HHS Public Access

Author manuscript

Pharmacogenomics J. Author manuscript; available in PMC 2012 December 01.

Published in final edited form as:

Pharmacogenomics J. 2012 June ; 12(3): 214–226. doi:10.1038/tpj.2011.8.

Discovery of Regulatory Elements in Human ATP-Binding Cassette Transporters through Expression Quantitative Trait Mapping

Pär Matsson, PhD[1,2], **Sook Wah Yee, PhD**[1], **Svetlana Markova, PhD**[1], **Kari Morrissey, BS**[1], **Gerard Jenkins, MD**[1], **Jiekun Xuan, MS**[1], **Eric Jorgenson, PhD**[1], **Deanna L. Kroetz, PhD**[1], and **Kathleen M. Giacomini, PhD**[1]

[1]Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, CA.

Abstract

ATP-Binding Cassette (ABC) membrane transporters determine the disposition of many drugs, metabolites and endogenous compounds. Coding region variation in ABC transporters is the cause of many genetic disorders, but much less is known about the genetic basis and functional outcome of ABC transporter expression level variation. We used genotype and mRNA transcript level data from human lymphoblastoid cell lines to assess population and gender differences in ABC transporter expression, and to guide the discovery of genomic regions involved in transcriptional regulation. Nineteen of 49 ABC genes were differentially expressed between individuals of African, Asian and European descent suggesting an important influence of race on expression level of ABC transporters. Twenty-four significant associations were found between transporter transcript levels and proximally located genetic variants. Several of the associations were experimentally validated in reporter assays. Through influencing ABC expression levels, these SNPs may affect disease susceptibility and response to drugs.

Keywords

ATP-Binding Cassette Transporters; Lymphoblastoid Cell Lines; Expression Quantitative Trait Mapping; Transcriptional Regulation; Population Differences

Introduction

ATP-Binding Cassette (ABC) transporters have diverse and important functions in maintaining cellular homeostasis of endogenous solutes and various xenobiotics including

Correspondence Kathleen M. Giacomini Department of Bioengineering and Therapeutic Sciences University of California San Francisco 1550 4th Street, Box 2911, San Francisco, CA 94143-2911 Tel: (415) 476-1936 Fax: (415) 514-4361 kathy.giacomini@ucsf.edu.

[2]Current address: Department of Pharmacy, Uppsala University, Uppsala, Sweden.

Supplementary information is available at *The Pharmacogenomics Journal*'s website.

Conflicts of interest None declared.

many clinically important drugs.[1-2] In humans, 49 ABC proteins have been identified and classified into seven subfamilies (ABCA-ABCG). Most of these are membrane transporters that utilize the energy from ATP hydrolysis to mediate efflux of solutes. The different members exhibit widely varying patterns of tissue and subcellular localization, and have substrate specificities that range from endogenous solutes such as cholesterol and bile salts (ABCA1, ABCB11, ABCG1, ABCG5 and ABCG8),[3-6] phospholipids (ABCB4)[7] and long-chain fatty acids (ABCD1 and ABCD2),[8-9] to drugs and their metabolites (ABCB1, ABCG2, and several members of the ABCC family).[2, 10-11]

The important physiological functions of ABC transporters are highlighted by the many genetic disorders resulting from mutations in these proteins, e.g., Tangier's disease (ABCA1),[4] progressive familial intrahepatic cholestasis (ABCB4 and ABCB11),[12-13] adrenoleukodystrophy (ABCD1),[9] sitosterolemia (ABCG5 and ABCG8),[3, 6] and cystic fibrosis (CFTR, also referred to as ABCC7).[14] While such severe clinical phenotypes are typically attributable to loss-of-function or reduced-function mutations, natural variation in transporter expression levels may also result in detrimental clinical conditions. For instance, low expression of ABCA1 is associated with reduced plasma levels of high-density lipoprotein, although less extreme than that observed in Tangier's disease patients.[15-16] Similarly, expression level variation in drug transporting ABCs such as ABCB1, ABCC2 and ABCG2 can affect intestinal absorption, biliary and renal excretion, and tissue distribution of substrate drugs, which in turn can lead to loss of pharmacological effects, or to drug-induced toxicities. Identification of regulatory elements in ABC transporter genes is therefore essential for understanding the factors that control the expression levels of these important transporters.

Though many cis-regulatory elements have been identified in the genome, discovery of such elements is difficult. The most common approach to the discovery of cis-regulatory regions has involved comparative genomic methods as a first step to identify conserved sequence motifs followed by further computational analyses to discover transcription factor binding sites.[17-19] While these methods have resulted in the identification of many regulatory regions, use of evolutionary conservation as a first step may not be reliable. In particular, most regulatory motifs are small, typically consisting of only a few base pairs, and are therefore difficult to identify using sequence comparisons.

Recently, several groups have used publicly available mRNA expression levels in lymphoblastoid cell lines (LBLs) along with genetic variation on a genome-wide scale to identify cis-regulatory variation in the human genome.[20-23] These studies have revealed global patterns of expression quantitative trait loci (eQTLs), including population differences, and support an abundance of polymorphisms in and nearby genes, which associate with gene expression levels. In this study, we leverage these publicly available datasets, with the specific aims of determining population and individual expression level differences in ABC transporters and to discover and functionally annotate cis regulatory regions that may harbor enhancer or silencer elements that regulate the expression of ABC transporters.

Materials and Methods

ABC Gene Expression Data

Gene expression data were downloaded from the Wellcome Trust Sanger Institute web site (http://www.sanger.ac.uk/humgen/genevar). The published expression data had been log-normalized using a quantile normalization method[24] across experimental replicates for each individual, and then by a median normalization method across all 270 HapMap individuals.[23] Only data for unrelated individuals were used in the analyses, resulting in a total of 60 US Caucasians of Northern and Western European origin (CEU), 60 Yoruba from Ibadan, Nigeria (YRI), 45 Japanese from the Tokyo area (JPT), and 45 Chinese from the Beijing area (CHB) in the final dataset. For genes where multiple probes were available, analyses were performed both using individual probes and using the mean expression across all probes. The differences were generally negligible and, for clarity, only the results based on mean probe intensities are presented in the paper. Probes were considered detectably expressed in LBLs if their average intensity was above the 25th percentile of all probes in that population.

Population Dependent Expression Analyses

ANOVA was used to assess population and sex dependent differences in gene expression, using either a single-variable (population) or multi-variable (population+sex) general linear model fit in the R statistics package (http://www.r-project.org). A Bonferroni-corrected $p < 0.001$ was considered significant.

Genotype Data and Tag SNP Determination

Genotype data from the r27 release of combined data from HapMap Phase II and III were downloaded from the International HapMap Consortium Web Site (http://www.hapmap.org). For each population, local SNPs, i.e., SNPs located within a region spanning from 50 kb upstream of the transcription start site to 50 kb downstream of the end of the longest transcript (in the UCSC Genome Browser human genome release hg18; http://genome.ucsc.edu) were extracted and used in subsequent analyses. Tag SNPs were determined for the CEU, YRI and combined JPT+CHB (ASI) populations separately, using the Tagger algorithm[25] in HaploView 4.1 (http://www.broadinstitute.org/haploview/haploview) with $r^2 \geq 0.8$ and minor allele frequency ≥ 0.05.

Association Analyses and Multiple Test Corrections

We used a linear model, as implemented in PLINK v1.05[26] (http://pngu.mgh.harvard.edu/purcell/plink), to fit normalized log scale expression data for each gene to each of its local tag SNPs. Genotypes were encoded as 0, 1 or 2 depending on the number of minor alleles present. Two types of multiple test corrections were used: permutation correction, in which the order of gene expression levels to genotypes was randomly permutated 25 000 times, and a per-gene Bonferroni correction. A corrected $p < 0.05$ was considered significant. The association analyses were performed in each of the three populations separately. To account for the Asian population being a combination of the Japanese and Han Chinese populations,

conditional permutations were performed such that genotypes and phenotypes were only combined within each subpopulation.

Multiple SNP associations using Partial Least Squares projection

Partial least squares (PLS) projection, as implemented in the PLS R package,[27] was used to approximate the total expression level variance for each gene explained by its local SNPs (R^2). To achieve comparable results for all genes and to minimize the risk of overfitting the regression models, we used only the first PLS component to estimate the total R^2. The relative contribution of each local SNP to the total R^2 was calculated as the fractional contribution of that SNP's variable importance (VIP) to the total VIP summed over all SNPs.

