Title

Integration of Cell Line and Clinical Trial Genome-Wide Analyses Supports a Polygenic Architecture of Paclitaxel-Induced Sensory Peripheral Neuropathy

Journal

Clinical Cancer Research, 19(2)

ISSN

1078-0432

Authors

Wheeler, Heather E
Gamazon, Eric R
Wing, Claudia
et al.

Publication Date

2013-01-15

DOI

10.1158/1078-0432.ccr-12-2618

Peer reviewed

Published in final edited form as:
Clin Cancer Res. 2013 January 15; 19(2): 491–499. doi:10.1158/1078-0432.CCR-12-2618.

Integration of cell line and clinical trial genome-wide analyses supports a polygenic architecture of paclitaxel-induced sensory peripheral neuropathy

Heather E. Wheeler[1], **Eric R. Gamazon**[2], **Claudia Wing**[1], **Uchenna O. Njiaju**[1], **Chidiamara Njoku**[1], **R. Michael Baldwin**[3], **Kouros Owzar**[4], **Chen Jiang**[4], **Dorothy Watson**[4], **Ivo Shterev**[4], **Michiaki Kubo**[5], **Hitoshi Zembutsu**[5], **Eric Winer**[6], **Clifford Hudis**[7], **Lawrence N. Shulman**[6], **Yusuke Nakamura**[1,5], **Mark J. Ratain**[1], **Deanna L. Kroetz**[3], **Nancy J. Cox**[2], and **M. Eileen Dolan**[1,*] **for the Cancer and Leukemia Group B**

[1]Section of Hematology/Oncology, Department of Medicine, University of Chicago, Chicago, IL 60637, USA

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

[3]Department of Bioengineering and Therapeutic Sciences, School of Pharmacy and Medicine, University of California, San Francisco, San Francisco, CA 94143, USA

[4]Department of Biostatistics and Bioinformatics, Duke University, Durham, NC 27710, USA

[5]Center for Genomic Medicine, RIKEN, Yokohama 230-0045, Japan

[6]Dana-Farber Cancer Institute, Boston, MA 02215, USA

[7]Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA

Abstract

Purpose—We sought to demonstrate the relevance of a lymphoblastoid cell line (LCL) model in the discovery of clinically relevant genetic variants affecting chemotherapeutic response by comparing LCL genome-wide association study (GWAS) results to clinical GWAS results.

Experimental Design—A GWAS of paclitaxel-induced cytotoxicity was performed in 247 LCLs from the HapMap Project and compared to a GWAS of sensory peripheral neuropathy in breast cancer patients (n=855) treated with paclitaxel in the Cancer and Leukemia Group B (CALGB) 40101 trial. Significant enrichment was assessed by permutation resampling analysis.

Results—We observed an enrichment of LCL cytotoxicity-associated single nucleotide polymorphisms (SNPs) in the sensory peripheral neuropathy-associated SNPs from the clinical trial with concordant allelic directions of effect (empirical $P = 0.007$). Of the 24 SNPs that overlap between the clinical trial ($P < 0.05$) and the preclinical cytotoxicity study ($P < 0.001$), 19 of them are expression quantitative trait loci (eQTLs), which is a significant enrichment of this functional class (empirical $P = 0.0447$). One of these eQTLs is located in *RFX2*, which encodes a member of the DNA-binding regulatory factor X family. Decreased expression of this gene by siRNA resulted in increased sensitivity of NS-1 (rat pheochromocytoma) cells to paclitaxel as measured by reduced neurite outgrowth and increased cytotoxicity, functionally validating the involvement of *RFX2* in nerve cell response to paclitaxel.

[*]Corresponding author: M. Eileen Dolan, Section of Hematology/Oncology, Department of Medicine, University of Chicago, 900 E. 57$_{th}$ St., Chicago, IL 60637, USA (773)-702-4441, edolan@medicine.bsd.uchicago.edu.

Disclosure: The authors have no conflicts of interest to declare.

Conclusions—The enrichment results and functional example imply that cellular models of chemotherapeutic toxicity may capture components of the underlying polygenic architecture of related traits in patients.

Keywords

pharmacogenomics; clinical trial; cell lines; paclitaxel; genome-wide association

Introduction

Paclitaxel is a tubulin-targeting agent, widely used in the treatment of malignant disease, including ovarian, breast, lung, and head and neck cancers. Its long-term use is often limited by sensory peripheral neuropathy, although the mechanism of this toxicity is poorly understood. In one recent large study of over 1500 breast cancer patients, severe (grade 3) sensory peripheral neuropathy occurred in 4% of patients treated with four cycles and 10% of patients treated with six cycles of single-agent paclitaxel (1). Currently, genetic prediction of which cancer patients may experience severe side effects induced by paclitaxel treatment is not possible (2), but several preliminary genetic associations have been made (3-8). If patients likely to experience such toxicities could be identified prior to beginning a paclitaxel regimen, patient care might be improved by implementing a reduced dose or an alternative treatment.

Because accruing large patient cohorts receiving the same drug regimen for discovery genome-wide and replication studies in oncology is challenging, several groups of investigators have employed lymphoblastoid cell lines (LCLs) as a discovery tool and for follow-up functional studies (9-13). LCLs are easy to experimentally manipulate, and the genetic background and expression environment is known. The LCL model also permits functional validation studies of candidate markers and genes discovered in both preclinical and clinical studies (14, 15). However, a critical question is how well this cell-based model generates clinically relevant markers and genes associated with patient response to drug. Recently, a few chemotherapeutic response SNPs discovered in LCLs have been replicated in patient populations by associating with phenotypes like tumor response and overall survival in patients receiving the same drug (16-19); however, these studies tested the individual variants most associated with the LCL phenotypes. We sought to understand to what extent the overall genetic architecture of patient response to chemotherapy can be captured by LCLs by investigating beyond just the top few signals. In contrast to previous studies that tested single SNPs, we use an enrichment method (20) to determine in a systematic manner whether top GWAS SNPs for paclitaxel-induced sensory peripheral neuropathy in breast cancer patients (3) are more likely to also be paclitaxel-induced cytotoxicity SNPs identified in LCLs than expected by chance.

In this study, we found that SNPs associated with patient paclitaxel-induced neuropathy are enriched for SNPs associated with paclitaxel-induced cytotoxicity in HapMap LCLs. This significant enrichment confirms that LCLs are a useful model in the study of a subset of shared genes involved in patient toxicity. The overlap SNPs are predominantly eQTLs as defined previously (21), therefore supporting an enriched functional role for these significant SNPs. We demonstrate a functional role for one eQTL host gene (*RFX2*) in paclitaxel toxicity, using a cellular model of peripheral neuropathy. These results are consistent with the hypothesis that the cell-based models capture components of the underlying genetic architecture for paclitaxel-induced sensory peripheral neuropathy.

