Published in final edited form as:

Pharmacogenomics. 2012 January ; 13(1): 55–70. doi:10.2217/pgs.11.121.

Lymphoblastoid cell lines in pharmacogenomic discovery and clinical translation

Heather E Wheeler and M Eileen Dolan*

Section of Hematology/Oncology, Department of Medicine, 900 East 57th St, University of Chicago, Chicago, IL 60637, USA

Abstract

The ability to predict how an individual patient will respond to a particular treatment is the ambitious goal of personalized medicine. The genetic make up of an individual has been shown to play a role in drug response. For pharmacogenomic studies, human lymphoblastoid cell lines (LCLs) comprise a useful model system for identifying genetic variants associated with pharmacologic phenotypes. The availability of extensive genotype data for many panels of LCLs derived from individuals of diverse ancestry allows for the study of genetic variants contributing to interethnic and interindividual variation in susceptibility to drugs. Many genome-wide association studies for drug-induced phenotypes have been performed in LCLs, often incorporating gene-expression data. LCLs are also being used in follow-up studies to clinical findings to determine how an associated variant functions to affect phenotype. This review describes the most recent pharmacogenomic findings made in LCLs, including the translation of some findings to clinical cohorts.

Keywords

β -blockers; acetaminophen; chemotherapy; cytotoxicity; gene expression; genome-wide association studies; HapMap; immunosuppressants; lymphoblastoid cell lines; pharmacogenomics; radiation; selective serotonin reuptake inhibitors; statins

The goal of personalized medicine is to maximize effective therapy response and to minimize adverse reactions by prescribing treatment based on a patient's genetic profile. Although *in vivo* studies in humans are the most relevant system to determine how an individual's genetic profile influences response to medication, this is not the most practical system, especially for toxic drugs. The use of human Epstein–Barr virus (EBV)-transformed lymphoblastoid cell lines (LCLs) has emerged as a promising model system in the study of the genetics of drug response. LCLs provide a cost-effective testing system where environmental factors such as drug dosage can be controlled. Genome-wide genotype [201–204] and gene expression [205–209] data, including next-generation sequencing (DNA and RNA-Seq) data, is publicly available for hundreds of established LCLs. The development

© 2012 Future Medicine Ltd

*Author for correspondence: Tel.: +1 773 702 4441, Fax: +1 773 702 9268 edolan@medicine.bsd.uchicago.edu.

For reprint orders, please contact: reprints@futuremedicine.com

Financial & competing interests disclosure

The authors are supported by the Pharmacogenetics of Anticancer Agents Research Group NIH/National Institute of General Medical Sciences grant U01GM61393, P50 CA125183, CA136765, R21CA139278 and the University of Chicago Cancer Research Center Pilot Funding. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

and initial implementation of this model system in the pharmacogenomics field has been reviewed extensively [1]. This review will focus primarily on the pharmacogenomic studies in LCLs that were published in the past few years.

The advantages of LCLs, which include the ease of experimental manipulation and a lack of the *in vivo* confounders present in clinical samples, are numerous. However, like any model system, there are limitations. A major limitation is that most drug-induced effects involve the interaction of different cell types and organs; thus, a single-model system cannot represent the complexity of drug effects in the human body. For example, LCLs do not express many of the CYP450 enzymes and therefore are not useful for pharmacokinetic studies, which instead are often performed in hepatocytes [2,3]. Other cell models, including fibroblasts and peripheral blood mononuclear cells are also used in pharmacogenetic studies [4,5]. However, unlike LCLs, these cell types do not have extensive catalogs of lines and genetic information available. Other limitations include nongenetic factors such as baseline growth rates, EBV copy numbers and ATP levels that may influence drug-induced phenotypes in LCLs [6]. LCL growth rate has been shown to be associated with chemotherapeutic-induced cytotoxicity and should be considered in all LCL analyses [7]. In addition, EBV transformation has been shown to alter apoptosis in response to certain drugs, which is important to consider when LCLs are used in pharmacogenomic studies, especially in studies of cancer drugs [8]. One recent study compared LCLs and primary B cells from the same individuals and found that EBV transformation affected the gene-expression profiles and promoter-methylation profiles of more than half of the genes measured [9]. However, most expression differences were of small magnitude (<1.5-fold) and the interindividual variation in gene-expression levels observed in the primary B cells was maintained in the LCLs [9]. Thus, expression quantitative trait loci (eQTLs) found in LCLs are likely to reflect the underlying regulatory variation of primary B cells and even of nonblood tissues. Several studies have provided empirical evidence that many eQTLs observed in LCLs are also observed in primary tissues, including the skin, liver and heart [10–12].

Importantly, because of potential *in vitro* confounders, pharmacogenomic SNP and gene associations discovered in the LCL model should be replicated in relevant tissues and in clinical populations. In addition to SNP discovery, LCLs are useful for functional follow-up studies. The mechanism of action of potential functional genes found in both clinical and LCL genome-wide association (GWA) studies can be further explored by molecular manipulation experiments in both LCLs and tumor cell lines. Demonstrating the utility of the LCL model, some SNPs associated with chemotherapeutic susceptibility in LCL discovery studies have recently been shown to associate with clinical phenotypes such as event-free survival and overall survival following treatment with the relevant chemotherapy; these important studies will be discussed in this review.

LCL population panels

The three main collections of LCLs that have been used in pharmacogenomics studies are large Centre d'Étude du Polymorphisme Humain (CEPH) pedigrees, International HapMap Project populations and Human Variation Panel populations (Table 1). The CEPH pedigrees comprise 48 multigenerational families each with an average of 14 individuals of northern and western European ancestry from Utah (CEU) [13]. SNP genotypes for approximately 5000 SNP markers in these CEPH families can be downloaded from the SNP Consortium Database of the SNP Consortium Linkage Map Project [14,201]. Global gene-expression data is also available for 14 of the families [15,206,209]. The cell lines are available through the Coriell Institute Cell Repositories (NJ, USA) [210]. Additional SNP genotypes can be imputed from HapMap data [16]. These large pedigrees have been used to demonstrate that

chemotherapeutic-induced cytotoxicity is a heritable trait amenable to genetic dissection [17,18].

Trios (mother–father–child) from some of these CEPH families were included in the International HapMap Project and comprise the CEU population [19–21,202]. The International HapMap Project was started in order to develop a human haplotype map that catalogs the common patterns of human DNA sequence variation across world populations. The HapMap Phase I/II Project published the genotypes of over 3.1 million SNPs from 270 apparently healthy individuals from four populations (CEU: 30 CEPH trios from Utah, USA; Yoruba from Ibadan, Nigeria [YRI]: 30 trios; Han Chinese from Beijing, China [CHB]: 45 unrelated; and Japanese from Tokyo, Japan [JPT]: 45 unrelated). Genotypes, phased haplotypes and linkage disequilibrium information can be viewed and downloaded from the HapMap Project website [22,202].

The recently released HapMap Phase III sample collection provides additional opportunities for applying cell-based approaches to the identification of pharmacogenetic loci. HapMap Phase III released 1.6 million SNP genotypes from 1184 samples from 11 diverse populations, which include additional individuals from CEU, YRI, CHB and JPT as well as seven additional populations (African ancestry in southwest USA [ASW]; Gujarati Indians in Houston, TX, USA [GIH]; Luhya in Webuye, Kenya [LWK]; Mexican ancestry in Los Angeles, CA, USA [MEX]; Chinese in metropolitan Denver, CO, USA [CHD]; Maasai in Kinyawa, Kenya [MKK]; Tuscans in Italy [TSI]) [20]. All of the HapMap LCLs are available through the Coriell Cell Repositories [210].

The complete genomes of many of these HapMap samples will be sequenced as part of the 1000 Genomes Project in order to capture additional common and rare variants that were missed in the HapMap Project. In 2010, the pilot phase of the 1000 Genomes Project was published [23]; data released include low coverage whole-genome sequencing data of 179 individuals (CEU, YRI, CHB and JPT), high-coverage whole-genome sequencing data of two trios (CEU and YRI) and exon-targeted sequencing data of 697 individuals from seven populations (CEU, TSI, LWK, YRI, CHB, CHD and JPT) [203].

Another collection of LCLs that has been used in pharmacogenomics studies is the Human Variation Panel. LCLs derived from 100 African–Americans (AA), 100 Caucasians (CA), 100 Han Chinese (HCA) and 100 Mexican–Americans are available through Coriell [210]. All of the individuals in this collection are unrelated. These 400 samples have been genotyped for approximately 940,000 SNPs and investigators can apply for access to the genotypes through the Database of Genotypes and Phenotypes [204].

