used in the International HapMap project was to allow us to utilize the genotype data generated as part of the HapMap project. Additionally, the genome wide gene expression data using Affymetrix Exon array was used to extract the expression levels of NT5C2 (GSE7761).

JPET Fast Forward. Published on June 28, 2011 as DOI: 10.1124/jpet.111.182873
This article has not been copyedited and formatted. The final version may differ from this version.

Cells were grown in an RPMI 1640 medium supplemented with 2 mM L-glutamine (Lonza Walkersville, Inc., Walkersville, MD) and 15% heat-inactivated serum at 37°C under 5% CO_2. DNA, RNA, and cytoplasmic fractions were extracted from the cell lines using standard protocols. Genomic DNA was used to discover novel genetic variants in the NT5C2 gene. Ara-C sensitivity was determined as described earlier (Hartford et al., 2009). Briefly, percent cell survival values were determined using alamarBlue (Biosource international, Camarillo, CA) after 72 hr exposure to 1, 5, 40 and 80 μM araC and survival curves were generated. The area under the survival curve (AUC) was calculated using the trapezoidal rule and AUC values were $\log_2$ transformed before statistical analysis.

Identification of Sequence Variations in the NT5C2 Gene:

All the 18 coding exons, the flanking intronic sequences as well as 1.5 kilobases of the 5'-UTR of the NT5C2 gene were PCR-amplified using primers and conditions listed in **Supplementary Table 1**. The primers were designed using PrimerSelect module of Lasergene v6.0 software (DNAStar) and synthesized at University of Minnesota, Biomedical Genomics center (BMGC). The primer sequences were verified using UCSC BLAT (http://genome.brc.mcw.edu/cgi-bin/hgBlat) to eradicate the possibility of amplification of any non-specific DNA sequences. Amplification was carried out in a 1x PCR buffer using 10 ng of genomic DNA, 10 pmol each of forward and reverse primers, 0.2 mM dNTPs, and 1.5 units of Taq polymerase (Expand High Fidelity PCR system; Roche). Prior to sequencing, unincorporated nucleotides and primers were removed by incubation with shrimp alkaline phosphatase and exonuclease I (USB, Cleveland, OH) for 30 min at 37°C, followed by inactivation at 80°C for 15 min. DNA Sequencing was carried out with an ABI Prism 3700 automated sequencer (Applied Biosystems, Foster City, CA) at BMGC using the PCR primers or internal primers (sequence available on request). Sequences were assembled using the Phred-Phrap-Consed package (University of Washington; Seattle, WA; http://droog.mbt.washington.edu/PolyPhred.html), which automatically detects the presence of heterozygous single-nucleotide substitutions by fluorescence based sequencing of PCR products (Nickerson et al., 1997) and Seqman, the Multiple Sequence Alignment module of Lasergene v6.0 software (DNA STAR, Inc.).

Bioinformatics Analysis:

SNPs identified by resequencing as well as from the HapMap database were analyzed using the following bioinformatic tools: 1) Splicing: SNPs at exon/intron junctions were screened using splice site prediction programs as Berkeley Drosophila Genome Project (http://www.fruitfly.org/seq_tools/splice.html) and the Web-based software Exonic splicing enhancer finder (ESE) (http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process_home), which screens for the potential splice sites and binding affinities for the four main serine/arginine (SR)-rich proteins, namely, SF2/ASF, SC35, SRp40, and SRp55. 2)Transcription factor binding sites: Promoter SNPs were evaluated for the loss/gain of known *cis*-regulatory motif binding sites using TRANSFAC Professional (phttps://portal.biobase-international.com/cgi-bin/portal/login.cgi). 3) Prediction of consequences on protein structure and/or function: Non-synonymous single nucleotide polymorphisms were evaluated using prediction programs such as SIFT (Sorting Intolerant From Tolerant, http://blocks.fhcrc.org/sift/SIFT.html), (Ng and Henikoff, 2001;Ng and Henikoff, 2003) and PolyPhen (Polymorphism Phenotyping, http://www.bork.embl-heidelberg.de/polyphen) (Ramensky et al., 2002;Sunyaev et al., 2000;Sunyaev et al., 2001).

Luciferase reporter Assays:

To study the functional effects of promoter SNPs on NT5C2 transcription, luciferase reporter gene constructs were created for wild type (WT) and mutant (MUT) alleles of 2 NT5C2 SNPs (rs11191612, rs10748839). HapMap genomic DNA samples with homozygous genotype for these SNPs were amplified using forward and reverse primers designed to capture the SNP loci of interest (listed in supplementary Table 2). The PCR product was cloned in pCR2.1-TOPO using TOPO-TA cloning kit (Invitrogen, Carlsbad CA) followed by restriction digestion and cloning into pGL3basic reporter vector (Promega).

Cos7 cells were maintained in DMEM medium containing 10% fetal bovine serum. Cells were plated at a density of $2x10^5$ cells/well in a 24 well plate. Cells were transfected with

JPET Fast Forward. Published on June 28, 2011 as DOI: 10.1124/jpet.111.182873
This article has not been copyedited and formatted. The final version may differ from this version.

luciferase reporter constructs (800 ng) and β gal expression vector (250 ng :psv-B gal) that encoded the beta galactosidase protein using Lipofectamine 2000 (Invitrogen, Carlsbad CA) at 1:2.5 ratio. Forty eight hours later the cells were harvested and lysed, followed by luciferase assay (Promega, Madison, WI), β-Galactosidase enzyme assay system (Promega, Madison, WI) and Bradford (Invitrogen, Long Island, NY) protein estimation following respective manufacturers' protocols. Results were reported as the ratio of firefly luciferase light units to β gal units or protein units, and values were also expressed as a percentage of the activity of the appropriate WT construct. All assays were performed in triplicate; i.e., three independent transfections were performed, and each experiment was repeated twice for a total of six independent determinations

Electrophoretic mobility shift assays (EMSA):

EMSAs were performed for NT5C2 promoter SNPs rs11191612 and rs10748839 to determine their potential effect on transcription regulation. Since the NT5C2 transcript has been reported to be highly expressed in bone marrow and cervix, nuclear extracts from HL60 and HeLa cells were chosen for use in EMSA assays. Briefly, oligonucleotides for rs11191612, A allele (5'- GCTGTGCGGTaAGAATTCTTA-3') and G allele (5'- GCTGTGCGGTgAGAATTCTTA-3 and for rs10748839, T allele (5'-CTCCTATACCtGGTGCCCACA-3') and C allele (5'- CTCCTATACCcGGTGCCCACA-3') were biotin labeled using non-radioactive Lightshift Chemiluminescent assay system (Pierce Biotechnology, Rockford, IL) as per manufacturer's instructions, each 20µl reaction contained 20 fmol of 3'-end labeled DNA target, 10µg of nuclear extract, 1 µg poly(dI-dC) and 1X binding buffer. A 200-fold molar excess of unlabeled specific DNA target was used to perform competition experiments. Reaction mixtures were incubated for 20 min at room temperature followed by electrophoresis on a 6% pre-cast DNA retardation gel (Invitrogen). The binding reactions were then transferred electrophoretically to a Biodyne B modified nylon membrane (Pierce, Rockford, IL). The transferred DNA was cross-linked at 120mJ/cm^2 for 1min. Subsequently, Biotin-labeled DNA-protein interactions were detected using non-radioactive Lightshift Chemiluminescent assay system (Pierce Biotechnology, Rockford, IL) and visualized following exposure to X-ray film.

NT5C2 expression constructs:

Full-length NT5C2 cDNA was amplified from one of the samples having WT sequence and cloned into a pET101 expression vector using a champion pET directional TOPO expression kit (Invitrogen, Long Island, NY), as per manufacturer's instructions. Site-directed mutagenesis was performed using a QuikChange site-directed mutagenesis kit (Stratagene, La Jolla, CA) to create NT5C2 3-mt, 47-mt and 136-mt expression constructs (numbers refer to amino acid positions). The empty pET101 vector, pET101-NT5C2 WT, pET101-NT5C2 3-mt, pET101-NT5C2 47-mt, and pET101-NT5C2 136-mt expression vectors were then transformed into the BL21 star Escherichia coli provided with the pET101 Directional TOPO Expression Kit (Invitrogen), following manufacturer's instructions. Briefly, NT5C2 expression in BL21 cells was induced by 1mM isopropyl -D-thiogalactoside, and cells were harvested after 5 hrs. The cell pellet was suspended in 50 mM Tris/HCl buffer, pH 7.4, containing 4 mM dithiothreitol, and the soluble fraction was prepared as per the manufacturer's instructions. The soluble fraction was checked for NT5C2 expression by western blotting before proceeding to activity assays.

Western Blotting:

BL21 cell cytosols were subjected to electrophoresis on 10% Tris–HCl acrylamide gel. Proteins were then transferred to nitrocellulose membranes and the membranes were incubated with mouse monoclonal anti-NT5C2 primary antibody (Abcam plc.) followed by secondary antibody. Immunoreactive proteins were detected using the ECL Western Blotting System (Amersham Pharmacia, Piscataway, New Jersey, USA) and visualized using BioRad ChemiDoc XRS+ system.

Diazyme 5'-Nucleotidase (5'-NT) Enzymatic Assay:

5'-Nucleotidase activity of the recombinant WT and mutant NT5C2 protein expressed in BL21 was measured using Diazyme 5'-Nucleotidase (5'-NT) Enzymatic Test Kit. The assay is based on the principle that enzymatic hydrolysis of 5'-monophosphate (5'-IMP) forms inosine which is converted to hypoxanthine by purine nucleoside phosphorylase (PNP), which, in turn is then converted to uric acid and hydrogen peroxide (H_2O_2) by xanthine oxidase (XOD). This H_2O_2 is then quantified by a Trinder reaction providing

JPET Fast Forward. Published on June 28, 2011 as DOI: 10.1124/jpet.111.182873
This article has not been copyedited and formatted. The final version may differ from this version.

average 5'-Nucleotidase activity when compared against a calibrator of known 5'-NT activity.

NT5C2 methylation:

CpG region sequence of the *NT5C2* was obtained from UCSC Golden Path Genome Browser Database. Methyl Primer Express (Applied Biosystems) was used to predict promoter CpG islands and to design primers for methylation specific PCR (MSP). Five different MSP primer sets (P1 to P5 in Fig 6) covering 5 predicted CpG rich sites were designed. For every site a methylated and an unmethylated primer set was synthesized at BMGC. All primers were between 15 and 25 bp in length. We used genomic DNA from 10 sensitive and 10 resistant unrelated HapMap cell lines from CEU and YRI panels for MSP. Sodium bisulfite modification of DNA was performed using the EZ DNA methylation kit (ZYMO research, Irvine, CA) according to manufacturer's protocol with minor modification.

The MSP reaction was carried out in 25 µl reaction mixture containing 2.5 µl of HotStar Taq10 × buffer, 0.5 µl of 10 mM deoxyribonucleotide triphosphate (dNTP) mix, 0.2 µl HotStar polymerase (Qiagen, Valencia, CA) 0.5 µM of each primer and 20–25 ng of bisulfite modified genomic DNA. PCR amplification consisted of initial activation step at 95^0C, followed by 34 cycles consisting of denaturation at 94^0C for 15 s, annealing at 54-57^0C (different for different primers) for 30 s, extension at 72^0C for 30 s, and final extension at 72^0C at 10min. The MSP products were separated by electrophoresis using 2% Agarose gel. The extent of methylation can only be judged qualitatively by this method. Methylation was considered negative (−) when no band was present, and positive (+) when a band was present. We also performed LINE1 methylation analysis on the same samples using Pyrosequencing for global methylation.

Samples from patients with AML:

Primary bone marrow samples were obtained from children newly diagnosed with AML enrolled on St. Jude AML97 (Rubnitz et al., 2009) or St Jude AML02 (Rubntiz et al, clinicaltrial.gov identifier) studies, after obtaining informed consent from them or from their parents/guardians, with assent from the patients, as appropriate. Genomic DNA

was extracted and samples were evaluated for *NT5C2* SNPs. This study and use of these samples were approved by the institutional review board at St. Jude.

Cohort I: St Jude AML97

The eligibility for the enrollment, treatment plan and clinical outcome of this clinical trial have been published elsewhere (Rubnitz et al., 2009). Briefly, patients were randomly assigned to receive either a daily short infusion or a continuous infusion of ara-C. Bone marrow aspirates were obtained at diagnosis in all patients and after day 1 and 2 of ara-C treatment as described previously (Crews et al., 2002). Leukemic cells were separated by Ficoll-Hypaque density-gradient centrifugation and intracellular ara-CTP levels were determined in samples obtained after day1 and day2 of ara-C treatment using high performance liquid chromatography as described earlier (Crews et al., 2002). The patient population included 58% white, 22% black, and 20% with other ethnic backgrounds. Event-free survival (EFS) and overall survival (OS) were estimated as described previously (Rubnitz et al., 2009).

Cohort II: St Jude AML02

Details of eligibility of enrollment, study design and clinical outcome for the AML02 protocol are described elsewhere (Rubnitz et al., 2010). Briefly, patients were randomized to receive induction therapy containing either high-dose cytarabine (3 g/m^2, given every 12 hours on days 1, 3, and 5) or low-dose cytarabine (100 mg/m^2) given every 12 hours on days 1-10 plus daunorubicin and etoposide as described earlier (Rubnitz et al., 2010). The patient population included 69.6% white, 18.7% black, and 11.7% with other ethnic backgrounds. Minimal residual disease (MRD), EFS and OS were estimated as described earlier (Rubnitz et al., 2010).

