Polymorphic Cytochrome P450 Enzymes (CYPs) and Their Role in Personalized Therapy

Sarah C. Preissner¹[®], Michael F. Hoffmann¹[®], Robert Preissner¹, Mathias Dunkel¹, Andreas Gewiess², Saskia Preissner^{1,3*}

1 Charité Universitätsmedizin Berlin, ECRC - Structural Bioinformatics Group, Deutsches Konsortium für Translationale Krebsforschung (DKTK), Berlin, Germany, 2 Institute for Laboratory Medicine, Berlin, Germany, 3 Charité Universitätsmedizin Berlin, Dental, Oral and Maxillary Medicine, CC3, Berlin, Germany

Abstract

The cytochrome P450 (CYP) enzymes are major players in drug metabolism. More than 2,000 mutations have been described, and certain single nucleotide polymorphisms (SNPs) have been shown to have a large impact on CYP activity. Therefore, CYPs play an important role in inter-individual drug response and their genetic variability should be factored into personalized medicine. To identify the most relevant polymorphisms in human CYPs, a text mining approach was used. We investigated their frequencies in different ethnic groups, the number of drugs that are metabolized by each CYP, the impact of CYP SNPs, as well as CYP expression patterns in different tissues. The most important polymorphic CYPs were found to be 1A2, 2D6, 2C9 and 2C19. Thirty-four common allele variants in Caucasians led to altered enzyme activity. To compare the relevant Caucasian SNPs with those of other ethnicities a search in 1,000 individual genomes was undertaken. We found 199 non-synonymous SNPs with frequencies over one percent in the 1,000 genomes, many of them not described so far. With knowledge of frequent mutations and their impact on CYP activities, it may be possible to predict patient response to certain drugs, as well as adverse side effects. With improved availability of genotyping, our data may provide a resource for an understanding of the effects of specific SNPs in CYPs, enabling the selection of a more personalized treatment regimen.

Citation: Preissner SC, Hoffmann MF, Preissner R, Dunkel M, Gewiess A, et al. (2013) Polymorphic Cytochrome P450 Enzymes (CYPs) and Their Role in Personalized Therapy. PLoS ONE 8(12): e82562. doi:10.1371/journal.pone.0082562

Editor: Daotai Nie, Southern Illinois University School of Medicine, United States of America

Received July 22, 2013; Accepted October 24, 2013; Published December 10, 2013

Copyright: © 2013 Preissner et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by: BMBF ImmunoTox, European Union Seventh Framework Programme SYNSYS (Synaptic Systems: dissecting brain function in health and disease), DKTK. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

* E-mail: saskia.preissner@charite.de

Introduction

Inter-individual variability of drug response and drug clearance is a complex and common problem in clinical practice [1]. Overlapping substrate specificity of enzymes, a multitude of single nucleotide polymorphisms (SNPs) [2] and variations between ethnic groups [3] make prediction of phenotypic drug response difficult. To avoid treatment failure and unnecessary toxicity, tailoring dosages and drug-cocktails for each individual is essential [4].

Differences in drug response can be attributed to variability in DNA sequences of specific genes which's products are crucial for drug metabolism. For instance, SNPs in phase 1 enzymes, such as cytochrome P450 oxidases (CYPs) [3], phase 2 enzymes, such as Uridine 5'-diphosphoglucuronosyltransferase (UGTs) [5], and absorptive and efflux transporters, such as ATP-binding cassette transporters (ABCtransporters) [4], have been previously reported. Characterization of these enzymes and the effects of minor allele variants on the metabolism of specific drugs have been described in the literature and have recently been compiled by our group into a comprehensive database called SuperCYP [6]. Phase I reactions include oxidation, reduction, hydrolysis and cyclization. Using oxygen and NADPH as a co-substrate, CYPs are the major enzymes responsible for catalyzing such reactions [7] and account for approximately 75% of total drug metabolism [8].

The Human Genome Project identified 57 human CYPs, which were classified into 18 families and 43 subfamilies based on sequence similarity [9]. CYP families 1, 2 and 3 are responsible for metabolism of drugs, xenobiotics and certain endogenous molecules [3] and hence are of particular relevance to this current study. Most CYPs metabolize more than one drug. Similarly, a drug is often metabolized by multiple CYPs. Drugs can also inhibit or induce CYP activity, either by directly interacting with the enzyme or altering its

[•] These authors contributed equally to this work.

expression. Characterization of these interactions is important to determine and predict compatible drug combinations [10]. Human CYPs are primarily membrane-associated proteins [11] that are ubiquitously expressed in most tissues. Highest expressions are generally found in liver tissue, but the distribution of particular CYPs varies [12], which indicates that the actual efficiency of a drug is likely to depend on CYP expression in the target tissue. There are significant interindividual differences in enzyme activity leading to distinct phenotypes. For example the most frequent phenotype of CYP 2D6 is the extensive-metabolizer (78.8%), followed by intermediate- (12.1%), poor (7.6%) and ultra-rapid metabolizers (1.5%) [13].

In addition to drug catabolism, many CYPs are responsible for activation of prodrugs, such as cancer therapeutics [14] and antipsychotics [15]. Prodrugs are pharmacologically inactive compounds that require activation via metabolic conversion [16], allowing control of where, when and how much drug activity occurs [17]. This is particularly important for chemotherapeutic drugs, where the active drug ideally only acts on tumor cells in order to reduce toxic side effects [18]. Prodrugs can be activated by photo irradiation [19], change in pH [20] or enzymatically [21], for instance by CYPs [22]. Polymorphisms in CYPs can result in ineffective or aberrant activation of prodrugs [22], which can lead to toxicity [4]. Fortunately, advances in genetic research have made genotyping of a large number of patients possible, leading to identification of SNPs that alter expression or activity of drug metabolizing enzymes [3]. In this study we set out to determine the most frequent CYP polymorphisms having the highest impact on drug metabolism in Caucasians. This knowledge could facilitate the development of tests for efficient genotyping of patients thus leading to a better and more personalized treatment.

Methods

Text mining

Information on drug metabolism can be found in more than 100,000 PubMed articles, yet limited data is available regarding the frequencies of SNPs in human CYPs. To identify relevant articles, a specific search tool was developed for text mining literature using Apache Lucene™ (http://lucenenet.apache.org) as a search engine library and LingPipe (http://alias-i.com/ lingpipe). Figure 1 summarizes the different methods used for the textmining approach. Complete Medline/PubMed data were downloaded from the NCBI FTP site in xml-format and then indexed. The indexed data was dynamically queried by a search engine written in Java that outputs an sql-file with the text mining hits, which served afterwards for manual validation. The search engine comprises several lists of synonyms for identifying entities, such as chemical compounds, biological targets, genes, cell types and polymorphisms, as well as interaction-related entities. If available, information on CYP polymorphism was extracted from the literature. Definitions and synonyms are included from UMLS® Metathesaurus®, that contains millions of biomedical and health related concepts, their synonymous names, and their relationships. As an example, the query for CYP2C19 was like:

(Abstract: CYP2C19** OR Title: CYP2C19**) AND (Abstract: population OR Title: population) AND (Abstract: effect OR Title: effect) AND (Abstract: frequenc* OR Title: frequenc*).

The term 'CYP2C19' was replaced through each human CYP and synonyms, as well as different ethnicity and outcome terms were used for 'population' and 'effect'. The positional distance between the different terms had to be restricted to reduce false positive hits, when terms occurred far from each other in the abstract. The records found were scored rulebased. The rules employed order, redundancy, distance, topic segmentation and sentence breaking for boundaries. For example, a distance \leq 7 between the CYP and the ethnicity and \leq 6 between the frequency and the CYP was given a score of 100. Greater distances and negative interaction words resulted in lower scores. Duplicates were removed and a team of scientists manually processed 1,037 papers found in PubMed for relevance to polymorphisms and their frequency in Caucasian populations. The team consisted of three medical scientists, with three years experience in validation of text mining results. During this time, they reviewed over 10,000 abstracts with the focus on CYPs. A weekly meeting took place to ensure and raise the quality of text mining and to discuss problems. The aim was to achieve a coherent review operation. The text mining validation tool is shown in Figure 2. CYP polymorphisms that occurred with a frequency of more than one percent in the Caucasian population were included in this study.

Localization of SNPs in a 3D CYP model

The evolutionary conservation taken from a multiple sequence alignment of CYPs was projected onto the 3D structure using CYP 2D6 as template (PDB ID: 3TDA). Frequent SNPs in the four most polymorphic CYPs (1A2, 2C9, 2C19 and 2D6) were labeled in the 3D model. The number of mutations was used to determine the thickness of the ribbon (Figure 3).