The HapMap and Human Variation Panel LCLs allow researchers to study the genetics of drug response in diverse world populations. Ancestry differences are often associated with variation in drug response [24]. Often these differences are found to be due to differing allele frequencies among populations in important pharmacokinetic or pharmacodynamic genes. For example, east Asians have much higher lung tumor-response rates upon treatment with an EGFR inhibitor than other cohorts, likely due to higher frequencies in east Asians of variants within *EGFR* that associate with lower gene-expression levels [25]. Nearly 90% of GWA studies have been completed in populations of European ancestry [26], and thus these LCLs provide an opportunity to perform GWA studies in non-European populations. Studies in populations of recent African ancestry are of particular importance since additional variants present at higher frequencies in African populations may be absent or rare in European and/or Asian populations [27,28]. It is unclear whether associations found in European and Asian populations can be consistently replicated in African populations: decreased linkage disequilibrium and gene–environment interactions could contribute to

lack of replication [28]. In one study, 49 SNPs that associated with high-density lipoprotein cholesterol levels in multiple European GWA and replication studies were tested for association in AA, American-Indians and Mexican-Americans, with 48, 45 and 57% of the tested SNPs replicating in the respective populations [29]. Overall, 16 SNPs were significantly associated with cholesterol levels across all three of the tested populations. The commonly associated SNPs are likely to be either linked to the functional SNP across populations or represent the functional SNP directly. For SNPs that do not replicate, linkage disequilibrium patterns, effect sizes and/or allele frequencies may be different across populations and thus it may be necessary to examine potentially functional gene regions in each population individually. Importantly, the inclusion of populations of diverse ancestry in genomic studies advances the goal of reducing health disparities [27].

In addition to the populations mentioned above, LCLs are available from over 40 additional world populations for use in pharmacogenomic studies. They are also available through Coriell and include multiple populations from South America, eastern Europe, the Middle East and the Pacific [210]. However, most of these additional populations have smaller sample sizes and have not been extensively genotyped.

Recent pharmacogenomic findings using the LCL model

Chemotherapeutics

Table 2 displays a summary of the pharmacogenomic results highlighted in this review. Most pharmacogenomic studies in the LCL model have focused on the genetics of chemotherapeutic susceptibility. Most chemotherapeutic agents produce a cytotoxic or apoptotic effect in immortalized cell lines, therefore measuring a pharmacologic phenotype in LCLs is straightforward. The drawback to using LCLs is that they are not tumor cells – a more appropriate system to study the role of genetic variation in tumor response. However, LCLs allow for the study of germline genetic variation that may contribute to drug toxicity and, to some extent, drug response. In a recent proof-of-principle study, Peters *et al.* used a two-stage approach to demonstrate that an association between genetic variation in *TYMS* and response to the chemotherapeutic 5-fluorouracil (5-FU) could be detected in the LCL model [30]. *TYMS* encodes thymidylate synthase, which functions to maintain the thymidine monophosphate pool necessary for DNA replication and repair; it is considered to be the main pharmacological target of 5-FU, a drug used primarily in the treatment of breast and colorectal cancers [31]. In this study, 5-FU-induced cytotoxicity was measured in 427 LCLs from the CEPH pedigrees. In the first stage of analysis the 46 SNPs in *TYMS* genotyped in the 30 family trios of HapMap Phase I/II were tested for association with cytotoxicity. Two of these SNPs were associated with 5-FU cytotoxicity ($p < 0.01$) in the discovery sample and were subsequently genotyped in the remainder of the CEPH pedigree LCLs for the second stage of analysis. Both SNPs remained significant when analyzed in the entire pedigree ($p < 0.001$) [30]. Thus, the LCL model was successful in detecting variants associated with 5-FU-induced cytotoxicity in a mechanistically supported candidate gene. The authors of this study caution that GWA studies to detect previously unknown variants associated with drug phenotypes in individual HapMap populations have limited statistical power due to the relatively small sample sizes [30]. One approach to reduce false positives is to incorporate additional sources of genomic information into the analysis, such as gene-expression data.

Although genome-wide gene-expression data has been previously included in linkage and association studies, the functional importance of incorporating gene-expression data into association studies of pharmacological phenotypes was recently demonstrated. Gamazon *et al.* showed that the top chemotherapeutic-induced cytotoxicity-associated SNPs from several GWA studies in HapMap CEU LCLs are more likely to be eQTLs than random sets of SNPs

in the genome, conditioned on minor allele frequency [32]. In addition, chemotherapy susceptibility SNPs are more likely to be eQTLs associated with the expression of ten or more genes, so called 'master regulators'. Interestingly, no enrichment was observed for nonsynonymous, untranslated region or splice site variants [32]. The eQTL data used in this study is publicly accessible through the SCAN database [2,205]. An important caveat is that all the SNPs in this analysis had minor allele frequencies >0.05 and different enrichment patterns may be detected for rare variants, once larger datasets become available. Fundamentally, this is important because based on the results of this enrichment study, variation in individual sensitivity to chemotherapeutic drugs may be more dependent on regulatory variants than traditionally studied exonic variants within candidate genes (i.e., DNA-repair genes and apoptosis-related genes) [32].

Demonstrating the potential importance of regulatory variation in pharmacokinetic genes, Matsson *et al.* tested the 45 ATP-binding cassette membrane transporter genes expressed in HapMap Phase I/II LCLs for *cis*-acting eQTLs [33]. ATP-binding cassette proteins transport various molecules, including many clinically important drugs, across extra- and intracellular membranes. Twenty four *cis*-eQTLs significantly associated with the expression of ten ATP-binding cassette genes and differential allelic expression was functionally confirmed for five of the 16 eQTLs tested via luciferase reporter assays [33].

As a means to incorporate gene-expression data as well as drug-induced cytotoxicity or other pharmacologic phenotypes, 'the triangle model' was first presented in 2007 [34]. The first arm of the triangle is a GWA study between SNPs and a particular drug-sensitivity measurement. Then in the second arm, eQTL analysis is performed on the most significant first arm GWA SNPs to find those that are also associated with the expression of transcripts. In the final arm, the expression of the eQTL target genes is tested for significant correlation to drug sensitivity [34].

The triangle model has been used to compare the pharmacogenomics of cytarabine arabinoside (ara-C) susceptibility between the HapMap Phase I/II CEU and YRI populations [35]. Ara-C is used predominantly in the treatment of acute myeloid leukemia and other hematologic malignancies; thus, LCLs are an excellent model system for this drug in particular. Using the triangle model, population-specific pharmacogenetic signatures consisting of four SNPs explaining 51% of the variability in ara-C cytotoxicity among the CEU, and five SNPs explaining 58% of the variation among the YRI were identified [35]. Associations between the expression of two genes (*GIT1* and *SLC25A37*) and ara-C were validated in an additional panel of LCLs. *GIT1* acts as a scaffold for certain intracellular signaling cascade proteins, including the MAP kinase pathway. Increased *GIT1* expression has been shown to increase MAP kinase signaling [36,37], which may result in increased apoptosis in response to ara-C [35]. *SLC25A37* is a member of the SLC25 solute carrier family. This carrier imports iron into mitochondria and is involved in heme biosynthesis [38]. In a study of leukemia cell lines, depletion of intracellular iron resulted in increased sensitivity to ara-C [39]. This ara-C LCL study also confirmed and extended the results of a previous study [40], finding that genetic variation in *DCK*, which encodes deoxycytidine kinase, influences its activity and expression. *DCK* phosphorylates ara-C to ara-CMP, the rate-limiting step in its activation to ara-CTP [40]. Genetic variation in *DCK* also predicts the variability observed in intracellular levels of the active metabolite ara-CTP [35].

The Pharmacogenomics And Cell database (PACdb) contains the results from this ara-C study, as well as the results from several other LCL-based pharmacogenomics studies, and thus provides a unique resource to the community [41,211]. Users can search for specific genes and SNPs of interest to determine if they have been found to associate with a particular drug phenotype. Currently, PACdb contains summary results for SNP genotype

versus cytotoxicity and gene expression versus cytotoxicity for cisplatin, carboplatin, daunorubicin, etoposide and ara-C, although it is likely to include additional drugs as data is published. PACdb also contains population-differential expression data and splicing-index data.

Following this, another group applied the triangle model to investigate the pharmacogenomics of ara-C, as well as the related drug gemcitabine, in LCLs from 60 CA, 54 AA and 60 HCA individuals of the Human Variation Panel [42]. Three eQTL target genes were selected from the triangle model analysis for functional validation. Knockdown of two of these genes, *VAV3* and *GPM6A*, was shown to desensitize both an LCL and a pancreatic cancer cell line to both gemcitabine and ara-C [42]. *VAV3* is a guanine nucleotide-exchange factor and a known proto-oncogene, and *GPM6A* is a membrane glycoprotein that might function as a chaperone in cancer cells; however, the mechanisms of their potential involvement in ara-C and gemcitabine response are unclear [42–44]. An additional study identified 102 copy number variants in these Human Variation Panel lines and found 11 copy number variants that associated with ara-C or gemcitabine cytotoxicity [45]. There was no overlap in the top SNPs reported from the triangle analysis of each ara-C study (35 SNPs in Hartford *et al.* and 30 SNPs in Li *et al.*), perhaps due to the limited sample sizes in both studies or to the different assays used to assess cytotoxicity [35,42].

Platinating agents, such as carboplatin and cisplatin, are used widely in the treatment of testicular, ovarian, lung, head and neck, and endometrial cancer. Another triangle model investigation took advantage of population differences in platinating-agent toxicity to identify predictive genetic variants [46]. First, the clinically observed result that east Asians are more sensitive to platinating agents was confirmed in HapMap LCLs [47]. The authors hypothesized that Asian populations might be enriched for platinating agent susceptibility SNPs and performed GWA studies for both carboplatin- and cisplatin-induced cytotoxicity using LCLs derived from the HapMap Phase I/II CHB and JPT populations. Top platinating agent SNPs were associated with the expression of multiple genes, highlighted by the histone H3 family, which was implicated in both drugs [46]. Platinating agents are known to form adducts on histones, which could act as a nuclear ‘reservoir’ of platinum compounds for further DNA-adduct formation [48]. Cisplatin treatment results in histone H3 phosphorylation in cancer cells, which is likely to affect chromatin stability [49]. Thirteen of the top SNPs also associated with either carboplatin or cisplatin cytotoxicity in a combined CEU and YRI population, suggesting cross-population effects [46].