In vitro sensitivity (MTT) assay of primary leukemic blast samples

In vitro sensitivity of leukemic cells obtained at diagnosis to cytarabine was determined in patients enrolled on the AML02 protocol using the 4-day 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-tetrazoliumbromide (MTT) cytotoxicity assay as described previously (Holleman et al., 2004; Lamba et al, 2011). If necessary, samples were enriched to achieve more than 80% leukemic blasts by the use of magnetic cell sorting (Miltenyi Biotech, Germany). The leukemic cells were exposed to 6 different concentrations of

JPET Fast Forward. Published on June 28, 2011 as DOI: 10.1124/jpet.111.182873
This article has not been copyedited and formatted. The final version may differ from this version.

cytarabine (0.002-2.5 ng/μL) or to culture media (untreated) in a 96-well plate format. After 96-h incubation, MTT was added and cell viability was measured 6 h later. The drug concentration lethal for 50% of treated cells (LC_{50}) was estimated using nonlinear curve fitting. It has been demonstrated previously that 96 h of drug incubation is necessary to achieve cytotoxic effect in primary leukemic cells.

A cytospin slide for untreated cells was prepared on days 0 and 4. The percentage of leukemic blasts in each sample was determined by Geimsa staining. Samples with 90% or more leukemic cells on day 0, with 70% or more leukemic cells in the untreated wells after 4 days of culture, and with an optical density greater than 0.05 arbitrary units were considered suitable for evaluation. For samples in which 50% cytotoxicity was not achieved even at the highest tested concentration or ara-C (i.e., complete resistance), the LC_{50} value was assigned as twice the highest concentration tested (Beesley et al., 2006). For cells with viabilities less than 50% at all tested concentrations (i.e., high sensitivity), the LC_{50} value was assigned as half of the lowest concentration tested.

NT5C2 mRNA expression in clinical AML blast samples

NT5C2 mRNA expression levels in diagnostic leukemic blasts were extracted from Affy U133A array data that was available from 137 patients (AML97, n=41; AML02, n=96). NT5C2 SNPs (n=16) representing major LD groups and that were potentially significant were genotyped in genomic DNA from AML97 (n=55) and AML02 (n=232) patients using Sequenom platform that uses MALDI-TOF based chemistry. The SNPs were selected based on analysis of HapMap data and criteria which prioritized TagSNPs, coding and regulatory SNPs.

Statistical Analysis:

The association of SNP genotype with NT5C2 expression and ara-C sensitivity measured as AUC in HapMap cell lines was explored using an analysis of variance model that uses a Toeplitz correlation matrix (Wolfinger, 1996) to account for correlation of a child's measurements to those of their parents. The Kruskal-Wallis test was used to compare expression levels and ara-C LC_{50} values across groups of AML patients classified by SNP genotype, clinical risk group, or race. Spearman correlation coefficients were used to

JPET Fast Forward. Published on June 28, 2011 as DOI: 10.1124/jpet.111.182873
This article has not been copyedited and formatted. The final version may differ from this version.

characterize the association of the number of minor alleles with MRD, NT5C2 expression and LC50 values. Event-free survival (EFS) was defined as the time elapsed from protocol enrollment to induction failure, relapse, secondary malignancy or failure with patients free of these events censored at last follow-up. Overall survival (OS) was defined as the time elapsed from protocol enrollment to death, with living patients censored at last follow-up. The statistic of Jung, Owzar, and George (Jung et al., 2005) was used to characterize the association of the number of minor alleles with EFS and OS. Cox regression models were used to explore associations of genotype with EFS and OS while accounting for previously identified prognostically important variables. Pleiotropic effects were identified using projection onto the most interesting statistical evidence (PROMISE; (Pounds et al., 2009; Pounds S, 2011). All tests are two-sided and no multiple-testing adjustments were performed in these exploratory analyses.

Results

Identification of Genetic Variants in NT5C2:

All the coding exons and the proximal promoter were sequenced in CEU (30 trios with European ancestry, n=90) and YRI (30 trios with African ancestry, n=90) samples. Fig 1A shows the snapshot from UCSC genome browser (http://genome.ucsc.edu/index.html?org=Human) of regions resequenced in NT5C2 along with key regions of functional importance (CpG site, 7x regulatory potential and conserved transcription factor binding sites). A total of 41 genetic variants were identified including 40 SNPs and one indel (insertion-deletion) by resequencing. Of these 41 SNPs, 2 were present in the promoter, 5 in exons, 30 in the intron and 4 in the 3'UTR (Table 1). Seventeen SNPs were found exclusively in the African ancestry group and 12 exclusively in the European ancestry group; 10 SNPs were common to both groups. Sixteen SNPs overlapped with SNPs present in the HapMap database. Of the exonic SNPs three were non-synonymous resulting in Thr3Ala, Lys47Arg, Gln136Arg changes.

Analysis of linkage disequilibrium on SNPs identified in the current study and from HapMap database (www.HapMap.org) was performed using Haploview. Based on the LD plot and results from tagger (Haploview) we were able to group the strongly linked SNPs

JPET Fast Forward. Published on June 28, 2011 as DOI: 10.1124/jpet.111.182873
This article has not been copyedited and formatted. The final version may differ from this version.

($r^2 > 0.8$) in CEU and YRI into nine and ten distinct groups, respectively (Fig 1B and 1C). Eight CEU SNPs and 19 YRI SNPs did not occur in linkage with other SNP/s.

Association of NT5C2 SNPs with its mRNA Expression and ara-C sensitivity in HapMap Cell Lines:

NT5C2 mRNA expression in 87 CEU (representing 29 trios) and 90 YRI (30 trios with African ancestry) samples was obtained from the publicly available database Gene expression omnibus (GSE7761). NT5C2 mRNA expression demonstrated significant inter-individual variability (CEU: 8.34 ~ 9.90; YRI: 8.29 ~ 9.91) and significant ethnic differences with African ancestry having significantly higher NT5C2 mRNA as compared to European ancestry (9.35+/- 0.03 vs. 9.21+/- 0.031, P = 0.002) (Fig.2). NT5C2 expression was positively correlated with ara-C AUC in CEU (coefficient estimate, r=0.197 p=0.04) but not in YRI (coefficient estimate, r=0.044, p=0.67) panels (Figs. 2B and 2C).

We analyzed the association of NT5C2 SNPs (identified by sequencing as well as from HapMap database) with its mRNA expression and cellular sensitivity to ara-C in HapMap cell lines (Table 2). Within CEU samples, SNPs represented by groups CEU.1, CEU.2 and CEU.3 ($p<0.05$) were significantly associated with ara-C cellular sensitivity measured as AUC (area under the survival curve; see materials and methods) in growth inhibition assays. Table 2 and Fig 3A demonstrates the association of the representative SNP within each group (rs10748839 for CEU.1; rs11191558 for CEU.2 and rs2274339 for CEU.3). In all these cases, the minor allele was associated with increased sensitivity to ara-C (as depicted by lower AUC). Two SNPs within CEU.1 (rs10748839 and rs4307650) were in the promoter and two (rs943035 and rs943036) were present in the 3' end of NT5C2 gene. The rest of the SNPs within CEU.1 were intronic. SNP rs11191548 within group CEU.2 was present in 3'UTR of NT5C2.
In the 87 CEU samples, NT5C2 mRNA expression was significantly associated with SNPs in group CEU.7 (two new SNPs identified by re-sequencing 78616 and 78626) as well as with SNPs rs2296569, rs2274341, rs1163849 and rs10786736 ($p<0.05$, Table 2, Fig 3B).
SNPs within groups CEU.6 (includes rs1598702, rs11191612 and rs1163238), CEU.10 (includes rs7092200 and rs3736922), CEU.11 (includes rs4917384 and rs1163073) and CEU.12 (represented by rs1163075 but includes 8 SNPs that occur in LD) were significantly

JPET Fast Forward. Published on June 28, 2011 as DOI: 10.1124/jpet.111.182873
This article has not been copyedited and formatted. The final version may differ from this version.

associated with both NT5C2 mRNA expression and sensitivity to ara-C (Table 2, Fig 3C). Both CEU.11 and CEU.12 groups represent SNPs present in the 5'UTR of NT5C2. Subjects homozygous for minor allele GG for rs11191612 (a promoter SNP within group CEU.6) had higher NT5C2 expression ($p<0.05$) as well as greater resistance to ara-C measured as area under the cell survival curve ($p<0.05$) in the cellular growth inhibition assays (Fig. 3C). For rs7092200, a 3'UTR SNP in group CEU.10 subjects with at least one T allele had higher ara-C AUC ($p=0.009$) and greater NT5C2 expression ($p=0.019$). Similarly for both rs1163087 (group CEU.10) and rs10786736 SNPs presence of G allele was also associated with higher NT5C2 expression as well as cellular resistance to ara-C ($p<0.05$, Fig. 3C and Table 2).
Among YRI samples rs11191549, a 3'UTR SNP within group YRI.1 was associated with ara-C sensitivity ($p=0.036$, Fig. 3A).

Although multiple SNPs demonstrated potential for alternate splicing when analyzed by bioinformatic tools, we could not find any evidence of alternate splicing using cDNA samples from selected HapMap cell lines representing different alleles of potential SNPs (data not shown).

In vitro Luciferase reporter assays

We used Transfac and Matinspector to determine if the NT5C2 promoter SNPs (rs11191612, rs10748839) associated with ara-C sensitivity and/or mRNA levels within HapMap cell lines were disrupting or creating any transcription factor binding sites. Results from Transfac analysis (version 10) demonstrating differences in the transcription binding sites for A and G allele of rs11191612 and T and C allele of rs10748839 are depicted in Figs. 4A and 4D.

NT5C2 –pGL3 reporter plasmids representing rs11191612 and rs10748839 SNPs were transfected into Cos7 cells. Transactivation potential of rs11191612 A and G allele and rs10748839 T and C allele were compared with respect to control pGL3 basic (without any NT5C2 promoter). As is shown in Fig. 4B, the NT5C2 reporter plasmid with variant (rs11191612:G allele) had a significantly increased (4 fold) transactivation potential as compared to WT (rs11191612:A allele) ($p=0.001$). This observation complements the

JPET Fast Forward. Published on June 28, 2011 as DOI: 10.1124/jpet.111.182873
This article has not been copyedited and formatted. The final version may differ from this version.

earlier observation of higher NT5C2 mRNA expression as well as greater AUC for ara-C survival curve in subjects homozygous for rs11191612 GG genotype as compared to subjects with homozygous for A allele (Fig. 3C). For rs10748839 polymorphism, transactivation potential of variant (C) allele was greater as compared to WT (A) allele (p=0.01) (Fig 4 E).

Electrophoretic mobility shift assays (EMSA)

EMSAs were performed to compare the binding efficiencies of the NT5C2 variant allele's rs11191612 and rs10748839 using nuclear extracts from HL60 and HeLa cell lines as described in materials and methods. For rs11191612 SNP, we observed differential protein binding for the WT and the variant allele. Overall the variant allele demonstrated weaker binding as compared to the WT allele for HeLa cell line but no shift was observed for HL 60 (Fig. 4C). For rs10748839 polymorphism the variant (C) allele demonstrated stronger binding to factors in the nuclear extracts as compared to the WT (T) allele for both HeLa and HL60 nuclear extracts (Fig. 4F). Additionally EMSA assays were performed for two promoter SNPs in LD block CEU.12 that were also present in a potentially regulatory region as predicted by 7x regulatory potential and DNA hypersensitivity sites (Supplementary Fig. 1A). As shown in Supplementary Fig1 B and C presence of variant allele resulted in reduced protein binding upon incubation with HeLa and THP-1 nuclear extracts.

Analysis of NT5C2 Activity

To determine the functional significance of 3 nonsynonymous coding variants we expressed NT5C2 wild-type (WT) and variant cDNAs (NT5C2.3mt, NT5C2.47mt NT5C2.136mt) in BL21 *E. coli* using pET101 expression vectors. The recombinant proteins were obtained 5 hr after IPTG induction and were analyzed for NT5C2 expression by western blotting using a NT5C2-C-ter antibody. Recombinant NT5C2 wild-type and mutant proteins were expressed in equivalent amounts, whereas no NT5C2 protein was detected with empty pET101 vector (inset in Fig. 2B). NT5C2 activity was measured using Diazyme 5'-Nucleotidase (5'-NT) Enzymatic Test Kit at different time points. As shown in Fig.5 the variant NT5C2 isoforms demonstrated no significant difference in NT5C2 activity as compared to the WT isoform.

NT5C2 Methylation

MSP was performed with all the 5 primer sets (for 5 putative CpG sites) on 10 most sensitive and 10 most resistant unrelated CEU and YRI cell lines. Commercially available control methylated and unmethylated DNA were used as positive controls (Epitect PCR Control DNA set, Qiagen) for all the 5 CpG sites. Both positive controls demonstrated amplification with respective methylated and unmethylated primer sets. None of CEU and YRI cell lines showed any band in methylated primer set while band was present in unmethylated primer set (Fig. 6B); indicating that in HapMap cell lines methylation does not contribute towards regulation of NT5C2 gene expression. To assess the global methylation status in these cell lines we performed LINE1 analysis using Pyrosequencing and found approximately 80% LINE1 methylation in these cell lines.

Association of NT5C2 SNPs with mRNA expression in leukemic cells from AML patients:

A total of 16 SNPs within NT5C2 were genotyped in DNA samples collected at diagnosis from AML patients. Four out of sixteen SNPs were not present in patient samples and two occurred with the minimum allele frequency of < 0.005 and were not included in the analysis. Since mRNA expression was determined in patients diagnostic leukemic blasts as described earlier, the data from two clinical trials was combined. In total, 137 subjects (41 from St. Jude AML97 and 96 from St. Jude AML02) had both NT5C2 genotype and microarray data available. NT5C2 expression levels did not differ significantly according to race (p=0.59) or risk group (p=0.26). We observed a significant association of group CEU.6 SNP represented by rs11598702 in Fig 7A (occurs in LD with promoter SNP rs11191612, described earlier) with expression of NT5C2 (p=0.0043, Table 3). In accordance with the results in the HapMap samples, the minor allele (A allele for rs11598702 and G for rs11191612) was associated with higher expression of NT5C2 in AML cells. For rs1163075 (SNP within group CEU.12), the C allele demonstrated higher NT5C2 mRNA expression levels in HapMap cell lines as well as in AML patient samples (CC vs. CT vs. TT, p=0.0012) (Fig. 7A and tables 2 and 3).