Materials and Methods

Cytotoxicity assays

HapMap LCLs from a population with Northern and Western European ancestry from Utah, USA (HAPMAPPT01, CEU, n = 77), a Yoruba population in Ibadan, Nigeria (HAPMAPPT03, YRI, n = 87), and an African-American population from the Southwest of the USA (HAPMAPPT07, ASW, n = 83) were treated with 12.5 nM paclitaxel and cytotoxicity was determined using an AlamarBlue (Invitrogen, Carlsbad, CA) cellular growth inhibition assay as described (22). The cytotoxicity phenotype used in the LCL GWAS was mean percent survival at 12.5 nM paclitaxel determined from six replicates from two independent experiments. Percent survival values for each cell line were $\log_2$-transformed prior to statistical analysis to form an approximately normal distribution in each population.

LCL genome-wide meta-analysis

A GWAS of paclitaxel-induced cytotoxicity was performed on each of the three populations separately. Greater than 2 million SNPs from HapMap r27 (minor allele frequency (MAF) > 0.05 within the panel, no Mendelian errors and in Hardy-Weinberg equilibrium (P > 0.001)) were tested for association with paclitaxel cytotoxicity in each population, using the quantitative trait disequilibrium test total association model (23). To control for population structure in the admixed ASW population, local ancestry at each genotyped SNP locus was estimated using HAPMIX (24) and to increase genome coverage of the ASW, ungenotyped markers were imputed using BEAGLE (25) as previously described (26). Genomic control lambda (λ_{GC}) values (27) were calculated for the GWAS of each population. Studies with λ_{GC} values greater than 1 were corrected for residual inflation of the test statistic by dividing the observed test statistic at each SNP by the λ_{GC} (27), and then the corresponding P-values were carried through the meta-analysis.

Using the software METAL, we combined SNP P-values across the three population studies, taking into account a study-specific weight (sample size) and direction of effect (positive or negative beta) (28). This approach converted the direction of effect and P-value observed in each study into a signed Z-score, such that very negative Z-scores indicate a small P-value and an allele associated with higher drug sensitivity, whereas large positive Z-scores indicate a small P-value and an allele associated with higher drug resistance. Z-scores for each SNP were combined across studies in a weighted sum, with weights proportional to the square-root of the sample size for each study (28).

Patient samples and GWAS

CALGB 40101 is a phase III trial comparing the efficacy of standard therapy cyclophosphamide and doxorubicin with single-agent paclitaxel as adjuvant therapy for breast cancer in women with 0-3 positive axillary lymph nodes. All study participants were enrolled in CALGB 40101 and gave their additional consent to participate in the pharmacogenetic companion study (CALGB 60202), which has been published (3). All patient research met state, federal and institutional review board guidelines. Germline DNA was isolated from 1,040 patients on the paclitaxel arm of CALGB 40101 and genotyped using the Illumina 610-Quad platform as described previously (3). Following QC analysis, genotypes were available for 520,679 SNPs. Principal components (PC) analysis identified 855 genetic Europeans that were used in a GWAS of sensory peripheral neuropathy (3). A dose-to-event analysis was performed, with an event defined as grade 2 or greater sensory peripheral neuropathy. The Cox score test, powered for additive genetic effects, was used to test these marginal associations. Only SNPs with minor allele frequencies (MAFs) > 0.05 in

the patient population and in Hardy-Weinberg equilibrium in the CEU ($P > 0.001$) were used in the LCL GWAS comparisons.

Enrichment analysis

We conducted a permutation resampling analysis (29) to test for an enrichment of cytotoxicity-associated SNPs (LCLs) among the paclitaxel-induced sensory peripheral neuropathy-associated SNPs (patients). To this end, the patient outcomes (cumulative dose and event indicator vectors) were randomly shuffled while keeping the genotype data fixed to preserve linkage disequilibrium. Based on this permutation replicate, the standardized Cox score statistics were recalculated for all the SNPs. This process was conducted 1000 times. For each of the 1000 permutation replicates, the number of SNPs that had $P < 0.05$ in the patient data, $P < 0.001$ in the LCL data, and the same direction of effect (the same allele associated with increased neuropathy and increased cytotoxicity) was calculated. The overlap distribution from the permutations was compared to the observed SNP overlap to generate an empirical P-value, calculated as the proportion of permutations in which the number of LCL/patient overlap SNPs exceeds the observed number. To test the robustness of our findings, we calculated an empirical P-value across a range of inclusion thresholds from $P < 0.001$ to $P < 0.1$. We also tested for enrichment of patient SNPs among the LCL SNPs by generating 1000 randomized SNP sets the same size and MAF distribution as the observed LCL data at a range of P-value thresholds to calculate empirical P-values. In addition to the paclitaxel LCL cytotoxicity data, we compared the patient sensory peripheral neuropathy data to LCL cytotoxicity GWAS data from capecitabine (30) and carboplatin (13) as negative controls.

To test for eQTL enrichment in the LCL, patient, and LCL/patient overlap SNPs, we generated 10,000 randomized SNP sets each of the same size as the observed set of LCL cytotoxicity ($P < 0.001$), patient neuropathy ($P < 0.05$), or LCL/patient overlap SNPs. The randomized SNP sets were matched on MAF distribution of the observed list and sampled (without replacement) from the set of SNPs on the Illumina 610-Quad platform, similar to the method of Gamazon et al. (31). We grouped the platform SNPs into discrete MAF bins of a width of 5%, from which the SNPs used in the simulations were selected. For each of the 10,000 sets, we determined the number of eQTLs ($P < 10^{-4}$) and calculated an empirical P-value for enrichment. The eQTLs were defined previously and are available in the SCAN database (21, 31).

Filtering procedure for functional analysis

First, we determined which of the LCL/patient overlap SNPs from the enrichment analysis were located in or near (within 2kb) gene transcripts (dbSNP build 129, human genome assembly build 36). Eleven of the 24 overlap SNPs were in or near genes and genotyping intensity plots for these SNPs in the patient data are available in Supplementary Fig. S1. Second, we determined which SNPs within genes were also eQTLs (31) and prioritized by which had the most target genes ($P < 10^{-4}$). We also tested whether the expression of the eQTL target genes associated with paclitaxel-induced cytotoxicity ($P < 0.05$) using previously published exon array data (32). A general linear model was constructed between gene expression and paclitaxel-induced cytotoxicity with growth rate (33) and population as covariates. A Toeplitz covariance structure with two diagonal bands was used to allow for familial dependencies in the data as previously described (9).

