

RESEARCH ARTICLE

Cytochrome P450-2D6: A novel biomarker in liver cancer health disparity

Zahraa I. Khamis^{1,2*}, Xiaodong Pang³, Zihan Cui⁴, Qing-Xiang Amy Sang^{1,5}, Jinfeng Zhang^{4*}

1 Department of Chemistry & Biochemistry, Florida State University, Tallahassee, Florida, United States of America, **2** Laboratory of Cancer Biology and Molecular Immunology, Department of Biochemistry, Faculty of Sciences-I, Lebanese University, Beirut, Lebanon, **3** Insilicom LLC, Tallahassee, Florida, United States of America, **4** Department of Statistics, Florida State University, Tallahassee, Florida, United States of America, **5** Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida, United States of America

* jinfeng@stat.fsu.edu (JZ); zik06@fsu.edu (ZIK)

**OPEN ACCESS**

Citation: Khamis ZI, Pang X, Cui Z, Sang Q-XA, Zhang J (2021) Cytochrome P450-2D6: A novel biomarker in liver cancer health disparity. PLOS ONE 16(10): e0257072. <https://doi.org/10.1371/journal.pone.0257072>

Editor: Narasimha Reddy Parine, King Saud University, SAUDI ARABIA

Received: November 2, 2020

Accepted: August 24, 2021

Published: October 1, 2021

Copyright: © 2021 Khamis et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting Information](#) files.

Funding: J.Z. was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number R01GM126558. ZIK was supported by a Lebanese University grant. QXAS was supported by CRC Multidisciplinary Support (MDS) grant from the Florida State University and an Endowed Chair Professorship in Cancer Research from anonymous donors. The funders had no role in

Abstract

Liver cancer morbidity and mortality rates differ among ethnic groups. In the United States, the burden of liver cancer in Asian Americans (AS) is higher compared to Caucasian Americans (CA). Research on liver cancer health disparities has mainly focused on environmental and socioeconomic factors yet has ignored the genotypic differences among various racial/ethnic groups. This lack of molecular level understanding has hindered the development of personalized medical approaches for liver cancer treatment. To understand the genetic heterogeneity of liver cancer between AS and CA, we performed a systematic analysis of RNA-seq data of AS and CA patients from The Cancer Genome Atlas (TCGA). We used four differential gene expression analysis packages; DESeq2, limma, edgeR, and Superdelta2, to identify the differentially expressed genes. Our analysis identified cytochrome P450-2D6 enzyme (CYP2D6) as the gene with the greatest differential expression with higher levels in AS compared to CA. To scrutinize the underlying mechanism of CYP2D6, Ingenuity Pathway Analysis (IPA) and Cytoscape were conducted and found hepatocyte nuclear factor-4 α (HNF4A) and interleukin-6 (IL6) in direct association with CYP2D6. IL6 is downregulated in AS compared to CA, while HNF4A is not significantly different. Herein, we report that CYP2D6 may serve as a putative biomarker in liver cancer health disparities. Its negative association with IL6 proclaims an intricate relationship between CYP2D6 and inflammation in the ethnic differences seen in AS and CA liver cancer patients. The goal of the present study was to understand how genetic factors may contribute to the interethnic variability of liver cancer prevalence and outcomes in AS and CA patients. Identifying ethnic-specific genes may help ameliorate detection, diagnosis, surveillance, and treatments of liver cancer, as well as reduce disease-related incidence and mortality rates in the vulnerable population.

study design, data collection and analysis, decision to publish, or preparation of the manuscript and only provided support in the form of research materials.

Competing interests: The authors have read the journal's policy and have the following competing interests: XP is a paid employee of Insilicom (<https://insilicom.com/>). There are no patents, products in development or marketed products associated with this research to declare. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

Introduction

Liver cancer is a major global health problem and a leading cause of death worldwide. The World Health Organization estimates that more than 1 million deaths will occur from liver cancer in 2030 [1]. In 2021, it has been estimated that 42,230 people will be diagnosed with liver cancer and 30,230 will die from the disease in the United States [2]. Liver cancer incidence continues to rise in 2020 at a faster pace compared to other cancers. The major etiologic factors for this cancer include metabolic disorders like diabetes and obesity, chronic infection with hepatitis B or C viruses, smoking, and excessive alcohol consumption [3]. Despite these seemingly preventable risk factors, the 5-year survival rate of liver cancer is only 18% in 2020, ranking it the second most fatal tumor after pancreatic cancer [2].

Disparities in liver cancer incidence and mortality rates are reported among populations with different races and ethnicities. Latinos have suffered from the highest increase in liver cancer incidence in the past years. Compared to Caucasian patients with different etiologies, Latinos presented with more advanced and severe features of the disease [4]. Other factors contributing to disparities include age, gender, socioeconomic status, access to health care, and geographic location. A mounting evidence suggests that genetic and epigenetic factors can predict an individual's susceptibility to developing liver cancer. In a Caucasian population, five genetic biomarkers have been identified as significant determinants of liver cancer risk [5]. Liver cancer disproportionately affects Asians and Caucasians in the United States with higher incidence among Asian groups [6]. This disparity has never been explored at the genetic level between the aforementioned groups. So, we opt to identify the genes differentially expressed in AS and CA liver cancer patients to improve disease screening, early cancer detection, and therapies tailored to individual's needs.

To determine differentially expressed genes (DEGs), we have performed an integrative computational analysis of gene expression data from The Cancer Genome Atlas (TCGA) of AS and CA liver cancer patients. The use of four different R/Bioconductor packages: DESeq2 [7], limma [8], edgeR [9], and Superdelta2 [10] maximized the statistical power of our study. We have found 19 common DEGs among the four packages with CYP2D6 on the top of the list. CYP2D6 exhibited increased expression in AS compared to CA. To gain more insights about the underlying molecular pathways associated with CYP2D6, IPA and Cytoscape plug-in, Cytoscape, were used to generate integrated pathways, networks, and upstream regulators, and to identify important hub genes, respectively [11]. Using these analyses, we have identified key genes that may decrease ethnic-driven health disparities and may serve as valuable prognostic and therapeutic targets that can stop the development of liver cancer (Fig 1).

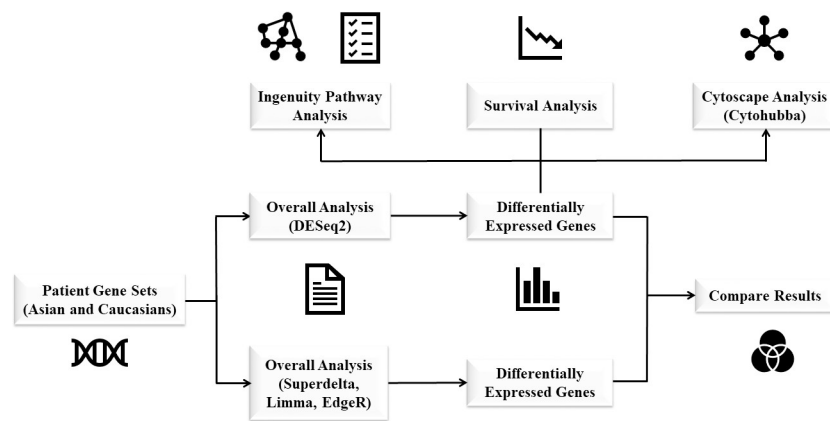


Fig 1. Work flowchart used for data analysis.

