Functional Evaluation of Genetic and Environmental Regulators of P450 mRNA Levels

Dazhi Wang1,2, Zhengwen Jiang2, Zhongyang Shen3, Hui Wang3, Beilan Wang2, Weihua Shou1,2, Hong Zheng3, Xun Chu1,2, Jinxiu Shi2, Wei Huang1,2

1 Rui Jin Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China, 2 Department of Genetics, Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center, Shanghai, China, 3 Organ Transplant Center, Tianjin First Central Hospital, Tianjin, China

Abstract

Variations in the activities of Cytochrome P450s are one of the major factors responsible for inter-individual differences in drug clearance rates, which may cause serious toxicity or inefficacy of therapeutic drugs. Various mRNA levels are a major factor for different activity of the major P450 genes. Although both genetic and environmental regulators of P450 gene expression have been widely investigated, few studies have evaluated the functional importance of cis- and trans-regulatory factors and environmental factors in the modulation of inter-individual expression variation of the P450 genes. In this study, we measured the mRNA levels of seven major P450 genes (CYP1A1, CYP1A2, CYP2C9, CYP2C19, CYP2D6, CYP3A4 and CYP3A5) in 96 liver biopsy samples from Chinese population. Both trans-acting (mRNA levels and non-synonymous SNPs of putative regulator genes) and cis-acting (gene copy number and functional SNPs) factors were investigated to identify the determinants of expression variation of these seven P450 genes. We found that expression variations of most P450 genes, regulator genes and housekeeping genes were positively correlated at the mRNA level. After partial correlation analysis using ACTB and GAPDH expression to eliminate the effect of global regulators, a UPGMA (Unweighted Pair Group Method with Arithmetic Mean) tree was constructed to reveal the effects of specific regulation networks potentially masked by global regulators. Combined with the functional analysis of regulators, our results suggested that expression variation at the mRNA level was mediated by several factors in a gene-specific manner. Cis-acting genetic variants might play key roles in the expression variation of CYP2D6 and CYP3A5, environmental inducers might play key roles in CYP1A1 and CYP1A2 variation and global regulators might play key roles in CYP2C9 variation. In addition, the functions of regulators that play less important roles in controlling expression variation for each P450 gene were determined.

Introduction

The Cytochrome P450 superfamily is a very large and diverse group of hemoproteins that is present in almost all living organisms. About 57 putative functional Cytochrome P450 genes and more than 58 pseudogenes have been identified in the human genome [1,2]. These P450 genes encode enzymes, such as mono-oxygenase, which play a crucial role in detoxification of xenogenous xenobiotics, decomposition of drugs and metabolism of many endogenous compounds, such as hormones, fatty acids, prostaglandin, cholesterol and vitamin D [3,4]. It is well recognized that inter-individual variability in activity of the P450 enzymes is a major factor responsible for inter-individual variations in drug clearance rates [3]. Cytochrome P450 enzymes, which play key roles in Phase I oxidative metabolism, have been extensively investigated due to their great variations in activity in the human population. The inter-individual variability in the total activity of P450s is primarily caused by polymorphisms that affect activity and expression. Strong correlations between expression level and enzyme activity have been observed for CYP1A1, CYP1A2, CYP3A4, CYP2C8, CYP2C9, CYP2D6 and CYP2B6, suggesting that the mRNA level may be a major determinant of the total activity of these P450 genes [6,7].

Ten-to a hundred-fold inter-individual differences in expression levels have been observed for most Cytochrome P450 genes [6,8]. Most of the P450s that metabolize exogenous compounds are highly expressed in the human liver, but some P450 forms are expressed at low levels in extrahepatic tissues. The mechanisms underlying the maintenance and regulation of the high expression levels of the P450 genes in the liver are not completely understood. Data have shown that the expression levels of some liver-enriched transcription factors, such as the Constitutive Androstane Receptor (CAR), Hepatic Nuclear Factor 4α (HNF4α) and P450 Oxidoreductase (POR), may determine the variability in the basal expression and activity of a broad range of P450 genes [8]. In addition, P450 gene-specific mechanisms have been identified. For example, activation of the aryl hydrocarbon receptor (AhR)-mediated pathway induces an increase in expression of the CYP1A1 and CYP1A2 genes [9,10]; for CYP2D6 and CYP2A6, copy number variation and some regulatory alleles have been reported.
to be important for expression regulation [11,12]. A polymorphism in CYP3A5, the CYP3A5*3 allele, is the major factor that modulates expression [13]. Despite the tremendous amount of research that has been performed to construct the regulatory network of P450 genes, few studies have evaluated the functional importance of various factors in controlling P450 expression variation among individuals.