siRNA

Neuroscreen-1 (NS-1) rat pheochromacytoma cells (Cellomics Inc., Pittsburg, PA) were maintained in NS-1 media (RPMI supplemented with 10% horse serum, 5% fetal calf serum and 1% L-glutamine). Cells were seeded at a density of 1×10^5 cells/ml on collagen I-

coated plates and induced to differentiate by adding 20 ng/ml nerve growth factor (NGF, BD Biosciences, Bedford, MA) to the media 24 hr prior to transfection. Cells for cytotoxicity assays were plated in 96-well collagen I-coated plates, whereas cells for expression quantification and neurite-outgrowth assays were plated in 6-well collagen I-coated plates. Pooled *Rfx2* siRNA (25 nM; Qiagen, Germantown, MD; S101639659, S101639666, S101639673, S101639680) or non-targeting control siRNA (Qiagen; 1027292) was transiently transfected into the NS-1 cells using DharmaFECT Reagent #1 (Dharmacon, Lafayette, CO). Quantitative reverse transcription PCR (qRT-PCR) was performed for *Rfx2* (Rn00501380_m1) and control gene *Gapdh*(4352338E) using TaqMan Gene Expression Assays (Applied Biosystems, Foster City, CA) 24 hr post-transfection in the neurite-outgrowth assays and 24 hr, 48 hr, 72 hr, and 96 hr post-transfection in the cytotoxicity assays to assess *Rfx2* knockdown in NS-1 cells. Expression of the potential *Rfx2* target genes *Cyp51* (Rn01526553_m1), *Bach1* (Rn01477344_m1), and *Cbara1* (Rn01644475_m1) was also measured by qRT-PCR at 24 hr post-*siRfx2* transfection. Each qRT-PCR was run in duplicate and individual samples were run in triplicate on each plate. Percent knockdown was calculated by dividing the relative *Rfx2* expression levels in the *siRfx2* sample by those in the non-targeting control sample.

Neurite-outgrowth assays

Twenty-four hours following siRNA transfection, transfection media was removed from the NS-1 cells and 0, 12.5, or 100 nM paclitaxel in NS-1 media (supplemented with 20 ng/ml NGF) was added to either the *siRfx2* or non-targeting control cells. After 24 hr in the presence of paclitaxel, phase-contrast images (10×) of the cells were taken using an Axiovert 200M inverted widefield fluorescence microscope (Zeiss, Oberkochen, Germany). At least 500 cells per treatment in six randomly chosen fields were imaged and the longest neurite per cell was measured using ImageJ (34) software. The entire experiment was done in duplicate and mean neurite lengths were normalized relative to the 0 nM drug treatment for each siRNA. Because tracing neurite lengths is somewhat qualitative, two scientists independently measured neurite lengths and the second scientist was blinded to siRNA/drug treatment. The mean of each set of measurements between the two scientists was assessed for significance by two-way ANOVA (factors: siRNA treatment and drug treatment) to determine if the *siRfx2* affected neurite length upon paclitaxel treatment.

NS-1 cytotoxicity assays

Twenty-four hr after siRNA transfection, transfection media was removed from the NS-1 cells and 0, 6.25, 12.5, 25, 50 or 100 nM paclitaxel in NS-1 media (supplemented with 20 ng/ml NGF) in triplicate was added to the either the *siRfx2* or non-targeting control cells. After 72 hr of paclitaxel treatment, ATP levels were measured using the CellTiter-Glo assay (Promega, Madison, WI) and percent survival curves were generated. The entire experiment was done in duplicate and two-way ANOVA was used to determine if the *siRfx2* significantly affected overall cytotoxicity upon paclitaxel treatment.

Results

Enrichment of LCL cytotoxicity SNPs in patient sensory peripheral neuropathy SNPs

We performed a genome-wide meta-analysis (see Materials and Methods) to test common SNPs for association with paclitaxel-induced cytotoxicity in LCLs. We compared the results from this analysis to those from clinical trial CALGB 40101, a GWAS of paclitaxel-induced sensory peripheral neuropathy in breast cancer patients (3). Neither study produced genome-wide significant results ($\alpha < 0.05$) nor did the very top SNPs match between the two studies (Fig. 1). However, through a permutation resampling analysis of the CALGB patient data, we found that the top sensory peripheral neuropathy-associated SNPs ($P < 0.05$) are

significantly enriched for SNPs associated with paclitaxel-induced cytotoxicity in LCLs ($P < 0.001$) with consistent allelic directions of effect (Fig. 2, empirical $P = 0.007$). The observed enrichment of 24 SNPs between the LCL and patient studies is likely paclitaxel-specific, due to the sensory peripheral neuropathy SNPs not being enriched for either capecitabine- or carboplatin-induced cytotoxicity SNPs, which were tested as negative controls (Fig. 2). Positional information and effect sizes of all 24 overlap SNPs in the LCL and patient data can be found in Supplementary Table S1. When the inclusion thresholds for overlap SNPs were relaxed and when the LCL SNPs were tested for enrichment of patient SNPs, the significant overlap was present at a range of P-value thresholds from 0.001 to 0.1, demonstrating the robustness of our findings (Supplementary Table S2).

Enrichment of eQTLs in LCL/patient overlap SNPs

We tested the top paclitaxel-induced LCL cytotoxicity SNPs ($P < 0.001$) and the top paclitaxel-induced patient sensory peripheral neuropathy SNPs ($P < 0.05$) for eQTL enrichment because these were the thresholds used in the primary overlap analysis. We compared the observed number of eQTLs at these thresholds to the number of eQTLs in 10,000 randomly selected minor allele frequency (MAF) matched SNP sets (for details, see Materials and Methods). Neither cytotoxicity-associated SNPs nor neuropathy-associated SNPs alone were enriched for eQTLs (Fig. 3). However, we found that the 24 paclitaxel LCL/patient overlap SNPs at these thresholds are enriched for eQTLs when compared to MAF-matched SNP sets (empirical $P = 0.0447$), potentially revealing an important role for this functional class in paclitaxel toxicity.

Prioritization of LCL/patient overlap SNPs for functional analysis

First, we determined that 11 of the 24 overlap SNPs from the enrichment analysis were located in or near (within 2kb) gene transcripts (Table 1). The relationship of these 11 SNPs with paclitaxel-induced sensory peripheral neuropathy in patients and LCL cytotoxicity is shown in Supplementary Fig. S2. Next, we determined which of these 11 SNPs within genes were also eQTLs (31). Of the eight eQTLs, we determined which had the most potential target genes at an arbitrary threshold of $P < 10^{-4}$. The SNP in *RFX2* had 18 target genes, more than any other of the eight eQTLs. In addition, we tested the expression of the target genes for association with paclitaxel-induced cytotoxicity adjusted for growth rate (see Materials and Methods). We found that expression of three of the *RFX2* target genes associated with paclitaxel-induced cytotoxicity (Table 1, Supplementary Table S3); therefore, we pursued evaluating *RFX2* in a model of neuropathy.