A more comprehensive study to detect crosspopulation pharmacogenetic variants has recently been undertaken. Using six HapMap populations of diverse ancestry totaling 608 LCLs (comprising YRI, ASW, CEU, CHB, JPT and CHD), meta-analyses of over 3 million SNPs for both carboplatin- and cisplatin-induced cytotoxicity were performed [50]. A local ancestry approach was taken to account for the admixture in the ASW population [28]. Cisplatin-susceptibility SNPs were enriched for carboplatin-susceptibility SNPs. Because most of the variants that associate with platinum-induced cytotoxicity are polymorphic across multiple world populations, they could be tested in follow-up studies in diverse clinical populations. Seven genes previously implicated in platinating-agent response, including *BCL2*, *GSTM1*, *GSTT1*, *ERCC2* and *ERCC6* were also implicated in the meta-analyses [50]. Inhibition of apoptosis by increased *BCL2* expression has been shown to lead to cisplatin resistance [51]. *GSTM1* and *GSTT1* are glutathione *S*-transferases involved in the detoxification of platinating agents [52]. *ERCC2* and *ERCC6* are involved in nucleotide-excision repair, which is the primary mechanism of platinum DNA-adduct repair [52,53].

The majority of pharmacogenomic studies of chemotherapeutics have used drug-induced cytotoxicity as the phenotype; however, drug-induced apoptosis has emerged as an additional

suitable phenotype [54,55]. One pilot study showed apoptosis as a detectable phenotype for pharmacogenomic studies in LCLs with greater variation between siblings compared with variation between monozygotic twins after treatment with paclitaxel, cisplatin, carboplatin and ara-C [55]. Larger, more comprehensive studies are now underway.

In another type of study using LCLs, methotrexate polyglutamate (MTXPG) accumulation was measured in both leukemia cells from acute lymphoblastic leukemia patients and HapMap CEU and YRI LCLs. MTXPG is the active metabolite of the chemotherapeutic methotrexate. The expression of the top seven genes associated with MTXPG levels in the leukemia cells accounted for more variation in MTXPG (46%) than the expression of the top seven genes associated with MTXPG in the LCLs (20%) [56]. Thus, the authors used LCLs to show that acquired variation in leukemia cells has a stronger influence on MTXPG accumulation than inherited variation.

Radiation

Similar to chemotherapy, radiation is widely used in the treatment of cancer and response to treatment varies widely among individuals [57]. In a study by Smirnov *et al.*, LCLs from 15 CEPH pedigrees with seven to nine children each were used to study the genetics of radiation-induced changes in gene expression [58]. Global gene expression was measured in these lines at baseline and at 2 and 6 h after radiation exposure. Radiation-induced changes in gene expression compared with baseline were treated as quantitative phenotypes and for more than 1200 of these phenotypes there was significant evidence of linkage to specific chromosomal regions ($\log\text{-arithm}_{10}$ of odds >3.4) [58]. Interestingly, most ($>99\%$) of the radiation-induced expression phenotypes were *trans*-regulated (linkage peak was >5 mb from the gene or on a different chromosome) [58]. Several genes within linkage peaks contained SNPs that associated with radiation-induced cytotoxicity and apoptosis measured in a subset of the LCLs, further supporting their involvement in radiation response. Additional molecular analysis in LCLs showed that knockdown of potential regulators by siRNA followed by irradiation altered the expression of radiation-responsive target genes for five of the 11 regulators attempted [58]. Two genes implicated by these siRNA experiments, *JUN* and *FAS*, both well-known oncogenes, were also implicated in radiation response by a recent study of low-dose radiation (0.01–0.1 Gy)-induced gene-expression changes in LCLs [59], although a much higher dose (10 Gy) was used in the initial study [58].

A third group used 277 LCLs from the Human Variation Panel that included AA, CA and HCA to investigate the genomics of radiation response [60]. Rather than measure global gene expression after irradiation, Niu *et al.* used basal gene-expression levels on a genome-wide scale in order to identify potential biomarkers predictive of radiation response. Assays for radiation-induced cytotoxicity were also performed in these LCLs to achieve an integrated SNP-expression-cytotoxicity analysis similar to the triangle model described above for chemotherapy drugs. Twenty three of the top genes from this analysis were chosen for functional screening. Knockdown of the genes by siRNA in four cancer cell lines revealed five genes that significantly altered radiation sensitivity in at least two cancer cell lines [60]. The functionally validated genes from this study did not overlap with those reported previously by Smirnov *et al.*, which indicates there are likely to be many genes and pathways involved in radiation response. Importantly, different phenotypes (i.e., radiation-induced expression changes vs baseline expression and 24 h vs 72 h cytotoxicity assays) were used in Smirnov *et al.* and Niu *et al.*, respectively [58,60].

Statins

Although most studies have evaluated therapies that result in cellular cytotoxicity (i.e., chemotherapy and radiation), LCLs have also been used to study statins, a class of drugs

prescribed for the prevention and treatment of cardiovascular disease. Statins lower low-density lipoprotein (LDL) cholesterol levels [61]. Response to this drug class is variable among individuals and worse among AA compared with individuals of European ancestry only [62]. Statins are competitive inhibitors of HMGCR, which catalyzes the rate-limiting step of cholesterol biosynthesis [61]. Previously, Krauss and colleagues used LCLs to demonstrate that the association of the *HMGCR* haplotype H7 with reduced response to simvastatin was due to statin-induced expression of an alternatively spliced transcript of *HMGCR* lacking exon 13 [63,64]. The *HMGCR* association accounts for a relatively small proportion of the total variance in statin response and thus, additional candidates have been tested, including the LDL receptor (*LDLR*). Haplotype L5 of *LDLR* was associated with smaller simvastatin-induced lipid reductions in black patients, but not white patients, participating in the Cholesterol and Pharmacogenetics trial [65]. Functional studies in LCLs derived from the black participants demonstrated that simvastatin reduced the protein expression of LDLR in haplotype L5 carriers, but not in noncarriers [65]. Mangravite *et al.* also showed that the combined presence of both *LDLR* L5 and *HMGCR* H7 in black patients was associated with a significantly attenuated cholesterol-reduction response compared with either noncarriers for both haplotypes or noncarriers for one of the haplotypes [65]. The greater prevalence of these two haplotypes in black patients compared with white patients partially explains the observed ancestry differences in statin response.

Selective serotonin reuptake inhibitors

Selective serotonin reuptake inhibitors are the most commonly used first-line drug treatment for major depression, yet 30–40% of patients fail to improve after several weeks of treatment [66,67]. Morag *et al.* screened 80 female LCLs from the National Laboratory for the Genetics of Israeli Populations collection for growth inhibition by the selective serotonin reuptake inhibitor paroxetine [68]. Whole-genome expression profiling was subsequently performed on the seven most sensitive and seven most resistant LCLs and *CHLI* expression was 6.3-fold lower in the most sensitive lines, which was confirmed by reverse-transcription quantitative PCR [68]. *CHLI* encodes a neuronal cell adhesion protein, which plays a central role in neurogenesis and brain circuitry maintenance [69]. Supporting a role for the gene in mental illness, mutations in *CHLI* have been implicated in schizophrenia [70,71]. However, further studies are needed to confirm *CHLI* expression as a biomarker of paroxetine response.

Additional pharmacogenomic studies

In addition to the drugs discussed in the previous examples, several pharmacogenomic investigations of other drug classes have also employed the LCL model. For example, mycophenolic acid is an immunosuppressant given to patients receiving solid-organ transplants [72]. A study comparing baseline gene expression to mycophenolic acid cytotoxicity in LCLs identified 35 significantly associated genes, four of which were functionally verified in subsequent siRNA experiments [72]. In another study, LCLs were used in a resequencing study of the glucocorticoid receptor gene – the target of glucocorticoid steroid immunosuppressants such as dexamethasone and prednisone [73]. The gene was sequenced using DNA from 240 Human Variation Panel LCLs and 108 polymorphisms (57 novel) were identified. One 5'-upstream SNP was found to be associated with glucocorticoid receptor gene expression and receptor number in the LCLs [73]. In a discovery GWA study, SNPs associated with cytotoxicity-induced by the reactive metabolite of acetaminophen in LCLs were identified [74]. Although acetaminophen is an over-the-counter pain reliever, its reactive metabolite can lead to acute liver failure in some patients. Further studies are needed to confirm these SNP associations and to elucidate their functions. Another pharmacogenetic study investigated β -blockers, which downregulate β -adrenergic receptor signaling to treat hypertension. This downregulation is mediated through

G protein-coupled receptors such as GRK2 (encoded by the gene *ADRBK1*) [75]. *ADRBK1* regions were sequenced in 48 individuals from the Human Variation Panel of LCLs. Putative functional SNPs were tested for mRNA expression differences in 96 LCLs and 12 leukocyte samples from hypertensives; however, no eQTLs were identified [76].