In this study, we systematically evaluated the determinants of P450 expression variation among individuals. We measured the absolute expression levels of three housekeeping genes (18S rRNA, GAPDH and ACTB), seven P450 genes (CYP1A1, CYP1A2, CYP2C9, CYP2C19, CYP2D6, CYPA4 and CYP3A5) and seven putative Cytochrome P450 regulator genes (USF1, CAR, PXR, HNF4A, HNF1A, AHR and ARNT) [for full names of these genes, refer to Table S1]. Pairwise correlation analysis was performed to explore the network of interaction among these genes. The gene-gene interactions and gene-environment interactions made it difficult to distinguish the effects of each regulatory factor. Therefore, allelic expression ratios (AERs) of CYP1A1, CYP1A2, CYP2C9, CYP2C19 and CYPA4 were measured using two SNP markers from each gene to determine the presence of cis-acting regulatory variants [14–17]. The copy number of CYP2D6 was determined, and the correlation between the copy number and mRNA level of CYP2D6 was examined. Fifteen SNPs located in the seven P450 genes and three regulator genes (Table S2) were typed and tested for association with the expression levels of the corresponding P450 genes.

**Results**

**Expression variations in the P450s and regulatory genes**

The mRNA levels of two putative housekeeping genes (GAPDH and ACTB), seven P450 genes (CYP1A1, CYP1A2, CYP2C9, CYP2C19, CYP2D6, CYPA4 and CYP3A5) and seven regulatory genes (USF1, CAR, PXR, HNF4A, HNF1A, AHR and ARNT) were determined by quantitative real-time PCR and normalized to 18S rRNA expression. Up to 8- to 10-fold inter-individual differences were observed for GAPDH and ACTB (Table S4). This result raised doubts about the applicability of GAPDH and ACTB as reference genes. Thus, 18S rRNA was selected as the reference gene based on two observations: 1) the 18S rRNA level had the lowest CV (coefficient of variation) as shown in Fig. 1 and 2) the 18S rRNA level is strongly correlated (r = 0.76) with the cDNA level of USF1, which had the second lowest CV and encodes a ubiquitous transcription factor.

In our samples, all regulatory genes were constitutively transcribed at low levels with low inter-individual variability (<7-fold), whereas most P450 genes showed much greater inter-individual variability (>13-fold), except for CYP2C9 (7.2-fold). The expression levels of the seven P450 genes were CYPA4>CYP2C9>CYP2D6>CYP1A2>CYP2C3>CYP2D6>CYP1A2>CYP2C9>CYP2A5 (Table S4). The expression variations of CYP1A1>CYP2D6>YCY>CYP2C9>CYP2C19>CYP2A5 and CYP1A2>CYP2C9 (Table S4). One sample with no CYP2D6 expression was excluded from the variation comparison due to the loss of both CYP2D6 alleles (as discussed later). In some individuals, extreme differences in expression levels were observed for CYPA4, CYP2D6 and CYP1A1 by comparison of the T5/B5 and Max/Min ratios. The T5/B5 ratio denotes the ratio of the average expression level of the top 5% of samples to that of the bottom 5% of samples. There was more than a 5-fold difference between these two ratios for CYPA4 (Table S4).

**Correlations among mRNA levels of P450 and regulator genes**

Interestingly, the expression levels of almost all of the genes (including GAPDH and ACTB) were strongly correlated with each other at the mRNA level (Table S5). Such correlations may be the result of powerful global regulators. Thus, we conducted a partial correlation analysis using the mRNA levels of both GAPDH and ACTB to eliminate or reduce these effects. After the partial correlation analysis, the degree of most correlations was largely reduced, and many were not significant (p>0.05) (Table S6). However, strong correlations were still observed between CYP1A1 and CYP1A2 (r = 0.57); CYP2C9 and CYP2A5 (r = 0.57); and among ARNT, HNF1A, AHR, PXR and CAR (r = 0.44). These correlations remained significant after adjustment for age and smoking history in multivariate logistic regression analysis (p<0.001). A UPGMA (Unweighted Pair Group Method with Arithmetic Mean) tree (Fig. 2) was constructed based on the partial correlation matrix, and three clusters were roughly defined. The first cluster was composed of five regulator genes (HNF1A, HNF4A, CAR, PXR and ARNT), the second contained CYP2C9, CYPA4 and CYP2C19, and the third included only CYP1A1 and CYP1A2.

**Copy number variation and expression level of CYP2D6**

The copy number of CYP2D6 varied from 0 to 9 among the 92 DNA samples (Fig. 3). Four samples were excluded due to inconsistency between measurements using different CYP2D6 primers or without mRNA. One sample containing a homozygous CYP2D6 deletion (CYP2D6<1) was identified. The mRNA level of CYP2D6 is plotted against copy number in Fig. 3. Higher copy numbers tended to increase the expression level in samples with ≤4 copies. A strong correlation (r = 0.63, p<0.0001) was observed between copy number and expression level for samples with no more than four copies of CYP2D6. The correlation remained significant after adjustment for age and smoking history in multivariate logistic regression analysis (p<0.001).