JPET Fast Forward. Published on June 28, 2011 as DOI: 10.1124/jpet.111.182873
This article has not been copyedited and formatted. The final version may differ from this version.

Additionally SNPs rs4917996, (representing CEU.1) and rs1926029 (representing CEU.3) were significantly associated with NT5C2 mRNA expression levels in leukemic cells from AML patients (CC vs. CT vs. TT, p<0.0002, Table 3).

Association of NT5C2 SNPs with *in vitro* ara-C sensitivity of primary leukemic cells from AML patients: Diagnostic blast ara-C cytotoxicity (measured as LC_{50}) was determined in 76 primary leukemic samples and demonstrated significant variation. The median LC_{50} value was 0.39 ng/µl (range 0.001-5.0 ng/µl); ara-C LC_{50} was not associated with race (p=0.27). Provisional risk group was significantly associated with ara-C sensitivity of leukemic blasts (p=0.001), hence we analyzed the association of SNPs in all patients, as well as within the high and standard provisional risk groups (provisional risk group). Interestingly, rs1163075 (a group CEU.12 SNP) that was associated with ara-C cytotoxicity as well as NT5C2 mRNA expression levels in HapMap cell lines was significant predictor of *in vitro* ara-C sensitivity in AML blasts. The presence of the C allele was associated with higher NT5C2 expression and was also associated with greater ara-C LC_{50} (p= 0.04) (Fig. 7B). Additionally, rs10786736 which was associated with NT5C2 mRNA expression as well as ara-C cytotoxicity in HapMap samples (Fig. 3C), was associated with ara-C sensitivity of primary AML leukemic blasts. In consensus with HapMap results, the GG genotype was associated with greater ara-C LC_{50} vs. the CG genotype (p=0.0051 Fig. 7B and table 3). A group CEU.2 SNP rs11191558 that was also associated with ara-C cytotoxicity in HapMap samples (Fig. 3A) was significantly associated with ara-C cytotoxicity in AML leukemic blasts. The median LC50 of leukemic blasts from patients with the GG genotype was higher versus in patients with the AG genotype (p=0.03, table 3).

Association of NT5C2 SNPs with clinical phenotype measures in AML patients: Within the St. Jude AML97 cohort, there was no significant association of NT5C2 SNPs with ara-CTP levels (day 1 and day 2) or clinical response. The heat-map in Fig. 8 shows the association pattern of the most significant SNPs (from data above) occurring with the MAF of greater than 0.10 with multiple parameters in AML patients. Interestingly, CEU.12 SNP rs1163075 which was associated with lower NT5C2 expression and ara-C AUC/LC_{50} in both CEU and primary AML samples, demonstrated

a trend towards association with slightly higher ara-CTP levels in the AML97 cohort and with better EFS and OS in both the AML97 and AML02 cohorts. CEU.11 SNP rs4917384 was associated significantly with day 22 MRD levels in the AML02 cohort (p=0.02). However, SNP rs1163075 did not show a significant association with EFS (p = 0.3) or OS (p = 0.3) in a Cox regression model that included predictors that were identified as clinically or statistically important in an earlier study (Rubnitz et al, 2009).

Pleiotropic Effects in HapMap Cell Lines and AML Patients

SNPs were screened for pleiotropic effects on NT5C2 expression and ara-C resistance in HapMap cell lines and for pleiotropic effects on NT5C2 expression, clinical response, and event-free survival in AML patients. In HapMap cell lines, the SNP rs1163075 showed a significant pleiotropic effect (p = 0.0053) characterized by a negative association of the minor allele with both NT5C2 expression and cytarabine resistance. Among AML patients, this SNP also showed a significant pleiotropic effect (p = 0.0233) characterized by the minor allele showing a negative association with NT5C2 expression and a positive association with both clinical response and event-free survival.

Discussion

NT5C2 is a cytosolic 5'-nucleotidase that has been implicated in the inactivation of ara-C by dephosphorylating ara-CMP to ara-C (Amici et al., 1997; Amici and Magni, 2002). Although reports suggest that expression of NT5C2 is a significant prognostic marker of worse clinical outcome in AML and MDS patients its role in ara-C response is still unclear (Mazzon et al., 2003; Galmarini, 2007). It has been hypothesized that because NT5C2 is involved in substrate dNTP cycles survival of highly proliferating NT5C2-expressing cells may favor disease recurrence, thus shortening disease free and overall survival (Galmarini et al., 2005; Galmarini, 2007).

It is still not clear whether the role of NT5C2 in response is due to its involvement in ara-CMP dephosphorylation or due to modification of cellular dNTP pools in leukemic cells or perhaps with the leukemic phenotype.

None-the-less there is evidence in the literature for the association of high NT5C2 expression with worse clinical response as well as with ara-C resistance in multiple cancer cell lines and experimental models. Hence we sought to determine if genetic polymorphisms within NT5C2 are associated with its expression and/or activity and therefore influence ara-C sensitivity. Since there are no data in the literature on sequencing of NT5C2 for SNP discovery, we sequenced NT5C2 gene in European (30 CEU trios) and African (30 YRI trios) ancestry HapMap panels to maximize the use of publicly available HapMap genotype data. Forty-one genetic variants including 1 in/del were identified by sequencing the coding and proximal promoter of NT5C2. Five coding SNPs were identified including three non-synonymous SNPs (Thr3Ala, Lys47Arg, Gln136Arg) which did not have any influence on NT5C2 activity when expressed as recombinant protein.

We observed strong linkage between multiple SNPs at the NT5C2 locus both in CEU and YRI samples (Figs.1B and 1C). We observed that NT5C2 expression was significantly higher in the YRI versus the CEU panel; however in the CEU cell lines expression of NT5C2 was directly correlated with ara-C cytotoxicity; this relationship was not observed in the YRI cell lines. This might be one of the factors contributing to lack of significant genotype-phenotype association observed in our analysis within the YRI samples (Fig. 2) (Aplenc et al., 2006). Within the CEU panel NT5C2 SNPs demonstrated significant association with its mRNA expression and *in vitro* ara-C sensitivity in both CEU HapMap cell lines and diagnostic leukemic blasts from AML patients (Figs. 3 and 7). Table 5 provides a summary of the most interesting genotype-phenotype association observed in CEU samples and AML patient samples for SNPs occurring with the minimum allele frequency of >0.10. Interestingly, the presence of minor alleles for SNPs within groups CEU.1-3, 7, 10-12 were associated with lower NT5C2 expression and ara-C sensitivity where as CEU.6 group SNPs were associated with increased expression and ara-C resistance. As indicated in Table 5, within each group we have SNPs that are linked and are present in 5'UTR or proximal promoter, exon, intron or 3'UTR. Therefore, this observed association could be due to any of the SNPs within a group, or driven primarily by promoter and 3'UTR SNPs, or due to small effects of multiple SNPs. Functional studies on two 5'UTR variants and our results indicate that although the group CEU.1 SNP

JPET Fast Forward. Published on June 28, 2011 as DOI: 10.1124/jpet.111.182873
This article has not been copyedited and formatted. The final version may differ from this version.

(rs10748839) has no effect on luciferase activation the variant allele (G) for CEU.6 group SNP rs11191612 was associated with increased luciferase activation (Fig. 4). This was in agreement with the association of group CEU.6 SNPs with NT5C2 mRNA expression and ara-C cytotoxicity in HapMap cell lines and in AML patient samples (Figs. 3C and 7). Li et al had reported in a genome wide analysis association of CEU.6 group SNP (rs11598702) with ara-C sensitivity (Li et al., 2009).

The most interesting SNPs in the present study include a group of 5'UTR SNPs with consistently demonstrated association with mRNA expression and ara-C cytotoxicity in HapMap and AML samples. Interestingly three SNPs within this LD group (i.e. rs7913461, rs1891292 and rs1891293) were present in a region of high 7X regulatory potential and ENCODE DNaseI hypersenstivity clusters (UCSC genome browser), there by indicating potentially regulatory significance.

Although our results did not demonstrate strong association of NT5C2 SNPs with clinical phenotype measures in AML patients, we observed a consistent pattern of association of CEU.12 SNP rs1163075 with clinical outcome in both AML97 and AML02 cohorts. In accordance with these results CEU.12 SNPs (rs7913461 and rs1891292) also demonstrated reduced binding for the variant oligo in EMSA assays (Supplementary Fig.1). The non-significant association observed with intracellular ara-CTP levels could be due to multiple factors such as small sample size, or contribution of other genes in the ara-C metabolic pathway. Alternatively, as suggested in previous studies NT5C2 might not be catalyzing the conversion of ara-CMP to ara-C. Although in the present study we have attempted to explore the impact of NT5C2 genetic variation the clinical cohorts were not designed to evaluate the pharmacogenetics of ara-C. We also acknowledge the fact that variability in the intracellular levels of ara-CTP is regulated by multiple enzymes in the ara-C activation pathway and we have previously shown association of DCK variants (Lamba et al., 2007). In order to achieve a full understanding of the genetic basis for the variability observed in intracellular concentration of ara-CTP, future studies in our lab will be directed toward comprehensive and simultaneous evaluation of other enzymes of relevance in the metabolic pathway of ara-C.

In summary, we have identified novel genetic variants at the NT5C2 locus by sequencing genomic DNA from two ethnic groups. NT5C2 SNPs are predictive of its expression and ara-C cytotoxicity not only in HapMap cell lines but also in AML patients.

We have further performed functional characterization of selected SNPs to further understand the underlying molecular mechanism. We also evaluated the clinical implication of NT5C2 polymorphisms for its association with clinical phenotypes in AML patients receiving ara-C-based therapy. However, the most significant SNPs identified in this study need to be confirmed in a larger patient population, to confidently arrive at the conclusion that SNPs within NT5C2, at least in part play a role in predicting the drug responsiveness and to guide individualized chemotherapy in AML patients receiving ara-C or other nucleoside containing therapies which are activated by the same metabolic pathway as ara-C.

Acknowledgements:

We would like to acknowledge the support provided by Biomedical Genomics Center (BMGC) and Minnesota Supercomputing Institute at University of Minnesota. We would also like to acknowledge help from pediatric oncology education (POE) scholars, Brittney Green and Jessica Gresham.

Author contributions: Participated in research design: *Mitra AK and Lamba JK*
Conducted experiments: *Mitra AK, Crews KR, Ghodke Y, Feldberg T and Lamba JK*
Contributed data: *Mitra AK, Crews KR, Gandhi V, Plunkett W, Campana D, Raimondi S, Downing J, Rubnitz J, Ribeiro R, Dolan ME, Harford C and Lamba JK*
Performed data analysis: *Mitra AK, Pounds S, Cao X and Lamba JK*
Wrote or contributed to the writing of the manuscript: *All Authors*

Reference List

Albertioni F, Herngren L, Juliusson G and Liliemark J (1994) Protein binding of 2-chloro 2'-deoxyadenosine (cladribine) in healthy subjects and in patients with leukaemia. *Eur J Clin Pharmacol* **46**:563-564.

Amici A, Emanuelli M, Magni G, Raffaelli N and Ruggieri S (1997) Pyrimidine nucleotidases from human erythrocyte possess phosphotransferase activities specific for pyrimidine nucleotides. *FEBS Lett* **419**:263-267.

Amici A and Magni G (2002) Human erythrocyte pyrimidine 5'-nucleotidase, PN-I. *Arch Biochem Biophys* **397**:184-190.

Aplenc R, Alonzo TA, Gerbing RB, Smith FO, Meshinchi S, Ross JA, Perentesis J, Woods WG, Lange BJ and Davies SM (2006) Ethnicity and survival in childhood acute myeloid leukemia: a report from the Children's Oncology Group. *Blood* **108**:74-80.

Beesley AH, Palmer ML, Ford J, Weller RE, Cummings AJ, Freitas JR, Firth MJ, Perera KU, de Klerk NH and Kees UR (2006) Authenticity and drug resistance in a panel of acute lymphoblastic leukaemia cell lines. *Br J Cancer* **95**:1537-1544.

Bianchi V, Pontis E and Reichard P (1986) Interrelations between substrate cycles and de novo synthesis of pyrimidine deoxyribonucleoside triphosphates in 3T6 cells. *Proc Natl Acad Sci U S A* **83**:986-990.

Crews KR, Gandhi V, Srivastava DK, Razzouk BI, Tong X, Behm FG, Plunkett W, Raimondi SC, Pui CH, Rubnitz JE, Stewart CF and Ribeiro RC (2002) Interim comparison of a continuous infusion

versus a short daily infusion of cytarabine given in combination with cladribine for pediatric acute myeloid leukemia. *J Clin Oncol* **20**:4217-4224.

Dumontet C, Fabianowska-Majewska K, Mantincic D, Callet BE, Tigaud I, Gandhi V, Lepoivre M, Peters GJ, Rolland MO, Wyczechowska D, Fang X, Gazzo S, Voorn DA, Vanier-Viornery A and MacKey J (1999) Common resistance mechanisms to deoxynucleoside analogues in variants of the human erythroleukaemic line K562. *Br J Haematol* **106**:78-85.

Galmarini CM (2007) What does over-expression of cN-II enzyme signify in haematological malignancies? *Leuk Res* **31**:1325-1326.

Galmarini CM, Cros E, Thomas X, Jordheim L and Dumontet C (2005) The prognostic value of cN-II and cN-III enzymes in adult acute myeloid leukemia. *Haematologica* **90**:1699-1701.

Galmarini CM, Graham K, Thomas X, Calvo F, Rousselot P, El JA, Cros E, Mackey JR and Dumontet C (2001a) Expression of high Km 5'-nucleotidase in leukemic blasts is an independent prognostic factor in adults with acute myeloid leukemia. *Blood* **98**:1922-1926.

Galmarini CM, Mackey JR and Dumontet C (2001b) Nucleoside analogues: mechanisms of drug resistance and reversal strategies. *Leukemia* **15**:875-890.