Functional validation of RFX2 in a paclitaxel-induced peripheral neuropathy model

We used Neuroscreen (NS-1) cells, a subclone of the rat pheochromocytoma cell line PC-12 that has previously been used as a research model for chemotherapy-induced neuropathy (35, 36), to test *Rfx2*, the rat ortholog of *RFX2*, for functional involvement in paclitaxel response. Using siRNA, we decreased expression of *Rfx2* resulting in increased sensitivity of NS-1 cells to paclitaxel, as measured by reduced neurite outgrowth and increased cytotoxicity (Fig. 4). The three *RFX2* SNP target genes whose expression associated with paclitaxel-induced cytotoxicity in LCLs are *CYP51A1*, *BACH1*, and *CBARA1* (Table 1, Fig. 5A-C, $P < 0.05$). We measured the expression of these three potential *Rfx2* target genes upon knockdown of *Rfx2* in NS-1 cells and found that the expression of one of the three genes, *Cyp51* (rat ortholog of *CYP51A1*), significantly decreased 24 hr post-transfection ($P < 0.05$), which is the expected direction of effect based on the LCL expression versus cytotoxicity data (Fig. 5D).

Discussion

We performed a GWAS of paclitaxel-induced cytotoxicity in LCLs and showed significant enrichment of the top cytotoxicity-associated SNPs in a clinical GWAS of paclitaxel-induced sensory peripheral neuropathy in breast cancer patients. This robust enrichment demonstrates that susceptibilities to increased cytotoxicity in LCLs and sensory peripheral neuropathy in breast cancer patients likely have some genetic mechanisms in common and supports the role of LCLs as a preclinical model for paclitaxel toxicity studies. Furthermore, the top SNPs that overlap between the two studies were enriched for eQTLs. This eQTL enrichment indicates that SNPs associated with paclitaxel-induced toxicity phenotypes may be functioning through gene regulatory mechanisms. Interestingly, neither GWAS alone was enriched for eQTLs. Thus, our integration method may be reducing noise and revealing important functional SNPs. An enrichment of eQTLs has previously been demonstrated in SNPs associated with six other chemotherapeutic drugs, which indicates that susceptibility to these drugs may depend on subtle gene expression differences across individuals (31).

The enrichment analyses were likely affected by the different linkage disequilibrium (LD) patterns among the populations studied. The LCL GWAS was a meta-analysis of African, African American and European populations while the patient GWAS was performed in Europeans. In the meta-analysis, SNPs that are associated with cytotoxicity in all populations are prioritized over those associated in only one of the populations. We may have missed identifying European-specific overlap alleles. However, because the population LD patterns differ and because African populations have shorter LD blocks, overlap SNPs are more likely to be functional SNPs rather than SNPs that simply tag a functional locus (37).

We functionally assessed the involvement of one overlap eQTL, *RFX2*, in the NS-1 neuropathy cell model. Paclitaxel has previously been shown to decrease neurite outgrowth in the parent clone of the NS-1 cell line (36). Here, we showed that decreased expression of *Rfx2* sensitizes NS-1 cells to paclitaxel by reducing neurite outgrowth and survival. This result validates our approach by demonstrating that patient neuropathy and LCL cytotoxicity overlap analyses can reveal genes mechanistically involved in paclitaxel response. Although most previous work on *RFX2* in mammalian cells describes its role in spermatogenesis (38, 39), several studies point to a potential role for the protein in sensory neurons. RFX2 and the related protein RFX1 have been shown to directly bind and regulate the transcription of *ALMS1* (40). Mutations in *ALMS1* cause the rare genetic disorder Alström syndrome, which is characterized by neurosensory degeneration, metabolic defects and cardiomyopathy (40). In addition, the regulatory factor X transcription factors present in *C. elegans* and *Drosophila*, which are called DAF-19 and RFX, respectively, regulate ciliated sensory neuron differentiation (41, 42).

Upon knockdown of *Rfx2* in NS-1 cells, the potential target gene *Cyp51* also decreased expression, which was the expected direction of effect based on the preliminary gene expression analysis in LCLs. However, *CYP51A1* does not contain an X-box RFX binding domain (43) in the promoter region (2kb upstream of the transcription start site), which means it is unlikely a direct target of *RFX2* and may instead be further downstream in the pathway. Alternatively, *RFX2* could be regulating an enhancer of *CYP51A1* that is further outside the gene region. CYP51A1 is a member of the cytochrome P450 superfamily of enzymes, which catalyze many reactions involved in the metabolism of drugs and endogenous compounds. Specifically, CYP51A1 is known to participate in the synthesis of cholesterol (44). CYP51A1 has not previously been implicated in paclitaxel metabolism (45).

In the CALGB GWAS, one of the top SNPs that associated with patient paclitaxel-induced sensory peripheral neuropathy (rs10771973, $P = 2.6 \times 10^{-6}$) is located in *FGD4* (3). Mutations in *FGD4* can cause the congenital peripheral neuropathy Charcot-Marie-Tooth disease type 4H and thus the gene is a plausible candidate for involvement in variation in peripheral neuropathy induced by paclitaxel. This SNP association was replicated in a second cohort of self-reported white breast cancer patients ($n = 154$, $P = 0.013$) and in a cohort of self-reported African American breast cancer patients ($n = 117$, $P = 6.7 \times 10^{-3}$) (3). However, this SNP was not associated with paclitaxel-induced cytotoxicity in LCLs ($P = 0.65$). *FGD4* is not expressed in LCLs (21) and thus the SNP is not expected to function in this model system. While our integrative approach can reveal variants and genes acting in paclitaxel response in both patients and LCLs, it does not identify genes potentially acting in patients that are not expressed in LCLs.

Effectively, LCLs have been used as an additional cohort to study the pharmacogenomics of various chemotherapeutics (16-19) because limited resources and *in vivo* confounders make obtaining large, homogeneous patient cohorts difficult. Here, we saw greater SNP overlap than expected by chance between SNPs associated with paclitaxel-induced cytotoxicity in LCLs and SNPs associated with paclitaxel-induced sensory peripheral neuropathy in patients at multiple P-value thresholds, which confirms a role for the LCL model in the analysis of at least a subset of genes involved in patient neurotoxicity. This significant enrichment among a relatively large number of top SNPs is consistent with an underlying polygenic architecture for paclitaxel-induced toxicity. Functional siRNA studies in the NS-1 neuropathy model validated the involvement of RFX2 in paclitaxel toxicity, supporting our multi-gene hypothesis. Our novel integrative enrichment approach that combines clinical and LCL GWAS results can be used to expand patient cohort sizes for any drug phenotype of interest, including other toxicities such as neutropenia, to find genes of potential impact that can be studied in cellular models.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors are grateful to the PAAR LCL core for maintaining and distributing lymphoblastoid cell lines.