CYP SNPs and 1,000 Genomes

The 1,000 Genomes Project (www.1000genomes.org) is an international initiative designed to provide full genomic sequence information from an ethnically diverse population [23]. CYP SNPs in 1,092 individuals were extracted using the online data slicer from the 1,000 Genomes Project (http:// browser.1000genomes.org). Frequency analysis focused on non-synonymous coding SNPs with a prevalence of one percent or higher in all genomes regardless of ethnicity. The search included the main 29 CYP alleles from "The Human Cytochrome P450 (CYP) Allele Nomenclature Database" (http://www.cypalleles.ki.se/) [24]. In addition, 16 CYP alleles not listed in the CYP allele database due to very heterogeneous distributions were included. The 1,000 Genomes Database includes SNP effect predictions on CYPs, calculated by PolyPhen [25], which predicts possible functional alterations in human proteins after amino acid substitution based on physical and comparative considerations [26].

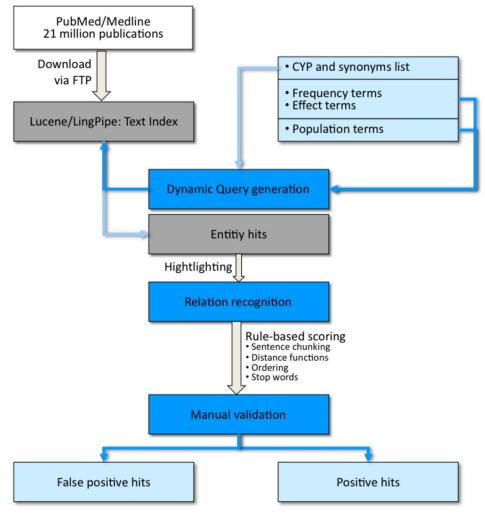


Figure 1. Flow-chart of the methods used for the text mining approach. doi: 10.1371/journal.pone.0082562.g001

Expression data

Affymetrix data was used to compare human CYP mRNA expression in 41 different types of tissue, further subcategorized into different regions of an organ, yielding a total of 65 tissue types. The series of datasets obtained from GEO (Gene Expression Omnibus, http://www.ncbi.nlm.nih.gov/ geo/) were originally generated from 10 post-mortem donors (5 females and 5 males), and represent normal human tissues (Series GSE3526) [27]. The 84 probe sets, which measure the expression level of CYPs were normalized and assigned to 40 types of CYPs. To display differences in expression, a heatmap was generated using Genesis software [28]. Relative expression was calculated as the intensity of the gene in the region minus the mean intensity of the gene in all regions then divided by the standard deviation. This heat-map served as data source for the CYP body map in which only two-fold decreased or increased values were considered.

Our work would not have been possible without the publicly available datasets mentioned above. We are grateful and honor the work of involved research groups.

Results

Frequencies of SNPs in CYPs

Analysis of the SNPs identified by text mining, showed that SNPs predominantly occurred in 3 polymorphic CYPs (2D6, 2A6 and 2B6) regardless of ethnic group. Only frequencies of known nucleotide changes were assessed to identify the extent of SNPs in CYPs. Figure 4 displays 9 CYPs, including 2D6 (114 SNPs), 2A6 (68 SNPs) and 2B6 (57 SNPs), which showed the highest number of SNPs. For other CYPs, the number of known SNPs was less than 22. CYP 2D6 is a major polymorphic CYP and, as expected, was the greatest contributor of polymorphic alleles in Caucasians.