Hartford CM, Duan S, Delaney SM, Mi S, Kistner EO, Lamba JK, Huang RS and Dolan ME (2009) Population-specific genetic variants important in susceptibility to cytarabine arabinoside cytotoxicity. *Blood* **113**:2145-2153.

Holleman A, Cheok MH, den Boer ML, Yang W, Veerman AJ, Kazemier KM, Pei D, Cheng C, Pui CH, Relling MV, Janka-Schaub GE, Pieters R and Evans WE (2004) Gene-expression patterns in drug-resistant acute lymphoblastic leukemia cells and response to treatment. *N Engl J Med* **351**:533-542.

Jung SH, Owzar K and George SL (2005) A multiple testing procedure to associate gene expression levels with survival. *Stat Med* **24**:3077-3088.

Kawasaki H, Carrera CJ, Piro LD, Saven A, Kipps TJ and Carson DA (1993) Relationship of deoxycytidine kinase and cytoplasmic 5'-nucleotidase to the chemotherapeutic efficacy of 2-chlorodeoxyadenosine. *Blood* **81**:597-601.

Kufe DW, Major PP, Egan EM and Beardsley GP (1980) Correlation of cytotoxicity with incorporation of ara-C into DNA. *J Biol Chem* **255**:8997-900.

Lamba JK, Crews K, Pounds S, Schuetz EG, Gresham J, Gandhi V, Plunkett W, Rubnitz J and Ribeiro R (2007) Pharmacogenetics of deoxycytidine kinase: identification and characterization of novel genetic variants. *J Pharmacol Exp Ther* **323**:935-945.

Lamba JK, Crews KR, Pounds SB, Cao X, Gandhi V, Plunkett W, Razzouk BI, Lamba V, Baker SD, Raimondi SC, Campana D, Pui CH, Downing JR, Rubnitz JE and Ribeiro RC (2011) Identification of predictive markers of cytarabine response in AML by integrative analysis of gene-expression profiles with multiple phenotypes. *Pharmacogenomics* **12**:327-339.

Li L, Fridley BL, Kalari K, Jenkins G, Batzler A, Weinshilboum RM and Wang L (2009) Gemcitabine and arabinosylcytosin pharmacogenomics: genome-wide association and drug response biomarkers. *PLoS One* **4**:e7765

Major PP, Egan EM, Beardsley GP, Minden MD and Kufe DW (1981) Lethality of human myeloblasts correlates with the incorporation of arabinofuranosylcytosine into DNA. *Proc Natl Acad Sci U S A* **78**:3235-3239.

Mazzon C, Rampazzo C, Scaini MC, Gallinaro L, Karlsson A, Meier C, Balzarini J, Reichard P and Bianchi V (2003) Cytosolic and mitochondrial deoxyribonucleotidases: activity with substrate analogs, inhibitors and implications for therapy. *Biochem Pharmacol* **66**:471-479.

Ng PC and Henikoff S (2001) Predicting deleterious amino acid substitutions. *Genome Res* **11**:863-874.

Ng PC and Henikoff S (2003) SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res* **31**:3812-3814.

Nickerson DA, Tobe VO and Taylor SL (1997) PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. *Nucleic Acids Res* **25**:2745-2751.

Pesi R, Turriani M, Allegrini S, Scolozzi C, Camici M, Ipata PL and Tozzi MG (1994) The bifunctional cytosolic 5'-nucleotidase: regulation of the phosphotransferase and nucleotidase activities. *Arch Biochem Biophys* **312**:75-80.

Pieters R, Huismans DR, Loonen AH, Peters GJ, Hahlen K, van der Does-van den Berg, van Wering ER and Veerman AJ (1992) Relation of 5'-nucleotidase and phosphatase activities with immunophenotype, drug resistance and clinical prognosis in childhood leukemia. *Leuk Res* **16**:873-880.

Pounds S, Cao X, CC, YJ, CD, EWE, PC-H and Relling MV (2011) Integrated Analysis of Pharmacokinetic, Clinical, and SNP Microarray Data using Projection onto the Most Interesting Statistical Evidence with Adaptive Permutation Testing. *International Journal of Data Mining and Bioinformatics* **5**:143-157.

Pounds S, Cheng C, Cao X, Crews KR, Plunkett W, Gandhi V, Rubnitz J, Ribeiro RC, Downing JR and Lamba J (2009) PROMISE: a tool to identify genomic features with a specific biologically interesting pattern of associations with multiple endpoint variables. *Bioinformatics* **25**:2013-2019.

Ramensky V, Bork P and Sunyaev S (2002) Human non-synonymous SNPs: server and survey. *Nucleic Acids Res* **30**:3894-3900.

Raza A, Gezer S, Anderson J, Lykins J, Bennett J, Browman G, Goldberg J, Larson R, Vogler R and Preisler HD (1992) Relationship of [3H]Ara-C incorporation and response to therapy with high-dose Ara-C in AML patients: a Leukemia Intergroup study. *Exp Hematol* **20**:1194-1200.

Rubnitz JE, Crews KR, Pounds S, Yang S, Campana D, Gandhi VV, Raimondi SC, Downing JR, Razzouk BI, Pui CH and Ribeiro RC (2009) Combination of cladribine and cytarabine is effective

for childhood acute myeloid leukemia: results of the St Jude AML97 trial. *Leukemia* **23**:1410-1416.

Rubnitz JE, Inaba H, Dahl G, Ribeiro RC, Bowman WP, Taub J, Pounds S, Razzouk BI, Lacayo NJ, Cao X, Meshinchi S, Degar B, Airewele G, Raimondi SC, Onciu M, Coustan-Smith E, Downing JR, Leung W, Pui CH and Campana D (2010) Minimal residual disease-directed therapy for childhood acute myeloid leukaemia: results of the AML02 multicentre trial. *Lancet Oncol* **11**:543-552.

Schirmer M, Stegmann AP, Geisen F and Konwalinka G (1998) Lack of cross-resistance with gemcitabine and cytarabine in cladribine-resistant HL60 cells with elevated 5'-nucleotidase activity. *Exp Hematol* **26**:1223-1228.

Sunyaev S, Ramensky V and Bork P (2000) Towards a structural basis of human non-synonymous single nucleotide polymorphisms. *Trends Genet* **16**:198-200.

Sunyaev S, Ramensky V, Koch I, Lathe W, III, Kondrashov AS and Bork P (2001) Prediction of deleterious human alleles. *Hum Mol Genet* **10**:591-597.

Suzuki K, Sugawara T, Oyake T, Uchiyama T, Aoki Y, Tsukushi Y, Onodera S, Ito S, Murai K and Ishida Y (2007) Clinical significance of high-Km 5'-nucleotidase (cN-II) mRNA expression in high-risk myelodysplastic syndrome. *Leuk Res* **31**:1343-1349.

Wang JJ, Selawry OS, Vietti TJ and Bodey GP, Sr. (1970) Prolonged infusion of arabinosyl cytosine in childhood leukemia. *Cancer* **25**:1-6.

Wolfinger RD (1996) Heterogeneous variance-covariance structures for repeated measures. *J Agric Biol- Environ Sci* **1**:205-230.

Footnotes:

This study was supported by National Institutes of Health (NIH), National Cancer Institute [R01CA132946] (Lamba) and by the American Lebanese Syrian Associated Charities (ALSAC).

JPET Fast Forward. Published on June 28, 2011 as DOI: 10.1124/jpet.111.182873
This article has not been copyedited and formatted. The final version may differ from this version.

Legends to Figure:

Figure.1 A) Snapshot from UCSC for NT5C2 locus. The regions sequenced in HapMap panels are shown by small boxes the first panel followed by **B) LD plot of NT5C2 in samples with European (CEU) and African (YRI) ancestry.** LD plots were generated in Haploview using genotype data from the present study and from HapMap in both CEU and YRI samples. The color scheme is: white when $r^2 = 0$; shades of grey light $0<r^2<1$; and black when $r^2 = 1$. **C) LD Groups within CEU and YRI.** SNPs that are linked $r^2 > 0.8$ (and picked by tagger program) are categorized in the same groups. MAF: minimum allele frequency for the group. SNPs without rs numbers are indicated by number corresponding to table 1. SNPs genotyped in AML patients are bold and have an asterisk. LD groups CEU 10-12 (grey) represent SNPs present upstream of NT5C2 gene.

Figure. 2 A) NT5C2 mRNA expression (extracted from GSE7761) in EBV- transformed lymphoblast cell lines derived from subjects with European (CEU) and African (YRI) ancestry. Median values for Log_2NT5C2 mRNA levels are indicated by a horizontal line for each ethnic group. **B and C) Correlation of NT5C2 expression with ara-C cytotoxicity in CEU and YRI cell lines.** Ara-C cytotoxicity was determined as described in materials and methods. AUC represents area under the survival curve. *p<0.05 and **p<0.01.

Figure. 3 A) NT5C2 SNPs associated with ara-C cytotoxicity in HapMap cell lines. Box plots for the association of representative NT5C2 SNPs (rs10748839 for CEU.1; rs11191558 for CEU.2; rs2274339 for CEU.3 and rs11191549 for YRI.1) with ara-C AUC in CEU and YRI samples. Plots show medians as a line between boxes which represent 1st and the 3rd

quartiles; the whiskers represent the range after excluding the outliers. The outliers are defined as data points which fall outside of the 1st and 3rd quartiles by more than 1.5 times the inter-quartile range. Circles falling outside the whiskers represent outliers. **B) NT5C2 SNPs associated with NT5C2 mRNA expression levels in HapMap cell lines.** Box plots for association of NT5C2 SNPs rs2296569, rs2274341, rs1163249 and 123193 (representing CEU.7 group) with its mRNA expression in CEU cell lines. **C) NT5C2 SNPs associated with its mRNA expression levels and ara-C cytotoxicity in HapMap cell lines.** Box plots for association of NT5C2 SNPs rs11191612 (representing CEU.6), rs7092200 (representing CEU.10), rs1163075 (representing CEU.12) and singleton SNP rs10786736 with its mRNA expression and ara-C AUC in CEU cell lines.

Figure 4. Functional characterization of NT5C2 promoter SNPs rs11191612 and rs10748839. A and D) Differences in transcription factor binding sites for WT and variant alleles of rs11191612 and rs10748839 identified by Transfac analysis. **B and E)** Luciferase reporter assays comparing transcriptional activation NT5C2-pGL3basic vetcors with WT or variant alleles of rs11191612 and rs10748839. Transfection efficiency was normalized to B gal activity. **C and F)** Representative EMSA gels comparing binding of oligos representing WT and variant alleles of promoter SNPs after incubation with HL60 and HeLa nuclear extracts.

Figure 5. Activity of recombinant human NT5C2 WT and mutant proteins. NT5C2 WT and amino acid variant (T3A, K47N and Q136R) isoforms were expressed in BL21E coli. The columns represent activity of NT5C2 isoforms which was determined using Diazyme 5'-Nucleotidase (5'-NT) Enzymatic Test Kit. Activity of NT5C2 amino acid

variants were compared with WT protein and no difference in activity was observed. Inset shows a representative western blot when equal amounts of total protein were loaded for WT and mutant NT5C2 proteins; all of the recombinant NT5C2 variants were expressed at equivalent levels.

Figure 6. A) NT5C2 promoter with CpG sites. Bisulfite modified genomic DNA sequence of the CpG region of the human NT5C2 gene. The position of the primers that were used for methylation specific PCR after bisulfite treatment are indicated by forward and reverse arrows for Primer 1-Primer 5 covering 5 different regions of CpG island. The bisulfite treatment converted the unmethylated cytosine to Uracil which is complimentary to adenosine. **B) Methylation analysis of 5 CpG sites in NT5C2 promoter.** Representative agarose gels showing methylation patterns in sensitive and resistant HapMap (CEU) cell lines for the NT5C2 primer1 to primer 5. M represents band for the methylated primers and U for unmethylated primers. The positive control methylated and unmethylated was used from Epitect PCR Control DNA set.

Figure 7A). Association of NT5C2 SNPs with mRNA expression in diagnostic AML leukemic blasts. NT5C2 mRNA expression levels in diagnostic leukemic blasts were extracted from Affy U133A array data from 137 patients (AML97, n=41; AML02, n=96) and was analyzed for association with NT5C2 SNPs. Box plots for association of NT5C2 SNPs rs11598702 (representing CEU.6), rs1163075 (representing CEU.12), rs4917996 (representing CEU.1), and rs1926029 (CEU.3) with its mRNA expression in diagnostic leukemic blasts from AML patients. **B) Association of NT5C2 SNPs with diagnostic blast ara-C cytotoxicity.** Ara-C cytotoxicity was determined by treating

JPET Fast Forward. Published on June 28, 2011 as DOI: 10.1124/jpet.111.182873
This article has not been copyedited and formatted. The final version may differ from this version.

diagnostic leukemic blasts with varying concentration of ara-C and LC_{50} values were calculated. Box plots represent association of NT5C2 SNPs (rs1078636; rs1163075 representing CEU.12 and rs11191558 representing CEU.3) in all patients.

Figure 8. Heatmap of association of 5 most interesting NT5C2 SNPs with multiple endpoints in HapMap and AML samples. Each row represents a SNP representing different LD groups and each column represents a phenotype. Blue represents association of minor allele with reduced levels of phenotype and red represents association of minor allele with increased value for phenotype. The colors are assigned by $\log_{10}$ p value as per accompanying color scale.

Table 1. Genetic variants identified by sequencing NT5C2 in the present study.