Clinical translation of LCL findings

The goal of pharmacogenomic studies is the ability to identify patients more likely to benefit from a particular treatment or those more likely to experience adverse events. Of fundamental importance is an understanding of how well LCLs recapitulate clinical data. Recently, a few studies have demonstrated the relevance of the LCL model by replicating findings from LCLs in patient cohorts. For example, variation in the expression of *FKBP5* was found to associate with ara-C-induced cytotoxicity in a genome-wide screen in LCLs and knockdown experiments in tumor cell lines confirmed the relationship [77]. *FKBP5* promotes the dephosphorylation of AKT, which can contribute to increased cytotoxicity by enhancing apoptosis in response to chemotherapeutic agents such as ara-C [78]. In a follow-up clinical study, SNPs within *FKBP5* were genotyped in a cohort of 187 pediatric acute myeloid leukemia patients treated with ara-C and two of the SNPs associated with both event-free and overall survival ($p < 0.05$) [79].

Asparaginase is used in the treatment of childhood acute lymphoblastic leukemia and a genome-wide approach was recently used to find genes and variants associated with cytotoxicity in asparaginase-treated HapMap LCLs [80]. Plausibly, the top associations from the unbiased genome-wide analysis were enriched for the aspartate metabolism pathway. The association of SNPs in aspartate metabolism genes with asparaginase IC₅₀ was confirmed in primary bone marrow samples from 54 patients with childhood acute lymphoblastic leukemia [80].

In another study, top SNPs from a triangle model analysis of carboplatin-induced cytotoxicity in LCLs were genotyped in 60 head and neck cancer patients treated with carboplatin [81]. Two linked SNPs associated with overall response to carboplatin-based induction chemotherapy in the patients (false discovery rate < 0.05). Hinting at a potential regulation mechanism, these two SNPs are associated with the expression of 20 genes in LCLs, including *SLC22A5* and *SLCO4C1*, which are organic cation/anion transporters known to affect platinating-agent uptake and clearance [81]. A similar study tested a top SNP found in an analysis of carboplatin-treated LCLs for association with outcome in patients from the Australian Ovarian Cancer Study [82]. The SNP rs1649942, which is a master regulator associated with the expression of 39 genes [32], associated with both progression-free and overall survival in 377 ovarian cancer patients treated with carboplatin and paclitaxel ($p < 0.05$) [82].

In another study, 157 top SNPs from a GWA study of cisplatin-induced cytotoxicity in Human Variation Panel LCLs were tested for association with overall survival in lung cancer patients [83]. Although not significant after Bonferroni correction, nine and ten SNPs associated with overall survival in patients treated with platinum for non-small-cell lung cancer and small-cell lung cancer, respectively ($p < 0.05$). Two genes targeted by these SNPs *in trans* were functionally validated through siRNA experiments in lung cancer lines: *DAPK3*, which is a leucine zipper-interacting kinase that acts as a tumor suppressor and regulates apoptosis, and *METTL6*, which belongs to the methyltransferase superfamily and might influence methylation-dependent chemoresistance [83].

These ara-C, asparaginase and platinum studies are promising because unbiased genome-wide discovery studies in LCLs were replicated in patient cohorts. However, all five studies

had relatively small patient sample sizes, so confirmation of the associations in larger cohorts would greatly strengthen the findings. Another concern is that the top findings in the cisplatin cytotoxicity GWA study performed in Human Variation Panel LCLs [83] do not overlap with the top findings of the same drug in HapMap LCLs [50]. Since neither study had genome-wide significant results, the lack of overlap could be arising from limited sample sizes. In addition, differences in linkage disequilibrium across populations as well as the use of different cytotoxicity assays between studies could lead to different SNP associations. Differences in nongenetic factors such as baseline growth rates could also account for some of the differences between the studies [6]. Combining multiple LCL results in a meta-analysis would help to alleviate potential power issues due to limited sample sizes. Streamlining cytotoxicity assay protocols and confounding variable measurements between laboratories may be necessary. In addition, follow-up functional studies of top candidates in LCLs and other relevant primary or tumor cell lines can help filter out false positives.

Accruing large patient cohorts receiving the same drug regimen for GWA studies is challenging, particularly in oncology. Even more difficult is performing a replication GWA in patients. Therefore, genetic variants in LCLs can be used as a discovery or replication set with clinical patient data. Top variants from an initial GWA study in LCLs can be tested for association in the patient samples as discussed previously. Alternatively, LCLs can be used to verify GWA findings from a clinical cohort and/or to perform follow-up functional studies in LCLs on the SNPs of interest. These studies are useful even for SNPs that do not reach genome-wide significance, perhaps due to limited sample size. This latter approach was used successfully to identify functionally relevant variants associated with musculoskeletal adverse events in white women receiving aromatase inhibitors for breast cancer treatment [84]. The top SNP (rs11849538) p-value was 6.67×10^{-7} , close to the genome-wide significance threshold of 1×10^{-7} , and the SNP was located 926 bp downstream of the *TCL1A* gene. Microarray data from the Human Variation Panel LCLs showed that this gene is highly and variably expressed. Interestingly, although rs11849538 is linked to many other SNPs, its minor allele was predicted to create an estrogen-response element by the TRANSFAC database [85,212]. Chromatin immunoprecipitation experiments in LCLs of known rs11849538 genotype transfected with ER α confirmed that ER α bound to the minor allele sequence, but not to the major allele. In addition, in LCLs following ER α transfection, expression of *TCL1A* was higher in those with the minor allele than in those without [84]. Potential designs for incorporating LCL and clinical data in studies of pharmacogenomic phenotypes are presented in Figure 1.

Future perspective

In the past few years, the use of LCLs in candidate gene and GWA studies of pharmacological phenotypes has expanded thanks to the availability of genotype data for many commercially available lines from diverse world populations. Discoveries made in LCLs will need to be tested for replication in clinical datasets and a few early examples have been successful [79–83]. LCLs will continue to be a useful model for follow-up functional studies of genes and variants found to be associated with pharmacological phenotypes in both clinical cohorts and LCL discovery studies [35,42,58,65,84].

As next-generation sequencing technologies become more affordable, the entire genomes of LCL panels and patient cohorts will be sequenced. Indeed, the 1000 Genomes Project aims to sequence at least 1000 individuals from the HapMap Project and make the data publicly available [23,86]. As reviewed here, pharmacological phenotypes have already been collected for many of these LCLs; however, initial analyses have been limited to the common SNP variation collected by the HapMap Project. Both novel common and novel

rare variants have been discovered in the pilot phase of the 1000 Genomes Project [23,203]; thus, reanalysis of already collected phenotypes may be fruitful. Data from the 1000 Genomes Project will also allow more comprehensive imputation, especially in populations of African descent, where most of the novel variants are being discovered.

Methods that combine rare variants from multiple individuals that are in the same gene into a score that can be tested for association with different pharmacological phenotypes are being developed. Some of these methods take into account information about whether a rare variant causes a stop codon, a nonconservative amino acid change or affects gene expression [87–90]. The importance of eQTLs in pharmacological phenotypes has been demonstrated [32], thus, it will be important to consider SNPs in regions outside of exons. In addition to scoring individual genes, combining information from molecular pathways, including pharmacokinetic and pharmacodynamic pathways, may elucidate additional associations. If differences in a particular gene affect a drug-response phenotype, it is probable that other genes in the same pathway also affect the phenotype. Therefore, rare variants in different genes of the same pathway could be combined to test for association of the pathway with a pharmacological trait of interest.

Researchers have used the LCL model to identify genes and variants that predict drug response and toxicity. In addition to extensive DNA-sequencing data becoming more widely available in the near future, other sources of genome-wide data, including methylation patterns, both coding and ncRNA expression, protein expression and protein–DNA-binding patterns in LCLs will be incorporated into pharmacogenomic studies. Similar to initial global mRNA studies, many initial investigations of miRNA [91,92] and DNA methylation [93,94] have been performed in LCLs. Differences in transcriptional regulation, including those caused by miRNA and DNA methylation, are likely to account for some of the variation in pharmacological phenotypes observed among individuals [95–97]. In addition, variable protein levels and post-translational modifications, which can be measured by newly developed high-throughput assays such as the microwestern array [98], are also likely to play functional roles in drug response and will be incorporated into pharmacogenomic studies. Similar to the enrichment of eQTLs found in chemotherapeutic-associated SNPs discovered in LCLs, clinical trait-associated SNPs are also enriched for eQTLs [99,100]. Thus, variants associated with complex traits may be more likely to alter the amount or timing of protein production than actually change the type of protein produced. Comparing drug-related GWA results to susceptibility GWA results of the diseases the drug is used to treat may elucidate pathways involved in both disease progression and treatment.

Replication of LCL findings in clinical studies in patients who experience a response or an adverse event are critical for understanding which drugs and clinical phenotypes are best represented within the LCL model. LCLs used in follow-up studies can help determine the function of a particular genetic variant in affecting the phenotype measured, potentially revealing new drug targets. Importantly, some significant associations may have small effect sizes and thus may not be clinically actionable, but will still inform researchers about drug mechanisms of action. However, algorithms that combine multiple genetic variants of small effect to explain a larger proportion of the variance could potentially be useful in the clinic. Once an association between a variant or variants and drug-induced phenotype is confidently established, basic researchers and clinicians will need to assess the potential benefits to patients of implementing a test for a particular variant or combination of variants prior to treatment. The Pharmacogenomics Research Network [213] routinely publishes a dynamic series of gene-based drug-dosing guidelines through the Clinical Pharmacogenetics Implementation Consortium [101]. For these guidelines to improve care, full clinical implementation will require widespread provider education, acceptance and automated decision support. Combinations of genetic variants that explain a large proportion of the

variance of a particular pharmacological phenotype will be the best candidates for clinical implementation, thus making personalized medicine a reality.