**Association analysis between genetic variants in P450s and regulator genes and mRNA levels of seven P450 genes**

Expression variations in the P450 genes may result from the different activities of regulator protein isoforms in addition to different protein levels. All three non-synonymous SNPs in the regulatory genes mentioned above with MAF (minor allele frequency)>5% in the Chinese population were genotyped in 96 liver samples. These three SNPs, two (rs1169288 and rs2464196) in HNF1A and one (rs2066853) in AHR, were not associated with the expression levels of any of the seven P450 genes examined, which indicates that the polymorphisms of HNF1A and AHR do not contribute to the expression variations of these seven P450 genes. In total, 12 marker SNPs in the P450 genes (Table S2) used for allelic expression imbalance (AEI) detection were tested for association with mRNA levels. Rs767746 (also known as 6968G>A in P450s, as an A>G substitution in intron 3 that causes aberrant splicing, was strongly correlated with the mRNA level of CYPA5 (p = 0.506, Fig. 4). The correlations remained significant after adjustment for age and smoking history in multivariate logistic regression analysis (p<0.001). Marginal associations were observed between rs17861162 and the CYPA4 mRNA level (p = 0.044); rs1934967 and the CYP2C9 mRNA level (p = 0.025) and rs33972293 and the CYPA4 mRNA level (p = 0.036) by the Kruskal-Wallis test but not by ANOVA. There was also a marginal association between rs4244825 and the CYP2C9 mRNA level (p = 0.047 detected by ANOVA but not by the Kruskal-Wallis test. These associations were no longer
significant after adjustment for age and smoking history in multivariate logistic regression analysis \( p_{adj} < 0.05 \).

**Discussion**

Expression variations of P450s and regulator genes

*CYPIA1* showed the largest inter-individual difference in expression (138-fold, Table S4), which is consistent with the observation that expression of *CYPIA1* was highly inducible by environmental inducers. The levels of inter-individual expression variability observed in this study for most of the *P450* genes and some regulatory genes were much lower than those found in several previous studies [6,8]. For example, an approximately 6-fold difference was detected in our samples for *CAR*, but Wortham et al. [8] reported a 145-fold difference. This study found differences of approximately 14- and 64-fold for *CYPIA2* and *CYP5A4*, respectively, whereas differences of 582- to 3638-fold were reported for *CYPIA2* and differences of 860- to 1281-fold were reported for *CYP3A4* by Wortham et al. [8] and Rodriguez-Antona et al. [6], respectively. Some studies have found levels of variability similar to our results, such as differences of 118-fold and approximately 15-fold for *CYP3A4* [2] and *CYPIA2* [18], respectively.

![Figure 1. Coefficients of variation of transcript molecule numbers in 96 Chinese liver biopsy samples for 17 genes tested.](https://doi.org/10.1371/journal.pone.0024900.g001)

![Figure 2. A UPGMA tree showing the relationships between 14 genes based on expression levels.](https://doi.org/10.1371/journal.pone.0024900.g002)
The inconsistency among reports may be attributed to differences in ethnic groups, environmental factors, RNA quality and experiment design. In our study, the samples were obtained from Chinese Han healthy liver donors and treated with the protocols recommended by Qiagen. Most primers were designed in a region close to the 3' end of the mRNA to amplify a short fragment of 60–180 bp in length, which should reduce the variation resulting from low RNA quality. Moreover, all SNPs in the sequences used for primer design were masked first by performing a BLAST search of the SNP database; therefore, the primers did not harbor any SNP, which prevented differences in amplification efficacies due to cDNA samples with different genotypes at SNPs located within primer sequences.

In our study, measures were taken to reduce the noise effects of non-genetic factors. Besides the genetic factors, the inter-individual mRNA expression variations of P450 may be due to differences in gender, age, food intake, drug treatment, smoking, alcohol using etc. All the subjects recruited in this study were males and between 30 to 45 years old. According to the donors' statements, no drugs were taken before the surgeries for at least 30 days. Alcohol using and smoking were prohibited 30 days prior to surgery, which avoided the confounding by alcohol using and smoking. Multivariate logistic regression was used to adjust for potential confounding factors including age and smoking history.

However, differences in the food intake of the 96 participants were not considered, which might be a resource of bias in our study. Additionally, the expression variation caused by rare alleles may not be evident in our limited sample size. Thus, further studies are needed to demonstrate the expression variations of the P450s in larger sample sizes to eliminate bias.

Coexpression analysis

Unexpectedly, strong correlations were detected among almost all genes examined in this study, including GAPDH and ACTB, which were not thought to be involved in the P450 regulation network (Table S5). Yang et al. [7] suggested that liver P450s are situated at the center of many endocrine and xenobiotic metabolic pathways, which require cross talk among numerous nuclear receptor networks. Alternatively, it is also reasonable to assume that this correlation may be caused by powerful global regulatory factors that control expression of most active genes, which would mask the true regulation network of P450s and the corresponding regulator genes. Such global regulatory factors are most likely basic transcription factors and ubiquitous chromatin modifiers. We performed a partial correlation analysis using the expression levels of GAPDH and ACTB as controls for the elimination of the presumed effect. Not surprisingly, 39% of the correlations were eliminated (Table S6). Based on the remaining “corrected” correlations, we performed a cluster analysis and constructed the UPGMA tree with MEGA 4 [19]. Finally, three clusters were roughly identified (Fig. 2).