Financial support: This study is supported by NIH/NCI Cancer Biology Training grant T32CA009594 and NRSA F32CA165823 (HEW), NIH/NIGMS Pharmacogenomics of Anticancer Agents grant U01GM61393 (MJR, NJC, MED, KO), NIH/NIGMS Pharmacogenetics of Membrane Transporters grant U01GM61390 (DLK), University of Chicago SPORE Grant NCI P50CA125183 (MED), NIH/NCI Medical Oncology Training grant T32CA009566 (UON), and the Biobank Japan Project funded by the Japanese Ministry of Education, Culture, Sports, Science and Technology. This work is part of the NIH Pharamcogenomics Research Network-RIKEN Center for Genomic Medicine Global Alliance.

The research for CALGB 60202 and 40101 was supported, in part, by grants from the National Cancer Institute (CA31946) to the Cancer and Leukemia Group B (Monica M. Bertagnolli, M.D., Chair) and to the CALGB Statistical Center (Daniel J. Sargent, Ph.D., CA33601). The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute.

References

1. Shulman LN, Cirrincione CT, Berry DA, Becker HP, Perez EA, O'Regan R, et al. Six Cycles of Doxorubicin and Cyclophosphamide or Paclitaxel Are Not Superior to Four Cycles As Adjuvant Chemotherapy for Breast Cancer in Women With Zero to Three Positive Axillary Nodes: Cancer and Leukemia Group B 40101. J Clin Oncol. 2012

2. Pachman DR, Barton DL, Watson JC, Loprinzi CL. Chemotherapy-induced peripheral neuropathy: prevention and treatment. Clin Pharmacol Ther. 2011; 90:377–87. [PubMed: 21814197]
3. Baldwin RM, Owzar K, Zembutsu H, Chhibber A, Kubo M, Jiang C, et al. A Genome-Wide Association Study Identifies Novel Loci for Paclitaxel-Induced Sensory Peripheral Neuropathy in CALGB 40101. Clin Cancer Res. 2012; 18:5099–109. [PubMed: 22843789]
4. Green H, Soderkvist P, Rosenberg P, Horvath G, Peterson C. mdr-1 single nucleotide polymorphisms in ovarian cancer tissue: G2677T/A correlates with response to paclitaxel chemotherapy. Clin Cancer Res. 2006; 12:854–9. [PubMed: 16467099]
5. Hertz DL, Motsinger-Reif AA, Drobish A, Winham SJ, McLeod HL, Carey LA, et al. CYP2C8*3 predicts benefit/risk profile in breast cancer patients receiving neoadjuvant paclitaxel. Breast Cancer Res Treat. 2012; 134:401–10. [PubMed: 22527101]
6. Leandro-Garcia LJ, Leskela S, Jara C, Green H, Avall-Lundqvist E, Wheeler HE, et al. Regulatory polymorphisms in beta-tubulin IIa are associated with paclitaxel-induced peripheral neuropathy. Clin Cancer Res. 2012; 18:4441–8. [PubMed: 22718863]
7. Leskela S, Leandro-Garcia LJ, Mendiola M, Barriuso J, Inglada-Perez L, Munoz I, et al. The miR-200 family controls beta-tubulin III expression and is associated with paclitaxel-based treatment response and progression-free survival in ovarian cancer patients. Endocr Relat Cancer. 2011; 18:85–95. [PubMed: 21051560]
8. Sissung TM, Mross K, Steinberg SM, Behringer D, Figg WD, Sparreboom A, et al. Association of ABCB1 genotypes with paclitaxel-mediated peripheral neuropathy and neutropenia. Eur J Cancer. 2006; 42:2893–6. [PubMed: 16950614]
9. Huang RS, Duan S, Bleibel WK, Kistner EO, Zhang W, Clark TA, et al. A genome-wide approach to identify genetic variants that contribute to etoposide-induced cytotoxicity. Proc Natl Acad Sci U S A. 2007; 104:9758–63. [PubMed: 17537913]
10. Li L, Fridley BL, Kalari K, Jenkins G, Batzler A, Weinshilboum RM, et al. Gemcitabine and arabinosylcytosin pharmacogenomics: genome-wide association and drug response biomarkers. PLoS One. 2009; 4:e7765. [PubMed: 19898621]
11. Watters JW, Kraja A, Meucci MA, Province MA, McLeod HL. Genome-wide discovery of loci influencing chemotherapy cytotoxicity. Proc Natl Acad Sci U S A. 2004; 101:11809–14. [PubMed: 15282376]
12. Wheeler HE, Dolan ME. Lymphoblastoid cell lines in pharmacogenomic discovery and clinical translation. Pharmacogenomics. 2012; 13:55–70. [PubMed: 22176622]
13. Wheeler HE, Gamazon ER, Stark AL, O'Donnell PH, Gorsic LK, Huang RS, et al. Genome-wide meta-analysis identifies variants associated with platinating agent susceptibility across populations. Pharmacogenomics J. 201110.1038/tpj.2011.38
14. Ingle JN, Schaid DJ, Goss PE, Liu M, Mushiroda T, Chapman JA, et al. Genome-wide associations and functional genomic studies of musculoskeletal adverse events in women receiving aromatase inhibitors. J Clin Oncol. 2010; 28:4674–82. [PubMed: 20876420]
15. Shukla SJ, Duan S, Wu X, Badner JA, Kasza K, Dolan ME. Whole-genome approach implicates CD44 in cellular resistance to carboplatin. Hum Genomics. 2009; 3:128–42. [PubMed: 19164090]
16. Huang RS, Johnatty SE, Gamazon ER, Im HK, Ziliak D, Duan S, 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; 17:5490–500. [PubMed: 21705454]
17. Mitra AK, Crews K, Pounds S, Cao X, Downing JR, Raimondi S, et al. Impact of genetic variation in FKBP5 on clinical response in pediatric acute myeloid leukemia patients: a pilot study. Leukemia. 2011; 25:1354–6. [PubMed: 21483441]
18. Tan XL, Moyer AM, Fridley BL, Schaid DJ, Niu N, Batzler AJ, et al. Genetic variation predicting Cisplatin cytotoxicity associated with overall survival in lung cancer patients receiving platinum-based chemotherapy. Clin Cancer Res. 2011; 17:5801–11. [PubMed: 21775533]
19. Ziliak D, O'Donnell PH, Im HK, Gamazon ER, Chen P, Delaney S, et al. Germline polymorphisms discovered via a cell-based, genome-wide approach predict platinum response in head and neck cancers. Transl Res. 2011; 157:265–72. [PubMed: 21497773]