Count	Score	PMID	Sentence		TR	RUE F.	ALSE	NOT SURE	Comm	ent
1	100	11014415	Because CYP2C19*2 is not able to explain metabolizers, other mutations (CYP2C19*4 to present in North Indians.							
2	100	10460072	Four individuals (1.0%) were predicted to be poor metabolizers (CYP2C19*2/*2), a significantly lower frequency compared to Middle European populations.			v			*2: 1.0%	
3	100	18240905	Among the 250 AJ individuals, the CYP2C9*1, *2, *3 and *5 allele frequencies were 0.772, 0.140, 0.086 and 0.002, respectively, and the genotypes were distributed into extensive- (60.8%), intermediate- (32.8%) and poor- (6.4%) metabolizer			✓			*1: 77.2%;	*2: 1.4%
4	100	21410749	Differences between Spaniards and Mestizo were detected in relation to the frequencies linked to either absent enzyme activity. CYP2 7.1%, respectively), or reduced CYP2A6 enz CYP2A6*9A (6.4 and 10.3%, respect	of the allele 2 <mark>A6*4A</mark> (4 a zyme activit	and	elat 7	ion	/alida	ation *4A: 7.1% c	lecreas
		b	ack to home		Sc	ientis	t 1		SAV	
Abs	strac	t♥		Results	na franuancias amo	on both Connin	rde and Equadoria	as correspond to those	Jump to	Mainhara Inur
	etween Sna		21410749	(P > 0.05).						
copulations. cuadorian M cyP2A6*4A	This study w Mestizo peop variant allele	as designed to ble regarding C1 es at the CYP2/	adorians in CYP2A6 allele frequencies: comparison with other investigate the potential differences between Spaniards and YP2A6*1A, CYP2A6*1B1, CYP2A6*1x2A, CYP2A6*0A, and A6 gene and also to compare the observed frequencies with thos oups. DNA from 234 Spaniard and 300 Ecuadorian subjects were	only exception being t prevalence in Spaniar previously found amor Spaniards and a Finni this work was similar (Table I. Frequenc	that of the CYP2A6 ards (4%) was the sa ong white Canadian hish population (1%) ($P > 0.05$) to that of	*4A allele assor ime than that pr people (1.2%) ([8] shows a bo oserved in our p	ciated with abolish reviously reported P < 0.001 [25]. In referine significant revious study in 10	ed enzyme activity. In in a French Caucasia addition, the compar- ce level (<i>P</i> = 0.07). Or 00 Spaniards [8].	er Caucasian populations : this regard, although the C n population [15], it was his ison of the CYP2A6*A fre- n the other hand, CYP2A6* e found in other population	CYP2A6*4A gher than that quencies betweer 4A frequency in
opulations. cu ado rian M YP2A 6*4 A revio usly rep nalyzed by e	This study w Mestizo peop variant allele ported in diff either PCR o	ras designed to ble regarding C es at the CYP2 ferent ethnic gro or PCR-restriction	investigate the potential differences between Spaniards and (P2A6*1A, CYP2A6*1B1, CYP2A6*1x2A, CYP2A6*9A, and A6 gene and also to compare the observed frequencies with those oups, DNA from 234 Spaniard and 300 Ecuadorian subjects werk on fragment length polymorphism. Differences between Spania	only exception being t previously found amo Spaniards and a Finni this work was similar (Table I. Frequence CYP2A6 Sp allele Stru	that of the CYP2A6 ards (4%) was the sa ong white Canadian nish population (1%) ($P > 0.05$) to that of cles of CYP2A6 all opulation panlard (this Ecu:	*4A allele assor ime than that pr people (1.2%) () (§) shows a bo observed in our p eles (%) observed adorian (this	ciated with abolish reviously reported P < 0.001 [25]. In referine significant revious study in 10	ed enzyme activity. In in a French Caucasia addition, the compar- ce level (<i>P</i> = 0.07). Or 00 Spaniards [8].	this regard, although the C n population [15], it was hig ison of the CYP2A6*4A free the other hand, CYP2A6*	CYP2A6*4A gher than that quencies betweer 4A frequency in
opulations. cu ado rian M YP2A6*4A revio usly rep nalyzed by e nd Mestizo	This study w Mestizo peop variant allele ported in diff either PCR o Ecuadoriar	vas designed to ole regarding C es at the CYP2/ ferent ethnic gro or PCR-restrictions were detect	investigate the potential differences between Spaniards and (P2A6*1A, CYP2A6*1B1, CYP2A6*1x2A, CYP2A6*9A, and A6 gene and also to compare the observed frequencies with those oups. DNA from 234 Spaniard and 300 Ecuadorian subjects werr on fragment length polymorphism. Differences between Spania ed in relation to the frequencies of the alleles linked to eithe	only exception being to previously found amously spaniards and a Finni this work was similar (Table I. Frequence CYP2A6 Spp allele stu CYP2A6-14 64.	that of the CYP2A6 ards (4%) was the sa ong white Canadian hish population (1%) ($P > 0.05$) to that of celes of CYP2A6 alle opulation panlard (this Ecu: udy) stud	*4A allele assoc ime than that pr people (1.2%) ([§] shows a bo oserved in our p eles (%) observed adorian (this iy)	ciated with abolish eviously reported P < 0.001) [25]. In rderline significant revious study in 11 red in this study	ed enzyme activity. In in a French Caucasia addition, the compar- te level (<i>P</i> = 0.07). Or 00 Spaniards [<u>8</u>]. compared with thos	this regard, although the C n population [15], it was high son of the CVP2A6*4 fre to the other hand, CVP2A6* e found in other population Japanese 3 16.4 (2444) [17]	CYP246*4A ther than that quencies betweer 4A frequency in ons African from Ghana 80.5 (420)
opulations. cuadorian M YP2A 6*4 A reviously rej nalyzed by e nd Mestizo bs ent enzy	This study w <u>Aestizo</u> peop variant allele ported in diff either PCR o <u>Ecuadoriar</u> me activity,	as designed to oble regarding C es at the CYP2/ ferent ethnic gro or PCR-restrictions were detects CYP2A6*4A (4	investigate the potential differences between Spaniards and (P2A6*1A, CYP2A6*1B1, CYP2A6*1x2A, CYP2A6*9A, and A6 gene and also to compare the observed frequencies with those oups, DNA from 234 Spaniard and 300 Ecuadorian subjects werk on fragment length polymorphism. Differences between Spania	only exception being t previously found amo Spaniards and a Finni this work was similar (Table I. Frequenc CYP246 sp allele stu CYP246*1A 64.	that of the CYP2A6 drds (4%) was the sa ong white Canadian ong white Canadian (P> 0.05) to that ob cles of CYP2A6 alle oppulation panlard (this Ecu: udy) stud 5.9 (488) 61.7	*4A allele assoc ime than that pr people (1.2%) ([§] shows a bo oserved in our p eles (%) observed adorian (this iy)	ciated with abolish eviously reported (P < 0.001) [25]. In rederline significance revious study in 11 red in this study Caucasian	ed enzyme activity. In na French Caucasia addision, the compar- te level (P = 0.07). Or 00 Spaniards [3]. compared with thos Chinese 27.2 (192) [34.5 (192) [this regard, although the C in population [15], it was hig son of the CVP2A6*44 free to the other hand, CVP2A6* e found in other population Japanese [3] 16.4 (2444) [17] -20.3 (368) [32] [3] 27.0 (368) [32]	CYP246*4A ther than that quencies between 4A frequency in ons African from Ghana 80.5 (420) [33] 11.9 (420)
opulations. ⁻ cuadorian N YP2A6*4A reviously rep nalyzed by o nd Mestizo bs ent enzy YP2A9*9A igher than th	This study w Mestizo peop variant allele ported in diff either PCR o Ecuadoriar me activity, (6.4 and 10. nose in Africa	as designed to ble regarding CV es at the CYP2/ ferent ethnic gro or PCR-restriction ns were detect CYP2A6*4A (.3%, respected ans or Caucasi	investigate the potential differences between Spaniards and YP2A6*1A, CYP2A6*1B1, CYP2A6*1x2A, CYP2A6*9A, and A6 gene and also to compare the observed frequencies with thor ours. DNA from 234 Spaniard and 300 Ecuadorian subjects werr on fragment length polymorphism. Differences between Spania ed in relation to the frequencies of the alleles linked to eithe 4 and 7.1%, respectively), or reduced CYP2A6 enzyme activi y). CYP2A6*4A and CYP2A6*9A frequencies in Ecuadorians we an groups and lower than those in Asian. This study provides, fo	only exception being to prevalence in Spaniar previously found amo Spaniards and a Finin this work was similar (Table I. Frequenc CYP2A6 Sp allele stu CYP2A6*1A 64.	that of the CYP2A6 drds (4%) was the sa ong white Canadian ong white Canadian (P> 0.05) to that ob cles of CYP2A6 alle oppulation panlard (this Ecu: udy) stud 5.9 (488) 61.7	*4A allele associate than that properties (1.2%) (1.2\%) (1	siated with abolish eviously reported P < 0.001) [25]. In derline significant revious study in 11 red In this study Caucasian 67.0 (926) [15] 33.5 (1416) [25] 1 (200) [8]–1.2 (2	ed enzyme activity. In n a French Caucasia addision, the compar- be level (P = 0.07). Or 00 Spaniards [5]. compared with thos Chinese 27.2 (162) [34.5 (162) [51.3 (226) 336) 6.7 (224) [2]	this regard, although the C population [13], was high son of the CYP2A6*4 the the other hand, CYP2A6*1 e found in other population Japanese 3 16.4 (2444) [17] -20.3 (268) [22] 3 27.0 (268) [23] 3 24.2 (128) [25] 4 4.4 (128) [25] 4 24.2 (128) [25]	CYP246'4A ther than that quencies between 4A frequency in ons African from Ghana 80.5 (420) [33] 11.9 (420) [33]
opulations. cu ado rian N (YP2A6*4A) reviously rep nalyzed by of nd Mestizo bs ent enzyn (YP2A9*9A) igher than the rst time, the	This study we destize peop variant allele ported in diff either PCR of Ecuadoriar me activity, (6.4 and 10. nose in Africe result of the	as designed to ble regarding C es at the CYP2/ ferent ethnic gro or PCR-restrictions were detect CYP2A6*4A (.3%, respected ans or Caucasi e analysis of CY	investigate the potential differences between Spaniards and (P2A6*1A, CYP2A6*1B1, CYP2A6*1x2A, CYP2A6*9A, and A6 gene and also to compare the observed frequencies with thor oups. DNA from 234 Spaniard and 300 Ecuadorian subjects werr on fragment length polymorphism. Differences between Spania ed in relation to the frequencies of the alleles linked to eithe 4 and 7.1%, respectively), or reduced CYP2A6 enzyme activi (). CYP2A6*4A and CYP2A6*9A frequencies in Ecuadorians we an groups and lower than those in Asian. This study provides, for P2A6 allele frequency in a South American population and	only exception being " previation to "Spaniar previation to "Spaniards and a Fino Spaniards and a Fino Spaniards and a Fino Spaniards and a Fino this work was similar 1 Table I. Frequence CYP2A6*1A 64. CYP2A6*1B1 30. CYP2A6*1B1 40.	that of the CYP246 indis (4%) was the same ong white Canadian insh population (1%) (P > 0.05) to that of cless of CYP248 alle opulation apaniard (this Ecu- udy) stud 4.9 (468) 61.7 0.9 (468) 7.1 (*4A allele association that prepare that that prepare (1.2%) ((1) (2) shows a bosserved in our preles (%) observed in our preles	siated with abolish eviously reported P < 0.001) [25] In rderline significanc revious study in 11 red in this study of Caucasian 67.0 (926) [15] 33.5 (1416) [25]	ed enzyme activity. In n a French Caucasia addision, the compar- be level (P = 0.07). Or 00 Spaniards [5]. compared with thos Chinese 27.2 (162) [34.5 (162) [51.3 (226) 336) 6.7 (224) [2]	the regard, athrough two is population [13, 24, 24, 24, 24, 24, 24, 24, 24, 24, 24	CYP246*4A ther than that quencies between 4A frequency in ons African from Ghana 80.5 (420) [33] 11.9 (420)
opulations. cu ado rian N (YP2A6*4A) reviously rep nalyzed by of nd Mestizo bs ent enzyn (YP2A9*9A) igher than the rst time, the	This study we destize peop variant allele ported in diff either PCR of Ecuadoriar me activity, (6.4 and 10. nose in Africe result of the	as designed to ble regarding CV es at the CYP2/ ferent ethnic gro or PCR-restriction ns were detect CYP2A6*4A (.3%, respected ans or Caucasi	investigate the potential differences between Spaniards and (P2A6*1A, CYP2A6*1B1, CYP2A6*1x2A, CYP2A6*9A, and A6 gene and also to compare the observed frequencies with thor oups. DNA from 234 Spaniard and 300 Ecuadorian subjects werr on fragment length polymorphism. Differences between Spania ed in relation to the frequencies of the alleles linked to eithe 4 and 7.1%, respectively), or reduced CYP2A6 enzyme activi (). CYP2A6*4A and CYP2A6*9A frequencies in Ecuadorians we an groups and lower than those in Asian. This study provides, for P2A6 allele frequency in a South American population and	only exception being " previation to "Spaniar previation to "Spaniards and a Fino Spaniards and a Fino Spaniards and a Fino Spaniards and a Fino this work was similar 1 Table I. Frequence CYP2A6*1A 64. CYP2A6*1B1 30. CYP2A6*1B1 40.	that of the CYP2A6 trids (4%) was the as ong white Canadian insh population (1%) (P> 0.05) to that of cies of CYP2A6 alide oppulation paniar (this Eccu udy) stud 1.9 (468) 61.7 0.9 (468) 7.1 (4 (468) 10.3	*4A allele associate that properties that provide that provide that provide that provide that the provide that the provided that the provi	siated with abolish eviously reported P < 0.001) [25]. In derline significant revious study in 11 red in this study Caucasian 67.0 (926) [15] 33.5 (1416) [25] 1 (200) [8]–1.2 (2 [25]–4.0 (926) [15]	de enzyme activity. In a French Caucasia addition, the compar- te level (P = 0.07). OC Spaniards []. compared with thos Chinese 27.2 (162) [34.5 (162) [-51.3 (226)] -15.1 (182) 15.6 (224) []	the regard, athrough the population (15, 41 was his population (15, 41 was his soon of the CVP2A64A fee the other hand, CVP2A64 Jupanese 10, 10, 41 (24, 41, 12), -20, 3 (366) (22) 20, 20, 30(56) (22) 20, 20, 20, 20, 20, 20, 20, 20, 20, 20,	CYP246'4A ther than that quencies between 4A frequency in ons African from Ghana 80.5 (420) [33] 11.9 (420) [33]

