

MECHANISMS OF VARIABILITY IN CYP2D6 METABOLISM:  
THE CONTRIBUTIONS OF POLYMORPHISMS, COPY NUMBER VARIATIONS  
AND microRNA

Anuradha Ramamoorthy

Submitted to the faculty of the University Graduate School  
in partial fulfillment of the requirements  
for the degree  
Doctor of Philosophy  
in the Department of Medical and Molecular Genetics,  
Indiana University

August 2010

Accepted by the Faculty of Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

---

David A. Flockhart, M.D., Ph.D., Chair

---

Todd C. Skaar, Ph.D.

Doctoral Committee

---

Tatiana Foroud, Ph.D.

April 7, 2010

---

Brittney-Shea Herbert, Ph.D.

---

Harikrishna Nakshatri, Ph.D.

## ACKNOWLEDGEMENTS

My Ph.D. journey would not have been possible without the support from both my teachers and family. I would first like to thank my mentor, Dr. David Flockhart, for giving me the opportunity to pursue my Ph.D. under his guidance. I have benefited immensely from his vast knowledge of pharmacogenetics. He has been a constant source of inspiration and motivation. I am also greatly indebted to my research mentor, Dr. Todd Skaar. His scientific ideas, guidance and support have been instrumental to the successful completion of my research. He ensured that I stayed on course and focused on my objective.

I would like to thank the members of my research committee, Dr. Brittney-Shea Herbert, Dr. Tatiana Foroud, and Dr. Harikrishna Nakshatri for their critical evaluations, insights and guidance. I am thankful to the members of Dr. Skaar's lab for all of their help. I am also thankful to the members of the Division of Clinical Pharmacology and the Department of Medical and Molecular Genetics for the assistance that they have provided to me throughout my time at Indiana University.

Lastly, I would like to thank my family for their constant support and encouragement in the pursuit of my personal and professional objectives. I am very thankful to my parents and brother for their unconditional love and support even from such a long distance. I am also thankful to my other set of parents for their constant words of encouragement and care packages. Finally, this work would not have been possible without the endless love, support and encouragement from my husband.

## ABSTRACT

### Mechanisms of variability in CYP2D6 metabolism: the contributions of polymorphisms, copy number variations and microRNA

Anuradha Ramamoorthy

Cytochrome P450 2D6 (CYP2D6) is an important drug metabolizing enzyme that is involved in the metabolism of 20-25% of commonly prescribed drugs. There is interindividual variability in CYP2D6 enzyme activity and this leads to compromised metabolism of many drugs. Genetic and environmental factors explain only a part of the interindividual variability; the other factors that contribute to this variability are largely unknown. Hence, it becomes important to study CYP2D6 to understand the endogenous and exogenous factors that control its activity.

The *specific objective* of this research was to determine the contribution of genetic and epigenetic factors in the regulation of CYP2D6 expression and activity. The *specific aims* were: (1) to identify the common *CYP2D6* polymorphisms in Vietnamese and Filipino women with breast cancer and evaluate its association with plasma concentrations of endoxifen (an active metabolite of the breast cancer therapeutic drug, tamoxifen); (2) to identify the *CYP2D6* copy number variations (CNVs) in these women and evaluate their association with endoxifen concentration; and (3) to identify microRNAs (miRNAs) that regulate the expression of CYP2D6 directly or indirectly.

The results of this study indicated that: (1) in Vietnamese and Filipino women, the reduced function allele *CYP2D6\*10* was frequent (~55%) and it was significantly associated with reduced endoxifen concentration; (2) in these women, only 39% carried two copies of the *CYP2D6* gene, the rest had a genomic imbalance for *CYP2D6*, primarily involving the *CYP2D6>(\*36)<sub>n</sub>-\*10* allele. However, carrying multiple copies of *CYP2D6\*36* allele did not significantly affect CYP2D6 activity, suggesting that multiple copies of a gene does not always translate to additive effects; and (3) microRNAs were identified to target HNF4A, a transcriptional factor that regulates CYP2D6 expression. These miRNAs are likely to play an important role in the indirect regulation of CYP2D6.

Taken together, these results emphasize on the role of polymorphisms, CNVs and possibly miRNAs in the regulation of CYP2D6. These clinically important biomarkers will help to improve the efficacy and reduce the side effects of many CYP2D6 substrate drugs and thus contribute to *personalization of drug therapy*.

David A. Flockhart, M.D., Ph.D., Chair

## TABLE OF CONTENTS

List of Tables.....	ix
List of Figures.....	xi
Abbreviations.....	xiv
Chapter One: Introduction and Literature Review	
Pharmacogenetics of interindividual variability in drug response.....	1
The cytochrome P450 drug metabolizing enzyme family.....	1
CYP2D6 and its role in drug metabolism.....	3
<i>CYP2D6</i> gene structure.....	6
CYP2D6 expression.....	8
Transcriptional regulation of CYP2D6 expression by HNF4A.....	9
CYP2D6 mediated drug metabolism.....	10
Tamoxifen as an endocrine therapy for breast cancer.....	11
Metabolism of tamoxifen by CYP2D6.....	13
Interindividual variability in CYP2D6 enzyme activity and CYP2D6 mediated drug metabolism.....	20
Causes of variability in CYP2D6 enzyme activity.....	21
Environmental factors.....	21
Genetic factors.....	22
<i>CYP2D6</i> gene polymorphisms.....	22
CYP2D6 copy number variations.....	28
Epigenetic factors.....	34
Introduction to microRNAs.....	36

MicroRNA gene processing.....	36
Functions of microRNAs.....	38
Pharmacological relevance of microRNAs.....	39
Drug induced changes in microRNAs.....	41
Polymorphisms affect mRNA-miRNA interactions.....	41
Bioinformatic algorithms predict mRNA-microRNAs interactions.....	42
Summary.....	45
Research aims and significance.....	45
<b>Chapter Two: Cytochrome P450 2D6 Variants Predict Endoxifen Concentration in Vietnamese and Filipino Breast Cancer Patients</b>	
Abstract.....	49
Introduction.....	50
Results.....	52
Discussion.....	63
Materials and Methods.....	66
<b>Chapter Three: CYP2D6 Copy Number Variations in Vietnamese and Filipino Breast Cancer Patients and the Effect of CYP2D6*36 Allele on Endoxifen Concentration.</b>	
Abstract.....	69
Introduction.....	70
Results.....	74
Discussion.....	84

Materials and Methods.....	89
Chapter Four: In silico and In vitro Identification of microRNAs that Post-transcriptionally Regulate HNF4A Expression.	
Abstract.....	91
Introduction.....	92
Results.....	95
Discussion.....	114
Materials and Methods.....	119
Chapter Five: Conclusion and Future Directions	
Effect of <i>CYP2D6</i> gene polymorphisms on tamoxifen metabolism.....	128
Effect of <i>CYP2D6</i> gene copy number variations on enzyme activity.....	135
Direct and indirect regulation of <i>CYP2D6</i> by microRNAs.....	139
Clinical significance.....	144
Future directions.....	146
Summary.....	148
References.....	150
Curriculum Vitae	



## LIST OF TABLES

Table 1.1.	Major CYP2D6 drug substrates and inhibitors.....	4
Table 1.2.	Summary of studies reporting an association between <i>CYP2D6</i> genotype and response to tamoxifen therapy.....	18
Table 1.3.	Summary of studies reporting either a lack of association or better clinical outcome due to <i>CYP2D6</i> gene variants after tamoxifen therapy.....	19
Table 1.4.	Consequences of selected <i>CYP2D6</i> gene polymorphisms.....	25
Table 1.5.	CYP2D6 enzyme activity scores based on <i>CYP2D6</i> allele and copy number variations.....	27
Table 1.6.	Consequences of <i>CYP2D6</i> gene copy number variations.....	30
Table 2.1.	Patient characteristics.....	53
Table 2.2.	<i>CYP2D6</i> allele frequencies in the Vietnamese and Filipino breast cancer patients.....	55
Table 2.3.	<i>CYP2D6</i> genotypes in Vietnamese and Filipino breast cancer patients.....	56
Table 2.4.	Plasma concentrations of tamoxifen and its metabolites in subjects with different CYP2D6 score.....	58
Table 3.1.	<i>CYP2D6</i> allelic frequencies in the different cohorts.....	77
Table 3.2.	Frequencies of <i>CYP2D6</i> diplotypes in the different cohorts.....	78
Table 3.3.	Plasma concentrations of tamoxifen and its metabolites in subjects with different <i>CYP2D6</i> diplotypes involved in <i>CYP2D6</i> (*36) <sub>n</sub> copy number variations.....	80

Table 4.1.	MicroRNAs that are predicted to target the 3'-UTR of CYP2D6 and HNF4A.....	96
Table 4.2.	HNF4A 3'-UTR SNPs predicted to destroy or create a miRNA target site.....	109
Table 4.3.	Minor allele frequency of HNF4A 3'-UTR SNP, rs11574744.....	114

## LIST OF FIGURES

Figure 1.1.	The liver expressed cytochrome P450s and their contribution to drug metabolism.....	2
Figure 1.2.	CYP2D6 genotype-phenotype correlation and nortriptyline dose adjustment for the different metabolizer groups.....	5
Figure 1.3.	Human <i>CYP2D</i> gene family and common variants in the <i>CYP2D6</i> gene.....	7
Figure 1.4.	HNF4A is a master transcriptional regulator of drug metabolizing enzymes.....	10
Figure 1.5.	Mechanism of action of tamoxifen in breast cancer cells.....	12
Figure 1.6.	Metabolism of tamoxifen by Phase I and II drug metabolizing enzymes..	14
Figure 1.7.	<i>CYP2D6</i> gene polymorphisms and concomitant use of CYP2D6 inhibitors are associated with plasma endoxifen concentration .....	15
Figure 1.8.	<i>CYP2D6</i> genotype is associated with clinical outcomes.....	16
Figure 1.9.	CYP2D6 mediated quantitative influences on pharmacokinetics of selected antidepressant drugs.....	21
Figure 1.10.	Polymorphisms in selected <i>CYP2D6</i> alleles.....	24
Figure 1.11.	Frequencies of selected <i>CYP2D6</i> alleles in different ethnic populations.....	26
Figure 1.12.	<i>CYP2D6</i> gene copy number variations in the Japanese and Chinese Coriell diversity panel.....	29
Figure 1.13:	Comparison of estimated CYP2D6 gene copy numbers using four	

	copy number assays.....	33
Figure 1.14.	Lack of association between <i>CYP2D6</i> mRNA and protein expression in the human livers.....	35
Figure 1.15.	Gene processing and mechanisms of action of microRNAs.....	37
Figure 2.1.	Association between <i>CYP2D6</i> gene score and tamoxifen and its metabolites concentrations.....	60
Figure 3.1.	Frequency of <i>CYP2D6</i> gene copy number variations in Vietnamese and Filipino breast cancer patients.....	75
Figure 3.2.	Correlation between <i>CYP2D6</i> allele copy numbers with tamoxifen metabolism in the combined cohort.....	83
Figure 4.1.	Schematic representation of the pIS-2D6 and pIS-HNF4A vectors.....	97
Figure 4.2.	Regulatory elements in the 3'-UTR of <i>CYP2D6</i> and <i>HNF4A</i> .....	98
Figure 4.3.	Regulation of pIS-2D6 and pIS-HNF4A by miRNAs.....	99
Figure 4.4.	Lack of regulation of pIS-2D6 by miRNAs.....	100
Figure 4.5.	Regulation of pIS-HNF4A by miRNAs.....	101
Figure 4.6.	MicroRNAs do not regulate the expression of <i>HNF4A</i> mRNA expression.....	103
Figure 4.7.	Regulation of <i>HNF4A</i> protein by miRNAs.....	104
Figure 4.8.	Regulation of <i>PXR</i> mRNA by miRNAs.....	105
Figure 4.9.	Transfection efficiency of synthetic miRNAs.....	106
Figure 4.10.	Expression of mature miRNAs in human cell lines and hepatocytes.....	107

Figure 4.11.	HNF4A 3'-UTR SNP predicted to destroy a miRNA binding site.....	110
Figure 4.12.	In vitro validation of SNP predictions.....	111
Figure 4.13.	The SNP rs11574744 is not predicted to affect the local secondary structure of HNF4A mRNA.....	113
Figure 5.1.	<i>CYP2D6</i> allele frequencies in the current Vietnamese and Filipino study are comparable to other Asian populations but not to Caucasian or African populations.....	131
Figure 5.2.	Association of <i>CYP2D6</i> genetic variants with tamoxifen metabolism in the combined Vietnamese and Filipino cohort based on regression analysis.....	133
Figure 5.3.	Schematic representation of CNVs involving some of the <i>CYP2D6</i> alleles.....	137
Figure 5.4.	Schematic representation of possible indirect regulation of <i>CYP2D6</i> by microRNAs targeting HNF4A.....	142
Figure 5.5.	Genetic and epigenetic factors regulate the expression and activity of <i>CYP2D6</i> .....	149

## ABBREVIATIONS

°C	degree Celsius
β-actin	beta-actin
μg	microgram
μl	microliter
3'-UTR	3'-untranslated region
4-OH-tam	4-hydroxytamoxifen
5'-UTR	5'-untranslated region
AI	aromatase inhibitor
bp	base pairs
cDNA	complementary DNA
cel-miR	<i>Caenorhabditis elegans</i> microRNA
CNV	copy number variation
CNVR	copy number variable region
CYP	cytochrome
CYP1B1	cytochrome P450 1B1
CYP2D6	cytochrome P450 2D6

CYP2D7	cytochrome P450 2D7
CYP2D8	cytochrome P450 2D8
CYP3A4	cytochrome P450 3A4
CYP450	cytochrome P450
DDI	drug-drug interaction
DME	drug metabolizing enzyme
DNA	deoxyribonucleic acid
<i>E. coli</i>	<i>Escherichia coli</i>
ECL	enhanced chemiluminescence
EM	extensive metabolizer
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
HNF4A	hepatic nuclear factor 4 alpha
HRP	horseradish peroxidase
hsa-miR	<i>Homo sapiens</i> microRNA
HWE	Hardy-Weinberg equilibrium
IM	intermediate metabolizer
MAF	minor allele frequency

Mb	Mega bases
MCS	multiple cloning site
MFE	minimum folding energy
miRNA	micro ribonucleic acid or microRNA
miRSNP	microRNA SNP
mRNA	messenger RNA
NDM	N-desmethyldamoxifen
ng	nanogram
nM	nanomolar
PCR	polymerase chain reaction
PM	poor metabolizer
PXR	pregnane X receptor
qRT-PCR	quantitative real-time polymerase chain reaction
RISC	RNA induced silencing complex
RNA	ribonucleic acid
RT	reverse transcription
SD	standard deviation



SEM	standard error of the mean
SERM	selective estrogen receptor modulator
siRNA	small interfering RNA
SNP	single nucleotide polymorphisms
snRNA	small nuclear RNA
SSRI	selective serotonin reuptake inhibitor
Tam	tamoxifen
UGT2B7	UDP-glucuronosyltransferase 2B7
UM	ultrarapid metabolizer

## CHAPTER ONE

### **Introduction and Literature Review**

#### **Pharmacogenetics of interindividual variability in drug response**

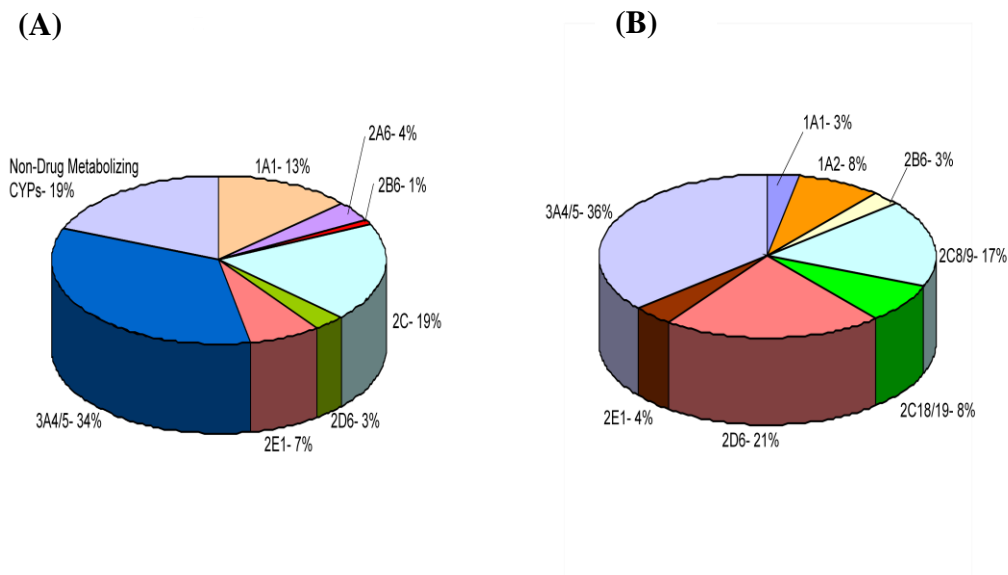
Genetic polymorphisms in drug metabolizing enzymes (DMEs), drug receptors and drug transporters influence an individual's ability to metabolize and eliminate drugs. These pharmacogenetic variations contribute to interindividual variability in drug metabolism and response. Among the DME genes, the greatest impact is due to polymorphisms in drug metabolizing Cytochrome P450s (CYP450s) such as CYP2A6, CYP2B6, CYP2C9, CYP2C19, CYP3A4, CYP3A5, and CYP2D6. Of these CYP450s, polymorphisms in *CYP2D6* gene have been extensively studied because of its clinical importance, especially in treating patients with drugs such as antidepressants, antipsychotics, anticancer agents, etc. For many CYP2D6 substrate drugs, screening an individual's *CYP2D6* genetic makeup and tailoring drug therapy based on that information can decrease the adverse events and increase both clinical and economic benefits (Laika et al., 2009).

#### **The cytochrome P450 drug metabolizing enzyme family**

The cytochrome P450 is a superfamily of heme-thiolate monooxygenase enzymes (EC 1.14.14.1) that is involved in the oxidative metabolism of a number of endogenous and exogenous compounds like steroids, drugs, carcinogens and mutagens. In humans, 57 functional CYP450 genes and 58 pseudogenes have been described (Nelson et al., 2004; Guengerich, 2008). The functional CYP450 genes are expressed in

the endoplasmic reticulum of a number of tissues (Nelson et al., 2004). The liver expressed CYP450s include many non-drug and drug metabolizing enzymes (Figure 1.1 A). Of the drug metabolizing CYP450s, the clinically important ones include phase I DMEs like CYP1A1, CYP1A2, CYP2B6, CYP2C8, CYP2C9, CYP2C19, CYP2E1, CYP3A4, CYP3A5 and CYP2D6 (Figure 1.1 B).

**Figure 1.1 A-B: The liver expressed Cytochrome P450s and their contribution to drug metabolism.**



Note: (A) Physical abundance of CYP450 enzymes in the human liver; (B) Percentage of clinically important drugs metabolized by the various CYP450 enzymes in the human liver. Source: redrawn from [www.pharmacy.wsu.edu/courses/PharS541/Docs/541.05.IV.Role%20of%20Induction%20&%20Inhibition.htm](http://www.pharmacy.wsu.edu/courses/PharS541/Docs/541.05.IV.Role%20of%20Induction%20&%20Inhibition.htm).

The drug metabolizing CYP450s are responsible for the metabolism of 40-45% of all marketed drugs and 70-80% of all phase I dependent drug metabolism (Ingelman-Sundberg and Rodriguez-Antona, 2005; Guengerich, 2008). The activity of these CYP450 enzymes determine: (1) the rate of biotransformation of a parent drug into active

and/or inactive metabolites, (2) the concentration of the drug and/or its metabolites that will be achieved in the body, and (3) the rate of elimination of the drug and its metabolites.

The activity of the drug metabolizing CYP450 enzymes, especially that of CYP2D6 is highly variable among individuals (Ingelman-Sundberg, 2005). Some of this variability is due to environmental (including drug-drug interactions) and genetic factors. These two factors lead to poor risk/benefit ratio for a number of CYP2D6 substrate drugs and can result in unexpected drug effects including, therapeutic failure and adverse drug reactions (Rau et al., 2004). Hence, for practical clinical purposes, it is important to understand the factors that contribute to variability in CYP2D6 mediated drug metabolism.

### **CYP2D6 and its role in drug metabolism**

Cytochrome P450 2D6 (CYP2D6) is a member of the CYP450 superfamily that is involved in the metabolism and bioactivation of a number of xenobiotics. CYP2D6 accounts for only ~3% of all CYP450 expression in the liver (Figure 1.1 A), but it is responsible for the metabolism of about 20-25% of commonly prescribed drugs (Ingelman-Sundberg, 2005; Figure 1.1 B). These include drugs such as antidepressants, neuroleptics, opiates, antiarrhythmics and antihypertensive agents (Table 1.1).

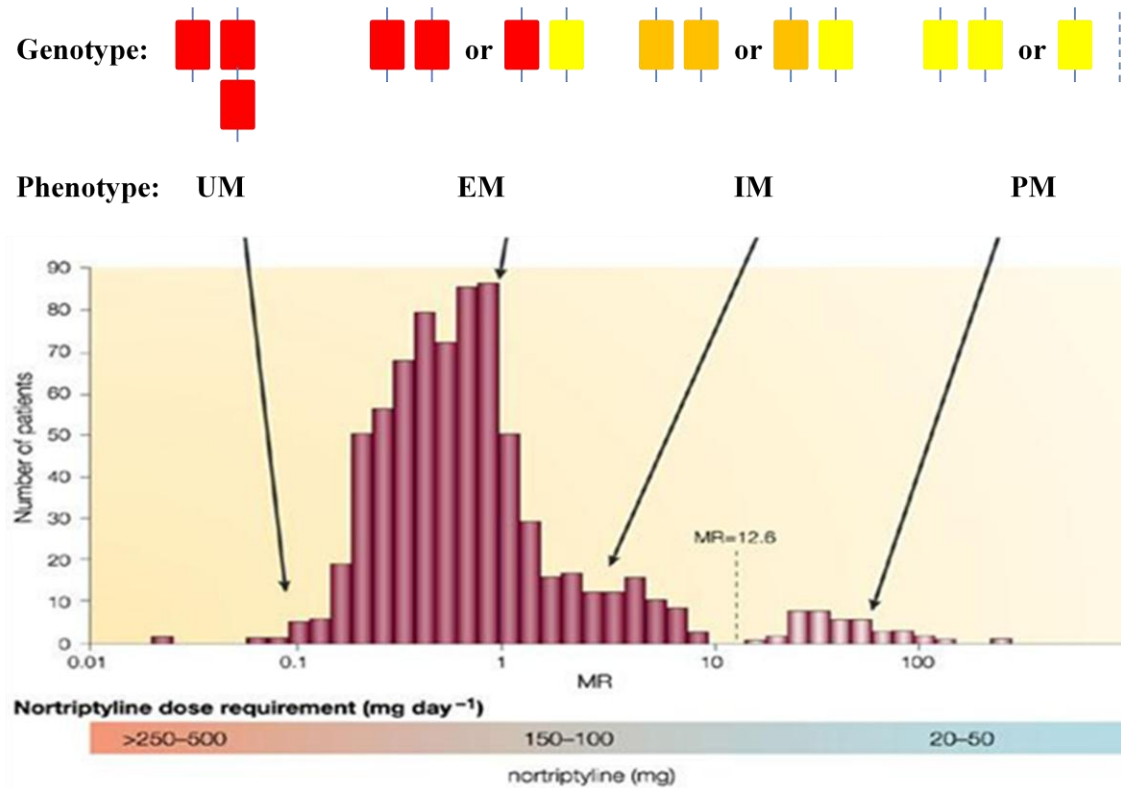
**Table 1.1: Major CYP2D6 drug substrates and inhibitors.**

SUBSTRATES		INHIBITORS
<p><b>Beta-blockers</b> S-metoprolol propafenone propranolol carvedilol timolol</p> <p><b>Antidepressants</b> clomipramine amitriptyline nortriptyline venlafaxine desipramine imipramine paroxetine fluoxetine</p> <p><b>SERM</b> tamoxifen</p> <p><b>Opioids</b> oxycodone codeine</p>	<p><b>Antipsychotics</b> zuclopenthixol perphenazine thioridazine haloperidol risperidone</p> <p><b>Others</b> dextromethorphan chlorpheniramine chlorpromazine promethazine amphetamine debrisoquine fluvoxamine phenacetin bufuralol lidocaine sparteine tramadol</p>	<p><b>Strong</b> bupropion paroxetine fluoxetine quinidine</p> <p><b>Moderate</b> terbinafine duloxetine sertraline</p> <p><b>Weak</b> amiodarone cimetidine</p>

Source: Redrawn from Cytochrome P450 drug interaction table (<http://medicine.iupui.edu/clinpharm/ddis/table.asp>).

There is considerable interindividual variability in CYP2D6 enzyme activity (Ingelman-Sundberg, 2005). Based on the enzyme activity, an individual can be classified as: (1) ultrarapid metabolizer (UM) with increased enzyme activity, (2) extensive metabolizer (EM) with 'normal' enzyme activity, (3) intermediate metabolizer (IM) with decreased enzyme activity, or (4) poor metabolizer (PM) with negligible enzyme activity (Figure 1.2).

**Figure 1.2: CYP2D6 genotype-phenotype correlation and nortriptyline dose adjustment for the different metabolizer groups.**



Note: Fully functional *CYP2D6* alleles are denoted by red boxes, decreased function alleles by orange boxes, null alleles by yellow boxes, and gene deletion by a dashed line. The associated phenotypes were determined by the urinary metabolic ratio (MR) of debrisoquine to 4-hydroxy-debrisoquine (interpreted as an *in vivo* index of CYP2D6 activity). MR = 12.6 is the cutoff point between subjects with 'poor metabolism' and subjects with intermediate or extensive metabolism. Source: modified from Myers, U.A., 2004.

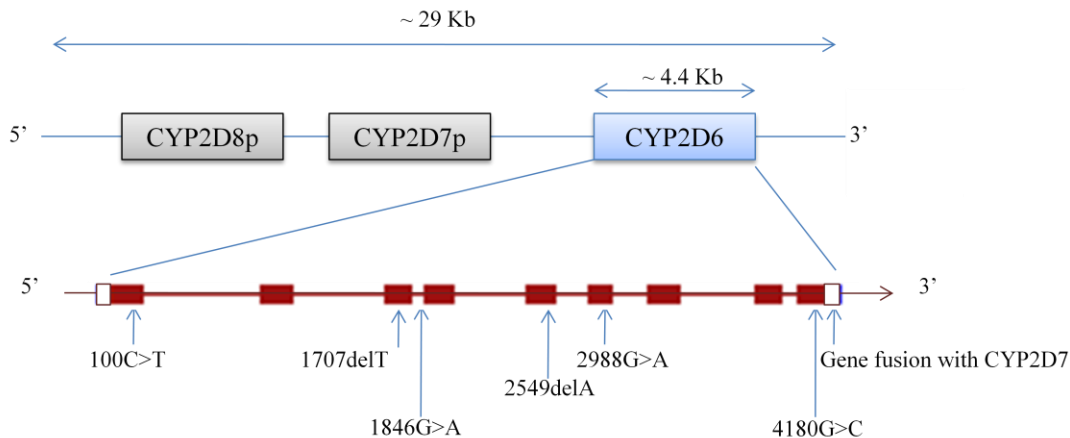
This variability in enzyme activity can affect drug metabolism and consequently, therapeutic response. For example, individuals who have increased CYP2D6 enzyme activity (UM) due to multiple copies of the functional gene can have therapeutic failure due to rapid drug metabolism. Hence, therapeutic plasma concentrations will not be achieved at ordinary drug dosages (Hedenmalm et al., 1997; Ingelman-Sundberg, 2004).

For example, when treating with the antidepressant drug, nortriptyline, the UMs require a much higher dose than the EMs, who in turn require a higher dose than the PMs (Figure 1.2). On the other hand, individuals who lack CYP2D6 enzyme activity (PM) due to lack of the functional copies of *CYP2D6* gene have either: (1) therapeutic failure due to decreased bioactivation of the prodrugs (e.g., tamoxifen), or (2) increased side effects due to slower clearance of CYP2D6 substrate drugs (e.g., desipramine, fluvoxamine, fluoxetine, citalopram, etc.; Ingelman-Sundberg et al., 1999; Ingelman-Sundberg, 2004).

### ***CYP2D6* gene structure**

*CYP2D6* belongs to the *CYP2D* gene family. The *CYP2D* gene cluster which is located on chromosome 22q13.1 (Kimura et al., 1989) consists of three genes: (1) *CYP2D6*, a 'functional' gene; (2) *CYP2D8*, a pseudogene that arose by gene conversion; and (3) *CYP2D7*, a nonfunctional gene that arose by gene duplication (Figure 1.3). In the *CYP2D* family, *CYP2D6* is the only gene that encodes a functional enzyme (Gaedigk et al., 2005). The wild type *CYP2D6* gene is ~4.4 kb long. It is composed of 9 exons and a 1491 bp long open reading frame (ORF) that translates into a protein with 497 amino acids (GenBank Accession No. M33388.1; Kimura et al., 1989). Many alternatively spliced transcript variants that encode different isoforms have also been reported (source: Entrez Gene - <http://www.ncbi.nlm.nih.gov/gene/1565>). The *CYP2D6* gene is highly polymorphic; these polymorphisms affect the pharmacokinetics, metabolism, safety, and efficacy of drugs that are metabolized by it (Ingelman-Sundberg, 2005; Figure 1.3).

**Figure 1.3: Human *CYP2D* gene family and common variants in the *CYP2D6* gene.**



Note: In the *CYP2D6* gene, the brown boxes represent the exons and the white boxes represent the untranslated regions (UTRs). Some of the *CYP2D6* representative variants are also shown; del denotes deletion; > denotes conversion.

The *CYP2D6* gene is not highly conserved across species. The rat and mouse orthologs equivalent to human *CYP2D6* are *cyp2d1* and *cyp2d22*, respectively. Among these three species, the drug metabolism profile is very different; they exhibit variability in substrate specificity and also enzyme inhibition (Bogaards et al., 2000). Among rats and mice, there is also considerable interstrain difference in the metabolic profile (Corchero et al., 2001). For example, debrisoquine, a *CYP2D6* substrate drug, is metabolized to 4-hydroxydebrisoquine by humans and Sprague-Dawley rats, but not by Dark Agouti rats, or by C57BL/6, DBA/2 and ddY strains of mice (Al-Dabbagh et al., 1981; Masubuchi et al., 1997). Hence, model organisms are generally considered to be inadequate for preclinical studies of *CYP2D6* substrate drugs. However, a humanized *CYP2D6* transgenic mouse model is now available and has been shown to be a good experimental model resembling the human extensive metabolizer phenotype for some drugs, including debrisoquine (Corchero et al., 2001).



## **CYP2D6 expression**

The CYP2D6 protein is expressed primarily in the endoplasmic reticulum and accounts for ~3% of the total CYP450 protein content in the human liver (Hakkola et al., 1994; Rodriguez-Antona et al., 2000; Figure 1.1 A). CYP2D6 mRNA and/or protein are also expressed at variable levels in extra-hepatic tissues including the circulating peripheral blood mononuclear cells, kidney, placenta, heart, brain, breast, lung and intestine (Romkes-Sparks et al., 1994; Carcillo et al., 1996; Guidice et al., 1997; McFayden et al., 1998; Siegle et al., 2001; Miksys et al., 2002). However, when compared to hepatic CYP2D6 expression, the extra-hepatic expression level is very low. Though *in vitro* studies have suggested a role for brain expressed CYP2D6 in the metabolism of endocannabinoids (Snider et al., 2008), the role of other tissues in CYP2D6 mediated drug metabolism is still not clear and is generally considered to be negligible.

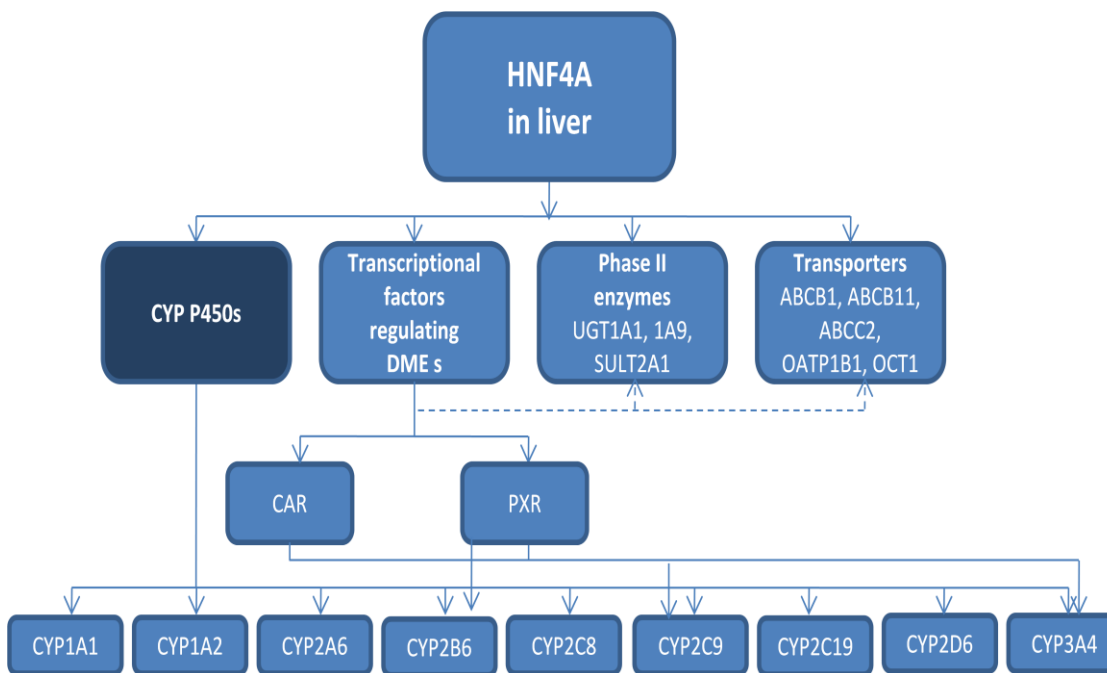
The CYP2D6 mRNA and protein expression do not appear to be induced by drugs and xenobiotics (Eichelbaum et al., 1986). CYP2D6 mRNA and protein expression appears to be mainly under genetic control, i.e., *CYP2D6* gene polymorphisms and gene copy numbers. However, CYP2D6 enzyme activity is induced by drugs like rifampin and dexamethasone (source: Cytochrome P450 drug interaction table: <http://medicine.iupui.edu/clinpharm/ddis/table.asp>) and also by pregnancy (Wadelius et al., 1997). The detailed mechanism underlying the induction during pregnancy is unclear. It was thought that the placenta or fetal liver may contribute to this metabolic induction. However, it has been shown that in the fetus, even though CYP2D6 mRNA is detectable, CYP2D6 protein expression is mostly undetectable. The CYP2D6 protein

concentration rises only a few days after birth (Treluyer et al., 1991). The mechanisms that trigger post-natal expression of CYP2D6 are still not known. Hence, it is possible that the developmental regulation of CYP2D6 expression may be due to posttranslational regulation by microRNAs.

### **Transcriptional regulation of CYP2D6 expression by HNF4A**

The hepatic expression of CYP2D6 is critical for the metabolism of the substrate drugs. This hepatic CYP2D6 gene expression is primarily regulated through liver-enriched transcription factors (LTF) such as hepatic nuclear factor 4 $\alpha$  (HNF4 $\alpha$ ; HNF4A; NR2A1; Corchero et al., 2001). Mutations in HNF4A have been shown to regulate CYP2D6 expression, and consequently affect CYP2D6 enzyme activity (Lee et al., 2008). HNF4A is considered to be a ‘master regulator’ that regulates the expression of many genes involved in drug metabolism like phase I enzymes (including CYP2D6), phase II enzymes, transporters and other transcriptional factors that regulate the CYP450 genes (Kamiyama et al., 2007; Figure 1.4). Apart from the regulation of drug metabolizing enzymes, HNF4A regulates several genes that are involved in other hepatic functions including, fatty acid and cholesterol metabolism, glucose metabolism, urea biosynthesis and liver differentiation (Kamiyama et al., 2007). Thus, regulation of HNF4A will affect many genes that are involved in important liver functions.

**Figure 1.4: HNF4A is a master transcriptional regulator of drug metabolizing enzymes.**



Note: Dashed line indicates that only some of the Phase II enzymes and transporters are specifically regulated by PXR and/or CAR.

Most of the data on CYP2D6 transcriptional regulation has been obtained from transient transfection of CYP2D6 in hepatoma cell lines. This is because the CYP2D6 enzyme expression is quickly lost in freshly isolated hepatocytes, and is neither maintained nor inducible in hepatoma cell lines (Cairns et al., 1996).

### **CYP2D6 mediated drug metabolism**

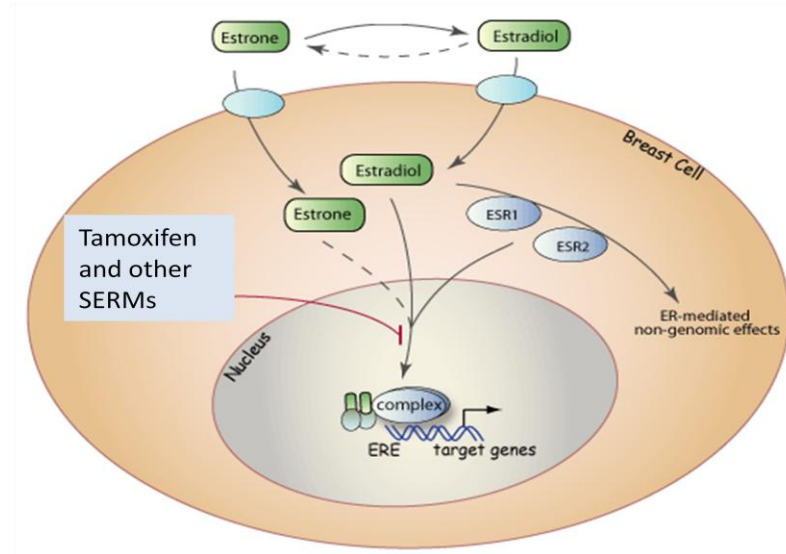
CYP2D6 is responsible for the metabolism of about 20-25% of commonly prescribed drugs (Ingelman-Sundberg, 2005). Some of these drugs include (Table 1.1): beta-blockers, antiarrhythmics, tricyclic antidepressants, selective serotonin reuptake inhibitors (SSRIs), opioids, antihypertensives, antitussives, antipsychotics and selective

estrogen receptor modulators (SERMs; e.g., tamoxifen). While, some of these CYP2D6 substrate drugs (e.g., nortriptyline) are metabolized into inactive metabolites, the other drugs (e.g., codeine and tamoxifen) are bioactivated into active metabolites.