Prioritizing SNPs for Experimental Validation

All tag SNPs significantly associated with ABC gene expression levels were considered for experimental follow-up, as were all HapMap SNPs within their linkage disequilibrium (LD) blocks ($r^2 \geq 0.8$). Multiple data sources were used to prioritize SNPs for experiments: a SNP was given a higher probability of being functional, i.e., driving the observed expression variation, if it was i) better correlated with the phenotype than were other SNPs in the LD block, ii) located in regions with higher *a priori* probabilities of being functional, such as proximal to the transcription start site, in the 3′-UTR, or in the first intron, iii) predicted to alter binding of transcription factors by the P-Match (http://www.gene-regulation.com), TF-Search (http://www.cbrc.jp/research/db/TFSEARCH.html) or the DeltaMatch software[28], or iv) located in a region conserved between humans and other placental mammals (based on the phyloP44wayPlacMammal track in the UCSC Genome Browser; http://genome.ucsc.edu).

Cloning of Putative Regulatory Elements

The regions surrounding SNPs selected for experimental validation were amplified by a PCR of pooled human genomic DNA (SOPHIE cohort, http://pharmacogenetics.ucsf.edu), using previously reported PCR protocols.[29-30] PCR primers were designed using the Primer 3 software (http://frodo.wi.mit.edu/ primer3) by inputting 350 bp upstream and downstream sequence relative to each SNP. The primers used to construct the putative regulatory regions are listed in Supplemental Table 3. The PCR-amplified regions (approximately 500 bp centered on the SNP) were purified by gel extraction (QIAquick Gel Extraction Kit, Qiagen, Germantown, MD) and cloned into pENTR-dTOPO vectors (Invitrogen, Carlsbad, CA). Following plasmid DNA purification (QIAprep Spin Miniprep Kit (Qiagen), proper orientation was confirmed by sequencing. Clones in the correct orientation were subcloned into the multiple cloning site of the pGL4.23[*luc2*/minP] vector (Promega, Madison, WI), upstream of a minimal TATA promoter and a *luc2* firefly luciferase reporter gene. The orientation and sequence of the inserts were then re-verified by sequencing. The use of pooled genomic DNA from individuals of different ethnicities enabled us to obtain both the reference and variant alleles from the same PCR amplification for most regions. Site-directed mutagenesis (QuikChange Site-Directed Mutagenesis Kit, Stratagene, Santa Clara, CA) was used to obtain the variant allele for the remaining regions. For follow-up studies of

negative regulatory elements, putative regulatory regions were re-amplified from the pGL4.23[*luc2*/minP] plasmids and subcloned into the pGL3 Promoter vector (Promega), which contains a strong SV40 promoter and a *luc*+ firefly luciferase reporter gene downstream from the multiple cloning site.

Reporter Assays for Determining Transcription Regulatory Effects

The human renal cell adenocarcinoma cell line ACHN was purchased from American Type Culture Collection (ATCC, Manassas, VA). The human hepatoblastoma cell lines HepG2 and Huh-7, human colorectal cell line HCT-116, and human renal adenocarcinoma cell line HEK293T were supplied by the UCSF Cell Culture Facility. All cell lines were maintained in a culture medium consisting of Dulbecco's modified Eagle's medium with 4 500 mg/L glucose, supplemented with 10% fetal bovine serum, 100 U/mL penicillin and 100 μg/mL streptomycin. In the reporter assay, the five cell lines were seeded into 48-well culture plates and were transfected 16 to 24 hours later using the Lipofectamine LTX transfection reagent (Invitrogen), according to the manufacturer's protocols. The pGL4.23[*luc2*/minP] vectors (200 ng), with or without the region of interest, were co-transfected with pGL4.74[*hRluc*/TK] vector(5 ng; to normalize for transfection efficiency) into each well using 1.0 – 1.5 μL Lipofectamine LTX transfection reagent. Cells were lysed with passive lysis buffer 24 hours after transfection, and were assayed for firefly and renilla luciferase activity in a GloMax 96 Microplate Dual Injector Luminometer, using the Dual-Luciferase Reporter Assay system kit (Promega). To account for differences in cell type, transfection efficiency and cell density, reporter activities were normalized as $\log_2$ [(firefly luciferase/renilla luciferase)/average of the negative control], where negative control refers to cells transfected with the pGL4.23[*luc2*/minP] vector or pGL3-Promoter vector not containing the PCR-amplified regions.

Accession Numbers

The expression data used in this study have been deposited previously at the Wellcome Trust Sanger Institute Web Site (http://www.sanger.ac.uk/humgen/genevar) and in the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo; accession GSE6536).

Results

Population dependent expression of ABC transporters

Our analyses followed the steps outlined in Fig. 1 and began with expression level analyses followed by association analyses of gene expression with genetic variants in and nearby each ABC transporter gene. For expression level analyses, first, we assessed the extent of inter-individual, population dependent and gender dependent variation in the expression levels of each ABC transporter (Fig. 1A). Genome-wide transcript levels, measured in LBLs from the 270 participants in the original HapMap European American (CEU), Han Chinese (CHB), Japanese (JPT) and Yoruba African (YRI) populations, were obtained from Stranger et al.[23] Probes were available for all 49 human ABC transporters (Table 1), and 45 of these were detectably expressed in at least one of the HapMap populations.

The inter-individual variance of ABC transporter transcript levels was slightly higher than the average for all detectable transcripts, although not statistically significant (1.2-1.5 × genome average; $p < 0.15$-0.5, depending on the population). Compared to within-population variation, population differences contributed significantly less to the overall variability (Supplemental Table 1). On average, 10% (range: 1-40%) of the total variance for the different transporters was related to population. Expression levels generally correlated well ($r^2 = 0.985$-0.997) between pairs of populations (Fig. 2). To increase power, expression level data from the Japanese and Chinese populations were combined in subsequent analyses (ASI population).

Of the 45 ABC transporters expressed in any population, 19 showed statistically significant population differences (Bonferroni-corrected $p < 0.001$, Fig. 3). However, these differences were generally modest, with an average difference in log normalized expression of 0.27±0.18 between the populations with the lowest and highest expression (corresponding to an average relative difference of 21%; range 5-81%) (Fig. 3). No major gender differences were observed (Supplemental Table 1); the only ABC transporter with a statistically detectable gender effect was ABCG1, which was expressed on average at 22% higher levels in females than in males (Bonferroni-corrected $p < 0.05$).

Associations between genotypes and gene expression

Next, we performed association analyses to identify local single nucleotide polymorphisms (SNPs) that associated with expression levels in each of the ABC transporters. (See Fig. 1B-D). Genome-wide analyses have suggested that the majority of regulatory polymorphisms are located within or proximal to the regulated genes,[20, 22-23, 31] and based on this we decided to include SNPs in a ±50 000 bp region surrounding each transporter to balance coverage of putative regulatory regions against the statistical power to detect associations.

We used genotype data from the combined first, second and third phases of the HapMap project, which resulted in an increased coverage of genetic variation compared to previous genome-wide association analyses for this expression dataset.[23] In total, 13 242, 13 120 and 13 335 SNPs with a population-specific minor allele frequency of at least 5% were examined in the CEU, YRI and ASI populations, respectively. The number of SNPs per gene varied between 79 and 963 depending on the transporter and the population, but when accounting for the large variation in gene size, the coverage was more similar among genes and populations: between 0.54 and 4.71 SNPs per kb, with an average of 1.46, 1.45 and 1.43 in the Asian, Caucasian, and African populations, respectively.

To increase our power to detect association signals, we exploited linkage disequilibrium patterns to limit the number of statistical tests. As generally observed, genetic diversity was largest in the Yoruba population, for which a total of 3 882 tag SNPs were needed to cover all common HapMap SNPs at an $r^2 \geq 0.8$. In the Caucasian and Asian populations, the corresponding numbers were 2 069 and 1 744. The genome-wide association signals obtained for these tag SNPs are shown in Figure 4. Each population harbored some SNPs that were associated with expression levels of individual ABC transporters at extraordinarily low p-values. After correcting for multiple tests using a stringent Bonferroni correction or the permutation of phenotype data to determine empirical *p*-values, 24 tag SNPs were

significantly associated with the expression levels of 10 different ABC transporters (Table 2). On average, the tag SNPs explained 21±8% (range: 11-38%) of the variation in the corresponding transcript levels. Figure 5A shows the most significant association for each of these 10 transporters.