Figure 2. Text mining validation tool. The table shows the text mining validation tool with columns for score, PubMedID, relation sentence and checkboxes for the validation. The SNPs are highlighted in blue, frequencies in green, effects in orange and the ethnicities in red. 'Scientist 1' reads the abstract and, if necessary, has access to full text. Afterwards, the relation has to be validated as 'true', 'false' or 'not sure'. If the relation is 'true', the relation is copied into the 'comment' field. These relations are copied into a new sql-file. If 'Scientist 1' activates the 'not sure' field, the relation has to be validated again by another scientist. doi: 10.1371/journal.pone.0082562.g002

Allelic frequency in CYPs of Caucasians

The PubMed search yielded articles on 34 different CYP alleles with an occurrence greater than one percent in the Caucasian population (Table 1), which may be indicative of altered substrate metabolism. CYP 2D6 and 2B6 possessed the largest number of alleles with a known impact on metabolism in Caucasian population (11 and 6, respectively). Maximum allele frequencies in CYP 2D6 varied from 20.7% to 32.4%. When considering all 34 alleles, the most frequent alleles were CYP 3A5*3C at 81.3% (decreased enzyme activity) followed by CYP 1A2*1F at 33.3% (increased enzyme activity). Furthermore, the major alleles leading to increased metabolism were CYP 2A6*1B (30.0%), 3A4*1B (17.0%), 1A1*2A (19.0%) and 2C19*17 (18.0%). In contrast, decreased metabolism was attributed to 2D6*2A (32.4%), 2D6*4 (20.7%) and 2C9*2 (16.0%). Carrying the 2A6*4 allele (1.0%) leads to an inactive enzyme with no detectable substrate metabolism.

Comparison among different ethnic groups revealed that frequencies differed considerably and displayed a

heterogeneous distribution of CYP alleles. For instance, in Asian and African populations, CYP2A6*2 possessed a frequency of 28.0% and 62.0%, respectively, whereas a frequency of 8.0% was observed in Caucasians. A more detailed table with additional information on CYP SNPs in Caucasians and other ethnic groups is available in Table S1.

Major drug metabolizing CYPs

Not all 57 human CYPs are involved in drug metabolism. The primary CYPs responsible for drug metabolism were determined by first ranking the CYPs according to the total number of drug substrates (Figure 5). Twelve CYPs accounted for 93.0% of drug metabolism, regarding to the entire number of 1.839 known drug-metabolizing-reactions in the SuperCYP database. CYP 1A2, 2D6, 2C9 and 2C19 were responsible for nearly 40.0% of drug metabolism and including CYP 3A4 even for 60.0%.

Since the described four CYPs are highly polymorphic and commonly occur in Caucasians, further detailed analyses were

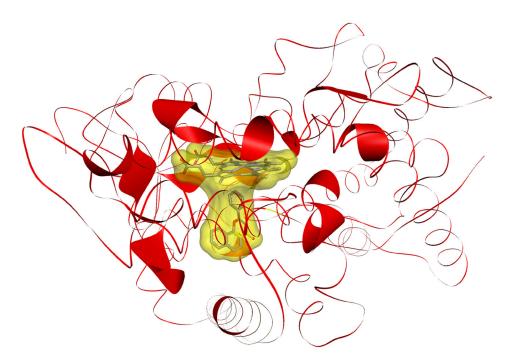


Figure 3. 3D structure of a cytochrome with heme and a ligand with localization of frequent SNPs. Frequent SNPs in the four most polymorphic CYPs (1A2, 2C9, 2C19 and 2D6) were labeled in this 3D model of CYP 2D6 (PDB ID: 3TDA). Therefore, the ribbon was enlarged in the appropriate positions. The binding-side (transparent, orange colored surface) contains an iron ion and a porphyrine ring (heme). Frequent mutations of CYPs at the binding-side occurred at the following positions: 67, 89, 107, 117, 118, 120, 125, 132, 151, 201, 227, 261, 325, 377, 382, 386, 410, 454, 456, 469 and 470. doi: 10.1371/journal.pone.0082562.g003

restricted to these four CYPs. On the overall CYP system, it is expected that these four CYPs would have the greatest impact on inter-individual variability of drug response. Although CYP 2A6 and 2B6 possess various relevant alleles in Caucasians, they do not cover a large range of drug interactions (51 and 74 substrates, respectively).

All SNPs in the four most polymorphic CYPs (1A2, 2C9, 2C19 and 2D6) influenced enzymatic activity due to localization in the substrate-binding cavity as shown in Figure 3.

Expression data

Because of high CYP expression levels in some tissues, an impact of CYP isoforms in particular tissues can be deduced. The work of Nishimura and colleagues demonstrated differences in CYP mRNA expression in various human tissues. For example, CYP 2F1, 4B1, 4F8, 11S, 11A, 11B1, 11B2, 19 and 24 are not expressed in the liver [29]. The current study confirmed these results and extended the findings, which are shown in Figure 6. Nishimura analyzed mRNA levels of 30 CYP isoforms in 11 tissue types. Similarly, the current study investigated the expression of 40 CYP isoforms in 41 tissue types. The liver was considered separately in the analysis in order to identify the differences between the other tissues. In 21 different tissues, a heterogeneous distribution of CYPs was observed. For instance, 39 different CYP isoforms showed higher mRNA expression in at least one or more tissue types. Significant differences were observed in the adrenal gland

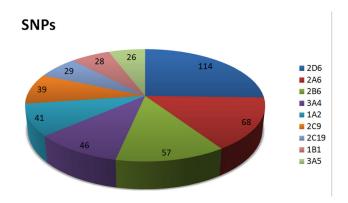


Figure 4. Number of known SNPs per CYP. The pie chart illustrates the number of known SNPs in different CYPs. doi: 10.1371/journal.pone.0082562.g004

cortex, which possessed 6-fold higher expression of CYP 11A1, 11B1 and 11B2 (compared to the mean expression). Interestingly, no other tissue showed high levels of expression of these three CYPs. Large differential expression compared to other tissues was also observed in the kidneys, where a 6-fold increase in CYP 4A22, a 5-fold increase in CYP 8B1 and 4-fold increases in 4V2, 4F2, 4A11 and 2B6 were noted. In addition, 5-fold higher expression of CYP 2C8 was found in lung, CYP 4F8 in prostate, 4F3 in bone, 2F1 in bronchial tubes and 2C8 in

Table 1. CYP SNP frequencies in Caucasians.

СҮР	Allele	Amino Acid	Caucasian (%)	Enzyme Activity	Test Drug	References (PMID)
1A1	*2A	I462V	19.0	increase	17β-estradiol	19514967
1A2	*1F	none	33.3	higher inducibility	Omeprazole	12534642 / 22299824
	*1D	none	4.82	decrease	Clozapine	12534642 / 20797314
2C9	*2	R144C	19.0% *1/*2 1.6% *2/*2 1.8 % *2/*3	decrease	Warfarin	15284536
	*3	1359L	9.0	decrease	Tolbutamide	11678789
2C19 *:	*2	Splicing I331V defect	16.0	decrease	Clopidogrel	10460072
	*17	I331V	18.0	increase	Omeprazole	21247447
2D6	*3	N166D; 259 Frameshift	2.04	decrease	Debrisoquine	9012401
	*4	P34S; L91M; H94R; Splicing defect; S486T	20.7	decrease	Dextromethorphan	9012401
	*4D	P34S; Splicing defect; S486T	3.4	decrease	Bufuralol	11266079
	*4L	P34S; Splicing defect; S486T	4.5	decrease	Bufuralol	11266079
	*5	CYP2D6 deleted	4.1	no enzyme		9511177
	*6	118Frameshift	1.3	nonfunctional		9511177
	*7	H324P	1.0	decrease	Sparteine	9089660
	*9	K281del	2.0	decrease	Sparteine	9511177
	*10	P34S; S486T	8.0	decrease	Metoprolol	9511177 / 11505219
	*41	R296C; Splicing defect; S486T	8.0	decrease (expression)		15289790
2A6	*12	10 aa substitutions	2.9	decreased (expression)		16041240
	*1B	none	32.6	increase	Caffeine	22850738
	*2	L160H	2.3	decrease	Nicotine	11259354
	*4	CYP2A6 deleted	1.0	no enzyme		11259354
	*9	(TATA box)	7.1	decreased	Nicotine	15475735
3A4	*17	F189S	2.0	decrease	Testosterone	11714865
	*1B	none	17.0	increase (transcription)	Tacrolimus	12692107
	*2	S222P	2.7	decrease	Nifedipine	10668853
3A5	*3C	Splicing defect	81.3	decrease	Sirolimus	17162466
	*3k / *10	Splicing defect; F446S	2.0	decrease	Nifedipine	12893984
3A7	*2	T409R	8.0	increase	Dehydroepiandrosterone	15903124
2B6	*2	R22C	5.3	increase	Artemether	21746968 / 12242601
	*5	R487C	14.0	decrease	Nirvanol	11470993
	*4	K262R	5.0	increase	Bupropion	14515060
	*6	Q172H; K262R	25.2	decrease (expression)	Cyclophosphamide	14515060
	*7	Q172H; K262R; R487C	3.0	decrease	7-ethoxy-4-trifluoromethylcoumarin	12242601 / 14551287
	*22	none	3.0	increase (transcription)		15722458

Polymorphisms that are relevant for Caucasians are shown here with CYP, allele, amino acid and frequencies and their effect on enzyme activity. The test drug and the PubMed ID complete the table. A more detailed table including gene information and frequencies in other ethnic groups can be found in Table S1.

doi: 10.1371/journal.pone.0082562.t001

stomach. Furthermore, CYP 2C18 showed a high level distribution restricted to the oral cavity, pharynx and esophagus. Two-fold lower expression was detected for CYP 2A1 in the esophagus, 2A7 in the prostate, as well as 2C9 and 2D6 in the spleen.