<https://doi.org/10.1371/journal.pone.0257072.g001>

Materials and methods

RNA-seq data from TCGA

The RNA-seq data of all liver hepatocellular carcinoma (LIHC) tumor samples of AS (159) and CA (185) were downloaded from The Cancer Genome Atlas (TCGA) (<http://cancergenome.nih.gov/>). Normalized gene expression counts were used, unless otherwise noted. The complete list of patient IDs is found in [S1 Table](#).

Differential gene expression analysis

To calculate the DEGs between AS and CA, with CA as the reference group, four different methods were used: DESeq2 (version 1.30) [7], limma (version 3.46.0) [8], edgeR (version 3.32.0) [9], and Superdelta2 (version 2.0) [10]. For each method, genes with the \log_2 fold-change greater than 1 and adjusted p-value smaller than 0.05 were selected as DEGs. Genes with a sum of 10 reads or fewer across all samples were removed before analysis. The Benjamini-Hochberg method was used for p-value adjustment [12]. For DESeq2, parametric fit type was used. Outliers replacement, outliers filtering, and independent filtering were on. For limma, RNA-seq data were first transformed with the voom function. The robust settings were applied while the test significance relative to a fold-change threshold was not. For edgeR, the robust settings were applied. For Superdelta2, offset of 0.5 was added to avoid infinite value of \log_2 of 0. The trimming portion was set to 0.2 when estimating the between-group test. The heatmap was generated using R package pheatmap (version 1.0.12).

Pathway and network analysis

The DEGs from DESeq2 were submitted to the IPA software to determine the relative enrichment of genes in different pathways using right-tailed Fisher's exact test. Based on the differential expression data, IPA analysis generated canonical pathways, upstream regulators, and top tox list. Hub genes were identified using the CytoHubba plug-in in the Cytoscape software.

Survival analysis

Survival analysis was performed using the survival (version 3.2–7) and survminer (version 0.4.8) packages in R.

Results

Genotypic profiling identifies 19 common differentially expressed genes

To uncover the genes that may contribute to disparities in liver cancer incidences between AS and CA, we used several analysis methods—DESeq2 [7], limma [8], edgeR [9], and Superdelta2 [10]—to identify the differentially expressed genes. An overlapping set of genes representing the strongest contributors to the disparities in liver cancer incidences was determined. For each method, we selected the significant genes with an absolute \log_2 fold-change >1 (\log_2 FC > 1) and adjusted p-value <0.05 as differentially expressed. The DESeq2 package revealed a set of 559 differentially expressed genes between AS and CA patients. Analysis from limma, edgeR, and Superdelta2 pipelines revealed 99, 74, and 182 differentially expressed genes, respectively ([S2 Table](#)). Only 19 differentially expressed genes are common among the four methods ([Fig 2](#)). [Table 1](#) lists the details of these genes. A negative \log_2 FC reveals a downregulation in AS compared to CA and a positive \log_2 FC shows an upregulation in AS compared to CA. Out of the 19 common DEGs, 6 genes are upregulated and 13 genes are downregulated.

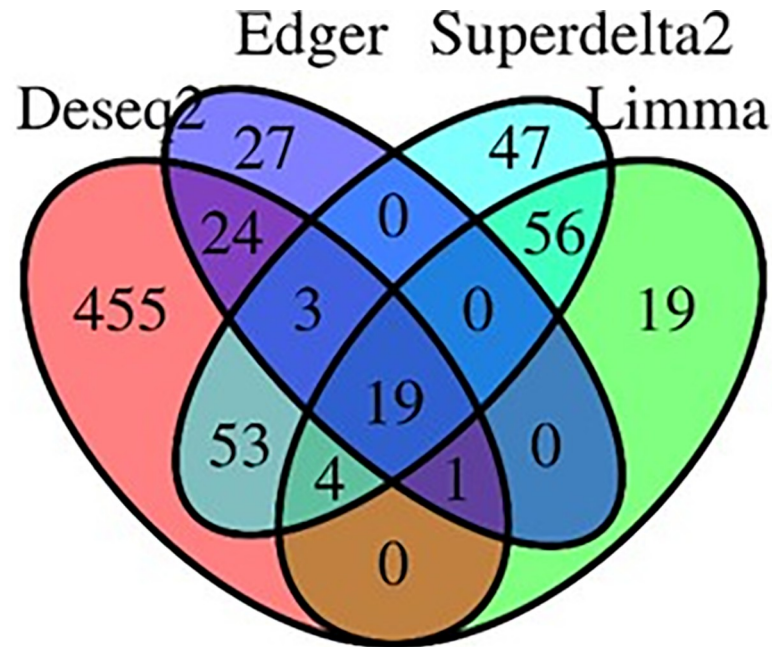


Fig 2. Venn diagram of DEGs. Four different methods (DESeq2, edgeR, limma, and Superdelta2) were utilized generating 19 common differentially expressed genes.

<https://doi.org/10.1371/journal.pone.0257072.g002>

Table 1. The common genes with differential expression.

Gene ID	log ₂ FC	p-value	padj
CYP2D6	1.82	8.67e-27	5.66e-23
SLC39A4	1.68	5.95e-14	4.22e-11
COL9A3	1.24	6.17e-9	7.49e-7
TSPAN10	1.2	7.58e-18	1.48e-14
DHRS2	1.18	0.0000429	0.000748
ACSM1	1.11	0.00000647	0.000176
GREM2	-1.01	0.000259	0.00284
ESR1	-1.01	0.00000566	0.000159
GSTA2	-1.03	0.0000469	0.000798
TAT	-1.03	0.0000219	0.00045
ADH1B	-1.03	5.77e-7	0.0000302
PPM1K	-1.04	1.4e-14	1.25e-11
SRD5A2	-1.1	0.00000495	0.000144
CFHR4	-1.11	5.09e-7	0.0000272
CPED1	-1.14	9.39e-12	3.67e-9
MFAP3L	-1.17	2.5e-10	5.32e-8
CLRN3	-1.28	0.00000105	0.000046
GNAO1	-1.38	2.16e-10	4.75e-8
SPATA18	-1.48	4.66e-13	2.34e-10

The list of the 19 DEGs that are common among the four analysis methods used. The genes are listed with their respective fold change (log₂ FC), the p-value, and the adjusted p-value (padj).

<https://doi.org/10.1371/journal.pone.0257072.t001>

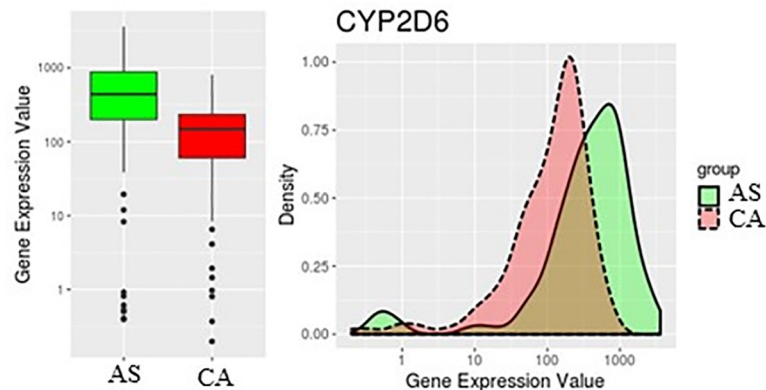


Fig 3. Expression levels of CYP2D6 in AS and CA groups. CYP2D6 is higher in AS than CA.