### **Tamoxifen as an endocrine therapy for breast cancer**

Tamoxifen is a clinically important drug that is metabolized by CYP2D6. Tamoxifen is primarily used in the hormonal therapy in breast cancer. Additionally, it is also used in the treatment of infertility, retroperitoneal fibrosis, gynecomastia and bipolar disorder. It is a selective estrogen receptor modulator (SERM) that has been repeatedly shown to improve the survival rate in breast cancer patients (Osborne, 1998). In breast cells, tamoxifen acts as an antagonist and competes with estradiol to bind to the estrogen receptors (ER) and induces conformational changes that inhibit the transcription of estrogen regulated genes (Figure 1.5). However, it is not a pure antagonist; it exhibits mixed agonist/antagonist activity in different tissues. In case of tissues that are weakly estrogenic (e.g., cardiovascular tissues and bone), it has some protective effect (Chang et al., 1996). But, in case of tissues that are strongly estrogenic (e.g., endometrium), it can lead to rare adverse side effects like endometrial cancer. Despite the common and rare side effects, the risk/benefit ratio is still favorable for tamoxifen therapy (Bonanni et al., 2007).

**Figure 1.5: Mechanism of action of tamoxifen in breast cancer cells.**



Note: Estradiol binds to the estrogen receptor (ER) and results in changes that bring about cellular proliferation. Tamoxifen is anti-estrogen that competes with estradiol for binding with the ER and blocks the proliferation. Source: modified from [www.pharmgkb.org/](http://www.pharmgkb.org/).

Worldwide, tamoxifen is the most commonly used hormonal therapy for breast cancer. Tamoxifen is still heavily used, even though some postmenopausal patients are now taking aromatase inhibitors (AIs) instead of tamoxifen (American Society of Clinical Oncology guidelines; Winer et al., 2005). In the U.S. alone, in 2007, ~1.6 million tamoxifen prescriptions were issued (source: [www.drugtopics.com](http://www.drugtopics.com)). Some of the reasons for the heavy use of tamoxifen include: (a) it is *the* standard endocrine therapy for premenopausal breast cancer, since AIs are not given to premenopausal women, (b) it is the most common endocrine therapy used for breast cancer risk reduction, (c) it is the only FDA approved endocrine therapy for the treatment of ductal carcinoma *in situ* (DCIS), (d) it is the only hormonal therapy approved for male breast cancer, (e) it is used in the treatment of patients who are intolerant or resistant to AI treatment, and (f) in the

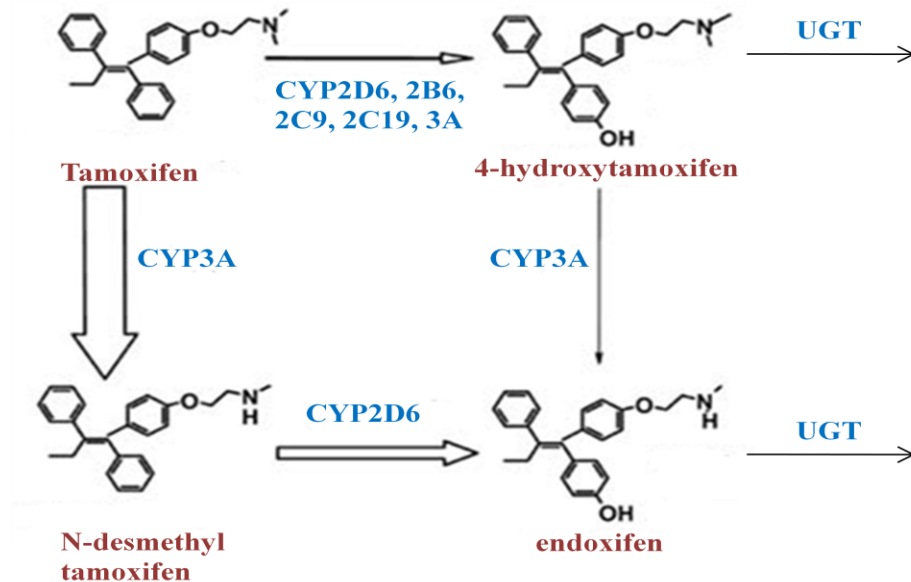
USA and around the world, its low cost makes it affordable to many women who cannot afford the AIs.

Although tamoxifen is a very effective therapy, the pharmacodynamic response to tamoxifen varies considerably between individuals (Lonning et al., 1992; Ingle et al., 1999). This variability is seen in both its effectiveness as well as its side effects (Osborne, 1998). Some of the side effects include hot flashes, mood disturbances, vaginal discharge and bleeding, invasive endometrial cancer and deep vein thrombosis. These side effects can be severe enough to cause discontinuation of the therapy. About 30-50% of women do not complete the recommended 5 years of therapy (Barron et al., 2007); one of the common reasons for discontinuation is the side effects (Rae et al., 2009). Consequently, a better understanding of the variability in the occurrence of these side effects would help to improve the compliance of this life saving drug.

### **Metabolism of tamoxifen by CYP2D6**

Tamoxifen is heavily metabolized by several of the cytochrome P450 drug metabolizing enzymes (Desta et al., 2004; Figure 1.6). CYP3A is involved in the metabolism of tamoxifen to N-desmethyltamoxifen (NDM). NDM is subsequently metabolized to a more active metabolite, endoxifen, by the polymorphic enzyme CYP2D6. Tamoxifen is also hydroxylated by a number of CYP450s (CYP3A, CYP2C9, CYP2C19, CYP2B6, and CYP2D6) to form 4-hydroxytamoxifen (4-OH-tam) which is further metabolized by CYP3A to endoxifen. Phase II drug metabolizing enzyme like UDP glucuronosyltransferase 2B7 (UGT2B7) may be involved in the subsequent clearance of these metabolites (Sun et al., 2007; Blevins-Primeau et al., 2009).

**Figure 1.6: Metabolism of tamoxifen by Phase I and II drug metabolizing enzymes.**

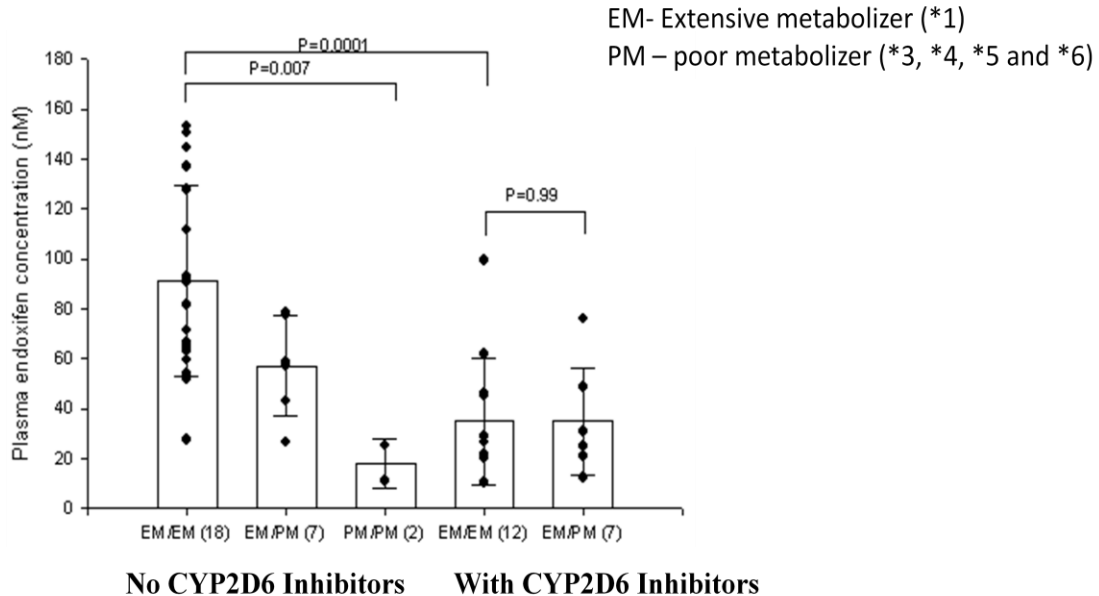


Note: Size of the arrow indicates the relative size of the pathway. Source: modified from Borges, et al., 2006.

Endoxifen is considered to be the most active metabolite of tamoxifen and is primarily generated by the activity of CYP2D6 (Desta et al., 2004; Johnson et al., 2004). Compared to the parent drug, tamoxifen, endoxifen is ~30-100-fold more potent in estrogen receptor binding, regulation of estrogen dependent gene expression and inhibition of breast cancer cell growth (Johnson et al., 2004). Endoxifen is similar to the extensively studied tamoxifen metabolite, 4-hydroxytamoxifen (4-OH-tam; Lim et al., 2005; Lim et al., 2006), but in breast cancer patients, it reaches ~10-fold higher concentration than 4-OH-tam (Stearns et al., 2003). However, there is over 10-fold interindividual variability in the concentration of this potent metabolite in breast cancer patients (Stearns et al., 2003; Jin et al., 2005). A part of this variability is due to *CYP2D6* gene variants and due to concurrent use of drugs that inhibit CYP2D6 enzyme activity (Figure 1.7; Stearns et al., 2003; Jin et al., 2005; Borges et al., 2006). Hence,

comprehensive *CYP2D6* genotyping, identification of *CYP2D6* gene copy numbers and evaluation of co-medication for *CYP2D6* inhibitors are essential to account for the variability in plasma endoxifen concentrations. However, there is considerable unexplained variability in the endoxifen concentrations even after correcting for these genetic and environmental factors.

**Figure 1.7: *CYP2D6* gene polymorphisms and concomitant use of *CYP2D6* inhibitors are associated with plasma endoxifen concentration.**



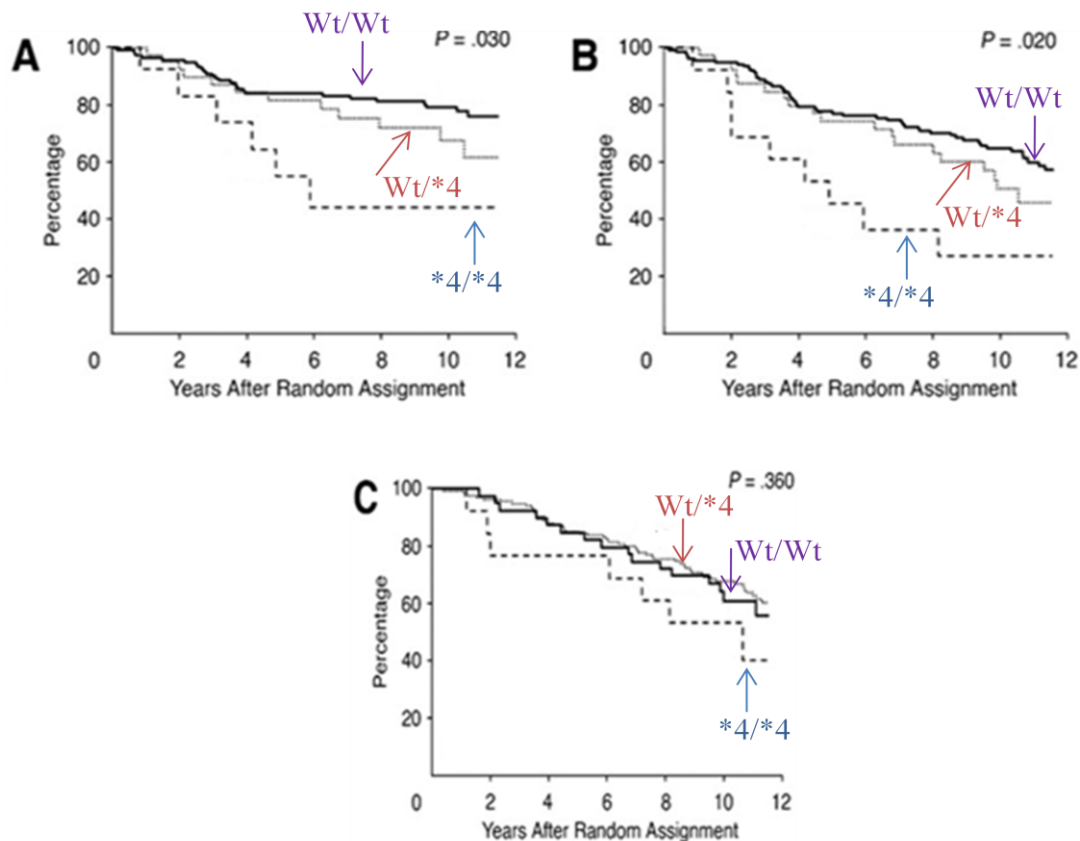
Note: Blood samples were collected after 4 months of tamoxifen treatment for identifying *CYP2D6* genotype and for the measurement of plasma endoxifen concentrations. Solid circles represent the individual plasma endoxifen concentrations; vertical bars represent the mean plasma endoxifen concentration for each group. Source: modified from Jin, et al., 2005.

The *CYP2D6* polymorphisms that reduce endoxifen concentration are also associated with reduced side effects and poorer clinical outcomes. A number of studies



performed primarily in Caucasian women (Goetz et al., 2005; Bonanni et al., 2006; Goetz et al., 2007; Gonzalez-Santiago et al., 2007; Schroth et al., 2007; Schroth et al., 2009) and a relatively few studies in Asian women (Lim et al., 2007; Kiyotani et al., 2008; Xu et al., 2008) have suggested that *CYP2D6* PM or IM allelic status is associated with poorer outcomes in prevention, adjuvant and metastatic settings (Figure 1.8; Table 1.2).

**Figure 1.8: *CYP2D6* genotype is associated with clinical outcomes.**



Note: Kaplan-Meier estimates of (A) relapse-free time; (B) disease-free survival; and (C) overall survival for patients with *CYP2D6*\*4/\*4 genotype (PM) and *CYP2D6*\*Wt/\*4 (IM) when compared to the wildtype *CYP2D6*\*1/\*1 genotype (EM). Source: modified from Goetz, et al., 2005.

Not all studies have confirmed the association between *CYP2D6* variant genotype and poor clinical outcomes (Nowell et al., 2005; Wegman et al., 2005; Wegman et al.,

2007; Okishiro et al., 2009; Table 1.3). Two of these studies (Wegman et al., 2005; Wegman et al., 2007) have reported that the PMs have a better outcome when compared to the EMs. One of the possible explanations is that the PMs (who generally have fewer side effects) continued with their tamoxifen therapy, while the EMs (who generally have more side effects) discontinued the therapy. Two other studies (Nowell et al., 2005; Okishiro et al., 2009) have reported a lack of association between *CYP2D6* gene variants and clinical outcome, i.e., *CYP2D6* PMs and IMs have the same outcome as the EMs. None of these studies have no data on drug compliance. About 30-50% of women do not complete the 5-year tamoxifen regimen (Barron et al., 2007); one of the reasons for discontinuation is increased side effects in the EMs when compared to the PMs (Rae et al., 2009). Another possible explanation for the differences in the results is that the variability in other related pathways may also be important. For example, plasma endoxifen concentration is dependent not only on its formation by *CYP2D6* but also on its clearance by phase II enzymes like *UGT2B7* (Blevins-Primeau et al., 2009). A missense mutation in *UGT2B7* (His268Tyr; *UGT2B7*\*2) results in decreased enzyme activity and consequently slower drug clearance (Sun et al., 2007). Hence, it is conceivable that *UGT2B7* alone or in combination with *CYP2D6* contributes to variability in tamoxifen response. Similarly, a polymorphism in the drug transporter *ABCC2* has been shown to be associated with poor clinical outcome, even though it is not associated with plasma concentrations of tamoxifen and its metabolites (Kiyotani et al., 2010). Hence, it is likely that other genes in related pathways may also account for some of the interindividual variability in drug metabolism and response.

**Table 1.2: Summary of studies reporting an association between *CYP2D6* genotype and response to tamoxifen therapy.**

Reference	Goetz et al., 2005	Schroth et al., 2007	Goetz et al., 2007	Xu et al., 2008	Kiyotani et al., 2008	Lim et al., 2007
Population	Caucasian			Asian		
Location	USA	Germany	USA	China	Japan	Korea
N	256	486	256	293	67	212
Setting	Adjuvant	Adjuvant	Adjuvant	Adjuvant	Adjuvant	Adjuvant, metastatic
Menopausal and tumor status	Postmenopausal and ER <sup>+</sup>	Postm./ER <sup>+</sup> , or ER <sup>-</sup>	Postm./ER <sup>+</sup>	Prem., or postm./ER <sup>+</sup> , or ER <sup>-</sup>	Prem., or postm./ER <sup>+</sup>	Prem., or postm./ER <sup>+</sup>
Analyzed subgroup		Tamoxifen monotherapy, ER <sup>+</sup>		Tamoxifen	CYP2D6*1 and *10	Metastatic disease
Subgroup N	190	206	180	152	58	21
Reference	CYP2D6*1/*1 or CYP2D6*1/*4	EM/EM or EM/IM	EM/EM + no CYP2D6 inhibitor	CYP2D6*1/*1 or *1/*10	CYP2D6*1/*1	CYP2D6*1/*1 or *1/*10
Comparator	CYP2D6*4/*4	EM/PM or IM/PM or IM/IM or IM/PM	PM/PM or any genotype + potent CYP2D6 inhibitor	CYP2D6*10/*10	CYP2D6*10/*10	CYP2D6*10/*10
Outcome	RFS	EFS	DFS	DFS	RFS	TDP
Univariate hazard ratio*	2.71 (1.15–6.41) <sup>  </sup>	Not reported	Not reported	Not reported	8.67 (0.24–19.79) <sup>  </sup>	3.69 (1.28–0.67) <sup>  </sup>
Multivariate hazard ratio*	1.85 (0.76–4.52)	1.89 (1.10–3.25) <sup>  </sup>	2.44 (1.27–4.69) <sup>  </sup>	4.7 (1.1–20.0) <sup>  </sup>	10.04 (1.17–6.27) <sup>  </sup>	3.68 (1.23–1.04) <sup>  </sup>

Note: \* 95% confidence intervals are indicated in brackets. <sup>||</sup> p < .05; <sup>†</sup> p < .01; ER<sup>+</sup>, estrogen ± progesterone positive tumor; ER<sup>-</sup>, estrogen ± progesterone negative tumor; N, number; DFS, disease-free survival; EFS, event-free survival; RFS, recurrence-free survival; RFT, relapse-free time; TDP, time to disease progression; Prem., premenopausal; Postm., postmenopausal. Source: Redrawn from Hoskins et al., 2009.

**Table 1.3: Summary of studies reporting either a lack of association or better clinical outcome due to CYP2D6 gene variants.**

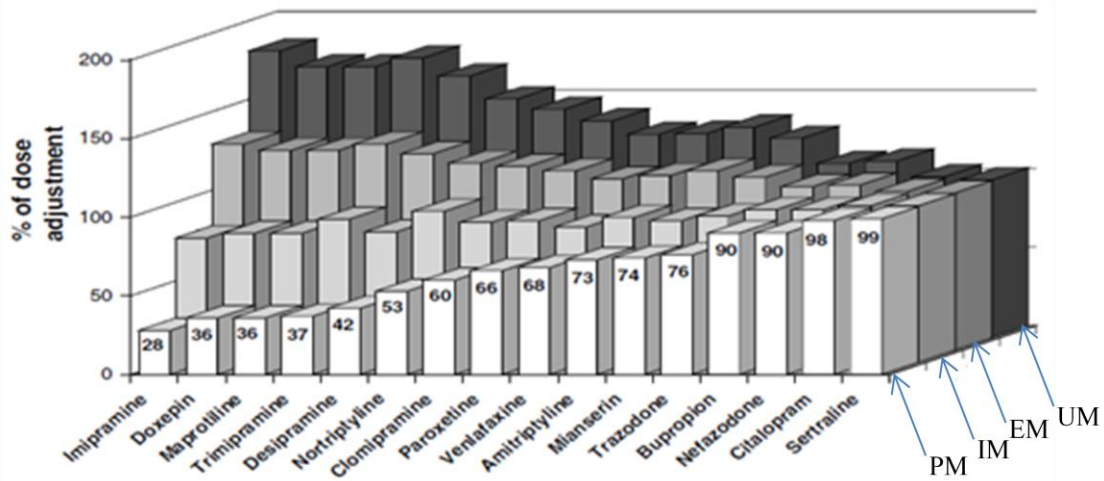
Reference	Wegman et al., 2007	Nowell et al., 2005	Wegman et al., 2005	Okishiro et al., 2009
<b>Population</b>	Caucasian			
<b>Location of study</b>	Sweden	USA	Sweden	Asian
<b>N</b>	677	337	226	173
<b>Setting</b>	Adjuvant	Adjuvant	Adjuvant	Adjuvant chemotherapy or goserelin <sup>§</sup>
<b>Menopausal and tumor status</b>	Postm./ ER <sup>+</sup>	ER <sup>+</sup> or ER <sup>-</sup>	Prem., or postm./ ER <sup>+</sup> , or ER <sup>-</sup>	Prem., or postm./ ER <sup>+</sup> , or ER <sup>-</sup>
<b>Analyzed subgroup</b>	2 yr tamoxifen	Tamoxifen	CYP2D6*1/*1	Tamoxifen monotherapy
<b>Subgroup N</b>	103	160	107	73
<b>Reference</b>	CYP2D6*1/*1	CYP2D6*1/*1	No tamoxifen	CYP2D6*1/*1 or *1/*10
<b>Comparator</b>	CYP2D6*1/*4 or *4/*4	CYP2D6*1/*4 or *4/*4	Tamoxifen	CYP2D6*10/*10
<b>Outcome</b>	RFS	PFS	DRFS	RFS
<b>Univariate hazard ratio<sup>*</sup></b>	0.87 (0.38–1.97) <sup>  </sup>	Not reported	Not reported	0.94 (0.34–2.60)
<b>Multivariate hazard ratio<sup>*</sup></b>	Not reported	0.67 (0.33–1.35)	0.28 (0.11–0.74) <sup>#</sup>	0.6 (0.18–1.92)

Note: <sup>\*</sup>95% confidence intervals are indicated in brackets. <sup>||</sup>p < 0.05; <sup>#</sup>p < 0.01; ER<sup>+</sup>, estrogen ± progesterone positive tumor; ER<sup>-</sup>, estrogen ± progesterone negative tumor; N, number; DRFS, distance recurrence free survival; RFS, recurrence-free survival; PFS, progression free survival; Prem., premenopausal; Postm., postmenopausal. Source: Redrawn from Hoskins et al., 2009.

## **Interindividual variability in CYP2D6 enzyme activity and CYP2D6 mediated drug metabolism**

There is considerable interindividual variability in CYP2D6 enzyme activity and this leads to variability in metabolism and disposition of the substrate drugs (Ingelman-Sundberg, 2005). Based on *CYP2D6* allelic make up and enzyme activity, an individual can be classified into one of the four different phenotype groups: poor (PM), intermediate (IM), extensive (EM), and ultrarapid (UM) metabolizer. In these four different phenotype groups, for a number of drugs, one drug dose size does not fit all. For example, for a number antidepressants (e.g., imipramine, paroxetine, nortriptyline, etc.), the dosing required to achieve the same plasma concentration can differ by 10-30 fold between individuals due to their CYP2D6 metabolizer status (Figure 1.9; Kirchheiner et al., 2004; Ingelman-Sundberg and Rodriguez-Antona, 2005).

**Figure 1.9: CYP2D6 mediated quantitative influences on pharmacokinetics of selected antidepressant drugs.**



Note: CYP2D6 mediated quantitative influences on pharmacokinetics of selected antipsychotic drugs expressed as percent dose adjustments. White bars represent *CYP2D6* poor metabolizer (PM), light gray – intermediate metabolizer (IM), dark gray – extensive metabolizer (EM), black – ultrarapid metabolizers (UM). Source: modified from Kirchheiner et al., 2004.

### **Causes of variability in CYP2D6 enzyme activity**

The primary factors that are currently known to cause interindividual variability in CYP2D6 enzyme activity include environmental and genetic factors.

#### **1. Environmental factors**

Environmental factors like diet (Aklillu et al., 2002) and drug-drug interactions (Lam et al., 2002) can influence CYP2D6 enzyme activity. Many clinically important drugs are CYP2D6 enzyme inhibitors (Table 1.1) and consequently affect the metabolism of other CYP2D6 substrate drugs. *In vivo* and *in vitro* studies using the CYP2D6 probe drug dextromethorphan have shown that certain selective serotonin reuptake inhibitors

(SSRIs; e.g., fluoxetine and paroxetine) can act as CYP2D6 enzyme inhibitors. These CYP2D6 inhibitor drugs can convert extensive metabolizers into poor metabolizers (Lam et al., 2002). In breast cancer patients taking tamoxifen, these SSRIs were routinely co-prescribed to treat depression and to alleviate hot flash symptoms. This results in lower plasma endoxifen concentration because these SSRIs inhibit CYP2D6 activity and thus interfere with tamoxifen metabolism (Jin et al., 2005). Thus, these inhibitors convert a CYP2D6 extensive metabolizer to a phenotypic poor metabolizer, thus creating a phenocopy of a poor metabolizer (Figure 1.7). The concurrent use of some of the SSRIs (e.g. paroxetine) during tamoxifen therapy is associated with increased risk of death from breast cancer (Kelly et al., 2010).

So far, there are no known xenobiotics or drugs that have been shown to induce the expression of CYP2D6 messenger RNA (mRNA) and/or protein (Eichelbaum et al., 1986).

## **2. Genetic factors**

CYP2D6 enzyme expression appears to be mainly under genetic control, i.e., gene polymorphisms and copy number variations (CNVs). These variants can affect CYP2D6 enzyme expression and activity.

### **a. *CYP2D6* gene polymorphisms**

*CYP2D6* is a highly polymorphic gene. So far, more than 78 alleles and allelic subvariants have been identified (Human Cytochrome P450 Allele Nomenclature Committee: [www.cypalleles.ki.se/cyp2d6.htm](http://www.cypalleles.ki.se/cyp2d6.htm)).

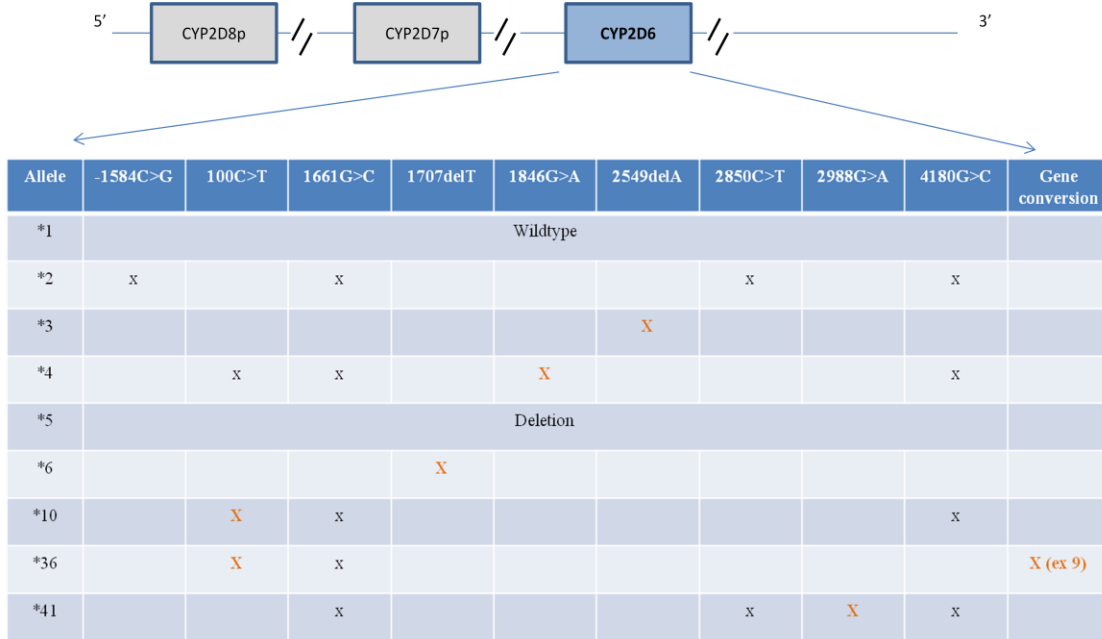
These alleles can be classified as:

- i. *Functional alleles*, with ‘normal’ enzyme activity (e.g., *CYP2D6*\*1, \*2, \*35, etc.)
- ii. *Reduced functional alleles*, with decreased enzyme activity (e.g., *CYP2D6*\*9, \*10, \*17, \*29, \*37, \*41, etc.), and
- iii. *Nonfunctional alleles*, with negligible enzyme activity (e.g., *CYP2D6*\*3-\*8, \*11-\*16, \*18-\*20, \*36, \*38, \*40, \*42, etc.).

A *CYP2D6* star-allele represents either a single genetic variant or a ‘haplotype’ that results in: amino acid substitution, post-transcriptional modification, post-translational modification, alteration in transcription, splicing, or translation (Robarge et al., 2007). *CYP2D6*\*1 is the wildtype reference sequence to which all the variants are compared. As new nucleotide variants (either a single variant or a combination of variants) are identified, a unique number is assigned (e.g., *CYP2D6*\*2, \*4, \*10, etc.; Figure 1.10). Different combinations of nucleotide changes within a named star-allele are defined by additional letters. For example, *CYP2D6*\*2A is characterized by the following nucleotide changes: -1584C>G, -1235A>G, -740C>T, -678G>A, gene conversion with *CYP2D7* in intron 1, **1661G>C**, **2850C>T** and **4180G>C**; the subvariant *CYP2D6*\*2B is characterized by 1039C>T, **1661G>C**, **2850C>T**, **4180G>C** and the subvariant *CYP2D6*\*2C is characterized by **1661G>C**, 2470T>C, **2850C>T**, and **4180G>C**. That is, *CYP2D6*\*2A-C share key mutations, but are further distinguished based on the additional mutations that they do not share.



**Figure 1.10: Polymorphisms in selected *CYP2D6* alleles.**



Note: **X** represents single nucleotide polymorphisms (SNP) that are used for genotyping; x represents some of the representative SNPs present in the given *CYP2D6* star-allele haplotype.

These polymorphisms can affect *CYP2D6* enzyme expression and activity. For example, individuals with *CYP2D6*\*1/\*10 and *CYP2D6*\*10/\*10 genotypes had 3-fold lower *CYP2D6* protein in their liver microsomes when compared to *CYP2D6*\*1/\*1 individuals (Shimada et al., 2001). *In vitro* expression studies have shown that *CYP2D6*\*10 enzyme had 50-100 fold lower efficiency in metabolizing dextromethorphan and fluoxetine when compared to *CYP2D6*\*1 (Yu et al., 2002). Consequently, these polymorphisms affect pharmacokinetics, metabolism, safety and efficacy of *CYP2D6* substrate drugs. Based on the *CYP2D6* allelic make up, an individual can be classified into different categories such as ultrarapid (UM), extensive (EM), intermediate (IM) or poor (PM) metabolizer (Table 1.4).

**Table 1.4: Consequences of selected *CYP2D6* gene polymorphisms.**

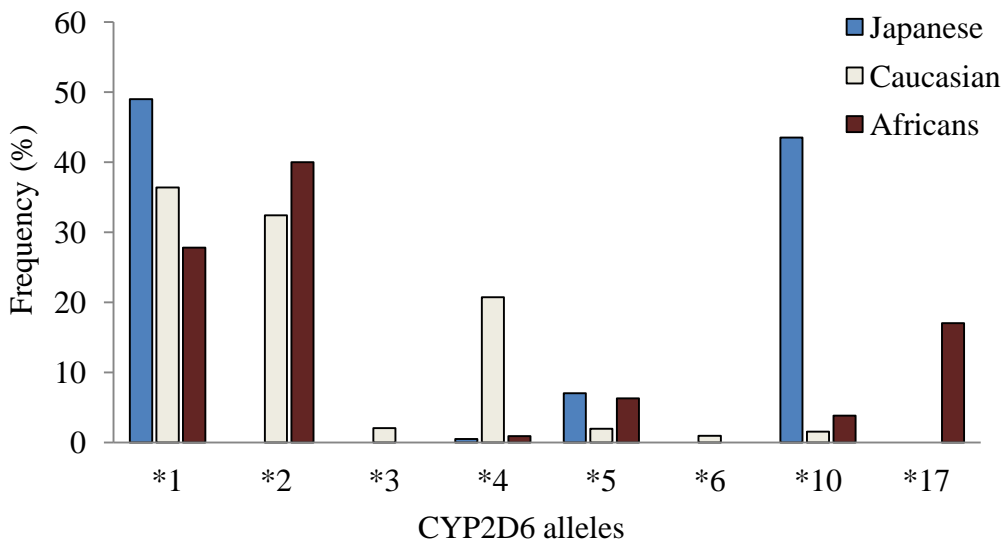
<i>CYP2D6</i> allele	Mutation <sup>a</sup>	Consequence of mutation	Enzyme <sup>b</sup>	Metabolizer status
*1	None	None	Normal	EM
*2	4180G>C	Amino acid change	Normal	EM
*17	1023C>T	Amino acid change	Unstable	IM
*41	2988G>A	Splicing defect	Unstable	IM
*10	100C>T	Amino acid change	Unstable	IM
*36	100C>T, gene conversion in exon 9	Amino acid changes	Inactive	PM
*3	A2549del	Frameshift	Inactive	PM
*4	1846G>A	Splicing defect	Inactive	PM
*5	Gene deletion	No enzyme	No enzyme	PM
*6	T1707del	Frameshift	Inactive	PM

Note: Some of the frequent *CYP2D6* alleles in the Caucasian, Asian and African populations are represented here. <sup>a</sup> mutation that is used for genotyping the corresponding *CYP2D6* allele; <sup>b</sup> *in vivo* enzyme activity (source: [www.cypalleles.ki.se/cyp2d6.htm](http://www.cypalleles.ki.se/cyp2d6.htm)); PM - poor metabolizer; IM - intermediate metabolizer; EM - extensive metabolizer.

There are considerable differences in the frequencies of *CYP2D6* gene polymorphisms in different ethnic populations (Figure 1.11). For example, *CYP2D6*\*4, *CYP2D6*\*3, *CYP2D6*\*5 and *CYP2D6*\*6 are present at a combined frequency of 20-30% in Caucasians (Bradford, 2002; Sistonen et al., 2007) and are the primary cause of poor metabolism in this population. However, in Asians and Africans, *CYP2D6*\*3-\*6 alleles are present in less than 7% frequency (Bradford, 2002). In Asians, the primary cause for reduced enzyme activity is *CYP2D6*\*10, an intermediate metabolizer allele that is present at a frequency of 40-60% (Veiga et al., 2009). However, in both Caucasians and

Africans, *CYP2D6*\*10 is present at a frequency of less than 5% (Bradford, 2002). In Africans, the primary cause of reduced enzyme activity are *CYP2D6*\*17 and *CYP2D6*\*29. Both these alleles are present in less than 3% frequency in both Caucasian and African populations (Bradford, 2002; Sistonen et al., 2007).

**Figure 1.11: Frequencies of selected *CYP2D6* alleles in different ethnic populations.**



Source: data from Veiga et al., 2009, Sachse et al., 1997 and Wennerholm et al., 1999.

For practical clinical purposes, *CYP2D6* enzyme activity is predicted based on the *CYP2D6* genotype scoring system (Table 1.5). Each of the nonfunctional *CYP2D6* allele which codes for either no protein (*CYP2D6*\*5) or for protein with no enzyme activity (e.g., *CYP2D6*\*3, \*4, \*6, etc.) is given a score of 0. A reduced function allele (e.g., *CYP2D6*\*9, \*10, \*17, \*41, etc.) which codes for protein with decreased enzyme activity is given a score of 0.5. A fully functional allele (e.g., *CYP2D6*\*1, \*2, \*35, etc.) which codes for protein with normal enzyme activity is given a score of 1. An individual's *CYP2D6* genotype score is the summation of the individual allelic scores.

**Table 1.5: CYP2D6 enzyme activity scores based on CYP2D6 alleles and copy number variations.**

<i>CYP2D6</i> allele	Metabolizer status	<i>CYP2D6</i> allele score
*3-*8, *11-*16, *18-*20, *36, *38, *40, *42, *44, *56, *4xN	PM	0
*10, *17, *37, *9, *29, *41, *45, *46	IM	0.5
*41xN, *45xN	EM <sup>a</sup>	1
*1, *2, *33, *35	EM	1
*1xN, *2xN, *35xN	UM	2

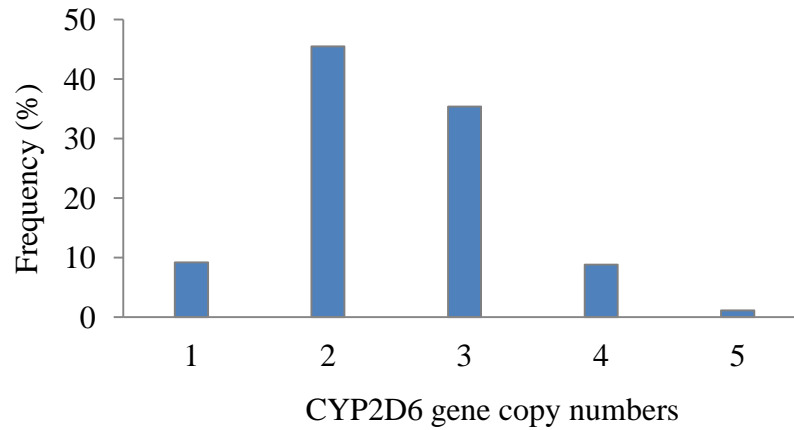
Note: This is an arbitrary scoring system based on the *CYP2D6* alleles present. A score of 0.5 does not mean that an allele has one-half or one-fourth the enzyme activity when compared to an allele with a score of 1 or 2, respectively. PM - poor metabolizer; IM - intermediate metabolizer; EM - extensive metabolizer; UM - ultrarapid metabolizer; <sup>a</sup> considered to be UM by other publications (Zanger et al., 2004; Zineh et al., 2004). Source: redrawn from Borges, S., et al., 2010.