CYP SNPs and 1,000 Genomes

The current study identified 199 non-synonymous coding SNPs with frequencies greater than one percent (Table S2). Compared to the "Human Cytochrome P450 Allele Nomenclature Database" (<u>http://www.cypalleles.ki.se/</u>), we found several SNPs in 1,000 Genomes not related to alleles defined and named in the Database. To elucidate the

difference between the 'The Human Cytochrome P450 Allele Nomenclature' and 1000genome data regarding new SNPs, we examined CYP2A6 exemplary. Table 2 summarizes SNPs most likely to alter enzyme activity [25]. It displays five SNPs, which can lead to an altered enzyme activity with frequencies between 1.4 and 5.1 %. Only I471T (rs5031016) is also contained in the CYP nomenclature and reflects the CYP2A6*36 allele. New updates have to be done to map a comprehensive CYP SNP data source.

With the potential to alter drug metabolism, the 72 listed SNPs occurred in 24 CYPs. The most frequent SNPs were CYP 4A11 rs112743 (42.6%; highly expressed in kidney tissue), CYP 4F11 rs1060463 (49.5%; highly expressed in bronchus tissue) and CYP 2A7 rs3869579 (46.7%; highly

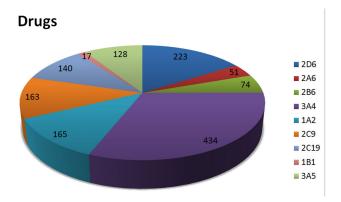


Figure 5. Number of drugs metabolized per CYP. The pie chart illustrates the number of drugs that can be metabolized by a specific CYP.

doi: 10.1371/journal.pone.0082562.g005

expressed in the pituitary gland but low in the prostate gland). Our study findings show that some CYPs are not only heterogeneously expressed, but also highly polymorphic.

Discussion

Genetic diversity and polymorphisms

Mutations in a CYP gene can lead to functional alterations, such as increased or decreased activity. If a mutant allele occurs at a frequency of at least one percent in a population, it is referred to as a pharmacogenetic polymorphism. Such polymorphisms can be discovered at the genotype level and/or the phenotype level based on altered function of the enzyme [30].

Individuals in a population can be stratified according to metabolic ratios of particular CYPs, which have great clinical relevance. For example, a CYP 2D6 poor metabolizer should not be administered codeine since the drug would have no effect. Conversely, a CYP 2D6 ultra-rapid metabolizer would likely suffer side effects from a normal dosage [31,32]. CYP 2D6 is a highly polymorphic CYP with at least 70 allelic variants [33] that can be categorized into four phenotypic classes. Overall CYP 2D6 expression in liver tissue is only approximately 2%, but hundreds of drugs are metabolized by this enzyme, including opiates, beta-blockers, anti-arrhythmics, tricyclic antidepressants, SSRIs, 5-HT3-antagonists and neuroleptics [34]. About 10% of the Caucasian population have difficulties in fully metabolizing these drugs [35], leading to harmful side effects [32,36]. Therefore, personalized prescriptions will become of great importance [37].

Personalized medicine

Since 2009, the Clinical Pharmacogenetics Implementation Consortium (CPIC) provides information on how genetic test results can be used to optimize drug therapy. The guidelines center on genes or on specific drugs. For some drugs, they also provide dosing guidelines for clinicians [38]. **Psychiatric drugs.** As most psychiatric drugs are metabolized by highly polymorphic CYP 2D6 and CYP 2C19, psychiatrists were first to propose the idea of CYP genotyping [39–41]. Three state hospitals in Kentucky recruited 4,532 psychiatric patients for genotyping of both CYPs with the help of DNA microarray technology.

Results from the current study were consistent with previous studies of allele frequency [35], demonstrating the importance of personalized prescription given that more than one tenth of patients are not likely to respond to standard treatment and suffer unwarranted toxicity. In the study performed by de Leon and colleagues, the dosage was adapted to the guidelines of Kirchheiner [15] for antipsychotics and antidepressants. The authors propose a numeric dosage adaptation system that reflects expression of CYP 2D6 and CYP 2C19.

Cardiovascular drugs. An important area of focus is stent implantation and/or inhibition of blood clots after an acute coronary syndrome (ACS) to prevent ischemic events. Therefore, antiplatelet agents are administered before and after percutaneous coronary intervention (PCI) to reduce the risk of ischemic events. Currently, the gold standard therapy is a combination of aspirin and clopidogrel [42,43]. Unfortunately, approximately 29% of people respond poorly to clopidogrel [44] and, therefore, have an increased risk for recurrent ischemic events after PCI [45]. Several different factors were discovered to contribute to the variability in clopidogrel response, including polymorphisms, impaired absorption or bioavailability, poor compliance and pre-existing conditions (increased body mass index, diabetes mellitus, ACS) [46]. In addition, clopidogrel is a prodrug that requires activation through the CYP system. The activated metabolite inhibits the ADP P2Y12 receptor [47]. Polymorphisms causing loss of function in the CYP system are associated with poor drug response. Most notably, the CYP 2C19*2 polymorphism was shown to lead to a 30% increased risk of major adverse cardiovascular events during treatment with clopidogrel [48-51]. Furthermore, the CYP 2B6*5 and P2Y12 polymorphisms are also associated with clopidogrel resistance [52]. In contrast, an enhanced response due to increased transcriptional activity occurs with the CYP 2C19*17 polymorphism, leading to increased risk of bleeding during clopidogrel therapy [53,54].

The CYP3A4*2 allele with a frequency of 2.7 % in Caucasian leads in vitro to reduced (six fold to nine fold) intrinsic clearance for nifedipine [55]. This could have a great influence on the tolerability of patients getting this dihydropyridine calcium channel blocker. Indications for nifedipine are widely distributed, e.g. Angina pectoris, Hypertonia, Achalasia and Raynaud's phenomenon, so the application is very common. An in vivo research regarding the alteration of nifedipine metabolism in CYP3A4*2 patients should be done, to possibly prevent toxic and/or increased side effects.

Previous findings described above, emphasize the importance of CYP polymorphisms and alternatively metabolized drugs in clinical practice. Prediction of CYP activity may be helpful to assess drug response. For instance, the (13)C-pantoprazole breath test, which measures CYP 2C19 activity, can detect clopidogrel resistance [56] and support use

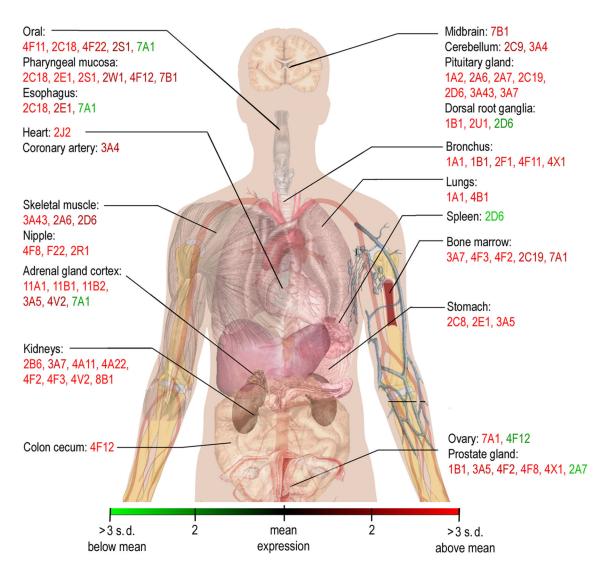


Figure 6. Body map of cytochrome P450 enzyme expression. A schematic map of the specific expression of different CYPs in human organs is presented. Expression values are relative to the mean expression in all organs. At least three-fold higher expression of a CYP in one organ is indicated by red text. At least three-fold lower expression is indicated by green text. A color spectrum for the expression values is illustrated in the provided scale. CYPs with average expression in one organ were not included.

doi: 10.1371/journal.pone.0082562.g006

of suitable drug alternatives like Ticagrelor (no activation required, metabolized via CYP 3A4).

