HHS Public Access

Author manuscript

Clin Pharmacol Ther. Author manuscript; available in PMC 2015 June 01.

Published in final edited form as:

Clin Pharmacol Ther. 2014 June ; 95(6): 644–652. doi:10.1038/clpt.2014.37.

Integrating Cell-Based and Clinical Genome-Wide Studies to Identify Genetic Variants Contributing to Treatment Failure in Neuroblastoma Patients

Navin Pinto[1,*], **Eric R. Gamazon**[2,*], **Nirav Antao**[3], **Jamie Myers**[3], **Amy L. Stark**[3], **Anuar Konkashbaev**[2], **Hae Kyung Im**[4], **Sharon J. Diskin**[5], **Wendy B. London**[6,7], **Susan M. Ludeman**[8], **John M. Maris**[5], **Nancy J. Cox**[2], **Susan L. Cohn**[1], and **M. Eileen Dolan**[3,†]

[1]Section of Pediatric Hematology/Oncology, Department of Pediatrics, The University of Chicago; Chicago, IL, 60637

[2]Section of Genetic Medicine, Department of Medicine, The University of Chicago; Chicago, IL, 60637

[3]Section of Hematology/Oncology, Department of Medicine, The University of Chicago; Chicago, IL, 60637

[4]Department of Health Studies; The University of Chicago; Chicago, IL, 60637

[5]Division of Oncology and Center for Childhood Cancer Research; The Children's Hospital of Philadelphia and Department of Pediatrics; Perelman School of Medicine at the University of Pennsylvania; Philadelphia, PA, 19104

[6]Boston Children's Hospital / Dana Farber Harvard Cancer Center; Harvard University; Boston, MA, 02215

[7]Children's Oncology Group Statistics and Data Center; Arcadia, CA, 91006

[8]Department of Basic and Social Sciences, Albany College of Pharmacy and Health Sciences, Albany, NY 12208

Abstract

High-risk neuroblastoma is an aggressive malignancy with high rates of treatment failure. We evaluated genetic variants associated with *in vitro* sensitivity to two derivatives of cyclophosphamide for association with clinical response in a separate replication cohort of neuroblastoma patients (n=2,709). Lymphoblastoid cell lines (LCLs) were exposed to increasing concentrations of 4-hydroperoxycyclophosphamide [4HC n=422] and phosphoramide mustard

†To whom correspondence should be addressed: M. Eileen Dolan, PhD, The University of Chicago, 900 E. 57th Street, Room 7100, Chicago, IL 60637, Phone – (773) 702-4441, Fax – (773) 702-9268, edolan@medicine.bsd.uchicago.edu.

*The first two authors contributed equally to this work.

Conflict of Interest

The authors have no relevant conflicts of interest to disclose pertaining to this manuscript.

Author Contributions

NJC, SLC, MED, NP, and ERG wrote manuscript. NJC, SLC, and MED designed the research. SJD, JMM, NP, ERG, NA, and JM performed research. SJD, JMM, NP, ERG, AS, AK, HKI, and WBL analyzed data. SML contributed new reagents/analytical tools.

[PM n=428] to determine sensitivity. Genome-wide association studies (GWAS) were performed to identify single nucleotide polymorphisms (SNPs) associated with 4HC and PM sensitivity. SNPs consistently associated with LCL sensitivity were analyzed for associations with event-free survival in patients. Two linked SNPs, rs9908694 and rs1453560, were found to be associated with PM sensitivity in LCLs across populations and were associated with event-free survival in all patients (P=0.01) and within the high-risk subset (P=0.05). Our study highlights the value of cell-based models to identify candidate variants that may predict response to treatment in patients with cancer.

Keywords

neuroblastoma; pharmacogenomics; cell-based models; *IKZF3*; *ZPBP2*; expression quantitative trait loci

INTRODUCTION

Neuroblastoma, a malignancy of the sympathetic nervous system, is the most common extra-cranial solid malignancy in children[1]. Clinical factors such as age at diagnosis and stage as well as biologic factors such as amplification of the *MYCN* proto-oncogene, tumor histology and tumor cell ploidy are used to classify patients as having a low-, intermediate- or high-risk of relapse and provide guidance on appropriate therapy[2]. High-risk neuroblastoma is characterized by patients ≥ 18 months of age with metastatic disease as well as by any patient with *MYCN* amplification and residual tumor after diagnosis[2]. Despite therapeutic advances such as megatherapy with autologous peripheral blood stem cell rescue[3,4] and targeted immunotherapy[5], overall survival remains poor for patients with high-risk disease. To date, no well-validated biomarkers exist to predict which high-risk patients will fail to respond to traditional high-risk therapies. Methods to identify these patients and divert them to alternative treatment strategies (thereby sparing them from ineffective treatment) are likely to improve outcomes for this devastating pediatric malignancy.

A growing body of evidence suggests that common genetic variation within the host genome (as opposed to the tumor) affects treatment outcome for patients with cancer treated with chemotherapy[6–8]. Nonetheless, pharmacogenomic studies in oncology have been impeded by the large numbers of patients needed to achieve adequate power to detect associations in a population treated with multidrug regimens. Human cell-based models, in which genetic variation, gene expression and response to a given drug (cytotoxicity, apoptosis, cell cycle arrest) can be measured in large cohorts, have emerged as an important discovery platform for genetic variants associated with chemotherapeutic sensitivity and toxicity[9–11]. The human lymphoblastoid cell line (LCL) repertoire generated by the International HapMap project[12] is a particularly useful cell-based resource because of its wealth of publicly available genotype information that can be used in genome-wide association studies (GWAS). The cell lines can be tested for genotype-phenotype associations under identical conditions without any of the confounders encountered *in vivo* and findings from the model can then serve as candidates for clinical validation.

Our group has used this model previously to identify genetic variants associated with paclitaxel-induced peripheral neuropathy[13], cisplatin response in head and neck cancer[14] as well as carboplatin response in ovarian cancer[15]. In this study, we evaluated active metabolites of the prodrug cyclophosphamide, an agent used to treat numerous malignancies, and a cornerstone of neuroblastoma treatment. 4-Hydroxycyclophosphamide is the active metabolite formed by oxidation of the parent drug and it leads to the therapeutically beneficial, DNA crosslinking agent, phosphoramide mustard (PM). We sought to determine the role of variants associated with *in vitro* sensitivity to two pre-activated metabolites of cyclophosphamide, 4-hydroperoxycyclophosphamide (4HC) and phosphoramide mustard cyclohexylamide salt (PM) in the LCL model, in both neuroblastoma risk group classification and response to therapy in a large cohort of children enrolled on a Children's Oncology Group (COG) biology study, ANBL00B1. We hypothesized that genetic variants associated with cellular resistance to cyclophosphamide would also be associated with poorer event-free survival in high-risk neuroblastoma patients.

RESULTS

Cellular Phenotyping

Figure 1A illustrates our analytic approach. LCLs were phenotyped for sensitivity to 4HC and PM. Specifically, CEU LCLs (n = 163 for 4HC; 164 for PM), YRI LCLs (176 for 4HC; 177 for PM) and ASW LCLs (n = 83) were exposed to increasing concentrations of 4HC and PM, respectively. Figure 2 illustrates the pharmacologic phenotype results for 4HC and PM by ethnic population. Comparisons between panels revealed the ASW population to be the most sensitive to the cytotoxic effects of 4HC, with a median AUC significantly lower than that in the CEU ($P = 1\times10^{-4}$) and the YRI ($P < 1\times10^{-4}$). No differences in sensitivity to 4HC were observed between the CEU and YRI populations ($P = 0.21$). No inter-population differences in sensitivity were observed for PM.