Position from Translation start site (ATG) as +1	rs number	SNP/ indel	Chromosomal position	Location in the gene	Amino Acid Change	Bioinformatic analyses	Allele frequency	
							CEU	YRI
-20080	rs7095304	G/A	104944785	5'-UTR		Loss of V$E2F_Q4 site	0.33	0.27
-18832	rs10748839	T/C	104943537	5'-UTR		Gain in V$VDR_Q3 site	0.39	0.30
-237	rs12261294	C/T	104924942	INTRON 1			0.00	0.07
-123		G/A	104924828	INTRON 1			0.00	0.05
-69		C/T	104924774	INTRON 1		Loss of BranchSite, Loss of SR protein Binding site	0.00	0.01
7	rs10883841	T/C	104924699	EXON 2	T3A	on the protein surface	0.12	0.00
35377		C/T	104889329	INTRON 2			0.00	0.01
35519	rs72846108	G/C	104889187	EXON 3	K47N	Possibly damaging (Polyphen), Affect Protein Function (SIFT)	0.01	0.00
68714		G/A	104855992	INTRON 4			0.37	0.26
73650	rs12262171	T/C	104851056	EXON 6	Q136R	Hydrogen Bond Lost (SNP3D)	0.00	0.01
73783	rs41287482	A/G	104850923	INTRON 6		Gain in SR protein Binding site	0.02	0.00
74663	rs2274339	A/T	104850043	INTRON 7			0.36	0.27
75886	rs73353837	C/T	104848820	INTRON 8			0.00	0.06
77498	rs10786737	A/C	104847208	INTRON 10		Gain in SR protein Binding site	0.17	0.07
78554	rs12412038	G/A	104846152	INTRON 10			0.06	0.00
78616	123193C	T/G	104846090	INTRON 10			0.02	0.00
78626	123203	T/G	104846080	INTRON 10			0.01	0.00
78639		T/C	104846067	INTRON 10			0.00	0.01
79046	rs1926029	C/T	104845660	INTRON 11		Loss of SR protein Binding site	0.34	0.27
79060	rs1926030	A/G	104845646	INTRON 11			0.40	0.31
81127		A/C	104843579	INTRON 14			0.00	0.04
81177		C/T	104843529	INTRON 13			0.00	0.01
81213	rs2274341	T/A	104843493	INTRON 13			0.00	0.08
81220		A/G	104843486	INTRON 13			0.00	0.03
81426	rs72843997	T/C	104843280	INTRON 13			0.00	0.01
82050	rs3837340		104842656	INTRON 14				
82054	rs66473231	G/A	104842652	INTRON 14			0.00	0.02
83315		C/T	104841391	INTRON 14			0.03	0.00

83320	rs11191553	G/T	104841386	INTRON 14		Gain in Branch Site	0.32	0.27
83415	rs17094683	G/T	104841291	INTRON 15			0.05	0.02
83431	rs10883830	C/T	104841275	INTRON 15		Loss of SR protein Binding site	0.35	0.26
83881	rs11191551	T/G	104840825	INTRON 15		Loss of BranchSite, Gain in SR protein Binding site	0.05	0.00
83893		T/C	104840813	INTRON 15			0.00	0.01
84084	rs3736922	C/T	104840622	INTRON 16		Gain in 5SS_U2_Human site , CHANGE in Score	0.44	0.29
84096	rs34758128	T/G	104840610	INTRON 16		Loss of Splice site, Gain in SR protein Binding site	0.13	0.00
84219		G/T	104840487	EXON 17	E440E		0.00	0.01
85248	rs3740387	C/T	104839458	EXON 18	D549D		0.40	0.30
85547		A/G	104839159	3'UTR			0.00	0.01
85572	rs12573221	A/C	104839134	3'UTR			0.03	0.03
85584		C/G	104839122	3'UTR			0.00	0.01
85600	rs10786736	G/C	104839106	3'UTR			0.09	0.00

Table 2. Association of NT5C2 SNPs with its mRNA expression and ara-C cytotoxicity in lymphoblast cell lines from subjects with European (CEPH) and African (YRI) ancestry.

		Phenotype I		**Phenotype II**	
		Ara-C area under the survival curve (AUC)		(Log) NT5C2 mRNA expression	
	Genotype	mean±SD	p value	mean±SD	p value
European Ancestry: CEPH					
Group CEU.1	CC (n=11)	2895.234 ± 759.255	CC vs. CT vs. TT, *p=0.014		
represented by 10748839 C>T	CT (n=45)	3450.095 ± 643.267	CC vs. CT+TT, *p=0.005		
	TT (n=29)	3659.039 ± 868.523			
Group CEU.2					
represented by 11191558 G>A	AG (n=10)	2911.575 ± 606.627	AG vs. GG, p=0.017		
	GG (n=75)	3521.309 ± 763.643			
Group CEU.3					
represented by 2274339 A>T	AA (n=30)	3704.852 ± 812.523	AT+TT vs. AA, p=0.043		
	GT (n=47)	3331.494 ± 724.121			
	TT (n=6)	3137.643 ± 788.639			
Group CEU.6					
represented by 11598702 C>T	CC (n=10	3980.334 ± 798.856	CC vs. CT vs. TT, p=0.05	9.371 ± 0.177	CC vs. CT vs. TT, *p=0.015
	CT (n=42)	3458.257 ± 712.271	CC vs. CT+TT, *p=0.035	9.224 ± 0.299	CC + CT vs.TT, *p=0.011
	TT (n=33)	3277.691 ± 778.746		9.131 ± 0.284	
Group CEU.6					
represented by rs11191612 A>G	GG (n=10)	3980.334 ± 798.856	AA+AG vs.GG, *p=0.035	9.371 ± 0.177	AA vs. AG vs. GG, p=0.02
	AG (n=42)	3440.149 ± 718.944		9.225 ± 0.300	AA+AG vs. GG, *p=0.046
	AA (n=33)	3300.738 ± 775.766		9.129 ± 0.282	AA vs. AG +GG, *p=0.017
Group CEU.7					
represented by 78616 T>G	GT (n=3)			8.889 ± 0.171	GT vs. TT, p=0.026
	TT (n=81)			9.225 ± 0.285	
Group CEU.10					
rs7092200 C>T	CC (n=15)	3033.064 ± 820.026	CC vs. CT vs. TT, *p=0.032	9.068 ± 0.286	CC vs. CT+TT, *p=0.019
	CT (n=46)	3492.019 ± 656.968	CC vs. CT+TT, *p=0.009	9.252 ± 0.292	
	TT (n=26)	3628.546 ± 871.160		9.203 ± 0.270	

Group CEU.11					
rs4917384 C>T	CC (n=37)	3579.244 ± 783.663	CC vs. CT vs. TT, *p=0.041	9.310 ± 0.231	CC vs. CT vs. TT, *p=0.009
in LD with rs163703	CT (n=43)	3306.684 ± 658.351		9.127 ± 0.302	CC vs. CT+ TT, *p=0.002
	TT (n=4)	4027.323 ± 962.587		9.195 ± 0.217	
Group CEU.12					
rs1163087 G>A	AA (n=15)	3319.396 ± 890.147	AA vs. AG vs. GG, p=0.059	9.149 ± 0.315	AA vs. AG vs. GG, *p=0.003
in LD with 8 other SNPs	AG (n=46)	3314.117 ± 679.138	AA+ AG vs. GG, p=0.017	9.135 ± 0.285	AA vs. AG vs. GG, *p=0.0006
	GG (n=25)	3759.956 ± 799.118		9.359 ± 0.224	
rs2296569 T>C	CC (n=4)			9.334 ± 0.241	CC+CT vs. TT, p =0.02
	CT (n=25)			9.298 ± 0.269	
	TT (n=57)			9.151 ± 0.291	
rs2274341 T>A	AA (n=7)			9.339 ± 0.208	AA vs. AT vs. TT*p=0.031
	AT (n=29)			9.284 ± 0.276	AA+AT vs TT, p=0.012
	TT (n=50)			9.140 ± 0.296	
rs1163249 G>A	AA (n=3)			9.088 ± 0.032	AA+AG vs. GG, p=0.047
	AG (n=33)			9.119 ± 0.315	
	GG (n=51)			9.269 ± 0.265	
rs10786736 G>C	CC (n=1)	4622.78	CC vs. CG vs. GG, *p=0.023	9.094	CC+CG vs. GG, *p=0.019
	CG (n=13)	2992.518 ± 664.375		9.042 ± 0.332	CC vs. CG vs. GG, *p=0.005
	GG (n=71)	3449.576 ± 769.551		9.205 ± 0.289	
<u>African Ancestry : YRI</u>					
Group YRI.1					
	CC (n=49)	2787.952 ± 525.570	CC vs. CT+TT, *p=0.036		
represented by rs11191549	CT (n=36)	3020.564 ± 642.470			
	TT (n=4)	3026.675 ± 429.907			

Table 3. Association of NT5C2 SNPs with its expression and ara-C cytotoxicity in diagnostic leukemic blasts from AML patients.

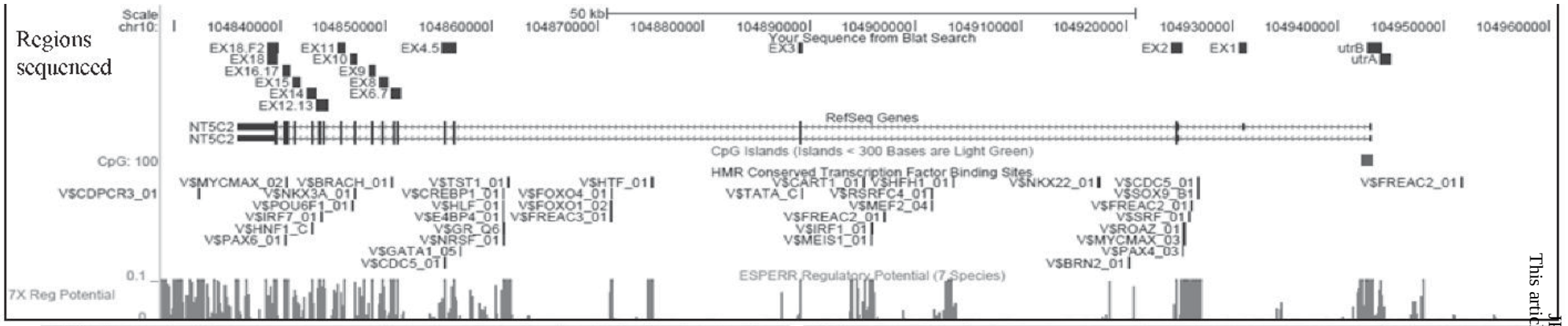
	Phenotype I Ara-C Cytotoxicity of diagnostic leukemic blasts (n=71, AML02)			Phenotype II Log NT5C2 mRNA expression in diagnostic leukemic blasts (n= 137, AML97; n=41 and AML02;n= 96)		
	Genotype	mean±SD	p value	Genotype	mean±SD	p value
CEU.12						
rs1163075	CC (n=18)	2.109±2.164	CC vs. CT vs. TT, *p=0.047	CC (n=36)	6.833±0.287	CC.vs.CT vs. TT,*p=0.0012
	TC (n=34)	1.129±1.696	CC vs. CT+ TT, *p=0.04	CT (n=73)	6.643±0.348	CC vs.CT + TT,*p=0.001
	TT (n=14)	1.235±2.051		TT (n=22)	6.586±0.297	
CEU.1						
rs4917996	AA (n=26)	1.622±1.952	AA vs. CA vs. CC, p = 0.245	AA (n=58)	6.809±0.262	AA vs. CA vs. CC, ***p=0.0001
	CA (n=31)	1.262±1.903		CA (n=57)	6.669±0.332	AA + CA vs. CC, **p=0.0006
	CC (n=11)	1.611±2.212		CC (n=20)	6.458±0.369	AA vs. CA + CC, **p=0.0005
CEU.6						
rs11598702	CC (n=9)	1.58±2.018	CC vs. CT vs. TT, p=0.3	CC (n=17)	6.974±0.200	CC vs. CT vs. TT *p=0.0043
	CT (n=24)	1.244±1.776		CT (n=53)	6.689±0.323	CC vs. CT + TT *p=0.0001
	TT (n=30)	1.414±2.049		TT (n=63)	6.644±0.334	
CEU.3						
rs1926029	CC (n=28)	1.167±1.673	CC vs. TC vs. TT, p=0.73	CC (n=69)	6.7926±0.279	CC vs. TC vs. TT, ***p<0.0001
	TC (n=30)	1.278±1.933		TC (n=50)	6.6061±0.347	CC vs. TC + TT, ***p=0.0001
	TT (n=7)	2.517±2.366		TT (n=11)	6.3808±0.358	CC + TC vs. TT, **p=0.0034
CEU.11						
rs4917384	CC (n=31)	2.015±2.153	CC vs. CT vs. TT, *p=0.041	CC (n=72)	6.724±0.340	CC vs. CT vs. TT, p=0.44
	TC (n=32)	0.937±1.606		CT (n=55)	6.679±0.327	
	TT (n=2)	2.60±3.392		TT (n=7)	6.673±0.308	
CEU.2						
rs11191558	AG (n=9)	0.232±0.247	AG vs. GG, *p=0.033	AG (n=20)	6.738±0.378	AG vs. GG, p=0.68
	GG (n=56)	1.630±2.019		GG (n=111)	6.689±0.322	
rs10786736	CG (n=10)	0.218±0.237	CG vs. GG, **p=0.0051	CC (n=1)	6.2403	CC vs. CG vs. GG, p=0.52
	GG (n=58)	1.670±2.037		CG (n=22)	6.690±0.388	
				GG(N=112)	6.705±0.322	

Table 4. Summary of NT5C2 SNPs showing most significant association in CEU HapMap cell lines and Leukemia cells from AML patients.