New *CYP2D6* alleles and diplotype combinations are continuing to be reported; this adds to the complexity of *CYP2D6* genotype-guided drug dosage decisions in the clinic. Co-medication with *CYP2D6* inhibitors also adds to the discordance. Hence, genotype-phenotype based scoring systems have been developed to help clinicians to individualize drug therapy (Zanger et al., 2004; Gaedigk et al., 2008; Borges et al., 2010). Potential confounders in implementing these scoring systems include some of the poorly understood factors like pregnancy that affect *CYP2D6* enzyme activity.

## **b. CYP2D6 copy number variations (CNVs)**

Although *CYP2D6* gene polymorphisms contribute to a large portion of interindividual variability in *CYP2D6* enzyme activity, *CYP2D6* gene copy number variations (CNVs) also contribute to this variability (Weinshilboum, 2003; Eichelbaum et al., 2006; Yu and Shao, 2009). Copy number variation is defined as the variation in the number of copies of a gene (or sequences of DNA) from the normal two copies in the genome of a specific individual. The *CYP2D6* gene often undergoes gene deletion or multiplication resulting in CNVs. The frequencies of these CNVs vary considerably among different ethnic populations (Bradford, 2002; Sistonen et al., 2007). For example, the frequency of the *CYP2D6* whole gene deletion (*CYP2D6*\*5) ranges from 1-7% in Caucasians, Africans and Asians populations (Bradford, 2002). Similarly, *CYP2D6* whole gene multiplication ranges from 0-13% in Caucasians and Africans (Bradford, 2002; Sistonen et al., 2007), while in Asians, *CYP2D6* gene multiplication is as high as 54% in Asians (Figure 1.12; Hosono et al., 2009). The genetics of *CYP2D6* CNV is further complicated by the variety of *CYP2D6* alleles that exist as multiple copies. While, the functional *CYP2D6*\*1 and \*2 duplications are frequent among all ethnicities (Bradford, 2002; Sistonen et al., 2007), the nonfunctional *CYP2D6*\*4 is more common among Caucasians (Bradford, 2002) and the reduced function *CYP2D6*\*36-\*10 allele multiplications are common in Asians (Johansson et al., 1994; Hosono et al., 2009).

**Figure 1.12: *CYP2D6* gene copy number variations in the Japanese and Chinese Coriell diversity panel.**



Source: data from Hosono et al., (Hosono et al., 2009).

The phenotypic consequence of CNVs depend on the *CYP2D6* allele involved in duplication (Table 1.6). For example, based on the metabolism of the *CYP2D6* probe drug dextromethorphan, it has been shown that individuals with *CYP2D6*\*2 duplications have increased enzyme activity when compared to carriers of a single copy of *CYP2D6*\*2 allele (Aklillu et al., 1996). However, individuals with *CYP2D6*\*10 duplications have similar enzyme activity as the individuals with a single copy of *CYP2D6*\*10 allele (Ishiguro et al., 2004). Thus, the carriers of multiple copies of functional *CYP2D6* alleles (\*1xN, \*2xN) are classified as UMs, the carriers of \*10x2 are still considered to be IMs and carriers of nonfunctional \*4x2 are still PMs (Table 1.6). Little is known about the phenotypic consequence of other multiple copy arrangements including \*36-\*10-\*10 or \*36-\*36-\*10, etc. Hence, the total number of *CYP2D6* copies detected by the CNV assays cannot reasonably be used to estimate the phenotypic consequence, without identifying the *CYP2D6* allele involved in the duplication.

**Table 1.6: Consequences of *CYP2D6* gene copy number variations.**

Variation	<i>CYP2D6</i> allele	Enzyme activity	Phenotype group	Ethnicity
Gene deletion	*5	None	PM	AA, CA, As
Gene multiplication	*4xN	None	PM	CA
	*1xN, *2xN, *35xN	Increased	UM	AA, CA, As
	*41xN, *45xN	Increased	EM <sup>a</sup>	CA, AA
	*36-*10	Decreased	IM	As, AA
	(*36) <sub>n</sub> -( <sup>*</sup> 10) <sub>n</sub>	Unknown	Unknown	As

Note: PM - poor metabolizer; IM - intermediate metabolizer; EM - extensive metabolizer; UM - ultrarapid metabolizer; AA - African Americans; CA - Caucasians; As - Asians; <sup>a</sup> considered to be UM in some publications (Zanger et al., 2004; Zineh et al., 2004) and EM in others (Borges et al., 2010). Source: [www.cypalleles.ki.se/cyp2d6.htm](http://www.cypalleles.ki.se/cyp2d6.htm).

The determination of *CYP2D6* gene copy numbers is complicated by the presence of two highly homologous pseudogenes, *CYP2D7* and *CYP2D8*. It is also complicated by the presence of *CYP2D6-CYP2D7* fusion genes. For example, *CYP2D6*\*36 includes the 5' end of the *CYP2D6* gene fused with the 3' end of the *CYP2D7* gene due to gene recombination. The genome wide association study (GWAS) arrays have relatively few *CYP2D6* probes and therefore are inefficient for accurate identification of allele specific *CYP2D6* CNVs. Furthermore, because of the large sizes of gene deletion and duplications, the *CYP2D6* copy number estimation commonly involves relatively difficult techniques like long-template PCR and southern blotting (Meijerman et al., 2007).

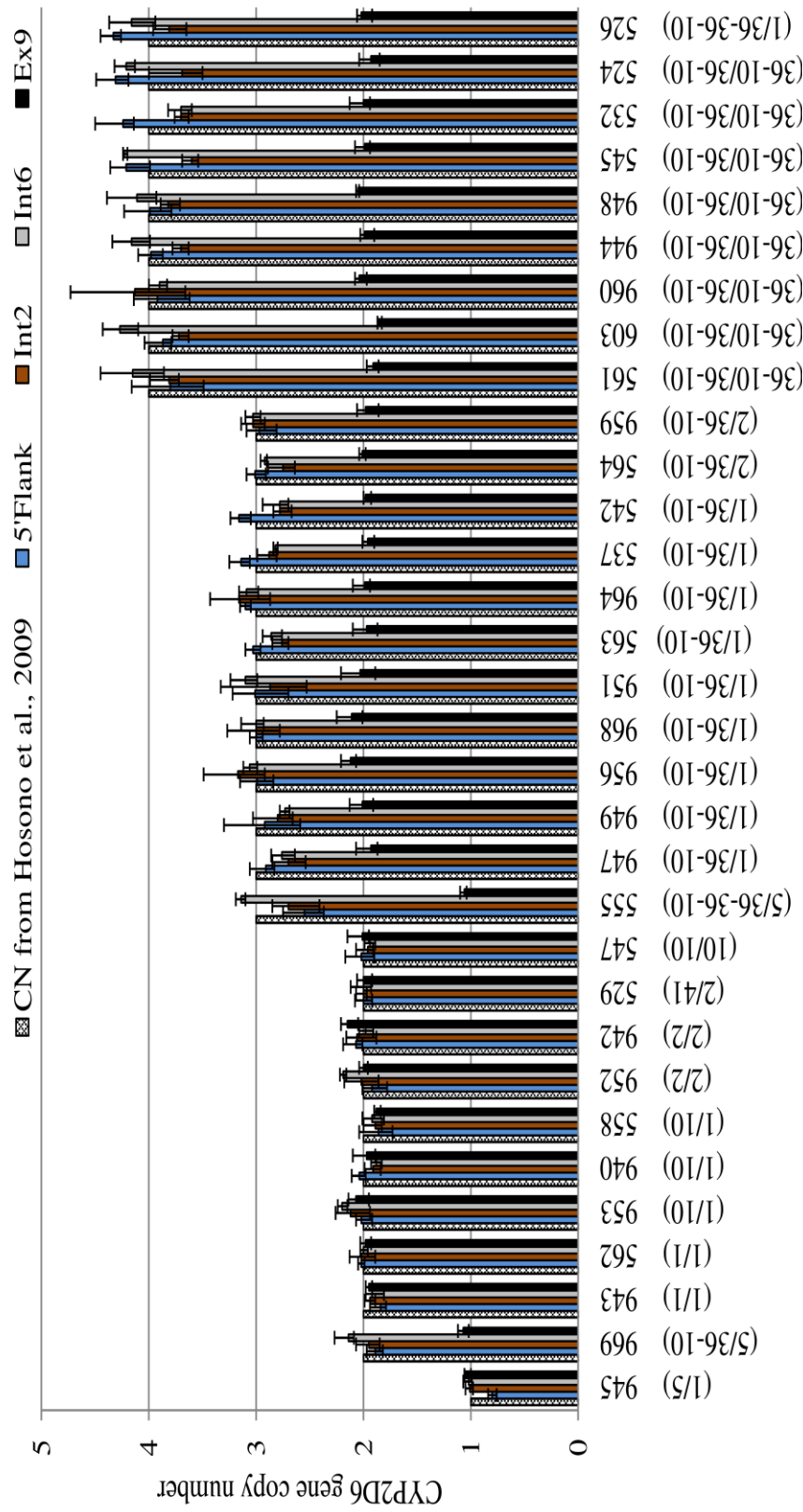
Recently, several qRT-PCR assays have been reported (Schaeffeler et al., 2003; Meijerman et al., 2007; Hosono et al., 2009). Commercial qRT-PCR assays have also become available; however, since the exact positions of the target sequences are not

provided for these assays, their ability to distinguish between specific alleles and fusion genes must be determined experimentally. For example, when *CYP2D6* gene copy number was estimated in 32 Asian individuals from the Coriell Japanese-Chinese human diversity DNA panel using four different qRT-PCR copy number assays, not all the assays yielded concordant results (Ramamoorthy et al., 2010). These copy number assays targeted different regions of the *CYP2D6* gene: (i) 5'-flanking region (5' flank) assay and (ii) intron 2 (Int2) assay, both available commercially from Applied Biosystems (Foster City, CA, USA), (iii) intron 6 (Int6) assay and (iv) exon 9 (Ex9) assay, both assays reported previously in the literature (Hosono et al., 2009). Three of the *CYP2D6* CNV assays (5'-flank, Int2, and Int6) estimated *CYP2D6* copy numbers that were concordant for all 32 subjects (Figure 1.13). However, the Ex9 assay results were concordant in only 10 of 32 samples. The 10 concordant samples did not contain any *CYP2D6*\*36 alleles and the 22 discordant samples contained at least one *CYP2D6*\*36 allele. Also, the Ex9 assay accurately quantified all of the non-*CYP2D6*\*36 alleles in all samples. The specific target site of the Ex9 assay was verified by sequencing the PCR amplicon and the result indicated that the assay targets a region of *CYP2D6* exon 9 that undergoes partial gene-conversion in the *CYP2D6*\*36 allele. These results indicate that the *CYP2D6* CNV assays can be specific for subsets of *CYP2D6* alleles and the decision of which CNV assays to use must include consideration of the alleles that are being targeted. In conclusion, all copy number assays may not yield similar results and have to be selected based on the region of the gene that they target. This will likely also be important for other genes as the genomic structure of the gene copies are being elucidated.



The presence of many copies of *CYP2D6* gene and the involvement of the different alleles (e.g., *CYP2D6*\*1, \*2, \*10, \*35, etc.) makes it difficult to accurately infer the diplotypes present in the two homologous chromosomes. In order to predict the complex haplotypes within a chromosome, bioinformatic algorithms like CNV phaser (Kato et al., 2008) can be used. This program uses data from genotyping assays and gene copy number assays, that is, *CYP2D6* allelic status and *CYP2D6* copy units, to predict the *CYP2D6* diplotype combinations.

Figure 1.13: Comparison of estimated *CYP2D6* gene copy numbers using four copy number assays.



Note: *CYP2D6* copy numbers (y-axis) were estimated using CNV assays that target the 5' flanking region (blue), intron 2 (brown), intron 6 (gray), and exon 9 (black). Also included are the known copy numbers (cross hatched) from Hosono et al., (Hosono et al., 2009). The sample identification numbers (x-axis) are those from the Coriell JCH diversity panel (each of the numbers are preceded by an NA18). The *CYP2D6* genotype is given within the parenthesis for each sample.

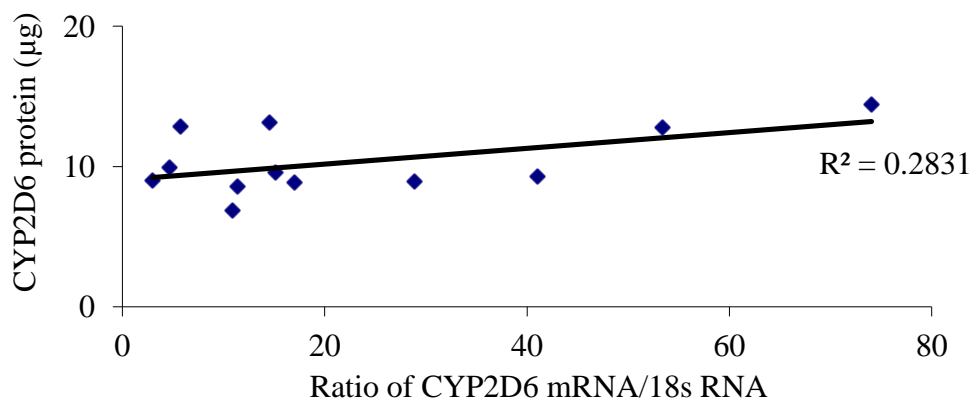
A number of other genes involved in drug metabolism also exhibit CNVs - *UGT2B17* and *UGT2B28* (McCarroll et al., 2006), *SULT1A1* (Hebbring et al., 2007), *GSTT1* and *GSTMI* (Bolt and Thier, 2006). In fact, CNVs are more common than originally thought before the human genome project (Iafate et al., 2004; Sebat et al., 2004). In the human genome, there are thousands of copy number variable regions (CNVRs) that range from 100 bp to several Mb (Redon et al., 2006). Studies have suggested an increasing role for CNVs in many diseases including cancer, developmental diseases, mental illness, autoimmune diseases and infectious diseases (Shrestha et al., 2009; Wain et al., 2009). However, though a recent genome-wide association study identified an association between gene CNVs and eight common human diseases (e.g., Crohn's disease, rheumatoid arthritis, type 1 diabetes and type 2 diabetes), it concluded that CNVs are unlikely to contribute greatly to the genetic basis of common human diseases (Craddock et al., 2010). The researchers speculated that SNPs rather than CNVs may be associated with the pathogenesis of these diseases. That is, even though CNVs are present frequently in individuals, their phenotypic consequence may depend on the sequence of the DNA that is duplicated.

### **3. Epigenetic factors**

Genetic and environmental factors only explain a part of the variability in CYP2D6 enzyme activity. Epigenetic factors like DNA methylation and microRNAs may explain some of this unexplained interindividual variability. For example, there is a lack of statistically significant correlation between CYP2D6 mRNA and protein

expression in human livers, even with the same *CYP2D6* genotype group (e.g., *CYP2D6*\*1/\*1; Figure 1.14; Borges et al., unpublished).

**Figure 1.14: Lack of association between *CYP2D6* mRNA and protein expression in the human livers.**



Note: The *CYP2D6* mRNA level was not correlated with the protein level in *CYP2D6* wildtype (\*1/\*1) livers (N = 11) in a panel of 54 human livers. The *CYP2D6* and 18s mRNA levels were measured by specific RT-PCR assays and the ratio was multiplied by  $10^6$  to simplify data presentation. The protein levels were quantified by western blots using *CYP2D6* specific antibody. Source: Borges et al., unpublished.

This poor correlation between *CYP2D6* mRNA and protein suggests that *CYP2D6* mRNA may be post-transcriptionally regulated by factors like p-bodies (sites in the cytoplasm where mRNA is sequestered until translation or degradation), or microRNAs, etc. Recent studies have shown that a number of CYP450 enzymes as well as some of their transcriptional factors are regulated by microRNAs (Tsuchiya et al., 2006; Takagi et al., 2008; Pan et al., 2009). *However, the role of miRNAs in the regulation of CYP2D6 is still not known.*

## **Introduction to microRNAs**

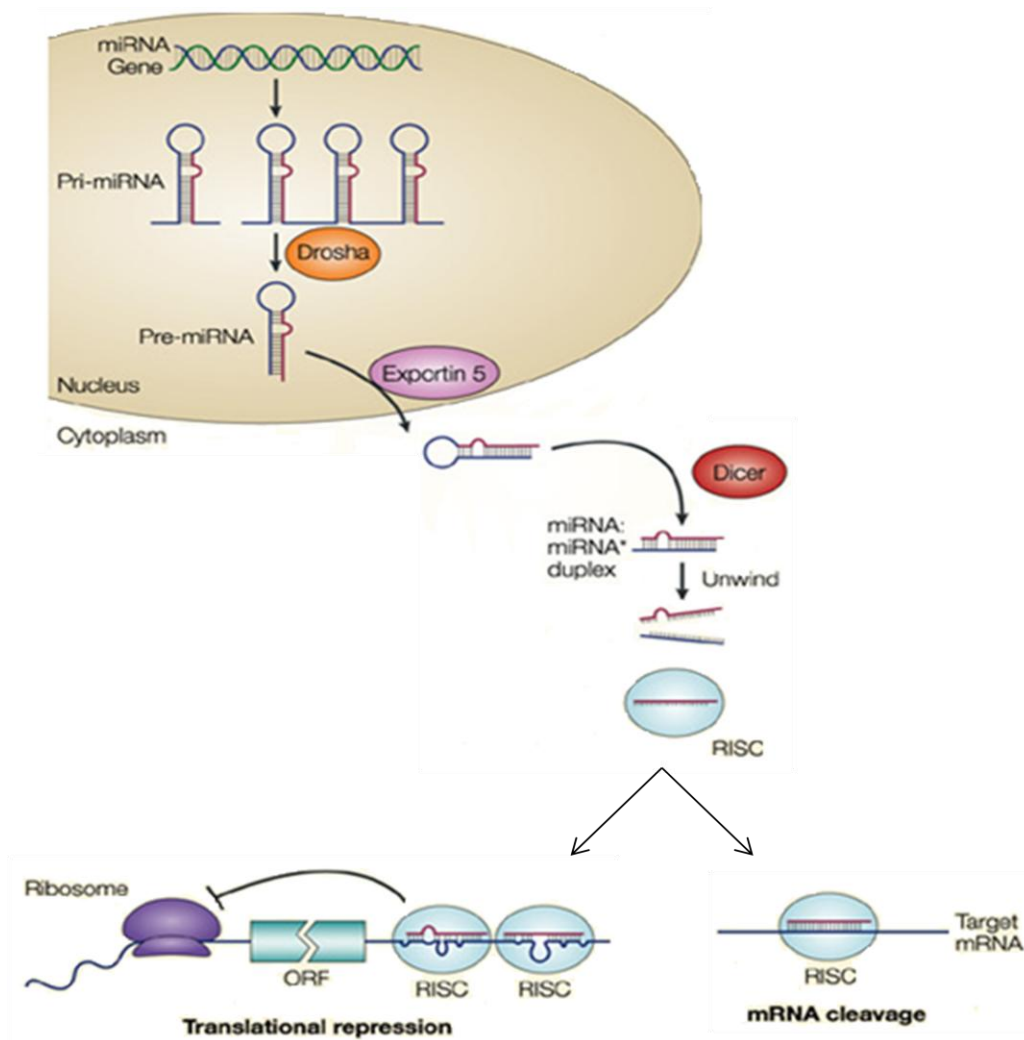
MicroRNAs (miRNAs) are small (18-25 nucleotides), noncoding RNAs that bind typically to the 3'-UTR of a messenger RNA (mRNA) and negatively regulate gene expression. These miRNAs regulate gene expression by one of the two following mechanisms: by blocking protein translation or by degrading the messenger RNA (Olsen and Ambros, 1999; Ambros et al., 2003; Figure 1.15). So far, 940 mature miRNAs have been reported in the human genome (miRBase Registry; version 15.0; Griffiths-Jones et al., 2006). These miRNAs are estimated to control 20-90% of the human genes (Lewis et al., 2005; Miranda et al., 2006). Each miRNA can regulate multiple protein coding genes; therefore, having hundreds of miRNAs offers an enormous potential for post-transcriptional regulation.

## **MicroRNA gene processing**

The initial steps of miRNA gene processing are similar to that of the protein coding genes; however, there are many differences in the later stages of processing. In the nucleus, genes encoding miRNAs are transcribed primarily by the RNA polymerase II enzyme to form pri-miRNA transcripts (200-300 bp precursors) that have a 5' cap and a 3' poly-A tail. These are then processed by the RNase III enzyme complex, Drosha-Pasha/DGCR8, to form 70-80 bp long pre-miRNAs (Gregory et al., 2004; Han et al., 2004). These pre-miRNAs are then exported by exportin5 to the cytoplasm (Lund et al., 2004) where they are processed by another RNase III enzyme, Dicer, to form double stranded, mature miRNA (18-25 bp). One strand of the mature miRNA is then incorporated into a RNA-induced silencing complex (RISC; Du and Zamore, 2005). This

RISC complex then targets the miRNA complementary regions within the mRNA 3'-UTR. This results in translational repression either by direct mRNA degradation or more commonly by blocking protein synthesis (Figure 1.15).

**Figure 1.15: Gene processing and mechanisms of action of microRNAs.**



Source: modified from He and Hannon, 2004.

## **Functions of microRNAs**

MicroRNAs are involved in a wide range of biological activities (John et al., 2004) including: cell differentiation, cell death, noncancerous diseases (diabetes, viral infection) and cancer (breast cancer, colon cancer, glioblastoma, etc.). MiRNAs can act as oncogenes or tumor suppressor genes and hence play an important role in tumorigenesis (Esquela-Kerscher and Slack, 2006). Amplification or loss of miRNA genes has been reported in a number of cancers subtypes; these altered miRNA expression patterns affect cell cycle and survival (Calin and Croce, 2006). Aberrant miRNA expression patterns act as ‘signatures’ and allow for better classification of tumor type, progression and prognosis (Chen and Stallings, 2007).

As more miRNAs are being identified and studied, newer functions are being recognized. Previously, miRNAs were thought to function by binding to the mRNA 3’-UTR and regulating gene expression by either mRNA degradation or by blocking protein synthesis (Bartel, 2004). New evidence now indicates that miRNAs are also involved in induction of gene expression. For example, hsa-miR-373 can bind to a highly complementary region in E-cadherin promoter sequence and readily induce gene expression (Place et al., 2008). Similarly, mmu-miR-10a can interact with the 5’-UTR of mRNAs encoding ribosomal proteins to enhance their translation (Orom et al., 2008). It has now been shown that miRNAs can also bind to coding regions and repress gene expression (Duursma et al., 2008); this mechanism may explain some of the differential expression seen in mRNA splice variants.

## Pharmacological relevance of microRNAs

An important role for miRNAs in drug metabolism and drug resistance is beginning to emerge. For example, ATP-binding cassette xenobiotic transporter ABCG2 (To et al., 2008), cytochrome P450 1B1 (Tsuchiya et al., 2006), cytochrome P450 3A4 (Pan et al., 2009) and the transcription factor pregnane X receptor (Takagi et al., 2008) have been shown to be regulated by miRNAs. This miRNA regulation of DMEs may explain some of the interindividual variability that is seen in drug response (Ingelman-Sundberg et al., 2007).

Cytochrome P450 1B1 (CYP1B1) is a member of the CYP450 superfamily that is constitutively expressed in many tissues including breast and ovary. In the estrogen target tissues, CYP1B1 catalyzes the biotransformation of 17 $\beta$ -estradiol to 4-hydroxyestradiol that causes DNA damage. In a number of cancers, including breast cancer, there is an increased expression of CYP1B1 protein. Recently, CYP1B1 has been shown to be post-transcriptionally regulated by hsa-miR-27b by blocking protein synthesis (Tsuchiya et al., 2006). This was the first CYP450 that was shown to be regulated by miRNAs.

Pregnane X receptor (PXR; NR1I2) is a transcription factor that regulates the expression of a number of DMEs and drug transporters. PXR expression is regulated by hsa-miR-148a post-transcriptionally by blocking protein synthesis (Takagi et al., 2008). This down-regulation in PXR expression affected both constitutive and inducible expression of Cytochrome P450 3A4 (CYP3A4). Even though bioinformatic analyses predicted that hsa-miR-148a can directly regulate CYP3A4, the *in vitro* experiments did not support the bioinformatic predictions. Hence, hsa-miR-148a appears to regulate



CYP3A4 expression not directly, but indirectly through the transcription factor, PXR.

This miRNA regulation of CYP3A4 may explain some of the interindividual variability seen in response to CYP3A4 substrate drugs.

CYP3A4 is a member of the CYP450 superfamily of DMEs that is involved in the metabolism of a number of clinically important drugs (e.g., antiarrhythmics, benzodiazepines, HIV antivirals, calcium channel blockers, etc.; Source: <http://medicine.iupui.edu/flockhart/table.htm>). Both CYP3A4 mRNA and protein expression is regulated by hsa-miR-27b (Pan et al., 2009). Vitamin D receptor (VDR/NR1H1), a transcriptional factor that regulates CYP3A4 expression is also post-transcriptionally regulated by hsa-miR-27b, thus contributing to an increased repression of CYP3A4 protein. This direct and indirect miRNA regulation of CYP3A4 may also explain some of the CYP3A4 mediated interindividual variability seen in drug response.

Apart from drug metabolizing enzymes and their transcriptional factors, a number of other pharmacologically important targets have been shown to be regulated by miRNAs. For example, estrogen receptor alpha (ER $\alpha$ ; ESR1), a target of several drugs, has been shown to be regulated by hsa-miR-206 (Adams et al., 2007). Hsa-miR-206 binds to the 3'-UTR of ER $\alpha$  and reduces its expression post-transcriptionally. The expression of hsa-miR-206 was also inhibited by ER $\alpha$  agonists suggesting the involvement of a mutually inhibitory feedback loop. This miRNA is also expressed at a higher level in ER $\alpha$  negative breast cancers (Iorio et al., 2005).

## **Drug induced changes in microRNAs**

MicroRNAs are influenced by drug or xenobiotic treatment. For example, treatment with carbon tetrachloride and acetaminophen caused changes in the expression of a number of miRNAs in rat liver (Fukushima et al., 2007). MiRNAs have also been shown to play an important role in drug response and in the development of resistance or sensitivity to drugs. For example:

(1) Suppressing the expression of hsa-miR-21 in a cholangiocarcinoma cell line increased its sensitivity to the chemotherapeutic agent gemcitabine (Meng et al., 2006).

(2) Hsa-miR-221 and hsa-miR-222 target ER $\alpha$  (Zhao et al., 2008) and p27<sup>Kip1</sup> (Miller et al., 2008) and confer *in vitro* resistance to tamoxifen.

(3) MiRNA expression profile is different between fulvestrant-resistant vs. -sensitive MCF7 cells and this may contribute to development of drug resistance (Xin et al., 2009).

(4) A parental gastric cancer cell line (BCL2), when compared to its multidrug resistant progeny had altered expression profile for a small set of miRNAs (Xia et al., 2008).

When some of those miRNAs (e.g., hsa-miR-15b and hsa-miR-16) were over expressed in the multidrug resistant progeny cells, they became sensitized to anticancer drugs, while inhibition of those miRNAs made the parental cell line multidrug resistant.

## **Polymorphisms affect mRNA-miRNA interactions**

MicroRNA SNPs (miRSNPs) are SNPs that occur either in the miRNA or their target site on the mRNAs. These miRSNPs can interfere with mRNA-miRNA interaction by destroying a miRNA target site or by creating a new miRNA target site. Such loss or gain of miRNA targeting by miRSNPs can result in the development of drug resistance.

For example, dihydrofolate reductase (DHFR), an important enzyme involved in the intracellular folate metabolism has been shown to be differentially regulated *in vitro* by a miRSNP (Mishra et al., 2007). In cancer chemotherapeutics, DHFR is a target of the drug methotrexate (MTX). Wildtype DHFR is regulated by hsa-miR-24 through a target site in the DHFR mRNA 3'-UTR. A SNP (829C>T) that occurs near the miRNA target site interferes with the normal pairing of hsa-miR-24 with DHFR mRNA. This results in DHFR over-expression and consequently, methotrexate resistance. Similarly, a SNP in the hsa-miR-206 target site in the ER $\alpha$  mRNA contributes to increased repression ER $\alpha$  protein (Adams et al., 2007). Such miRSNPs may have great clinical application (Mishra et al., 2007).

MiRSNPs can thus act as a loss or gain of function mutation that result in loss of target repression or gain of illegitimate repression, respectively. This can in turn lead to the development of drug resistance. The regulation of the CYP450s by these miRSNPs may explain some of the interindividual variability seen in drug response. A number of bioinformatic algorithms including PolymiRTS (Bao et al., 2007) and Patrocles (Georges et al., 2006) databases have been developed to identify the physiological effect of miRSNPs in the mRNA 3'-UTR.

### **Bioinformatic algorithms predict mRNA-microRNAs interactions**

A number of bioinformatic algorithms are available for predicting mRNA-miRNA interactions. Most of these prediction algorithms have some common rules. These include: (1) 'Seed sequence' - perfect or near perfect complementarity of 7 bases (base pair 2-8 in the 5' region of the miRNA) between mRNA and miRNA, (2) restricted

sequence searches of the 3'-UTRs of mRNAs, and (3) homology between two or three animal genomes.

Some of the important bioinformatic programs include:

1. **miRanda:** miRanda algorithm (John et al., 2004) is based on sequence complementarity and uses position-specific rules for the identification of mRNA-miRNA target pairs. This program requires strict conservation across species (human, mouse and rat; target position within  $\pm 10$  residues of each other); it filters out the predicted target sites that are not well conserved. The program also places emphasis on combinatorial regulation of one mRNA by many miRNAs.

2. **miRBase Targets:** miRBase Targets (Griffiths-Jones et al., 2006) looks for complementarity between mRNA and miRNAs starting at the 5' end of the miRNA. It requires strict complementarity in the 'seed region' and does not predict alignments that have more than one mismatch in this region. Species conservation is important; the target site in the 3'-UTR has to be conserved and detected at the same position by a specific miRNA in at least two orthologous species.

3. **TargetScan:** In order to predict the miRNA targets, TargetScan program (Lewis et al., 2003) relies on 'seed region' identification, energy of binding between the RNA heteroduplex, and mRNA conservation across species (human, mouse and rat). Earlier versions of the program required strict species conservation. However, the later version allows the users to choose the type of conservation: (1) highly conserved, (2) conserved, and (3) poorly conserved.

4. **PicTar:** PicTar algorithm (Krek et al., 2005) predicts miRNAs based on evolutionary conservation of the 3'-UTR sequences across related species. The minimal

number of species is user defined. The CYP450 genes do not appear to be a part of this program's database probably because they are not highly conserved.

5. **RNA22:** RNA22 (Miranda et al., 2006) is a pattern-based program that first identifies the miRNAs target sites in a user defined RNA sequence and then identifies the corresponding miRNA(s). Unlike other programs, RNA22 does not look for species conservation. Hence, it also identifies the miRNA binding sites that are not conserved across species. Also, as the input sequence is user defined, hence this program offers the flexibility of identifying miRNAs that target other regions of a gene – promoter, 5'-UTR, or coding region.

The differences in these algorithms are due to the approaches that are used to deal with the 'loop-outs' that are formed by non-complementary regions, non-Watson and Crick base pairing and the stringency of conservation across species. Four of these programs (miRanda, miRBase Targets, TargetScan and PicTar) use evolutionary conservation parameter. The use of the evolutionary conservation parameter might decrease the noise, but may potentially fail to identify a lot of 'true' mRNA-miRNA interaction because of the stringent species conservation requirement (Miranda et al., 2006). This can be overcome by supplementing with a program like RNA22 that does not require phylogenetic conservation. Because of these inherent differences, different programs often predict different miRNAs to target the same mRNA. Hence, using many programs in conjunction and applying a filter to select miRNAs will result in a comprehensive list of predicted miRNAs that can be tested *in vitro*.

## **Summary**

CYP2D6 is a key drug metabolizing enzyme that is involved in the metabolism of a number of clinically important drugs. There is considerable interindividual variability in CYP2D6 enzyme activity that leads to compromised drug metabolism, which in turn affects the clinical outcome. This variability is due in part to *CYP2D6* gene variants and due to drug-drug interactions that results in CYP2D6 enzyme inhibition. Although these factors explain part of the interindividual variability in CYP2D6 enzyme activity, substantial unexplained variability still remains. Post-transcriptional regulatory factors like microRNAs may explain some the variability; very little is known about the factors contributing to post-natal and post-transcriptional regulation of CYP2D6.

## **Research aims and significance**

### **a. Research aims**

Cytochrome P450 2D6 (CYP2D6) enzyme is involved in the metabolism of 20-25% of drugs commonly used in the clinics, including tamoxifen (Ingelman-Sundberg, 2005). There is considerable interindividual variability in CYP2D6 enzyme activity; between different individuals, the CYP2D6 metabolic activity can vary up to a 1000-fold (Dahl et al., 1992). This results in substantial variability in both beneficial and adverse side effects of CYP2D6 metabolized drugs. Part of this variability can be explained by environmental factors and genetic factors. The environmental factor includes diet and drug-drug interactions. Many clinically important drugs are CYP2D6 enzyme inhibitors (Table 1.1) and consequently affect the metabolism of other CYP2D6 substrate drugs. The genetic factors include both polymorphisms and copy number variations in the

*CYP2D6* gene. *CYP2D6* is a highly polymorphic gene with more than 78 allelic and allelic subvariants reported so far ([www.cypalleles.ki.se/cyp2d6.htm](http://www.cypalleles.ki.se/cyp2d6.htm)). These allelic variants result in proteins with no enzyme activity (PM), reduced enzyme activity (IM), ‘normal’ enzyme activity (EM), or, increased enzyme activity (UM). Deletion and multiplication of *CYP2D6* gene has been reported; the phenotypic consequence of these CNVs will depend on the *CYP2D6* alleles involved. There is considerable variability in the frequencies of both *CYP2D6* polymorphisms and gene copy numbers between different ethnicities.

The genetic and environmental factors only explain a part of the interindividual variability; the other factors that contribute to variability in CYP2D6 enzyme activity are largely unknown. Hence, it becomes important to study CYP2D6 to understand the endogenous and exogenous factors that control its activity.

The *overall hypothesis* is that genetic and epigenetic factors contribute to the variability seen in CYP2D6 mediated drug response. The *specific objective* of this research is to determine the contributions of *CYP2D6* genotype, copy number polymorphism and microRNA to the variability in CYP2D6 mediated drug metabolism. Since hepatic CYP2D6 activity determines the therapeutic exposure to several commonly used drugs (including tamoxifen), a better understanding of the regulation of this enzyme will help explain the variable efficacy and side effects. The *specific objective* will be tested through the following *specific aims*:

**Specific aim 1: To identify common *CYP2D6* genetic variants in Vietnamese and Filipino women with breast cancer and to test the association of these variants with plasma endoxifen concentrations.** The working hypothesis is that in Vietnamese

and Filipino women on tamoxifen therapy, genetic polymorphisms in the *CYP2D6* gene are associated with reduced plasma endoxifen concentration.

**Specific aim 2: To identify *CYP2D6* gene copy number variations in Vietnamese and Filipino women with breast cancer and to test the association of these copy number variants with plasma endoxifen concentrations.** The working hypothesis is that high *CYP2D6* copy number polymorphisms exist in Vietnamese and Filipino women and the presence of these CNVs is associated with altered *CYP2D6* enzyme activity and consequently changes in plasma endoxifen concentration.

**Specific aim 3: To identify microRNAs that target *CYP2D6*.** The working hypothesis is that the variability in the enzyme activity within a genotype may be due to direct or indirect miRNA regulation of *CYP2D6*. That is, in subjects with the same *CYP2D6* genotype, miRNAs regulate *CYP2D6* expression by suppressing *CYP2D6* mRNA translation and thus contribute to the interindividual variability in *CYP2D6* expression.

## **b. Significance**

Published literature has shown that there is interindividual variability in *CYP2D6* enzyme activity. A part of the interindividual variability in drug response is due to environmental and genetic factors. However, the cause of much of the variability remains unknown. This *CYP2D6* mediated interindividual variability has clinical consequences. While some individuals that poorly metabolize (PM) *CYP2D6* substrate drugs have increased risk of adverse side effects from the parental drug, other individuals that metabolize these drugs faster (UM) show poor therapeutic response or increased



toxicity due to high concentration of the metabolites. It has been estimated that the cost of treating patients with extreme CYP2D6 enzyme activity (UM and PM) is on average \$4,000-6,000 per year greater than the cost of treating patients in the EM and IM groups (Chou et al., 2000).

This research project aims to understand the contributions of genetic polymorphisms, copy number variations and microRNAs in the regulation of CYP2D6 enzyme activity. Upon completion of the first and second specific aims, the contribution of *CYP2D6* polymorphisms and copy number variations to CYP2D6 enzyme activity and consequently tamoxifen metabolism in Vietnamese and Filipino women with breast cancer can be determined. This is important because considerable interethnic variability in *CYP2D6* variants have been reported; the phenotypic consequence of *CYP2D6* gene variants has not been studied in these two Asian populations. Also, tamoxifen is more commonly used than AIs in many developing countries in Asia. Hence, these results obtained in the Vietnamese and Filipino population will help to personalize tamoxifen therapy. Upon completion of the third specific aim, the role of microRNAs in the regulation of CYP2D6 would be elucidated for the first time.