Neuroblastoma Cohort

Genotypes for 3,404 neuroblastoma patients enrolled on ANBL00B1 were utilized. After previously described[16] sample-based quality control filtering, 2,709 patients formed the analytic cohort (Table 1).

GWAS of 4HC and PM Sensitivity in CEU

Figure 1B summarizes the results of the GWAS of 4HC and PM sensitivity in the CEU population. No variants reached the genome-wide significant threshold of 5×10^{-8} for either metabolite. Nevertheless, we observed that the top SNPs were enriched for expression quantitative trait loci (eQTLs, empirical P <0.001) [17] and that limiting the association analysis to these eQTLs improves our power to detect associations. Thus, the top SNPs from this analysis (Supplementary Table 1) are enriched for SNPs previously shown to be functionally important for their association with gene expression.

Validation of CEU Variants

No variants associated with 4HC sensitivity in the CEU could be validated in both the YRI/ASW populations and the neuroblastoma cohort. On the other hand, a variant, rs9908694, within intron 5 of the gene *IKZF3* was associated with PM sensitivity in the CEU ($P = 5.62\times10^{-5}$), YRI ($P = 0.049$), ASW ($P = 0.05$) as well as high-risk neuroblastoma ($P = 9.67\times10^{-4}$, when compared to patients with low- or intermediate-risk neuroblastoma). These results are summarized in Figure 3. Supplementary Table 1 shows the top population-specific SNP associations with the PM phenotype. The SNP is a common genetic variation in all panels examined here and has a much higher MAF in the African panels (YRI [MAF = 0.31] or ASW [MAF = 0.16]) than in CEU (MAF = 0.05). Supplemental Figure 1 shows the SNP's global allele frequency distribution, as represented in the Human Genome Diversity, showing that the risk, ancestral allele T has substantially higher frequency in Sub-Saharan African panels than in the rest of the world populations examined. A chromosome 3 SNP rs9863764, intronic to glutamate receptor metabotropic 7 (GRM7), was associated with PM sensitivity in CEU (P = 8.08×10−5) and showed, at best, only very modest association with event-free survival (P = 0.09), but was not associated in YRI (P = 0.21) or ASW (P = 0.96).

Impact of rs9908694 on Neuroblastoma Event-Free Survival

In Kaplan-Meier survival analyses, rs9908694 was associated with event-free survival (EFS) in the entire neuroblastoma cohort ($P = 0.01$) as well as in the high-risk subpopulation ($P = 0.05$) with the proportion of African ancestry used as a covariate. Risk allele frequency in the neuroblastoma cohort was 5.1%. Of note, patients with the risk allele, which confers resistance to PM, had less favorable EFS (Figure 4).

Enrichment of SNPs associated with outcome among SNPs associated with PM sensitivity

We evaluated whether SNPs in the CEU population associated with PM sensitivity at the *P* value threshold less than or equal to 1×10^{-4} harbor SNPs associated with event-free survival in neuroblastoma patients and found a significant excess relative to expectation (see Methods) (Supplemental Figure 2).

Functional Annotation

Within the SCANdb database[18], rs9908694 was associated with expression of 7 transcripts in YRI LCLs (Figure 5): *ORMDL3* ($P = 1\times10^{-5}$), *DDIT3* ($P = 5\times10^{-5}$), *NEU3* ($P = 6\times10^{-5}$), *SLC28A2* ($P = 6\times10^{-5}$), *MYC* ($P = 7\times10^{-5}$), *IGLV6-57* ($P = 1\times10^{-4}$) and *ASNS* ($P = 1\times10^{-4}$). In addition, rs9908694 was associated with the expression of its host gene, *IKZF3* ($P = 0.03$).

The SNP is a strong enhancer from a Hidden Markov Model study[19] that models the combinatorial patterns of observed histone modifications and shows a ChIP-Seq peak on the basis of Histone H3 acetyl K27 (H3K27ac) as well as Histone H3 mono methyl K4 (H3K4me1) data (Broad ENCODE group) in the CEU cell line GM12878. The SNP is in perfect LD ($r^2 = 1$, in all populations CEU, YRI, and ASW) with a SNP rs1453560 (located 1 kb 5' of *ZPBP2*) that is in a DNAse hypersensitive site in a variety of cell types including several YRI cell lines (GM18507, GM19238, GM19239, GM19240) and is highly

conserved (from GERP and SiPhy scores). In particular, the SNP rs1453560 is associated with PM sensitivity in the CEU ($P = 5.62\times10^{-5}$), YRI ($P = 0.05$), and ASW ($P = 0.05$) as well as with high-risk neuroblastoma ($P = 9.67\times10^{-4}$); furthermore, rs1453560 is associated with EFS in the entire neuroblastoma cohort ($P = 0.01$) as well as in the high-risk subpopulation ($P = 0.05$).

Effect of IKZF3 siRNA knockdown

siRNA of *IKZF3* mRNA resulted in a significant decrease in host gene expression at all measured time points ($P = 4.66\times10^{-14}$, 0.0165, 6.25×10^{-13}, for 5, 29, and 53 hours, respectively). In addition, the expression of several of the predicted eQTL targets of our *IKZF3* SNP was altered by *IKZF3* siRNA. As outlined in Figure 6, significant alterations when compared to nontargeting siRNA were seen at 53 hours post-transfection for *ASNS* ($P = 0.03$) and *IGLV6-57* ($P = 4.84\times10^{-6}$) and at 29 hours for *MYC* ($P = 0.02$).

DISCUSSION

In this study, we used the International HapMap CEU LCL resource to identify variants associated with sensitivity to an activated form of cyclophosphamide and its metabolite PM. A variant associated with *in vitro* PM sensitivity (rs9908694) in CEU LCLs was replicated across 2 other LCL panels ($P \leq 0.05$). This SNP was associated with both high-risk neuroblastoma status and with EFS in cohort of 2,709 children with neuroblastoma. Furthermore, in children with high-risk neuroblastoma, those with the rs9908694 risk allele had poorer EFS than patients homozygous for the non-risk allele, suggesting the SNP's impact on survival is independent of neuroblastoma risk group classification. This variant was associated with the expression of the host gene and with seven other transcripts in LCLs[18]. Three of these seven genes were significantly altered by siRNA knockdown of the SNP's host gene, lending further evidence to the functional significance of *IKZF3* variation on downstream genes.

Cyclophosphamide is an integral component of both intermediate- and high-risk neuroblastoma therapies as well as in the treatment of many other cancers. Cyclophosphamide is a prodrug converted by a number of P450 enzymes to 4-hydroxycyclophosphamide[20]. 4-Hydroxycyclophosphamide spontaneously (non-enzymatically) leads to the formation of acrolein, a urotoxic metabolite with no known antitumor properties, and the active alkylating compound phosphoramide mustard[21,22]. 4-Hydroperoxycyclophosphamide (4HC) is a synthetic, pre-activated analog of cyclophosphamide, which under aqueous conditions, spontaneously converts to 4-hydroxycyclophosphamide. This metabolite then provides for the formation of phosphoramide mustard (PM) through the usual metabolic pathway of cyclophosphamide. Since 4HC, like cyclophosphamide, produces PM and acrolein simultaneously, we chose to include authentic PM in our experiments. It was expected that this would allow for an understanding of the germline genetic variants associated with the toxicity and efficacy of the parent cyclophosphamide (via 4HC) distinct from those associated with the ultimate metabolite of therapeutic consequence (PM). In addition, genetic variants within enzymes responsible for the bioactivation, metabolism and clearance of cyclophosphamide were not

able to explain interindividual variability in the clearance of cyclophosphamide or 4-hydroxycyclophosphamide, suggesting that variants outside of the established pharmacokinetic pathway may be responsible for this interindividual variability[23]. Our inability to validate variants associated with *in vitro* 4HC sensitivity within the neuroblastoma survival analysis may be due to the fact that this metabolite generates both toxic and therapeutic compounds whereas PM only contributes to the therapeutic effect of cyclophosphamide and is therefore a purer representation of response to therapy.