Notably, for some transporters, multiple significant tag SNPs were detected in a single population. For example, five significant associations were detected in ABCA1 in the CEU population, with r^2 ranging from 0.21 to 0.38. Since a cutoff of r^2 0.8 was used to define LD blocks, intermediate intercorrelations often still exist among selected tags (e. g., the five significant tags in ABCA1 exhibit pairwise r^2 = 0.03-0.51 and D'=0.31-1.0). Consequently, the tag SNPs are not completely independent, and partial correlations can result in cases in which a true causal effect of one tag is detected as a weaker effect in another. We therefore adopted a multivariate approach, using partial least squares (PLS) projection, to estimate the overall contribution of all local SNPs to the expression variation.

For ABC transporters having at least one significant association, the full set of local SNPs explained 25±10% (range: 13-49%) of the total variance in expression levels (Supplemental Table 2). Very similar results were obtained when only tag SNPs were included in the models, indicating that the PLS projection methodology accounts for SNP intercorrelation caused by linkage disequilibrium. The multi-SNP analyses typically recaptured the same relative importance of local SNPs as that determined through single-SNP associations: the average correlation between PLS-derived SNP weights and correlation coefficients derived from the single-SNP analyses was r^2=0.92±0.06. In general, the total expression variance explained by all local SNPs was only moderately larger than that explained by the most significant single SNP (Supplemental Fig. 1) suggesting that for most of the studied genes, a single regulatory locus explains all of the expression variation related to the interrogated local SNPs.

Population dependent associations

We examined whether SNPs that associated with expression levels in any one population also explained the observed population differences. For seven of the SNPs associated with within-population expression variation, the allele frequencies were highly correlated with population average transcript levels (r^2≥0.95; <0.1). In particular, the significantly higher expression of ABCA12 in the CEU compared to the YRI and ASI individuals could be attributed to a SNP (rs2888327) private to the Caucasian population (Fig. 5B). Similarly, population average expression levels of ABCC10 were well explained by the allele frequencies of rs12195350 (r^2=1.0.; p<0.002), with expression levels CEU>ASI>YRI and allele frequencies CEU>ASI>YRI (data not shown).

Detection of transcriptional regulatory regions

SNPs that associated with expression level of the transporters were used to identify putative regulatory regions. These regions were then experimentally validated in medium throughput functional studies using a transient transfection reporter system (see Fig. 1E-H). We prioritized all SNPs in LD with each significant tag based on multiple criteria, including their association with additional phenotypes, their location in genomic regions with higher *a*

priori potential for regulatory function (e.g., near the transcription start site, in UTRs, or in the first few introns of the gene of interest), conservation to other placental mammals, or predicted alteration of transcription factor motifs (Fig. 1E; Materials and Methods).

A total of 16 SNPs in 7 ABC transporters were selected for experimental verification. For each of these, a region of approximately 500 bp centered on the SNP was cloned from human genomic DNA into a luciferase reporter vector, upstream of a minimal TATA promoter. The constructs were transiently transfected on three to five separate occasions into five different cell lines, derived from human colorectal epithelium (HCT-116), renal epithelium (ACHN, HEK293T), or hepatocytes (HepG2, Huh-7) (Fig. 6A). For most constructs, similar results were obtained across the panel of cell lines. Five of the selected 16 variants resulted in a more than two-fold increased luciferase activity in multiple cell lines, three of which showed more than four-fold increased activity in specific cell types. Notably, several of the cloned regions resulted in a markedly reduced luciferase activity (Fig. 6A). Since the weak promoter is not specifically designed to study negative regulatory effects, we subcloned these six regions into an alternative vector upstream of the stronger SV40 promoter. Two of the regions gave a more than four-fold reduced luciferase activity also for the strong promoter (Fig. 6B), suggesting a binding of repressive regulatory factors to these regions. In summary, guided by statistical associations between genetic polymorphisms and expression levels, we identified and validated a total of five regions that were designated enhancer or silencer regions, and two that were designated moderate enhancers. Through affecting the expression levels of proximally located ABC transporters, these regulatory regions may influence diverse physiological and pharmacological functions including cholesterol efflux to high density lipoprotein,[3-4] peroxisomal transport of very-long-chain fatty acids,[9, 32] presentation of cytosolic peptide antigens to MHC class I molecules in the endoplasmatic reticulum,[33-34] and resistance to anti-cancer drugs such as docetaxel and vinorelbine[35-36] (Table 3).

For the seven regulatory regions (five of which resulted in >2-fold enhanced luciferase activity and two of which resulted in a >4-fold reduced luciferase activity), we compared luciferase activities of the reference and variant alleles to determine the effect of the variants on transcription. Significant differences were observed for 5 of the allelic variants (Table 4). For most of these, statistical differences were specific to one or two of the tested cell lines. Four polymorphisms showed the same directionality as that observed in the association with mRNA expression in LBLs, although for two of these the difference did not reach statistical significance. Notably, the SNP rs17091297 in ABCD1 showed large differences across all tested cell lines, reproducing the directionality of the LBL association: the major allele resulted in an almost eight-fold lower reporter activity than the minor allele, mirroring the lower average LBL gene expression for the major allele (Fig 7). In contrast, the SNPs rs12227668 in ABCD2 and rs9349256 in ABCC10 showed the opposite directionality to that expected from LBL expression levels (Table 4); these results were consistent across multiple experiments and cell lines. Computationally predicted transcription factor binding affinities were consistently greater for the genetic variants that showed the higher reporter activity (Table 4), suggesting that the observed allelic differences are real, and that the

different directionality could be the result of combinatorial regulation by additional regions in vivo, that were not included in the cloned regions.

Thus, by using statistical associations between SNPs and expression levels to guide the experiments, five regions were identified that resulted in a markedly altered reporter activity, and an additional six having intermediate effects. To assess if this was more than expected by chance, we estimated the total number of regulatory regions in the examined ABC transporters based on a combination of species conservation data and experimental data for DNAse I hypersensitivity or histone 3-lysine 4 monomethylation (H3K4me1). Each of these measures have been shown to correlate with regulatory function,[37-40] but used separately they are likely to overestimate the number of regulatory active sites. We used the fraction of the interrogated genomic DNA that had two or more of these characteristics as an estimate of true regulatory function, resulting in a background chance of hitting a regulatory region of 6.4% (Supplemental Table 4). In contrast, our rate of 29-69%, depending on the cutoff used to call a significantly altered reporter activity, is significantly higher ($p < 0.0002–1\times10^{-14}$).

Discussion

Predisposition to disease and the response to drugs are highly variable in human populations,[41-42] with expression level differences in various genes contributing to interindividual differences in disease risk and drug response[43-44]. We assessed the influence of ethnicity, gender, and local genetic variation on the expression levels of ABC transporters in lymphoblastoid cell lines, and used the detected expression quantitative trait loci (eQTL) to guide experimental assays of putative regulatory genomic regions.

For 19 of the 49 ABC genes in the human genome, ethnicity was significantly associated with expression levels. Since ABC transporters are known to affect the disposition of many drugs, metabolites and endogenous ligands,[1-2] such population-dependent expression can ultimately lead to differential susceptibility to disease or outcome of drug therapy among populations. For example, we observed that the expression level of the hepatic bile salt efflux pump (ABCB11/BSEP), which when mutated is associated with familial, drug-induced and pregnancy-induced intrahepatic cholestasis,[13, 45-46] was significantly lower in the Asian than in the European and African samples. Though our data cannot be directly extrapolated to expression level differences in the liver, if such population differences in hepatic expression levels of ABCB11 occurred, our data would suggest that Asians may be more susceptible to intrahepatic cholestasis mediated by BSEP. Intriguingly, the incidence of xxx is higher in Asians than in European women, and the possible connection with population dependent BSEP expression thus warrants further studies.

The largest difference among populations was observed for ABCG1, with expression levels on average 1.8 fold higher in the Asian compared to the Caucasian samples. Notably, in contrast to the several cases of statistically significant population dependent expression, ABCG1 was the only ABC transporter for which gender was significantly associated with expression. ABCG1 mediates the efflux of cellular cholesterol to high density lipoprotein (HDL),[5, 47] and the higher average expression levels in women, and in the Asian population,

suggests that ABCG1 expression might contribute to the relatively higher HDL plasma levels observed in these populations.[48-49] Again, some caution in extrapolation of our data in lymphoblastoid cell lines to liver must be exercised.