In cluster I, strong correlations remained among CAR, PXR, HNF1A, HNF4A and ARNT (Fig. 2), which suggested that there was co-regulation or interaction among these genes. HNF1A,
which is regulated by *HNF4A* [20,21], displayed the three strongest correlations with *CAR* (r = 0.67), *PXR* (r = 0.66) and *ARNT* (r = 0.63), indicating that *HNF1A* may be important for regulating the expression of *CAR*, *PXR* and *ARNT*. Of all of the correlations between *HNF4A* and other regulator genes, the correlation between *HNF4A* and *HNF1A* was the strongest (r = 0.545), which is consistent with the observation that *HNF1A* is regulated by *HNF4A*. Regulation of *PXR* by *HNF1A* is supported by a previous report that a putative *HNF1A* binding site is present in the human *PXR* promoter and a 6-bp deletion at this site diminished the activity of hPAR-2, one of the *PXR* transcripts [22]. However, limited data have uncovered a relationship between *HNF1A* and *CAR* or *ARNT*. Here, we propose a simplified regulation network among *CAR*, *PXR*, *HNF1A*, *HNF4A* and *ARNT*, in which *HNF4A* regulates the expression of *HNF1A*, and *HNF1A* further regulates the expression of *CAR*, *PXR* and *ARNT*. Other interactions among *HNF4A*, *ARNT*, *CAR* and *PXR* are possible, and more data must be collected to evaluate this hypothesis. Clusters II and III showed co-expression of the *P450* genes. Because *P450s* are evolved from a common ancestor gene, the correlations among the *P450* genes [1,7,23] might result from the shared regulatory factors. *CYP1A1* and *CYP1A2* in cluster III are orientated in a head-to-head configuration on 15q22-q24. The 23.3 kb intergenic region might harbor certain element with a role in co-regulation of the expression of *CYP1A1* and *CYP1A2* [24]. However, the correlation did not entirely agree with the sequence similarity [1]. The effects of specific regulatory factors on each *P450* gene are discussed in detail below.

**CYP1A1 and CYP1A2**

*CYP1A2* is constitutively expressed in the liver. *CYP1A1* is induced by various types of environmental chemicals, such as PAHs (polycyclic aromatic hydrocarbons), HAAs (heterocyclic aromatic amines amides) and HAHs (halogenated aromatic hydrocarbons) via the ligand-activated AHR-ARNT pathway. A strong correlation (p = 0.81) between the expression levels of *CYP1A1* and *CYP1A2* demonstrated that the two *CYP1A* genes are co-regulated. However, the mRNA levels of *CYP1A1* and *CYP1A2* were not correlated with those of most of the regulator genes, except for the weak correlations observed between *CYP1A1* and *USF1* (p = 0.037) and between *CYP1A2* and *AHR* (p = 0.025), *CAR* (p = 0.024) and *USF1* (p = 0.009), which indicate that environmental inducers involved in activation of the AHR-ARNT pathway may be an important determinant for the variability in mRNA levels of *CYP1A1* and *CYP1A2*. Quantitation of the active and total AHR protein levels may provide direct evidence to support this hypothesis. Cis-acting variants may also account for the variations, albeit it to a small extent, as described in Fig. 5. These cis-acting variants were estimated to induce inter-individual expression differences of only <2.5-fold for *CYP1A1* and <1.7-fold for *CYP1A2*, compared with a total variability in mRNA level of 138-fold for *CYP1A1* and 14-fold for *CYP1A2*. After exclusion of

---

Figure 4. Correlation between genotypes at rs776746 and the expression level of CYP3A5. A is the wild-type allele and G is the mutant allele (CYP3A5*3), which results in aberrant splicing. The natural logarithm of the mRNA data was used for ANOVA analysis. p denotes the Spearman’s correlation coefficient. The Boxplot displays the minimal, lower quartile, median, upper quartile, maximal points and outliers (marked by circle and sample No.).

doi:10.1371/journal.pone.0024900.g004
four outlier samples for the mRNA levels of CYP1A2, the inter-individual expression difference was about 9-fold, and thus the cis-acting genetic variants may account for up to 20% of the inter-individual variability. Therefore, the major determinants for variability in the expression levels of CYP1A1 and CYP1A2 may include environmental factors involved in activation of the AHR-ARNT pathway, global regulators as described above and, to a lesser extent, cis-acting variants.