Additional observed effects

Apart from altered drug metabolism, CYP polymorphisms were also potentially associated with neoplastic growth, adverse psychological behavior and other diseases. In women, polymorphisms in CYP 1A1 seemed to increase susceptibility to genital cancers [57,58]. Conversely, the CYP 2D6*4 polymorphism has been shown to have a protective effect against breast cancer [59]. Furthermore, 2C19*2, 2D6*4, 2D6*10 and 1A1*2A have been associated with increased risk of head and neck squamous cell carcinoma [60].

In addition to the role that CYP polymorphisms play in pathological processes and susceptibility to certain diseases, recent genome-wide association studies (GWASs) have demonstrated an association between increased coffee consumption and SNPs rs2472297-T (located between CYP1A1 and CYP1A2) and rs6968865 (next to aryl hydrocarbon receptor) [61]. Huo et al. (2012) determined that certain SNPs are associated with increased susceptibility to schizophrenia [62], while Peñas-Lledó and colleagues found a positive association between the extent of active CYP 2D6 and frequency of suicide attempts, providing evidence that CYP diversity may need to be accounted for in clinical practice [63].

Table 2. Differentially expressed CYPs and their SNP frequencies.

СҮР	ID	Mutation	Amino Acid	Global frequency (%)	CYP	ID	Mutation	Amino acid	Global frequency (%)
3A43	rs45450092	435G>T	M145I	1.5	2F1	rs144315434	1172T>C	L391P	5.4
	rs45621431	825G>A	M275I	2.3		rs146029724	1330A>C	M444L	6.5
	rs680055	1018C>G	P340A	13.4	2W1	rs61746347	557G>A	R186H	2.8
	rs78548296	389G>A	R130Q	1.0		rs117826462	547C>G	L183V	1.4
4A11	rs1126743	1374C>G	I458M	42.6	4F11	rs1060463	1271G>A	R424Q	49.5
	rs4926581	553G>T	V185F	28.1		rs148197835	538C>T	R180C	4.2
	rs61736429	1525C>T	L509F	2.1		rs57519667	436C>T	R146C	1.6
	rs62618709	553G>T	V185F	1.1	2C18	rs115091705	431G>A	R144H	1.7
4A22	rs112604161	181G>A	G61R	1.6		rs117111102	370C>T	R124W	1.4
	rs113777592	553G>T	V185F	29.6		rs2281891	1154C>T	T385M	19.3
	rs2056900	388G>A	G130S	29.9		rs41286880	1004G>A	R335Q	2.5
	rs4926600	1525C>T	L509F	12.9		rs79500998	1324C>T	R442C	1.0
	rs61507155	311A>T	Y104F	6.6	7A1	rs8192875	1039G>A	D347N	1.6
	rs61736431	1154C>T	P385L	1.2	4F12	rs16995378	47C>T	T16M	7.6
4B1	rs12094024	986A>C	Y329S	2.2		rs57578760	808G>C	V270L	3.7
	rs2297809	1123C>T	R375C	18.3		rs76142062	88C>A	L30I	3.4
	rs4646487	517C>T	R173W	16.8	11B1	rs11775687	562C>T	P188S	5.7
	rs59694031	1109G>C	C370S	4.0		rs9657020	593C>T	T198M	12.7
4F2	rs2074900	515C>T	Thr172I	25.4	4F3	rs118159249	1420G>A	A474T	1.1
	rs2108622	1297G>A	V433M	20.9	2A7	rs111390860	988C>T	R330W	1.0
	rs3093153	554G>T	G185V	3.7		rs184466431	1301G>T	R434L	1.2
	rs3093200	1555C>A	L519M	8.4		rs3869579	778C>T	R260C	46.7
5A1	rs13306050	1372C>T	R458C	3.3		rs60711313	1259T>C	I420T	3.2
	rs13306052	679GA	V227M	1.4		rs75152309	1106A>T	K369M	6.6
	rs6952940	544C>T	P182S	2.4		rs78754793	244G>C	A82P	2.4
2A6	rs5031017	1436G>T	G479V	1.4	2S1	rs34971233	1397C>T	P466L	1.1
	rs5031016	1412T>C	I471T	5.1	2E1	rs6413419	535G>A	V179I	7.2
	rs28399499	983T>C	1328T	2.3		rs28969387	1370A>T	H457L	6.3
	rs8192709	64C>T	R22C	4.5	2C19	rs17884712	431G>A	R144H	1.4
	rs28399499	383T>C	I128T	2.3	8A1	rs5626	706C>T	R236C	3.7
2C8	rs11572103	805A>T	1269F	16.4	2D6	rs2982054	986G>A	R329H	31.2
	rs1058930	792C>G	I264M	4.1		rs1058172	941G>A	R314H	7.9
	rs11572103	805A>T	1269F	16.4		rs59421388	859G>A	V287M	5.3
2C9	rs28371686	1080C>G	D360E	2.3		rs1065852	100C>T	P34S	25.9
	rs28371685	1003C>T	R335W	2.0	1A1	rs4646422	134GA	G45D	6.7
	rs2256871	752A>G	H251R	4.0		rs17861094	233T>C	I78Thr	8.3

Data was extracted from the 1,000 Genomes Project site (<u>http://www.1000genomes.org/</u>) [23]. A more detailed table can be found in the supplemental material. Possibly CYP-activity damaging SNPs are included in Table S2.

doi: 10.1371/journal.pone.0082562.t002

Diversity of expression in human tissues

Variable expression of functionally distinct CYP isoforms across different tissue types indicates that certain isoforms play specific roles in a tissue-dependent manner. Figure 6 provides an illustrative overview of CYP expression in the human body. Such knowledge may be useful for development of new prodrugs activated by a specific CYP highly expressed in the preferentially targeted tissue, ultimately leading to increased bioavailability at the target site and reduced side effects. On the other hand, variable expression of CYPs in different tissues may adversely affect drug efficacy in some tissues. Such a case could occur if drugs undergo an inactivation through a higher expressed CYP in their target tissues. Regardless, further clinical investigation is required.

Even polymorphic CYP isoforms show a heterogeneous tissue distribution. In particular, CYP 1A2, 2C19 and 2D6 are highly expressed in the pituitary gland. Furthermore, highest expression of CYP 2C9 was detected in the cerebellum, while greatest expression of 2D6 and 2B6 were found in skeletal muscle and kidneys. The influence of mutations in CYPs in particular organs remains to be determined and requires further investigation.

Differential distribution of CYPs may have an influence on specific side effects of drugs. For example, cyclophosphamide (CPA) therapy can lead to development of hyponatremia. CPA is a prodrug converted by CYP 2B6 into the active form [64].

The hyponatremia is the result of increased expression of aquaporins 1 and 7, which is induced by CPA [65]. CYP 2B6 has high expression in kidneys, indicating that a higher level of active CPA is likely to occur in the kidneys and lead to the undesirable side effect.

Conclusions

In summary, the current study identified four major CYPs (1A2, 2D6, 2C9 and 2C19) and 34 polymorphic alleles with a significant impact on the drug metabolism in the Caucasian population. Once genomic testing becomes part of routine analysis, this data enables prediction of complications in drug therapy and development of a personalized treatment regimen, where drug dosages are based on an individual's specific CYP profile [6]. Ultimately, this approach may prevent treatment failures and avoid unnecessary side effects. Another interesting field could be the consideration of CYP polymorphisms in clinical trials. Potentially, it would decrease the failures if information of potential polymorphisms in different ethnic groups was included. Findings from the current study will be included in the SuperCYP database.

With the aim of assessing the effects of CYP polymorphisms on chemotherapy and establishing a cost efficient method to detect relevant CYP polymorphisms, a retrospective study in leukemia cells from pediatric patients is currently under way [66].

References

- Zhou SF, Liu JP, Chowbay B (2009) Polymorphism of human cytochrome P450 enzymes and its clinical impact. Drug Metab Rev 41: 89-295. doi:10.1080/03602530902843483. PubMed: 19514967.
- Raether W, Seidenath H, Loewe H (1978) Action of p-(4-amidinophenoxy)-benzaldehyde-p-amidino-phenylhydrazone dihydrochloride on Leishmania donovani infections in the golden hamster. Ann Trop Med Parasitol 72: 543-547. PubMed: 216324.
- McGraw J, Waller D (2012) Cytochrome P450 variations in different ethnic populations. Expert Opin Drug Metab Toxicol 8: 371-382. doi: 10.1517/17425255.2012.657626. PubMed: 22288606.
- Deenen MJ, Cats A, Beijnen JH, Schellens JH (2011) pharmacogenetic variability in drug transport and phase I anticancer drug metabolism part 2. Oncologist 16: 820-834. doi:10.1634/theoncologist.2010-0259. PubMed: 21632461.
- Grover S, Kukreti R (2012) Functional Genetic Polymorphisms from Phase-II Drug Metabolizing. Enzymes - CNS Neuroscience Ther 18: 705-706. doi:10.1111/j.1755-5949.2012.00343.x.
- Preissner S, Kroll K, Dunkel M, Senger C, Goldsobel G et al. (2010) SuperCYP: a comprehensive database on Cytochrome P450 enzymes including a tool for analysis of CYP-drug interactions. Nucleic Acids Res 38: D237-D243. doi:10.1093/nar/gkp970. PubMed: 19934256.
- McLean KJ, Sabri M, Marshall KR, Lawson RJ, Lewis DG et al. (2005) Biodiversity of cytochrome P450 redox systems. Biochem Soc Trans 33: 796-801. doi:10.1042/BST0330796. PubMed: 16042601.
- Guengerich FP (2008) Cytochrome p450 and chemical toxicology. Chem Res Toxicol 21: 70-83. doi:10.1021/tx700079z. PubMed: 18052394.
- Ingelman-Sundberg M (2005) The human genome project and novel aspects of cytochrome P450 research. Toxicol Appl Pharmacol 207: 52-56. doi:10.1016/j.taap.2005.01.030. PubMed: 15993453.
- Flockhart DA, Oesterheld JR (2000) Cytochrome P450-mediated drug interactions. Child Adolesc Psychiatr Clin N Am 9: 43-76. PubMed: 10674190.
- Berka K, Hendrychová T, Anzenbacher P, Otyepka M (2011) Membrane position of ibuprofen agrees with suggested access path entrance to cytochrome P450 2C9 active site. J Phys Chem A 115: 11248-11255. doi:10.1021/jp204488j. PubMed: 21744854.