The *rationale* for undertaking these studies is that a better understanding of the factors that regulate CYP2D6 will help to better predict the interindividual variability in CYP2D6 mediated drug metabolism. This in turn, will lead to improvements in safety and efficacy of drugs metabolized by CYP2D6, including tamoxifen therapy for breast cancer. Ultimately, the new information generated from this research, when added to already existing literature on CYP2D6 pharmacogenetics, will contribute to the personalization of drug therapy.

## CHAPTER TWO

### **Cytochrome P450 2D6 Variants Predict Endoxifen Concentration in Vietnamese and Filipino Breast Cancer Patients.**

The work presented in this chapter has been prepared in part for submission as 'Cytochrome P450 2D6 and UDP-Glucuronosyl transferase 2B7 genetic variants predict tamoxifen active metabolite, endoxifen, in Vietnamese and Filipino breast cancer patients' by Zereusenay Desta, Anuradha Ramamoorthy, Todd Skaar, Michael Daum, Faouzi Azzouz, Lang Li, David Flockhart, Richard Love, et al.

#### **Abstract**

Tamoxifen is a hormonal therapy that is used globally in the treatment and prevention of breast cancer. Tamoxifen is metabolized to a more active metabolite, endoxifen, by the drug metabolizing enzyme cytochrome P450 2D6 (CYP2D6). Genetic polymorphisms in *CYP2D6* reduce the enzyme activity and consequently reduce endoxifen concentration. These *CYP2D6* variants are also associated with poor clinical outcome. There is considerable interethnic variability in the frequencies of *CYP2D6* gene polymorphisms. In 93 Vietnamese and 144 Filipino premenopausal women diagnosed with breast cancer, the association between *CYP2D6* genotype and endoxifen concentration was analyzed at 4 months after initiation of tamoxifen therapy (20 mg/day orally). In the combined Vietnamese and Filipino cohort, the reduced function allele *CYP2D6\*10* was the most common allele with a frequency of 55.8%. To evaluate the correlation between *CYP2D6* polymorphisms and tamoxifen metabolism, these women were divided into four groups based on their CYP2D6 genotype score. In these four different score groups: (1) EM/EM, (2) EM/IM, (3) IM/IM or EM/PM, and (4) IM/PM, the mean ( $\pm$ SD) plasma endoxifen concentrations were  $145\pm 139$ ,  $106\pm 71$ ,  $66\pm 63$  and

61±31 nM, respectively and the endoxifen/N-desmethyldoxifen ratios were 0.53±1.61, 0.17±0.11, 0.09±0.14 and 0.06±0.02, respectively. Both the endoxifen concentration and endoxifen/N-desmethyldoxifen ratio were significantly different ( $p < .005$ ) among these four groups in the combined population. In conclusion, *CYP2D6* gene variants are associated with lowered plasma concentration of endoxifen in the Vietnamese and the Filipino populations.

## **Introduction**

Tamoxifen is a selective estrogen receptor modulator (SERM) that is used globally in the hormonal therapy of estrogen receptor (ER) positive breast cancers in both pre- and post-menopausal women. It has been repeatedly shown to improve both the disease-free and overall survival in breast cancer patients (Osborne, 1998; Clarke, 2006). Although tamoxifen is a very effective therapy, the pharmacodynamic response to tamoxifen varies considerably between individuals (Lonning et al., 1992; Ingle et al., 1999).

Tamoxifen is metabolized by several of the cytochrome P450 (CYP450) enzymes into a number of metabolites with varying pharmacological activities (Desta et al., 2004). Of these enzymes, CYP2D6 is primarily responsible for the generation of one of the more active metabolite, endoxifen (Desta et al., 2004). Endoxifen is ~30-100 fold more potent than the parent drug tamoxifen (Johnson et al., 2004). It is equipotent to another metabolite, 4-hydroxytamoxifen, but it is present in much higher plasma concentration (Lee et al., 2003; Stearns et al., 2003; Johnson et al., 2004; Jin et al., 2005; Borges et al., 2006). Endoxifen is now being developed as a new therapeutic drug for breast cancer.

Pre-clinical study using cancer cell lines and rodent models has shown that endoxifen can act as a standalone drug and inhibit cellular proliferation (Ahmad et al., 2010); human clinical trials have not been reported yet.

In breast cancer patients undergoing tamoxifen therapy, the plasma concentration of endoxifen is affected by polymorphisms in the *CYP2D6* gene (Stearns et al., 2003; Jin et al., 2005; Borges et al., 2006). *CYP2D6* is a highly polymorphic gene (Ingelman-Sundberg, 2005); more than 78 alleles and allelic subvariants have been described so far (Human Cytochrome P450 Allele Nomenclature Committee: [www.cypalleles.ki.se/cyp2d6.htm](http://www.cypalleles.ki.se/cyp2d6.htm)). The frequencies of many of the *CYP2D6* alleles vary with ethnicity (Bradford, 2002; Sistonen et al., 2007). In Caucasians, the primary cause of poor metabolizer (PM) status is the *CYP2D6*\*4 allele, followed by other alleles including *CYP2D6*\*3, \*5 and \*6; the combined frequency of these alleles is about 20-30% (Bradford, 2002; Sistonen et al., 2007). However, in Asians, *CYP2D6*\*3-\*6 alleles are present in about 5-6% (Bradford, 2002; Sistonen et al., 2007). In Asians, the primary cause of reduced *CYP2D6* activity is the *CYP2D6*\*10 allele. The frequency of the *CYP2D6*\*10 allele is 40-60% in the Asians, but is less than 1% in the Caucasians (Bradford, 2002; Sistonen et al., 2007).

The *CYP2D6* gene polymorphisms that lead to reduced (IM) or poor (PM) enzyme activity are associated with reduced plasma endoxifen concentration (Stearns et al., 2003; Jin et al., 2005; Borges et al., 2006) and poor clinical outcomes (Goetz et al., 2005; Bonanni et al., 2006; Goetz et al., 2007; Gonzalez-Santiago et al., 2007; Schroth et al., 2007; Schroth et al., 2009). Published data on *CYP2D6* genotype-endoxifen concentration or *CYP2D6* genotype-outcome is largely from Caucasian populations. To

date, little information is available from the different Asian populations (Wain et al., 2009). Studies performed in Korea (Lim et al., 2007), Japan (Kiyotani et al., 2008) and China (Xu et al., 2008) have suggested that *CYP2D6\*10* is associated with poorer clinical outcome. Not all studies have confirmed these findings in both Caucasian and Asian populations (Nowell et al., 2005; Wegman et al., 2005; Wegman et al., 2007; Okishiro et al., 2009).

This study is set up as a prospective clinical trial for breast cancer patients undergoing adjuvant tamoxifen therapy in two Asian countries - Vietnam and the Philippines in order to study the association of *CYP2D6* gene polymorphism, tamoxifen metabolism and clinical outcome. As a part of the initial analyses, the association between *CYP2D6* genetic variants and the concentrations of tamoxifen and its metabolites were determined.

## **Results**

### **Patient characteristics:**

The breast cancer patient cohort consisted of 93 Vietnamese and 144 Filipino premenopausal women. These women were diagnosed with clinical stage II-III hormone receptor positive breast cancer by biopsy. They were treated with surgical oophorectomy and tamoxifen. There were no statistically significant difference in demographic characteristics between different the *CYP2D6* genotype groups (Table 2.1).

**Table 2.1: Patient characteristics**

Demographics	<i>CYP2D6</i> score group							
	Score = 2.0		Score = 1.5		Score = 1.0		Score = 0.5	
	Vietnamese	Filipino	Vietnamese	Filipino	Vietnamese	Filipino	Vietnamese	Filipino
Mean age – years (S.D.)	42.5 (4.8)	42.1 (4.1)	44.3 (3.7)	43.3 (3.6)	42.8 (4.4)	42.1 (5.5)	38.7 (7.4)	40.7 (4.2)
Mean weight – Kg (S.D.)	49.2 (5.53)	56.8 (10.9)	55.0 (11.0)	56.2 (8.79)	50.9 (6.44)	47.7 (6.59)	51.0 (5.57)	57.8 (11.4)
Mean BMI (S.D.)	20.7 (2.43)	24.5 (3.86)	20.3 (2.27)	24.2 (3.60)	21.3 (2.72)	23.6 (4.03)	21.4 (1.18)	25.1 (3.86)

### ***CYP2D6* variants and allele frequencies**

Genotyping was performed for 7 different *CYP2D6* alleles including those that are common in Asian and Caucasian population, *CYP2D6*\*2-\*6, \*10, and \*41. The *CYP2D6* alleles that did not carry any of the variations mentioned above were designated as *CYP2D6*\*1. The functional *CYP2D6* alleles (*CYP2D6*\*1 and \*2) were designated as extensive metabolizers (EM), reduced function alleles (*CYP2D6*\*10 and \*41) were designated as intermediate metabolizers (IM), and nonfunctional alleles (*CYP2D6*\*4 and \*5) were designated as poor metabolizers (PM).

The *CYP2D6* allele frequencies for both the Vietnamese and the Filipino populations are listed in Table 2.2. The most frequent allele present in both the populations was the IM allele, *CYP2D6*\*10. The *CYP2D6* poor metabolizer alleles that are frequent in Caucasians, *CYP2D6*\*3, \*4, \*6 were absent in both the populations. Most of the allelic frequencies, except *CYP2D6* \*2 in the Filipino population, were in accordance with the Hardy-Weinberg equilibrium (Table 2.2). The frequencies of the genotype combinations, along with their predicted metabolizer phenotype and score are given in Table 2.3.

**Table 2.2: *CYP2D6* allele frequencies in the Vietnamese and Filipino breast cancer patients.**

<i>CYP2D6</i> allele	Frequency (%)		
	Vietnamese cohort	Filipino cohort	Combined cohort
*10	59.1	53.6	55.8
*1	25.3	30.0	28.1
*2	9.7	12.1 <sup>a</sup>	11.2 <sup>a</sup>
*5	4.8	2.1	3.2
*41	1.1	2.1	1.7
*4	0	0	0
*3	0	0	0
*6	0	0	0

Note: <sup>a</sup> *CYP2D6* allele not in Hardy-Weinberg equilibrium (HWE). The HWE was not calculated for *CYP2D6*\*1 and other alleles (*CYP2D6*\*3, \*4 and \*6) that were not present in the Vietnamese and/or Filipino populations.



**Table 2.3: *CYP2D6* genotypes in Vietnamese and Filipino breast cancer patients.**

CYP2D6 metabolizer status	Score	CYP2D6 genotype	Frequency (%)		
			Vietnamese cohort	Filipino cohort	Combined cohort
EM/EM	2	*1/*1	6.5	1.4	3.4
		*1/*2	3.2	8.6	6.4
		*2/*2	1.1	7.9	5.2
EM/IM	1.5	*1/*10	30.1	47.1	40.3
		*2/*10	12.9	0	5.2
		*1/*41	1.1	1.4	1.3
EM/PM	1	*1/*5	3.2	0	1.3
		*2/*5	1.1	0	0.4
IM/IM	1	*10/*10	34.4	26.4	29.6
		*10/*41	1.1	2.9	2.1
PM/IM	0.5	*5/*10	5.4	4.3	4.7

Note: EM – extensive metabolizer with *CYP2D6* functional allele; IM – intermediate metabolizer with *CYP2D6* reduced function allele; PM – poor metabolizer with *CYP2D6* null allele; Undetermined – failure to make a *CYP2D6* genotype call.

#### **Association between *CYP2D6* variants and plasma tamoxifen metabolites**

*CYP2D6* genotype and tamoxifen metabolite concentration data was available for 88 individuals in the Vietnamese cohort and 138 individuals in the Filipino cohort. Only these individuals were included for subsequent genotype to phenotype association analyses.

In the entire cohort of Vietnamese patients, the mean plasma concentrations ( $\pm$  SD) of tamoxifen was  $385 \pm 281$  nM, N-desmethyldtamoxifen (NDM) was  $640 \pm 459$  nM, 4-hydroxytamoxifen (4-OH-tam) was  $4.4 \pm 3.8$  nM, and endoxifen was  $80.9 \pm 63.9$  nM. Similarly, in the entire cohort of Filipino patients, the mean plasma concentrations ( $\pm$  SD) of tamoxifen was  $617 \pm 790$  nM, NDM was  $1059 \pm 1249$  nM, 4-OH-tam was  $5.2 \pm 4.4$  nM, and endoxifen was  $105.5 \pm 96.3$  nM.

For the purpose of this analysis, individuals were classified into four predicted phenotypic groups based on their genotype score (Table 2.3): EM/EM (score = 2), EM/IM (score = 1.5), EM/PM and IM/IM (score = 1), and IM/PM (score = 0.5). The association between the CYP2D6 score groups and tamoxifen and its metabolites concentrations were analyzed by ANOVA. No significant associations were observed between the four CYP2D6 score groups and the concentrations of the parent drug tamoxifen, or two other metabolites, NDM and 4-OH-tam in the Filipino or the combined cohorts. However in the Vietnamese cohort, significant association ( $p < .05$ ) was observed between the four CYP2D6 score groups and 4-OH-tam concentrations, but not for the parent drug tamoxifen or NDM concentrations (Figure 2.1 A-C; Table 2.4 A-C).

The plasma concentration of endoxifen ( $p \leq .001$ ; Figure 2.1 D; Table 2.4 A-C) was significantly associated with CYP2D6 score groups in the Vietnamese and the combined cohort, but not in the Filipino cohort. Similarly, the endoxifen/NDM ratio ( $p < .05$ ; Figure 2.1 E; Table 2.4 A-C) were significantly associated with CYP2D6 score groups in the Vietnamese, Filipino and the combined Vietnamese and Filipino cohorts. The wildtype *CYP2D6* alleles (\*1 and \*2; score group 2) had higher plasma endoxifen concentration and higher endoxifen/NDM ratio when compared to the other CYP2D6

score groups with one or more variant alleles. However, there is still considerable interindividual variability in plasma endoxifen concentrations and endoxifen/NDM ratio even within the same CYP2D6 score group (Figure 2.1 D and E). Based on the regression analyses, these *CYP2D6* gene variants explain 15%, 8% and 10% of the variability in the endoxifen concentration, and 4%, 6% and 30% of the variability in the endoxifen/NDM ratio, in the Vietnamese, Filipino and the combined cohorts, respectively.

**Table 2.4 A-C: Plasma concentrations of tamoxifen and its metabolites in subjects with different CYP2D6 score.**

**(A) Vietnamese cohort**

CYP2D6 score	N	Mean concentration, nM ( $\pm$ SD)				
		tamoxifen	N-desmethyl tamoxifen	4-hydroxy tamoxifen	endoxifen	endoxifen/NDM ratio
2	10	349 $\pm$ 264	458 $\pm$ 403	6.3 $\pm$ 5.1	125.2 $\pm$ 90.1	1.27 $\pm$ 2.98
1.5	40	406 $\pm$ 226	668 $\pm$ 393	4.8 $\pm$ 2.9	100.6 $\pm$ 62.2	0.18 $\pm$ 0.12
1	34	336 $\pm$ 281	609 $\pm$ 502	3.1 $\pm$ 2.7	53.3 $\pm$ 47.5	0.12 $\pm$ 0.20
0.5	4	562 $\pm$ 322	925 $\pm$ 434	4.8 $\pm$ 2.5	61.8 $\pm$ 23.5	0.07 $\pm$ 0.02
<b>p-value</b>		<b>.88</b>	<b>.29</b>	<b>.01<sup>a</sup></b>	<b>.001<sup>a</sup></b>	<b>&lt; .03<sup>a</sup></b>

**(B) Filipino cohort**

CYP2D6 score	N	Mean concentration, nM ( $\pm$ SD)				
		tamoxifen	N-desmethyl tamoxifen	4-hydroxy tamoxifen	endoxifen	endoxifen/NDM ratio
2	25	671 $\pm$ 788	1003 $\pm$ 1046	6.1 $\pm$ 5.5	152.2 $\pm$ 155.8	0.23 $\pm$ 0.14
1.5	67	487 $\pm$ 477	818 $\pm$ 695	4.9 $\pm$ 3.3	109.0 $\pm$ 75.5	0.17 $\pm$ 0.11
1	40	817 $\pm$ 1150	1487 $\pm$ 1908	5.3 $\pm$ 5.5	77.1 $\pm$ 72.4	0.07 $\pm$ 0.04
0.5	6	519 $\pm$ 455	1141 $\pm$ 654	4.6 $\pm$ 2.9	61.1 $\pm$ 37.9	0.05 $\pm$ 0.01
<b>p-value</b>		<b>.45</b>	<b>.09</b>	<b>.58</b>	<b>.08</b>	<b>&lt; .0001<sup>a</sup></b>

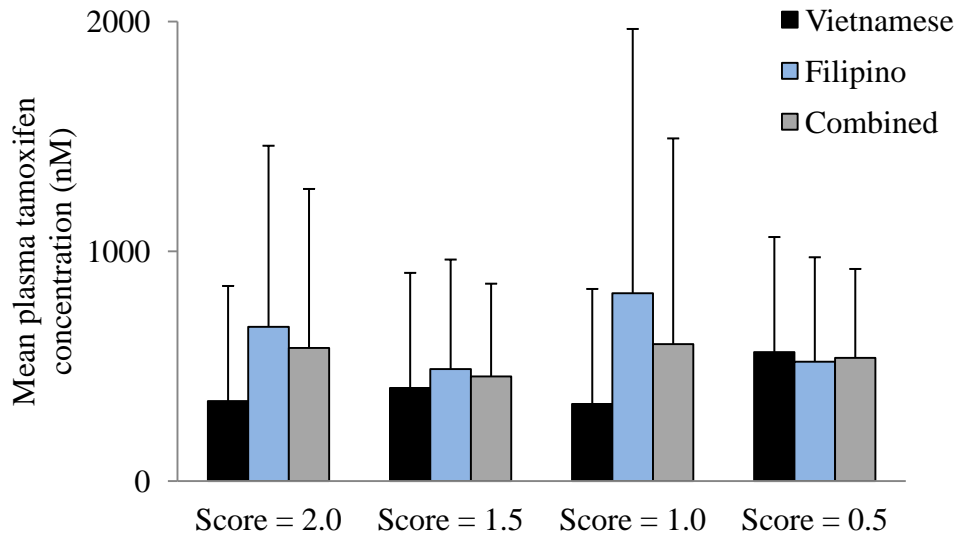
**(C) Combined cohort**

CYP2D6 score	N	Mean concentration, nM ( $\pm$ SD)				
		tamoxifen	N-desmethyl tamoxifen	4-hydroxy tamoxifen	endoxifen	endoxifen/NDM ratio
2	35	579 $\pm$ 692	848 $\pm$ 937	6.1 $\pm$ 5.3	144.5 $\pm$ 139.4	0.53 $\pm$ 1.61
1.5	107	456 $\pm$ 403	762 $\pm$ 602	4.8 $\pm$ 3.1	105.9 $\pm$ 70.7	0.17 $\pm$ 0.11
1	74	596 $\pm$ 895	1083 $\pm$ 1501	4.3 $\pm$ 4.5	66.2 $\pm$ 63.0	0.09 $\pm$ 0.14
0.5	10	536 $\pm$ 387	1005 $\pm$ 559	4.7 $\pm$ 2.6	61.4 $\pm$ 31.4	0.06 $\pm$ 0.02
<b>p-value</b>		<b>.63</b>	<b>.11</b>	<b>.06</b>	<b>&lt; .0001<sup>a</sup></b>	<b>&lt; .005<sup>a</sup></b>

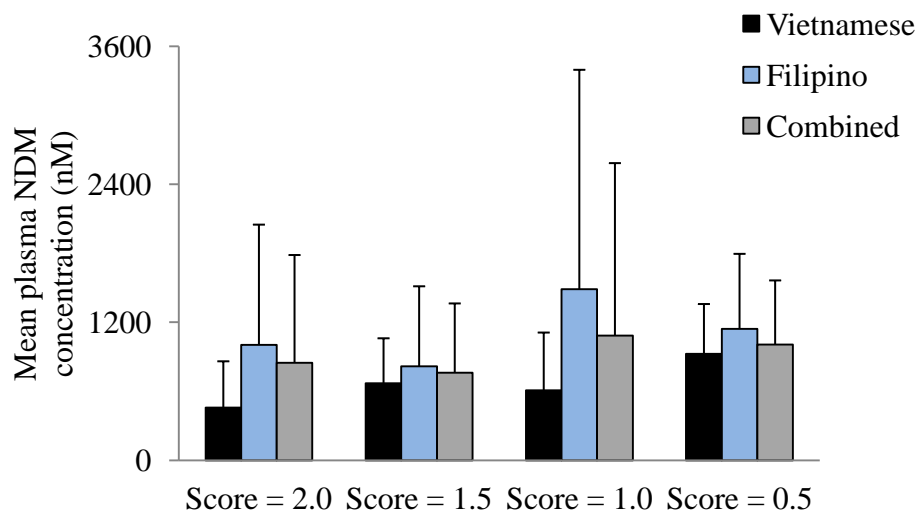
Note: Mean plasma concentration  $\pm$  SD for tamoxifen and its metabolites, N-desmethyltamoxifen, 4-hydroxytamoxifen, endoxifen, endoxifen/NDM ratio in (A) the Vietnamese cohort (N = 88 individuals), (B) the Filipino cohort (N = 138 individuals), and (C) the combined cohorts (N = 226 individuals) based on the CYP2D6 score; <sup>a</sup> p-value < .05 based on ANOVA.

**Figure 2.1 A-E: Association between *CYP2D6* gene score and tamoxifen and its metabolites concentrations.**

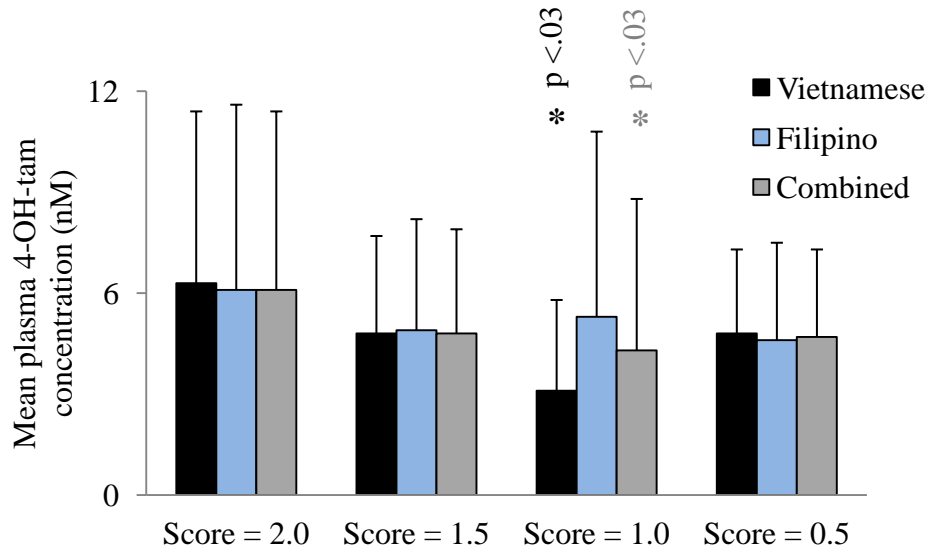
**(A) Tamoxifen**



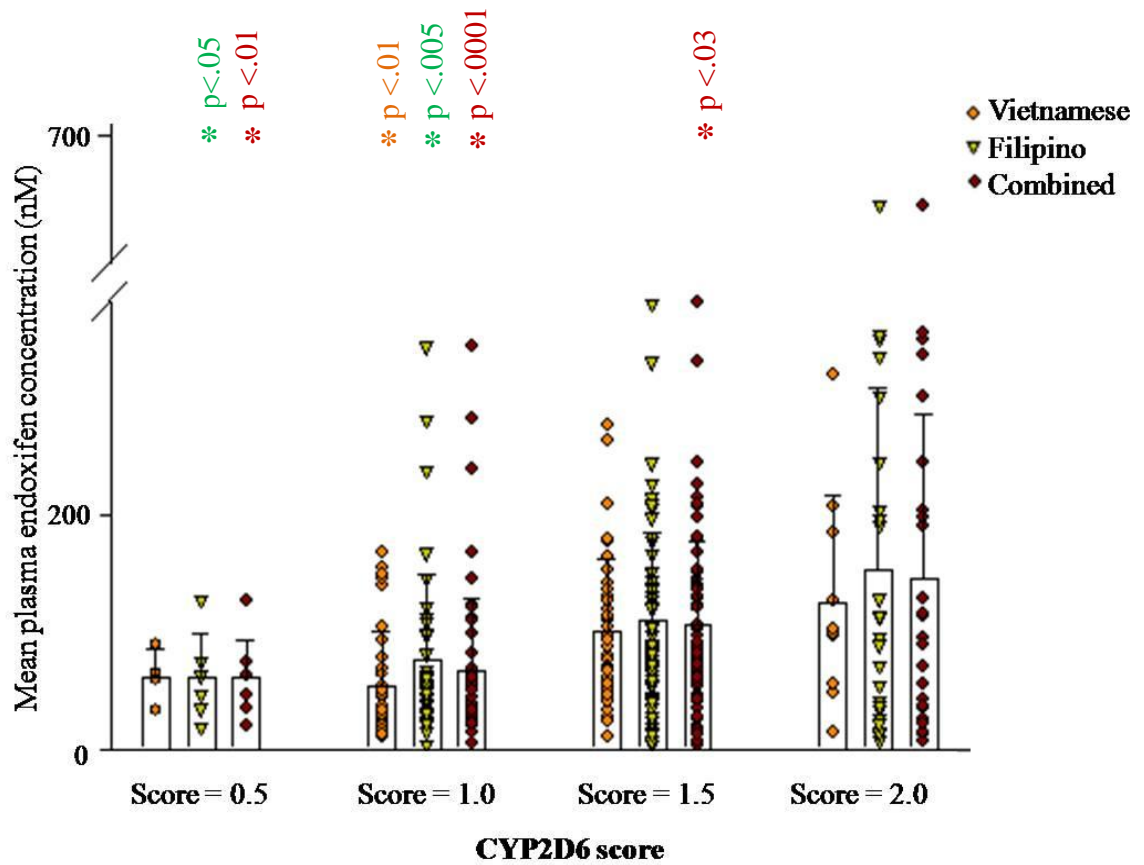
**(B) N-desmethyltamoxifen**



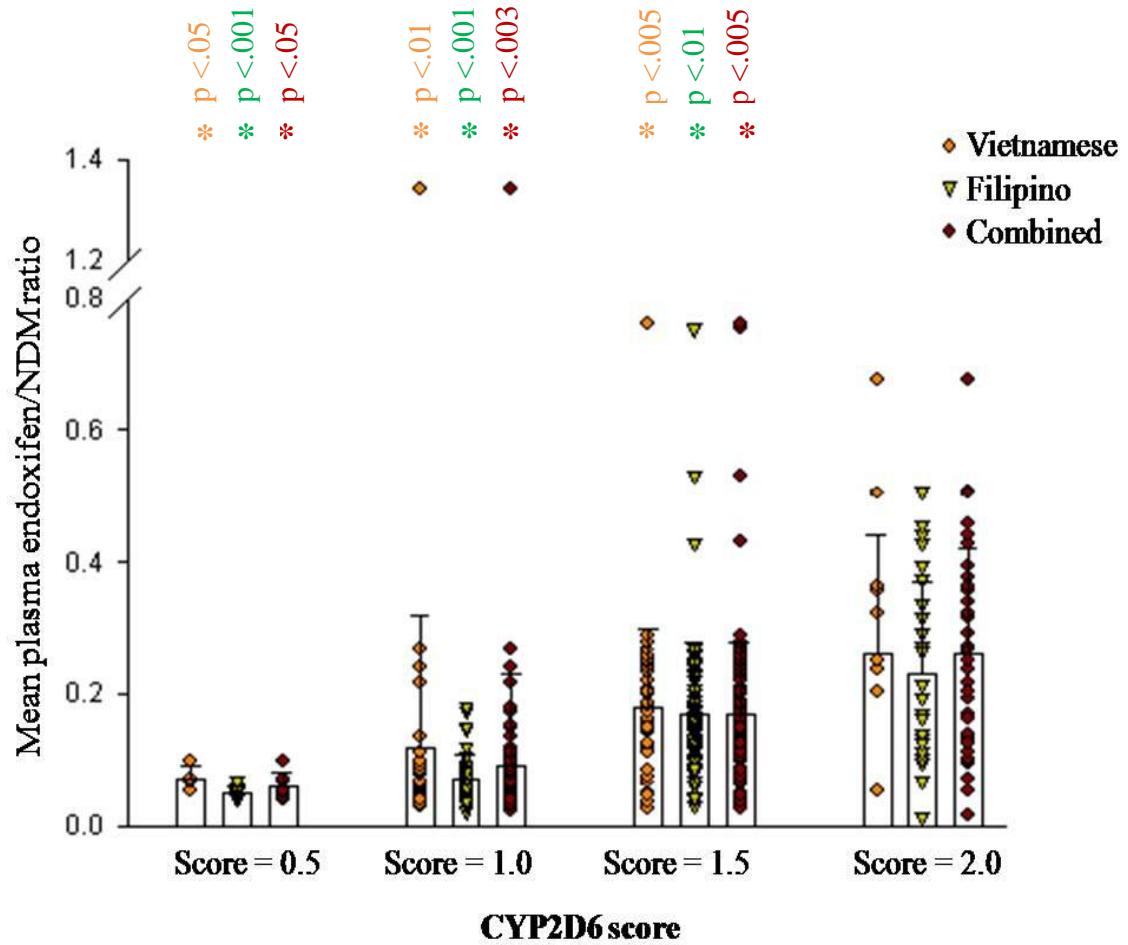
**(C) 4-hydroxytamoxifen**



**(D) Endoxifen**



**(E) Endoxifen/N-desmethyltamoxifen ratio**



Note: The effect of CYP2D6 score groups on tamoxifen and its metabolites concentration in Vietnamese (N = 88 individuals), Filipino (N = 138 individuals), and combined cohorts (N = 226 individuals) after tamoxifen therapy (20 mg/day) for four months. The solid bars represent mean + SD of (A) tamoxifen; (B) N-desmethyltamoxifen; (C) 4-hydroxytamoxifen; (D) endoxifen concentration in the Vietnamese (orange circle), Filipino (green triangle), combined cohort (brown diamond) based on the CYP2D6 score; (E) endoxifen/N-desmethyltamoxifen ratio (with one outlier removed) in the Vietnamese (orange diamond), Filipino (green triangle), and combined cohort (brown diamond); \* indicates  $p < .05$  on a pair-wise comparison with the CYP2D6 score group 2.0 (wildtype).

## Discussion

Worldwide, tamoxifen is the standard endocrine therapy for estrogen receptor positive breast cancer. However, efficacy and side effects of this therapy are not predictable for individual patients. Recent data suggest that genetic differences among individuals in tamoxifen metabolizing enzymes, primarily, *CYP2D6* may account for significant variability in response to this drug (Goetz et al., 2005; Bonanni et al., 2006; Goetz et al., 2007; Lim et al., 2007; Schroth et al., 2007; Kiyotani et al., 2008; Xu et al., 2008; Schroth et al., 2009). Most of *CYP2D6* genotype-clinical outcome studies have been primarily focused on the Caucasian population poor metabolizer allele, *CYP2D6\*4* (Goetz et al., 2005; Bonanni et al., 2006; Goetz et al., 2007; Schroth et al., 2007; Schroth et al., 2009). These studies showed that when compared to EMs, the PMs had fewer adverse events but, increased frequency of cancer recurrence. Similarly in Asian women, three studies have reported an association between the *CYP2D6* intermediate metabolizer allele, *CYP2D6\*10* and poor clinical outcome (Lim et al., 2007; Kiyotani et al., 2008; Xu et al., 2008). Together, these studies have suggested that in the prevention, adjuvant and metastatic settings, PMs and IMs have an increased rate of cancer recurrence and decreased rate of survival. However, these results have not been confirmed by all the studies (Nowell et al., 2005; Wegman et al., 2005; Wegman et al., 2007; Okishiro et al., 2009). In these studies reporting a negative or a lack of association between *CYP2D6* genotype and clinical outcome, one reason for the discordant result may be that, the EMs discontinued the therapy while, the PMs continued with their therapy. The EMs can have increased side effects when compared to the PMs and this may lead to discontinuation of therapy in the EM group (Rae et al., 2009). However, this explanation cannot be



confirmed as these studies do not report data on drug compliance, side effects and concentrations of tamoxifen and its metabolites. Hence, in order to understand the association between *CYP2D6* gene variants, endoxifen concentration and the clinical outcome, controlled prospective trials are needed.

A prospective clinical trial is being conducted in Vietnamese and Filipino premenopausal women with breast cancer who are on adjuvant tamoxifen therapy. In this study, the association between *CYP2D6* polymorphisms and metabolism of tamoxifen was analyzed. In this study population, the frequency of *CYP2D6* intermediate metabolizers (IM) was higher than the other metabolizer groups (Table 2.1 and 2.2). The reduced function allele *CYP2D6\*10* was the most frequent allele (54-59%) in both the populations. The other alleles that are common in Caucasian population, *CYP2D6\*3*, *\*4*, *\*6* and *\*41* were present in less than 2% frequency (Table 2.1 and 2.2). The frequency of these *CYP2D6* allelic variants in our present study of Vietnamese and Filipino women were similar to the other Asian populations previously reported (Kubota et al., 2000; Cai et al., 2007; Kiyotani et al., 2008; Lee et al., 2009; Veiga et al., 2009).

Since the *CYP2D6* polymorphisms influence the metabolism of tamoxifen to endoxifen in other populations (Stearns et al., 2003; Jin et al., 2005; Borges et al., 2006), the association of the *CYP2D6* polymorphisms with plasma endoxifen concentration in the Vietnamese and Filipino populations was investigated. The results suggested that the reduced function variants, primarily *CYP2D6\*10*, were significantly associated with reduced metabolism of tamoxifen to its active metabolite, endoxifen (Figure 2.1 D and E; Table 2.4). In the combined Vietnamese and Filipino cohorts, both endoxifen concentration and endoxifen/NDM ratio were significantly different ( $p < .005$ ) between

the four CYP2D6 score based groups. These results are consistent with our previous studies in Caucasian women (Stearns et al., 2003; Jin et al., 2005; Borges et al., 2006) and also studies reported by others in Caucasian (Gjerde et al., 2008) and Asian (Lim et al., 2007) women showing that *CYP2D6* genetic variants are associated with reduced plasma endoxifen concentration. The concentrations of both the parent drug tamoxifen and the metabolite NDM were not significantly different among the different score groups. However, the concentration of another metabolite, 4-OH-tam was significantly different in the Vietnamese cohort. This metabolite is generated in part, by the activity of CYP2D6 (Desta et al., 2004) and this result is consistent with another previous study reported in Korean population (Lim et al., 2007).

This study shows that plasma endoxifen concentration is associated with *CYP2D6* genotype; however there is still considerable variability within each *CYP2D6* genotype group (Figure 2.1 D and E). Based on the analyses, *CYP2D6* variants explain approximately 8-15% of the variability in the endoxifen concentration, and 4-30% of the variability in the endoxifen/NDM ratio, in these two populations. Although *CYP2D6* genotype appears to explain a fairly small amount of the variability, it is reproducible and it appears to be important in clinical outcomes. In some cases, this interindividual variability may be explained by other rare *CYP2D6* alleles or other genes involved in endoxifen clearance (e.g., UGT2B7) that were not analyzed.

In conclusion, this study provides the first comprehensive analysis of the association between *CYP2D6* polymorphisms and the concentration of endoxifen in Vietnamese and Filipino women. In many developing countries including Vietnam and the Philippines, tamoxifen is preferred over aromatase inhibitors (AI) because of its low

cost. Hence, a better understanding of the interindividual variability in the benefits and side effects would help to improve the effectiveness and compliance of this life saving drug.

## **Materials and Methods**

**Patients:** Premenopausal women with breast cancer were recruited from Vietnam (Hanoi K Hospital; N = 93) and Philippines (Philippine General Hospital; N = 144).

These women were newly diagnosed with clinical stage II-III hormone receptor positive breast cancer. These women were treated with tamoxifen (20 mg/day orally) adjuvant to surgical oophorectomy. Tamoxifen (Nolvadex) was supplied by AstraZeneca. The study protocol was approved by the respective institutional review boards of all the participating institutions in three countries (Vietnam, Philippines and USA). The patients gave their informed consent to participate in this study.

**Sample processing:** The blood samples were collected after four months of therapy. Based on the half-life of tamoxifen, by four months, the plasma concentration is expected to reach the steady-state. The collected specimen were centrifuged and separated under light-protected conditions into plasma, serum and leukocytes. The specimens were stored at -70°C and shipped in liquid nitrogen canisters to Indiana University, where the samples were analyzed for *CYP2D6* genotype and plasma endoxifen concentrations.

**DNA extraction and quantification:** Genomic DNA was isolated from the leukocytes using the QIAamp DNA Blood Maxi Kit (Qiagen, Valencia, CA, USA) and stored at -80°C until use. The DNA concentrations were measured by Quant-iT DNA

Broad Range Kit (Invitrogen, Carlsbad, CA, USA). The DNA was diluted to 10 ng/μl and used for performing the *CYP2D6* genotyping assays.