The SNP rs9908694 lies within intron 5 of *IKZF3* on chromosome 17q21, and unbalanced gains of this chromosomal region are found frequently in clinically aggressive neuroblastoma[24]. *IKZF3* is one of five genes in the Ikaros family of zinc finger DNA binding transcription factors important in the regulation of lymphocyte differentiation[25]. Although the host gene (*IKZF3*) for one of our predictive SNPs has not been previously implicated in either neuroblastoma pathogenesis or cyclophosphamide response, the expression of the host gene and several other transcripts appears to be affected by the presence of the (eQTL) rs9908694 risk allele, which may provide insight into the mechanism of this SNP's negative impact on event-free survival seen in our cohort. Interestingly, three of the seven genes targeted by rs9908694 – *DDIT3*, *ORMDL3* and *ASNS* – are thought to be involved in the endoplasmic reticulum stress response[26–28], a process that is often dysregulated in human cancers including neuroblastoma leading to inherent resistance to cell death[29,30]. The SNP is in perfect LD, in all populations examined here, with a SNP rs1453560 (1 kb upstream of *ZPBP2*) that is in a DNAse hypersensitive site, which may provide future insights into the role of this locus in treatment failure.

Integrative epidemiology studies have previously linked disease susceptibility variants to those associated with treatment response[31], and our findings may serve as an additional example of this concept. The *IKZF3-ZPBP2* locus (importantly, the SNP rs1453560 from our study) has been implicated, in a recent large-scale multiethnic study, as a susceptibility locus for systemic lupus erythematosus[32], which is often treated with cyclophosphamide. Furthermore, the SNP rs9908694 has been found to be nominally associated with glycosylated hemoglobin in a genome-wide study in a British population[33]. Aberrant glycosylation occurs in human cancers[34] and has been implicated in studies of neuroblastoma[35].

High-risk neuroblastomas are clinically aggressive tumors that account for a disproportionate amount of deaths due to pediatric cancer[1]. Identifying those patients at risk for treatment failure may lead to improvements in treatment delivery for patients with cancer with little chance of cure. However, identifying and avoiding the component(s) responsible for treatment failure in multi-modality regimen can prove difficult. The methods outlined here support that polymorphisms in the *IKZF3-ZPBP2* locus contribute to neuroblastoma treatment failure likely through resistance to cyclophosphamide and its metabolites.

Our study highlights the value of cell-based models to identify candidate variants that may predict response to treatment in patients with cancer. Drawbacks of this approach include the weight of validation in other ethnic populations with markedly different linkage disequilibrium patterns. Another drawback is that patients on this study were enrolled on a

biobanking protocol separately from an available treatment protocol and most, but not all patients enrolled on the protocol were treated on a Children's Oncology Group clinical trial. Our approach could be easily adapted to other cytotoxic chemotherapies for a variety of malignancies. Clinical validation of the most highly significant SNPs from cell-based models may provide important insights into the genetic architecture of human response to chemotherapy, and may inform the interpretation of results from past clinical studies and the design of future studies[48]. Prospective validation of the variant we identified and its impact on survival will be needed before clinical implementation.

METHODS

Drugs and Lymphoblastoid cell lines

Cyclophosphamide is a prodrug requiring activation to 4-hydroxycyclophosphamide by the cytochrome P450 system[20]; however, LCLs show low expression of the P450 enzymes that are required for this oxidation. 4-Hydroperoxycyclophosphamide (4HC) serves as a synthetic equivalent of 4-hydroxycyclophosphamide; under aqueous conditions, it spontaneously converts to 4-hydroxycyclophosphamide and, ultimately, phosphoramide mustard (PM)[20] . The synthesis of 4HC has been described[36]; PM, as the cyclohexylammonium salt, was a gift from the Drug Synthesis and Chemistry Branch, Division of Cancer Treatment, National Cancer Institute.

HapMap LCLs derived from Caucasian individuals (n= 163 for 4HC and 164 for PM) from Utah in the United States with northern/western European Ancestry (CEU), Yoruba individuals (n= 176 for 4HC and 177 for PM) from Ibadan, Nigeria (YRI) and African-American individuals (n= 83 for both 4HC and PM) from the Southwestern United States (ASW), were used for drug sensitivity phenotyping. LCLs were purchased from the Coriell Institute for Medical Research (Camden, NJ) and cultured in RPMI 1640 media (Cellgro, Herndon, VA) containing 15% heat-inactivated bovine growth serum (Hyclone, Logan, UT) and 20 mM L-glutamine.

Cellular Phenotyping

Cell lines were diluted three times per week at a concentration of 300,000 to 350,000 cells/mL and incubated at 37°C with 5% CO_2/95% humidified air. Cell growth inhibition for a given cell line was evaluated both at concentrations of 0, 1, 5, 10, 20 and 40 μM/L for 4HC and 0, 5, 10, 25, 50, 100 and 200 μMl/L of PM in separate experiments. Aqueous solutions of 4HC and PM are unstable[37]; thus, they were made fresh, as needed, kept cool (ice bath) and used as quickly as possible (maximum lag time between dissolution and use: 30 min). Cell lines in the exponential growth phase with >85% viability (Vi-Cell XR viability analyzer, Beckman Coulter, Fullerton, CA) were plated in triplicate in 96-well plates (Corning, Corning, NY) 24 hours before drug exposure. LCLs were exposed to drug suspended in PBS (Life Technologies, Carlsbad, CA) for 72 hours. Cell growth inhibition was measured by the Alamar Blue® assay (Life Technologies) at wavelengths of 570 and 600 nm (Synergy-HT multidetection plate reader, BioTek, Winooski, VT). The area under the curve (AUC) was determined for each LCL through the trapezoidal rule and $\log_2$-

transformed for further data analysis. Cell lines were excluded from analysis for poor growth in culture or failure to achieve a reproducible cellular sensitivity to drug.

GWAS of cellular sensitivity to 4HC and PM

CEU genotypes (n= 180) were downloaded from the HapMap Consortium release 27. Log_2-transformed AUC values were used as phenotype. A GWAS was performed in CEU LCLs on 4HC and PM sensitivity (separately) (figure 1a). Greater than 2 million SNPs (MAF >5%, no Mendelian errors, and in Hardy-Weinberg equilibrium) was tested for association with either 4HC or PM.

We conducted GWAS of 4HC and PM sensitivity (separately) in the two African panels, YRI and ASW. Since ASW contains samples of recent admixture, we used the proportion of African ancestry, derived from Principal Component Analysis[38], in these samples as a covariate in the association analysis.

Those SNPs that achieved a nominal level of significance in the CEU (defined as $P \leq 1\times10^{-4}$) were tested for association in the YRI ($P < 0.05$ and concordant direction). Those SNPs that passed this "validation" step were also tested for association in the ASW population. SNPs that survived this multi-step validation analysis in independent sets of LCLs were then taken forward for validation in the clinical dataset.