<https://doi.org/10.1371/journal.pone.0257072.g003>

CYP2D6 as a potential biomarker in Asian Americans with liver cancer

Table 1 shows CYP2D6 as the highest significantly differentially expressed gene in our data set. The CYP2D6 (Cytochrome P450 Family 2 Subfamily D Member 6) gene encodes a member of the cytochrome P450 superfamily of enzymes and is involved in the metabolism of drugs and other endogenous and exogenous molecules [13]. CYP2D6 (\log_2 FC = 1.82) is overexpressed in AS compared to CA (Fig 3). This is in agreement with our previous report about breast cancer health disparities where AS breast cancer patients experience higher levels of CYP2D6 compared to their CA counterparts [14].

Ingenuity pathway analysis identifies HNF4A as an upstream regulator of CYP2D6

To decipher the role of CYP2D6 in the disparity of liver cancer incidence and outcome between AS and CA, we imported the DESeq2 differentially expressed dataset of 559 genes into the Ingenuity Pathway Analysis (IPA) tool. IPA uses the Ingenuity Knowledge Database to map DEGs into networks, canonical pathways, and a top tox list to determine functional relationships. The IPA tox list analyses revealed a significant enrichment of cytochromes P450 enzymes, xenobiotic metabolism, and fatty acid metabolism. Table 2 displays the most significantly altered pathways in our dataset based on the tox list tool analysis. Heatmap of DEGs in the top five tox pathways shows our gene of interest, CYP2D6, to be highly associated with P1, P3, and P4 pathways which directly correlates with CYP2D6 function in fatty acid metabolism and xenobiotics (Fig 4). Additional IPA network analysis of CYP2D6 shows its association with HNF4A (Fig 5A). HNF4A is a highly expressed receptor in the liver and is known to

Table 2. The top toxicological pathways.

Name	Pathway	P-value
P1	Cytochrome P450 panel- substrate is a xenobiotic (human)	9.05 E-04
P2	Increases cardiac proliferation	1.36 E-03
P3	Fatty acid metabolism	3.25 E-03
P4	Cytochrome P450 panel- substrate is a xenobiotic (Rat)	3.72 E-03
P5	Cardiac hypertrophy	1.01 E-03

The pathways are predicted by Ingenuity Pathway Analysis tool based on the differentially expressed genes taken from DESeq2 pipeline.

<https://doi.org/10.1371/journal.pone.0257072.t002>

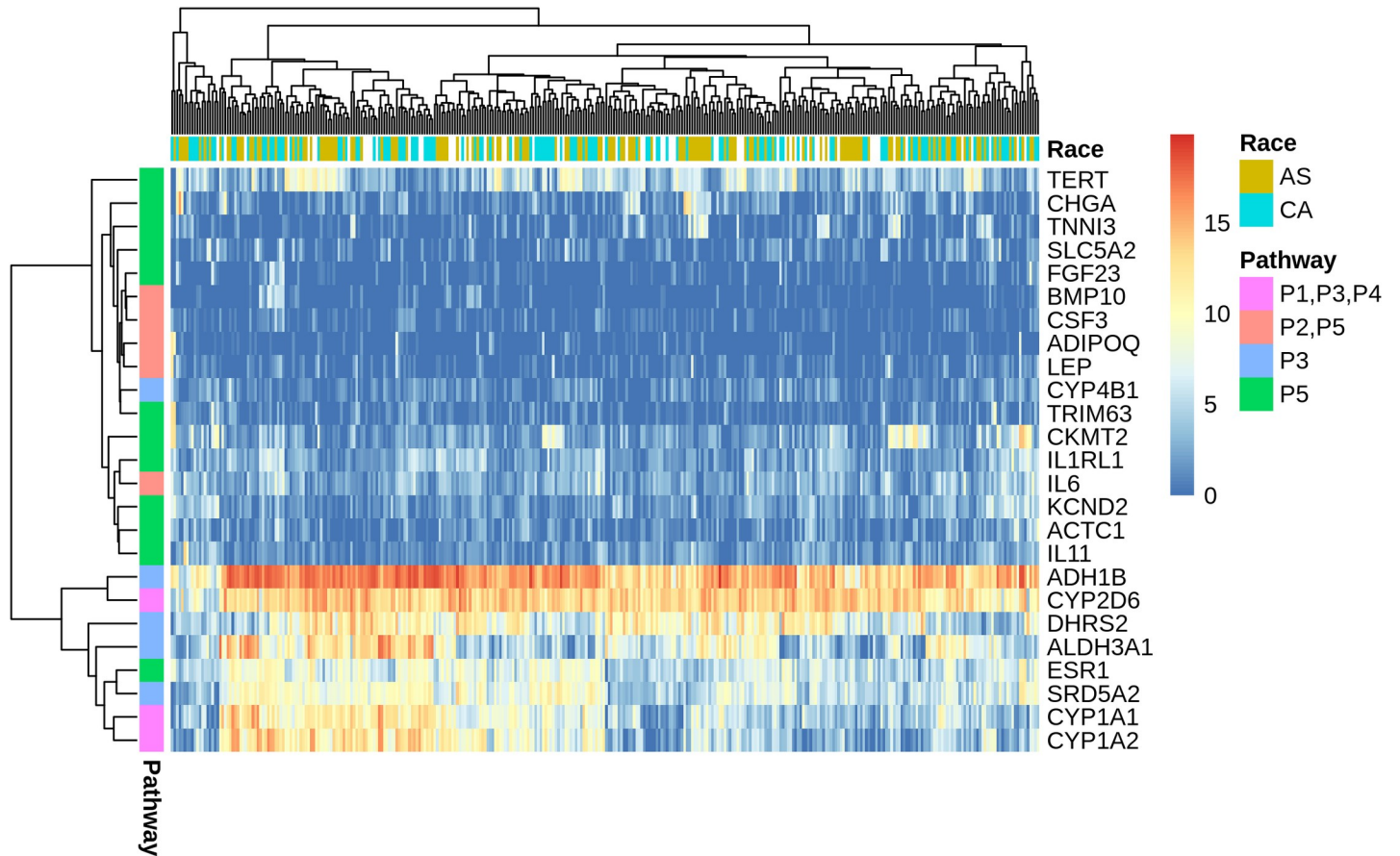


Fig 4. Heatmap of \log_2 FC value of differentially expressed genes in the top 5 tox pathways. CYP2D6 has high expression and close association with P1, P3, and P4 pathways pertinent to its normal biological functions in xenobiotics and fatty acid metabolism.