Supporting Information

Table S1. Extended table of CYP SNP frequencies in Caucasians and other ethnics. The table includes the content of Table 1 and further information (other ethnics and nucleotide changes) that have been extracted by text mining. (XLSX)

Table S2. Extended table of differentially expressed CYPs and their SNP frequencies. The extended table lists possibly CYP-activity damaging SNPs. CYP SNPs with frequencies (%) greater than one percent were included. The data was extracted from the 1,000 Genomes Project site (<u>http://www.1000genomes.org/</u>).

(XLSX)

Author Contributions

Conceived and designed the experiments: RP. Performed the experiments: SCP SP MFH AG. Analyzed the data: SCP SP MFH RP MD. Contributed reagents/materials/analysis tools: MD AG. Wrote the manuscript: SP RP SCP MFH.

- Culhane AC, Schwarzl T, Sultana R, Picard KC, Picard SC et al. (2010) GeneSigDB--a curated database of gene expression signatures. Nucleic Acids Res 38: D716-D725. doi:10.1093/nar/gkp1015. PubMed: 19934259.
- de Leon J, Susce MT, Johnson M, Hardin M, Maw L et al. (2009) DNA microarray technology in the clinical environment: the AmpliChip CYP450 test for CYP2D6 and CYP2C19 genotyping. CNS Spectr 14: 19-34. PubMed: 19169185.
- Patterson LH, McKeown SR, Robson T, Gallagher R, Raleigh SM et al. (1999) Antitumour prodrug development using cytochrome P450 (CYP) mediated activation. Anticancer Drug Des 14: 473-486. PubMed: 10834269.
- Kirchheiner J, Nickchen K, Bauer M, Wong ML, Licinio J et al. (2004) Pharmacogenetics of antidepressants and antipsychotics: the contribution of allelic variations to the phenotype of drug response. Mol Psychiatry 9: 442-473. doi:10.1038/sj.mp.4001494. PubMed: 15037866.
- 16. Albert A (1958) Chemical aspects of selective toxicity. Nature 182: 421-422. doi:10.1038/182421a0. PubMed: 13577867.
- Tanabe K, Ishizaki J, Ando Y, Ito T, Nishimoto S (2012) Reductive activation of 5-fluorodeoxyuridine prodrug possessing azide methyl group by hypoxic X-irradiation. Bioorg Med Chem Lett 22: 1682-1685. doi:10.1016/j.bmcl.2011.12.106. PubMed: 22248856.
- Guengerich FP, Shimada T (1998) Activation of procarcinogens by human cytochrome P450 enzymes. Mutat Res 400: 201-213. doi: 10.1016/S0027-5107(98)00037-2. PubMed: 9685642.
- Dormán G, Prestwich GD (2000) Using photolabile ligands in drug discovery and development. Trends Biotechnol 18: 64-77. doi:10.1016/ S0167-7799(99)01402-X. PubMed: 10652511.
- Simplício AL, Clancy JM, Gilmer JF (2007) Beta-aminoketones as prodrugs with pH-controlled activation. Int J Pharm 336: 208-214. doi: 10.1016/j.ijpharm.2006.11.055. PubMed: 17197138.
- Topf N, Worgall S, Hackett NR, Crystal RG (1998) Regional 'pro-drug' gene therapy: intravenous administration of an adenoviral vector expressing the E. coli cytosine deaminase gene and systemic administration of 5-fluorocytosine suppresses growth of hepatic metastasis of colon carcinoma. Gene Ther 5: 507-513. doi:10.1038/ sj.gt.3300611. PubMed: 9614575.

- Huttunen KM, Mähönen N, Raunio H, Rautio J (2008) Cytochrome P450-activated prodrugs: targeted drug delivery. Curr Med Chem 15: 2346-2365. doi:10.2174/092986708785909120. PubMed: 18855665.
- Genomes Project C (2010) A map of human genome variation from population-scale sequencing. Nature 467: 1061-1073. doi:10.1038/ nature09534. PubMed: 20981092.
- 24. Sim SC, Ingelman-Sundberg M (2010) The Human Cytochrome P450 (CYP) Allele Nomenclature website: a peer-reviewed database of CYP variants and their associated effects. Hum Genomics 4: 278-281. doi: 10.1186/1479-7364-4-4-278. PubMed: 20511141.
- Ramensky V, Bork P, Sunyaev S (2002) Human non-synonymous SNPs: server and survey. Nucleic Acids Res 30: 3894-3900. doi: 10.1093/nar/gkf493. PubMed: 12202775.
- Worth CL, Preissner R, Blundell TL (2011) SDM--a server for predicting effects of mutations on protein stability and malfunction. Nucleic Acids Res 39: W215-W222. doi:10.1093/nar/gkr363. PubMed: 21593128.
- Roth RB, Hevezi P, Lee J, Willhite D, Lechner SM et al. (2006) Gene expression analyses reveal molecular relationships among 20 regions of the human CNS. Neurogenetics 7: 67-80. doi:10.1007/ s10048-006-0032-6. PubMed: 16572319.
- Sturn A, Quackenbush J, Trajanoski Z (2002) Genesis: cluster analysis of microarray data. Bioinformatics 18: 207-208. doi:10.1093/ bioinformatics/18.1.207. PubMed: 11836235.
- Nishimura M, Yaguti H, Yoshitsugu H, Naito S, Satoh T (2003) Tissue distribution of mRNA expression of human cytochrome P450 isoforms assessed by high-sensitivity real-time reverse transcription PCR. Yakugaku Zasshi 123: 369-375. doi:10.1248/yakushi.123.369. PubMed: 12772594.
- Hasler JA (1999) Pharmacogenetics of cytochromes P450. Mol Aspects Med 20: 12-25. 10575648.
- Johnson M, Markham-Abedi C, Susce MT, Murray-Carmichael E, McCollum S et al. (2006) A poor metabolizer for cytochromes P450 2D6 and 2C19: a case report on antidepressant treatment. CNS Spectr 11: 757-760. PubMed: 17008819.
- Laine K, Kytölä J, Bertilsson L (2004) Severe adverse effects in a newborn with two defective CYP2D6 alleles after exposure to paroxetine during late pregnancy. Ther Drug Monit 26: 685-687. doi: 10.1097/00007691-200412000-00016. PubMed: 15570195.
- Ingelman-Sundberg M, Sim SC, Gomez A, Rodriguez-Antona C (2007) Influence of cytochrome P450 polymorphisms on drug therapies: pharmacogenetic, pharmacoepigenetic and clinical aspects. Pharmacol Ther 116: 496-526. doi:10.1016/j.pharmthera.2007.09.004. PubMed: 18001838.
- de Leon J, Armstrong SC, Cozza KL (2006) Clinical guidelines for psychiatrists for the use of pharmacogenetic testing for CYP450 2D6 and CYP450 2C19. Psychosomatics 47: 75-85. doi:10.1176/appi.psy. 47.1.75. PubMed: 16384813.
- Bradford LD (2002) CYP2D6 allele frequency in European Caucasians, Asians, Africans and their descendants. Pharmacogenomics 3: 229-243. doi:10.1517/14622416.3.2.229. PubMed: 11972444.
- Papp-Jámbor C, Jaschinski U, Forst H (2002) Cytochrome P450 enzymes and their role in drug interactions. Anaesthesist 51: 2-15. doi: 10.1007/s101-002-8365-5. PubMed: 11963295.
- de Leon J (2009) The future (or lack of future) of personalized prescription in psychiatry. Pharmacol Res 59: 81-89. doi:10.1016/j.phrs. 2008.10.002. PubMed: 18996200.
- Relling MV, Klein TE (2011) CPIC: Clinical Pharmacogenetics Implementation Consortium of the Pharmacogenomics Research Network. Clin Pharmacol Ther 89: 464-467. doi:10.1038/clpt.2010.279. PubMed: 21270786.
- Roots I, Laschinski G, Arjomand-Nahad F, Kirchheiner J, Schwarz D et al. (2007) Genotype and phenotype relationship in drug metabolism. Ernst Schering Res Found Workshop. pp. 81-100. PubMed: 17117716.
- Seeringer A, Kirchheiner J (2008) Pharmacogenetics-guided dose modifications of antidepressants. Clin Lab Med 28: 619-626. doi: 10.1016/j.cll.2008.05.006. PubMed: 19059066.
- Kirchheiner J, Brøsen K, Dahl ML, Gram LF, Kasper S et al. (2001) CYP2D6 and CYP2C19 genotype-based dose recommendations for antidepressants: a first step towards subpopulation-specific dosages. Acta Psychiatr Scand 104: 173-192. doi:10.1034/j. 1600-0447.2001.00299.x. PubMed: 11531654.
- 42. Mehta SR, Yusuf S, Peters RJ, Bertrand ME, Lewis BS et al. (2001) Effects of pretreatment with clopidogrel and aspirin followed by longterm therapy in patients undergoing percutaneous coronary intervention: the PCI-CURE study. Lancet 358: 527-533. doi:10.1016/ S0140-6736(01)05701-4. PubMed: 11520521.
- 43. Steinhubl SR, Berger PB, Mann JT 3rd, Fry ET, DeLago A et al. (2002) Early and sustained dual oral antiplatelet therapy following percutaneous coronary intervention: a randomized controlled trial.