CYP2C9

CYP2C9 is the most abundant CYP2C subfamily isozyme in human liver [25–27]. Expression of CYP2C9 has been reported to be regulated by a wide range of liver-enriched transcription factors, such as CAR, PXR and HNF4A [28]. Significant correlations between the mRNA levels of CYP2C9 and all seven regulator genes, especially CAR, PXR, ARNT and HNF1A (p>0.5), were observed, which provides evidence that expression of putative P450 regulator genes contributes to the inter-individual variability of CYP2C9 expression levels. Although HNF4A has been reported to regulate CYP2C9 expression, our results suggest that HNF1A may play a more important role in the regulation of CYP2C9 expression than HNF4A because a stronger correlation was observed between the mRNA levels of CYP2C9 and HNF1A than between those of CYP2C9 and HNF4A (Table S4). In addition, our study shows for the first time that ARNT may be an important regulator of CYP2C9. Strong correlations among CYP2C9, CYP2C19 and CYP3A4 support the reports that these three genes are co-regulated by CAR and PXR [29]. No significant common cis-acting genetic variants were detected for CYP2C9 (Fig. 5), which is consistent with previously reported data [17]. In conclusion, of the seven P450 genes examined, CYP2C9 displayed the least expression variation in 96 Chinese liver biopsy samples, and the variability in expression was mainly determined by expression of P450 regulator genes, such as CAR, PAR, HNF1A and ARNT, in addition to the global regulators as described above. Environmental inducers may also account for some variation in CYP2C9 expression by activating CAR or PXR.

CYP2C19

Up to 45-fold inter-individual differences were detected for CYP2C19 among the 96 Chinese liver samples. Only low correlations between CYP2C19 and several regulator genes (PXR, CAR, ARNT and HNF1A) were observed. The genetic variant CYP2C19*2 at rs4244285, which results in aberrant splicing, may account for up to 2-fold allelic expression difference (Fig. 5); however, only a weak association was detected between rs4233285 and the mRNA levels of CYP2C19 by ANOVA. Therefore, the inter-individual variability in the mRNA levels of CYP2C19 (up to 45-fold) may be largely determined by factors other than cis-acting genetic variants or the mRNA levels of regulator genes.

CYP2C19 is highly induced by drugs, such as rifampin, via activation of the CAR/PXR-mediated pathway [29]. Strong correlations were found among the mRNA levels of three target genes of CAR/PXR: CYP2C19, CYP2C9 and CYP3A4. However, slight correlations between the mRNA levels of CYP2C19 and CAR and PXR suggest the importance of CAR/PXR activation by CYP2C19 inducers in modulating the expression of CYP2C19. Quantitation of the active and total CAR or PXR protein levels may provide further evidence for this hypothesis. Therefore, the major determinants for the inter-individual variability in the mRNA levels of CYP2C19 may include global regulators and environmental inducers for activation of the CAR/PXR pathway. To a lesser extent, cis-acting genetic variants or the mRNA levels of CAR and PXR may account for a small part of the expression variation of CYP2C19.
CYP2D6

A substantial inter-individual difference was identified (up to 68-fold without inclusion of a sample homozygous for CYP2D6 deletion) at the mRNA level for CYP2D6. Slight correlations were observed between CYP2D6 and several regulator genes (PXR, CAR, HNF4A and ARNT) at the mRNA level, demonstrating that the mRNA levels of regulator genes contribute to the variability in expression of CYP2D6 (Table S6), albeit to a much lesser extent compared with CYP2C9. The global factor that affected expression of most of the other genes played a less important role in CYP2D6 expression (Table S6). The copy number of functional CYP2D6 was the major determinant of the variability in the mRNA levels of CYP2D6 (Fig. 3). The P450 regulator genes and global regulators may make some contributions but to lesser extents.

The correlation between the copy number and mRNA level of CYP2D6 was not consistent when the copy number is more than four in this study. The inconsistency may attribute to two reasons. First, the inconsistency may be caused by the insufficient sample size. Thirteen samples with more than four copies of CYP2D6 were investigated in current study. Even larger sample size is needed in the further study to validate whether this inconsistency is true. Second, the CYP2D6 gene duplications include functional, partly functional and nonfunctional genes, which could cause splicing defect or produce unstable mRNA [30–32]. It is possible that the samples in our study with more than four CYP2D6 copies contain partly functional or nonfunctional gene copies, which results in this inconsistency. Further long range PCR and clone sequencing are needed to clarify this issue.

CYP3A4

High expression levels and high inter-individual expression differences were observed in CYP3A4. The slight correlations observed between CYP3A4 and several regulator genes (PXR, CAR, HNF1A and ARNT) at the mRNA level indicated that regulator genes have moderate effects on CYP3A4 expression variation (Table S6). CYP3A4 is highly induced by drugs, such as rifampin, via activation of the CAR/PXR-mediated pathway. Similarly to CYP2C19, the determinants for the expression variability of CYP3A4 may be global regulators, environmental inducers that activate CAR/PXR and, to a lesser extent, the mRNA levels of the CAR and PXR genes. Some rare cis-acting genetic variants might induce up to 2-fold inter-individual expression differences; however, only a very small proportion of samples (4%) was affected (Fig. 5). Thus, these genetic variants contribute little to the variability in the mRNA levels of CYP3A4. In summary, the determinants for the expression variability of CYP3A4 may be global regulators, environmental inducers that activate CAR/PXR and, to a lesser extent, the mRNA levels of the CAR and PXR genes.