<https://doi.org/10.1371/journal.pone.0257072.g004>

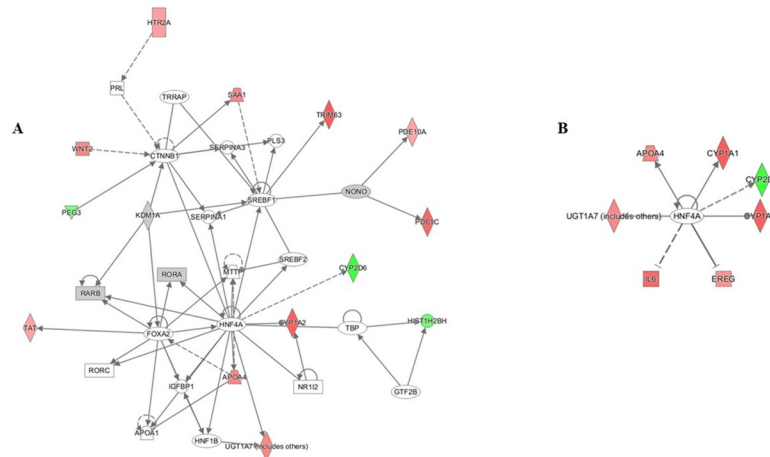


Fig 5. IPA analysis of differentially expressed genes. (A) Network analysis reveals a direct interaction between CYP2D6 and HNF4A. (B) Upstream network analysis of HNF4A shows CYP2D6 and IL6 as major players.

<https://doi.org/10.1371/journal.pone.0257072.g005>

Table 3. Upstream network analysis of HNF4A.

Symbol	Entrez Gene Name	Log Ratio	p-value
APOA4	apolipoprotein A4	-1.37	0.0000648
CYP1A1	cytochrome P450 family 1 subfamily A member 1	-1.83	3.26E-08
CYP1A2	cytochrome P450 family 1 subfamily A member 2	-1.8	0.00000922
CYP2D6	cytochrome P450 family 1 subfamily D member 6	1.82	8.67E-27
EREG	Epiregulin	-1.23	0.000397
IL6	interleukin 6	-1.72	6.05E-14
UGT1A7	UDP glucuronosyltransferase family 1 member A	-1.35	0.00406

The molecules correlated with HNF4A when network analysis upstream of HNF4A is performed using IPA. Expression levels and p-values are shown.

<https://doi.org/10.1371/journal.pone.0257072.t003>

regulate liver development and metabolic functions [15, 16]. To gain more biological insight into the role of HNF4A, an upstream regulator analysis was performed. IPA depends on experimentally-observed biological findings available in the literature to predict upstream regulatory molecules based on input genes [17]. Interestingly, our analysis identified HNF4A as an upstream regulator of the DEGs in our dataset. Therefore, this gene may contribute to the observed gene expression changes seen in AS and CA liver cancer patients in this study.

Cytohubba analysis identifies IL6 as a hub gene among DEGs in liver cancer

In search for the underlying mechanism that may explain the disparity between AS and CA liver cancer patients, we have performed an IPA network analysis on the upstream regulator, HNF4A. The interacting molecules include enzymes (CYP1A1, CYP1A2, CYP2D6, UGT1A7), a growth factor (EREG), a cytokine (IL6), and a transporter (APOA4) (Fig 5B) (Table 3). Based on previous reports, IL6 cytokine expression is inversely correlated with the activity of cytochrome P450 drug metabolizing enzymes [18], the category in which the most significant DEG, CYP2D6, falls. Thus, IL6 is further pursued. Cytoscape analysis of our DESeq2 differentially expressed genes using cytohubba plugin has revealed IL6 as the central node with the highest degree of connectivity (Fig 6). Therefore, the highly connected IL6 node is an important biological hub gene in our dataset (Table 4). IL6 is downregulated in AS compared to CA liver cancer patients (\log_2 FC = -1.72). Shi et al. has reported a similar decrease in IL6 expression in AS breast cancer patients compared to their CA counterparts supporting our results [19]. Hence, IL6 and CYP2D6 expression levels are inversely proportional in accordance with the known effects of IL6 on cytochrome P450 enzymes [18].

Survival analysis

Survival analysis indicates significant differences in survival between AS and CA (p-value = 0.00014) (Fig 7). Although they have a greater liver cancer incidence rate than CA, AS patients display a much higher 5-year survival rate of ~80% compared to CA, which is ~45%.

Discussion

Liver cancer is a leading cause of cancer-related deaths worldwide. Like all types of cancer, liver cancer incidence, prevalence, and mortality rates vary substantially between different populations [20]. These health disparities are caused by a combination of tumor biology, environmental, social, and economic factors. The contribution of the genotypic differences among various populations is underexplored [21, 22]. In a retrospective study of racial and ethnic

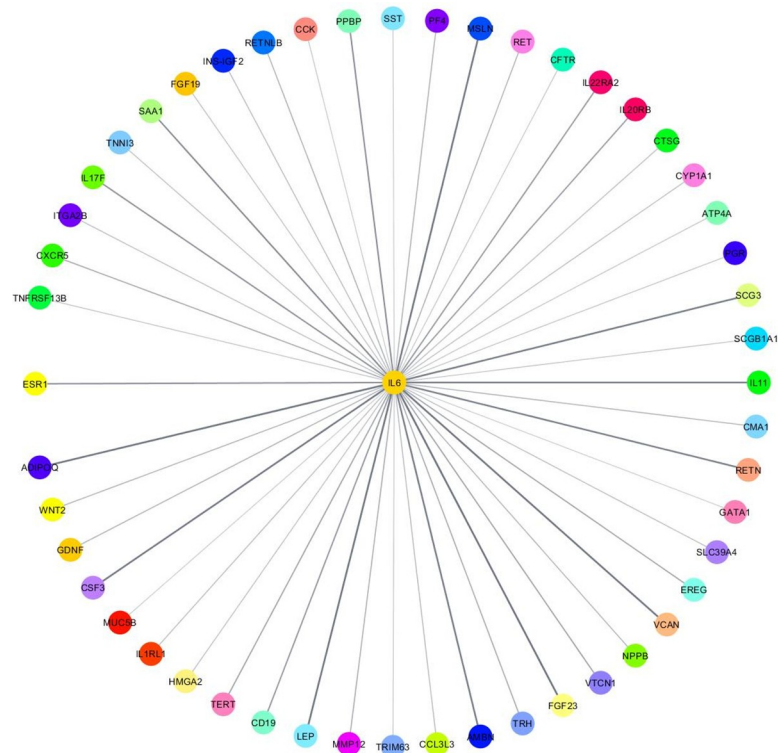


Fig 6. IL6 as a hub gene. IL6 has the highest degree of connectivity.

<https://doi.org/10.1371/journal.pone.0257072.g006>

minorities diagnosed with liver cancer, Rich et al. have reported that variations in liver cancer outcomes are associated with delays in diagnosis and receipt of curative treatment [23]. Additionally, evaluation of liver cancer incidence by Yang et al. in California based on race/ethnicity, socioeconomic status, and geographic area has found that Asians/Pacific Islanders (APIs) and Hispanics have the highest incidence rate [24]. Among US populations, Asians have a higher liver cancer burden than other racial groups. However, in our study, we have found that AS have a much higher survival rate than CA (Fig 7). The etiology of this disparity has

Table 4. Top 10 essential genes in our DEGs dataset.