References

Papers of special note have been highlighted as:

▪ of interest

▪▪ of considerable interest

- 1▪. Welsh M, Mangravite L, Medina MW, et al. Pharmacogenomic discovery using cell-based models. *Pharmacol Rev.* 2009; 61(4):413–429. Comprehensive review of pharmacogenomic findings in lymphoblastoid cell lines (LCLs) prior to 2009. [PubMed: 20038569]
- 2▪. Gamazon ER, Zhang W, Konkashbaev A, et al. SCAN: SNP and copy number annotation. *Bioinformatics.* 2010; 26(2):259–262. Useful database that hosts expression quantitative trait loci and copy number variation–quantitative trait loci results in LCLs. [PubMed: 19933162]
3. Zhou SF, Liu JP, Chowbay B. Polymorphism of human cytochrome P450 enzymes and its clinical impact. *Drug Metab Rev.* 2009; 41(2):89–295. [PubMed: 19514967]
4. van Baarsen LG, Vosslander S, Tijssen M, et al. Pharmacogenomics of interferon-beta therapy in multiple sclerosis: baseline IFN signature determines pharmacological differences between patients. *PLoS ONE.* 2008; 3(4):e1927. [PubMed: 18382694]
5. Wertz IE, Kusam S, Lam C, et al. Sensitivity to antitubulin chemotherapeutics is regulated by *MCL1* and *FBW7*. *Nature.* 2011; 471(7336):110–114. [PubMed: 21368834]
6. Choy E, Yelensky R, Bonakdar S, et al. Genetic analysis of human traits *in vitro*: drug response and gene expression in lymphoblastoid cell lines. *PLoS Genet.* 2008; 4(11):e1000287. [PubMed: 19043577]
7. Stark AL, Zhang W, Mi S, et al. Heritable and non-genetic factors as variables of pharmacologic phenotypes in lymphoblastoid cell lines. *Pharmacogenomics J.* 2010; 10(6):505–512. [PubMed: 20142840]
8. Liu MT, Chen YR, Chen SC, et al. Epstein–Barr virus latent membrane protein 1 induces micronucleus formation, represses DNA repair and enhances sensitivity to DNA-damaging agents in human epithelial cells. *Oncogene.* 2004; 23(14):2531–2539. [PubMed: 14716302]
9. Caliskan M, Cusanovich DA, Ober C, Gilad Y. The effects of EBV transformation on gene expression levels and methylation profiles. *Hum Mol Genet.* 2011; 20(8):1643–1652. [PubMed: 21289059]
10. Bullaughey K, Chavarria CI, Coop G, Gilad Y. Expression quantitative trait loci detected in cell lines are often present in primary tissues. *Hum Mol Genet.* 2009; 18(22):4296–4303. [PubMed: 19671653]
11. Ding J, Gudjonsson JE, Liang L, et al. Gene expression in skin and lymphoblastoid cells: refined statistical method reveals extensive overlap in *cis*-eQTL signals. *Am J Hum Genet.* 2010; 87(6):779–789. [PubMed: 21129726]
12. Schadt EE, Molony C, Chudin E, et al. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol.* 2008; 6(5):e107. [PubMed: 18462017]
13. Dausset J, Cann H, Cohen D, Lathrop M, Lalouel JM, White R. Centre d'Etude du Polymorphisme Humain (CEPH): collaborative genetic mapping of the human genome. *Genomics.* 1990; 6(3):575–577. [PubMed: 2184120]
14. Thorisson GA, Stein LD. The SNP Consortium website: past, present and future. *Nucleic Acids Res.* 2003; 31(1):124–127. [PubMed: 12519964]
15. Morley M, Molony CM, Weber TM, et al. Genetic analysis of genome-wide variation in human gene expression. *Nature.* 2004; 430(7001):743–747. [PubMed: 15269782]
16. Burdick JT, Chen WM, Abecasis GR, Cheung VG. *In silico* method for inferring genotypes in pedigrees. *Nat Genet.* 2006; 38(9):1002–1004. [PubMed: 16921375]

17. Dolan ME, Newbold KG, Nagasubramanian R, et al. Heritability and linkage analysis of sensitivity to cisplatin-induced cytotoxicity. *Cancer Res.* 2004; 64(12):4353–4356. [PubMed: 15205351]
18. Watters JW, Kraja A, Meucci MA, Province MA, McLeod HL. Genome-wide discovery of loci influencing chemotherapy cytotoxicity. *Proc Natl Acad Sci USA.* 2004; 101(32):11809–11814. [PubMed: 15282376]
19. Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P. A haplotype map of the human genome. *Nature.* 2005; 437(7063):1299–1320. [PubMed: 16255080]
20. Altshuler DM, Gibbs RA, Peltonen L, et al. Integrating common and rare genetic variation in diverse human populations. *Nature.* 2010; 467(7311):52–58. Discusses HapMap Phase III findings. [PubMed: 20811451]
21. Frazer KA, Ballinger DG, Cox DR, et al. A second generation human haplotype map of over 3.1 million SNPs. *Nature.* 2007; 449(7164):851–861. Discusses HapMap Phase I/II findings. [PubMed: 17943122]
22. Thorisson GA, Smith AV, Krishnan L, Stein LD. The International HapMap Project Web site. *Genome Res.* 2005; 15(11):1592–1593. [PubMed: 16251469]
23. The 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature.* 2010; 467(7319):1061–1073. Describes novel variation uncovered by the 1000 Genomes Project, which may lead to new pharmacogenomic discoveries in the near future. [PubMed: 20981092]
24. O'Donnell PH, Dolan ME. Cancer pharmacogenetics: ethnic differences in susceptibility to the effects of chemotherapy. *Clin Cancer Res.* 2009; 15(15):4806–4814. [PubMed: 19622575]
25. Calvo E, Baselga J. Ethnic differences in response to epidermal growth factor receptor tyrosine kinase inhibitors. *J Clin Oncol.* 2006; 24(14):2158–2163. [PubMed: 16682734]
26. Rotimi CN, Jorde LB. Ancestry and disease in the age of genomic medicine. *N Engl J Med.* 2010; 363(16):1551–1558. [PubMed: 20942671]
27. Need AC, Goldstein DB. Next generation disparities in human genomics: concerns and remedies. *Trends Genet.* 2009; 25(11):489–494. [PubMed: 19836853]
28. Wheeler HE, Gorsic LK, Welsh M, et al. Genome-wide local ancestry approach identifies genes and variants associated with chemotherapeutic susceptibility in African-Americans. *PLoS ONE.* 2011; 6(7):e21920. [PubMed: 21755009]
29. Dumitrescu L, Carty CL, Taylor K, et al. Genetic determinants of lipid traits in diverse populations from the Population Architecture using Genomics and Epidemiology (PAGE) study. *PLoS Genet.* 2011; 7(6):e1002138. [PubMed: 21738485]
30. Peters EJ, Kraja AT, Lin SJ, et al. Association of thymidylate synthase variants with 5-fluorouracil cytotoxicity. *Pharmacogenet Genomics.* 2009; 19(5):399–401. [PubMed: 19339911]
31. Pinedo HM, Peters GF. Fluorouracil: biochemistry and pharmacology. *J Clin Oncol.* 1988; 6(10):1653–1664. [PubMed: 3049954]
32. Gamazon ER, Huang RS, Cox NJ, Dolan ME. Chemotherapeutic drug susceptibility associated SNPs are enriched in expression quantitative trait loci. *Proc Natl Acad Sci USA.* 2010; 107(20):9287–9292. [PubMed: 20442332]
33. Matsson P, Yee SW, Markova S, et al. Discovery of regulatory elements in human ATP-binding cassette transporters through expression quantitative trait mapping. *Pharmacogenomics J.* 2011 (Epub ahead of print). 10.1038/tpj.2011.8
34. Huang RS, Duan S, Bleibel WK, et al. A genome-wide approach to identify genetic variants that contribute to etoposide-induced cytotoxicity. *Proc Natl Acad Sci USA.* 2007; 104(23):9758–9763. First study to describe the triangle model, integrating genetic variation, gene expression and pharmacological phenotypes in LCLs. [PubMed: 17537913]
35. Hartford CM, Duan S, Delaney SM, et al. Population-specific genetic variants important in susceptibility to cytarabine arabinoside cytotoxicity. *Blood.* 2009; 113(10):2145–2153. [PubMed: 19109566]
36. Yin G, Zheng Q, Yan C, Berk BC. GIT1 is a scaffold for ERK1/2 activation in focal adhesions. *J Biol Chem.* 2005; 280(30):27705–27712. [PubMed: 15923189]
37. Lu Z, Xu S. ERK1/2 MAP kinases in cell survival and apoptosis. *IUBMB Life.* 2006; 58(11):621–631. [PubMed: 17085381]