CYP3A5

No correlations were observed between the mRNA levels of CYP3A5 and the regulatory genes. A strong association (p<0.0001, Fig. 4) was identified between the mRNA levels of CYP3A5 and rs776746, which is the CYP3A5*3 allele that induces aberrant splicing of CYP3A5. The unstable mRNA produced by aberrant splicing decreases the expression level. Thus, this cis-acting element is another determinant of CYP3A5 expression variation in addition to global regulators.

Based on the data reported here, we propose that at least four factors are determinants of the expression variation of P450 genes (Table 1): I) global regulators that affect the expression levels of most active genes, which are most likely basic transcription factors and ubiquitous chromatin modifiers; II) P450-specific regulators; III) environmental inducers responsible for the activation of P450-specific regulators and IV) genetic variants of the P450 genes, including gene duplication or deletion. In our study, we weighted the functional importance of each factor on the expression variation of each P450 gene.

In summary, we functionally evaluated several factors that control the variability in the mRNA levels of CYP1A1, CYP1A2, CYP2C9, CYP2C19, CYP2D6, CYP3A4 and CYP3A5. Moreover, the variations in the mRNA levels of 7 P450 genes characterized in 96 Chinese liver samples in this study provide useful data on the liver P450s in Asian populations, which traditionally have not been analyzed in European-based studies.

Although strong correlations were observed between the mRNA level and certain variants (CYP3A5 and rs776746; CYP2D6 and the copy number), such genetic variation is not good enough to be used clinically in the individual patient. The activities of drug-metabolizing enzymes are regulated at many levels except for genetic factors. Further studies should be done before the application of the research conclusions in clinical practice.

**Materials and Methods**

**Ethics Statement**

All participants gave written informed consent. The acquisition of all data using samples from these participants was approved by the Ethics Review Committee of the Chinese National Human Genome Center at Shanghai.

**Human Liver Tissues**

A total of 96 human liver biopsy samples were obtained in Tianjin First Center Hospital from donor livers that were morphologically normal and tested negative for HIV (human immunodeficiency virus) and hepatitis. All the subjects recruited in this study were males and between 30 to 45 years old. According to the donors’ statements, no drugs were taken before the surgeries for at least 30 days. Alcohol using and smoking were prohibited 30 days prior to surgery. About 50–200 mg of tissue for each sample was immediately submerged in RNA Later (Sigma-Aldrich) after collection from live livers during liver transplant surgery. After storage overnight in RNA Later at 4°C, all samples were stored at −70°C until the DNA and RNA extractions were performed.

**Isolation of DNA and Total RNA from Human Liver Tissues**

Tissues in RNALater were thawed at room temperature, washed with DEPC-treated water and homogenized in 1 ml Trizol (Invitrogen) using a Pro200 homogenizer (Pro Scientific Inc.). Both DNA and total RNA were extracted according to the manufacturer’s protocol. Extracted DNA and total RNA were quantitated using a Biophotometer (Eppendorf). Total RNA was purified by DNase I (Takara) treatment followed by phenol/chloroform extraction and ethanol precipitation. Purified RNA was dissolved in DEPC-treated water, quantitated, diluted to approximately 1 μg/μl and stored at −70°C.

**Absolute Quantitation of mRNA level by Real-time PCR**

All primers for real-time quantitative PCR were designed to specifically amplify fragments of 80–180 bp in length inside single exons using the online program Primer3 (http://frodo.wi.mit.edu/ primer3/input.htm). The sequence for each primer is listed in Table S3. Reverse transcription reactions were prepared in a total volume of 20 μl containing 1 × RT buffer, 0.5 mM dNTPs, 0.5 μM poly24dT, 5 μM N9, 20 U RNase inhibitor (BIO BASIC INC.) and 100 U M-MLV(H-) reverse transcriptase (Promega). The
measurements obtained using different primer pairs were averaged to obtain the quantity of CYP2D6 molecules. The CYP2D6 copy number was estimated by rounding the ratio of the quantity of CYP2D6 molecules to that of RPPH1.