**Genotyping the *CYP2D6* variants:** *CYP2D6* is a highly polymorphic gene with more than 78 different alleles and allelic subvariants. In our patient population, the genotyping was performed for *CYP2D6*\*2, \*3, \*4, \*5, \*6, \*10 and \*41 alleles. These alleles were selected because they are expected to cover >95% of the known reduced functional and nonfunctional alleles in the Asian and Caucasian populations (Bradford, 2002; Sistonen et al., 2007; Veiga et al., 2009). Genotyping was performed for \*2-\*4, \*6, \*10 and \*41 using the predeveloped TaqMan Genotyping Assays (Applied Biosystems, Foster City, CA, USA) following the manufacturer's instructions. The dbSNP rs id and the assay catalog numbers are: rs1135840 (C\_\_27102414\_10), rs35742686 (C\_\_32407232\_50), rs3892097 (C\_\_27102431\_D0), rs1065852 (C\_\_11484460\_40) and rs28371725 (C\_\_34816116\_20). The *CYP2D6*\*5 allele was assayed by using long range PCR (XL-PCR) as described previously (Gaedigk et al., 2008). Briefly, a 2.9 Kb PCR product for *CYP2D6*\*5 was generated using the following primers: forward primer: 5'CTCCAGCCTCCACCAGTCCAG and reverse primer: 5'CAGGCATGAGCTAAGGCACCCAGAC. The wildtype *CYP2D6* allele (i.e., one without any deletions) does not produce a PCR product for these primers. A 3.8 kb IL-10 (forward primer: 5'GCATGCACAGCTCAGCACTGC and reverse primer: 5'GCCACCCTGATGTCTCAGTTTCG) was used as an internal control (present in all the samples) to assure that the DNA quality was sufficient to perform the long template PCR. The long range PCR was carried out using JumpStart REDAccuTaq LA DNA polymerase (Sigma, St. Louis, MO, USA) in a 16 μl reaction volume containing 20 ng of

the genomic DNA. The annealing temperature was 68°C and extension time was 4.5 min. The PCR products (3 µl) were analyzed by 1% agarose gel electrophoresis. For the purpose of this study, *CYP2D6* alleles that did not carry any of the variations mentioned above were designated as *CYP2D6\*1*. The alleles were designated as *CYP2D6\*10* if they had the 100T variant, but not the 1846A mutation. Although, 100T is present in *CYP2D6\*36*, in this study, we did not distinguish between *CYP2D6\*10* and *\*36*. As a quality control measure, 10% of randomly selected samples were re-genotyped for each SNP.

***CYP2D6* gene score:** Based on the expected enzyme activity, the *CYP2D6* alleles were assigned a score (Borges et al., 2010). Fully functional *CYP2D6* alleles (*\*1* and *\*2*) were assigned a score of 1, alleles associated with reduced enzyme activity (*\*10* and *\*41*) were assigned a score of 0.5 and the nonfunctional alleles (*\*3*-*\*6*) were assigned a score of 0. The bi-allelic score for *CYP2D6* genotype was the summation of the two individual alleles.

**Assays of tamoxifen and metabolites:** Tamoxifen and its metabolites concentrations were measured by Desta et al., using routine LC-MS/MS method, as a part of the collaborative ongoing study.

**Statistical analysis:** Chi-square tests were performed for the calculation of Hardy-Weinberg equilibrium for all the alleles. The comparison between *CYP2D6* gene score and the plasma concentrations of tamoxifen and its metabolites, N-desmethyltamoxifen, 4-hydroxytamoxifen, endoxifen and the derivative endoxifen/NDM ratio was performed using ANOVA and regression analysis. A p-value of < .05 was considered to be statistically significant.

## CHAPTER THREE

### ***CYP2D6* Copy Number Variations in Vietnamese and Filipino Breast Cancer**

#### **Patients and the Effect of *CYP2D6*\*36 Allele on Endoxifen Concentration.**

The work presented in this chapter has been prepared in part for submission as '*CYP2D6* copy number variations in Vietnamese and Filipino breast cancer subjects: effect of *CYP2D6*\*36 on tamoxifen metabolites' by Anuradha Ramamoorthy, Michael Daum, Zereusenay Desta, Faouzi Azzouz, David Flockhart, Richard Love, Todd Skaar, et al.

#### **Abstract**

The cytochrome P450 2D6 (*CYP2D6*) is an important drug metabolizing enzyme that is involved in the metabolism of 20-25% of commonly prescribed drugs. Copy number variations (CNVs) in the *CYP2D6* gene can affect enzyme activity and consequently, metabolism of many drugs. Multiple *CYP2D6* alleles are involved in CNVs and the enzyme activity depends on the allele involved. CNVs involving *CYP2D6*\*10 and \*36 alleles are common in Japanese and Chinese population; however, their frequencies in many other Asian populations are unknown. To determine *CYP2D6* gene copy numbers, quantitative RT-PCR assays were performed in a clinical trial involving the Vietnamese (N = 93) and Filipino (N = 144) breast cancer patients undergoing tamoxifen therapy. The results indicated that only 39% of the individuals had two copies of the *CYP2D6* gene; 3.8% had only one copy and 57.2% had more than 2 copies of the *CYP2D6* gene. The most common allele involved in CNV was *CYP2D6*\*36-\*10, with a frequency of 31.4%. To determine the phenotypic consequence of these *CYP2D6* CNVs, the plasma concentration of endoxifen in individuals with *CYP2D6*\*10, *CYP2D6*\*36-\*10 and *CYP2D6*\*36-\*36-\*10 alleles were compared. The results suggest that there is a trend of decreased enzyme activity with increased numbers

of *CYP2D6*\*36 copies. However, the activity of *CYP2D6*(\*36)<sub>n</sub>-\*10 allele was not significantly different from that of the *CYP2D6*\*10 allele. The *CYP2D6*(\*36)<sub>n</sub>-\*10 allele does not appear to affect the metabolizer status of an individual, i.e., like the *CYP2D6*\*10 carriers, the individuals with *CYP2D6*(\*36)<sub>n</sub>-\*10 alleles can also be classified as intermediate metabolizers. In conclusion, the Vietnamese and Filipino women have a high *CYP2D6* CNV and this frequently involves the *CYP2D6*(\*36)<sub>n</sub>-\*10 allele. This *CYP2D6*(\*36)<sub>n</sub> CNV allele does not appear to reduce or increase the function of the adjacent *CYP2D6*\*10 allele in these two Asian populations.

## **Introduction**

The cytochrome P450 2D6 (*CYP2D6*) is a member of the cytochrome P450 superfamily of drug metabolizing enzymes that is involved in the metabolism of 20-25% of commonly prescribed drugs (Ingelman-Sundberg, 2005). The *CYP2D6* is a highly polymorphic gene; over 78 alleles and allelic subvariants have been described so far (Human Cytochrome P450 Allele Nomenclature Committee: [www.cypalleles.ki.se/cyp2d6.htm](http://www.cypalleles.ki.se/cyp2d6.htm)). These polymorphisms alter *CYP2D6* enzyme activity and consequently, affect the pharmacokinetics, efficacy and side effects of drugs that are metabolized by it (Ingelman-Sundberg, 2005; Eichelbaum et al., 2006). There is considerable interethnic variability in the population incidence of these polymorphisms (Bradford, 2002; Sistonen et al., 2007). For example, in Caucasians, the polymorphisms that result in complete loss of activity (poor metabolizer - PM), *CYP2D6*\*3-\*6 are present at a relatively high frequency (20-30%) when compared to the Asians (5-6%) (Bradford, 2002; Sistonen et al., 2007). In Asians, the most common allele is

*CYP2D6\*10*, a reduced function allele (intermediate metabolizer - IM) that is present at a frequency of 40-60% (Johansson et al., 1994; Garcia-Barcelo et al., 2000). Currently, the most widely used method (Fukuda et al., 1999; Xu et al., 2008) for genotyping *CYP2D6\*10* involves testing for the presence of the 100C>T variant. However, the 100T nucleotide is also present in *CYP2D6\*36*, a null allele (Chida et al., 2002; Gaedigk et al., 2006; Sakuyama et al., 2008) which is also present in a high frequency in Asians (Hosono et al., 2009). In fact, in Japanese, *CYP2D6\*36-10* gene duplication arrangement was more frequent than *\*10* alone [30% vs. 6% (Soyama et al., 2006) and 24% vs. 9% (Hosono et al., 2008)]. Since the *CYP2D6\*10* and *\*36* alleles have different functional activities, it is important to distinguish between these alleles for predicting an individual's *CYP2D6* phenotype.

Although polymorphisms in *CYP2D6* account for a large portion of the interindividual variability seen in *CYP2D6* mediated drug metabolism, copy number variations (CNV) in the *CYP2D6* gene also appear to be important contributors of variability (Weinshilboum, 2003; Eichelbaum et al., 2006; Yu and Shao, 2009). These *CYP2D6* CNVs are composed of both whole gene deletions and multiplications. The frequencies of these CNVs also have considerable interethnic variability. The frequencies of *CYP2D6\*5* gene deletion (1-7%), *CYP2D6\*1-1* gene duplication (0.2-13%) and *\*2-2* gene duplication (0.5-8%) are comparable in Africans, Caucasians and Asians (Bradford, 2002; Sistonen et al., 2007). However, there is considerable difference in cases of other *CYP2D6* alleles. For example, *CYP2D6\*4-4*, *\*35-35* and *\*41-41* gene duplications are common (0.6-15%) in the Caucasians, while the *CYP2D6\*36-10*



and \*10-\*10 gene duplications are common (20-30%) in the Asians (Bradford, 2002; Soyama et al., 2006; Sistonen et al., 2007; Hosono et al., 2009).

These *CYP2D6* gene CNVs have important phenotypic consequence. For example, *CYP2D6*\*5, i.e., gene deletion results in poor metabolizer (PM) status that severely compromises an individual's ability to metabolize *CYP2D6* substrate drugs. However, the phenotypic consequence of the *CYP2D6* multiple copies depends on the allele that is duplicated (source: [www.cypalleles.ki.se/cyp2d6.htm](http://www.cypalleles.ki.se/cyp2d6.htm)). For example: (a) the carriers of multiple copies of functional *CYP2D6* alleles (\*1xN, \*2xN) have increased enzyme activity and are classified as ultrarapid metabolizers (UM); (b) the carriers of \*36-\*10 or multiple copies of \*10x2 do not have increased enzyme activity when compared to the carriers of \*10 and therefore are still considered to be intermediate metabolizers (IM); and (c) carriers of *CYP2D6*\*36x2 and *CYP2D6*\*4x2 have negligible enzyme activity and are still considered to be poor metabolizers (PM). However, little is known about the phenotypic consequence of other multiple copy arrangements such as \*36-\*10-\*10 and \*36-\*36-\*10.

The *CYP2D6* gene polymorphisms and copy number variations can affect enzyme activity and hence the metabolism of a number of *CYP2D6* substrate drugs including tamoxifen. Tamoxifen is used worldwide in the prevention and treatment of hormone-receptor positive breast cancer. There is considerable interindividual variability in tamoxifen metabolism and response (Lonning et al., 1992; Ingle et al., 1999). Tamoxifen is metabolized into a more active metabolite, endoxifen by the activity of *CYP2D6* (Desta et al., 2004). Polymorphisms and copy number variations in *CYP2D6* gene are associated with reduced plasma endoxifen concentration (Stearns et al., 2003; Jin et al.,

2005; Borges et al., 2006) and poor clinical outcomes (Goetz et al., 2005; Bonanni et al., 2006; Goetz et al., 2007; Gonzalez-Santiago et al., 2007; Lim et al., 2007; Schroth et al., 2007; Kiyotani et al., 2008; Xu et al., 2008; Schroth et al., 2009). However, not all studies have confirmed the association between *CYP2D6* genotype and clinical outcomes (Nowell et al., 2005; Wegman et al., 2005; Wegman et al., 2007; Okishiro et al., 2009). One of the possible explanations for the discordant results in these studies is that the PMs (who generally have fewer side effects) continued with their tamoxifen therapy, while the EMs (who generally have more side effects) discontinued the therapy. About 30-50% of women do not complete the 5 year tamoxifen regimen (Barron et al., 2007); one of the reasons for discontinuation of the therapy is increased side effects in the EMs when compared to the PMs (Rae et al., 2009). However, this cannot be confirmed as these studies do not have data on drug compliance and have not measured the plasma concentrations of tamoxifen and its metabolites. Hence, prospective controlled trials are needed to understand the association between *CYP2D6* gene variants, endoxifen concentration and the clinical outcome of tamoxifen therapy.

Currently, a prospective clinical trial is being conducted involving the Vietnamese and Filipino breast cancer patients undergoing adjuvant tamoxifen therapy. This trial will ultimately be able to address the association between *CYP2D6* variants and clinical outcomes. For the first part of this study, the association between *CYP2D6* gene variants and endoxifen concentration was previously determined (Chapter Two). In that study, the most common allele was *CYP2D6\*10* (~55% frequency) and it was associated with lower plasma endoxifen concentration.

In this current study, based on the evolutionary background of *CYP2D6* in two other Asian populations – Chinese and Japanese (Hosono et al., 2009), the hypothesis was that the Vietnamese and Filipino women may also have frequent *CYP2D6* copy number variations and this may affect the enzyme activity. To date, little is known about the presence of *CYP2D6* CNVs or *CYP2D6\*36* in the Vietnamese and Filipino population. Hence, the objective of this study was to determine the presence of CNVs and the functional consequence of *CYP2D6\*36* CNVs in the Vietnamese and Filipino populations.

## **Results**

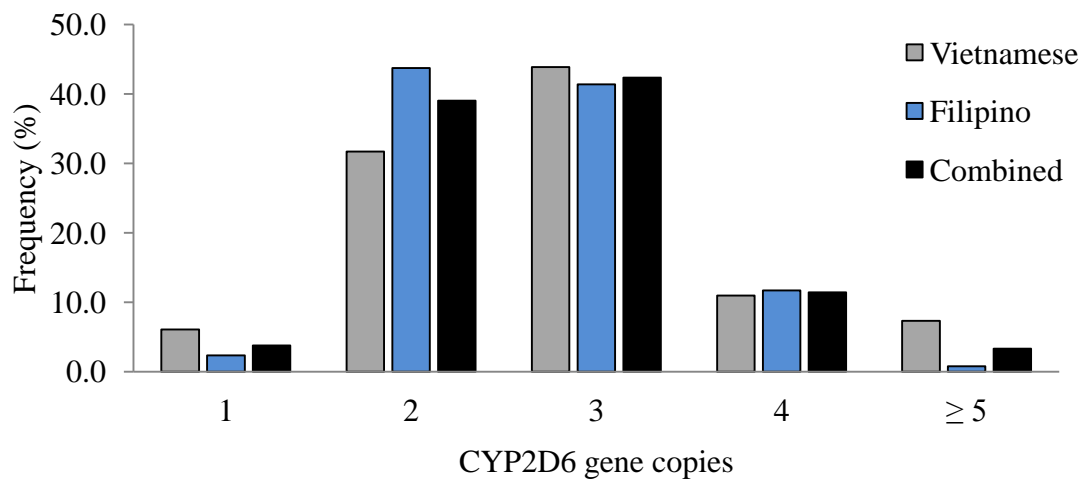
### **Estimation of *CYP2D6* gene copy number variations**

In order to estimate the *CYP2D6* gene copy numbers, *CYP2D6* copy number assays were performed in 237 (N = 93 Vietnamese and N = 144 Filipino) breast cancer patients who were undergoing tamoxifen therapy. Seventeen of these samples were not included in subsequent analyses because: (a) samples failed to amplify, or (b) failure to accurately estimate *CYP2D6* gene copy numbers (e.g., ambiguous copy number reads, or lack of concordance in the *CYP2D6* copy numbers estimated by 5' flank and Int2 assays), or (c) failure to identify the allele involved in the copy number variation. Reliable *CYP2D6* allele and copy number information was available for 210 samples; only these were included in the subsequent analyses.

The results indicated that in our combined cohort, only 39% of the individuals had 2 copies of the *CYP2D6* gene. The rest of the individuals carried a genomic imbalance for *CYP2D6* gene; 3.8% of the individuals had only 1 copy due to gene deletion

(*CYP2D6*\*5), 42.5% had 3 copies, 11.4% had 4 copies, and 3.3% had 5 or more copies of the *CYP2D6* gene (Figure 3.1). The frequencies of these CNVs were comparable for both Vietnamese and Filipino populations (Figure 3.1).

**Figure 3.1: Frequency of *CYP2D6* gene copy number variations in Vietnamese and Filipino breast cancer patients.**



Note: *CYP2D6* gene copy number was estimated using 5' flank qRT-PCR CNV assay along with an internal control, RNase P TaqMan Copy Number Reference Assay. For copy number estimation, each sample was analyzed in triplicate.

### Estimation of *CYP2D6* alleles and diplotypes

The *CYP2D6* diplotypes were estimated using the CNV phaser software (Kato et al., 2008). Three different analyses were performed:

(1) The Vietnamese population (N = 82) and the Filipino population (N = 128) were analyzed separately using the CNV phaser.

(2) Because of *CYP2D6* allelic and CNV frequencies were comparable between the Vietnamese and Filipino population, both the cohorts were combined (N = 210) for the second set of analysis using the CNV phaser.

(3) Since the *CYP2D6* allelic and CNV frequencies in the Vietnamese and Filipino population were comparable to that of the Japanese and Chinese populations reported previously (Hosono et al., 2009), all the four populations were combined (N = 653) by extrapolating data from Hosono et al., (Hosono et al., 2009) for all the *CYP2D6* alleles that were included in our genotype analysis (N = 443).

The results indicated that in the three different analyses, the probability scores for the different diplotypes *were comparable* with the exception of some very rare diplotypes. Since the probability scores were similar, the results from only the combined cohort (N = 210) is reported here (Table 3.1 and 3.2). The most frequent *CYP2D6* alleles involved in copy number variations (Table 3.1) were \*36-\*10 (31.4%), \*5 (3.6%) and \*36-\*36-\*10 (2.4%). The *CYP2D6*\*36 allele was almost always present along with a copy of a \*10 allele (99.8% of the cases). Of the diplotypes with copy number variations (Table 3.2), \*1/\*36-\*10 was the most frequent (24.3%), followed by \*10/\*36-\*10 (11%) and \*36-\*10/\*36-\*10 (9%).

**Table 3.1: *CYP2D6* allelic frequencies in the different cohorts.**

<i>CYP2D6</i> allele	Frequency %		
	Vietnamese cohort	Filipino cohort	Combined cohort
*36-*10 <sup>a</sup>	32.9	30.5	31.4
*1	27.4	28.5	28.1
*10	17.7	21.9	20.2
*2	9.1	11.7	10.7
*5 <sup>a</sup>	4.9	2.7	3.6
*36-*36-*10 <sup>a</sup>	4.3	1.2	2.4
*41	1.2	2.0	1.7
*36-*10-*10 <sup>a</sup>	0.6	0.4	0.5
*1-*1	0.6	0.0	0.2
*10-*10 <sup>a</sup>	0.0	0.4	0.2
*2-*2 <sup>a</sup>	0.0	0.4	0.2
*36-*36 <sup>a</sup>	0.6	0.0	0.2
*36-*36-*36-*10 <sup>a</sup>	0.0	0.4	0.2
*36-*36-*36-*36-*36-*10 <sup>a</sup>	0.6	0.0	0.2

Note: <sup>a</sup> *CYP2D6* alleles that are involved in copy number polymorphism.

**Table 3.2: Frequencies of *CYP2D6* diplotypes in the different cohorts.**

CYP2D6 diplotype	Probability Score <sup>a</sup>	N (%)		
		Vietnamese cohort	Filipino cohort	Combined cohort
*1/*36- <sup>b</sup> *10	1	14 (17.1)	37 (28.9)	51 (24.3)
*1/*10	1	12 (14.6)	19 (14.8)	31 (14.8)
*10/*36- <sup>b</sup> *10	1	11 (13.4)	12 (9.4)	23 (11.0)
*36- <sup>b</sup> *10/*36- <sup>b</sup> *10	0.88	7 (8.5)	12 (9.4)	19 (9.0)
*1/*2	1	3 (3.7)	10 (7.8)	13 (6.2)
*10/*10	1	1 (1.2)	9 (7.0)	10 (4.8)
*2/*2	1	0	9 (7.0)	9 (4.3)
*1/*1	1	5 (6.1)	2 (1.6)	7 (3.3)
*2/*36- <sup>b</sup> *10	1	7 (8.5)	0	7 (3.3)
*36- <sup>b</sup> *10/*36- <sup>b</sup> *36- <sup>b</sup> *10	0.95	5 (6.1)	0	5 (2.4)
*5/*36- <sup>b</sup> *10	1	1 (1.2)	3 (2.3)	4 (1.9)
*1/*5 <sup>b</sup>	1	3 (3.7)	0	3 (1.4)
*5/*10 <sup>b</sup>	1	1 (1.2)	2 (1.6)	3 (1.4)
*10/*41	1	0	3 (2.3)	3 (1.4)
*2/*10	1	3 (3.7)	0	3 (1.4)
*1/*41	1	1 (1.2)	1 (0.8)	2 (1.0)
*2/*5 <sup>b</sup>	1	1 (1.2)	1 (0.8)	2 (1.0)
*1/*36- <sup>b</sup> *36- <sup>b</sup> *10	1	0	2 (1.6)	2 (1.0)
*36- <sup>b</sup> *10/*41 <sup>b</sup>	1	1 (1.2)	1 (0.8)	2 (1.0)
*5/*36- <sup>b</sup> *36- <sup>b</sup> *10 <sup>b</sup>	1	1 (1.2)	1 (0.8)	2 (1.0)
*1/*1- <sup>b</sup> *1 <sup>b</sup>	1	1 (1.2)	0	1 (0.5)
*1/*36- <sup>b</sup> *36- <sup>b</sup> *36- <sup>b</sup> *36- <sup>b</sup> *36- <sup>b</sup> *10 <sup>b</sup>	1	1 (1.2)	0	1 (0.5)
*10/*10- <sup>b</sup> *10 <sup>b</sup>	1	0	1 (0.8)	1 (0.5)
*10/*36- <sup>b</sup> *10- <sup>b</sup> *10 <sup>b</sup>	1	0	1 (0.8)	1 (0.5)
*2/*2- <sup>b</sup> *2 <sup>b</sup>	1	0	1 (0.8)	1 (0.5)
*2/*36- <sup>b</sup> *36- <sup>b</sup> *10 <sup>b</sup>	1	1 (1.2)	0	1 (0.5)
*36- <sup>b</sup> *10/*36- <sup>b</sup> *36 <sup>b</sup>	0.93	1 (1.2)	0	1 (0.5)
*36- <sup>b</sup> *36- <sup>b</sup> *36- <sup>b</sup> *10/*36- <sup>b</sup> *10 <sup>b</sup>	0.61	0	1 (0.8)	1 (0.5)
*5/*36- <sup>b</sup> *10- <sup>b</sup> *10 <sup>b</sup>	1	1 (1.2)	0	1 (0.5)

Note: <sup>a</sup> Probability score reported from the CNV phaser analysis of combined cohort (N = 210); <sup>b</sup> *CYP2D6* diplotypes that are involved in copy number polymorphisms.

## Correlation between *CYP2D6* allele copy numbers and the plasma concentration of tamoxifen and its metabolites

More than 50% of the samples had at least one copy of *CYP2D6*\*36 allele (Table 3.1 and 3.2). Hence, the association between *CYP2D6*\*36 gene copy numbers and the plasma endoxifen concentration was determined. Tamoxifen and its metabolite concentration between were compared in the following groups:

- (a) \*1/\*10, \*1/\*36-*\*10* and \*1/\*36-*\*36-*\*10**,
- (b) \*10/\*10, \*10/\*36-*\*10*, \*36-*\*10*/\*36-*\*10* and \*36-*\*10*/\*36-*\*36-*\*10**, and
- (c) \*5/\*10, \*5/\*36-*\*10* and \*5/\*36-*\*36-*\*10**.

For these comparisons, there was no significant difference in plasma endoxifen concentration or endoxifen/N-desmethyldamoxifen concentration ratio (Table 3 A-C; Figure 3.2 A and B). These results suggest that *CYP2D6*\*36 copy number variation is not significantly associated with increased or decreased *CYP2D6* enzyme activity. Similarly, when the activity of *CYP2D6*\*10 was compared with both \*36-*\*10* and \*36-*\*36-*\*10** alleles combined together [i.e., *CYP2D6*(\*36)<sub>n</sub>-*\*10*], the results indicated that the activity of (\*36)<sub>n</sub>-*\*10* was not significantly different from that of \*10. Similarly, the concentrations of the parent drug tamoxifen and the other metabolites 4-hydroxytamoxifen and N-desmethyldamoxifen were not associated with *CYP2D6* enzyme activity (Table 3.3 A-C).



**Table 3.3 A-C: Plasma concentrations of tamoxifen and its metabolites in subjects with different *CYP2D6* diplotypes involved in *CYP2D6*(\*36)<sub>n</sub> copy number variations.**

**(A) Combined cohort**

<i>CYP2D6</i> diplotype <sup>a</sup>	N	Mean concentration, nM (± SD)				
		Tamoxifen	N-desmethyl tamoxifen	4-hydroxy tamoxifen	Endoxifen	Endoxifen/NDM ratio
*1/ *10	34	463 ± 396	729 ± 531	5.1 ± 3.0	118.2 ± 74.1	0.19 ± 0.13
*1/ *36-*10	58	439 ± 355	790 ± 648	4.5 ± 2.9	95.6 ± 62.1	0.15 ± 0.06
*1/ *36-*36-*10	2	667 ± 40	943 ± 9	5.0 ± 1.3	117.9 ± 28.7	0.12 ± 0.03
*10/ *10	10	1038 ± 1228	1734 ± 1995	6.4 ± 6.3	95.7 ± 83.2	0.07 ± 0.02
*10/ *36-*10	21	602 ± 1128	1076 ± 20037	4.2 ± 5.7	65.2 ± 75.6	0.13 ± 0.26
*36-*10/ *36-*10	18	701 ± 908	1305 ± 1398	4.9 ± 3.9	67.7 ± 56.9	0.07 ± 0.04
*36-*10/ *36-*36-*10	5	325 ± 342	663 ± 571	3.2 ± 3.8	41.0 ± 37.6	0.06 ± 0.01
*5/ *10	3	652 ± 659	1309 ± 993	5.5 ± 3.5	78.4 ± 47.4	0.07 ± 0.02
*5/ *36-*10	4	446 ± 295	844 ± 322	3.6 ± 1.9	48.6 ± 20.9	0.06 ± 0.01
*5/ *36-*36-*10	2	567 ± 379	1057 ± 390	4.8 ± 4.2	62.4 ± 39.1	0.06 ± 0.02

**(B) Vietnamese cohort**

CYP2D6 diplotype <sup>a</sup>	N	Mean concentration, nM ( $\pm$ SD)				
		Tamoxifen	N-desmethyl tamoxifen	4-hydroxy tamoxifen	Endoxifen	Endoxifen/ NDM ratio
*1/ *10	15	459 $\pm$ 237	699 $\pm$ 313	5.6 $\pm$ 3.0	129.7 $\pm$ 71.5	0.22 $\pm$ 0.17
*1/ *36-*10	20	367 $\pm$ 227	655 $\pm$ 480	4.1 $\pm$ 2.5	80.1 $\pm$ 49.8	0.15 $\pm$ 0.08
*1/ *36-*36-*10	0					
*10/ *10	1	81	176	1.0	16.0	0.09
*10/ *36-*10	10	273 $\pm$ 147	480 $\pm$ 455	2.8 $\pm$ 1.9	59.9 $\pm$ 57.4	0.21 $\pm$ 0.36
*36-*10/ *36-*10	6	469 $\pm$ 506	763 $\pm$ 688	4.4 $\pm$ 4.5	51.6 $\pm$ 52.6	0.06 $\pm$ 0.02
*36-*10/ *36-*36-*10	5	325 $\pm$ 342	663 $\pm$ 571	3.2 $\pm$ 3.8	41.0 $\pm$ 37.6	0.06 $\pm$ 0.01
*5/ *10	1	138	329	2.6	32.8	0.10
*5/ *36-*10	1	788	907	2.7	64.6	0.07
*5/ *36-*36-*10	1	835	1333	7.8	90.1	0.07

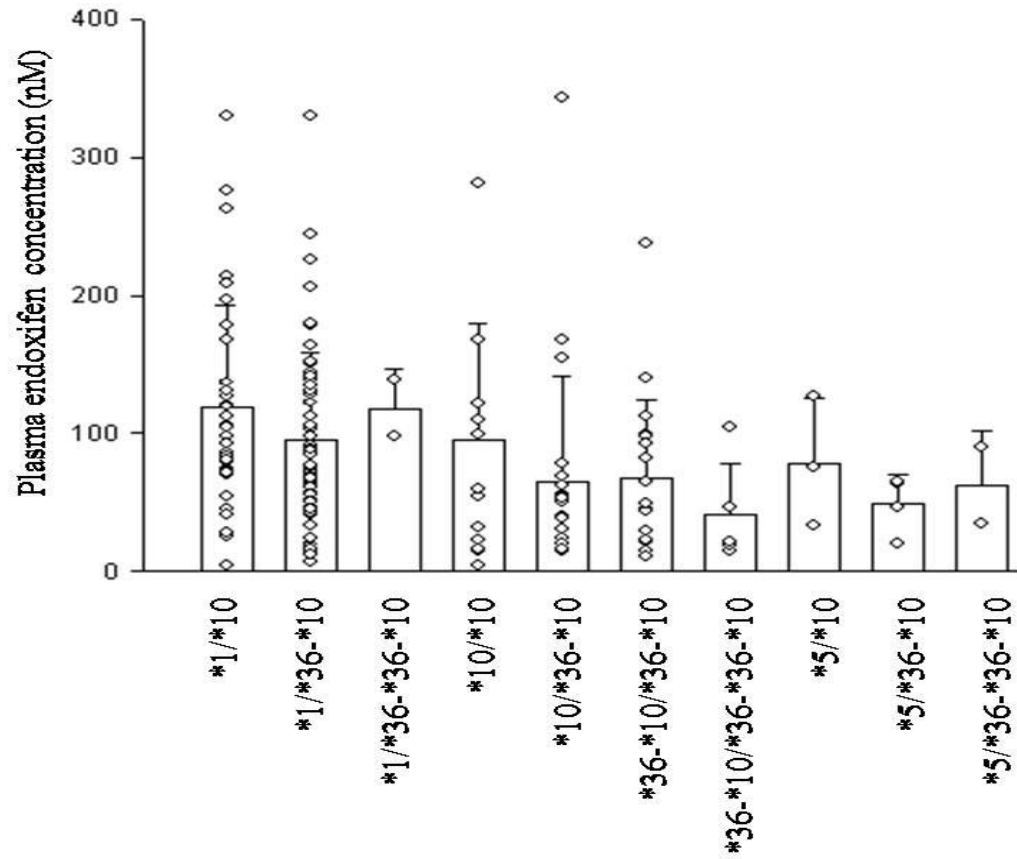
**(C) Filipino cohort**

CYP2D6 diplotype <sup>a</sup>	N	Mean concentration, nM ( $\pm$ SD)				
		Tamoxifen	N-desmethyl tamoxifen	4-hydroxy tamoxifen	Endoxifen	Endoxifen/ NDM ratio
*1/ *10	19	466 $\pm$ 493	753 $\pm$ 662	4.7 $\pm$ 3.0	109.1 $\pm$ 76.7	0.17 $\pm$ 0.09
*1/ *36-*10	38	476 $\pm$ 405	860 $\pm$ 717	4.8 $\pm$ 3.1	103.8 $\pm$ 66.9	0.14 $\pm$ 0.05
*1/ *36-*36-*10	2	667 $\pm$ 40	943 $\pm$ 9	5.0 $\pm$ 1.3	117.9 $\pm$ 28.7	0.12 $\pm$ 0.03
*10/ *10	9	1144 $\pm$ 1252	1907 $\pm$ 2035	7.0 $\pm$ 6.4	104.6 $\pm$ 83.1	0.07 $\pm$ 0.02
*10/ *36-*10	11	901 $\pm$ 1522	1617 $\pm$ 2676	5.4 $\pm$ 7.7	70.0 $\pm$ 91.8	0.06 $\pm$ 0.03
*36-*10/ *36-*10	12	817 $\pm$ 1056	1576 $\pm$ 1602	5.1 $\pm$ 3.7	75.8 $\pm$ 59.4	0.08 $\pm$ 0.05
*36-*10/ *36-*36-*10	0					
*5/ *10	2	910 $\pm$ 686	1799 $\pm$ 729	7.0 $\pm$ 3.5	101.2 $\pm$ 37.0	0.06
*5/ *36-*10	3	333 $\pm$ 229	824 $\pm$ 392	4.0 $\pm$ 2.2	43.2 $\pm$ 21.9	0.05 $\pm$ 0.01
*5/ *36-*36-*10	1	299	781	1.9	34.8	0.04

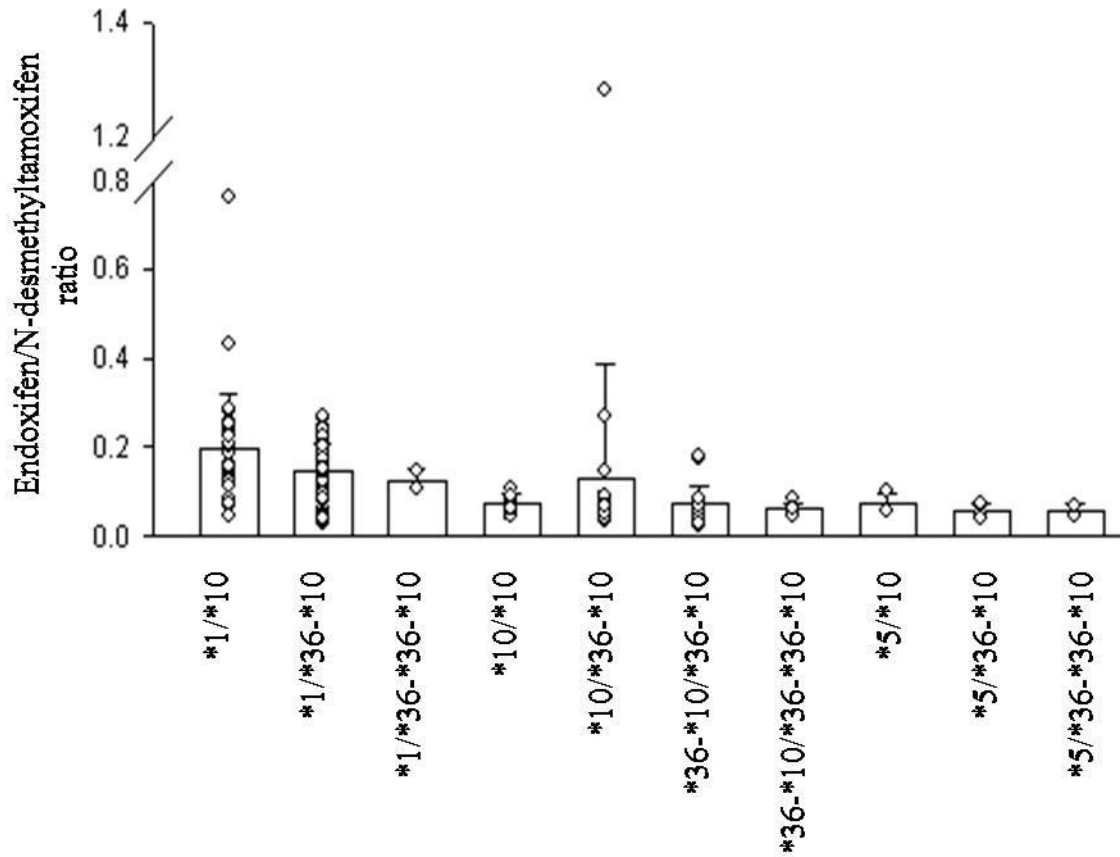
Note: (A) Combined cohort (N = 163); (B) Vietnamese cohort (N = 61); and (C) Filipino cohort (N = 102); <sup>a</sup> \*1 is inclusive of both \*1 and \*2 alleles.

Figure 3.2: Correlation between *CYP2D6* allele copy numbers with tamoxifen metabolism in the combined cohort.

(A) Endoxifen concentration



**(B) Endoxifen/NDM ratio**



Note: The effect of CYP2D6 score groups on endoxifen concentration in the Combined cohort (N = 157 individuals) after tamoxifen therapy (20 mg/day) for four months. The solid bars represent mean + SD of (A) endoxifen concentration and (B) endoxifen/N-desmethyltamoxifen ratio based on the CYP2D6 genotype and copy number variations.

**Discussion**

The drug metabolizing enzyme cytochrome P450 2D6 (*CYP2D6*) is involved in the metabolism of 20-25% of commonly prescribed drugs (Ingelman-Sundberg, 2005). Both *CYP2D6* gene polymorphisms and copy number variations (CNVs) are known to affect the enzyme activity and response to *CYP2D6* substrate drugs (Ingelman-Sundberg, 2005; Eichelbaum et al., 2006). The *CYP2D6* CNVs involve either deletion or multiplication of the entire gene. These *CYP2D6* CNVs are present in all ethnicities, but

the alleles involved in the CNV are different among the different ethnicities (Bradford, 2002; Sistonen et al., 2007). For example, CNVs involving *CYP2D6\*36-10* allele has been reported to be present in a high frequency in Asians, specifically Japanese and Chinese (Hosono et al., 2009). However, the prevalence of *CYP2D6* CNV in the Vietnamese and Filipino populations is not known.