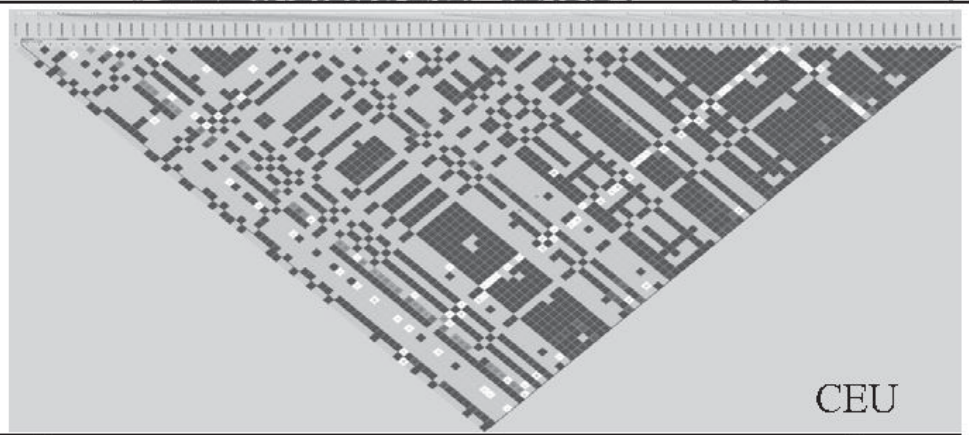
LD Group	SNP Location				HapMap samples		AML patients Diagnostic Blasts			MAF (HapMap)
	Promoter/ 5'UTR	Intronic	Coding SNP	3'UTR	mRNA Expression	AraC AUC	SNPs genotyped AML	mRNA Expression	AraC IC50	
CEU.1	4	8	1	0	⇣	↓	rs4917996	↓	⇣	0.400
CEU.3	5	20	0	1	⇣	↓	rs11191558	↓	⇣	0.353
CEU.6	1	2	0	0	↑	↑	rs11598702	↑		0.358
CEU.11	2	0	0	0	↓	↓	rs4917384		↓	0.259
CEU.12	9	0	0	0	↓	↓	rs1163075	↓	↓	0.442

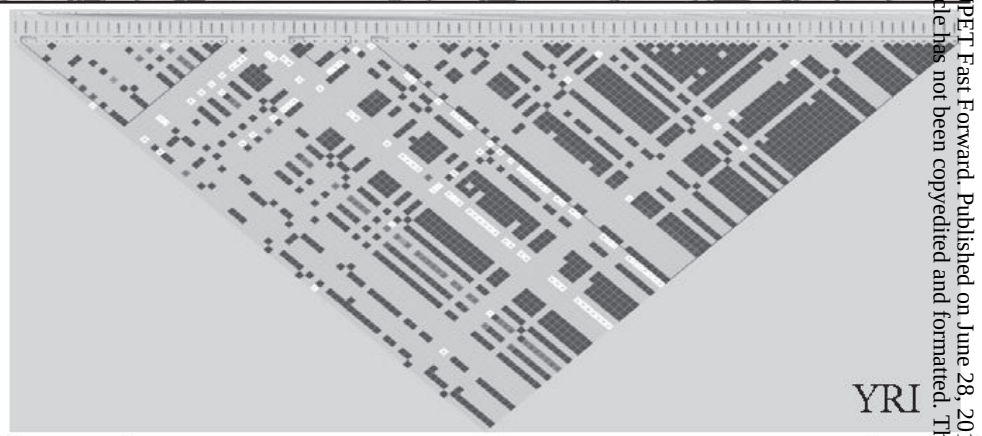
Downward arrows indicate association of minor allele with reduced expression and in vitro ara-C cytotoxicity,
Upward arrows indicate association with increased expression and ara-C cytotoxicity.
Solid arrows indicate statistically significant association and dashed arrows indicate trend not reaching statistical significance.
MAF: minimum allele frequency. Distribution of SNPs in gene within each LD group is also included.

1A)



1B)





1C)

Group/ Block	Alleles Captured	MAF
CEU.1	**rs4917996***, rs6584540, **rs3740387***, rs10748839, rs746293, rs3977751, rs7067970, rs7896547, rs7920251, rs2066323, rs4307650, rs1926030, rs10786740	0.400
CEU.2	rs11191593, rs11191557, rs12413046, rs12412038, rs12220743, rs10883832, rs11191582, rs12219304, rs11191560, **rs11191558***, rs732998, rs9633712, rs12220375, rs12217501, rs11191551, rs11191555, rs17094683, rs11191580, rs11191548	0.058
CEU.3	rs3740386, rs7897663, rs10883830, rs12355120, rs7095304, rs10883843, rs11191549, rs3781281, rs7917650, rs12414232, rs10509766, rs2148198, rs1935323, rs2274339, rs2281877, rs6584539, rs1541213, rs11191578, rs11191553, rs10883839, **rs1926029***, rs10786741, rs3781282, rs10786744, rs12775302, *68714*	0.353
CEU.4	rs7908960,*83315*, rs12573199, rs12573200, **rs12573221***	0.025
CEU.5	rs11191606, rs10786737, rs11191554, rs12773916, **rs10883841***	0.125
CEU.6	**rs11598702***, **rs1163238*,** rs11191612	0.358
CEU.7	*78616, 78626*	0.018
CEU.8	rs11191561, rs10509765	0.125
CEU.9	rs12261294, rs12240508	0.179
CEU.10	rs7092200,rs3736922	0.442
CEU.11	rs4917384,rs1163073	0.312
CEU.12	rs7913461,rs1163087,rs1712517,rs1891292,rs1712510, rs1891293,rs1163075,rs2506	0.425

Group/ Block	Alleles Captured	MAF
YRI.1	rs2281877, rs2066323, rs7897663, rs3781281, rs7095304, rs3781282, rs3977751, **rs1926029***, rs12414232, rs7896547, rs10883830, rs10883839, rs11191553, rs11191549, rs3736922, rs2274339, rs4307650, **rs3740387***, *68714*, **rs4917996***, rs1935323, rs10509766, rs12355120, rs6584539, rs10786741, rs1541213, rs7917650, rs10883843, rs6584540, rs746293, rs11191578, rs10748839, rs10786740, rs2148198, rs10786744, rs1926030, rs12775302, rs7067970, rs7920251, **rs1163238***, rs3740386	0.2
YRI.2	rs12220375, **rs11191558***, rs11191560, rs12412038, rs10883832, **rs10786736***, rs11191557, rs12217501, rs11191555, rs12413046	0.0
YRI.3	rs732998, rs12219304, rs11191593, **rs11191548**, rs9633712	0.0
YRI.4	**rs12573221***, rs7908960, rs12573200, rs12573199	0.025
YRI.5	rs7067741, **rs7916854***, rs9919482	0.093
YRI.6	rs66473231, *81220*	0.025
YRI.7	**rs12262171***, rs11191550	0.017
YRI.8	rs12240508, rs12261294	0.073
YRI.9	rs10786737, rs11191554	0.075
YRI.10	rs10509765, rs11191561	0.150

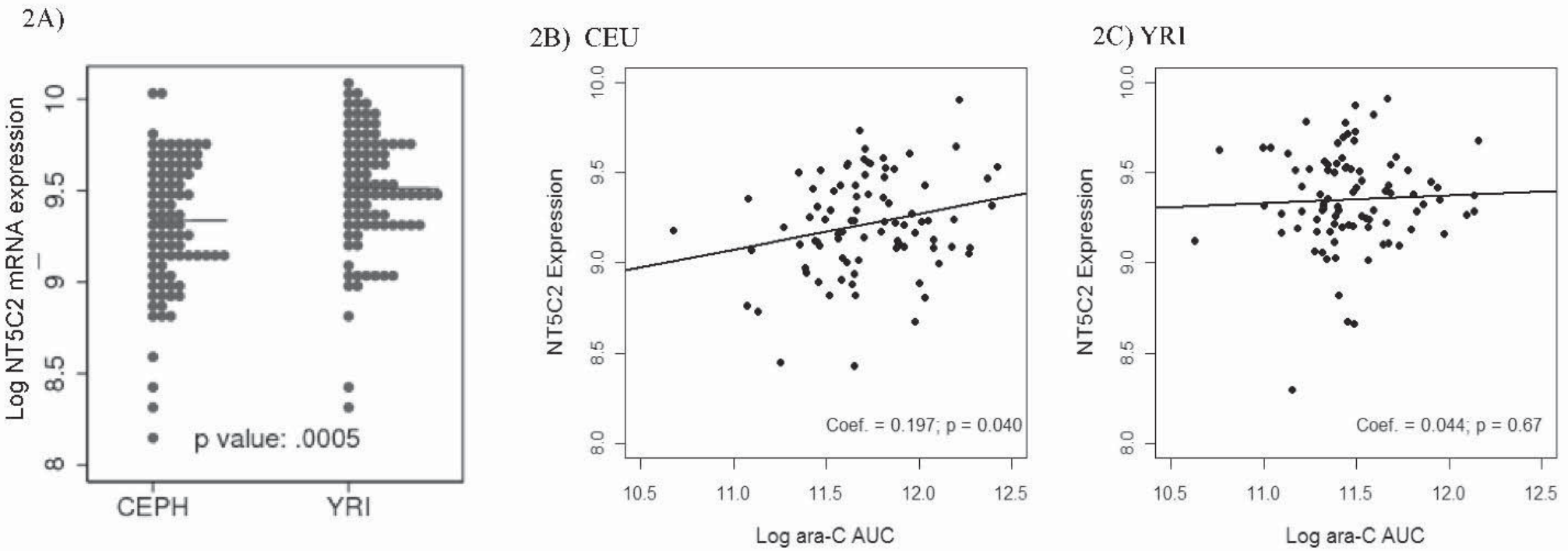
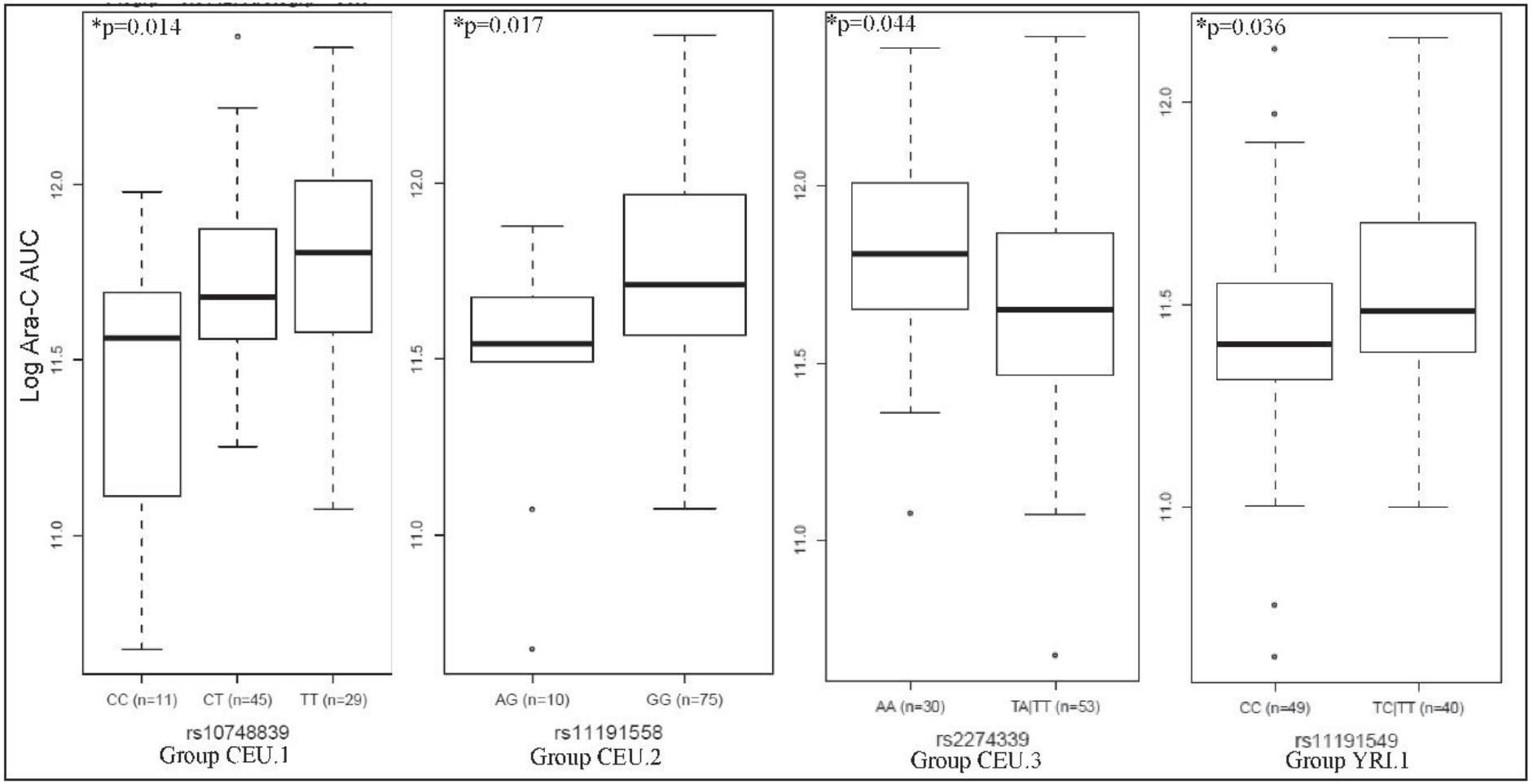
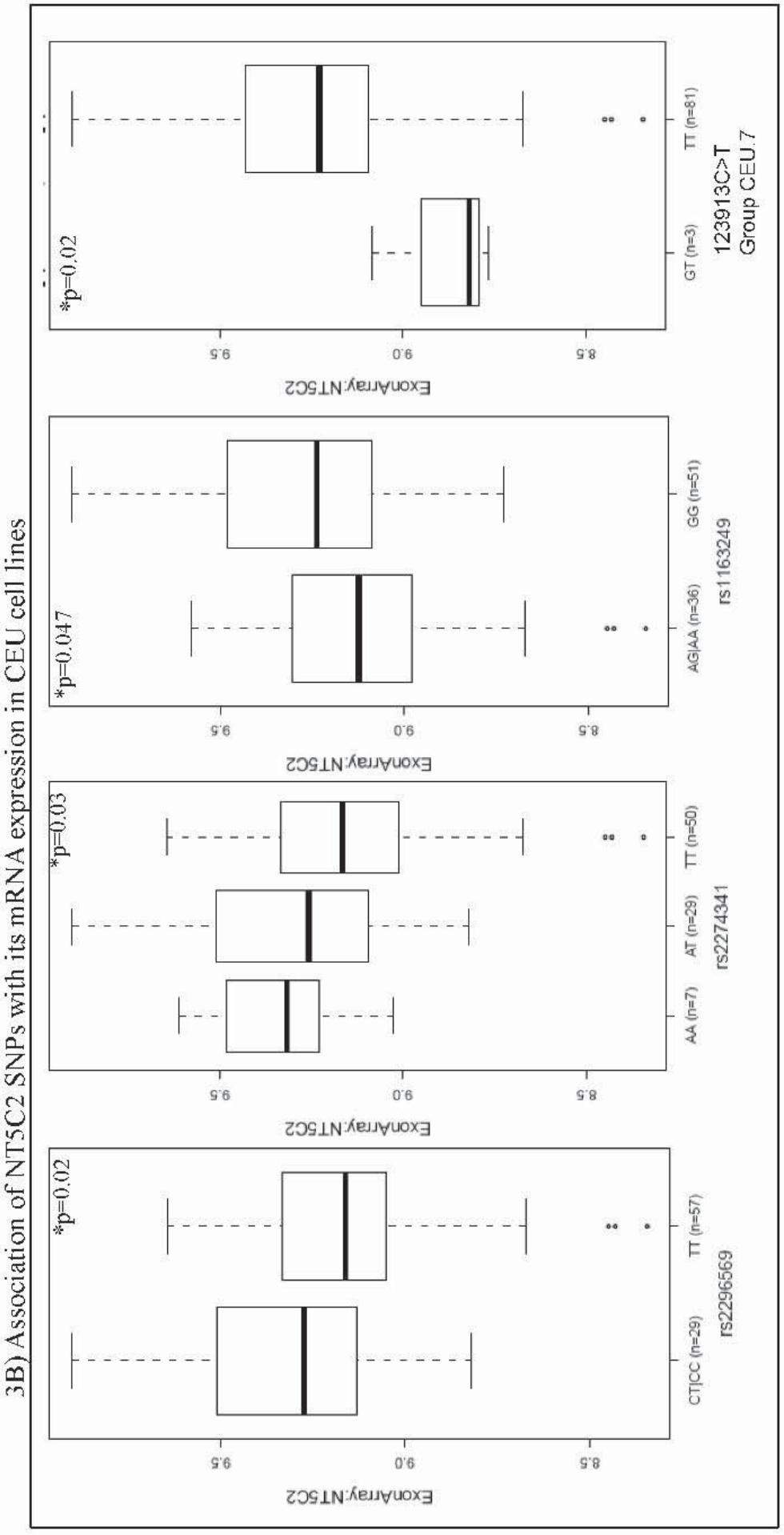