Neuroblastoma patient cohort and quality control of patient data

After institutional review board approval by participating institutions and informed consent were obtained, children diagnosed with neuroblastoma, ganglioneuroblastoma, or ganglioneuroma (maturing type) were enrolled on COG ANBL00B1 (NCT00904241) between 2001 and 2009. ANBL00B1 was a neuroblastoma biology registry, and selection of therapy was at the discretion of treating physicians and families. The majority of patients would have also been enrolled on a therapeutic clinical trial through the COG. Patients with available genotype, risk group classification and outcome data formed the analytic cohort. Methods to confirm the diagnosis, assignment of stage, analysis of tumor biology (ploidy, *MYCN* amplification and histology), and assignment of self-reported race have been previously described[39].

DNA samples from patients enrolled on COG ANBL00B1 were genotyped using three Illumina platforms: 550v1, 550v3, and Human610-Quad, as previously described[40–43]. Included in this analysis are those SNPs which were genotyped on all three platforms, had call rates > 95% and had minor allele frequencies > 5%.

The sample-based quality control pipeline for this cohort has previously been described[16]. Briefly, extensive QC analyses were performed on the genotype data, including detection of sex incompatibilities, mis-specified relationships, and duplications. Call rates were estimated by individual and by SNP in order to determine average heterozygosity (across all SNPs) for each sample and to evaluate genotype distributions for each SNP in relation to expected Hardy-Weinberg equilibrium (HWE)[40–43].

Association of PM sensitivity variants with high-risk neuroblastoma phenotype and event-free survival

SNPs that achieved significance in the CEU population and were validated in the YRI and ASW populations were tested for association ($P < 0.05$) with the high-risk neuroblastoma phenotype (Figure 1). The risk allele in the LCL model was evaluated for its (additive) effect on high-risk disease in patients. As the patient cohort includes both Caucasians and African Americans, we used the proportion of African ancestry in the association analysis, as previously described[16]. SNPs that achieved significance in both the cell-based model and the neuroblastoma cohort were tested for their impact on event-free survival among all patients in the analytic cohort as well as in the subset of high-risk patients alone. For each such SNP, log-rank comparison of event-free survival by SNP genotype was performed.

Enrichment Analyses

We conducted a permutation resampling analysis to test for an enrichment of 4HC and PM cytotoxicity-associated SNPs identified in the LCL model among SNPs associated with the high-risk phenotype in the neuroblastoma cohort. Using a previously described method[44], the risk group classification (high-risk vs non-high-risk phenotype) was randomly shuffled while keeping the genotype data fixed thereby preserving linkage disequilibrium. Based on this permuted replicate, logistic regression analyses were conducted for all of the SNPs. This process was repeated 1000 times. For each of the 1000 replicates, the number of SNPs with $P \leq 0.05$ in the patient data and $P \leq 0.0001$ in the LCL data, with concordant direction of effect was calculated. The distribution of the overlap count from the permutations was compared to the actual overlap count to generate an empirical P value. The significance of the excess of high-risk neuroblastoma associated SNPs among cellular sensitivity associated SNPs was calculated as the proportion of permutations in which the number of actual overlap SNPs matches or exceeds the observed overlap SNP count. We considered the most highly ranked SNPs in the LCL data ($P \leq 0.0001$) and generated a quantile-quantile plot from their association with high-risk neuroblastoma.

Impact of validated variant on LCL gene expression

Baseline gene expression was evaluated in 89 YRI and 87 CEU LCLs using the Affymetrix GeneChip Human Exon 1.0ST Array (Affymetrix, Santa Clara, CA) as previously described[45]. SNPs validated across the cell-based platform and the clinical trial were tested for effect on gene expression within the LCL model using data from the SCANdb database (http://www.scandb.org)[18].

In an effort to further functionally characterize the loci that harbor SNPs associated with drug sensitivity in the cell-based model and with event-free survival in patients, we downloaded and utilized the public ENCODE data assayed in a variety of cell types (including the CEU cell line GM12878 and the YRI cell lines GM18507, GM19238, GM19239, and GM19240) to determine the chromatin states of potentially regulatory regions as well as identify DNAse hypersensitivity sites[19].

siRNA knockdown of IKZF3

LCLs GM18856 and GM19093 were counted (Vi-Cell, Beckman Coulter, Fullerton, CA, USA) and diluted to 5×10^5 cells/ml 24 hours prior to nucleofection. Cells (viability >85%) were nucleofected using Lonza's Cell Line 96-well Nucleofector Kit SF (Lonza Inc, Basel, Switzerland). The reaction was performed according to the manufacturer's protocol using 1×10^6 cells per well and 2000 nM final concentration of AllStars Negative Control siRNA (Qiagen, Valencia, CA, USA) or a pool of IKZF3 FlexiTube siRNA (Hs_IKZF3_1, Hs_IKZF3_2, Hs_IKZF3_3 and Hs_ZNFN1A3_5)(Qiagen, Valencia, CA, USA). The cells were nucleofected using the DN-100 program on an Amaxa Nucleofector 96-well Shuttle (Lonza Inc, Basel, Switzerland). After a 10 minute rest, 85uL pre-warmed media was added to each transfected well and cells were allowed to rest for an additional 5 minutes. Cells were plated for mRNA pelleting in 24-well flat bottom plates. Cells were harvested at 5, 29 and 53 hours after nucleofection, washed in ice-cold PBS and centrifuged to remove PBS. All pellets were stored at −80°C until RNA isolation.

Quantitative real-time PCR

See Supplemental Methods.

Statistical Analysis of IKZF3 Knockdown

Significance across multiple cell lines for the effect of IKZF3 knockdown on the target genes' expression was evaluated using a mixed effects model combining the two cell lines and using cell line id as random effect. Log2 relative expression for a given time and gene was tested whether different from zero. The R statistical software and the lme4 package were used to fit the mixed effects model[46,47].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Research reported in this publication was supported by the National Center For Advancing Translational Sciences of the National Institutes of Health under Award Number UL1TR000430, NIH/NIGMS Pharmacogenomics of Anticancer Agents Grant U01GM61393, by the University of Chicago Breast Cancer SPORE Grant NCI P50 CA125183, NIH/NCI RO1 CA136765 and NIH/NCI R01 CA 016783. In addition, this research is supported by the St. Baldrick's Foundation and the Cancer Research Foundation. The authors are grateful to the Pharmacogenomics of Anticancer Agents Cell Line Core at the University of Chicago for assistance in maintaining the cell lines. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors would like to thank Shannon Delaney for her help in formatting the *IKZF3* siRNA results.