### SNP Genotyping

In total, 15 SNPs in 9 genes were selected and typed in 96 DNA samples extracted from liver tissue samples using the Multiplex SNaPshot kit (Applied Biosystems Inc.) (Table S2). For the CYP1A1, CYP1A2, CYP2C9, CYP2C19 and CYP3A4 genes, two SNPs within the gene with low linkage disequilibrium (LD) ($r^2<0.5$) and high minor allele frequency (MAF) (>20%) were selected. LD and MAF were calculated from genotyping data released by the HapMap project (http://www.hapmap.org). These ten SNPs were also typed in heterozygous cDNA samples as markers to determine allelic expression ratios. Rs776476 in CYP3A5 contains the CYP3A5*5 allele, which causes improper splicing and results in a nonfunctional, truncated protein. The selected SNPs in HNF1A and AHR are non-synonymous (rs1169288, rs2464196 and rs2066853). Three multiple PCR and extension panels were designed for the SNaPshot reactions. A 10 μl mixture was prepared for each multiple PCR reaction and included 1 x HotStarTaq buffer, 2.8 mM Mg<sup>2+</sup>, 0.2 mM dNTP, 0.1–0.3 μM primers, 0.3 U HotStarTaq polymerase (Qiagen Inc.) and 1 μl template DNA. The cycling program was 95°C for 15 min; 11 cycles of 94°C for 20 s, 62°C –0.5°C/cycle for 40 s and 72°C for 80 s; 25 cycles of 94°C for 20 s, 56°C for 30 s and 72°C for 80 s and 72°C for 5 min. The PCR products were purified with the Exo I and SAP enzymes, and subsequent SNaPshot extension was carried out as described in the manufacturer’s protocol. Extension products were separated by gel electrophoresis and analyzed using the ABI3730XL (Applied Biosystems) system. Data collected on the ABI3730XL system were analyzed with Genemapper 4.0.

### Measurement of allelic expression ratios

All heterozygous cDNA samples and 5 heterozygous DNA samples for each marker SNP in the CYP1A1, CYP1A2, CYP2C9, CYP2C19 and CYP3A4 genes were subjected to single SNaPshot reactions. Four-fold dilutions of the cDNA samples were used for amplification of fragments containing intronic SNPs, and twenty
The allelic expression imbalance (AEI) assay was carried out by detecting the expression ratio of two alleles in heterozygous cDNA samples in which the expression level of the two alleles were controlled by each other. Fig. 6 illustrates the basic concept of this assay using the single fluorescent-labeled nucleotide extension method (SNaPshot, ABI). Theoretically, this assay can significantly reduce the noise that results from variations in activity of regulatory factors and improve the sensitivity of detection.

Statistical analysis

SPSS VERSION 15.0 (SPSS Inc., Chicago, IL) was used for statistical analyses and for BOXPILOT construction. Both the Kruskal-Wallis test and ANOVA were performed to examine the associations between genotype and mRNA content. Pearson’s or Spearman’s rho correlation coefficients between mRNA levels or between the copy number and mRNA level of CYP2D6 were identified using bivariate regression analysis. Multivariate logistic regression was used to adjust for age and smoking history. The UPGMA tree for seven P450s and seven regulator genes was constructed using the MEGA4 program based on the pairwise Spearman’s rho correlation coefficient data.

Supporting Information

Table S1 Abbreviation of gene names. (DOC)
Table S2 Characterization of SNP information and primer sequence for multiplex SNaPshot reactions. (DOC)
Table S3 Primers for quantitative real time PCR. (DOC)
Table S4 Inter-individual variation in the mRNA level of housekeeping, P450 and regulatory genes (Normalized to 18SrRNA). (DOC)
Table S5 Correlations among the mRNA levels of two housekeeping genes, seven P450 genes and seven regulator genes. (DOC)
Table S6 Partial correlations among P450 genes and regulator genes by controlling on GAPDH and ACTB in mRNA level. (DOC)

Author Contributions

Conceived and designed the experiments: WH. Performed the experiments: DZW ZWJ JXS. Contributed reagents/materials/analysis tools: WH DZW ZWJ ZYS HW BLW WHS HZ. Analyzed the data: DZW ZWJ. Wrote the paper: DZW ZWJ XG WH.

References

1. Nelson DR, Zeldin DC, Hoffman SM, Maltas IJ, Wain HM, et al. (2004) Comparison of cytochrome P450 (CYP) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes and alternative-splice variants. Pharmacogenomics 14: 1–10.
2. Koch I, Weil R, Wolbold R, Brockmoller J, Hustert E, et al. (2002) Interindividual variability and tissue-specificity in the expression of cytochrome P450 3A mRNA. Drug Metab Dispos 30: 1108–1114.
3. Gonzalez EJ, Nebert DW (1990) Evolution of the P450 gene superfamily: animal-plant ‘warfare’, molecular drive and human genetic differences in drug oxidation. Trends Genet 6: 182–186.
4. Martinez-Jimenez GP, Castell JV, Gomez-Lechon MJ, Jover R (2006) Transcriptional activation of CYP2C9, CYP1A1, and CYP1A2 by hepatocyte nuclear factor 4alpha requires coactivators peroxisomal proliferator activated receptor-gamma coactivator 1alpha and steroid receptor coactivator 1. Mol Pharmacol 70: 1681–1692.
5. Nebert DW, Vesell ES (2004) Advances in pharmacogenomics and individualized drug therapy: exciting challenges that lie ahead. Eur J Pharmacol 500: 267–285.
6. Rodriguez-Antona C, Donato MT, Pareja E, Gomez-Lechon MJ, Castell JV (2001) Cytochrome P-450 mRNA expression in human liver and its relationship with enzyme activity. Arch Biochem Biophys 393: 308–315.
7. Yang X, Zhang B, Molony C, Chudin E, Hao K, et al. (2010) Systematic genetic and genomic analysis of cytochrome P450 enzyme activities in human liver. Genome Res 20: 1020–1036.
8. Worhain M, Czerwinska M, He L, Parkinson A, Wan YJ (2007) Expression of constitutive androstane receptor, hepatic nuclear factor 4 alpha, and P450 oxidoreductase genes determines interindividual variability in basal expression and activity of a broad scope of xenobiotic metabolism genes in the human liver. Drug Metab Dispos 35: 1700–1710.
9. Li W, Harper PA, Tang BK, Okey AB (1998) Regulation of cytochrome P450 enzymes by aryl hydrocarbon receptor in human cells: CYP1A2 expression in the LS180 colon carcinoma cell line after treatment with 2,3,7,8-tetrachlorodibenzo-p-dioxin or 3-methylcholanthrene. Biochem Pharmacol 56: 599–612.