Figure 2

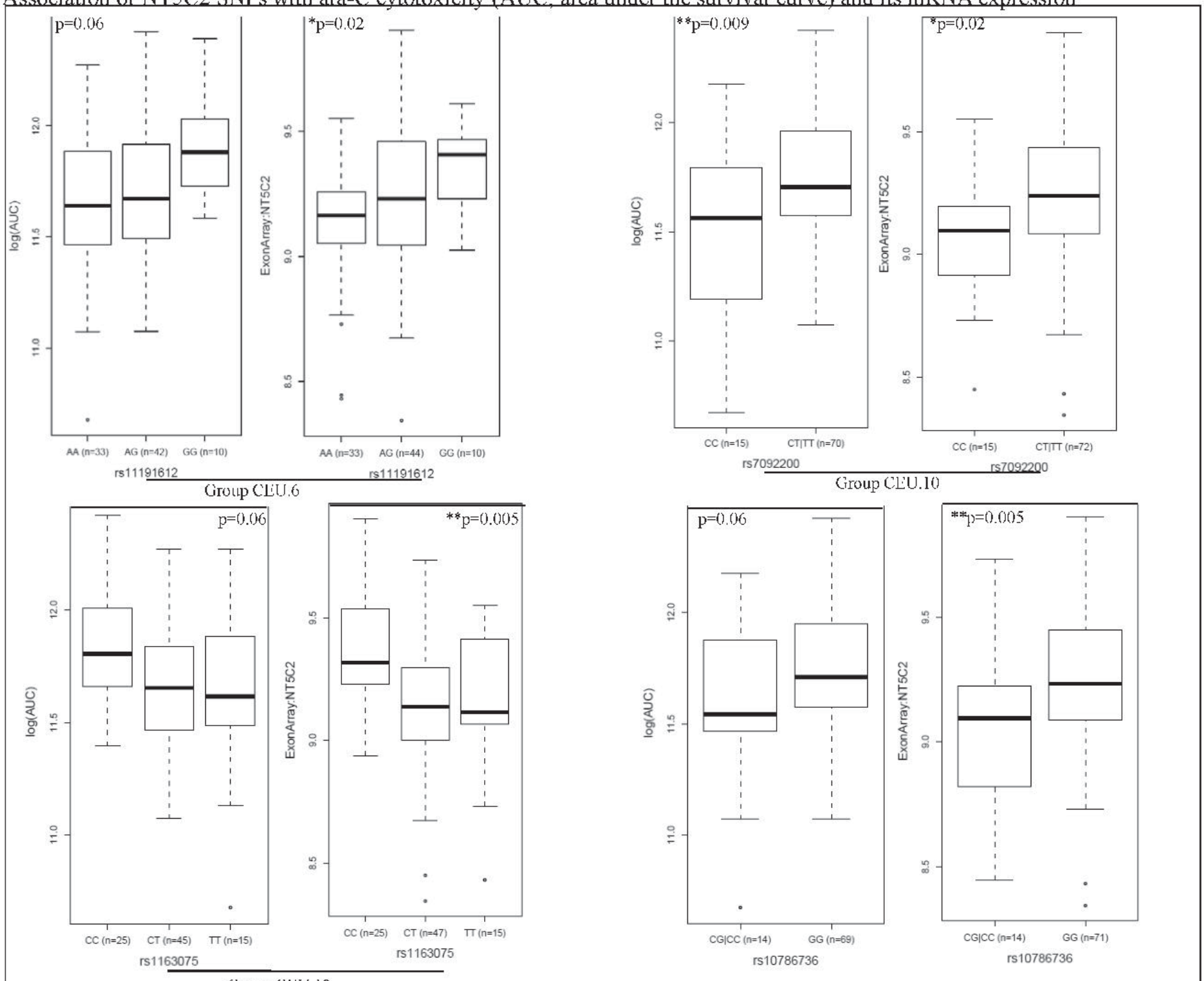
3A) Association of NT5C2 SNPs with ara-C cytotoxicity (AUC: area under the survival curve) in HapMap cell lines



JPET Fast Forward. Published on June 28, 2011 as DOI: 10.1124/jpet.111.182873
This article has not been copyedited and formatted. The final version may differ from this version.



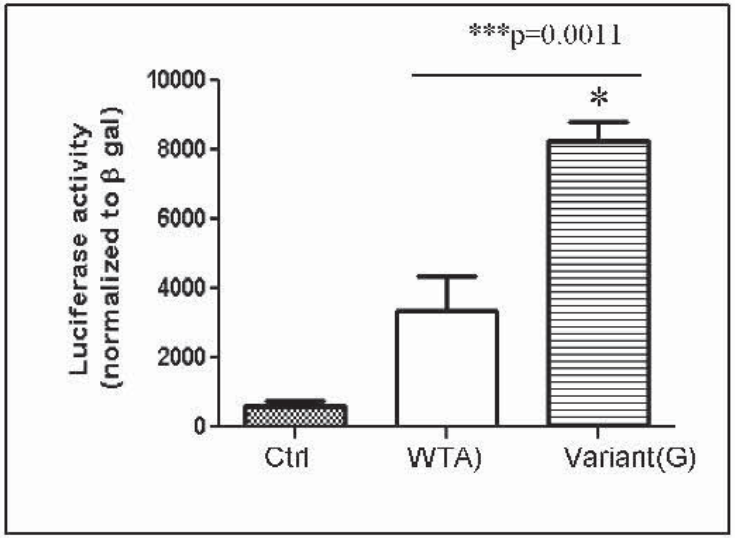
3C) Association of NT5C2 SNPs with ara-C cytotoxicity (AUC; area under the survival curve) and its mRNA expression



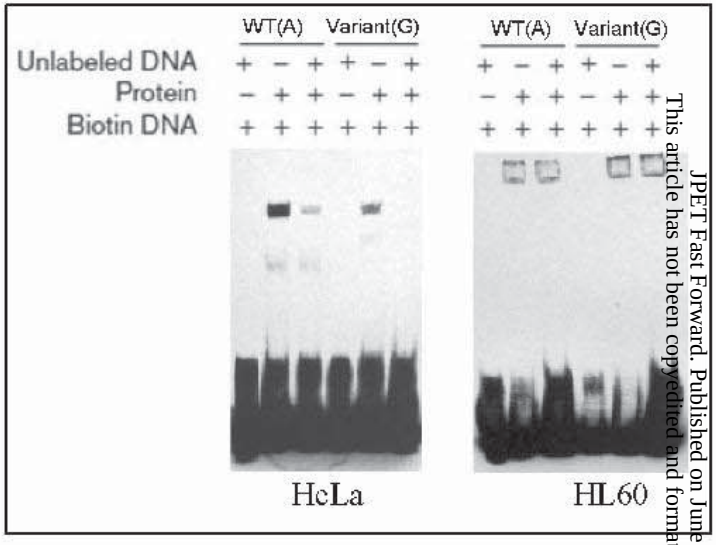
4A) Transfac analysis

rs11191612	
A allele	**G allele**
CDP	TBX5
C/EBPalpha	TCF11
C/EBPdelta	Xvent-1
3 sites for E2F	HSF2
FOXJ2	Nkx2-5
IPF1	p300
OCT4	Pax-8
POU3F2	RBP-Jkappa
STAT	Spz1
Tel-2	2 sites for SREBP

4B) Luciferase reporter assay in Cos 7



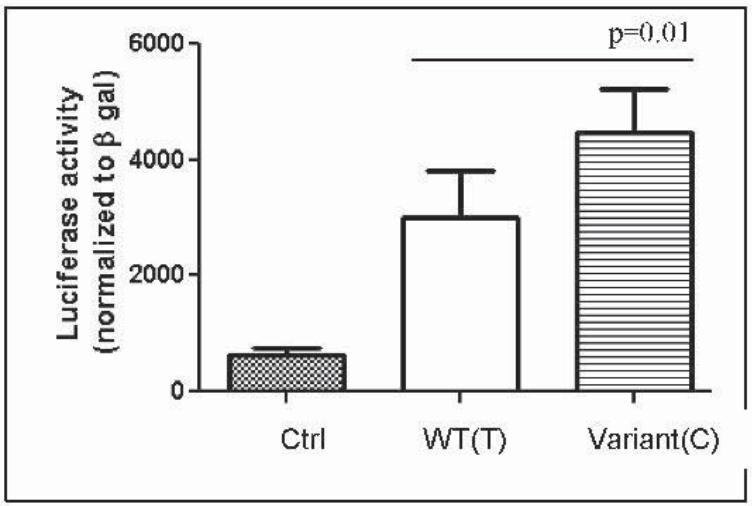
4C) EMSA



4D) Transfac analysis

rs10748839	
T allele	**C allele**
CDP	Pax-4
Hand1:E47	AP-2
PPAR	CACD
Tal-1beta:E47	TBX5
TBX5	v-Myb
TTF-1	
YY1	

4E) Luciferase reporter assay in Cos 7



4F) EMSA

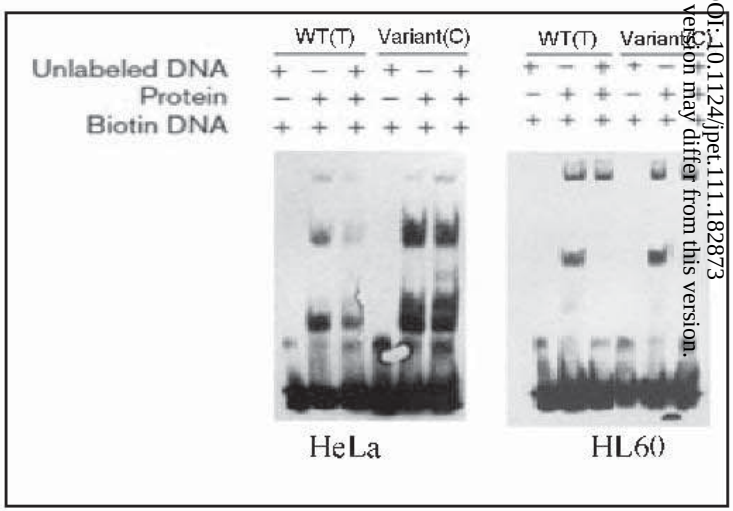
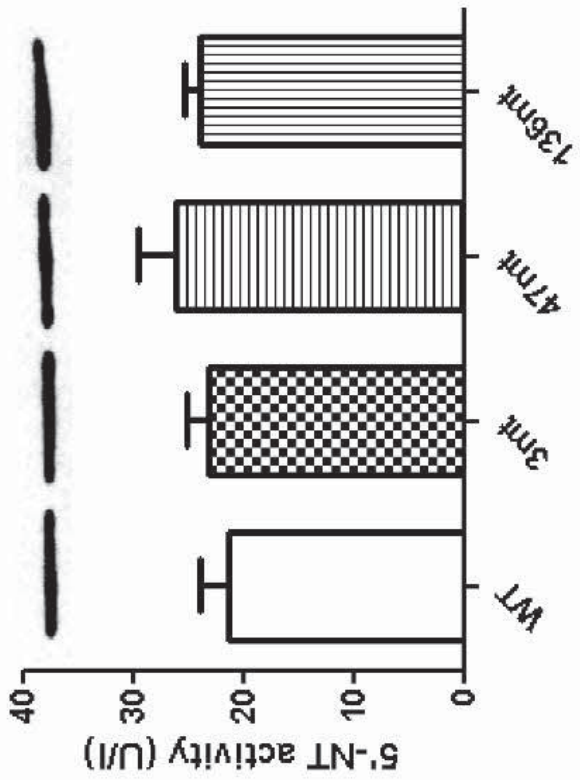


Figure 4

JPET Fast Forward. Published on June 28, 2011 as DOI: 10.1124/jpet.111.182873
This article has not been copyedited and formatted. The final version may differ from this version.