20. Cox NJ, Gamazon ER, Wheeler HE, Dolan ME. Clinical translation of cell-based pharmacogenomic discovery. Clin Pharmacol Ther. 2012; 92:425–7. [PubMed: 22910437]
21. Gamazon ER, Zhang W, Konkashbaev A, Duan S, Kistner EO, Nicolae DL, et al. SCAN: SNP and copy number annotation. Bioinformatics. 2010; 26:259–62. [PubMed: 19933162]
22. Njiaju UO, Gamazon ER, Gorsic LK, Delaney SM, Wheeler HE, Im HK, et al. Whole-genome studies identify solute carrier transporters in cellular susceptibility to paclitaxel. Pharmacogenet Genomics. 2012; 22:498–507. [PubMed: 22437668]
23. Abecasis GR, Cookson WO, Cardon LR. Pedigree tests of transmission disequilibrium. Eur J Hum Genet. 2000; 8:545–51. [PubMed: 10909856]
24. Price AL, Tandon A, Patterson N, Barnes KC, Rafaels N, Ruczinski I, et al. Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. PLoS Genet. 2009; 5:e1000519. [PubMed: 19543370]
25. Browning BL, Browning SR. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. Am J Hum Genet. 2009; 84:210–23. [PubMed: 19200528]
26. Wheeler HE, Gorsic LK, Welsh M, Stark AL, Gamazon ER, Cox NJ, et al. Genome-wide local ancestry approach identifies genes and variants associated with chemotherapeutic susceptibility in African Americans. PLoS One. 2011; 6:e21920. [PubMed: 21755009]
27. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999; 55:997–1004. [PubMed: 11315092]
28. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010; 26:2190–1. [PubMed: 20616382]
29. 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]
30. O'Donnell PH, Stark AL, Gamazon ER, Wheeler HE, McIlwee BE, Gorsic L, et al. Identification of novel germline polymorphisms governing capecitabine sensitivity. Cancer. 2012; 118:4063–73. [PubMed: 22864933]
31. 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; 107:9287–92. [PubMed: 20442332]
32. Duan S, Huang RS, Zhang W, Bleibel WK, Roe CA, Clark TA, et al. Genetic architecture of transcript-level variation in humans. Am J Hum Genet. 2008; 82:1101–13. [PubMed: 18439551]
33. Stark AL, Zhang W, Mi S, Duan S, O'Donnell PH, Huang RS, et al. Heritable and non-genetic factors as variables of pharmacologic phenotypes in lymphoblastoid cell lines. Pharmacogenomics J. 2010; 10:505–12. [PubMed: 20142840]
34. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012; 9:671–5. [PubMed: 22930834]
35. Geldof AA. Nerve-growth-factor-dependent neurite outgrowth assay; a research model for chemotherapy-induced neuropathy. J Cancer Res Clin Oncol. 1995; 121:657–60. [PubMed: 7593129]
36. Verstappen CC, Postma TJ, Geldof AA, Heimans JJ. Amifostine protects against chemotherapy-induced neurotoxicity: an in vitro investigation. Anticancer Res. 2004; 24:2337–41. [PubMed: 15330181]
37. Teo YY, Small KS, Kwiatkowski DP. Methodological challenges of genome-wide association analysis in Africa. Nat Rev Genet. 2010; 11:149–60. [PubMed: 20084087]
38. Horvath GC, Kistler WS, Kistler MK. RFX2 is a potential transcriptional regulatory factor for histone H1t and other genes expressed during the meiotic phase of spermatogenesis. Biol Reprod. 2004; 71:1551–9. [PubMed: 15229132]
39. Wolfe SA, Wilkerson DC, Prado S, Grimes SR. Regulatory factor X2 (RFX2) binds to the H1t/TE1 promoter element and activates transcription of the testis-specific histone H1t gene. J Cell Biochem. 2004; 91:375–83. [PubMed: 14743396]
40. Purvis TL, Hearn T, Spalluto C, Knorz VJ, Hanley KP, Sanchez-Elsner T, et al. Transcriptional regulation of the Alstrom syndrome gene ALMS1 by members of the RFX family and Sp1. Gene. 2010; 460:20–9. [PubMed: 20381594]

41. Swoboda P, Adler HT, Thomas JH. The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in C. elegans. Mol Cell. 2000; 5:411–21. [PubMed: 10882127]

42. Dubruille R, Laurencon A, Vandaele C, Shishido E, Coulon-Bublex M, Swoboda P, et al. Drosophila regulatory factor X is necessary for ciliated sensory neuron differentiation. Development. 2002; 129:5487–98. [PubMed: 12403718]

43. Gajiwala KS, Chen H, Cornille F, Roques BP, Reith W, Mach B, et al. Structure of the winged-helix protein hRFX1 reveals a new mode of DNA binding. Nature. 2000; 403:916–21. [PubMed: 10706293]

44. Halder SK, Fink M, Waterman MR, Rozman D. A cAMP-responsive element binding site is essential for sterol regulation of the human lanosterol 14alpha-demethylase gene (CYP51). Mol Endocrinol. 2002; 16:1853–63. [PubMed: 12145339]

45. Oshiro C, Marsh S, McLeod H, Carrillo M, Klein T, Altman R. Taxane Pathway. Pharmacogenet Genomics. 2009; 19:979–83. [PubMed: 21151855]

Translational Relevance

Lymphoblastoid cell lines (LCLs) have been used in chemotherapeutic pharmacogenomic marker discovery due to their ease of experimental manipulation, extensive genotype catalogs, and lack of the *in vivo* confounders present in clinical samples. One important question is how well these cell-based models generate clinically relevant SNPs associated with patient toxicity. We compared genome-wide association study results of paclitaxel-induced cytotoxicity in LCLs and paclitaxel-induced peripheral neuropathy in breast cancer patients. We observed significant overlap between the clinical and LCL studies, thus confirming a role for the LCL model in the analysis of at least a subset of genes involved in patient paclitaxel response. One overlap gene, *RFX2*, was functionally validated in a nerve cell model of paclitaxel response. Peripheral neuropathy is an often dose-limiting toxicity induced by paclitaxel treatment. If physicians could predict which patients are more likely to experience this severe toxicity, lower doses or alternative treatments could be prescribed.