10. Neberth DW, Dalton TP, Okey AB, Gonzalez FJ (2004) Role of aryl hydrocarbon receptor-mediated induction of the CYP1 enzymes in environmental toxicity and cancer. J Biol Chem 279: 23437–23450.

11. Rebleck TR (2000) More about: modification of clinical presentation of prostate tumors by a novel genetic variant in CYP1A1. J Natl Cancer Inst 92: 76.

12. Pitarque M, von Richter O, Oke B, Berkkan H, Oscarson M, et al. (2001) Identification of a single nucleotide polymorphism in the TATA box of the CYP2A6 gene: impairment of its promoter activity. Biochem Biophys Res Commun 284: 435–440.

13. Kuehl P, Zhang J, Lin Y, Lamba J, Ausem M, et al. (2001) Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression. Nat Genet 27: 303–309.

14. Yan H, Yuan W, Velculescu VE, Vogelstein B, Kinzler KW (2002) Allelic variation in human gene expression. Science 297: 1143.

15. Martin J, Cleak J, Willis-Owen SA, Flint J, Shizuru JK, et al. (2008) Polymorphisms affecting gene transcription and mRNA processing in provocative candidate genes: detection through allelic expression imbalance. Mol Psychiatry 12: 421–422.

16. Cirulli ET, Goldstein DB (2007) In vitro assays fail to predict in vivo effects of regulatory polymorphisms. Hum Mol Genet 16: 1931–1939.

17. Johnson AD, Zhang Y, Papp AC, Pinsouneault JK, Lim JE, et al. (2008) Characterization of CYP2C19 and CYP2C9 from human liver: respective roles in microsomal tolbutamide, S-mephenytoin, and omeprazole hydroxylations. Arch Biochem Biophys 453: 16–26.

18. Lee CR, Peiper JA, Frye RF, Hinderliter AL, Blaisdell JA, et al. (2003) Tolbutamide, flurbiprofen, and losartan as probes of CYP2C9 activity in humans. J Clin Pharmacol 43: 84–91.

19. Kumar V, Wahlstrom JL, Rock DA, Warren GJ, Gorman LA, et al. (2006) CYP2C9 inhibition: impact of probe selection and pharmacogenetics on in vitro inhibition profiles. Drug Metab Dispos 34: 1966–1973.

20. Chen Y, Kisling G, Negishi M, Goldstein JA (2005) The nuclear receptors constitutive androstane receptor and pregnane X receptor cross-talk with hepatic nuclear factor 1α to synergistically activate the human CYP2C9 promoter. J Pharmacol Exp Ther 314: 1125–1133.

21. Parkinson A, Mutha DR, Johnson C, Dwyer A, Carroll KM (2004) The effects of gender, age, ethnicity, and liver cirrhosis on cytochrome P450 enzyme activity in human liver microsomes and inducibility in cultured human hepatocytes. Toxicol Appl Pharmacol 199: 193–209.

22. Frohlich V, Thissen JD, Rimington I, Oakey AB, Stockwell BR, et al. (2003) Characterization of a single nucleotide polymorphism in the putative HNF1-binding site of the human CYP1A1 and CYP1A2 genes. Mol Pharmacol 69: 1924–1930.

23. Lasker JM, Wester MR, Aramounhattee E, Raucy JL (1998) Characterization of CYP2C19 and CYP2C9 from human liver: respective roles in microsomal tolbutamide, S-mephenytoin, and omeprazole hydroxylations. Arch Biochem Biophys 353: 16–28.

24. Ueda R, Ikemuki H, Nagata K, Kimura S, Gonzalez FJ, et al. (2006) A common regulatory region function bidirectionally in transcriptional activation of the human CYP3A4 and CYP3A5 genes. Mol Pharmacol 69: 1924–1930.