Previous studies have indicated that genetic differences account for a substantial part of observed gene expression differences among populations.[50-51] This was also observed here for some of our genes, with population differences in allele frequencies generally well correlated with population average expression. Particularly, the SNP rs2888327 was only observed in the CEU population, which had a higher average expression of the proximally located ABCA12 gene than the ASI and YRI populations (Figure 5B). On the whole, however, the contribution of population and gender was small compared to the total ABC gene expression variation, consistent with other studies of population effects on gene expression,[50-52] and also with the observation that 85-95% of human genetic variation is due to within-population variation.[53-54]

With current advances in sequencing techniques, access to near-complete personal genomes is not far from reality. However, for these large-scale datasets to be truly useful, much work is needed in functionally annotating the human genome. So far, our knowledge of non-coding regions is greatly lagging behind that of protein coding regions. Especially little is known about the extent and function of distantly located regulatory elements such as enhancers and silencers. Evolutionary constraints on non-coding sequences have been successfully used to identify regions with regulatory function.[17-19] However, not all evolutionarily conserved regions have detectable functions in in vitro and in vivo enhancer assays.[18-19, 55] Furthermore, many regulatory elements are likely located in regions not detectable using comparative genomics.[56-57] Alternatively, DNAse I hypersensitivity,[37-38] binding of enhancer-associated proteins such as the acetyltransferase and transcriptional coactivator p300,[40, 55] and epigenetic modifications such as histone methylation and acetylation[39-40] have been used to predict the genome-wide locations of regulatory regions. While highly promising, additional experimental validation of predicted regulatory regions is needed to determine the predictive value of these methods.

Here, we used a complementary approach to identify putative regulatory regions, through the association of genetic polymorphism with expression levels of ABC transporters. We identified 24 tag SNPs that were significantly associated with ABC transporter expression levels in lymphoblastoid cell lines. These were in linkage disequilibrium with a total of 104 interrogated SNPs with a minor allele frequency >5% in at least one population. After prioritizing the initial list of possible eQTLs based on association strength, genomic location, predicted alteration of transcription factor binding sites and species conservation, 16 regions within a ±50 000 bp range surrounding 7 ABC transporters were selected for experimental validation. This eQTL-centered approach led to the identification of five genomic regions that resulted in at least fourfold increased or decreased reporter activity, and six additional regions with at least twofold altered activities. Allele-specific regulatory function was shown for five of the regions. Genetic variation in these regions can thus contribute to differential expression among individuals, and can ultimately affect drug response or susceptibility to disease.

Previous analyses of expression quantitative traits in lymphoblastoid cell lines[20, 22] have focused on global properties of regulatory polymorphisms, and have not included experimental verification of identified associations. In our gene family focused effort, 5 to 11 of 16 tested associations were experimentally replicated depending on the reporter assay cutoff used. However, five regions did not alter reporter activity. While significant signals in the reporter assay suggest causative regulatory effects of the examined regions, interpreting the lack of signal is less straight-forward. It could be a consequence of false positives in the association analysis, but could also result from linkage disequilibrium between the examined region and the actual causative site. In addition, combinatorial regulation of gene expression in vivo is difficult to replicate in an artificial system that does not include the larger genomic context.

We recently used a combination of comparative genomics and transcription factor binding site predictions to select 50 putative regulatory regions in important hepatic drug transporters from the ABC and SLC families.[58] The enhancer activities of these regions were assayed in mice using hydrodynamic tail vein injection, which results in a highly localized expression in the liver. Using this technique, 12/50 (24%) of the prioritized regions showed experimental activity in vivo. Notably, however, none of the regions with experimental regulatory impact in the present study were detectably conserved among placental mammals, suggesting that eQTL mapping provides additional information compared to comparative genomics approaches.

The present work thus shows the general utility of expression level quantitative trait loci in defining genomic regions with regulatory potential, and, specifically, demonstrates transcriptional regulation of human ABC transporters by several hitherto uncharacterized genomic regions. Genetic and epigenetic variation in the identified regions can thus lead to altered transport of drugs and their metabolites, and of endogenous solutes such as cholesterol, fatty acids and peptide fragments. Ultimately, this can cause inter-individual variation in the pharmacological response to drugs and in important physiological processes, including cellular lipid metabolism, lysosomal protein degradation and in antigen presentation on the cell surface.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by National Institutes of Health grant GM61390. PM was supported in part by a grant from the IF Foundation for Pharmaceutical Sciences, Sweden.

Funding sources This work was supported by National Institutes of Health grant GM61390. PM was supported in part by a grant from the IF Foundation for Pharmaceutical Sciences, Sweden.

References

1. Glavinas H, Krajcsi P, Cserepes J, Sarkadi B. The role of ABC transporters in drug resistance, metabolism and toxicity. Curr Drug Deliv. 2004; 1(1):27–42. [PubMed: 16305368]

2. Schinkel AH, Jonker JW. Mammalian drug efflux transporters of the ATP binding cassette (ABC) family: an overview. Adv Drug Deliv Rev. 2003; 55(1):3–29. [PubMed: 12535572]
3. Berge KE, Tian H, Graf GA, Yu L, Grishin NV, Schultz J, et al. Accumulation of dietary cholesterol in sitosterolemia caused by mutations in adjacent ABC transporters. Science. 2000; 290(5497): 1771–1775. [PubMed: 11099417]
4. Brooks-Wilson A, Marcil M, Clee SM, Zhang LH, Roomp K, van Dam M, et al. Mutations in ABC1 in Tangier disease and familial high-density lipoprotein deficiency. Nat Genet. 1999; 22(4):336–345. [PubMed: 10431236]
5. Klucken J, Buchler C, Orso E, Kaminski WE, Porsch-Ozcurumez M, Liebisch G, et al. ABCG1 (ABC8), the human homolog of the Drosophila white gene, is a regulator of macrophage cholesterol and phospholipid transport. Proc Natl Acad Sci U S A. 2000; 97(2):817–822. [PubMed: 10639163]
6. Lee MH, Lu K, Hazard S, Yu H, Shulenin S, Hidaka H, et al. Identification of a gene, ABCG5, important in the regulation of dietary cholesterol absorption. Nat Genet. 2001; 27(1):79–83. [PubMed: 11138003]
7. Ruetz S, Gros P. Phosphatidylcholine translocase: a physiological role for the mdr2 gene. Cell. 1994; 77(7):1071–1081. [PubMed: 7912658]
8. Lombard-Platet G, Savary S, Sarde CO, Mandel JL, Chimini G. A close relative of the adrenoleukodystrophy (ALD) gene codes for a peroxisomal protein with a specific expression pattern. Proc Natl Acad Sci U S A. 1996; 93(3):1265–1269. [PubMed: 8577752]
9. Watkins PA, Gould SJ, Smith MA, Braiterman LT, Wei HM, Kok F, et al. Altered expression of ALDP in X-linked adrenoleukodystrophy. Am J Hum Genet. 1995; 57(2):292–301. [PubMed: 7668254]
10. Kruh GD, Belinsky MG. The MRP family of drug efflux pumps. Oncogene. 2003; 22(47):7537–7552. [PubMed: 14576857]
11. Vlaming ML, Lagas JS, Schinkel AH. Physiological and pharmacological roles of ABCG2 (BCRP): recent findings in Abcg2 knockout mice. Adv Drug Deliv Rev. 2009; 61(1):14–25. [PubMed: 19118589]
12. Deleuze JF, Jacquemin E, Dubuisson C, Cresteil D, Dumont M, Erlinger S, et al. Defect of multidrug-resistance 3 gene expression in a subtype of progressive familial intrahepatic cholestasis. Hepatology. 1996; 23(4):904–908. [PubMed: 8666348]
13. Strautnieks SS, Bull LN, Knisely AS, Kocoshis SA, Dahl N, Arnell H, et al. A gene encoding a liver-specific ABC transporter is mutated in progressive familial intrahepatic cholestasis. Nat Genet. 1998; 20(3):233–238. [PubMed: 9806540]
14. Kerem BS, Buchanan JA, Durie P, Corey ML, Levison H, Rommens JM, et al. DNA marker haplotype association with pancreatic sufficiency in cystic fibrosis. Am J Hum Genet. 1989; 44(6): 827–834. [PubMed: 2567116]
15. Frikke-Schmidt R, Nordestgaard BG, Jensen GB, Tybjaerg-Hansen A. Genetic variation in ABC transporter A1 contributes to HDL cholesterol in the general population. J Clin Invest. 2004; 114(9):1343–1353. [PubMed: 15520867]
16. Slatter TL, Jones GT, Williams MJ, van Rij AM, McCormick SP. Novel rare mutations and promoter haplotypes in ABCA1 contribute to low-HDL-C levels. Clin Genet. 2008; 73(2):179–184. [PubMed: 18199144]
17. Nobrega MA, Ovcharenko I, Afzal V, Rubin EM. Scanning human gene deserts for longrange enhancers. Science. 2003; 302(5644):413. [PubMed: 14563999]
18. Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, Shoukry M, et al. In vivo enhancer analysis of human conserved non-coding sequences. Nature. 2006; 444(7118):499–502. [PubMed: 17086198]
19. Visel A, Prabhakar S, Akiyama JA, Shoukry M, Lewis KD, Holt A, et al. Ultraconservation identifies a small subset of extremely constrained developmental enhancers. Nat Genet. 2008; 40(2):158–160. [PubMed: 18176564]
20. Dixon AL, Liang L, Moffatt MF, Chen W, Heath S, Wong KC, et al. A genome-wide association study of global gene expression. Nat Genet. 2007; 39(10):1202–1207. [PubMed: 17873877]