Figure 5



6A)

CGGTGTTTTATCGTTTGAAGATTTTTAGATTCGCGGGATAGATAGACGTAGCGCGGTTAGGGAGCGTCGGTTTCGTC
P1 F

GAATTTTGCGGCGGAGTTGGGGGACGGAATAGCGCGTCGGTTTTTGGTCGGGCGCGAGTAGAGGGGCGGCGCGGGGC
P1 R

GGTGTCGGGCGGCGGGGGCGGGGAGGGCGCGCGTTGAGGCGGTTGCGGTAGTTGCGCGTTGGGATTGTTGCGGTGCG

TTGGAGTCGGTGAGTGCGGCGGGTTTCGTGGGCGGGCGGGGAAGGAGCGGGCGAACGGGGAGTTAGGTGGGGAGGTG
P5 F

CGCGGTCGAATACGCGATTTTAGGGTAGCGGTCGTCGTTTGTTTTCGTTTTCGGTTTTTAGGGATTATGGTTTTAAG
P5 R

GTTTTTTTTTTTAGGTCGAGTTTTCGTTAGCGCGTTCGCGGACGTCGTACGTTTTTTAATTTTTTCGGTGCGGTTTT
P2 F

CGTTTTACGGTAGTTTTTTTTTTCGCGTCGTCGCGACGTTTTCGGTTTACGTTACGTTCGCGTTATTGTTTTTTAGT
P2 R P3 F

CGTCGGTTCGTTGGGATTTCGTTTTAGTTTTTTAGCGCGGTCGTTCGTCGCGTTTCGTTTTTTTTTTTTTTTTTTT

ACGTTGGTGGTAGTTTTTTTTTTAGTTTTGTTGACGTTTTTTAGTTGATTTTAGGTTGTTTTCGGTTATTTTTGAGG

TCGTTTTTTCGTTTTTTGTAAAAGTTTTTTCGTTTTTAGAGTTTTTCGGTCGTTTTTCGAGAAGTTAAGGGGTTTTT
P3 R

GGTGGAGCGTTAGGTTGATAGCGTTTTAGTAGGATCGCGAGAAACGGTTAGTTCGTTTGTTATACGTTCGTGTGTAA
P4 F

TATTTTTTGGAAAGGTTATAGCGAATTGTAAGTCGTTATAAAAAATATTCGGATAGTTTTTGGTTAGGGGGTTGAGT

CGTTTAGTGGCGGTTTATTTTTTTTTTTTTTCGGGGTTTTTGAAATCGGAGAGGGCGAGATGTCG
P4 R

6B)

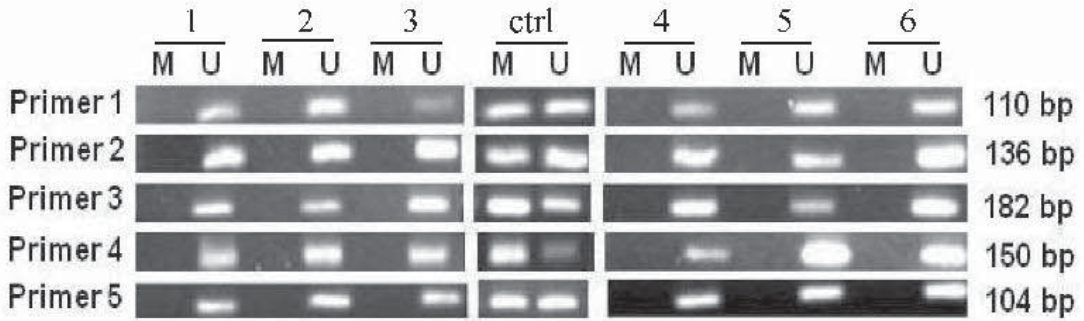
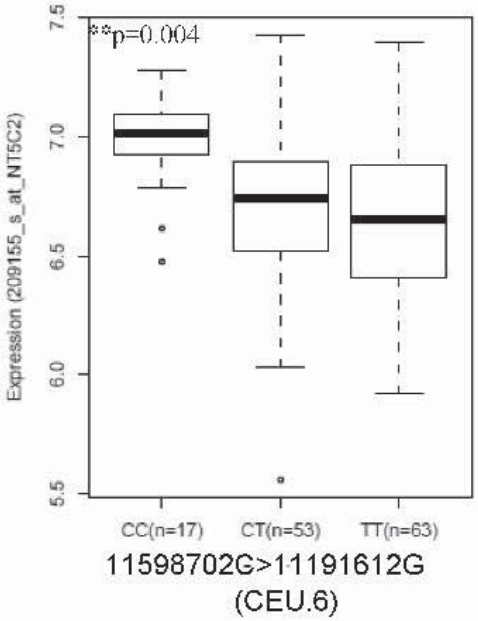
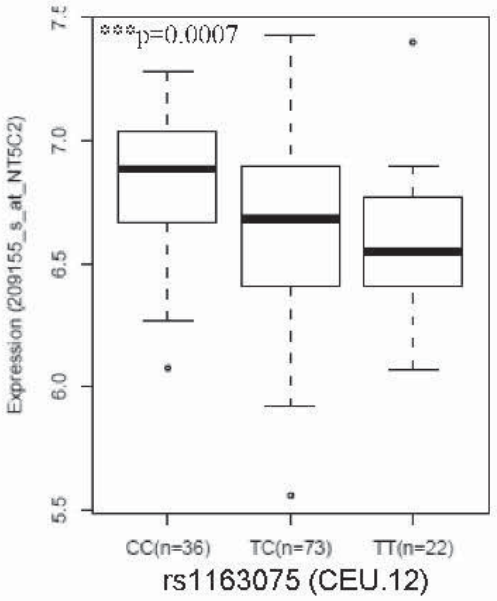
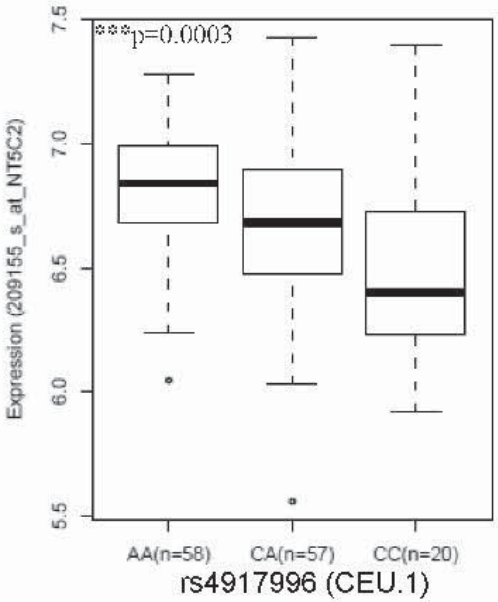


Figure 6

7A) Association of NT5C2 SNPs with its mRNA expression in diagnostic leukemia blast cells from AML patients







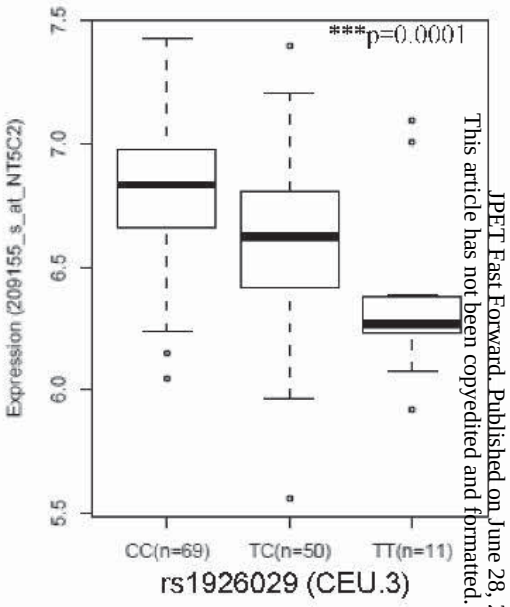


Figure 7A

7B) Association of NT5C2 SNPs with diagnostic leukemia blast cell ara-C cytotoxicity

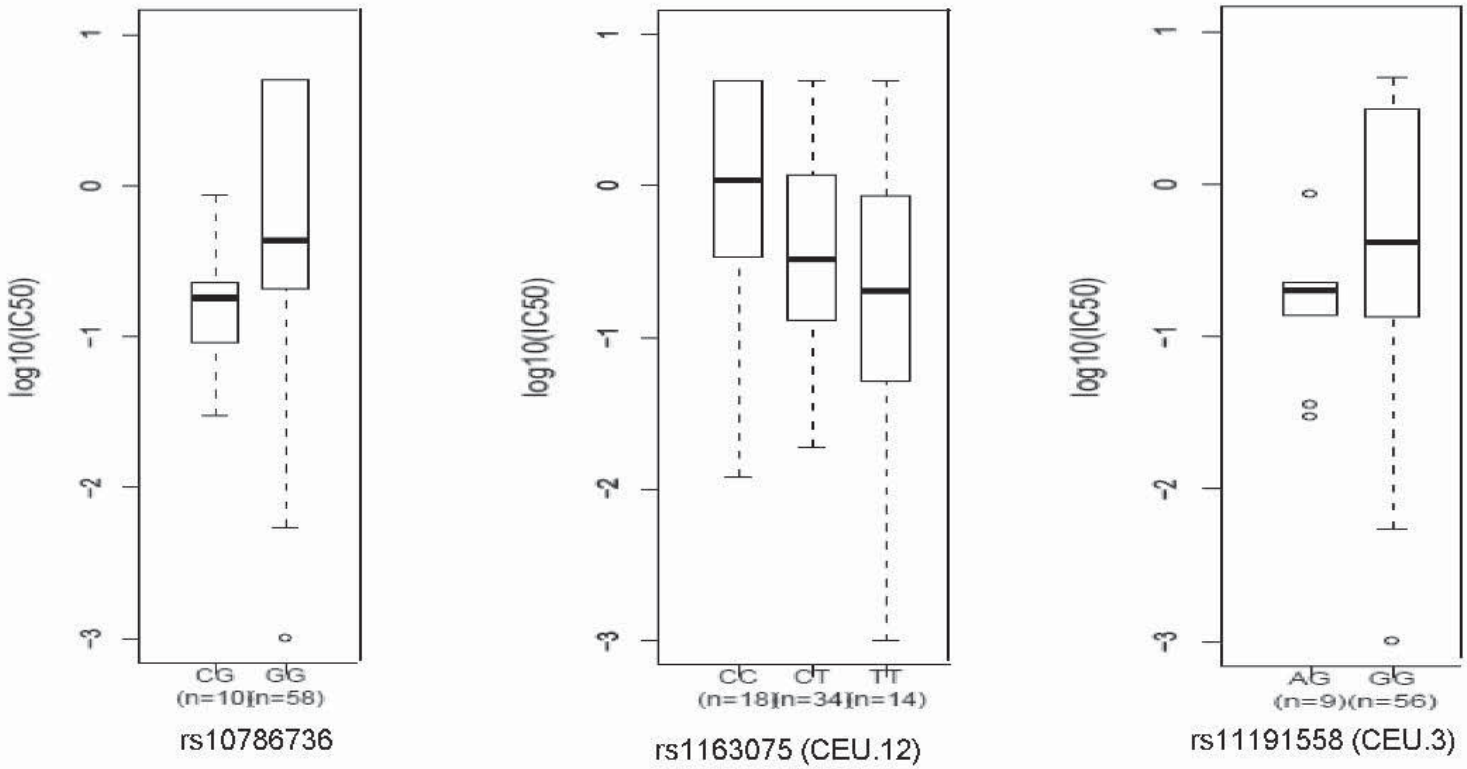


Figure 7B

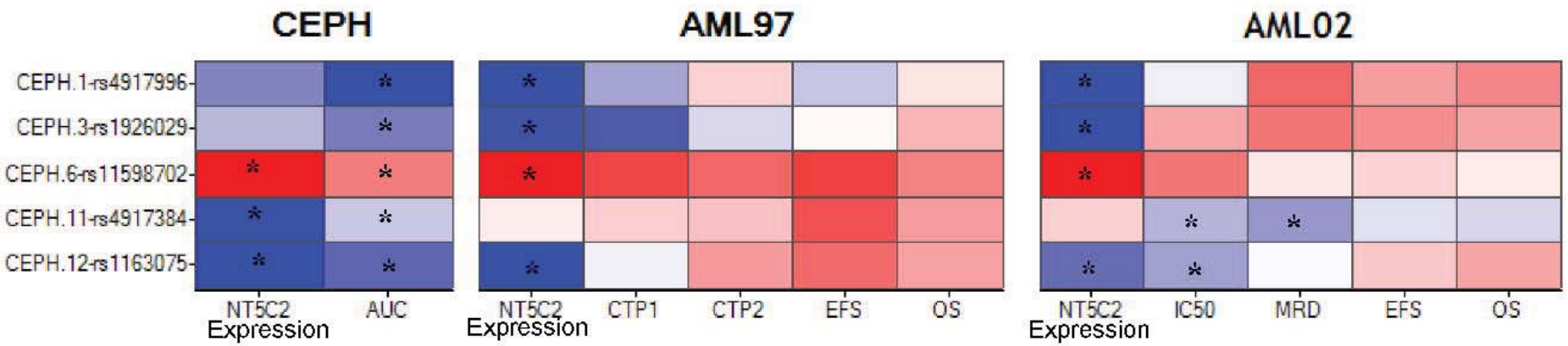


Figure 8