JAMA 288: 2411-2420. doi:10.1001/jama.288.19.2411. PubMed: 12435254.

- Dupont AG, Gabriel DA, Cohen MG (2009) Antiplatelet therapies and the role of antiplatelet resistance in acute coronary syndrome. Thromb Res 124: 6-13. doi:10.1016/j.thromres.2009.08.005. PubMed: 19324398.
- 45. Matetzky S, Shenkman B, Guetta V, Shechter M, Beinart R et al. (2004) Clopidogrel resistance is associated with increased risk of recurrent atherothrombotic events in patients with acute myocardial infarction. Circulation 109: 3171-3175. doi:10.1161/01.CIR. 0000130846.46168.03. PubMed: 15184279.
- Angiolillo DJ, Fernandez-Ortiz A, Bernardo E, Alfonso F, Macaya C et al. (2007) Variability in individual responsiveness to clopidogrel: clinical implications, management, and future perspectives. J Am Coll Cardiol 49: 1505-1516. doi:10.1016/j.jacc.2006.11.044. PubMed: 17418288.
- Savi P, Herbert JM (2005) Clopidogrel and ticlopidine: P2Y12 adenosine diphosphate-receptor antagonists for the prevention of atherothrombosis. Semin Thromb Hemost 31: 174-183. doi:10.1055/ s-2005-869523. PubMed: 15852221.
- Hulot JS, Collet JP, Silvain J, Pena A, Bellemain-Appaix A et al. (2010) Cardiovascular risk in clopidogrel-treated patients according to cytochrome P450 2C19*2 loss-of-function allele or proton pump inhibitor coadministration: a systematic meta-analysis. J Am Coll Cardiol 56: 134-143. doi:10.1016/j.jacc.2009.12.071. PubMed: 20620727.
- Sofi F, Giusti B, Marcucci R, Gori AM, Abbate R et al. (2011) Cytochrome P450 2C19*2 polymorphism and cardiovascular recurrences in patients taking clopidogrel: a meta-analysis. Pharmacogenomics J 11: 199-206. doi:10.1038/tpj.2010.21. PubMed: 20351750.
- Pettersen AA, Arnesen H, Opstad TB, Seljeflot I (2011) The influence of CYP 2C19*2 polymorphism on platelet function testing during single antiplatelet treatment with clopidogrel. Thromb J 9: 4. doi: 10.1186/1477-9560-9-4. PubMed: 21426546.
- Mega JL, Simon T, Collet JP, Anderson JL, Antman EM et al. (2010) Reduced-function CYP2C19 genotype and risk of adverse clinical outcomes among patients treated with clopidogrel predominantly for PCI: a meta-analysis. JAMA 304: 1821-1830. doi:10.1001/jama. 2010.1543. PubMed: 20978260.
- Kassimis G, Davlouros P, Xanthopoulou I, Stavrou EF, Athanassiadou A et al. (2012) CYP2C19⁺2 and other genetic variants affecting platelet response to clopidogrel in patients undergoing percutaneous coronary intervention. Thromb Res 129: 441-446. doi:10.1016/j.thromres. 2011.07.022. PubMed: 21831410.
- 53. Sibbing D, Koch W, Gebhard D, Schuster T, Braun S et al. (2010) Cytochrome 2C19*17 allelic variant, platelet aggregation, bleeding events, and stent thrombosis in clopidogrel-treated patients with coronary stent placement. Circulation 121: 512-518. doi:10.1161/ CIRCULATIONAHA.109.885194. PubMed: 20083681.
- 54. Dai ZL, Chen H, Wu XY (2012) Relationship between cytochrome P450 2C19*17 genotype distribution, platelet aggregation and bleeding risk in patients with blood stasis syndrome of coronary artery disease treated with clopidogrel. Zhong Xi Yi Jie He Xue Bao 10: 647-654.
- 55. Sata F, Sapone A, Elizondo G, Stocker P, Miller VP et al. (2000) CYP3A4 allelic variants with amino acid substitutions in exons 7 and 12: evidence for an allelic variant with altered catalytic activity. Clin Pharmacol Ther 67: 48-56. doi:10.1067/mcp.2000.104391. PubMed: 10668853.
- Tazaki J, Jinnai T, Tada T, Kato Y, Makiyama T et al. (2012) Prediction of clopidogrel low responders by a rapid CYP2C19 activity test. J Atheroscler Thromb 19: 186-193. doi:10.5551/jat.10009. PubMed: 22166969.
- Sergentanis TN, Economopoulos KP, Choussein S, Vlahos NF (2012) Cytochrome P450 1A1 (CYP1A1) gene polymorphisms and ovarian cancer risk: a meta-analysis. Mol Biol Rep, 39: 9921–30. PubMed: 22733497.
- Mrozikiewicz PM, Grześkowiak E, Seremak-Mrozikiewicz A, Bogacz A, Barlik M et al. (2011) Importance of CYP1A1 polymorphism and its transcriptional regulation in ovarian and endometrial cancer. Ginekol Pol 82: 925-932. PubMed: 22384629.
- Fernández-Santander A, del Saz Sánchez M, Tejerina Gómez A, Moya Bandres F (2012) CYP2D6*4 allele and breast cancer risk: is there any association? Clin Transl Oncol 14: 157-159. doi:10.1007/ s12094-012-0776-4. PubMed: 22301406.
- Ruwali M, Parmar D (2010) Association of functionally important polymorphisms in cytochrome P450s with squamous cell carcinoma of head and neck. Indian J Exp Biol 48: 651-665. PubMed: 20929050.
- Sulem P, Gudbjartsson DF, Geller F, Prokopenko I, Feenstra B et al. (2011) Sequence variants at CYP1A1-CYP1A2 and AHR associate

with coffee consumption. Hum Mol Genet 20: 2071-2077. doi: 10.1093/hmg/ddr086. PubMed: 21357676.

- 62. Huo R, Tang K, Wei Z, Shen L, Xiong Y et al. (2012) Genetic polymorphisms in CYP2E1: association with schizophrenia susceptibility and risperidone response in the Chinese Han population. PLOS ONE 7: e34809. doi:10.1371/journal.pone.0034809. PubMed: 22606226.
- Peñas-Lledó EM, Blasco-Fontecilla H, Dorado P, Vaquero-Lorenzo C, Baca-García E et al. (2012) CYP2D6 and the severity of suicide attempts. Pharmacogenomics 13: 179-184. doi:10.2217/pgs.11.146. PubMed: 22141351.
- Jounaidi Y, Chen CS, Veal GJ, Waxman DJ (2006) Enhanced antitumor activity of P450 prodrug-based gene therapy using the low

Km cyclophosphamide 4-hydroxylase P450 2B11. Mol Cancer Ther 5: 541-555. doi:10.1158/1535-7163.MCT-05-0321. PubMed: 16546968.

- Kim S, Jo CH, Park JS, Han HJ, Kim GH (2011) The role of proximal nephron in cyclophosphamide-induced water retention: preliminary data. Electrolyte. Blood Press 9: 7-15.
- 66. Shalapour S, Hof J, Kirschner-Schwabe R, Bastian L, Eckert C et al. (2011) High VLA-4 expression is associated with adverse outcome and distinct gene expression changes in childhood B-cell precursor acute lymphoblastic leukemia at first relapse. Haematologica 96: 1627-1635. doi:10.3324/haematol.2011.047993. PubMed: 21828124.