Rank	Name	Score
1	IL6	50
2	SST	33
3	CCK	28
4	SAA1	27
5	PVALB	25
6	GRIN1	25
7	GPR17	24
8	LEP	23
9	ADCY1	22
10	LPAR3	22

Cytoscape analysis using Cytohubba-plugin reveals IL6 as the highest interconnected gene based on the degree method.

<https://doi.org/10.1371/journal.pone.0257072.t004>

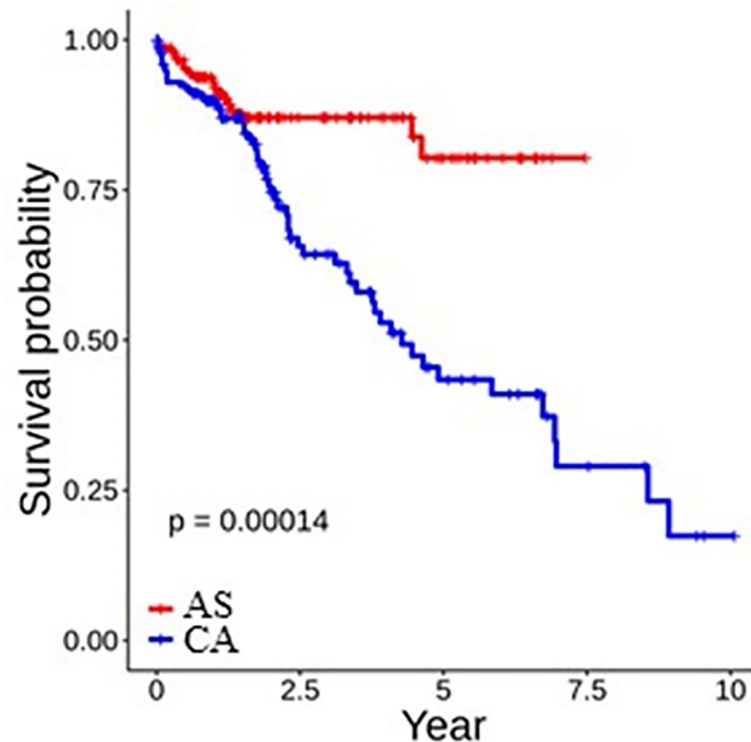


Fig 7. Survival curve of AS and CA groups. AS have higher survival rate than CA.

<https://doi.org/10.1371/journal.pone.0257072.g007>

never been investigated at the genetic level. Therefore, we have examined the genotypic differences between AS and CA patients to determine the genetic basis of the racial/ethnic specific patterns of liver cancer incidence. To our knowledge, this is the first study to identify differentially expressed genes in AS and CA that may potentially drive the observed health disparity.

In this study, we analyzed differentially expressed genes (DEGs) in AS versus CA liver cancer tumor samples. To make the results more robust, we used four differential expression analysis packages, DESeq2, limma, edgeR, and Superdelta2, and identified 19 common DEGs (Fig 2 and Table 1). The major drug-metabolizing enzyme, CYP2D6, displayed the highest significant change (\log_2 FC = 1.82) (Fig 3). This gene is expressed in various tissues, including those of the liver, kidney, intestine, and breast [25]. Although it constitutes ~1.5% of total hepatic CYP enzymes, CYP2D6 metabolizes ~30% of commonly used drugs [26]. Located on chromosome 22q13, CYP2D6 gene is highly polymorphic, with more than 90 described allelic variants. These variants—mainly due to single nucleotide polymorphisms (SNPs), gene deletions, or gene duplications—usually result in abolished or altered activity of the CYP2D6 enzyme [13]. Depending on the function of each allele in CYP2D6 gene, the enzyme can have four different phenotypes; poor metabolizer (PM), intermediate metabolizer (IM), extensive metabolizer (EM), and ultra-rapid metabolizer (UM) [27]. The distribution of the allelic variants shows a substantial interindividual differences. Caucasians and African Americans populations have higher PM frequency than the Chinese population highlighting the influence of CYP2D6 genotypic variation on its phenotype and thus, its expression and function [28]. The genetic polymorphism of CYP2D6 in various populations, together with its role in the metabolism of a large percentage of xenobiotics and administered drugs should be taken into consideration in clinical settings while assessing the therapeutic and toxic dose response of medications prescribed for patients.

The relationship between CYP2D6 allelic variants and cancer incidence is still debatable due to conflicting reports [29, 30]. CYP2D6 has been extensively studied in breast cancer due to its ability to convert tamoxifen to the more active metabolite, endoxifen [31]. Patients with poor metabolizer phenotype for CYP2D6 show lower concentration of endoxifen [27]. With respect to the liver, few reports have claimed a significant association between CYP2D6 polymorphism and the development of liver cancer. One study has showed that individuals who are homozygous for functional CYP2D6 alleles are at an increased risk of developing the disease [32, 33]. Also, individuals with hepatitis C virus (HCV) are reported to have lower activity of CYP2D6 in the liver [34]. However, assessment of CYP2D6 in liver cancer patients with different racial/ethnic backgrounds has not been reported.

Given CYP2D6 significant role in drug clearance, evaluation of CYP2D6 in liver cancer patients of diverse backgrounds must be prioritized to ensure the proper application of pharmaceuticals in precision medicine. In our analysis, CYP2D6 shows the greatest significant differential expression in the more vulnerable population. It has increased expression levels in AS than CA liver cancer patients (Fig 3), which is in accordance with our previous findings in breast cancer [14]. Therefore, CYP2D6 may serve as a putative biomarker in liver cancer health disparities. To understand the role of this gene in our dataset, we performed an IPA functional analysis. The top tox list tool of IPA identified the most significant pathways and gene lists in our input data to be cytochrome P450 enzymes, xenobiotic metabolism, and fatty acid metabolism (Table 2). Heatmap generation of DEGs in these pathways revealed high expression and clustering of CYP2D6 to these pathways, consistent with its biological roles in liver metabolism (Fig 4).

Additional IPA network analysis reveals that HNF4A acts as a regulator of CYP2D6 expression (Fig 5A). HNF4A is a member of the nuclear receptor superfamily that binds DNA elements and regulates gene transcription. Mainly expressed in the liver, it is essential for the proper maintenance of hepatic development and metabolic functions. Mice that are unable to express HNF4A have altered lipid metabolism and increased mortality, which underscores the key regulatory role of HNF4A in liver physiology [16, 35, 36]. Studies indicate that HNF4A is indeed a regulator of CYP2D6 [37] with HNF4A expression and CYP2D6 enzyme activity significantly correlated [38, 39]. Although the expression of HNF4A was not significantly different between AS and CA liver cancer patients in our study, our IPA analysis did identify HNF4A as an upstream regulator of CYP2D6.