REFERENCES

1. Maris JM, Hogarty MD, Bagatell R, Cohn SL. Neuroblastoma. Lancet. 2007; 369:2106–2120. [PubMed: 17586306]
2. Cohn SL, et al. The International Neuroblastoma Risk Group (INRG) classification system: an INRG Task Force report. J Clin Oncol. 2009; 27:289–297. [PubMed: 19047291]
3. Matthay KK, et al. Long-term results for children with high-risk neuroblastoma treated on a randomized trial of myeloablative therapy followed by 13-cis-retinoic acid: a children's oncology group study. J Clin Oncol. 2009; 27:1007–1013. [PubMed: 19171716]

4. Berthold F, et al. Myeloablative megatherapy with autologous stem-cell rescue versus oral maintenance chemotherapy as consolidation treatment in patients with high-risk neuroblastoma: a randomised controlled trial. Lancet Oncol. 2005; 6:649–658. [PubMed: 16129365]
5. Yu AL, et al. Anti-GD2 antibody with GM-CSF, interleukin-2, and isotretinoin for neuroblastoma. N Engl J Med. 2010; 363:1324–1334. [PubMed: 20879881]
6. Pinto N, Cohn SL, Dolan ME. Using germline genomics to individualize pediatric cancer treatments. Clin Cancer Res. 2012; 18:2791–2800. [PubMed: 22589487]
7. O'Donnell PH, Ratain MJ. Germline pharmacogenomics in oncology: decoding the patient for targeting therapy. Mol Oncol. 2012; 6:251–259. [PubMed: 22321460]
8. Gonzalez-Angulo AM, Hennessy BT, Mills GB. Future of personalized medicine in oncology: a systems biology approach. J Clin Oncol. 2010; 28:2777–2783. [PubMed: 20406928]
9. Zhang W, Huang RS, Dolan ME. Cell-based Models for Discovery of Pharmacogenomic Markers of Anticancer Agent Toxicity. Trends Cancer Res. 2008; 4:1–13. [PubMed: 21499559]
10. Welsh M, et al. Pharmacogenomic discovery using cell-based models. Pharmacol Rev. 2009; 61:413–429. [PubMed: 20038569]
11. Cox NJ, Gamazon ER, Wheeler HE, Dolan ME. Clinical translation of cell-based pharmacogenomic discovery. Clin Pharmacol Ther. 2012; 92:425–427. [PubMed: 22910437]
12. The International HapMap Project. Nature. 2003; 426:789–796. [PubMed: 14685227]
13. Wheeler HE, et al. Integration of cell line and clinical trial genome-wide analyses supports a polygenic architecture of paclitaxel-induced sensory peripheral neuropathy. Clin Cancer Res. 2012
14. Ziliak D, et al. Germline polymorphisms discovered via a cell-based, genome-wide approach predict platinum response in head and neck cancers. Transl Res. 2011; 157:265–272. [PubMed: 21497773]
15. Huang RS, et al. Platinum sensitivity-related germline polymorphism discovered via a cell-based approach and analysis of its association with outcome in ovarian cancer patients. Clin Cancer Res. 2011; 17:5490–5500. [PubMed: 21705454]
16. Gamazon ER, et al. Trans-population Analysis of Genetic Mechanisms of Ethnic Disparities in Neuroblastoma Survival. J Natl Cancer Inst. 2012
17. Gamazon ER, Huang RS, Cox NJ, Dolan ME. Chemotherapeutic drug susceptibility associated SNPs are enriched in expression quantitative trait loci. Proc Natl Acad Sci U S A. 2010; 107:9287–9292. [PubMed: 20442332]
18. Gamazon ER, et al. SCAN: SNP and copy number annotation. Bioinformatics. 2010; 26:259–262. [PubMed: 19933162]
19. Ernst J, Kellis M. Discovery and characterization of chromatin states for systematic annotation of the human genome. Nat Biotechnol. 2010; 28:817–825. [PubMed: 20657582]
20. Ludeman SM. The chemistry of the metabolites of cyclophosphamide. Curr Pharm Des. 1999; 5:627–643. [PubMed: 10469895]
21. Cox PJ. Cyclophosphamide cystitis--identification of acrolein as the causative agent. Biochem Pharmacol. 1979; 28:2045–2049. [PubMed: 475846]
22. Fleming RA. An overview of cyclophosphamide and ifosfamide pharmacology. Pharmacotherapy. 1997; 17:146S–154S. [PubMed: 9322882]
23. Ekhart C, et al. Influence of polymorphisms of drug metabolizing enzymes (CYP2B6, CYP2C9, CYP2C19, CYP3A4, CYP3A5, GSTA1, GSTP1, ALDH1A1 and ALDH3A1) on the pharmacokinetics of cyclophosphamide and 4-hydroxycyclophosphamide. Pharmacogenet Genomics. 2008; 18:515–523. [PubMed: 18496131]
24. Bown N, et al. 17q gain in neuroblastoma predicts adverse clinical outcome. U.K. Cancer Cytogenetics Group and the U.K. Children's Cancer Study Group. Med Pediatr Oncol. 2001; 36:14–19. [PubMed: 11464868]
25. Morgan B, et al. Aiolos, a lymphoid restricted transcription factor that interacts with Ikaros to regulate lymphocyte differentiation. Embo J. 1997; 16:2004–2013. [PubMed: 9155026]
26. Oyadomari S, Mori M. Roles of CHOP/GADD153 in endoplasmic reticulum stress. Cell Death Differ. 2004; 11:381–389. [PubMed: 14685163]

27. Cantero-Recasens G, Fandos C, Rubio-Moscardo F, Valverde MA, Vicente R. The asthma-associated ORMDL3 gene product regulates endoplasmic reticulum-mediated calcium signaling and cellular stress. Hum Mol Genet. 2010; 19:111–121. [PubMed: 19819884]
28. Balasubramanian MN, Butterworth EA, Kilberg MS. Asparagine synthetase: regulation by cell stress and involvement in tumor biology. Am J Physiol Endocrinol Metab. 2013; 304:E789–E799. [PubMed: 23403946]
29. Schleicher SM, Moretti L, Varki V, Lu B. Progress in the unraveling of the endoplasmic reticulum stress/autophagy pathway and cancer: implications for future therapeutic approaches. Drug Resist Updat. 2010; 13:79–86. [PubMed: 20471904]
30. Rossi M, Sayan AE, Terrinoni A, Melino G, Knight RA. Mechanism of induction of apoptosis by p73 and its relevance to neuroblastoma biology. Ann N Y Acad Sci. 2004; 1028:143–149. [PubMed: 15650240]
31. Spitz MR, Wu X, Mills G. Integrative epidemiology: from risk assessment to outcome prediction. J Clin Oncol. 2005; 23:267–275. [PubMed: 15637390]
32. Lessard CJ, et al. Identification of IRF8, TMEM39A, and IKZF3-ZPBP2 as susceptibility loci for systemic lupus erythematosus in a large-scale multiracial replication study. Am J Hum Genet. 2012; 90:648–660. [PubMed: 22464253]
33. Strachan DP, et al. Lifecourse influences on health among British adults: effects of region of residence in childhood and adulthood. Int J Epidemiol. 2007; 36:522–531. [PubMed: 17255346]
34. Hakomori S. Glycosylation defining cancer malignancy: new wine in an old bottle. Proc Natl Acad Sci U S A. 2002; 99:10231–10233. [PubMed: 12149519]
35. Long PM, Stradecki HM, Minturn JE, Wesley UV, Jaworski DM. Differential aminoacylase expression in neuroblastoma. Int J Cancer. 2011; 129:1322–1330. [PubMed: 21128244]
36. Zon G, et al. NMR spectroscopic studies of intermediary metabolites of cyclophosphamide. A comprehensive kinetic analysis of the interconversion of cis- and trans-4-hydroxycyclophosphamide with aldophosphamide and the concomitant partitioning of aldophosphamide between irreversible fragmentation and reversible conjugation pathways. J Med Chem. 1984; 27:466–485. [PubMed: 6708049]
37. Flowers JL, et al. Evidence for a role of chloroethylaziridine in the cytotoxicity of cyclophosphamide. Cancer Chemother Pharmacol. 2000; 45:335–344. [PubMed: 10755323]
38. Price AL, et al. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006; 38:904–909. [PubMed: 16862161]
39. Henderson TO, et al. Racial and ethnic disparities in risk and survival in children with neuroblastoma: a Children's Oncology Group study. J Clin Oncol. 2011; 29:76–82. [PubMed: 21098321]
40. Maris JM, et al. Chromosome 6p22 locus associated with clinically aggressive neuroblastoma. N Engl J Med. 2008; 358:2585–2593. [PubMed: 18463370]
41. Capasso M, et al. Common variations in BARD1 influence susceptibility to high-risk neuroblastoma. Nat Genet. 2009; 41:718–723. [PubMed: 19412175]
42. Wang K, et al. Integrative genomics identifies LMO1 as a neuroblastoma oncogene. Nature. 2011; 469:216–220. [PubMed: 21124317]
43. Diskin SJ, et al. Common variation at 6q16 within HACE1 and LIN28B influences susceptibility to neuroblastoma. Nat Genet. 2012; 44:1126–1130. [PubMed: 22941191]
44. Shterev ID, Jung SH, George SL, Owzar K. permGPU: Using graphics processing units in RNA microarray association studies. BMC Bioinformatics. 2010; 11:329. [PubMed: 20553619]
45. Huang RS, et al. Identification of genetic variants contributing to cisplatin-induced cytotoxicity by use of a genomewide approach. Am J Hum Genet. 2007; 81:427–437. [PubMed: 17701890]
46. R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2010. http://www.R-project.org/.
47. Bates, Douglas; Maechler, Martin; Bolker, Ben. lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-0. 2012. http://CRAN.R-project.org/package=lme4
48. Cox NJ, Gamazon ER, Wheeler HE, Dolan ME. Clinical translation of cell-based pharmacogenomic discovery. Clin Pharmacol Ther. 2012; 92:425–427. [PubMed: 22910437]