38. Shaw GC, Cope JJ, Li L, et al. Mitoferrin is essential for erythroid iron assimilation. *Nature*. 2006; 440(7080):96–100. [PubMed: 16511496]
39. Leardi A, Caraglia M, Selleri C, et al. Desferioxamine increases iron depletion and apoptosis induced by ara-C of human myeloid leukaemic cells. *Br J Haematol*. 1998; 102(3):746–752. [PubMed: 9722302]
40. Lamba JK, Crews K, Pounds S, et al. Pharmacogenetics of deoxycytidine kinase: identification and characterization of novel genetic variants. *J Pharmacol Exp Ther*. 2007; 323(3):935–945. [PubMed: 17855478]
41. Gamazon ER, Duan S, Zhang W, et al. PACdb: a database for cell-based pharmacogenomics. *Pharmacogenet Genomics*. 2010; 20(4):269–273. Useful database that hosts cytotoxicity genome-wide association (GWA) results in LCLs. [PubMed: 20216476]
42. Li L, Fridley BL, Kalari K, et al. Gemcitabine and arabinosylcytosin pharmacogenomics: genome-wide association and drug response biomarkers. *PLoS ONE*. 2009; 4(11):e7765. [PubMed: 19898621]
43. Shin BK, Wang H, Yim AM, et al. Global profiling of the cell surface proteome of cancer cells uncovers an abundance of proteins with chaperone function. *J Biol Chem*. 2003; 278(9):7607–7616. [PubMed: 12493773]
44. Trenkle T, McClelland M, Adlkofer K, Welsh J. Major transcript variants of VAV3, a new member of the VAV family of guanine nucleotide exchange factors. *Gene*. 2000; 245(1):139–149. [PubMed: 10713454]
45. Kalari KR, Hebring SJ, Chai HS, et al. Copy number variation and cytidine analogue cytotoxicity: a genome-wide association approach. *BMC Genomics*. 2010; 11:357. [PubMed: 20525348]
46. O'Donnell PH, Gamazon E, Zhang W, et al. Population differences in platinum toxicity as a means to identify novel genetic susceptibility variants. *Pharmacogenet Genomics*. 2010; 20(5):327–337. [PubMed: 20393316]
47. Millward MJ, Boyer MJ, Lehnert M, et al. Docetaxel and carboplatin is an active regimen in advanced non-small-cell lung cancer: a Phase II study in Caucasian and Asian patients. *Ann Oncol*. 2003; 14(3):449–454. [PubMed: 12598352]
48. Wu B, Droge P, Davey CA. Site selectivity of platinum anticancer therapeutics. *Nat Chem Biol*. 2008; 4(2):110–112. [PubMed: 18157123]
49. Wang D, Lippard SJ. Cisplatin-induced post-translational modification of histones H3 and H4. *J Biol Chem*. 2004; 279(20):20622–20625. [PubMed: 15010460]
50. Wheeler HE, Gamazon ER, Stark AL, et al. Genome-wide meta-analysis identifies variants associated with platinating agent susceptibility across populations. *Pharmacogenomics J*. 2011 (Epub ahead of print). Largest discovery GWA study in LCLs to date with over 600 LCLs. 10.1038/tpj.2011.38
51. Eliopoulos AG, Kerr DJ, Herod J, et al. The control of apoptosis and drug resistance in ovarian cancer: influence of *p53* and *Bcl-2*. *Oncogene*. 1995; 11(7):1217–1228. [PubMed: 7478541]
52. Marsh S, McLeod H, Dolan E, et al. Platinum pathway. *Pharmacogenet Genomics*. 2009; 19(7): 563–564. [PubMed: 19525887]
53. Rabik CA, Dolan ME. Molecular mechanisms of resistance and toxicity associated with platinating agents. *Cancer Treat Rev*. 2007; 33(1):9–23. [PubMed: 17084534]
54. Shukla SJ, Duan S, Badner JA, Wu X, Dolan ME. Susceptibility loci involved in cisplatin-induced cytotoxicity and apoptosis. *Pharmacogenet Genomics*. 2008; 18(3):253–262. [PubMed: 18300947]
55. Wen Y, Gorsic LK, Wheeler HE, Ziliak DM, Huang SR, Dolan EM. Chemotherapeutic-induced apoptosis: a phenotype for pharmacogenomics studies. *Pharmacogenet Genomics*. 2011; 21(8): 476–488. [PubMed: 21642893]
56. French D, Yang W, Cheng C, et al. Acquired variation outweighs inherited variation in whole genome analysis of methotrexate polyglutamate accumulation in leukemia. *Blood*. 2009; 113(19): 4512–4520. [PubMed: 19066393]
57. West CM, Elliott RM, Burnet NG. The genomics revolution and radiotherapy. *Clin Oncol (R Coll Radiol)*. 2007; 19(6):470–480. [PubMed: 17419040]
58. Smirnov DA, Morley M, Shin E, Spielman RS, Cheung VG. Genetic analysis of radiation-induced changes in human gene expression. *Nature*. 2009; 459(7246):587–591. [PubMed: 19349959]

59. Wyrobek AJ, Manohar CF, Krishnan VV, et al. Low dose radiation response curves, networks and pathways in human lymphoblastoid cells exposed from 1 to 10cGy of acute gamma radiation. *Mutat Res.* 2011; 722(2):119–130. [PubMed: 21497671]
60. Niu N, Qin Y, Fridley BL, et al. Radiation pharmacogenomics: a genome-wide association approach to identify radiation response biomarkers using human lymphoblastoid cell lines. *Genome Res.* 2010; 20(11):1482–1492. [PubMed: 20923822]
61. LaRosa JC, He J, Vupputuri S. Effect of statins on risk of coronary disease: a meta-analysis of randomized controlled trials. *JAMA.* 1999; 282(24):2340–2346. [PubMed: 10612322]
62. Simon JA, Lin F, Hulley SB, et al. Phenotypic predictors of response to simvastatin therapy among African-Americans and Caucasians: the Cholesterol and Pharmacogenetics (CAP) Study. *Am J Cardiol.* 2006; 97(6):843–850. [PubMed: 16516587]
63. Krauss RM, Mangravite LM, Smith JD, et al. Variation in the 3-hydroxyl-3-methylglutaryl coenzyme A reductase gene is associated with racial differences in low-density lipoprotein cholesterol response to simvastatin treatment. *Circulation.* 2008; 117(12):1537–1544. [PubMed: 18332269]
64. Medina MW, Gao F, Ruan W, Rotter JJ, Krauss RM. Alternative splicing of 3-hydroxy-3-methylglutaryl coenzyme A reductase is associated with plasma low-density lipoprotein cholesterol response to simvastatin. *Circulation.* 2008; 118(4):355–362. [PubMed: 18559695]
65. Mangravite LM, Medina MW, Cui J, et al. Combined influence of *LDLR* and *HMGCR* sequence variation on lipid-lowering response to simvastatin. *Arterioscler Thromb Vasc Biol.* 2010; 30(7):1485–1492. [PubMed: 20413733]
66. Souery D, Oswald P, Massat I, et al. Clinical factors associated with treatment resistance in major depressive disorder: results from a European multicenter study. *J Clin Psychiatry.* 2007; 68(7):1062–1070. [PubMed: 17685743]
67. Kirchheiner J, Nickchen K, Bauer M, et al. Pharmacogenetics of antidepressants and antipsychotics: the contribution of allelic variations to the phenotype of drug response. *Mol Psychiatry.* 2004; 9(5):442–473. [PubMed: 15037866]
68. Morag A, Pasmanik-Chor M, Oron-Karni V, Rehavi M, Stingl JC, Gurwitz D. Genome-wide expression profiling of human lymphoblastoid cell lines identifies *CHLI* as a putative SSRI antidepressant response biomarker. *Pharmacogenomics.* 2011; 12(2):171–184. [PubMed: 21332311]
69. Montag-Sallaz M, Baarke A, Montag D. Aberrant neuronal connectivity in *CHLI*-deficient mice is associated with altered information processing-related immediate early gene expression. *J Neurobiol.* 2003; 57(1):67–80. [PubMed: 12973829]
70. Chen QY, Chen Q, Feng GY, et al. Case-control association study of the close homologue of L1 (*CHLI*) gene and schizophrenia in the Chinese population. *Schizophr Res.* 2005; 73(2–3):269–274. [PubMed: 15653271]
71. Sakurai K, Migita O, Toru M, Arinami T. An association between a missense polymorphism in the close homologue of L1 (*CHLI*, *CALL*) gene and schizophrenia. *Mol Psychiatry.* 2002; 7(4):412–415. [PubMed: 11986985]
72. Wu TY, Fridley BL, Jenkins GD, Batzler A, Wang L, Weinshilboum RM. Mycophenolic acid response biomarkers: a cell line model system-based genome-wide screen. *Int Immunopharmacol.* 2011; 11(8):1057–1064. [PubMed: 21396482]
73. Niu N, Manickam V, Kalari KR, et al. Human glucocorticoid receptor alpha gene (*NR3C1*) pharmacogenomics: gene resequencing and functional genomics. *J Clin Endocrinol Metab.* 2009; 94(8):3072–3084. [PubMed: 19435830]
74. Moyer AM, Fridley BL, Jenkins GD, et al. Acetaminophen-NAPQI hepatotoxicity: a cell line model system genome-wide association study. *Toxicol Sci.* 2011; 120(1):33–41. [PubMed: 21177773]
75. Penela P, Murga C, Ribas C, Tutor AS, Peregrin S, Mayor F Jr. Mechanisms of regulation of G protein-coupled receptor kinases (GRKs) and cardiovascular disease. *Cardiovasc Res.* 2006; 69(1):46–56. [PubMed: 16288730]