In the HNF4A upstream regulator network, a set of DEGs are in direct association with HNF4A (Fig 5B); most importantly, interleukin-6 (IL6), a known effector of cytochrome P450 enzymes. IL6 is a pleiotropic cytokine that acts through its receptor (IL6R) to elicit a wide variety of effects ranging from acute phase response (APR) to proliferation, differentiation, and other biological processes [40, 41]. It is negatively associated with cytochrome P450 (CYP450) enzyme activity and HNF4A expression [18, 42]. In vitro exposure of human hepatocytes in cell culture to IL6 caused suppression of CYP450 enzymes [18]. The IL6-mediated suppression of liver CYP450 enzymes has been an issue in the drug industry because most prescribed drugs are CYP-metabolized. So, efforts have been made to block this cytokine using an anti-IL6 receptor antibody to reduce the bioavailability of CYP-metabolized medications [18, 43]. With respect to CYP2D6, primary cultures of human hepatocytes treated with IL6 caused a time- and dose-dependent decrease in CYP2D6 expression [44]. Other proinflammatory cytokines like tumor necrosis factor α (TNF α) and interleukin 1 β (IL1 β) increased as the nonalcoholic fatty liver disease progressed. This increase in TNF α and IL1 β levels was accompanied with a decrease in CYP enzymes, among which was CYP2D6. The authors implied that the elevated expression in cytokines may have caused the observed decrease in CYP enzymes activity [45]. The drug metabolizing CYP450 enzymes are known to induce or inhibit inflammation based on their ability to metabolize a variety of molecules. The role of CYP450 enzymes in

inflammation is reviewed extensively by Christmas P. [46]. Our finding, which shows an inverse expression of IL6 and CYP2D6, is consistent with the literature. IL6, with a \log_2 FC = -1.72, is downregulated in AS compared to CA. The importance of IL6 in our liver cancer health disparity dataset was further validated using Cytohubba plug-in, which identified IL6 as a hub gene with the highest degree of connectivity (Fig 6). Collectively, these results may indicate a relationship between inflammation and CYP enzymes; in particular, CYP2D6.

Based on their role in the metabolism of therapeutic drugs, xenobiotics, and endogenous substrates, CYP450 enzymes play major roles in inflammation and cancer. Generally, inflammation decreases the expression and activity of CYP450 enzymes due to the release of proinflammatory cytokines from immune cells. In the context of tumor microenvironment, the modulation of CYP450 levels in the presence of proinflammatory cytokines can induce tumorigenesis and affect response to chemotherapy [47]. Thus, it is indispensable to elaborate the inflammation-CYP450-cancer axis to determine drug-drug interactions and proper drug dosages to avoid adverse effects and therapeutic failure. Our results position CYP2D6 at the intersection between inflammation and liver cancer in AS and CA patients. The high levels of CYP2D6 in AS and IL6 in CA suggest variability in liver cancer etiology. Elevated IL6 expression in CA patients may imply an initial hepatic inflammation due to HBV/HCV that may have progressed to liver cancer; while the increased levels of CYP2D6 in AS may be due to another underlying problem like alcohol or metabolic syndrome. Therefore, it is of paramount importance to delineate the predominant cause of liver cancer in distinct populations, and to investigate how the underlying cause can affect the expression and activity of CYP2D6, and the release of proinflammatory cytokines, mainly IL6.

Our findings suggest an intricate interplay between the genes CYP2D6, HNF4A, and IL6 in AS and CA liver cancer patients. Research in the past decade has shown that these three genes are highly polymorphic, leading to variability in the function of the respective gene products. In a Korean population, the genetic polymorphism G60D in HNF4A resulted in low activity of CYP2D6 [48]. Another study in *Schistosoma haematobium*-infected patients from Africa found that certain polymorphic variants of CYP2D6 and IL6 were highly associated with morbidity and severity of the disease [49]. Thus, we contemplated that the relationship between CYP2D6, HNF4A, and IL6 was governed by single nucleotide polymorphisms (SNPs), specifically in CYP2D6. The availability of different allelic variants in each locus may dictate how each gene will interact with the others and regulate the observed disparity in AS and CA liver cancer patients. Surprisingly, SNP analysis revealed only one variant (ID: rs72657357) associated with IL6 expression with a frequency of 0.0239 in CA and 0.0179 AS (S3 Table). Neither CYP2D6 nor HNF4A had SNPs related to the health disparity between the two ethnic groups obviating the implication of CYP2D6 polymorphism in the inequalities of liver cancer in AS and CA. Therefore, it is the gene expression level of CYP2D6 that contributes to the liver cancer health disparity.

Liver cancer prevalence in AS is higher than any racial/ethnic population. However, survival analysis of our dataset showed a significant difference in survival between AS and CA patients. AS had a prolonged 5-year survival rate of 80%, while CA had a 45% rate (Fig 7). This result suggests that increased CYP2D6 expression levels are associated with better liver cancer survival rate; however, multiple biological, environmental, social, economic, cultural, and other confounding variables associated with liver cancer survival remain to be investigated.

Conclusions

The dismal prognosis of liver cancer necessitates the development of new therapeutics that can target all cancer patients regardless of their differences. With prevalence of liver cancer health

disparities, efforts are tailored to individualize disease origin, diagnosis, and treatment for each patient to optimize better responses to therapy. To eliminate health disparities, we need to understand how society, environment, lifestyle, and ancestral factors affect liver cancer incidence and outcome. This is initiated by a comprehensive profiling of the genome. Herein, we explored how genetic differences in AS and CA patients might influence liver cancer incidence and survival. Being the highest significantly DEG among the two populations, CYP2D6 may serve as a putative biomarker for liver cancer disparities. The higher expression of CYP2D6 in AS can be attributed to higher disease risk, better survival outcome, or different etiology profiles. Deciphering how CYP2D6 affects incidence or survival rates warrants further investigation. It is also imperative to dissect the inherent causes of liver cancer in AS and CA, and their effect on CYP2D6 activity and inflammatory responses. Once elucidated, we will have a better understanding of the way CYP2D6 modulates liver cancer and inflammation; thus, leading to better assessment of drugs interactions and conduct of efficient therapies with minimal side effects. Although these preliminary results require further validation, CYP2D6 and IL6 might have significant clinical impact in the management of liver cancer incidence and application of personalized treatment regimens.

Supporting information

S1 Table. Patient IDs and clinical information.

(XLSX)

S2 Table. Differentially expressed genes in AS and CA liver cancer patients. Results from all methods (DESeq2, edgeR, limma, and Superdelta2) are displayed.

(XLSX)

S3 Table. Single nucleotide polymorphism analysis. SNPs identified in AS and CA liver cancer patients.

(XLSX)

Acknowledgments

We would like to thank Mr. Justine Dees for the final edits.

Author Contributions

Conceptualization: Zahraa I. Khamis, Jinfeng Zhang.

Data curation: Xiaodong Pang, Zihan Cui.

Formal analysis: Zahraa I. Khamis, Jinfeng Zhang.

Funding acquisition: Zahraa I. Khamis, Qing-Xiang Amy Sang, Jinfeng Zhang.

Investigation: Zahraa I. Khamis, Xiaodong Pang.

Methodology: Zahraa I. Khamis, Xiaodong Pang, Jinfeng Zhang.

Project administration: Zahraa I. Khamis.

Resources: Jinfeng Zhang.

Software: Jinfeng Zhang.

Supervision: Jinfeng Zhang.

Writing – original draft: Zahraa I. Khamis.

Writing – review & editing: Zahraa I. Khamis, Qing-Xiang Amy Sang, Jinfeng Zhang.