The most widely used method of genotyping *CYP2D6\*10* is by the presence of 100T mutation (Fukuda et al., 1999; Xu et al., 2008). However, this genotyping method does not distinguish between *CYP2D6\*10* and *CYP2D6(\*36)<sub>n</sub>-10* as the 100T polymorphism is also present in *CYP2D6\*36*. This *CYP2D6\*36* allele is thought to have arisen from *CYP2D6\*10* that had undergone recombination with *CYP2D7* pseudogene, specifically in the exon 9 region. The *CYP2D6\*36-10* allele was first described in the Chinese population as a tandem 44kb fragment by *Xba*I digestion (Johansson et al., 1994). This *CYP2D6\*36-10* haplotype is different from the *CYP2D6\*10-10* duplication haplotype (Soyama et al., 2006). *In vitro* studies using recombinant *CYP2D6\*36* (expressed singly and not in tandem as *CYP2D6\*36-10*) showed that the protein content was decreased and the enzyme activity was negligible suggesting that the *CYP2D6\*36* allele belonged to the PM group (Hanioka et al., 2006; Sakuyama et al., 2008). These *in vitro* studies have suggested that because *CYP2D6\*36* had extremely low enzyme activity, the enzyme activity of *CYP2D6\*36-10* would be comparable to that of *CYP2D6\*10*. *In vivo* studies involving *CYP2D6\*36-36* duplication allele in one Japanese (Chida et al., 2002) and *CYP2D6\*36* allele in one Asian and four African American (Gaedigk et al., 2006) individuals concluded that carriers of the *CYP2D6\*36* allele were PMs. Thus, the carriers of *CYP2D6\*36* and *CYP2D6\*36-36* are considered

to be PMs, while the carriers of *CYP2D6*\*36-*\*10* are considered to be IMs. However, not much is known about the phenotypic consequence of other complex arrangements including the  $(*36)_n$ -*\*10* allele. Recent studies in Asian women from Korea (Lim et al., 2007), China (Xu et al., 2008) and Japan (Kiyotani et al., 2008) who were on adjuvant tamoxifen therapy have suggested that the *CYP2D6*\**10* allele is associated with poor clinical outcomes. However, these studies did not distinguish between *CYP2D6*\**10* and copy number variations involving *CYP2D6*( $*36$ )<sub>n</sub>-*\*10*, a CNV that is very common in Japanese and Chinese individuals (Hosono et al., 2009).

This study is the first comprehensive study to analyze: (1) the *CYP2D6* CNVs in Vietnamese and Filipino population and (2) the functional consequence of *CYP2D6*\*36 copy number variations. In a prospective clinical trial involving 93 Vietnamese and 144 Filipino women on adjuvant tamoxifen therapy, the presence of *CYP2D6* gene CNVs was investigated. The results indicated that in the combined cohort, only 39% of the individuals carried 2 copies of the *CYP2D6* gene. Approximately 61% of the individuals had a genomic imbalance for *CYP2D6* - 3.2% had a *CYP2D6* gene deletion (*CYP2D6*\*5) and 57.2% of the individuals had three or more copies of the *CYP2D6* gene (Figure 3.1). These results are comparable to the *CYP2D6* CNVs reported in Japanese and Chinese individuals (Hosono et al., 2009).

The results of the CNV phaser program (Kato et al., 2008) that was used to estimate the *CYP2D6* diplotypes indicated that 34.2% of the individuals had allelic CNVs involving the *CYP2D6*( $*36$ )<sub>n</sub>-*\*10* duplication alleles. In most of these individuals (99.8% of the cases), the *CYP2D6*( $*36$ )<sub>n</sub> allele was almost always present along with a

copy of a *CYP2D6\*10* allele (Table 3.1). Diplotypes involving these CNV alleles were also fairly common with a frequency of about 50% (Table 3.2).

In order to determine the phenotypic consequence of these CNVs involving the *CYP2D6*(\*36)<sub>n</sub> alleles, tamoxifen and its metabolite concentration between different alleles: \*10 vs. \*36-\*10 vs. \*36-\*36-\*10 and also \*10 vs. (\*36)<sub>n</sub> -\*10 were compared. There was no significant difference in plasma endoxifen concentration or endoxifen/NDM ratio between the different groups indicating that *CYP2D6*\*36 copy number variations were not significantly associated with increased or decreased *CYP2D6* enzyme activity (Table 3.3 A-C; Figure 3.2 A and B). The enzyme activity of *CYP2D6*\*36-\*10 and *CYP2D6*\*36-\*36-\*10 is comparable to that of *CYP2D6*\*10. That is, based on tamoxifen metabolism, these *CYP2D6*(\*36)<sub>n</sub> CNV alleles are still null alleles and consequently, can be considered as poor metabolizers. However, there appears to be a trend of decreased enzyme activity with an increase in the number of *CYP2D6*\*36 copies. That is, in some cases, plasma endoxifen concentration and endoxifen/NDM ratio of \*10 allele is greater than \*36-\*10 allele, which in turn is greater than that of \*36-\*36-\*10 allele (Table 3.3 A-C; Figure 3.2 A and B). Further confirmation of these results in other Asian populations will add to the literature. However, even after correcting for these copy number variations, there is considerable unexplained interindividual variability (Figure 3.2 A and B).

A number of genes involved in the metabolism of drugs and xenobiotics are also known to exhibit gene copy number variations - *UGT2B17* and *UGT2B28* (McCarroll et al., 2006), *SULT1A1* (Hebbring et al., 2007), *GSTT1* and *GSTM1* (Bolt and Thier, 2006). These CNVs can have functional consequence; however, CNVs involving the *SULT1A1*



gene were not found to be associated with the concentration of tamoxifen or its metabolites (Gjerde et al., 2008).

In the human genome, copy number variations are fairly common (Iafrate et al., 2004; Sebat et al., 2004). Thousands of copy number variable regions (CNVR) that ranges from 100 bp to several Mb have been reported (Redon et al., 2006). Recent studies have suggested an increasing role for CNVs in many diseases including cancer, developmental diseases, mental illness autoimmune diseases and infectious diseases (Shrestha et al., 2009; Wain et al., 2009). However, our knowledge of the functional consequences of these CNVs is still very poor. The phenotypic consequence of the individual copies of a gene within the CNVRs is not clearly elucidated. The results of this study suggest that *all copies of a gene are not created equal*. Multiple copies of a gene do not always translate to additive effects.

Overall, this study indicated that *CYP2D6* gene copy number variations exist in a high frequency in the Vietnamese and Filipino populations. These copy number variations primarily involve the *CYP2D6*(\*36)<sub>n</sub>-\*10 alleles. This study is the first large scale study to report that based on the plasma endoxifen concentration, the *CYP2D6*(\*36)<sub>n</sub>-\*10 enzyme activity is not significantly different from that of *CYP2D6*\*10. Hence, in order to determine the phenotypic consequence, in addition to determining the total *CYP2D6* copy numbers, it is important to determine which alleles have altered copy numbers. The high frequency of CNVs involving the *CYP2D6*\*36 null allele in these Asian populations raises some interesting questions about the evolution of such complex *CYP2D6* gene rearrangements.

## Materials and Methods

### Estimation of *CYP2D6* gene copy numbers and the presence of *CYP2D6\*36*

**allele:** The *CYP2D6* gene copy numbers were estimated using a quantitative RT-PCR CNV assay - 5' flank assay (Hosono et al., 2009) along with an internal control RNase P TaqMan Copy Number Reference Assay (ABI, Forest City, CA, USA). For copy number estimation, each sample was analyzed in triplicate using the TaqMan Genotyping PCR Master Mix (ABI, Forest City, CA, USA), in a 15  $\mu$ L reaction including 20 ng of genomic DNA (5 ng/ $\mu$ l) in a Step-One Plus PCR instrument (Applied Biosystems, Forest City, CA, USA). The final concentrations of 5' flank assay primers and hydrolysis probes were 900 nmol/L and 200 nmol/L, respectively. The cycling conditions used were: 95°C for 10 min for initial denaturation and enzyme activation, followed by 40 cycles each of 95°C for 15 s and 60°C for 1 min. The relative quantitation (RQ) was performed using the CopyCaller Software (Applied Biosystems, Forest City, CA, USA), following the comparative  $\Delta\Delta$ CT method. Two Coriell panel samples - NA18529 (2 copies) and NA18968 (3 copies) were used as internal controls for *CYP2D6* gene CNVs (Hosono et al., 2009). The assay was repeated in 10% of randomly selected samples to check for concordance and reproducibility. Additionally, another qRT-PCR assay, Int 2 assay (Hosono et al., 2009) was performed in 20-30% of the samples to check for concordance with the 5' flank assay.

Another copy number assay, Hs00010001\_cn (Ex9; ABI, Forest City, CA, USA) was used to estimate the number *CYP2D6\*36* alleles. This assay specifically targets exon 9 region of *CYP2D6* gene. Since the *CYP2D6\*36* allele undergoes gene conversion with *CYP2D7* in exon 9, the Ex9 assay does not amplify *CYP2D6\*36* allele. The

difference in copy numbers between the 5' flank assay and this Ex9 assay was used to indirectly estimate the copy number of *CYP2D6*\*36 allele. This approach has been verified using 32 DNA samples from the Japanese and Chinese Coriell diversity panel (Ramamoorthy et al., 2010). The assay was repeated using 10% of randomly selected samples to check for concordance and reproducibility.

**Statistical analysis:** In order to identify the *CYP2D6* diplotypes, the CNV phaser software was used (Kato et al., 2008). For this analysis, all the *CYP2D6* alleles (*CYP2D6*\*1, \*2, \*3, \*4, \*5, \*6, \*10, \*36 and \*41) were included. As described previously (Hosono et al., 2009), the copy number data was rounded to an integer and the diplotypes with the highest diplotype probability score are reported.

The comparison between the different *CYP2D6* diplotype groups and the plasma concentrations of tamoxifen and its metabolites, N-desmethyltamoxifen, 4-hydroxytamoxifen, endoxifen and also endoxifen/NDM ratio were analyzed by ANOVA and Wilcoxon sum-rank test using the SPSS software (version 13.0; SPSS Inc, Chicago, IL, USA). A p-value of < .05 was considered to be statistically significant.

## CHAPTER FOUR

### ***In silico* and *In vitro* Identification of microRNAs that Post-transcriptionally Regulate HNF4A Expression.**

The work presented in this chapter has been prepared in part for submission as '*In silico* and *in vitro* identification of microRNAs regulating hepatic drug metabolism: post-transcriptional regulation of HNF4A by microRNAs' by Anuradha Ramamoorthy, Lang Li, David Flockhart and Todd Skaar.

#### **Abstract**

Cytochrome P450 2D6 (CYP2D6) enzyme activity is highly variable among individuals; this variability is not completely explained by known environmental and genetic factors. To further understand this variability, the role of microRNAs (miRNAs) in the regulation of CYP2D6 expression was investigated. MiRNA identification algorithms were used to predict the miRNAs that regulate CYP2D6 and its transcriptional factor, hepatic nuclear factor 4 $\alpha$  (HNF4A). The results suggested that both genes are targets of miRNAs. Using these bioinformatic predictions, *in vitro* functional validation studies were performed. For CYP2D6, none of the five miRNAs that were tested regulated its expression. For HNF4A, five out of six miRNAs that were tested by luciferase assays, down-regulated its activity by 20-40%. In HepG2 cells, two of the miRNAs (hsa-miR-34a and hsa-miR-449a) down-regulated HNF4A protein expression by 30-60%. These two miRNAs did not reduce the levels of HNF4A mRNA, but did down-regulate the level a HNF4A down-stream target, pregnane X receptor (PXR) mRNA by 30-40%, suggesting that the miRNAs are blocking HNF4A protein synthesis. Of these miRNAs, the highest expressed miRNA in hepatocytes was hsa-miR-34a. Additionally, bioinformatic algorithms were used to identify polymorphisms in the

HNF4A mRNA 3'-untranslated region (3'-UTR) that might affect normal mRNA-miRNA interaction. *In vitro* luciferase assays validated this prediction for a SNP (rs11574744) in the HNF4A 3'-UTR; the luciferase activity of the variant was two-fold higher than the wild-type. These findings demonstrate a potential role for miRNAs in the regulation of genes involved in drug metabolism. Identification of endogenous hepatic miRNAs that regulate CYP2D6 directly or indirectly should be useful for understanding the clinical interindividual variability in drug metabolism and efficacy.

## **Introduction**

The cytochrome P450 2D6 (CYP2D6) enzyme metabolizes 20-25% of commonly prescribed drugs including many antidepressants, beta-blockers and anticancer agents (Ingelman-Sundberg, 2005). The *CYP2D6* gene is highly polymorphic and these polymorphisms affect enzyme activity (Ingelman-Sundberg, 2005; Eichelbaum et al., 2006). However, even within the same *CYP2D6* genotype group [e.g., extensive metabolizers (EM) or intermediate metabolizers (IM)], there is still unexplained interindividual variability (Borges et al., 2006). A part of this variability may be due to co-medication with drugs that act as CYP2D6 inhibitors (Stearns et al., 2003; Jin et al., 2005). While genetic and environmental factors explain only a part of the interindividual variability in CYP2D6 enzyme activity, substantial unexplained variability still remains. Epigenetic mechanisms may also contribute to the variability (Ingelman-Sundberg et al., 2007); however, little is known about which factors are important.

MicroRNAs are small (18-25 nucleotides), noncoding RNAs that regulate gene expression post-transcriptionally. In animals, miRNAs typically bind to the 3'-

untranslated region (3'-UTR) of the messenger RNAs (mRNAs) and negatively regulate gene expression either by blocking protein translation or by degrading the mRNA (Olsen and Ambros, 1999; Ambros et al., 2003). As more miRNAs are identified and studied, newer target sites and functions are being recognized. For example, it has now been shown that miRNAs can also bind to coding regions and repress gene expression (Duursma et al., 2008); this mechanism may explain some of the differential expression seen in mRNA splice variants. MiRNAs also appear to be involved in the induction of gene expression; this induction occurs through binding to complementary regions in the promoter (Place et al., 2008) and the 5'-UTR (Orom et al., 2008).

In the human genome, 940 mature miRNAs have been reported so far (miRBase Registry; version 15.0; Griffiths-Jones et al., 2006). Bioinformatic programs predict that miRNAs can regulate 20-90% of the mRNA transcripts (Lewis et al., 2005; Xie et al., 2005; Miranda et al., 2006). Each miRNA can regulate multiple genes and each gene can be regulated by multiple miRNAs; therefore, these miRNAs form a broad and complex regulatory network.

MicroRNAs are involved in a wide range of biological activities including cell differentiation, cell death, cancer and noncancerous human diseases (John et al., 2004). Emerging evidence indicates that these miRNAs also regulate genes involved in drug disposition. Those studies have focused on genes such the ATP-binding cassette transporter *ABCG2* (To et al., 2008), pregnane X receptor (*PXR*; Takagi et al., 2008), cytochrome P450 1B1 (*CYP1B1*; Tsuchiya et al., 2006) and cytochrome P450 3A4 (*CYP3A4*; Pan et al., 2009). Others have also speculated that interindividual variability in

CYP450 expression and drug response may be due to the action of miRNAs (Ingelman-Sundberg et al., 2007).

The activity of miRNAs can be affected by single nucleotide polymorphisms (SNPs) that occur either in the miRNA or in the miRNA target site on the mRNA. Such SNPs are called miRSNPs (Mishra et al., 2007). These miRSNPs can alter miRNA gene processing and/or the normal mRNA-miRNA interactions. Thus, these SNPs can create new miRNA target sites or destroy old target sites. Hence, these miRSNPs may also contribute to the interindividual variability in the CYP450 enzyme expression and activity.

The overall goal of the present study was to investigate the role of miRNAs in the regulation of CYP2D6 enzyme expression. The *HNF4A* gene was also included in this analysis as it regulates the expression of *CYP2D6* (Cairns et al., 1996; Corchero et al., 2001). *HNF4A* is an important transcriptional ‘master regulator’ that regulates the expression of a number of drug metabolizing genes including phase I enzymes, phase II enzymes, transporters and transcriptional factors that regulate the CYP450 genes (Kamiyama et al., 2007). Thus, regulation of *HNF4A* by miRNAs would likely affect many genes that are involved in drug metabolism. Hence, the substantial unexplained variability in CYP2D6 enzyme activity may be due to direct regulation of CYP2D6 by miRNAs, as well as indirect regulation of it through *HNF4A*.

In this study, the central hypothesis is that endogenous miRNAs regulate the expression of CYP2D6 and its transcription factor *HNF4A*. MiRNA regulation of *HNF4A* may explain some of the as yet unexplained variability in CYP2D6 expression. To test this hypothesis, bioinformatic analyses were first performed to predict miRNAs

that target CYP2D6 and HNF4A, followed by *in vitro* functional validation studies.

Lastly, polymorphisms in the miRNA target sites on HNF4A 3'-UTR that are predicted to alter the mRNA-miRNA interactions were identified. Collectively, these results suggest that miRNAs are likely to play an important role in the regulation of drug metabolism.

## **Results**

### **Bioinformatic predictions to identify miRNAs**

Six bioinformatic algorithms were used to identify miRNAs that are predicted to target CYP2D6 and HNF4A: the algorithms were miRanda (John et al., 2004), miRBase Targets (Griffiths-Jones et al., 2006), TargetScan (Lewis et al., 2003), PicTar (Krek et al., 2005), RNA22 (Miranda et al., 2006) and PITA (Kertesz et al., 2007). These algorithms predicted that both the genes were miRNA targets (Table 4.1). There was considerable variability in the number of miRNAs that were predicted by the different programs to target CYP2D6 and HNF4A. The *CYP2D6* gene has a 75 bp long 3'-UTR and was predicted to be targeted by 7 different miRNAs. The *HNF4A* gene has a 1724 bp long 3'-UTR and was predicted to be targeted by 350 different miRNAs.



**Table 4.1: MicroRNAs that are predicted to target the 3'-UTR of CYP2D6 and HNF4A.**

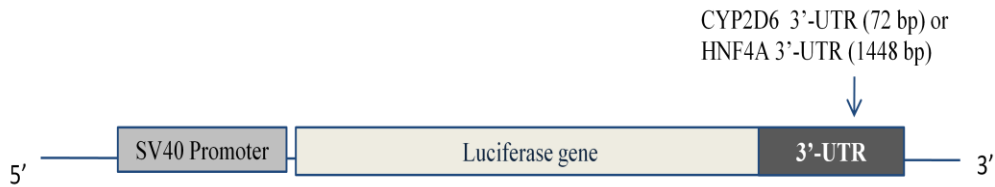
Genes and Reference id. <sup>a</sup>	3'-UTR length (bp)	Bioinformatic Programs <sup>b,c</sup>					Total no. of unique miRNAs <sup>d</sup>	Overlap <sup>e</sup>
		miRBase Targets	miRNA -da	Target Scan	RNA22	PITA		
CYP2D6 (NM_000106)	75	5	0	3	1	2	7	4
HNF4A (NM_000457)	1724	10	13	112	251	63	350	99

Note: <sup>a</sup> RefSeq Gene id from UCSC Genome browser in parenthesis; <sup>b</sup> PicTar predictions are not included in the table as CYP2D6 and HNF4A do not appear to be a part of the program's database; <sup>c</sup> Versions of the bioinformatic programs (retrieved on 07/27/2008): (1) miRBase Targets using miRBase release 11.0, (2) miRanda and TargetScan version 4.2 uses miRBase release 10.0, (3) miRBase Registry version 10.0 was used in RNA22 algorithm, and (4) PITA uses miRBase release 9.0; <sup>d</sup> Total number of unique miRNAs predicted by all the programs; <sup>e</sup> Total number of miRNAs predicted to target the genes by at least 2 programs.

### ***In vitro* functional studies to test the bioinformatic predictions**

In the first functional study, 3'-UTRs of *CYP2D6* and *HNF4A* genes were cloned into the 3'-UTR of the luciferase gene of the pIS-0 luciferase reporter construct (Figure 4.1). These plasmids with CYP2D6 and HNF4A 3'-UTR are referred to as pIS-2D6 and pIS-HNF4A, respectively.

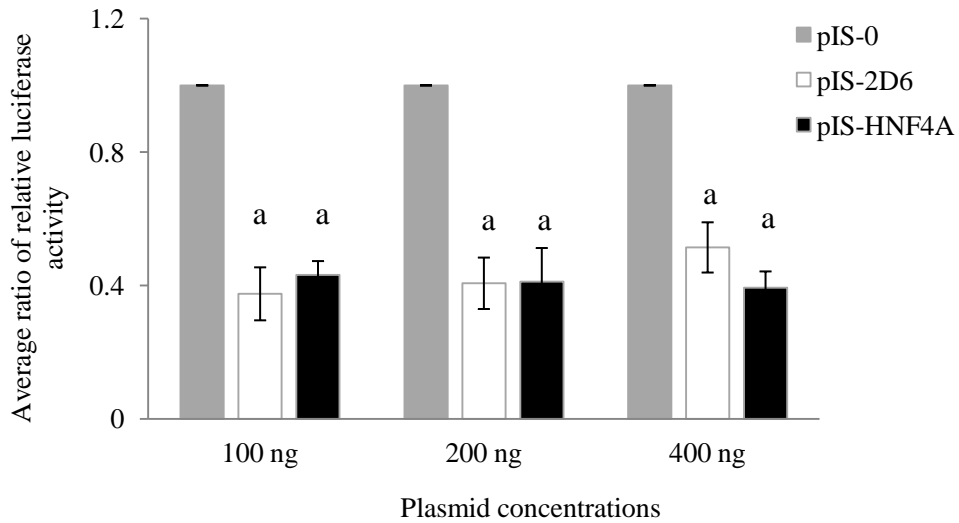
**Figure 4.1: Schematic representation of the pIS-2D6 and pIS-HNF4A vectors.**



Note: In pIS-2D6 vector, 72 bp of the full length 75 bp CYP2D6 3'-UTR was cloned and in pIS-HNF4A vector, 1448 bp of the full length 1724 bp HNF4A 3'-UTR was cloned.

In the *in vitro* studies, both pIS-2D6 and pIS-HNF4A plasmids significantly reduced ( $p < .01$ ) luciferase activity, relative to the pIS-0 (control) plasmid (Figure 4.2). The luciferase activities of pIS-2D6 and pIS-HNF4A plasmids were approximately 55% and 60 % lower than the pIS-0 control plasmid. The effect of both the CYP2D6 and HNF4A 3'-UTRs were consistent across three concentrations of transfected plasmids, but was not significantly different (Figure 4.2). These results indicated that the 3'-UTRs of both the *CYP2D6* and *HNF4A* genes have repressive activities, possibly through miRNA targeting.

**Figure 4.2: Regulatory elements in the 3'-UTR of CYP2D6 and HNF4A.**

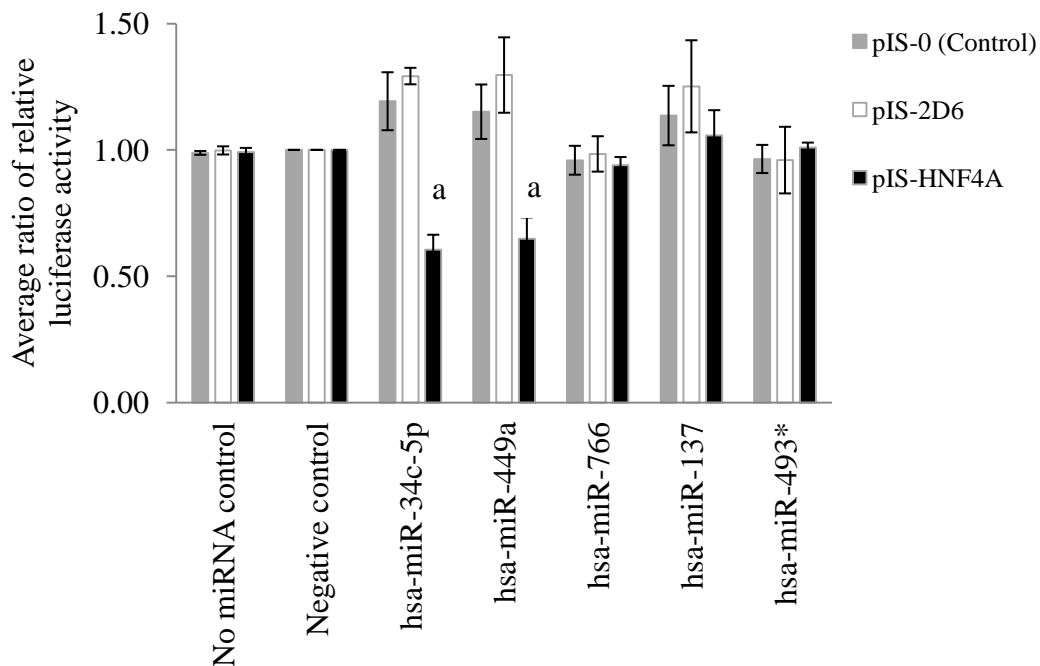


Note: HeLa, a cervical cancer cell line was transfected with pIS-0 (control), or pIS-2D6, or pIS-HNF4A plasmids at three different concentrations (100, 200 and 400 ng per well). *Renilla* luciferase plasmid was used as an internal control. Dual luciferase assays were performed 24 hours after transfection. Data are expressed as the pIS-2D6 and pIS-HNF4A luciferase activities corrected for *Renilla* luciferase and normalized to the pIS-0 within each experiment (mean  $\pm$  SEM; n = 3 independent experiments performed in triplicates). <sup>a</sup> indicates  $p < .01$  in pIS-2D6 and pIS-HNF4A when compared to the control pIS-0.

The next set of studies focused on testing the effect of individual miRNAs on the pIS-2D6, pIS-HNF4A and the control pIS-0 plasmids. Candidate miRNAs were selected based on two criteria: (1) the miRNA predicted by two or more of the bioinformatic algorithms and (2) favorable energy of binding ( $\Delta G$ ) between the miRNA and the target sequence on the mRNA. Among those that fit these criteria were hsa-miR-493\* (formerly referred to as hsa-miR-493-5p) and hsa-miR-137 which were predicted to target CYP2D6, and hsa-miR-34c-5p, hsa-miR-449a and hsa-miR-766 which were predicted to target HNF4A. These miRNAs were selected for the initial *in vitro* analysis. These miRNAs were cotransfected with the pIS plasmids (pIS-0, or pIS-2D6, or pIS-HNF4A). None of these miRNAs reduced the pIS-2D6 luciferase activity (Figure 4.3).

However, relative to the negative control miRNA (cel-miR-67), hsa-miR-34c-5p and hsa-miR-449a significantly reduced ( $p < .01$ ) the luciferase activity of the pIS-HNF4A plasmid by 40% and 35%, respectively. In contrast, hsa-miR-766 did not regulate the expression of pIS-HNF4A. Hsa-miR-34c-5p and hsa-miR-449a caused a slight unexplained increase in the pIS-0 and pIS-2D6 plasmids (Figure 4.3).

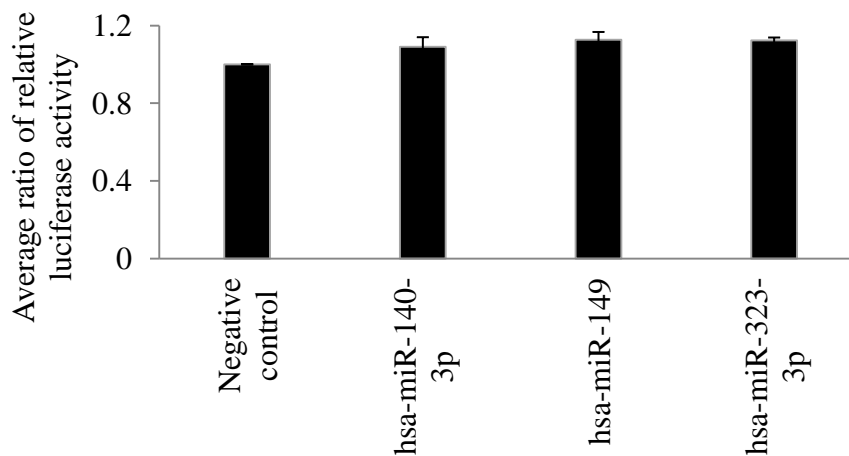
**Figure 4.3: Regulation of pIS-2D6 and pIS-HNF4A by miRNAs.**



Note: HeLa cells were cotransfected with 4  $\mu$ g of pIS-0, or pIS-2D6, or pIS-HNF4A luciferase plasmids, along with *Renilla* reporter plasmid for normalization. These cells were also transfected with 30 nM of miRNA Mimics. Data are expressed as the pIS-HNF4A or pIS-2D6 luciferase activity corrected for *Renilla* luciferase and normalized to pIS-0 within each experiment (mean  $\pm$  SEM;  $n = 3$  independent experiments performed in triplicates). <sup>a</sup> indicates  $p < .05$  when compared to the control the negative control (cel-miR-67) transfection.

In another set of transfections, three additional miRNAs (hsa-miR-140-3p, hsa-miR-149 and hsa-miR-323-3p) that were predicted by TargetScan version 5.1 to regulate the expression of CYP2D6 were tested. None of these miRNAs affected the pIS-2D6 luciferase activity (Figure 4.4).

**Figure 4.4: Lack of regulation of pIS-2D6 by miRNAs.**

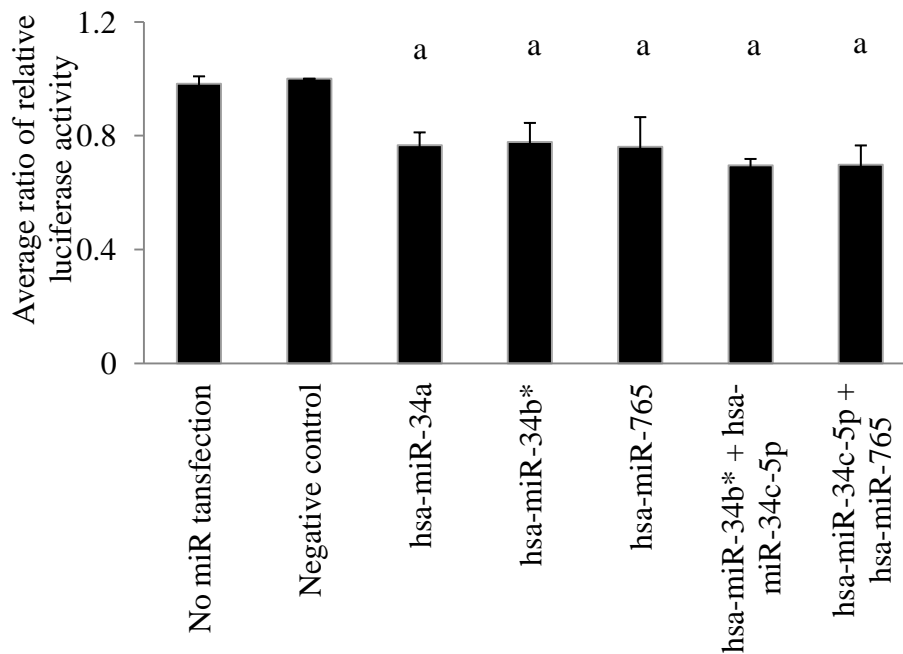


Note: HeLa cells were cotransfected with 4  $\mu$ g of pIS-2D6 luciferase plasmids, along with *Renilla* reporter plasmid for normalization. These cells were also transfected with 30 nM of miRNA Mimics. Data are expressed as the pIS-2D6 luciferase activity corrected for *Renilla* luciferase and normalized to negative control (cel-miR-67) within each experiment (mean  $\pm$  SEM; n = 3 independent experiments performed in triplicates).

Similarly, three additional miRNAs were tested for interaction with pIS-HNF4A. These miRNAs were: hsa-miR-34a (same family as hsa-miR-34c-5p), hsa-miR-34b\* (which is a part of a single large precursor miRNA transcript that makes hsa-miR-34c-5p and hsa-miR-34a) and hsa-miR-765. When compared to the negative control miRNA (cel-miR-67), hsa-miR-34a, hsa-miR-34b\* and hsa-miR-765 significantly reduced ( $p \leq .05$ ) the pIS-HNF4A luciferase activity by 23%, 22%, and 24%, respectively (Figure 4.5). Since multiple miRNAs can simultaneously interact with a target mRNA (Krek et al.,

2005), hsa-miR-34c-5p was cotransfected with hsa-miR-34b\* and hsa-miR-34c-5p was cotransfected with hsa-miR-765. Even though these cotransfections significantly reduced the luciferase activity ( $p < .05$ ), no additive or synergistic effects were observed by cotransfecting these miRNAs. In the case of the cotransfection of hsa-miR-34c-5p with the hsa-miR-34b\*, the lack of additional or synergistic effect may be because of competition, as they are predicted to share one target site in common.

**Figure 4.5: Regulation of pIS-HNF4A by miRNAs.**

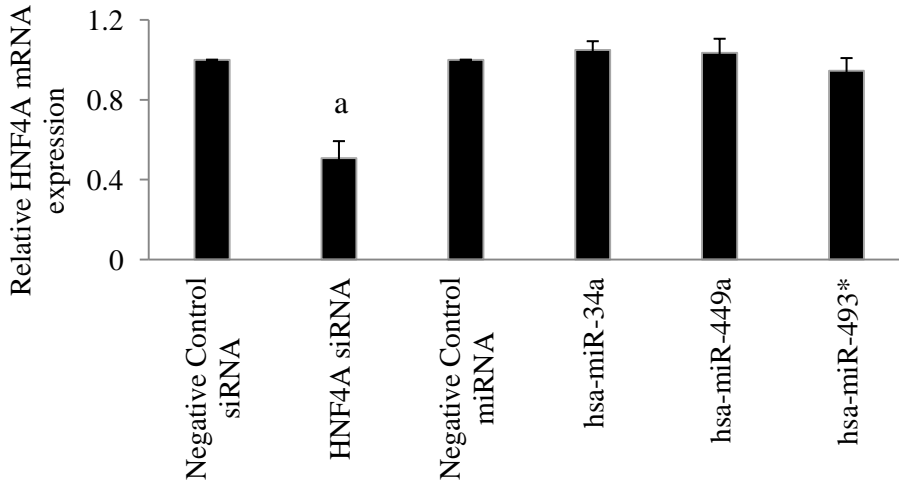


Note: HeLa cells were cotransfected with 4  $\mu$ g of pIS-HNF4A luciferase plasmid, along with *Renilla* reporter plasmid for normalization. These cells were also transfected with miRNA Mimics individually (cel-miR-67, hsa-miR-34a, hsa-miR-34b\*, or hsa-miR-765; 30 nM) or together (hsa-miR-34b\* + hsa-miR-34c-5p; or hsa-miR-34c-5p + hsa-miR-765; 15 nM each). All data are expressed as the pIS-HNF4A luciferase activity corrected for *Renilla* luciferase and normalized to negative control (cel-miR-67) within each experiment (mean  $\pm$  SEM;  $n = 3$  independent experiments performed in triplicates). <sup>a</sup> indicates  $p \leq .05$  when compared to the control the negative control (cel-miR-67) transfection.

### **MicroRNAs regulate HNF4A mRNA and protein expression *in vitro***

In order to investigate whether these miRNAs regulate HNF4A mRNA and protein expression, HepG2 (a liver cancer cell line) was transfected with hsa-miR-34a, hsa-miR-449a, hsa-miR-493\* and the negative control, cel-miR-67. Both hsa-miR-34a and hsa-miR-449a were predicted to target HNF4A at two positions and they both target the same locations: (1) positions 164-171, and (2) positions 254-260 of HNF4A 3'-UTR corresponding to the miRNA seed sequence. Hsa-miR-493\* is not predicted to target HNF4A and hence was used as a control. Small interfering RNAs (siRNA) for HNF4A and a negative control siRNA described previously (Iwazaki et al., 2008) were also used as process controls. The results indicated that only HNF4A siRNA significantly reduced HNF4A mRNA expression by about 50% ( $p = .05$ ), while none of the miRNAs down-regulated HNF4A mRNA expression (Figure 4.6).

**Figure 4.6: MicroRNAs do not regulate the expression of HNF4A mRNA expression.**

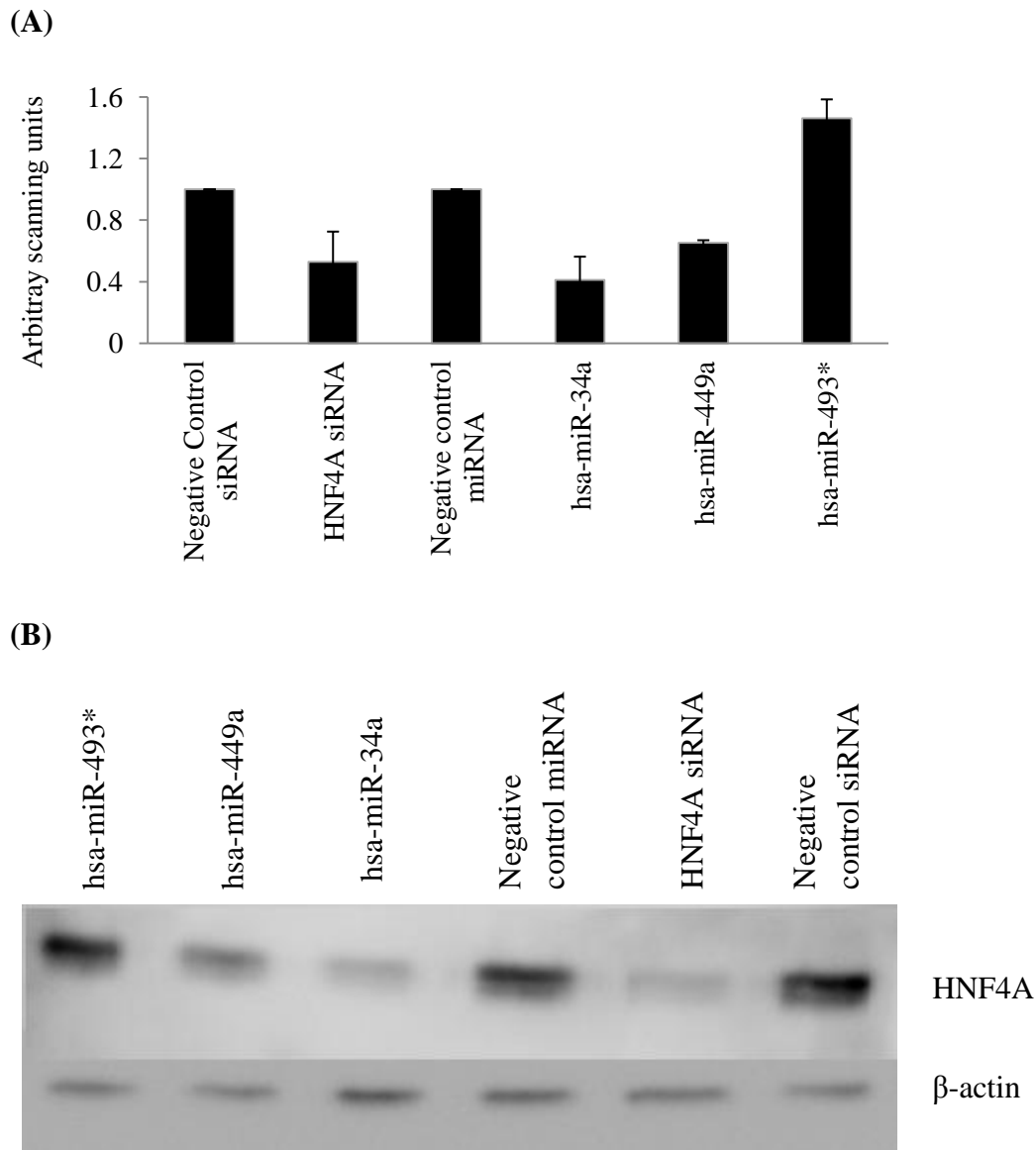


Note: HepG2 cells were transfected with miRNAs (hsa-miR-34a, hsa-miR-449a, hsa-miR-493\*, or the negative control cel-miR-67; 100 nM) or siRNAs (HNF4A siRNA, or negative control siRNA; 100 nM). At 72 hours after transfection, RNA was isolated. RT-PCR assays were performed for HNF4A mRNA using glyceraldehyde 3-phosphate dehydrogenase (GAPDH) as the internal control. MiRNA transfections were normalized to miRNA control cel-miR-67 and siRNA transfections were normalized to negative control siRNA. The relative mRNA expression was calculated as  $2^{-(\Delta\Delta CT)}$  (mean  $\pm$  SEM; n = 3 independent experiments performed in triplicates). <sup>a</sup> indicates p = .05 when compared to the control the negative control siRNA transfection.