21. Huang RS, Duan S, Shukla SJ, Kistner EO, Clark TA, Chen TX, et al. Identification of genetic variants contributing to cisplatin-induced cytotoxicity by use of a genomewide approach. Am J Hum Genet. 2007; 81(3):427–437. [PubMed: 17701890]
22. Stranger BE, Forrest MS, Clark AG, Minichiello MJ, Deutsch S, Lyle R, et al. Genome-wide associations of gene expression variation in humans. PLoS Genet. 2005; 1(6):e78. [PubMed: 16362079]
23. Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, Beazley C, et al. Population genomics of human gene expression. Nat Genet. 2007; 39(10):1217–1224. [PubMed: 17873874]
24. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics. 2003; 19(2):185–193. [PubMed: 12538238]
25. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005; 21(2):263–265. [PubMed: 15297300]
26. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81(3): 559–575. [PubMed: 17701901]
27. Mevik BH, Wehrens R. The pls package: Principal component and partial least squares regression in R. J Stat Softw. 2007; 18(2)
28. Williamson, DW. PhD thesis. University of California San Francisco; San Francisco, CA: 2008. Genetic and bioinformatic approaches to identify polymorphic modulators of transcription factor binding and disease phenotypes including HIV-1 viremia.
29. Tahara H, Yee SW, Urban TJ, Hesselson S, Castro RA, Kawamoto M, et al. Functional genetic variation in the basal promoter of the organic cation/carnitine transporters OCTN1 (SLC22A4) and OCTN2 (SLC22A5). J Pharmacol Exp Ther. 2009; 329(1):262–271. [PubMed: 19141711]
30. Yee SW, Shima JE, Hesselson S, Nguyen L, De Val S, Lafond RJ, et al. Identification and characterization of proximal promoter polymorphisms in the human concentrative nucleoside transporter 2 (SLC28A2). J Pharmacol Exp Ther. 2009; 328(3):699–707. [PubMed: 19098160]
31. Veyrieras JB, Kudaravalli S, Kim SY, Dermitzakis ET, Gilad Y, Stephens M, et al. High-resolution mapping of expression-QTLs yields insight into human gene regulation. PLoS Genet. 2008; 4(10):e1000214. [PubMed: 18846210]
32. van Roermund CW, Visser WF, Ijlst L, van Cruchten A, Boek M, Kulik W, et al. The human peroxisomal ABC half transporter ALDP functions as a homodimer and accepts acyl-CoA esters. FASEB J. 2008; 22(12):4201–4208. [PubMed: 18757502]
33. Bahram S, Arnold D, Bresnahan M, Strominger JL, Spies T. Two putative subunits of a peptide pump encoded in the human major histocompatibility complex class II region. Proc Natl Acad Sci U S A. 1991; 88(22):10094–10098. [PubMed: 1946428]
34. Perria CL, Rajamanickam V, Lapinski PE, Raghavan M. Catalytic site modifications of TAP1 and TAP2 and their functional consequences. J Biol Chem. 2006; 281(52):39839–39851. [PubMed: 17068338]
35. Bessho Y, Oguri T, Ozasa H, Uemura T, Sakamoto H, Miyazaki M, et al. ABCC10/MRP7 is associated with vinorelbine resistance in non-small cell lung cancer. Oncol Rep. 2009; 21(1):263–268. [PubMed: 19082471]
36. Hopper-Borge E, Chen ZS, Shchaveleva I, Belinsky MG, Kruh GD. Analysis of the drug resistance profile of multidrug resistance protein 7 (ABCC10): resistance to docetaxel. Cancer Res. 2004; 64(14):4927–4930. [PubMed: 15256465]
37. Boyle AP, Davis S, Shulha HP, Meltzer P, Margulies EH, Weng Z, et al. High-resolution mapping and characterization of open chromatin across the genome. Cell. 2008; 132(2):311–322. [PubMed: 18243105]
38. Sabo PJ, Humbert R, Hawrylycz M, Wallace JC, Dorschner MO, McArthur M, et al. Genomewide identification of DNaseI hypersensitive sites using active chromatin sequence libraries. Proc Natl Acad Sci U S A. 2004; 101(13):4537–4542. [PubMed: 15070753]
39. Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, et al. Genomic maps and comparative analysis of histone modifications in human and mouse. Cell. 2005; 120(2): 169–181. [PubMed: 15680324]

40. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet. 2007; 39(3):311–318. [PubMed: 17277777]

41. Yasuda SU, Zhang L, Huang SM. The role of ethnicity in variability in response to drugs: focus on clinical pharmacology studies. Clin Pharmacol Ther. 2008; 84(3):417–423. [PubMed: 18615002]

42. Giacomini KM, Krauss RM, Roden DM, Eichelbaum M, Hayden MR, Nakamura Y. When good drugs go bad. Nature. 2007; 446(7139):975–977. [PubMed: 17460642]

43. 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(20): 9287–9292. [PubMed: 20442332]

44. 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

45. Pauli-Magnus C, Meier PJ, Stieger B. Genetic determinants of drug-induced cholestasis and intrahepatic cholestasis of pregnancy. Semin Liver Dis. 2010; 30(2):147–159. [PubMed: 20422497]

46. Funk C, Ponelle C, Scheuermann G, Pantze M. Cholestatic potential of troglitazone as a possible factor contributing to troglitazone-induced hepatotoxicity: in vivo and in vitro interaction at the canalicular bile salt export pump (Bsep) in the rat. Mol Pharmacol. 2001; 59(3):627–635. [PubMed: 11179459]

47. Vaughan AM, Oram JF. ABCG1 redistributes cell cholesterol to domains removable by high density lipoprotein but not by lipid-depleted apolipoproteins. J Biol Chem. 2005; 280(34):30150–30157. [PubMed: 15994327]

48. Japan Arteriosclerosis Longitudinal Study-Existing Cohorts Combine (JALS-ECC): rationale, design, and population characteristics. Circ J. 2008; 72(10):1563–1568. [PubMed: 18728337]

49. Lloyd-Jones D, Adams RJ, Brown TM, Carnethon M, Dai S, De Simone G, et al. Heart disease and stroke statistics--2010 update: a report from the American Heart Association. Circulation. 2010; 121(7):e46–e215. [PubMed: 20019324]

50. Spielman RS, Bastone LA, Burdick JT, Morley M, Ewens WJ, Cheung VG. Common genetic variants account for differences in gene expression among ethnic groups. Nat Genet. 2007; 39(2): 226–231. [PubMed: 17206142]

51. Zhang W, Duan S, Kistner EO, Bleibel WK, Huang RS, Clark TA, et al. Evaluation of genetic variation contributing to differences in gene expression between populations. Am J Hum Genet. 2008; 82(3):631–640. [PubMed: 18313023]

52. Storey JD, Madeoy J, Strout JL, Wurfel M, Ronald J, Akey JM. Gene-expression variation within and among human populations. Am J Hum Genet. 2007; 80(3):502–509. [PubMed: 17273971]

53. Excoffier L, Hamilton G. Comment on "Genetic structure of human populations". Science. 2003; 300(5627):1877. author reply 1877. [PubMed: 12817127]

54. Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky LA, et al. Genetic structure of human populations. Science. 2002; 298(5602):2381–2385. [PubMed: 12493913]

55. Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, et al. ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature. 2009; 457(7231):854–858. [PubMed: 19212405]

56. Cooper GM, Brown CD. Qualifying the relationship between sequence conservation and molecular function. Genome Res. 2008; 18(2):201–205. [PubMed: 18245453]

57. McGaughey DM, Vinton RM, Huynh J, Al-Saif A, Beer MA, McCallion AS. Metrics of sequence constraint overlook regulatory sequences in an exhaustive analysis at phox2b. Genome Res. 2008; 18(2):252–260. [PubMed: 18071029]

58. Kim, MJ.; Skewes-Cox, P.; Fukushima, H.; Hesselson, S.; Nguyen, L.; Eshragh, JL., et al. In Vivo Pharmacogenomic Characterization of Enhancers of Liver Membrane Transporters. 2010. In Manuscript

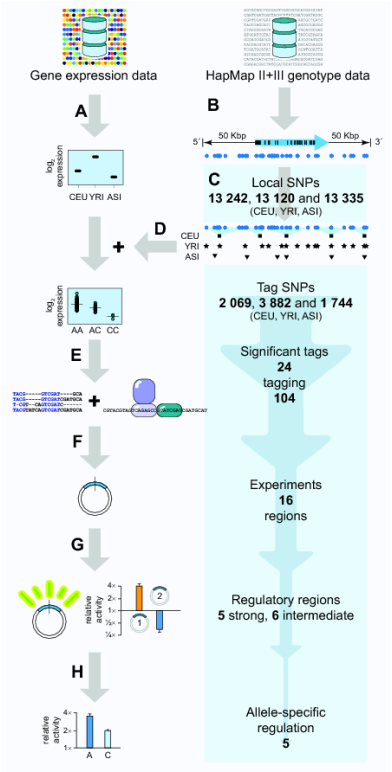

Fig. 1.

Schematic of the workflow applied to identify regulatory regions and SNPs in ABC genes. The number of SNPs assessed in each analysis step is shown in the shaded panel to the right. A: Gene expression data for ABC genes were downloaded and extracted from the Wellcome Trust Sanger Institute GeneVar database. Expression levels for all 210 unrelated individuals were used to determine population differences in ABC gene expression. B: Genotype data was downloaded from the HapMap database, and all SNPs within 50 kb of each ABC gene was used in the further analyses. C: Tag SNPs were determined for each of the CEU, YRI and ASI (JPT+CHB) populations separately. D: Associations between ABC expression levels and SNPs were determined in each population. E: Tag SNPs significantly associated with expression levels, together with all HapMap SNPs in linkage disequilibrium ($r^2 \geq 0.8$) with them, were prioritized for experimental follow-up based on individual association significance, SNP location, species conservation, and predicted altered transcription factor binding. F: The ~500 bp regions surrounding each selected SNPs were cloned into a luciferase reporter vector (pGL4.23[*luc2*/minP] vector). G: Luciferase activity was assessed in five different cell lines (ACHN, HEK293T, HCT-116, HepG2 and Huh-7) for each reporter construct. Constructs resulting in decreased luciferase activity compared to the negative control were subsequently cloned into an alternative vector carrying a strong SV40 promoter (pGL3-Promoter vector). H: Allele-specific effects on luciferase activity was determined for all significant regulatory regions.

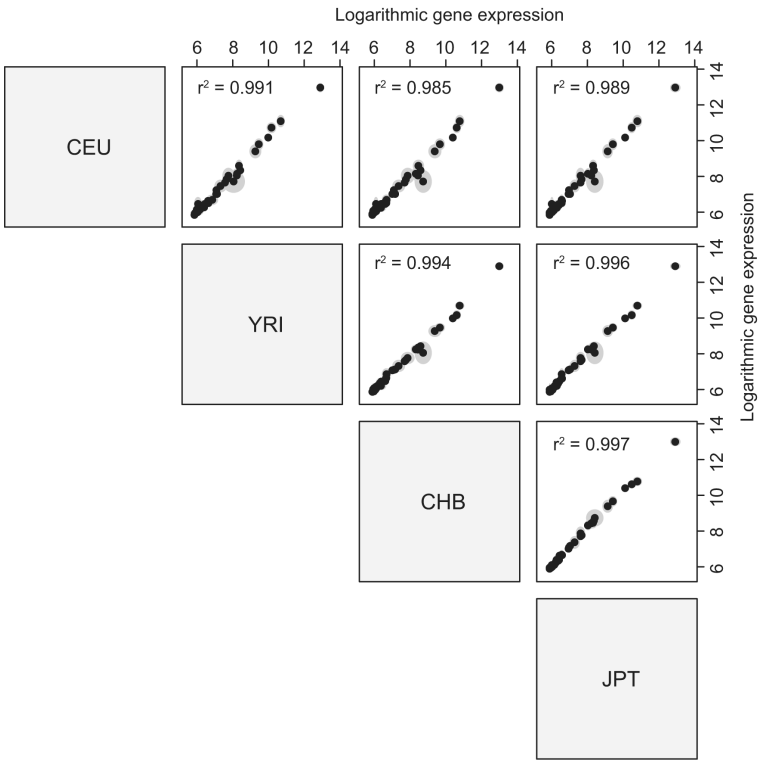

Fig. 2.

Comparisons of expression of ABC transporter genes in lymphoblastoid cell lines among four populations. The correlations of normalized probe intensities are shown across all pairs of populations. Probe intensities are shown as filled circles, and the standard error of the mean as grey ellipses. For most probes, the standard errors were smaller than the size of the plot symbols.

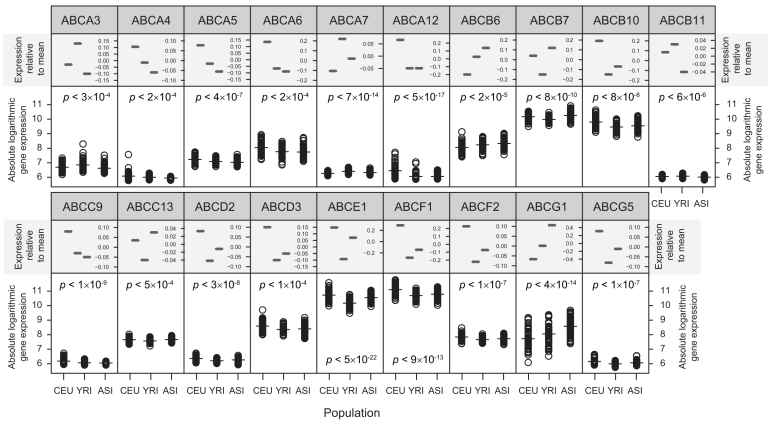

Fig. 3.

ABC transporter genes exhibiting significant population differences in expression levels in lymphoblastoid cell lines. ANOVA was used to determine differential expression between the CEU, YRI and ASI populations. The plot shows ABC genes with significant differences after correcting for multiple tests (Bonferroni corrected $p < 0.001$). Bottom panels show $\log_2$ normalized expression values for each population. Top panels show relative differences compared to the overall mean across all populations, in original untransformed units.