Study Highlights

What is the current knowledge on the topic?

- High-risk neuroblastomas are clinically aggressive tumors that account for a disproportionate amount of deaths.
- Identifying and avoiding the chemotherapeutic(s) that are ineffective in a multi-modality regimen can prove difficult.

What question did the study address?

- We aimed to identify the genetic variants associated with cyclophosphamide sensitivity in a human cell-based model and evaluate whether those variants are also associated with event-free survival in a clinical study that included nearly 3,000 children with neuroblastoma.

What this study adds to our knowledge

- Polymorphisms in the *IKZF3-ZPBP2* locus contribute to neuroblastoma treatment failure likely through resistance to phosphoramide mustard, the metabolite of cyclophosphamide thought to be responsible for its antitumor effect.

How this might change clinical pharmacology and therapeutics

- Our study highlights the value of cell-based models to complement clinical trials aimed at identifying and prioritizing genetic variants associated with non-response in rare diseases such as neuroblastoma for which large scale GWAS of survival are difficult to replicate. Our approach could easily be adapted to other chemotherapies for a variety of malignancies.

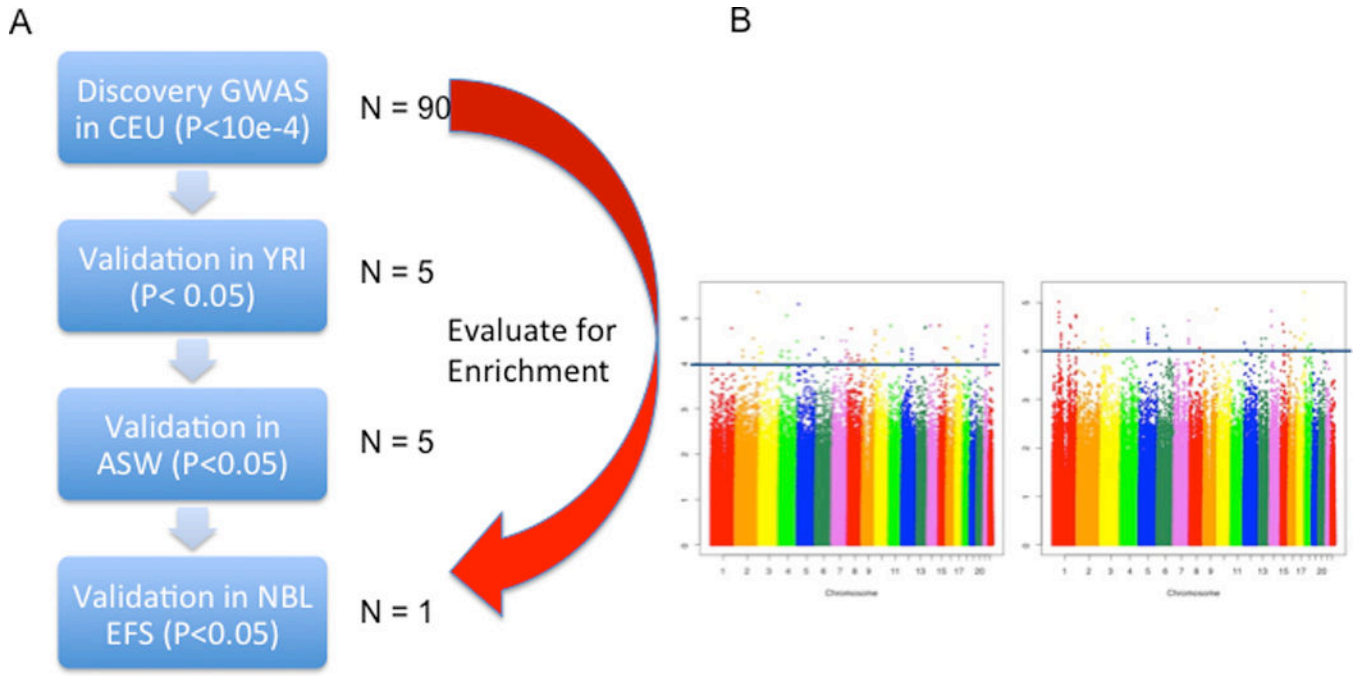

Figure 1.

(**A)** Analytic approach. SNPs with suggestive significance in the CEU LCL population were validated in both the YRI and ASW LCL populations before final validation in patients with neuroblastoma. **(B)** Genome-wide association results of cellular sensitivity to 4HC and PM in the CEU. The Manhattan plot shows the results of GWAS of log transformed AUC to 4HC and PM (n=180 individual samples).

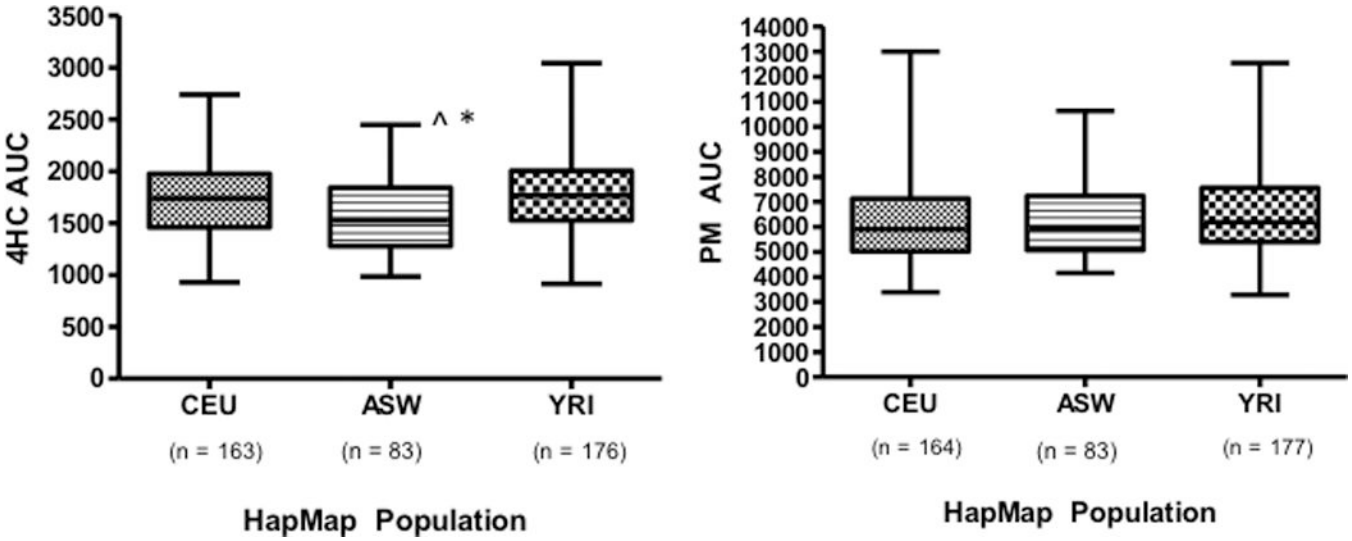

Figure 2.