76. Lobmeyer MT, Wang L, Zineh I, et al. Polymorphisms in genes coding for GRK2 and GRK5 and response differences in antihypertensive-treated patients. *Pharmacogenet Genomics*. 2011; 21(1): 42–49. [PubMed: 21127457]
77. Li L, Fridley B, Kalari K, et al. Gemcitabine and cytosine arabinoside cytotoxicity: association with lymphoblastoid cell expression. *Cancer Res*. 2008; 68(17):7050–7058. [PubMed: 18757419]
78. Pei H, Li L, Fridley BL, et al. *FKBP51* affects cancer cell response to chemotherapy by negatively regulating Akt. *Cancer Cell*. 2009; 16(3):259–266. [PubMed: 19732725]
- 79••. Mitra AK, Crews K, Pounds S, et al. Impact of genetic variation in *FKBP5* on clinical response in pediatric acute myeloid leukemia patients: a pilot study. *Leukemia*. 2011; 25(8):1354–1356. One of the first replications of an LCL pharmacogenetic discovery in patient samples. [PubMed: 21483441]
- 80••. Chen SH, Yang W, Fan Y, et al. A genome-wide approach identifies that the aspartate metabolism pathway contributes to asparaginase sensitivity. *Leukemia*. 2011; 25(1):66–74. One of the first replications of an LCL pharmacogenetic discovery in patient samples. [PubMed: 21072045]
- 81••. Ziliak D, O'Donnell PH, Im HK, et al. Germline polymorphisms discovered via a cell-based, genome-wide approach predict platinum response in head and neck cancers. *Transl Res*. 2011; 157(5):265–272. One of the first replications of an LCL pharmacogenetic discovery in patient samples. [PubMed: 21497773]
- 82••. Huang RS, Johnatty SE, Gamazon ER, 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(16):1–11. One of the first replications of an LCL pharmacogenetic discovery in patient samples.
- 83••. Tan XL, Moyer AM, Fridley BL, et al. Genetic variation predicting cisplatin cytotoxicity associated with overall survival in lung cancer patients receiving platinum-based chemotherapy. *Clin Cancer Res*. 2011; 17(17):5801–5811. One of the first replications of an LCL pharmacogenetic discovery in patient samples. [PubMed: 21775533]
- 84••. Ingle JN, Schaid DJ, Goss PE, et al. Genome-wide associations and functional genomic studies of musculoskeletal adverse events in women receiving aromatase inhibitors. *J Clin Oncol*. 2010; 28(31):4674–4682. Example of how LCLs can be used to functionally validate findings from patient GWA studies. [PubMed: 20876420]
85. Matys V, Fricke E, Geffers R, et al. TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res*. 2003; 31(1):374–378. [PubMed: 12520026]
86. Kaiser J. DNA sequencing. A plan to capture human diversity in 1000 genomes. *Science*. 2008; 319(5862):395. [PubMed: 18218868]
87. Li B, Leal SM. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *Am J Hum Genet*. 2008; 83(3):311–321. [PubMed: 18691683]
- 88•. Liu JZ, McRae AF, Nyholt DR, et al. A versatile gene-based test for genome-wide association studies. *Am J Hum Genet*. 2010; 87(1):139–145. Describes a gene-based GWA approach that combines SNPs within a gene to test for phenotypic association. [PubMed: 20598278]
89. Tatonetti NP, Dudley JT, Sagreiya H, Butte AJ, Altman RB. An integrative method for scoring candidate genes from association studies: application to warfarin dosing. *BMC Bioinformatics*. 2010; 11(Suppl 9):S9. [PubMed: 21044367]
90. Zhu X, Feng T, Li Y, Lu Q, Elston RC. Detecting rare variants for complex traits using family and unrelated data. *Genet Epidemiol*. 2010; 34(2):171–187. [PubMed: 19847924]
91. Gamazon ER, Im HK, Duan S, et al. Exptarget: an integrative approach to predicting human microRNA targets. *PLoS ONE*. 2010; 5(10):e13534. [PubMed: 20975837]
92. Wang L, Oberg AL, Asmann YW, et al. Genome-wide transcriptional profiling reveals microRNA-correlated genes and biological processes in human lymphoblastoid cell lines. *PLoS ONE*. 2009; 4(6):e5878. [PubMed: 19517021]
93. Bell JT, Pai AA, Pickrell JK, et al. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biol*. 2011; 12(1):R10. [PubMed: 21251332]

94. Sugawara H, Iwamoto K, Bundo M, Ueda J, Ishigooka J, Kato T. Comprehensive DNA methylation analysis of human peripheral blood leukocytes and lymphoblastoid cell lines. *Epigenetics*. 2011; 6(4):508–515. [PubMed: 21304276]
95. Baer-Dubowska W, Majchrzak-Celinska A, Cichocki M. Pharmacoeugenetics: a new approach to predicting individual drug responses and targeting new drugs. *Pharmacol Rep*. 2011; 63(2):293–304. [PubMed: 21602587]
96. Mishra PJ, Banerjee D, Bertino JR. MiRSNPs or MiR-polymorphisms, new players in microRNA mediated regulation of the cell: introducing microRNA pharmacogenomics. *Cell Cycle*. 2008; 7(7):853–858. [PubMed: 18414050]
97. Zhang W, Huang RS, Dolan ME. Integrating epigenomics into pharmacogenomic studies. *Pharmgenomics Pers Med*. 2008; 2008(1):7–14. [PubMed: 20622972]
98. Ciaccio MF, Wagner JP, Chu CP, Lauffenburger DA, Jones RB. Systems analysis of EGF receptor signaling dynamics with microwestern arrays. *Nat Methods*. 2010; 7(2):148–155. [PubMed: 20101245]
99. Nica AC, Montgomery SB, Dimas AS, et al. Candidate causal regulatory effects by integration of expression QTLs with complex trait genetic associations. *PLoS Genet*. 2010; 6(4):e1000895. [PubMed: 20369022]
100. Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet*. 2010; 6(4):e1000888. [PubMed: 20369019]
101. Relling MV, Gardner EE, Sandborn WJ, et al. Clinical Pharmacogenetics Implementation Consortium guidelines for thiopurine methyltransferase genotype and thiopurine dosing. *Clin Pharmacol Ther*. 2011; 89(3):387–391. [PubMed: 21270794]

Websites

201. The SNP Consortium Database. <http://snpdata.cshl.edu>
202. International HapMap Project. www.hapmap.org
203. The 1000 Genomes Project. www.1000genomes.org
204. Human Variation Panel dbGaP Access. www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000211.v1.p1
205. SCAN Database. www.scandb.org
206. Gene Expression Omnibus (CEPH pedigrees GEO accession: GSE1485). www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE1485
207. Gene Expression Omnibus (Human Variation Panel GEO accession: GSE24277). www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE24277
208. eQTL resources at the Pritchard Lab. <http://eqtl.uchicago.edu/Home.html>
209. CEU RNA-Seq Data. http://jungle.unige.ch/rnaseq_CEU60
210. Coriell Cell Repositories. <http://ccr.coriell.org>
211. PACdb. www.pacdb.org
212. TRANSFAC Database. www.gene-regulation.com
213. Pharmacogenomics Research Network. <http://pgrn.org>

Executive summary

Model system

- The use of human lymphoblastoid cell lines (LCLs) has emerged as a promising model system in the study of the genetics of drug response.
- Advantages of LCL experiments include environmental control, cost-effectiveness, unlimited cell supply and the ability to test drugs with a narrow therapeutic index.
- Limitations of LCLs include *in vitro* confounders such as growth rate and differing patterns of gene expression from the tissue of interest; therefore, replicating findings in patient cohorts is necessary.

LCL panels with genotypes available

- Centre d'Étude du Polymorphisme Humain pedigrees.
- International HapMap Project populations.
- Human Variation Panel populations.

Recent pharmacogenomic findings using the LCL model

- Chemotherapeutics: integration of genome-wide genotype, genome-wide gene expression and drug-induced phenotype data from LCLs has been used to find variants and genes associated with chemotherapeutic-induced cytotoxicity in diverse populations. Pharmacologic variants associated with chemotherapeutic susceptibility in LCLs are enriched for expression quantitative trait loci (eQTLs).
- Radiation: genome-wide studies incorporating either radiation-induced or baseline gene expression in LCLs identified variants and genes associated with radiation-induced cytotoxicity and apoptosis.
- Statins: functional studies in LCLs derived from participants in the Cholesterol and Pharmacogenetics trial demonstrated that simvastatin reduced the protein expression of LDL receptor only in carriers of a particular LDL receptor haplotype.
- Selective serotonin reuptake inhibitors: whole-genome expression profiling was performed on LCLs most sensitive and most resistant to paroxetine and most significantly, *CHLI* expression was 6.3-fold lower in the most sensitive lines.
- LCLs have also been used in pharmacogenomic studies of immunosuppressants, pain relievers and β -blockers.

Clinical translation of LCL findings

- SNPs within *FKBP5*, a gene chosen for analysis based on cytarabine arabinoside LCL studies, were associated with both event-free and overall survival in a cohort of 187 pediatric acute myeloid leukemia patients treated with cytarabine arabinoside.
- The top pathway (aspartate metabolism) associated with asparaginase cytotoxicity in a genome-wide LCL study was replicated in leukemic cells from 54 childhood acute lymphoblastic leukemia patients.