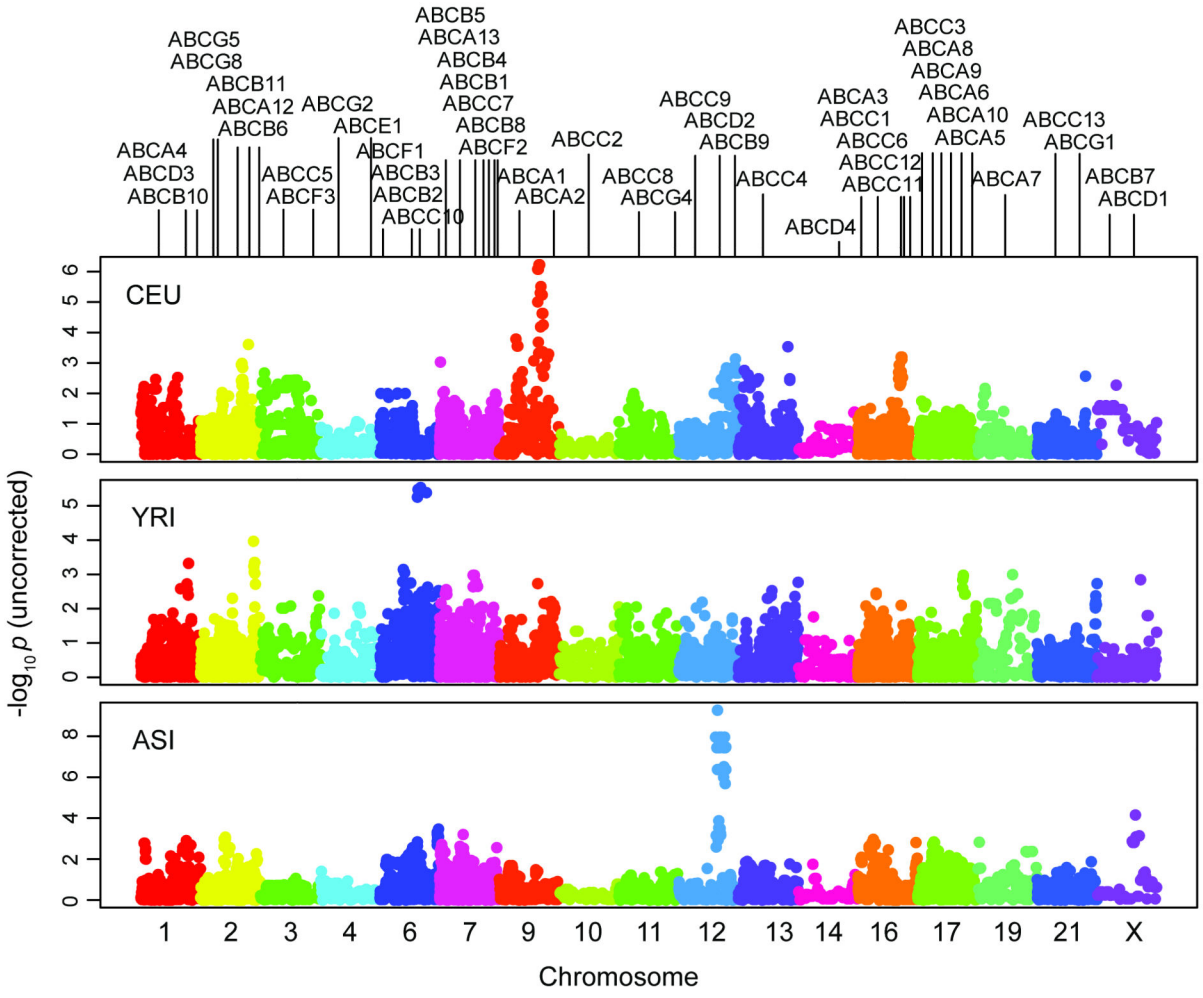

Fig. 4.

Genome-wide distribution of loci associated with ABC gene expression. The uncorrected significance of each SNP's association with expression levels in lymphoblastoid cell lines in the CEU, YRI and ASI populations is shown as a function of relative genomic position, divided per chromosome. The relative locations of human ABC genes are shown in the top panel.

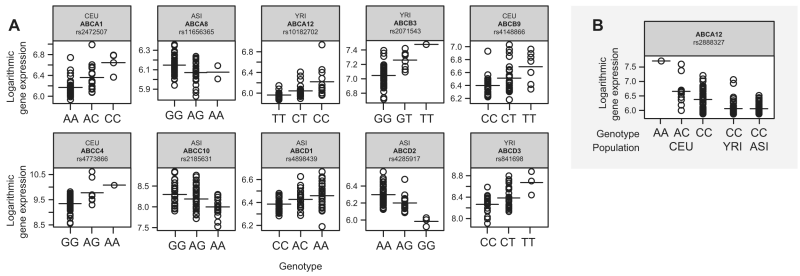

Fig. 5.

SNPs that significantly associate with expression levels of ABC transporters in lymphoblastoid cell lines from three populations. A: The most significant association across all populations is shown for each ABC gene with a significant local SNP association. B: Population differences in ABCA12 expression (see Fig. 3) can be attributed to a private SNP in the CEU population. Genotypes are shown for the coding strand.

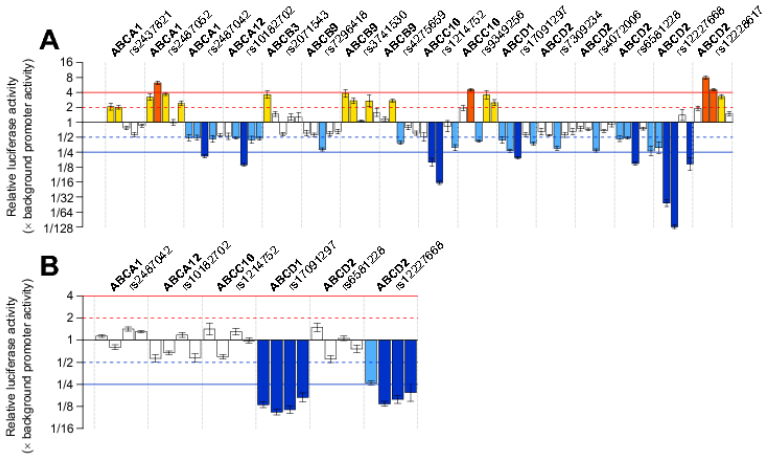

Fig. 6.

Regulation of luciferase reporter activity by genomic regions surrounding SNPs in or near ABC transporter genes. A: Putative regulatory elements were cloned into the weak promoter luciferase vector pGL4.23[*luc2*/minP]. B: Regions resulting in negative regulation in the weak promoter assay were assayed in an alternative vector containing the strong SV40 promoter (pGL3[Promoter]). Luciferase activity is shown relative to that of a vector only containing the promoter and the luciferase gene. The activity of each region was assessed in multiple cell lines (from left to right, for each construct: ACHN, HCT-116, HEK293T, HepG2 and Huh-7; the latter was only used in the weak promoter assay). Red, yellow, light blue, and dark blue bars: luciferase activity >4×, >2×, <0.5×, and <0.25× background promoter activity, respectively. Values are shown as mean ± S. E. M. from three to five separate experiments, each performed in triplicate.

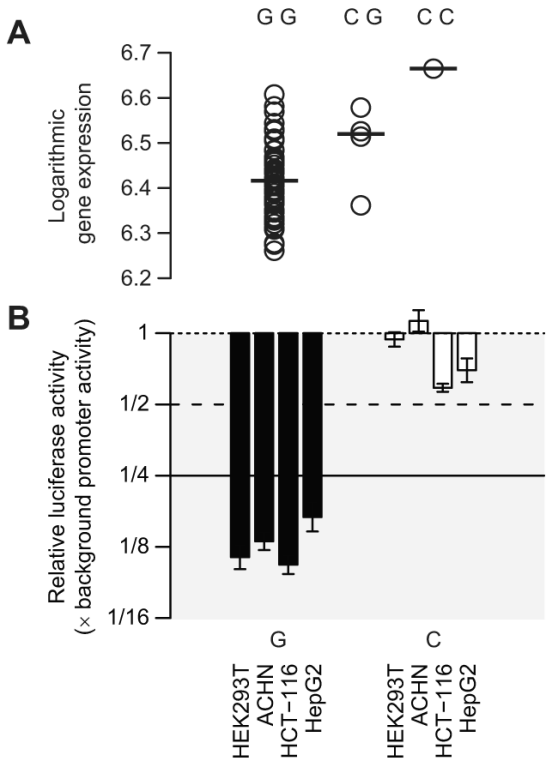

Fig. 7.

Association of a SNP in ABCD1 with expression levels in lymphoblastoid cell lines from the YRI population, and with luciferase reporter activity. A: The SNP rs17091297 is correlated with ABCD1 mRNA levels in the YRI population (Bonferroni-corrected $p < 0.05$). B: The major allele results in decreased luciferase activity compared to the control, when inserted upstream of a strong SV40 promoter. Luciferase activity is shown relative to the background promoter activity. Values are shown as mean ± S. E. M. from four separate experiments, each performed in triplicate.

Pharmacogenomics J. Author manuscript; available in PMC 2012 December 01.

Table 1

Number of genes expressed in lymphoblastoid cell lines from HapMap populations in the seven ABC transporter families.