Inter-population differences in sensitivity to 4HC and PM. The ASW population was more sensitive than either the CEU (^ $P = 0.0001$) or the YRI population (*$P = 0.0001$). There were no differences in sensitivities for PM between the three HapMap panels.

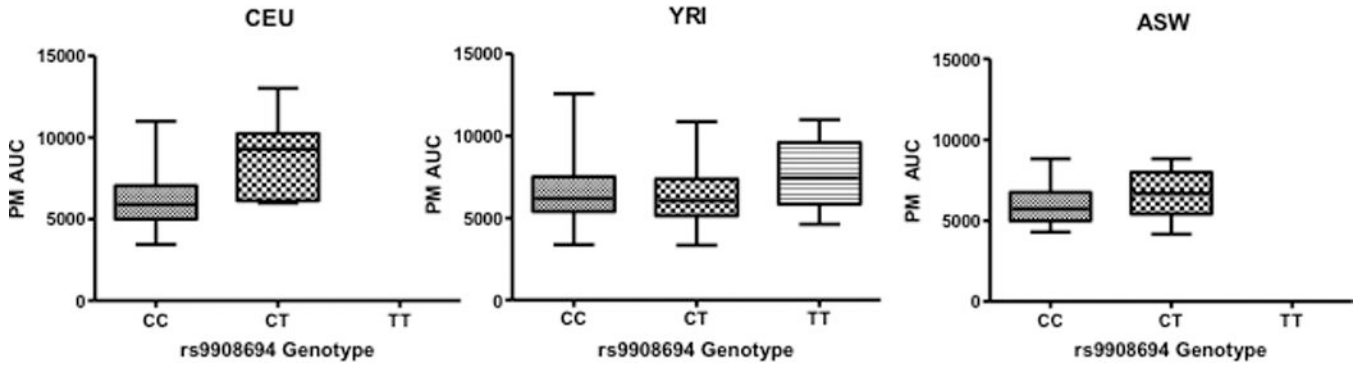

Figure 3.

PM sensitivity by rs9908694 genotype. The T allele of rs9908694, within intron 5 of the gene *IKZF3* is associated with PM sensitivity in the CEU ($P = 5.62\times10^{-5}$), YRI ($P = 0.049$), ASW ($P = 0.05$) population of LCLs.

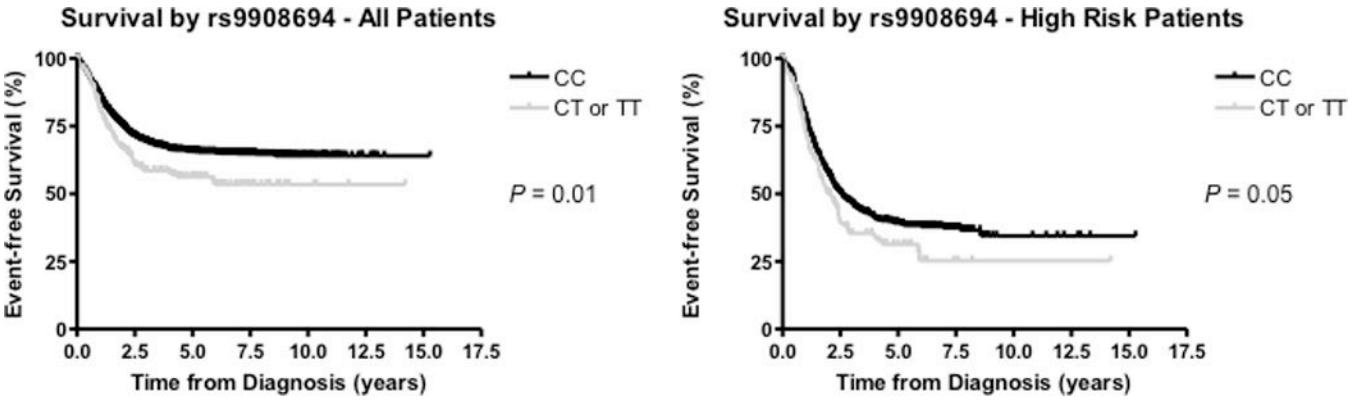

Figure 4.

Event-free survival by rs9908694 genotype. The T allele, which confers resistance to PM in the LCL model, is associated with inferior event-free survival in the entire neuroblastoma cohort ($P = 0.01$) and within the subset of high-risk neuroblastoma patients ($P = 0.05$).