References

1. WHO. Projections of mortality and causes of death, 2016 to 2060. https://www.who.int/healthinfo/global_burden_disease/projections/en/.
2. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2021. *CA Cancer J Clin*. 2021; 71: 7–33. <https://doi.org/10.3322/caac.21654> PMID: 33433946
3. Makarova-Rusher OV, Altekruse SF, McNeel TS, Ulahannan S, Duffy AG, Graubard BI, et al. Population attributable fractions of risk factors for hepatocellular carcinoma in the United States. *Cancer*. 2016; 122: 1757–1765. <https://doi.org/10.1002/ncr.29971> PMID: 26998818
4. Kuffinec GN, Levy R, Kieffer DA, Medici V. Hepatocellular carcinoma and associated clinical features in Latino and Caucasian patients from a single center. *Ann Hepatol*. 2019; 18: 177–186. <https://doi.org/10.5604/01.3001.0012.7910> PMID: 31113588
5. De Mattia E, Cecchin E, Polesel J, Bignucolo A, Roncato R, Lupo F, et al. Genetic biomarkers for hepatocellular cancer risk in a caucasian population. *World J Gastroenterol*. 2017; 23: 6674–6684. <https://doi.org/10.3748/wjg.v23.i36.6674> PMID: 29085212
6. Islami F, Miller KD, Siegel RL, Fedewa SA, Ward EM, Jemal A. Disparities in liver cancer occurrence in the United States by race/ethnicity and state. *CA Cancer J Clin*. 2017; 67: 273–289. <https://doi.org/10.3322/caac.21402> PMID: 28586094
7. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014; 15: 550. <https://doi.org/10.1186/s13059-014-0550-8> PMID: 25516281
8. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015; 43: e47. <https://doi.org/10.1093/nar/gkv007> PMID: 25605792
9. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010; 26: 139–140. <https://doi.org/10.1093/bioinformatics/btp616> PMID: 19910308
10. Liu Y, Zhang J, Qiu X. Super-delta: a new differential gene expression analysis procedure with robust data normalization. *BMC Bioinformatics*. 2017; 18: 582. <https://doi.org/10.1186/s12859-017-1992-2> PMID: 29268715
11. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol*. 2014; 8: S11. <https://doi.org/10.1186/1752-0509-8-S4-S11> PMID: 25521941
12. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B*. 1995; 57: 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
13. Agundez JAG. Cytochrome P450 gene polymorphism and cancer. *Curr Drug Metab*. 2004; 5: 211–24. <https://doi.org/10.2174/1389200043335621> PMID: 15180491
14. Li Y, Steppi A, Zhou Y, Mao F, Miller PC, He MM, et al. Tumoral expression of drug and xenobiotic metabolizing enzymes in breast cancer patients of different ethnicities with implications to personalized medicine. *Sci Rep*. 2017; 7: 4747. <https://doi.org/10.1038/s41598-017-04250-2> PMID: 28684774
15. Sladek FM, Zhong WM, Lai E, Darnell JE. Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. *Genes Dev*. 1990; 4: 2353–2365. <https://doi.org/10.1101/gad.4.12b.2353> PMID: 2279702
16. Hayhurst GP, Lee YH, Lambert G, Ward JM, Gonzalez FJ. Hepatocyte nuclear factor 4alpha (nuclear receptor 2A1) is essential for maintenance of hepatic gene expression and lipid homeostasis. *Mol Cell Biol*. 2001; 21: 1393–403. <https://doi.org/10.1128/MCB.21.4.1393-1403.2001> PMID: 11158324
17. Krämer A, Green J, Pollard J, Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics*. 2014; 30: 523–530. <https://doi.org/10.1093/bioinformatics/btt703> PMID: 24336805
18. Dickmann LJ, Patel SK, Rock DA, Wienkers LC, Slatter JG. Effects of Interleukin-6 (IL-6) and an Anti-IL-6 Monoclonal Antibody on Drug-Metabolizing Enzymes in Human Hepatocyte Culture. *Drug Metab Dispos*. 2011; 39: 1415–1422. <https://doi.org/10.1124/dmd.111.038679> PMID: 21555507
19. Shi Y, Steppi A, Cao Y, Wang J, He MM, Li L, et al. Integrative Comparison of mRNA Expression Patterns in Breast Cancers from Caucasian and Asian Americans with Implications for Precision Medicine. *Cancer Res* 2017; 77:423–433. <https://doi.org/10.1158/0008-5472.CAN-16-1959> PMID: 28069798