In contrast to the HNF4A mRNA expression, HNF4A protein expression was down-regulated by both the siRNA and miRNAs; HNF4A siRNA reduced the HNF4A protein expression by 40%, while hsa-miR-449a down-regulated the HNF4A protein by about 30%, and hsa-miR-34a by about 60% (Figure 4.7 A and 4.7 B).



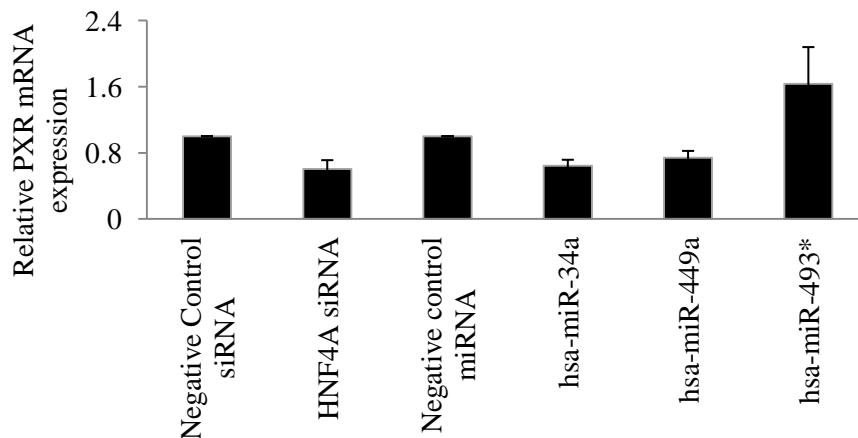
**Figure 4.7: Regulation of HNF4A protein by miRNAs.**



Note: HepG2 cells were transfected with miRNAs (hsa-miR-34a, hsa-miR-449a, hsa-miR-493\*, or the negative control cel-miR-67; 100nM) or siRNAs (HNF4A siRNA, or negative control siRNA; 100 nM). At 72 hours after transfection, nuclear protein was isolated and western blot assays were performed. Beta-actin ( $\beta$ -actin) was used as the internal control for normalization. MiRNA transfections were normalized to miRNA negative control cel-miR-67, and siRNA transfections were normalized to negative control siRNA. (A) The blots were scanned and the band density was measured (mean  $\pm$  SEM; n = 3 independent experiments); (B) Representative western blot performed using anti-HNF4A antibody and anti- $\beta$ -actin antibody.

To determine whether this down-regulation of HNF4A resulted in altered expression of downstream HNF4A target genes, PXR mRNA expression was measured. PXR is a target of HNF4A transcription factor that has been used as a marker of HNF4A activity (Iwazaki et al., 2008). The expression of PXR mRNA was reproducibly down-regulated by approximately 30% by hsa-miR-449a, by approximately 40% by both hsa-miR-34a and HNF4A siRNA (Figure 4.8). In contrast, hsa-miR-493\* did not down-regulate HNF4A mRNA, PXR mRNA, or HNF4A protein expression. In HepG2 cells, CYP2D6 mRNA expression was barely detectable by RT-PCR and did not appear to be influenced by HNF4A expression, probably due to its low basal expression levels (data not shown).

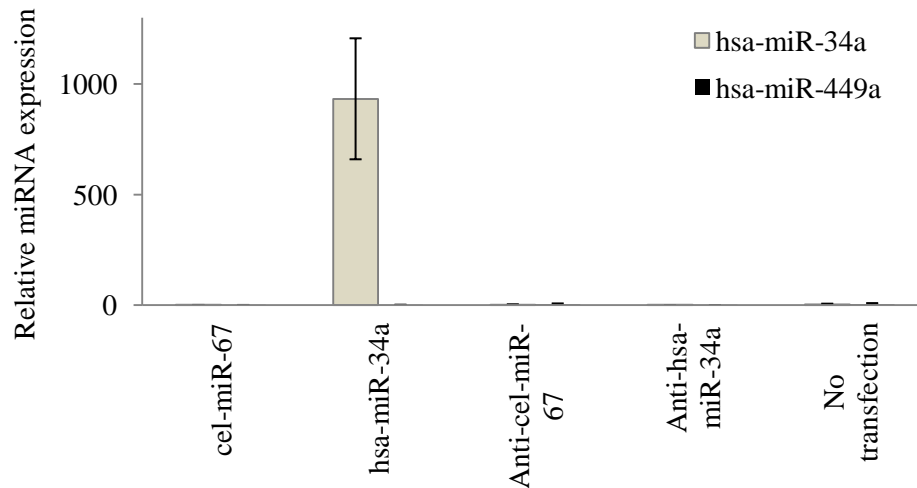
**Figure 4.8: Regulation of PXR mRNA by miRNAs.**



Note: HepG2 cells were transfected with miRNAs (hsa-miR-34a, hsa-miR-449a, hsa-miR-493\*, or cel-miR-67; 100 nM) or siRNAs (HNF4A siRNA, or negative control siRNA; 100 nM). At 72 hours after transfection, RNA was isolated. RT-PCR assays were performed with GAPDH as the internal control. MiRNA transfections were normalized to miRNA control cel-miR-67 and siRNA transfections were normalized to negative control siRNA. The relative mRNA expression was calculated as  $2^{-(\Delta\Delta CT)}$  (mean  $\pm$  SEM; n = 3 independent experiments performed in triplicates).

In order to assess the transfection efficiency of the miRNA Mimics, quantitative RealTime-PCR were performed in HeLa cells transfected with the synthetic miRNAs for hsa-miR-34a and cel-miR-67. The results indicated that transfection of hsa-miR-34a mimic resulted in approximately 900-fold increase in the concentration of mature miRNA when compared to the control (Figure 4.9).

**Figure 4.9: Transfection efficiency of synthetic miRNAs.**



Note: HeLa cells were transfected with synthetic Mimics and Inhibitors for hsa-miR-34a and negative control cel-miR-67. Total RNA including the small RNAs was isolated and reverse transcribed using hsa-miR-34a, or hsa-miR-449a, or U6 small nuclear RNA (snRNA) specific primers. RT-PCR experiments were performed as described under Methods section. The relative expression of the miRNAs was normalized to U6 snRNA. The relative mRNA expression was calculated as  $2^{-(\Delta\Delta CT)}$  (mean  $\pm$  SEM; n = 3 independent experiments performed in triplicates).

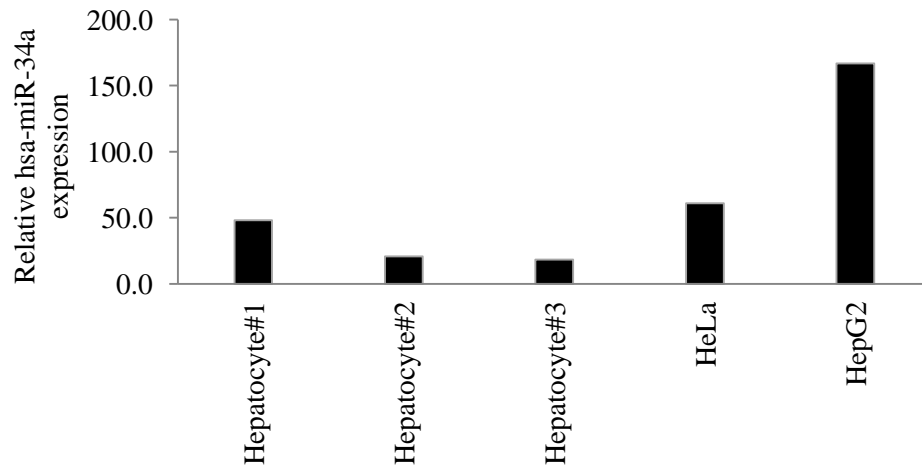
### Expression of miRNAs in human hepatocytes

To determine if the miRNAs that appear to target HNF4A are expressed in hepatocytes, their expression in primary hepatocytes isolated from three individual subjects was measured. Additionally, their expression in HeLa and HepG2 cells was also

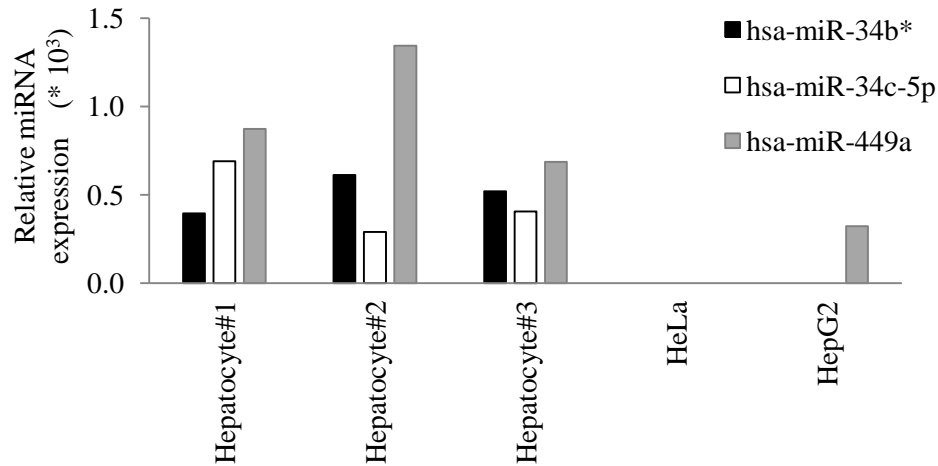
measured. Quantitative Real-Time PCR using specific TaqMan miRNA assays were performed using U6 small nuclear RNA (snRNA) as endogenous control. Hsa-miR-34a was easily detectable in all 3 hepatocyte preparations, as well as in HeLa and HepG2 cells (Figure 4.10 A). The other three hsa-miRNAs (hsa-miR-34b\*, hsa-34c-5p and hsa-miR-449a) were also expressed in hepatocytes, although at an apparent lower level (Figure 4.10 B). Neither hsa-miR-34b\*, nor hsa-34c-5p were detectable in HeLa or HepG2 cells. Hsa-miR-449a was expressed in HepG2, but not detectable in HeLa cells.

**Figure 4.10: Expression of mature miRNAs in human cell lines and hepatocytes.**

(A)



(B)



Note: Total RNA, including the small RNAs, was isolated from three different human hepatocytes preparations. RT-PCR assays were performed with U6 snRNA as the internal control. The relative miRNA expression was calculated as  $2^{-(\Delta CT)}$ . Values were multiplied by  $10^3$  to simplify data presentation. The values shown in the graph were obtained from triplicate assays in a single PCR experiment. (A) Expression of hsa-miR-34a; (B) Expression of hsa-miR-34b\*, hsa-miR-34c-5p and hsa-miR-449a.

### SNPs in the miRNA target sites

SNPs in the miRNA target sites on the mRNA 3'-UTR can alter the normal miRNA-mRNA interactions. The dbSNP database ([www.ncbi.nlm.nih.gov/projects/SNP](http://www.ncbi.nlm.nih.gov/projects/SNP)) was used to identify SNPs in the 3'-UTRs of *CYP2D6* and *HNF4A* genes. The dbSNP database indicated that SNPs are present in the 3'-UTR of *HNF4A* gene. Two programs, PolymiRTS (Bao et al., 2007) and Patrocles (Georges et al., 2006) were used to determine if any of these *HNF4A* 3'-UTR SNPs that exist in miRNA target sites are predicted to affect the base-pairing between the *HNF4A* mRNA and the predicted miRNAs. The results indicated the presence of 5 SNPs in the *HNF4A* gene that were predicted to alter miRNA targeting (Table 4.2). The analysis

predicted that SNPs destroyed 3 miRNA target sites and created 2 new target sites for miRNAs.

**Table 4.2: HNF4A 3'-UTR SNPs predicted to destroy or create a miRNA target site.**

dbSNP rs id	Minor allele frequency	Seed sequence [Ancestral/ Derived allele]	Old target destroyed	New target created
rs11574744	0-2.5 and 0-4.0% <sup>a</sup>	AC[T/A]GCCA	-34a, -34c, -449, -449b	
rs11574745	0-2.5%	A[C/T]CTTCA	-493-3p	
rs6130615	11-61% and 30% <sup>a</sup>	TAATG[C/T]G		-323
rs6103734	0% <sup>b</sup>	GTCA[G/A]GA	-378	
rs6103735	0% <sup>b</sup>	TTAA[G/A]GA		-302b*

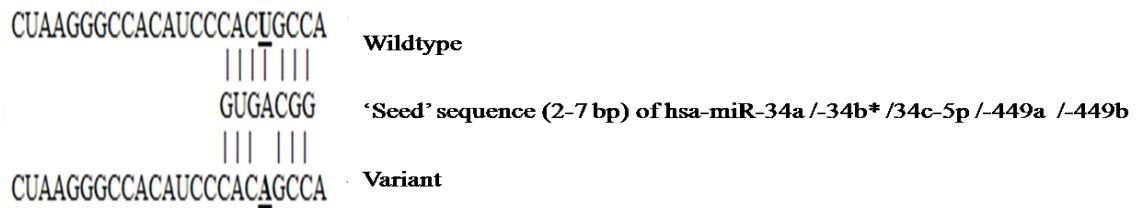
Note: <sup>a</sup> Seattle SNPs database was used in addition to dbSNP to identify the minor allele frequency (MAF); <sup>b</sup> The MAF appears to be 0% in over 200 individuals that have been genotyped for these 2 SNPs.

Since there were no reported SNPs reported in the 3'-UTR of the *CYP2D6* gene (data from 'Human Cytochrome P450 Allele Nomenclature Committee and dbSNP), the *CYP2D6* 3'-UTR was resequenced in 48 Caucasian and 48 African American DNA samples from the Coriell diversity panel. No polymorphisms were detected in the *CYP2D6* 3'-UTR in these two populations.

### ***In vitro* validation of SNP predictions**

In order to test the hypothesis that germline variations can alter the miRNA targeting of HNF4A, a SNP (rs11574744; T>A) in the HNF4A 3'-UTR was selected for further *in vitro* validation (Figure 4.11). This SNP was predicted by both PolymiRTS and Patrocles algorithms to destroy a miRNA binding site (Table 4.2).

**Figure 4.11: HNF4A 3'-UTR SNP predicted to destroy a miRNA binding site.**



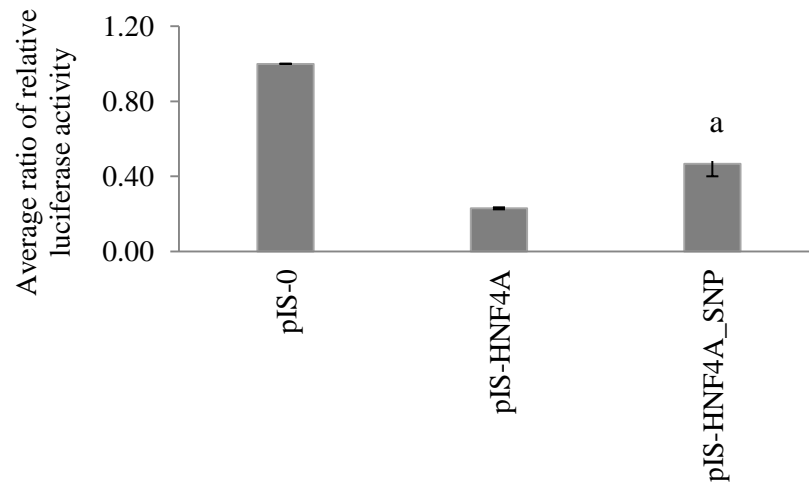
Note: Predicted miRNA interaction with wildtype and variant (rs11574744; T>A) in the HNF4A 3'-UTR.

In the first functional study, the activities of pIS-0, pIS-HNF4A and pIS-HNF4A\_SNP plasmids were compared. The luciferase activity from the plasmid with the variant HNF4A was two-fold higher than the wild type pIS-HNF4A plasmid (Figure 4.12 A;  $p < .05$ ). In the second functional study, the plasmids (pIS-0 or pIS-HNF4A or pIS-HNF4A\_SNP) were cotransfected with hsa-miR-34a, hsa-miR-449a and negative control (cel-miR-67). In the pIS-HNF4A plasmid, when compared to the negative control miRNA (cel-miR-67), hsa-miR-34a, and hsa-miR-449a reduced the pIS-HNF4A luciferase activity by 25% and 28% respectively; while in the pIS-HNF4A\_SNP plasmid, the luciferase activity was reduced by only 6% and 9%, respectively (Figure 4.12 B). The luciferase activity was significantly higher ( $p < .05$ ) in the pIS-HNF4A\_SNP plasmid

transfected cells, compared to the pIS-HNF4A, when transfected with either hsa-miR-34a or hsa-miR-449a (Figure 4.12 B).

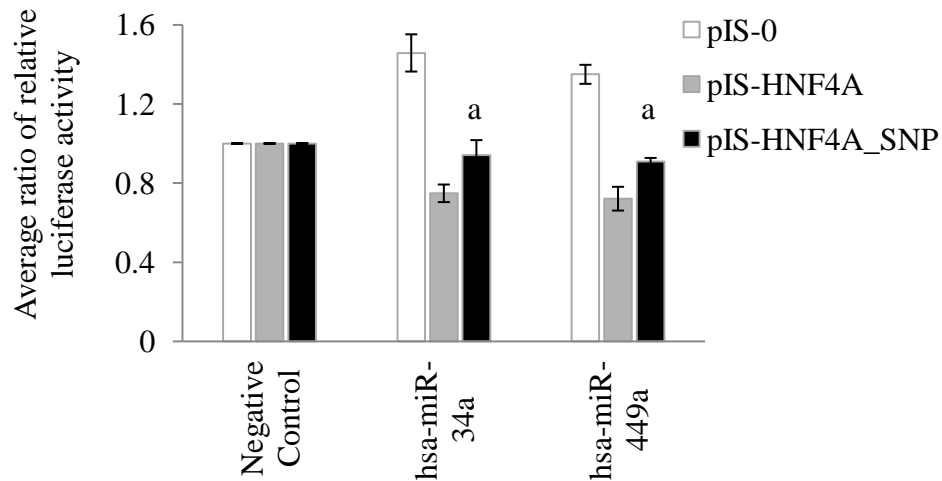
**Figure 4.12: *In vitro* validation of SNP predictions.**

(A)





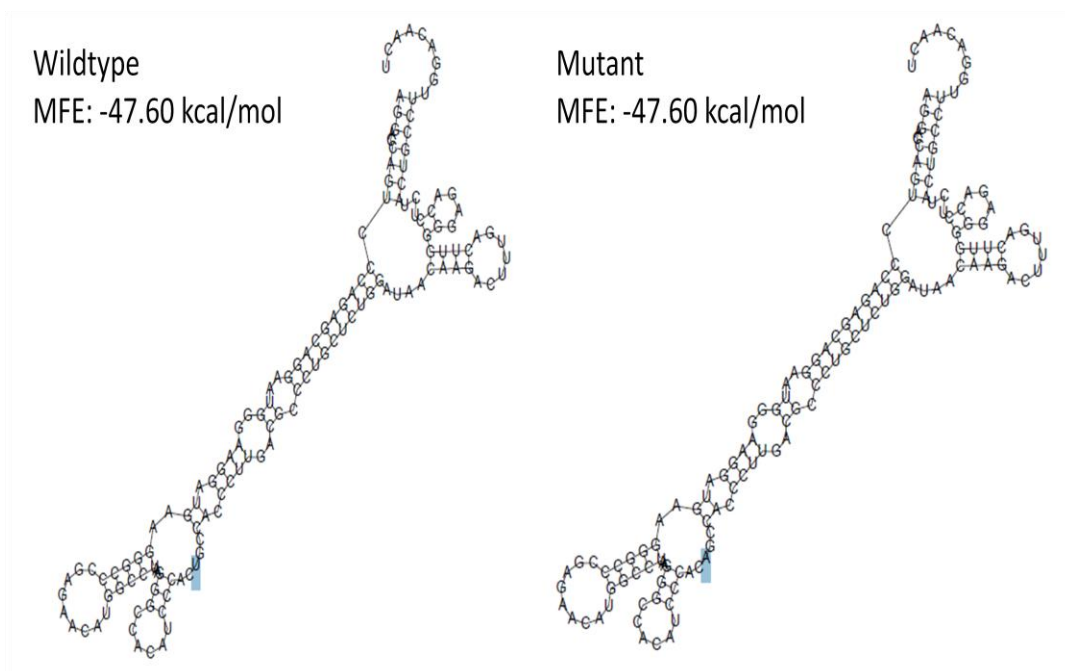
(B)



Note: (A) Transfection of HeLa cell line with 200 ng of control pIS-0 plasmid or pIS-HNF4A or pIS-HNF4A\_SNP constructs along with *Renilla* luciferase plasmid as internal control. Dual luciferase assays were performed at 24 hours. Data are expressed as the pIS-HNF4A and pIS-HNF4A\_SNP luciferase activity corrected for *Renilla* luciferase and normalized to the pIS-0 within each experiment (mean  $\pm$  SEM; n = 3 independent experiments performed in triplicates); (B) HeLa cells were cotransfected with 4  $\mu$ g of pIS-0, pIS-HNF4A, or pIS-HNF4A\_SNP luciferase constructs, along with *Renilla* reporter plasmid for normalization. The cells were also transfected with miRNA Mimics (hsa-miR-34a or hsa-miR449a; 30 nM). All data are expressed as the pIS-HNF4A or pIS-HNF4A\_SNP luciferase activity corrected for *Renilla* luciferase and normalized to negative control (cel-miR-67) within each experiment (mean  $\pm$  SEM; n = 3 independent experiments performed in triplicates). <sup>a</sup> indicates p < .05 when compared to the wildtype pIS-HNF4A plasmid.

RNAFold program (Gruber et al., 2008) was used to determine if the rs11574744 SNP changed the predicted mRNA secondary structure. A 70 bp nucleotide flank on either side of the SNP (Kertesz et al., 2007) was included to assess the minimum folding energy and secondary structure. The MFE was the same for both the wildtype and SNP sequences (Figure 4.13). Similarly, no differences were observed when the flanking sequence was extended to 200 bp on either side of the SNP.

**Figure 4.13: The SNP rs11574744 is not predicted to affect the local secondary structure of HNF4A mRNA.**



Note: Secondary structures of wildtype and variant (highlighted in blue) HNF4A mRNAs were predicted using RNAFold (Gruber et al., 2008). A 70 bp flanking either side of the SNP (rs11574744) was used to assess the minimum folding energy and secondary structure.

To determine the genotype frequency of this SNP in different ethnicities, a custom TaqMan assay (ABI, Forest City, CA, USA) was designed to genotype the Coriell human diversity panel comprising of Caucasian, African American and Asian DNA samples. The results of our genotyping suggest that the SNP is present only in African American samples at a minor allele frequency (MAF) of ~3% (Table 4.3).

**Table 4.3: Minor allele frequency of HNF4A 3'-UTR SNP, rs11574744.**

<b>Ethnicity</b>	<b>No of samples genotyped</b>	<b>SNP present</b>	<b>Minor allele frequency (%)</b>
Asian - Chinese	43	No	-
Asian - Japanese	44	No	-
Caucasian	94	No	-
African American	89	Yes	3.4

## **Discussion**

Interindividual variability in drug metabolism remains a significant contributor to differences in drug efficacy and toxicity. Some of this variability is due to known genetic variations and environmental factors that inhibit the enzymes or alter their expression levels. However, the mechanisms that underlie much of the variability are yet unknown. The studies presented here are the first steps in determining the contribution of miRNA to the regulation of drug metabolizing enzymes and ultimately their contribution to variability in drug metabolism.

The CYP2D6 enzyme metabolizes 20-25% of commonly prescribed drugs (Ingelman-Sundberg, 2005). The *CYP2D6* gene polymorphisms and CYP2D6 enzyme inhibiting drugs lead to variability in CYP2D6 enzyme activity and consequently, affect drug metabolism (Ingelman-Sundberg, 2005; Eichelbaum et al., 2006). While, genetic and environmental factors explain only a part of the interindividual variability in CYP2D6 enzyme activity, substantial unexplained variability still remains. For example, even within the same *CYP2D6* genotype group (e.g., extensive metabolizers - EM, or

intermediate metabolizers - IM) there is still unexplained interindividual variability (Borges et al., 2006). The hypothesis is that the substantial unexplained variability in CYP2D6 enzyme activity may be due to direct regulation of CYP2D6 by miRNAs, as well as indirect regulation of it through the regulation of the transcriptional factor, HNF4A.

The results of the bioinformatic analyses suggested that CYP2D6 is likely to be regulated by miRNAs (Table 4.1). In addition to directly targeting CYP2D6, the miRNAs were also predicted to target CYP2D6 indirectly by targeting one of its transcriptional regulators, HNF4A. The HNF4A is an important transcriptional ‘master regulator’ that regulates the expression of genes that metabolize drugs and xenobiotics; these genes include phase I enzymes, phase II enzymes, transporters and transcriptional factors (Kamiyama et al., 2007). Thus, regulation of HNF4A by miRNAs would likely affect many genes involved in drug disposition. HNF4A is also expressed in the kidney, intestine and pancreas; in those tissues, it controls lipid (Hayhurst et al., 2001) and glucose metabolism (Stoffel and Duncan, 1997). Nine different isoforms of HNF4A have been reported (Harries et al., 2008). Since the predominate isoform (isoform 2) expressed in the liver is a full length isoform with 3’-UTR (Ihara et al., 2005), it is one that is likely to be regulated by miRNAs. Recent studies have shown that PXR, another transcription factor that regulates the CYP450 gene expression, is also a target of miRNAs (Takagi et al., 2008).

As with many bioinformatic predictions, there was substantial variability between different algorithms in the miRNAs that are predicted target both the genes (Table 4.1). Part of this variability may be due to the different miRBase Registry versions that are

used by each algorithm; they ranged from versions 9 to 11. The variability may also be due to differences in the algorithms; these include differences in parameters such as, degree of complementarity, differences in UTR annotations and species conservation. For example, three of the programs (miRanda, miRBase Targets and PicTar) use evolutionary conservation parameter. Since the *CYP2D6* gene is not highly conserved across species (Nelson et al., 2004), this may contribute to the variability between the different algorithms. The total number of miRNAs predicted by the algorithms will continue to change as more miRNAs are being discovered and as the prediction algorithms are fine-tuned. The total number of miRNAs predicted is a starting point for subsequent analysis; expression of these miRNAs can also vary from tissue to tissue.

Based on the *in silico* analyses, further *in vitro* studies were performed. None of the five miRNAs that were predicted to target *CYP2D6* affected the expression of pIS-2D6 plasmid (Figure 4.3 and 4.4). For HNF4A, five of the miRNAs regulated the pIS-HNF4A luciferase plasmids (Figure 4.3 and 4.5), and at least two of those also regulate the HNF4A protein expression (Figure 4.7). Since the miRNAs did not reduce the HNF4A mRNA levels (Figure 4.6), they are most likely to be regulating HNF4A expression by blocking HNF4A protein synthesis rather than degradation of mRNA. Since *CYP2D6* is strongly regulated by HNF4A, the miRNAs that target HNF4A likely also affect the expression of *CYP2D6* and other downstream HNF4A target genes. This is supported by the results showing that the HNF4A targeting miRNAs suppressed the expression of another downstream target, PXR. Although the bioinformatic analysis indicated that PXR may also be a direct target of these miRNAs, the targets contained several mismatches indicating that they are more likely to block PXR translation rather

than cause mRNA degradation. There was no observable effect of those miRNAs on CYP2D6; the likely reason may be because CYP2D6 is expressed at very low levels in the HepG2 cell line and, consequently, not strongly regulated by HNF4A in these cells.

Of the miRNAs that target HNF4A, hsa-miR-34a expression appears to be higher than the other miRNAs in the primary hepatocytes preparations. It is possible that the miRNAs that target HNF4A may be down-regulated to allow full HNF4A expression during normal conditions and then specifically up-regulated when HNF4A needs to be repressed. Additional studies will be required to determine if any of these miRNAs are regulated by environmental factors that control HNF4A and CYP450 gene expression.

MicroRNA functions can be altered by genetic variants that affect the miRNA binding to the target mRNA; these variants are called miRSNPs (Mishra et al., 2007). MiRSNPs have been shown to alter miRNA function and hence enzyme activity. This has been shown to be the case for genes such as dihydrofolate reductase (Mishra et al., 2007) and the estrogen receptor  $\alpha$  (Adams et al., 2007). Genetic variants in the 3'-UTRs of drug metabolizing genes have not typically been given high priority in functional studies; however, based on these analyses, they may contribute to interindividual variability in the expression of the drug metabolizing enzymes. Some studies have reported on the SNPs in the 3'-UTRs of CYP450 genes that may be associated with altered phenotypes; these include CYP19A1 (Dunning et al., 2004) and CYP2A6 (Wang et al., 2006). It is conceivable that these SNPs may be a target of miRNAs. In our analyses, few SNPs were identified in *HNF4A* gene 3'-UTR that are predicted to alter the miRNA targets (Table 4.2). The *in vitro* luciferase assay results for HNF4A suggest that such genetic variants can cause altered mRNA-miRNA interactions (Figure 4.12 A and

B). Further functional studies will be required to determine the clinical phenotypic consequences of the miRSNPs. Both the programs that were used (PolymiRTS and Patrocles) only predict an effect if the SNPs in the ‘seed’ region of the miRNA target sites. However, SNPs in ‘non-seed’ regions can also affect miRNAs that bind either upstream or downstream to the SNP (Mishra et al., 2007). Neither of these programs predict such loss or gain of such mRNA-miRNA interactions. Similarly, SNPs in the mature miRNAs and pre-miRNA may also affect the mRNA-miRNA interaction. Since very few of the miRNA genes have been resequenced in depth, the genetic variants in those genes are not well characterized. Therefore, these SNPs were not included in this analysis. There were no SNPs in the CYP2D6 3’-UTR. The absence of any SNPs in this region suggests that it is highly conserved and may be important for CYP2D6 mRNA regulation; however, the short length (75 bp) makes it impossible to draw any definite conclusions about the conservation.

These findings confirm a potential role for miRNAs in the regulation of CYP450s involved in drug metabolism. It is conceivable that other genes, including Phase II enzymes, drug targets, and other drug transporters are also targets of the miRNAs. Recent studies have shown that other proteins involved in drug metabolism (Tsuchiya et al., 2006; Takagi et al., 2008; To et al., 2008; Pan et al., 2009) are subject to miRNA regulation. These results, taken together with previous findings suggest a complex regulatory mechanism for drug metabolizing enzymes by miRNAs. The identification of the endogenous hepatic miRNAs that regulate CYP2D6 directly or indirectly should help us understand the variability in therapeutic efficacy and toxicity for patients to several commonly used drugs. Further, identifying polymorphisms that alter the drug

metabolizing mRNA-miRNA interactions would likely be a clinically important biomarker for guiding the use of CYP2D6 metabolized drugs. Ultimately, these novel biomarkers would help improve the efficacy and reduce the side effects of the commonly prescribed drugs.

## **Materials and Methods**

**Cell culture:** HeLa (a cervical cancer cell line) and HepG2 (a liver cancer cell line) that were used in our *in vitro* transfection assays were obtained from ATCC (American Type Culture Collection; Manassas, VA, USA). All tissue culture reagents were purchased from Invitrogen (Carlsbad, CA, USA). HeLa and HepG2 cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum and 1% Penicillin-Streptomycin. The cells were maintained at 37°C in a humidified atmosphere containing 5% CO<sub>2</sub>.

**Generation of luciferase reporter gene constructs:** The pIS-0 vector (Yekta et al., 2004; Addgene plasmid #12178) was used to study 3'-UTR function. The 3'-UTR of CYP2D6 (72 bp of the full length 75 bp; nucleotides 5837-5908 of GenBank id M33388.1) was synthesized as an oligonucleotide with SacI and NheI restriction sites and subcloned into the pIS-0 vector using those restriction sites (GenScript, Piscataway, NJ, USA). The 3'-UTR of HNF4A (1448 bp of the full length 1724 bp; nucleotides 1558-3005 of NM\_000457) was amplified using genomic DNA from the Coriell panel using primers (Integrated DNA Technologies, Coralville, IA, USA) with NheI and SacI restriction sites (FP: 5'GGTGTTGAGCTCCTAAGAGAGCACCTGGTGA and RP: 5'GGGTTTGCTAGCGGAGACCTGGGTTCAAG; the restriction sites are italicized and



the HNF4A sequence is underlined). The PCR product was cloned into the TOPO TA vector (Invitrogen, Carlsbad, CA, USA) and the insert sequence was verified by DNA sequencing. The sequencing data revealed that the clone had the “variant” rs322210 (C allele); since the reported frequency of this allele is 55% (dbSNP), it appears to be the more common allele. Therefore, this clone was used. The insert was then subcloned into pIS-0 vector using the NheI and SacI restriction sites and transformed into DH5 $\alpha$  competent cells. Colonies with the inserts were identified by restriction digestion and the sequence was verified by direct DNA sequencing. Plasmids were purified using the Plasmid Maxi Kit (Qiagen, Valencia, CA, USA). The DNA concentrations were determined using QuantIT DNA Broad Range kit (Invitrogen, Carlsbad, CA, USA) and verified on the BioAnalyzer 2100 (Agilent Technologies, Palo Alto, CA, USA). The control vector is referred to as pIS-0 and the vectors with HNF4A and CYP2D6 3'-UTR sequences inserted are referred to as pIS-HNF4A and pIS-2D6, respectively.

A plasmid with the variant HNF4A 3'-UTR (rs11574744) was created by site-directed mutagenesis (GenScript, Piscataway, NJ, USA). Presence of this mutation was confirmed by resequencing the plasmid. This mutant plasmid is referred to as pIS-HNF4A\_SNP.

**Transfection:** For luciferase assays,  $0.9 \times 10^5$  HeLa cells were seeded into each well of a 24 well plate. The cells were transfected with one of the three pIS plasmids (pIS-0, or pIS-2D6, or pIS-HNF4A) at three different concentrations of the plasmids (100, or 200, or 400 ng each). *Renilla* luciferase reporter plasmid pGL4.74 (hLuc-TK) was used as a transfection control; the ratio of pIS: pGL4.74 was 50:1. Transfection was performed using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) following the

manufacturer's instructions using Opti-MEM (Invitrogen, Carlsbad, CA, USA) and culture media without any antibiotics. At 24 hours after transfection, the cells were harvested and dual luciferase assays (Promega, Madison, WI, USA) were performed as per manufacturer's instructions. Transfections for the site-directed mutagenesis experiments were performed using the same protocol using 200 ng of pIS-0, or pIS-HNF4A, or pIS-HNF4A\_SNP plasmids.

For luciferase and miRNA cotransfection experiments,  $1.5 \times 10^6$  HeLa cells were seeded in a T-25 flask. At 24 hours, the cells were then transfected with 4  $\mu$ g of pIS-0, or pIS-2D6, or pIS-HNF4A plasmid, along with 80 ng of *Renilla* luciferase reporter plasmid as transfection control. At 24 hours after transfection, the cells were trypsinized and counted using a Z1 Particle Coulter Counter (Beckman Coulter, Fullerton, CA, USA). Approximately  $1 \times 10^5$  cells were seeded into each well of a 24 well plate. The cells were reverse transfected (that is, plated onto the well containing the transfection mix) with either 30 nM of the miRNA, or combinations of miRNAs or the negative control (miRIDIAN Mimics from Dharmacon, Lafayette, CO, USA) using Lipofectamine 2000 transfection reagent (Invitrogen, Carlsbad, CA, USA). At 24 hours post miRNA transfection, the cells were harvested and dual luciferase assays were performed. Cotransfections for the site-directed mutagenesis experiment were performed with the same protocol using the pIS-0, or pIS-HNF4A, or pIS-HNF4A\_SNP plasmids and specific miRNAs. All transfections and luciferase assays were performed in triplicates over three different days.

In order to verify that the transfected synthetic miRIDIAN Mimics (Dharmacon, Lafayette, CO, USA) were efficiently taken up by the cells, HeLa cells ( $2 \times 10^5$  cells per

well in 96-well plates) were transfected with 30nM of either hsa-miR-34a mimic or negative control mimic (cel-miR-67). RNA isolation and reverse transcription was performed using microRNA Cell-to-Ct kit (ABI, Forest City, CA, USA) and specific TaqMan MicroRNA Assays (ABI, Forest City, CA, USA) for hsa-miR-34a, U6 snRNA (endogenous control) and hsa-miR-449a (a non-specific control) following the manufacturer's instructions. Specific details of the PCR are in the 'Quantification of miRNAs' section. The relative quantities of miRNA were calculated using the  $\Delta\Delta CT$  method using U6 snRNA and negative control (cel-miR-67) transfections as controls. The relative miRNA expressions are reported as  $2^{-(\Delta\Delta CT)}$  (Kreuzer et al., 1999). The miRNA transfections were replicated on three separate days and the RT-PCR was done in triplicates for each of the transfections.