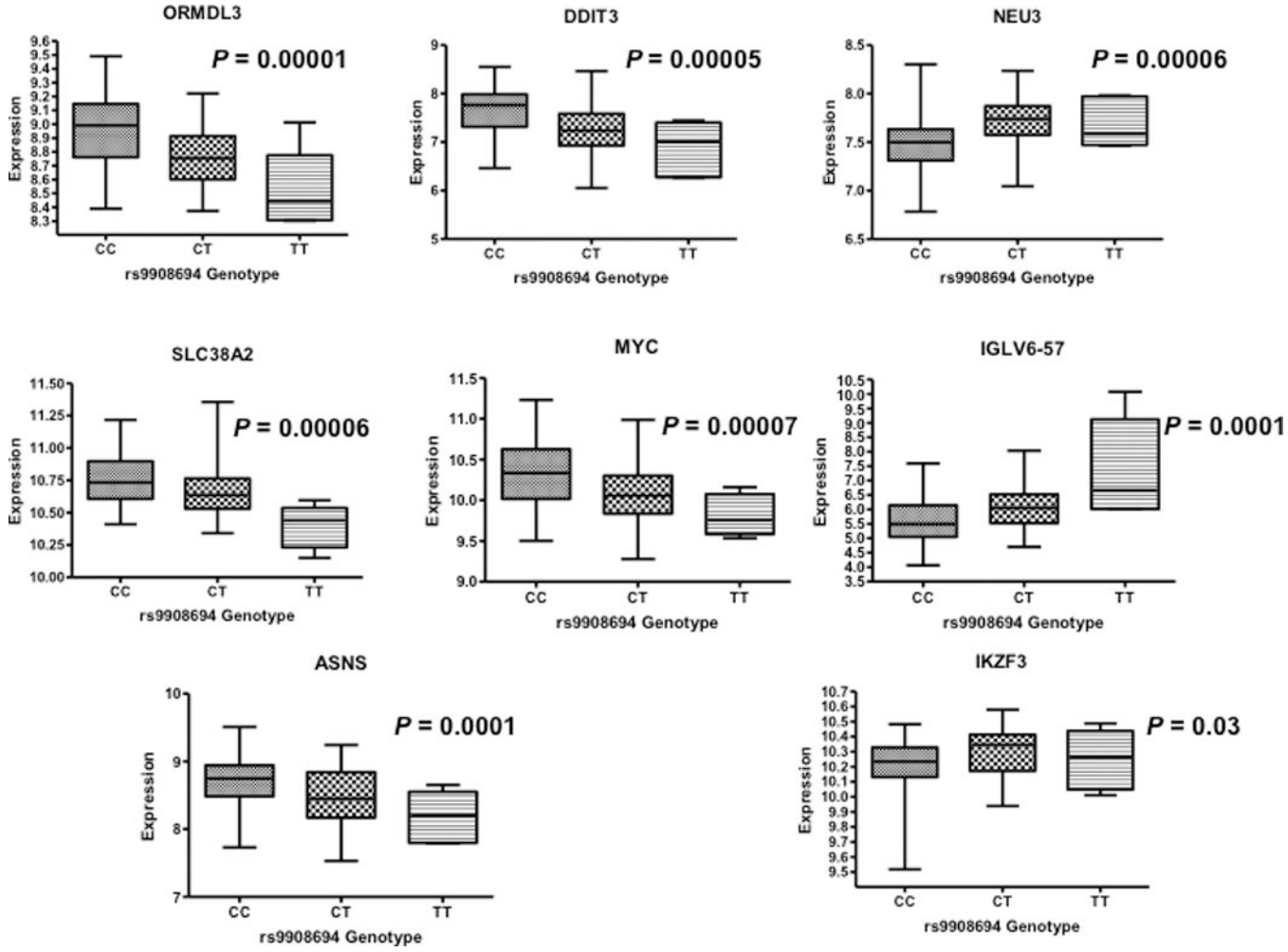

Figure 5.

Association between rs9908694 genotype and expression measured on the Affymetrix GeneChip Human Exon 1.0ST Array in LCLs of YRI ancestry. The genotypes are plotted against baseline gene expression of *ORMDL3* ($P = 1\times10^{-5}$), *DDIT3* ($P = 5\times10^{-5}$), *NEU3* ($P = 6\times10^{-5}$), *SLC28A2* ($P = 6\times10^{-5}$), *MYC* ($P = 7\times10^{-5}$), *IGLV6-57* ($P = 1\times10^{-4}$) and *ASNS* ($P = 1\times10^{-4}$). In addition, rs9908694 was associated with the expression of its host gene, *IKZF3* ($P = 0.03$).

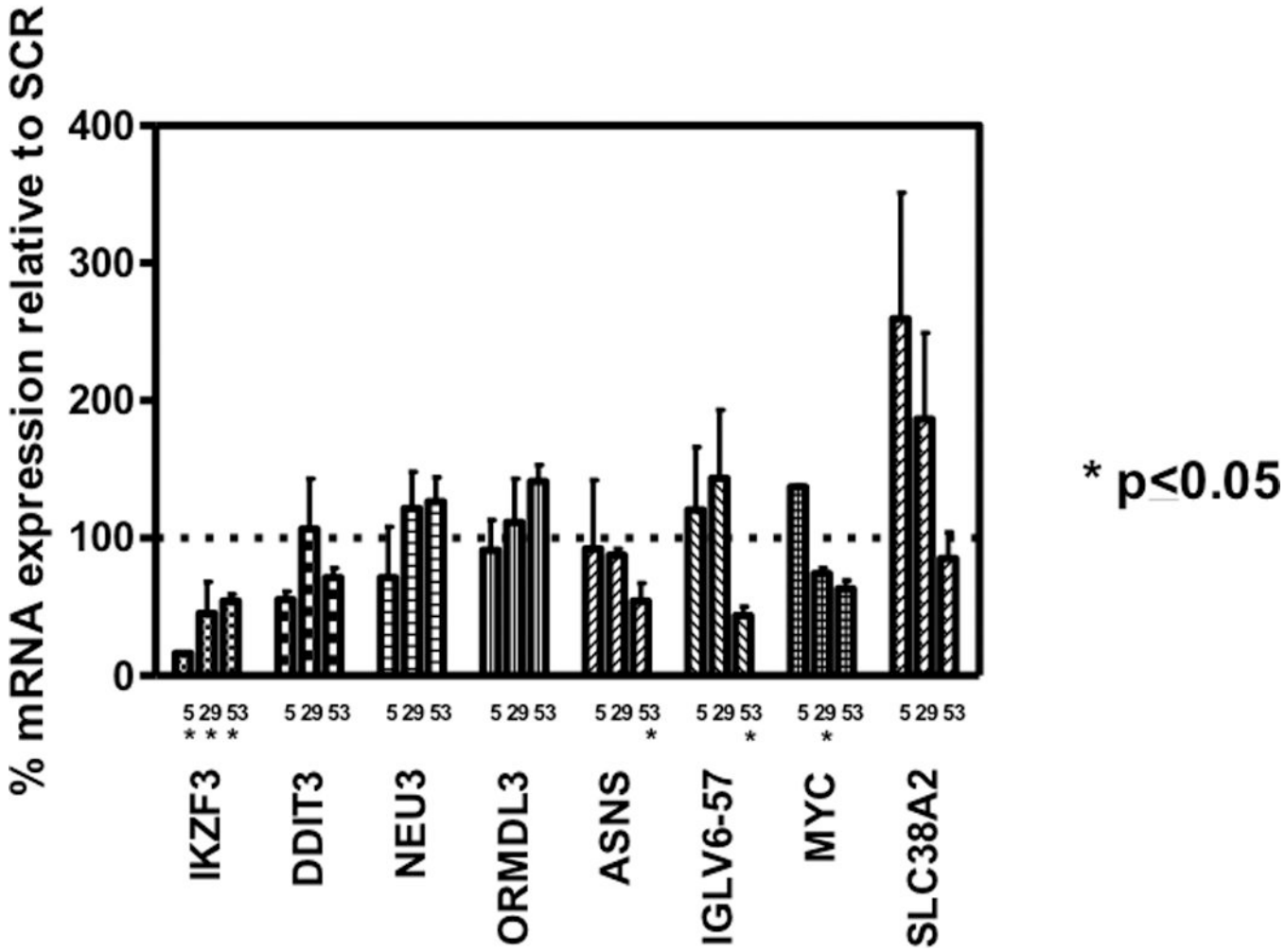

Figure 6.

Effects of siRNA knockdown of *IKZF3* on the host gene and 7 predicted eQTL targets. Significant alterations when compared to nontargeting siRNA controls were seen at 53 hours post-transfection for *ASNS* ($P = 0.03$) and *IGLV6-57* ($P = 4.84\times10^{-6}$) and at 29 hours for *MYC* ($P = 0.02$).

Table 1

Clinical characteristics of the analytic cohort.

Clinical Characteristic	**No (%) 2,709 (100%)**
Age (months)	
< 18	1285 (47.4)
>18	1424 (52.6)
Stage	
1, 2, 3, 4s	1459 (54.2)
4	1233 (45.8)
Sex	
Female	1235 (45.6)
Male	1474 (54.4)
***MYCN* Status**	
Not amplified	2067 (76.3)
Amplified	481 (17.8)
Unknown	161 (5.9)
Ploidy	
Hyperdiploid	1642 (60.6)
Diploid	805 (29.7)
Unknown	262 (9.7)
Histology	
Favorable	976 (36)
Unfavorable	837 (30.9)
Unknown	262 (9.7)
Risk Group	
Low	960 (35.4)
Intermediate	537 (19.8)
High	1212 (44.7)
Self Reported Race, Ethnicity	
White, Not Hispanic	2112 (80)
Black, Not Hispanic	230 (8.5)
Other	367 (13.5)