20. Wallace TA, Martin DN, Ambs S. Interactions among genes, tumor biology and the environment in cancer health disparities: examining the evidence on a national and global scale. *Carcinogenesis*. 2011; 32: 1107–1121. <https://doi.org/10.1093/carcin/bgr066> PMID: 21464040
21. Khemlina G, Ikeda S, Kurzrock R. The biology of Hepatocellular carcinoma: implications for genomic and immune therapies. *Mol Cancer*. 2017; 16: 149. <https://doi.org/10.1186/s12943-017-0712-x> PMID: 28854942
22. Ha J, Yan M, Aguilar M, Tana M, Liu B, Frenette CT, et al. Race/Ethnicity-specific Disparities in Hepatocellular Carcinoma Stage at Diagnosis and its Impact on Receipt of Curative Therapies. *J Clin Gastroenterol*. 2016; 50:423–301. <https://doi.org/10.1097/MCG.0000000000000448> PMID: 26583267
23. Rich NE, Hester C, Odewole M, Murphy CC, Parikh ND, Marrero JA, et al. Racial and Ethnic Differences in Presentation and Outcomes of Hepatocellular Carcinoma. *Clin Gastroenterol Hepatol*. 2019; 17: 551–559.e1. <https://doi.org/10.1016/j.cgh.2018.05.039> PMID: 29859983
24. Yang B, Liu JB, So SK, Han SS, Wang SS, Hertz A, et al. Disparities in hepatocellular carcinoma incidence by race/ethnicity and geographic area in California: Implications for prevention. *Cancer*. 2018; 124: 3551–3559. <https://doi.org/10.1002/cncr.31598> PMID: 30113700
25. Ray B, Ozcagli E, Sadee W, Wang D. CYP2D6 haplotypes with enhancer single-nucleotide polymorphism rs5758550 and rs16947 (*2 allele): implications for CYP2D6 genotyping panels. *Pharmacogenet Genomics*. 2019; 29: 39–47. <https://doi.org/10.1097/FPC.0000000000000363> PMID: 30520769
26. Mizutani T. PM frequencies of major CYPs in Asians and Caucasians. *Drug Metab Rev*. 2003; 35: 99–106. <https://doi.org/10.1081/dmr-120023681> PMID: 12959412
27. Hwang GS, Bhat R, Crutchley RD, Trivedi MV. Impact of CYP2D6 polymorphisms on endoxifen concentrations and breast cancer outcomes. *Pharmacogenomics J*. 2018; 18: 201–208. <https://doi.org/10.1038/tpj.2017.36> PMID: 28762370
28. Neafsey P, Ginsberg G, Hattis D, Sonawane B. Genetic polymorphism in cytochrome P450 2D6 (CYP2D6): Population distribution of CYP2D6 activity. *J Toxicol Environ Health B Crit Rev*. 2009; 12: 334–61. <https://doi.org/10.1080/10937400903158342> PMID: 20183526
29. He X, Feng S. Role of Metabolic Enzymes P450 (CYP) on Activating Procarcinogen and their Polymorphisms on the Risk of Cancers. *Curr Drug Metab*. 2015; 16: 850–63. Available: <http://www.ncbi.nlm.nih.gov/pubmed/26652254> <https://doi.org/10.2174/138920021610151210164501> PMID: 26652254
30. Hertz DL, Kidwell KM, Hilsenbeck SG, Oesterreich S, Osborne CK, Philips S, et al. CYP2D6 genotype is not associated with survival in breast cancer patients treated with tamoxifen: results from a population-based study. *Breast Cancer Res Treat*. 2017; 166: 277–287. <https://doi.org/10.1007/s10549-017-4400-8> PMID: 28730340
31. Ahern TP, Hertz DL, Damkier P, Ejlersen B, Hamilton-Dutoit SJ, Rae JM, et al. Cytochrome P-450 2D6 (CYP2D6) Genotype and Breast Cancer Recurrence in Tamoxifen-Treated Patients: Evaluating the Importance of Loss of Heterozygosity. *Am J Epidemiol*. 2017; 185: 75–85. <https://doi.org/10.1093/aje/kww178> PMID: 27988492
32. Agúndez JA, Olivera M, Ladero JM, Rodríguez-Lescure A, Ledesma MC, Díaz-Rubio M, et al. Increased risk for hepatocellular carcinoma in NAT2-slow acetylators and CYP2D6-rapid metabolizers. *Pharmacogenetics*. 1996; 6: 501–12. <https://doi.org/10.1097/00008571-199612000-00003> PMID: 9014199
33. Agúndez JA, Ledesma MC, Benítez J, Ladero JM, Rodríguez-Lescure A, Díaz-Rubio E, et al. CYP2D6 genes and risk of liver cancer. *Lancet (London, England)*. 1995; 345: 830–1. [https://doi.org/10.1016/s0140-6736\(95\)92965-7](https://doi.org/10.1016/s0140-6736(95)92965-7) PMID: 7898230
34. Ho MT, Kelly EJ, Bodor M, Bui T, Kowdley KV, Ho RJY. Novel cytochrome P450-2D6 promoter sequence variations in hepatitis C positive and negative subjects. *Ann Hepatol*. 2011; 10: 327–332. [https://doi.org/10.1016/s1665-2681\(19\)31545-5](https://doi.org/10.1016/s1665-2681(19)31545-5) PMID: 21677335
35. Chen WS, Manova K, Weinstein DC, Duncan SA, Plump AS, Prezioso VR, et al. Disruption of the HNF-4 gene, expressed in visceral endoderm, leads to cell death in embryonic ectoderm and impaired gastrulation of mouse embryos. *Genes Dev*. 1994; 8: 2466–2477. <https://doi.org/10.1101/gad.8.20.2466> PMID: 7958910
36. Walesky C, Apte U. Role of Hepatocyte Nuclear Factor 4 α (HNF4 α) in Cell Proliferation and Cancer. *Gene Expr*. 2015; 16: 101–108. <https://doi.org/10.3727/105221615X14181438356292> PMID: 25700366
37. Corchero J, Granvil CP, Akiyama TE, Hayhurst GP, Pimprale S, Feigenbaum L, et al. The CYP2D6 humanized mouse: effect of the human CYP2D6 transgene and HNF4 α on the disposition of debrisoquine in the mouse. *Mol Pharmacol*. 2001; 60: 1260–7. <https://doi.org/10.1124/mol.60.6.1260> PMID: 11723233

38. Yang X, Zhang B, Molony C, Chudin E, Hao K, Zhu J, et al. Systematic genetic and genomic analysis of cytochrome P450 enzyme activities in human liver. *Genome Res.* 2010; 20: 1020–1036. <https://doi.org/10.1101/gr.103341.109> PMID: 20538623
39. Pan X, Ning M, Jeong H. Transcriptional Regulation of CYP2D6 Expression. *Drug Metab Dispos.* 2017; 45: 42–48. <https://doi.org/10.1124/dmd.116.072249> PMID: 27698228
40. Kishimoto T. Interleukin-6: discovery of a pleiotropic cytokine. *Arthritis Res Ther.* 2006; 8 Suppl 2: S2. <https://doi.org/10.1186/ar1916> PMID: 16899106
41. Giannitrapani L, Soresi M, Balasus D, Licata A, Montalto G. Genetic association of interleukin-6 polymorphism (-174 G/C) with chronic liver diseases and hepatocellular carcinoma. *World J Gastroenterol.* 2013; 19: 2449–55. <https://doi.org/10.3748/wjg.v19.i16.2449> PMID: 23674845
42. Babeu JP, Boudreau F. Hepatocyte nuclear factor 4-alpha involvement in liver and intestinal inflammatory networks. *World J Gastroenterol.* 2014; 20: 22–30. <https://doi.org/10.3748/wjg.v20.i1.22> PMID: 24415854
43. Kim S, Östör AJK, Nisar MK. Interleukin-6 and cytochrome-P450, reason for concern? *Rheumatol Int.* 2012; 32: 2601–2604. <https://doi.org/10.1007/s00296-012-2423-3> PMID: 22451032
44. Yang Q, Doshi U, Li N, Li AP. Effects of Culture Duration on Gene Expression of P450 Isoforms, Uptake and Efflux Transporters in Primary Hepatocytes Cultured in the Absence and Presence of Interleukin-6: Implications for Experimental Design for the Evaluation of Downregulatory Effects of Biotherapeutics. *Curr Drug Metab.* 2012; 13: 938–946. <https://doi.org/10.2174/138920012802138570> PMID: 22475270
45. Fisher CD, Lickteig AJ, Augustine LM, Ranger-Moore J, Jackson JP, Ferguson SS, et al. Hepatic Cytochrome P450 Enzyme Alterations in Humans with Progressive Stages of Nonalcoholic Fatty Liver Disease. *Drug Metab Dispos.* 2009; 37: 2087–2094. <https://doi.org/10.1124/dmd.109.027466> PMID: 19651758
46. Christmas P. Role of Cytochrome P450s in Inflammation. *Adv Pharmacol.* 2015; 74: 163–92. <https://doi.org/10.1016/bs.apha.2015.03.005> PMID: 26233907
47. Stipp MC, Acco A. Involvement of cytochrome P450 enzymes in inflammation and cancer: a review. *Cancer Chemother Pharmacol.* 2021; 87:295–309. <https://doi.org/10.1007/s00280-020-04181-2> PMID: 33112969
48. Jiang F, Yeo CW, Lee SS, Oh MK, Ghim JL, Shon JH, et al. Effect of HNF4α genetic polymorphism G60D on the pharmacokinetics of CYP2D6 substrate tolterodine in healthy Korean individuals. *Pharmacogenet Genomics.* 2013; 23: 175–179. <https://doi.org/10.1097/FPC.0b013e32835de25e> PMID: 23292115
49. Cardoso R, Lacerda PC, Costa PP, Machado A, Carvalho A, Bordalo A, et al. Estrogen Metabolism-Associated CYP2D6 and IL6-174G/C Polymorphisms in *Schistosoma haematobium* Infection. *Int J Mol Sci.* 2017; 18: 2560. <https://doi.org/10.3390/ijms18122560> PMID: 29182577