Family	*n* genes interrogated (*n* in genome)	*n* genes expressed in CEU[1]	*n* genes expressed in YRI[1]	*n* genes expressed in ASI[1]	*n* genes expressed in any population[1]
ABCA	12 (12)	11	10	10	11
ABCB	11 (11)	10	10	9	10
ABCC	13 (13)	12	12	12	12
ABCD	4 (4)	4	4	4	4
ABCE	1 (1)	1	1	1	1
ABCF	3 (3)	3	3	3	3
ABCG	5 (5)	4	3	4	4
Total	49 (49)	45	43	43	45

[1] at 25th percentile of all interrogated probes or above

Pharmacogenomics J. Author manuscript; available in PMC 2012 December 01.

Table 2

Summary of SNPs significantly associated with ABC gene expression in lymphoblastoid cell lines from three populations.

Gene	Chromo-some	Tag SNP	Relative position[1]	*n* tagged SNPs	Alleles (M/m)[2, 3]	MAF CEU[4]	MAF YRI[4]	MAF ASI[4]	Most significant in population[5]	Directionality[3]	r^2	Association (unadjusted *p*)[5]	Association (Bonferroni *p*)[5]
ABCA1	9	rs2472507	In gene	4	T/G	0.20	0.06	0.35	CEU	M < m	0.38	8.3×10^{-7}	9.2×10^{-5}
ABCA1	9	rs2487030	−15 752	10	G/A	0.20	*n. p.*	0.38	CEU	M < m	0.33	5.8×10^{-6}	6.4×10^{-4}
ABCA1	9	rs2515629	In gene	1	T/C	0.12	0.19	0.10	CEU	M < m	0.24	1.6×10^{-4}	0.018
ABCA1	9	rs4149297	In gene	2	T/C	0.08	0.09	n. p.	CEU	M < m	0.23	2.8×10^{-4}	0.031
ABCA1	9	rs2000068	−27 872	2	T/C	0.38	0.17	0.36	CEU	M < m	0.21	4.3×10^{-4}	0.048
ABCA8	17	rs11656365	In gene	3	G/A	0.45	0.08	0.16	ASI	M > m	0.11	1.9×10^{-3}	0.049
ABCA12	2	rs2888327	In gene	1	G/T	0.13	*n. p.*	*n. p.*	CEU	M < m	0.22	2.5×10^{-4}	0.013
ABCA12	2	rs10182702	In gene	1	A/G	0.16	0.42	*n. p.*	YRI	M < m	0.25	1.1×10^{-4}	0.015
ABCB3/TAP2	6	rs2071543	−5 082	3	C/A	0.16	0.09	0.18	YRI	M < m	0.34	3.4×10^{-6}	2.4×10^{-4}
ABCB3/TAP2	6	rs9357155	−3 301	1	C/T	0.16	0.07	0.15	YRI	M < m	0.32	5.7×10^{-6}	4.0×10^{-4}
ABCB9	12	rs4148866	In gene	3	G/A	0.37	*n. p.*	0.44	CEU	M < m	0.18	1.3×10^{-3}	0.016
ABCB9	12	rs12425009	−42 292	7	A/G	0.34	*n. p.*	0.43	CEU	M > m	0.16	2.1×10^{-3}	0.027
ABCC4	13	rs4773866	In gene	1	G/A	0.09	0.07	*n. p.*	CEU	M < m	0.22	3.0×10^{-4}	0.049
ABCC10	6	rs2185631	In gene	1	G/A	0.49	0.13	0.44	ASI	M > m	0.14	3.5×10^{-4}	0.0045
ABCC10	6	rs12195350	−34 129	4	A/G	0.17	0.42	0.26	ASI	M > m	0.13	5.5×10^{-4}	0.0072
ABCC10	6	rs2487663	In gene	1	G/A	0.20	0.49	0.31	ASI	M > m	0.13	8.0×10^{-4}	0.010
ABCC10	6	rs9394952	In gene	9	A/G	0.46	*n. p.*	0.34	ASI	M > m	0.11	1.7×10^{-3}	0.022
ABCD1	X	rs4898439	17 452	1	C (A)	0.19	0.31	0.46	ASI	M < m	0.13	7.9×10^{-4}	0.0055
ABCD1	X	rs4898437	5 772	4	T (C)	0.18	0.31	0.42	ASI	M < m	0.12	1.4×10^{-3}	0.0096
ABCD1	X	rs17091297	−9 471	1	G (C)	*n. p.*	0.05	*n. p.*	YRI	M < m	0.18	1.4×10^{-3}	0.043
ABCD2	12	rs4285917	44 681	20	T (C)	0.16	0.05	0.15	ASI	M > m	0.33	1.1×10^{-8}	1.4×10^{-7}
ABCD2	12	rs4284427	In gene	10	A (G)	0.33	*n. p.*	0.33	ASI	M > m	0.16	3.1×10^{-4}	0.0040
ABCD2	12	rs10783969	−39 188	13	C (A)	0.13	0.41	0.33	CEU	M < m	0.16	2.2×10^{-3}	0.040
ABCD3	1	rs841698	34 435	1	C (T)	*n. p.*	0.32	n. p.	YRI	M < m	0.21	4.8×10^{-4}	0.028

[1] Negative positions are relative to the transcription start site and positive positions to the end of the longest transcript in the UCSC Genome Browser hg18 release.

[2] Genotypes are given for the coding strand.

[3] M: major allele, m: minor allele.

[4] *n. p.*, SNP not detected in population.

[5] All genotype dependent expression and association data refer to the population with the strongest association.

Table 3

Regulatory regions in ABC transporter genes identified experimentally. Regions were designated enhancers and silencers if they resulted in a more than fourfold increased or decreased reporter activity compared to the negative control, or moderate enhancers if the increase was more than twofold (see Figure 6).

Gene	Transporter function	SNP	Cloned region[1]	Designated as
ABCA1	Cholesterol transport	rs2487052	−216 - +245	Enhancer
ABCB3/TAP2	Peptide transport, antigen presentation	rs2071543	−230 - +232	Moderate enhancer
ABCB9	Lysosomal peptide transport	rs3741530	−247 - +238	Moderate enhancer
ABCC10	Drug resistance	rs9349256	−232 - +186	Enhancer
ABCD1	Peroxisomal fatty acid transport	rs17091297	−233 - +256	Silencer
ABCD2	Peroxisomal fatty acid transport	rs12227668	−275 - +189	Silencer
ABCD2	Peroxisomal fatty acid transport	rs12228617	−303 - +194	Enhancer

[1] Positions are given on the coding strand, relative to the location of the SNP.

Table 4

Direction of association of selected major (M) versus minor (m) alleles of ABC transporter genes with expression levels in lymphoblastoid cell lines and with lucifierase activities. In four cases, the direction of association of expression level agreed with the reporter assay, e.g., the major allele of SNP rs2071543 was associated with a lower expression level and reporter activity.

Gene	SNP	Alleles (M/m)[1, 2]	Association (unadjusted *p*)	Association directionality[2]	Association with expression of flanking gene (Unadjusted *p*)	Reporter assay significant cell line[3]	Reporter assay directionality[2]	Predicted TF binding directionality[2, 4, 5]
ABCA1	rs2487052	G/A	8.3×10^{-7}	M < m	-	*n. s.*	M < m	*n. a.*
ABCB3/TAP2	rs2071543	C/A	3.4×10^{-6}	M < m	ABCB2/TAP1 ($p<0.02$)	AC	M < m	*n. a.*
ABCB9	rs3741530	A/C	2.1×10^{-3}	M > m	ARL6IP4 ($p<0.009$)	AC	M < m	*n. a.*
ABCC10	rs9349256	A/G	1.9×10^{-3}	M > m	-	HG, HU	M < m	M < m (HNF-4)
ABCD1	rs17091297	G/C	1.4×10^{-3}	M < m	IDH3G ($p<0.05$)	AC, HE, HC, HG	M < m	M < m (AP-4)
ABCD2	rs12227668	A/G	1.1×10^{-8}	M > m	-	AC, HE, HC	M < m	M < m (HSF)
ABCD2	rs12228617	C/A	3.0×10^{-7}	M > m	C12orf40 ($p<0.04$)	*n. s.*	M > m	*n. a.*

[1] Genotypes are given for the coding strand.

[2] M: major allele, m: minor allele.

[3] AC: ACHN, HC: HCT-116, HE: HEK293T, HG: HepG2, HU: Huh-7, *n. s.*: not significant.

[4] Altered transcription factor binding predicted using TF-Search (http://www.cbrc.jp/research/db/TFSEARCH.html).

[5] *n. a.*, transcription factor sites with allele-specific binding were not detected within the cloned region