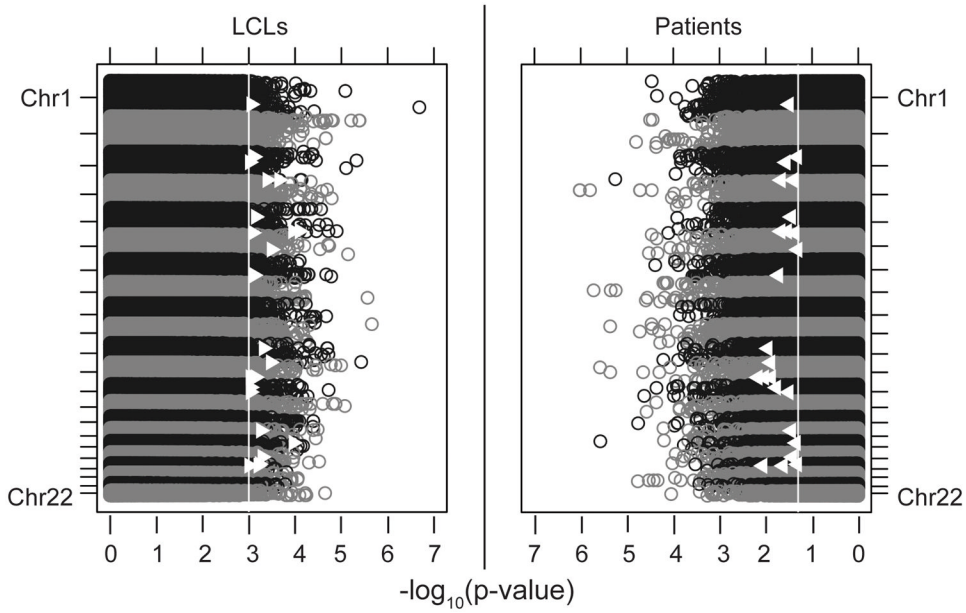

Figure 1. Comparison of individual genome-wide association study (GWAS) results

Left: paclitaxel-induced cytotoxicity in LCLs. Right: paclitaxel-induced sensory peripheral neuropathy in patients. White lines represent the overlap thresholds used in the primary enrichment analysis ($P < 0.001$ for LCLs and $P < 0.05$ for patients) and white triangles represent the 24 overlap SNPs at these thresholds.

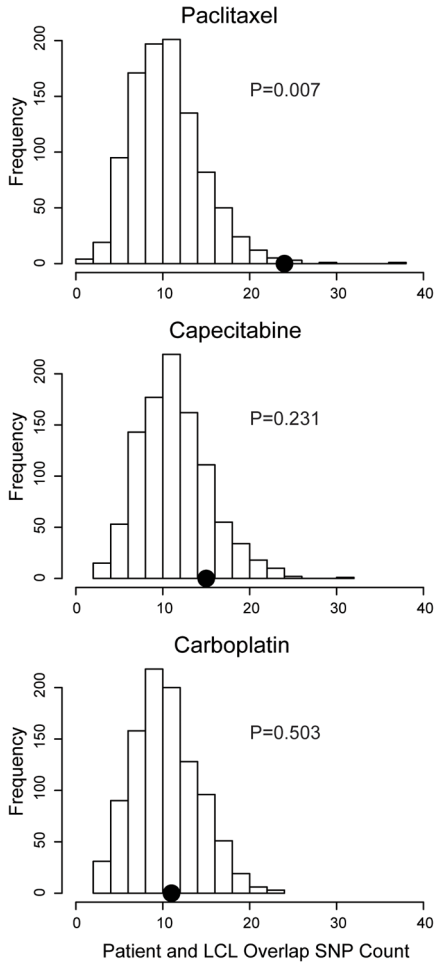

Figure 2. Patient paclitaxel-induced sensory peripheral neuropathy SNPs are enriched for SNPs associated with paclitaxel-induced cytotoxicity in LCLs

Distribution of chemotherapeutic-induced cytotoxicity SNP ($P < 0.001$) count in 1000 permutations of neuropathy phenotype-genotype connections ($P < 0.05$). The dot is the observed SNP overlap at these thresholds. Of the three drug studies tested (paclitaxel, capecitabine, carboplatin), only paclitaxel-induced cytotoxicity SNPs were significantly enriched in the patient GWAS (empirical $P = 0.007$).

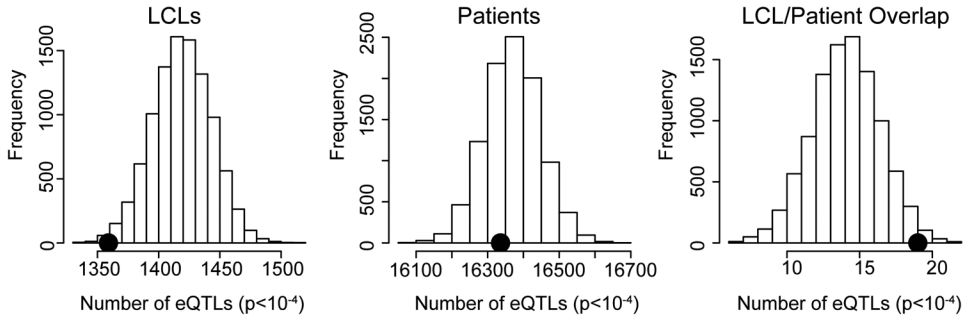

Figure 3. SNPs associated with both patient paclitaxel-induced sensory peripheral neuropathy and LCL paclitaxel-induced cytotoxicity are enriched for expression quantitative trait loci (eQTLs)

Distribution of eQTL ($P < 10^{-4}$) count in 10,000 simulations, each matching the MAF distribution of either LCL paclitaxel SNPs ($P < 0.001$), patient paclitaxel SNPs ($P < 0.05$), or the set of 24 LCL/patient overlap SNPs at these P-value thresholds. Neither the LCL paclitaxel SNPs nor the patient paclitaxel SNPs alone were enriched for eQTLs, but the overlap SNP set between the two GWAS was enriched for eQTLs (empirical $P = 0.0447$).