- An eQTL discovered in LCL studies of carboplatin cytotoxicity associated with overall response to carboplatin-based induction chemotherapy in 60 head and neck cancer patients.
- An eQTL discovered in carboplatin-treated LCLs also associated with progression-free and overall survival in 377 ovarian cancer patients treated with carboplatin and paclitaxel.
- Two eQTLs discovered in cisplatin-treated LCLs suggestively associated with overall survival in lung cancer patients treated with platinating agents.
- Molecular studies in LCLs, including chromatin immunoprecipitation, overexpression and knockdown experiments, were used successfully to identify a functionally relevant variant associated with musculoskeletal adverse events in white women receiving aromatase inhibitors.
- Close collaboration will be necessary between basic researchers and clinicians to translate LCL findings into the clinic.

Figure 1. Potential research design flows for pharmacogenomics studies that utilize the lymphoblastoid cell lines model

GWA: Genome-wide association; LCL: Lymphoblastoid cell line.

Table 1

Lymphoblastoid cell line panels with genome-wide genotype and gene-expression data available.

Panel	Ancestry of Samples	No. of LCLs	No. of SNP genotypes	SNP genotyping platform	Genome-wide gene expression data	Gene expression platform	Ref.
CEPH pedigrees	European ancestry from Utah (includes HapMap CEU)	676 total (48 families)	5000 (impute millions more from HapMap CEU)	Multiple platforms	Some CEPH families	Affymetrix Genome Focus array	[14,15,201,206,210]
HapMap Phase I/II	CEU YRI CHB JPT	30 trios 30 trios 45 unrelated 45 unrelated	3.1 million	Perlegen custom high-density oligonucleotide arrays	CEU and YRI exon microarray and RNA-Seq	Affymetrix GeneChip Human Exon 1.0 ST Array and Illumina Genome Analyzer II (RNA-Seq)	[2,21,202,203,205,208–210]
HapMap Phase III	CEU YRI CHB JPT ASW GIH LWK MEX CHD MKK TSI	30 trios 30 trios 45 unrelated 45 unrelated 11 trios, 24 duos and 9 unrelated 90 unrelated 90 unrelated 30 trios 90 unrelated 30 trios, 90 unrelated 90 unrelated	1.6 million	Affymetrix Human SNP array 6.0 and Illumina Human1M beadchip			[20,202,203,210]
Human Variation Panel	AA CA HCA MXA	100 unrelated 100 unrelated 100 unrelated 100 unrelated	940,000	Affymetrix Human SNP array 6.0	AA, CA and HCA microarray	Affymetrix U133 Plus 2.0 GeneChip	[42,204,207,210]

AA: African-Americans; ASW: African ancestry in southwest USA; CA: Caucasians; CEPH: Centre d'Étude du Polymorphisme Humain; CEU: Northern and western European ancestry from Utah; CHB: Han Chinese from Beijing, China; CHD: Chinese in metropolitan Denver, CO, USA; GIH: Gujarati-Indians in Houston, TX, USA; HCA: Han Chinese from Los Angeles, CA, USA; JPT: Japanese from

Tokyo, Japan; LCL: Lymphoblastoid cell line; LWK: Luhya in Webuye, Kenya; MEX: Mexican ancestry in Los Angeles, CA, USA; MKK: Maasai in Kinyawa, Kenya; MXA: Mexican-Americans from Los Angeles, CA, USA; TSI: Tuscans in Italy; YRI: Yoruba from Ibadan, Nigeria.

Table 2

Recent pharmacogenomic findings using the lymphoblastoid cell line model (since 2009).

Drug	Type of study	LCL panels used	Findings	Ref.
5-FU	Candidate gene	CEPH pedigrees	<i>TYMS</i> SNPs are associated with 5-FU cytotoxicity	[30]
Multiple chemotherapy drugs	Genome-wide association studies	HapMap Phase I/II CEU	Chemotherapeutic-induced cytotoxicity-associated SNPs are significantly enriched for eQTLs	[32]
None	Candidate gene	HapMap Phase I/II CEU, CHB, JPT and YRI	Identified and functionally validated <i>cis</i> -eQTLs in ABC membrane transporter genes	[33]
Ara-C	Genome-wide triangle model	HapMap Phase I/II CEU and YRI	Population-specific pharmacogenetic signatures, validation of association between expression of <i>CKK</i> , <i>GIT1</i> and <i>SLC25A37</i> and ara-C cytotoxicity	[35]
Ara-C and gemcitabine	Genome-wide triangle model	Human Variation Panel AA, CA and HCA	Knockdown of two eQTL targets from the triangle analysis, <i>VAV3</i> and <i>GPM6A</i> , desensitized both an LCL and a pancreatic cancer cell line to both drugs	[42]
Carboplatin and cisplatin	Genome-wide triangle model	HapMap Phase I/II CHB, JPT, CEU and YRI	Clinical observation that east Asians are more sensitive to platinum was confirmed, histone H3 family genes associated with cytotoxicity induced by both drugs	[46]
Carboplatin and cisplatin	Genome-wide meta-analysis	HapMap Phase I/II/III YRI, ASW, CHD, CHB, JPT and CEU	Cross-population variants associated with platinum cytotoxicity were identified, confirmed involvement of <i>BCL2</i> , <i>GSTM1</i> , <i>GSTT1</i> , <i>ERCC2</i> and <i>ERCC6</i> in platinum response	[50]
Radiation	Genome-wide linkage analysis	CEPH pedigrees	Radiation-induced changes in expression of 1200 genes were linked to specific chromosomal regions, several regions contained SNPs that associated with radiation-induced cytotoxicity and apoptosis	[58]
Radiation	Genome-wide triangle model	Human Variation Panel AA, CA and HCA	Knockdown of the top 23 genes in four cancer cell lines revealed five genes that significantly altered radiation sensitivity in at least two cancer cell lines	[60]
Simvastatin	Candidate gene study	LCLs derived from black participants in the Cholesterol and Pharmacogenetics trial	<i>LDLR</i> haplotype 5 associated with smaller drug-induced lipid reductions in black patients, simvastatin reduced the protein activity of <i>LDLR</i> in LCLs from haplotype L5 carriers, but not in noncarriers	[65]
Puroxetine	Genome-wide expression profiling on most sensitive and most resistant LCLs	Female LCLs from the National Laboratory for the Genetics of Israeli Populations collection	<i>CHL1</i> expression was 6.3-fold lower in the most sensitive lines, which was confirmed by real time-quantitative PCR	[68]
Ara-C	Follow-up candidate gene study in patients	Initial finding: Human Variation Panel AA, CA and HCA	<i>FKBP5</i> SNPs associated with event-free and overall survival in 187 pediatric AML patients treated with Ara-C	[77,79]
Asparaginase	Genome-wide pathway analysis in LCLs and follow-up in patients	HapMap Phase I/II CEU, YRI	The top pathway (aspartate metabolism) associated with asparaginase cytotoxicity in the LCL study was replicated in 54 childhood ALL patients	[80]
Carboplatin	Genome-wide triangle model and follow-up study in patients	HapMap Phase I/II CEU, YRI, CHB and JPT	An eQTL discovered in LCLs also associated with overall response to carboplatin in 60 head and neck cancer patients	[81]

Drug	Type of study	LCL panels used	Findings	Ref.
Carboplatin	Genome-wide triangle model, replication in LCLs and follow-up study in patients	HapMap Phase I/II/III CEU	An eQTL discovered in carboplatin-treated LCLs also associated with progression-free and overall survival in 377 ovarian cancer patients treated with carboplatin and paclitaxel	[82]
Cisplatin	Genome-wide association study in LCLs, follow-up study in patients and functional study in LCLs	Human Variation Panel AA, CA and HCA	Two eQTLs discovered in cisplatin-treated LCLs suggestively associated with overall survival in lung cancer patients treated with platinating agents, siRNA experiments in LCLs validated the involvement of the eQTL target genes, <i>DAPK3</i> and <i>MEITL6</i> , in cisplatin response	[83]
Aromatase inhibitors	Genome-wide association study in patients followed by functional studies in LCLs	Human Variation Panel AA, CA and HCA	A SNP in an estrogen response element downstream of <i>TCLL1A</i> associated with musculoskeletal adverse events in white women receiving aromatase inhibitors, functional studies in LCLs confirmed that ER α bound to the minor allele and increased <i>TCLL1A</i> expression, but did not bind to the major allele sequence	[84]

5-FU: 5-fluorouracil; AA: African-Americans; ABC: ATP-binding cassette; ALL: Acute lymphoblastic leukemia; AML: Acute myeloid leukemia; Ara-C: Cytarabine arabinoside; ASW: African ancestry in southwest USA; CA: Caucasians; CEPH: Centre d'Étude du Polymorphisme Humain; CEU: Northern and western European ancestry from Utah; CHB: Han Chinese from Beijing, China; CHD: Chinese in metropolitan Denver, CO, USA; eQTL: Expression quantitative trait locus; HCA: Han Chinese from Los Angeles, CA, USA; JPT: Japanese from Tokyo, Japan; LCL: Lymphoblastoid cell line; LDLR: Low-density lipoprotein receptor; YRI: Yoruba from Ibadan, Nigeria.