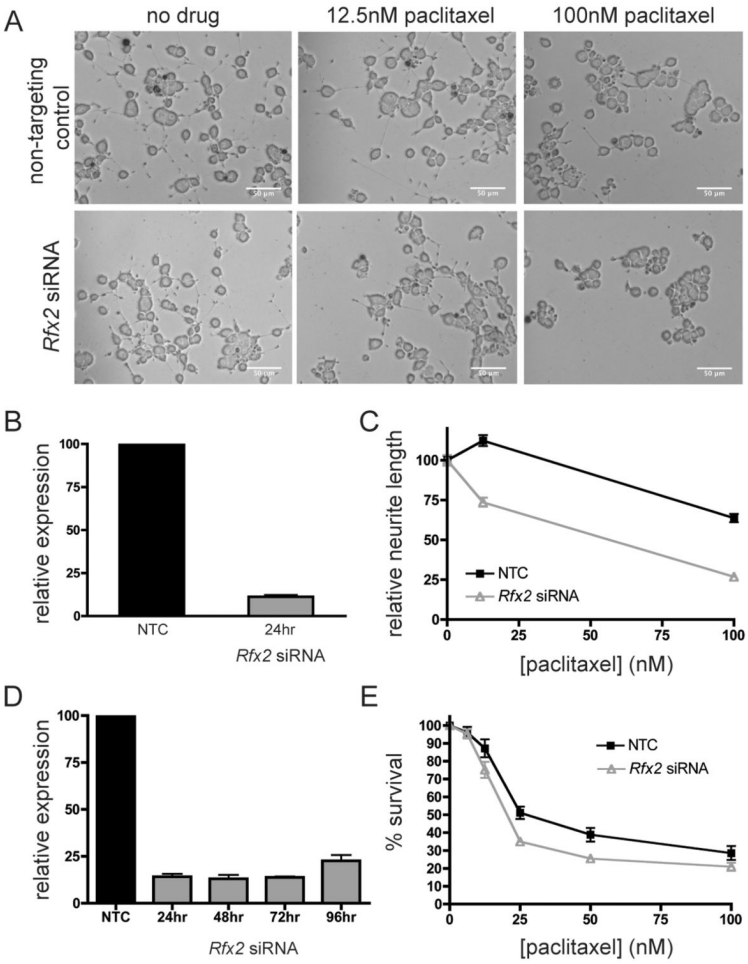

Figure 4. Functional validation of *RFX2* in paclitaxel response using a peripheral neuropathy cell model

(A) Representative micrographs comparing neurite lengths of NS-1 cells upon siRNA knockdown of *Rfx2* and treatment with paclitaxel (10× phase-contrast). (B) Relative gene expression 24 hr post-transfection in the two neurite length experiments. NTC = non-targeting control. (C) Decreased expression of *Rfx2* causes decreased neurite length of differentiating NS-1 cells ($P < 10^{-4}$) 24 hr post-paclitaxel treatment (48 hr post-transfection). Error bars represent the SEM of the longest relative neurite length of at least 500 cells in each of two independent experiments. (D) Relative gene expression 24-96 hr post-transfection in the two cytotoxicity experiments. (E) Decreased expression of *Rfx2* causes decreased survival (increased cytotoxicity, $P < 10^{-4}$) of differentiating NS-1 cells measured by CellTiter-Glo 72 hr post-paclitaxel treatment (96 hr post-transfection). Error bars represent the SEM of survival in two independent experiments with three replicates each.

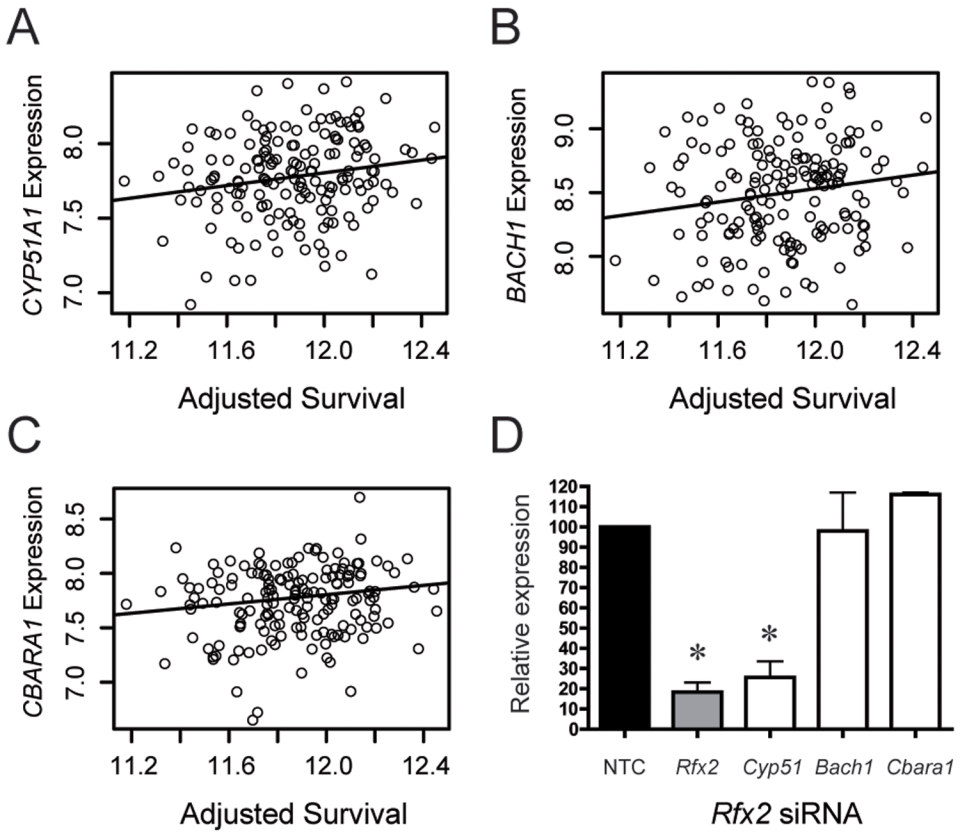

Figure 5. Target genes of *RFX2* eQTL rs7254081 in paclitaxel response

Increased baseline expression of the rs7254081 target genes (A) *CYP51A1*, (B) *BACH1*, and (C) *CBARA1* associate with increased cellular survival (adjusted for growth rate) of LCLs treated with 12.5nM paclitaxel ($P < 0.05$). (D) *Rfx2* siRNA in NS-1 cells decreases the expression of *Rfx2* and *Cyp51*, but not *Bach1* and *Cbara1*, compared to the non-targeting control (NTC) 24 hr post-transfection. *$P < 0.05$. Error bars represent the SEM of relative gene expression in two independent experiments with three replicates each.

Table 1

Paclitaxel-induced LCL cytotoxicity (P < 0.001) and paclitaxel-induced patient sensory peripheral neuropathy (P < 0.05) overlap SNPs located in genes.

SNP	LCL cytotoxicity P-value	Patient sensory peripheral neuropathy P-value	Gene	eQTL	# target genes	Target genes associated with LCL paclitaxel-induced cytotoxicity[a] (P < 0.05)
rs7254081	5.9E-04	4.8E-02	*RFX2*	x	18	*CYP51A1, BACH1, CBARA1*
rs7642318	2.2E-04	3.9E-02	*TMEM44*	x	6	
rs10933663	4.0E-04	2.1E-02	*TMEM44*	x	4	
rs8002545	9.2E-04	3.1E-02	*DIS3*	x	3	
rs4782010	5.5E-04	3.5E-02	*XYLT1*	x	2	
rs11111539	7.9E-04	6.7E-03	*C12orf42*	x	1	
rs7306825	7.2E-04	9.3E-03	*C12orf42*	x	1	
rs8069856	1.1E-04	4.5E-02	*RICH2*	x	1	
rs4868011	8.2E-04	4.2E-02	*KCNIP1*			
rs10778237	9.3E-04	1.3E-02	*C12orf42*			
rs323285	5.1E-04	3.7E-02	*KIAA1328*			

[a] adjusted for growth rate