**Bioinformatic analysis to predict microRNAs:** Six different web-based bioinformatic programs were used to identify the miRNAs predicted to target CYP2D6 and HNF4A. The programs included miRanda (John et al., 2004), miRBase Targets (Griffiths-Jones et al., 2006), TargetScan (Lewis et al., 2003), PicTar (Krek et al., 2005), RNA22 (Miranda et al., 2006) and PITA (Kertesz et al., 2007). The parameter settings were either default or those used in the publications describing the tools. For RNA22, which is a downloadable program with user defined mRNA and miRNA sequences, *CYP2D6* gene reference sequence id was identified from the Human Cytochrome P450 Allele Nomenclature Committee ([www.cypalleles.ki.se/cyp2d6.htm](http://www.cypalleles.ki.se/cyp2d6.htm)) and the 3'-UTR sequence was identified using the UCSC Genome Browser (<http://genome.ucsc.edu>). The mature miRNA sequences (version 10.0) were downloaded from the miRBase Sequence database (Griffiths-Jones et al., 2006).

**Identification of SNPs located in the 3'-UTR:** SNPs from the 3'-UTR were obtained from NCBI SNP database (dbSNP; [www.ncbi.nlm.nih.gov/projects/SNP](http://www.ncbi.nlm.nih.gov/projects/SNP)) and UCSC Genome Browser database. The minor allele frequencies (MAF) were obtained from dbSNP database for CYP2D6 and HNF4A genes. Additionally, Seattle SNPs database (<http://pga.gs.washington.edu>) was used for identifying the MAF for HNF4A.

**Bioinformatic analysis to predict the effect of SNPs on mRNA-miRNA interactions:** Two programs - PolymiRTS (Bao et al., 2007) and Patrocles (Georges et al., 2006) were used to predict the effect of miRSNPs in the 3'-UTR of CYP2D6 and HNF4A mRNA on the mRNA-miRNA interaction.

**RNA isolation:** Fresh human hepatocytes were isolated by Vitacyte LLC, (Indianapolis, IN, USA) from liver specimens that were collected after Indiana University's Institutional Review Board (IRB) approval. These hepatocytes were flash frozen until RNA isolation. Total RNA, including the small RNA was isolated using the miRNeasy kit (Qiagen, Valencia, CA, USA) following the manufacturer's instructions, except, the phase separation step was done using the maXtract tubes (Qiagen, Valencia, CA, USA). The on-column DNase treatment step was included in the RNA isolation procedure and was done using the DNase set (Qiagen, Valencia, CA, USA). The RNA yield was determined using the Quant-iT RNA Broad Range assay kit (Invitrogen, Carlsbad, CA, USA) in the Qubit fluorometer (Invitrogen, Carlsbad, CA, USA). RNA quality/integrity was assessed using RNA 6000 Nano Labchip in the BioAnalyzer 2100 (Agilent Technologies, Palo Alto, CA, USA).

**Quantification of miRNAs:** The expression of miRNAs was analyzed using specific TaqMan MicroRNA Assays (ABI, Forest City, CA, USA) following the

manufacturer's instructions in a StepOne Plus real time PCR instrument (ABI, Forest City, CA, USA). Briefly, total RNA, including the small RNA, was analyzed for specific miRNA expression in a two-step protocol. The first step, reverse transcription, was performed using a miRNA specific primer which confers both specificity and sensitivity. In the second step, Real-Time PCR was performed with TaqMan probes that amplify only the target mature miRNA. The PCR steps included a 10 min polymerase enzyme activation at 95°C, 50 cycles each of denaturation at 95°C for 15 s, and annealing/extension at 60°C for 1 min. Relative quantities of miRNA were calculated using the  $\Delta$ CT method using U6 snRNA as endogenous control. That is, the relative miRNA expression was calculated as  $2^{-(\Delta CT)}$  and was multiplied by  $10^3$  to simplify data presentation.

**Genotyping of HNF4A 3'-UTR SNP:** A custom TaqMan SNP genotyping assay (ABI, Forest City, CA, USA) was developed for HNF4A 3'-UTR SNP (rs11574744). The primer and probe sequences were: CCCGAGAACATGGCCTAAGG as forward primer, CCAGAGCAGGGCGTCAA as reverse primer, VIC-ATCCCACAGCCACCC as probe 1 and FAM-ATCCCACTGCCACCC as probe 2 (variant sequence is underlined). Genotyping was performed with the Coriell DNA Diversity panel genomic DNA samples (African American, Caucasian and Asian) in a StepOne Plus Real-Time PCR System (ABI, Forest City, CA, USA) using the recommended genotyping PCR conditions i.e., an initial denaturation for 10 min at 95°C, followed by 40 cycles of denaturation for 15 s at 92°C and annealing/extension for 1 min at 60°C. The genotyping assay results were confirmed by resequencing 96 DNA samples (48 Caucasian and 48 African American) from the Coriell biorepository. DNA resequencing was performed

using the commercial DNA sequencing services of Polymorphic DNA Technologies Inc (Alameda, CA, USA).

**Resequencing of CYP2D6 3'-UTR:** In order to resequence the CYP2D6 3'-UTR, CYP2D6 PCR amplicon was generated using a forward primer (2D6 int6x2F: 5'AGGAGGCAAGAAGGAGTGTTCAGG) that binds to intron 6 and a reverse primer (2D6 R/2: 5'ACTGAGCCCTGGGAGGTAGGTAG) that binds downstream of the *CYP2D6* gene (Gaedigk et al., 2007). These primers specifically amplify 1668 bp of the *CYP2D6* gene and not the *CYP2D7* or *CYP2D8* gene sequences. To sequence the 3'-UTR in the long template PCR fragment, a forward primer (E9seq: 5'CTTCCGTGGAGTCTTGCAGG; Gaedigk, A., personal communication) that binds to exon 9, and/or a reverse primer (UTR Seq RP: 5'CTCAGCCTCAACATACCCCT) that binds downstream of the 3'-UTR were used. Resequencing was performed using DNA from the Coriell Diversity Panels, including 48 Caucasian and 48 African American samples. All sequencing of the *CYP2D6* gene was done at the DNA Sequencing Core Facility of Indiana University.

**Quantification of HNF4A protein and mRNA:** In a 6-well plate,  $1 \times 10^6$  HepG2 cells/ml were reverse transfected with 100 nM of synthetic miRNA Mimics (hsa-miR-34a, hsa-miR-449a, hsa-miR-493\* and control cel-miR-67; miRIDIAN microRNA Mimics from Dharmacon, Lafayette, CO, USA) using siPORT NeoFX transfection reagent (Ambion, Austin, TX, USA) following the manufacturer's specifications. As described previously (Iwazaki et al., 2008), a Genome-Wide siRNA for human HNF4A (Hs\_HNF4A\_9) and an AllStars Negative Control siRNA (Qiagen, Valencia, CA, USA) were also used as process controls. At 72 hours of transfection, the cells were harvested

for HNF4A protein and RNA analyses. The transfections were repeated on three separate days.

For protein analyses, nuclear protein extract was isolated using NucBuster protein extraction kit (Novagen, Madison, WI, USA). Protein concentrations were determined using bicinchoninic acid (BCA) reagent protein assay kit (Pierce, Bradford, IL, USA). The nuclear protein lysate (20  $\mu$ g) was electrophoresed using a 4-20% Tris-Glycine gel (Invitrogen, Carlsbad, CA, USA) and transferred to a Immobilon polyvinylidene fluoride (PVDF) transfer membrane (Millipore, Bedford, MA, USA) using Novex semi-dry blotter (Invitrogen, Carlsbad, CA, USA). The membranes were blocked using 3% milk and then incubated overnight at 4°C with mouse anti-human HNF4A primary antibody (Perseus Proteomics, Tokyo, Japan) at a dilution of 1:5000, followed by horseradish peroxidase (HRP) conjugated ImmunoPure goat anti-mouse secondary antibody (Pierce, Bradford, IL, USA) at a dilution of 1:10000 for 1 hour. HRP conjugated beta-actin antibody (Abcam, Cambridge, MA, USA) at a dilution of 1:5000 was used as the loading control. All the three antibodies were diluted in Starting Block (T20) blocking buffer (Pierce, Bradford, IL, USA). Protein bands were developed using a SuperSignal enhanced chemiluminescence kit (Pierce, Rockford, IL, USA). The protein bands were visualized using a LAS-1000 plus system (Fujifilm, Tokyo, Japan) and quantified using the Multi Gauge software (Fujifilm, Tokyo, Japan).

For RNA analyses, total RNA, including the small RNA, was isolated as described above. The RNA yield was determined using the Quant-iT RNA Broad Range assay kit (Invitrogen, Carlsbad, CA, USA) in the Qubit fluorometer (Invitrogen, Carlsbad, CA, USA). The cDNA was generated from 1  $\mu$ g of total RNA using Reverse

Transcription System (Promega, Madison, WI, USA) following the manufacturer's instructions. The expression of HNF4A and GAPDH (endogenous control) mRNAs was analyzed using specific TaqMan Gene Expression Assays (ABI, Forest City, CA, USA) following the manufacturer's instructions in a StepOne Plus real time PCR instrument (ABI, Forest City, CA, USA). Relative quantity of HNF4A mRNA was calculated using the  $\Delta\Delta$ CT method using GAPDH and negative controls (cel-miR-67 and negative control siRNA) as reference controls. The RT-PCR was done in triplicate for each of the three transfections.

**Statistical analysis:** Statistical analyses were carried out as described in the figure legends, by ANOVA using SPSS software (version 13.0; SPSS Inc, Chicago, IL, USA). A p-value of  $p \leq 0.05$  was considered statistically significant.



## CHAPTER FIVE

### **Conclusion and Future Directions**

Cytochrome P450 2D6 (CYP2D6) is an important drug metabolizing enzyme that is involved in the metabolism of ~20-25% of commonly prescribed drugs (Ingelman-Sundberg, 2004). There is considerable interindividual variability in CYP2D6 mediated enzyme activity and consequently, there is variability in both drug metabolism and response (Ingelman-Sundberg, 2004). Only a part of this interindividual variability is explained by environmental and genetic factors (Ingelman-Sundberg, 2001; Eichelbaum et al., 2006); the cause of much of the variability is still unknown. Hence, it becomes important to study the endogenous and exogenous factors that regulate CYP2D6 enzyme activity.

The *overall goal* of this research was to gain insight into the contribution of genetic and epigenetic factors that drive the interindividual variability seen in CYP2D6 mediated drug metabolism and response. The *specific objective* of this research was to determine the role of *CYP2D6* genotype, copy number variations and microRNA that result in the variability seen in CYP2D6 mediated drug metabolism.

#### **Effect of *CYP2D6* gene polymorphisms on tamoxifen metabolism**

*CYP2D6* is a highly polymorphic gene (Ingelman-Sundberg, 2005). These polymorphisms lead to interindividual variability in CYP2D6 enzyme activity (Ingelman-Sundberg, 2005). This variability in enzyme activity results in the creation of extreme phenotype groups [poor metabolizers (PM) and ultrarapid metabolizers (UM)]; these groups have an increased risk for either therapeutic failure or adverse drug reaction. The

cost of treating patients with these extreme CYP2D6 phenotypes (UM and PM) is estimated to be on an average, \$4,000-6,000 greater than the cost of treating patients in the extensive metabolizer (EM) and intermediate metabolizer (IM) groups (Chou et al., 2000). Hence, prospective *CYP2D6* genotyping can help predict therapy outcome and the occurrences of adverse events (Zanger et al., 2004). Prospective *CYP2D6* genotyping is estimated to be beneficial in the treatment of ~10% of all clinically used drugs (Ingelman-Sundberg, 2005).

One of the important drug that is biotransformed by CYP2D6 is tamoxifen. Tamoxifen is used globally as a selective estrogen receptor modulator in the hormonal therapy of estrogen receptor (ER) positive breast cancers in both pre- and post-menopausal women. However, there is considerable interindividual variability in response to tamoxifen (Lonning et al., 1992; Ingle et al., 1999).

CYP2D6 is primarily responsible for the sequential generation of the active tamoxifen metabolite, endoxifen (Desta et al., 2004). Endoxifen concentration is influenced by polymorphisms in the *CYP2D6* gene (Stearns et al., 2003; Jin et al., 2005; Borges et al., 2006). *CYP2D6* gene polymorphisms are also associated with clinical outcomes of tamoxifen therapy (Goetz et al., 2005; Bonanni et al., 2006; Goetz et al., 2007; Gonzalez-Santiago et al., 2007; Lim et al., 2007; Schroth et al., 2007; Kiyotani et al., 2008; Xu et al., 2008). Published data on *CYP2D6* genotype-endoxifen concentration or *CYP2D6* genotype-outcome is largely from Caucasian women in the US and Europe. To date, three studies have been reported from Asia; these studies performed in Korea (Lim et al., 2007), Japan (Kiyotani et al., 2008) and China (Xu et al., 2008) have suggested that *CYP2D6\*10* allele is associated with poorer clinical outcome. However,

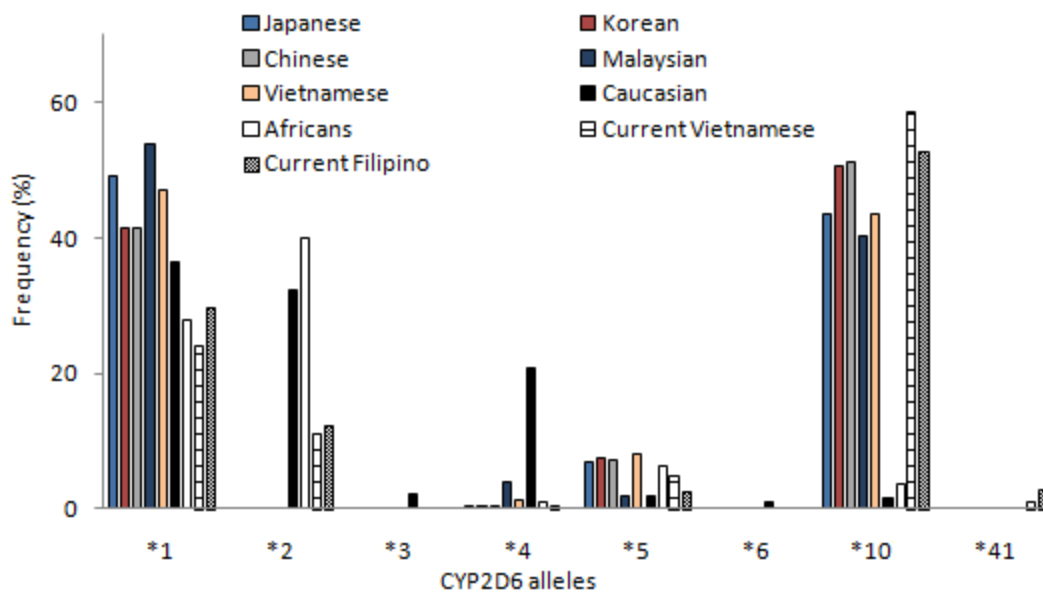
not all studies have confirmed this association between *CYP2D6* variant genotype and clinical outcomes (Nowell et al., 2005; Wegman et al., 2005; Wegman et al., 2007; Okishiro et al., 2009). Hence, prospective trials are needed to understand the association between *CYP2D6* gene variants, endoxifen concentration and clinical outcome.

In this study (Chapter Two), an association between *CYP2D6* genetic variants and the concentration of tamoxifen and its metabolites was analyzed in breast cancer patients undergoing adjuvant tamoxifen therapy in two Asian countries - Vietnam (N = 93) and the Philippines (N = 144). In both these Asian populations, *CYP2D6\*10* was the most frequent *CYP2D6* allele with a frequency of 59.1% and 53.6% in the Vietnamese and Filipino populations, respectively. The *CYP2D6* alleles that are frequent in Caucasians (*CYP2D6\*4*, *\*3*, *\*6* and *\*41*) were present in less than 3% in both the populations. The *CYP2D6* allelic frequencies in the Vietnamese and Filipino populations were comparable to that of other Asian populations previously reported (Kubota et al., 2000; Cai et al., 2007; Kiyotani et al., 2008; Lee et al., 2009; Veiga et al., 2009; Figure 5.1).

These *CYP2D6* polymorphisms that result in *CYP2D6* IM or PM status were significantly associated with impaired metabolism of tamoxifen; both endoxifen concentration and endoxifen/N-desmethyltamoxifen ratio were significantly different ( $p < .01$ ) when compared to the wildtype *CYP2D6* allele. These results are consistent with previous results from our group (Stearns et al., 2003; Jin et al., 2005; Borges et al., 2006) and also results reported by others in Caucasian (Gjerde et al., 2008) and Asian women (Lim et al., 2007). In these Vietnamese and Filipino women, the primary allele that was involved in decreased endoxifen concentration was the *CYP2D6\*10* allele. It is an intermediate metabolizer (IM) allele with reduced function. This allele is primarily

characterized by the ‘tag’ SNP 100C>T which results in an amino acid change P34C (Johansson et al., 1994). This amino acid change occurs in the conserved tetrapeptide, PPGP sequence (a proline rich region) that is important for folding the CYP450 enzymes. As a result, the CYP2D6\*10 enzyme becomes less stable and also decreases the affinity to CYP2D6 substrate drugs (Yu et al., 2002).

**Figure 5.1: *CYP2D6* allele frequencies in the current Vietnamese and Filipino study are comparable to other Asian populations but not to Caucasian or African populations.**



Source: Veiga, S., et al., 2009; Sachse, C., et al., 1997; Wennerholm, A., et al., 1999; current study involving Vietnamese and Filipino breast cancer patients (Chapter Two).

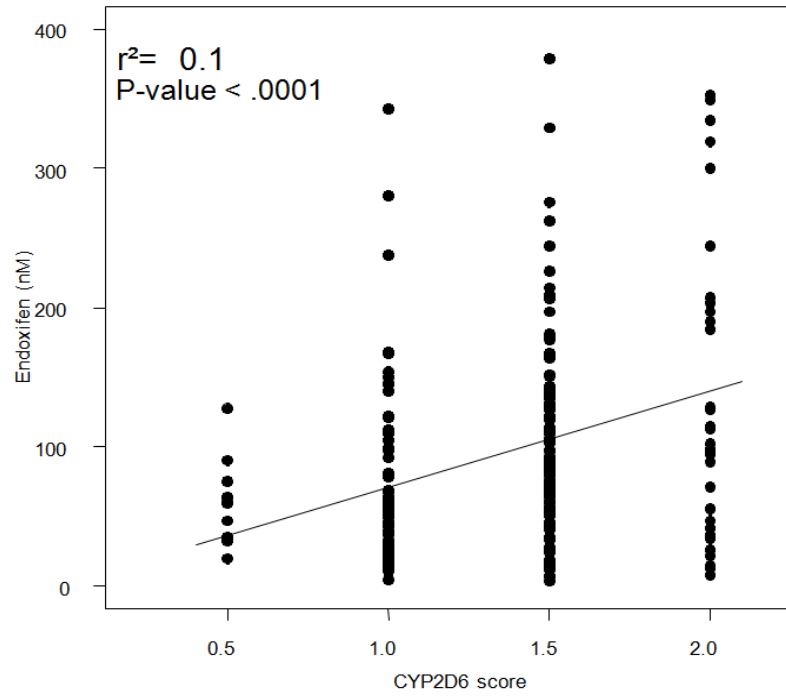
In summary, this study provides the first comprehensive analysis of the association between *CYP2D6* polymorphisms and the concentration of endoxifen in the Vietnamese and Filipino population. Like other ethnicities, the *CYP2D6* genetic

polymorphisms appear to affect endoxifen concentrations in these two Asian populations. In both the Vietnamese and Filipino populations, the *CYP2D6\*10* allele is the primarily responsible for the reduced plasma endoxifen concentration. Since several studies have shown that endoxifen appears to contribute to the activity of tamoxifen, these results indicate that *CYP2D6* genotyping may be useful predicting the efficacy of tamoxifen in Vietnamese and Filipino women. Furthermore, because CYP2D6 inhibiting drugs create a phenocopy of the *CYP2D6* poor metabolizer phenotype, they also would likely reduce endoxifen concentrations and ultimately increase breast cancer recurrence in these women. Although the final conclusions will depend on the recurrence outcomes from the ongoing clinical trial, the conclusion based on the current results is that it would be prudent for these women avoid CYP2D6 inhibitor drugs.

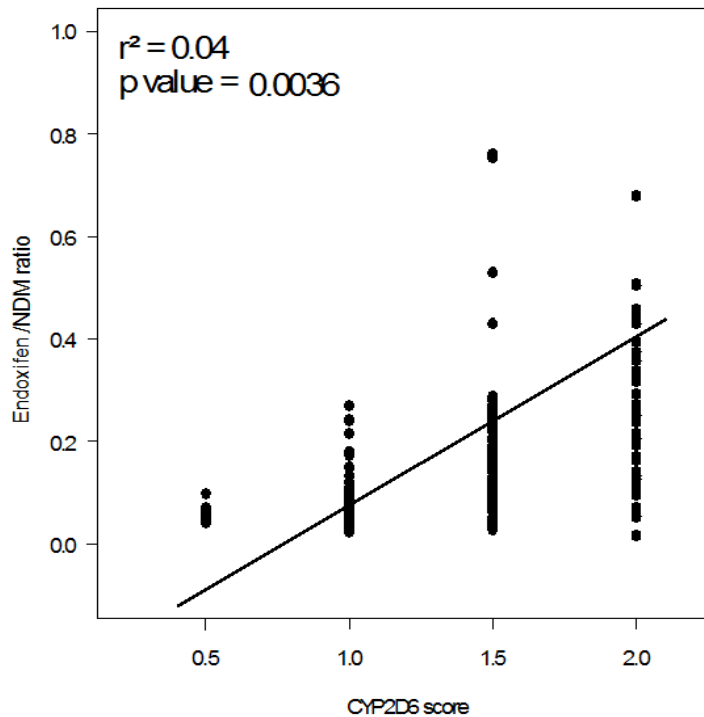
Even though the *CYP2D6* gene polymorphisms were significantly associated with endoxifen concentrations, considerable interindividual variability exists even after correcting for this factor (Figure 2.1 D and E). For example, within *CYP2D6\*10* group, there is over a 10-fold difference in plasma endoxifen concentration and endoxifen/NDM ratio. In the combined Vietnamese and Filipino cohort, the *CYP2D6* variants explain about 10% of the variability in the endoxifen concentration and 4% of the variability in the endoxifen/NDM concentration ratio (Figure 5.2 A and B).

**Figure 5.2: Association of *CYP2D6* genetic variants with tamoxifen metabolism in the combined Vietnamese and Filipino cohort based on regression analysis.**

**(A) Endoxifen concentration**



**(B) Endoxifen/NDM ratio**



Note: The association of *CYP2D6* gene variants with (A) endoxifen concentration and (B) endoxifen/NDM ratio in combined Vietnamese and Filipino cohort was measured by regression analysis. Individuals were classified into the different score groups based on their *CYP2D6* genotype (data in Chapter Two). The endoxifen concentrations and endoxifen/NDM ratios were log transformed for normalization.

A part of this interindividual variability may be explained by other *CYP2D6* rare alleles that were not tested (e.g., \*14, \*21, \*27, \*39, \*47, etc.). This variability may also be due differences in enzyme activity because of *CYP2D6* gene copy number variations. Hence, the next specific aim (Chapter Three) was designed to assess the association of *CYP2D6* gene copy numbers on plasma endoxifen concentrations.

## Effect of CYP2D6 gene copy number variations on enzyme activity

One of the primary sources of variation in the human genome is the copy number variations (CNV) involving deletion or multiplication of DNA segments (Sebat et al., 2004). Recent studies have suggested an increasing role for CNVs in many diseases including cancer, developmental diseases, mental illness, autoimmune diseases and infectious diseases (Shrestha et al., 2009; Wain et al., 2009). A number of drug metabolizing enzymes also exhibit gene copy number variations. These include *UGT2B17*, *UGT2B28* (McCarroll et al., 2006), *SULT1A1* (Hebring et al., 2007), *GSTT1*, *GSTM1* (Bolt and Thier, 2006) and *CYP2D6* (Bradford, 2002; Sistonen et al., 2007).

The *CYP2D6* gene CNVs affect *CYP2D6* enzyme activity and consequently, alter drug metabolism and response, including both safety and efficacy. These *CYP2D6* CNVs can either lead to reduced metabolism (gene deletion) or increased metabolism (>2 functional copies). This phenotypic variability in *CYP2D6* enzyme activity arises from the variety of *CYP2D6* alleles that exist in CNVs. For example, (a) the carriers of *CYP2D6*\*5 (gene deletion) have no enzyme activity and are considered to be PMs; (b) the carriers of multiple copies of functional *CYP2D6* alleles (\*1xN, \*2xN) have increased enzyme activity and are classified as UMs; (c) the carriers of \*10x2 do not have increased enzyme activity when compared to \*10 and are *still* considered to be IMs; and (d) carriers of *CYP2D6*\*36x2 and *CYP2D6*\*4x2 have negligible enzyme activity and are *still* considered to be PMs. However, little is known about the phenotypic consequence of other complex multiple copy arrangements including \*36-\*10-\*10 or \*36-\*36-\*10, etc.



In different ethnic populations (Caucasians, Asians and Africans), different *CYP2D6* alleles are involved in gene multiplication and there is also variability in the frequencies of these *CYP2D6* CNVs (Bradford, 2002; Sistonen et al., 2007). For example, CNVs involving the \*36-\*10 allele are frequent in Japanese and Chinese population, but not in Caucasians or Africans (Gaedigk et al., 2006; Hosono et al., 2009). To date, not much is known about the presence of *CYP2D6* gene CNVs in the Vietnamese and Filipino populations. Hence, in the prospective clinical trial involving the Vietnamese and Filipino breast cancer patients, quantitative RT-PCR assays were performed to identify the *CYP2D6* gene CNVs and the presence of *CYP2D6*\*36 alleles. The *objective* was to investigate the association of *CYP2D6* CNVs with the steady-state concentrations of tamoxifen and its metabolites.

The results of this study indicated that in the combined cohort, only 39% of the individuals carried 2 copies of the *CYP2D6* gene. Approximately 61% of the individuals had a genomic imbalance for *CYP2D6*: 3.8% had a *CYP2D6* gene deletion (*CYP2D6*\*5) and 57.2% of the individuals had more than two copies of the *CYP2D6* gene. These results are comparable to the *CYP2D6* CNVs reported in the Japanese and Chinese populations (Hosono et al., 2009). The CNV phaser program (Kato et al., 2008) was used to estimate the *CYP2D6* diplotypes. The results indicated that 34.2% of the individuals had CNVs involving one or more copies of *CYP2D6*\*36 allele along with a copy of *CYP2D6*\*10 allele, i.e., *CYP2D6*(\*36)<sub>n</sub>-\*10 allelic arrangement. In most of these individuals (99.8% of the cases), the *CYP2D6*(\*36)<sub>n</sub> allele was present along with a copy of a *CYP2D6*\*10 allele (Figure 5.3). Approximately 50% of the diplotypes in the combined populations involved *CYP2D6*(\*36)<sub>n</sub> allelic CNVs.

Figure 5.3: Schematic representation of CNVs involving some of the *CYP2D6* alleles.

CYP2D6 allele	Frequency %	Schematic representation
*1	27.4	
*10	17.7	
*36-*10	32.9	
*36-*36-*10	4.3	
*36-*36	0.6	
*36-*10-*10	0.6	
*36	0	

Note: Schematic representation of *CYP2D6* alleles showing the arrangements of copy number variations involving the *CYP2D6*\*36 allele. The shaded region in *CYP2D6*\*36 represents gene recombination with *CYP2D7* in the exon 9 region. The two highly homologous pseudogenes - *CYP2D7p* and *CYP2D8p* that are present 5' of the *CYP2D6* gene are also shown.

In order to determine the phenotypic consequence of the CNVs involving the *CYP2D6*(\*36)<sub>n</sub> allele, tamoxifen and its metabolite concentration were compared between following: \*10 vs. \*36-\*10 vs. \*36-\*36-\*10 and also \*10 vs. (\*36)<sub>n</sub>-\*10. The results indicated that, based on the endoxifen concentration, there is a trend of decreased enzyme activity with an increase in the number of *CYP2D6*\*36 copies. However, both plasma endoxifen concentration and endoxifen/NDM ratio were not significantly different between these groups. Thus, multiple copies of the *CYP2D6*(\*36)<sub>n</sub> CNV alleles do not appear to increase the *CYP2D6* activity. Further studies with larger sample numbers will be required to confirm the possible effect of the *CYP2D6*\*36 multiple copies on the adjacent *CYP2D6*\*10 allele.

In summary, this study is the first comprehensive study to analyze: (1) the *CYP2D6* CNVs in the Vietnamese and Filipino population and (2) the functional consequence of CNVs involving the *CYP2D6*\*36 allele. Like the other two Asian populations, Japanese and Chinese (Hosono et al., 2009), the Vietnamese and Filipino populations also have frequent *CYP2D6* gene CNVs. This specifically involves the *CYP2D6*(\*36)<sub>n</sub>-\*10 allele. These *CYP2D6* gene CNVs involving the *CYP2D6*(\*36)<sub>n</sub> allele neither increases nor decreases the *CYP2D6* enzyme activity. That is, the enzyme activities of *CYP2D6*\*36-\*10 and *CYP2D6*\*36-\*36-\*10 alleles are comparable to that of *CYP2D6*\*10. Based on tamoxifen metabolism, these *CYP2D6*(\*36)<sub>n</sub> CNV alleles are still null alleles and consequently, can be considered as poor metabolizers.

These results suggest that multiple copies of a gene do not always translate to additive effects. Even though CNVs are frequent in the human genome, the phenotypic consequence may depend on the sequence of the DNA that is duplicated. In case of the

*CYP2D6*\*36 allele, the gene polymorphisms, rather than the gene copy number variations, appear to be associated with *CYP2D6* enzyme activity. Thus, in the Asian populations where CNVs involving the *CYP2D6*\*36 allele is frequent, in order to estimate the phenotypic consequence, it is not enough to estimate just the total number of copies of the *CYP2D6* gene; it is necessary to identify the *CYP2D6* alleles involved in gene multiplication.

There was considerable interindividual variability even after correcting for *CYP2D6* copy number variations (Figure 3.2). Some of the variability may be explained by genes that were not analyzed (eg. other genes involved in drug metabolism). Plasma endoxifen concentration is dependent not only on its formation by *CYP2D6* enzyme but also on its clearance by phase II enzymes including UDP-glucuronosyltransferase 2B7 (UGT2B7; Blevins-Primeau et al., 2009). Hence, it is conceivable that UGT2B7 alone or in combination with *CYP2D6* contributes to the interindividual variability in tamoxifen metabolism. A part of this interindividual variability may also be due other additional factors like epigenetic mechanisms; however, little is known about which factors are important. Hence, the next specific aim (Chapter Four) was designed to assess the role of microRNAs (miRNA) in the direct and indirect regulation of *CYP2D6* expression.

### **Direct and indirect regulation of *CYP2D6* by microRNAs**

In this study, the *specific aim* was to identify microRNAs (miRNAs) that are predicted to target *CYP2D6*. The working *hypothesis* is that the variability in *CYP2D6* enzyme activity within a genotype group may be due to direct miRNA regulation of

CYP2D6 or indirect miRNA regulation of hepatic nuclear factor 4 $\alpha$  (HNF4A), a transcriptional factor that regulates CYP2D6.

MicroRNAs are small (18-25 nucleotides), noncoding RNAs that post-transcriptionally regulate gene expression. These miRNAs are predicted to regulate the expression of 20-90% of human transcripts (Lewis et al., 2005; Xie et al., 2005; Miranda et al., 2006). In animals, miRNAs typically bind to the 3'-untranslated region (3'-UTR) of the mRNAs and negatively regulate gene expression by blocking protein translation or by degrading the mRNA (Olsen and Ambros, 1999; Ambros et al., 2003). As more miRNAs are identified and studied, newer target sites and functions are being recognized, including coding region, promoter and 5'-UTR.

MicroRNAs are involved in the regulation of a number of genes involved in drug disposition - the ATP-binding cassette xenobiotic transporter (*ABCG2*; To et al., 2008), Pregnane X Receptor (*PXR*; Takagi et al., 2008), Cytochrome P450 1B1 (*CYP1B1*; Tsuchiya et al., 2006) and Cytochrome P450 3A4 (*CYP3A4*; Pan et al., 2009). Others have also speculated that interindividual variability in CYP450 expression and drug response may be due to the action of miRNAs (Ingelman-Sundberg et al., 2007).

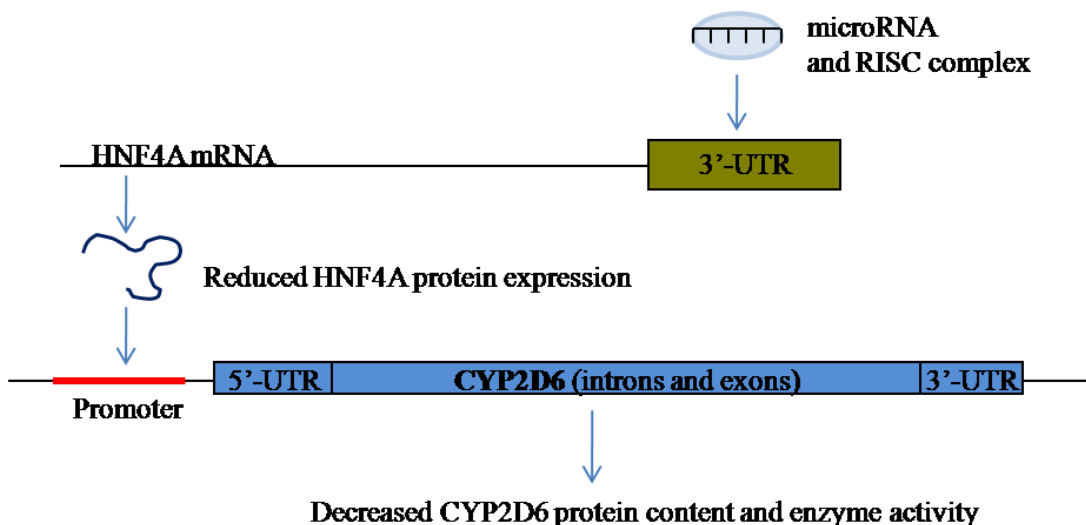
In this study, bioinformatic analyses were first performed to predict the miRNAs that target CYP2D6. The results suggested that CYP2D6 is likely to be regulated by miRNAs (Table 4.1). However, the results of the *in vitro* luciferase functional validation studies suggested that none of the five miRNAs tested (hsa-miR-493\*, hsa-miR-137, hsa-miR-140-3p, hsa-miR-149 and hsa-miR-323-3p) interacted with CYP2D6 3'-UTR. Bioinformatic analyses of the CYP2D6 5'-UTR and coding region using RNA22 algorithm predicted that a number of miRNAs target these regions. It is conceivable that

the miRNAs regulate CYP2D6 expression by targeting these regions. Further *in vitro* experiments are needed to validate these predictions.

MicroRNAs can regulate transcription factors (e.g., PXR, VDR) and thus, indirectly regulate the expression of CYP450 enzymes (Takagi et al., 2008; Pan et al., 2009). Hence, in the next step of the study, the miRNA regulation of an important transcriptional regulator of CYP2D6 -- HNF4A was investigated (Figure 5.4). HNF4A is transcriptional 'master regulator' that regulates the expression of a number of drug metabolizing genes including phase I enzymes, phase II enzymes, transporters and transcriptional factors that regulate the CYP450 genes (Kamiyama et al., 2007; Figure 1.4). Thus, regulation of HNF4A by miRNAs would likely affect many genes that are involved in drug disposition.

Bioinformatic analyses were performed to predict miRNAs that target HNF4A. The results indicated that many miRNAs may be involved in the regulation of HNF4A expression (Table 4.1). Based on these predictions, *in vitro* studies were performed in HeLa and HepG2 cell lines. Five of the miRNAs (hsa-miR-34a, hsa-miR-34b\*, hsa-miR-34c-5p, hsa-miR-449a and hsa-miR-765) regulated the expression of pIS-HNF4A luciferase plasmids. Based on the expression of these miRNAs in primary human hepatocytes, two of these miRNAs (hsa-miR-34a and hsa-miR-449a) were selected for further *in vitro* studies. Both these miRNAs down-regulated HNF4A protein expression. HNF4A mRNA expression was not affected, suggesting that they are most likely regulating the HNF4A expression by blocking HNF4A protein synthesis rather than degrading HNF4A mRNA.

**Figure 5.4: Schematic representation of possible indirect regulation of CYP2D6 by microRNAs targeting HNF4A.**



Note: microRNAs can regulate the expression of HNF4A. This can in turn lead to indirect regulation of the expression of CYP2D6 mRNA, protein and enzyme activity.

Since CYP2D6 is strongly regulated by HNF4A, the miRNAs that target HNF4A likely also affect its expression. There was no observable effect of those miRNAs on CYP2D6, however, this was likely because CYP2D6 is expressed at basal levels in the HepG2 cell line and consequently, these cells are not a good model for studying the down-regulation of CYP2D6 expression. In contrast, the expression level of PXR, another downstream target of HNF4A which is expressed highly in the HepG2 cells, was suppressed by the HNF4A miRNA.

MicroRNA functions can be altered by miRSNPs that affect the miRNA binding to the target mRNA (Mishra et al., 2007). These miRSNPs have been shown to alter the enzyme activity of dihydrofolate reductase (Mishra et al., 2007) and estrogen receptor  $\alpha$  (Adams et al., 2007). Hence, if miRSNPs are identified for HNF4A, these would serve as novel biomarkers of CYP2D6 activity. These biomarkers would help predict adverse

drug reactions and improve drug efficacy. The bioinformatic analyses predicted that miRSNPs do exist in HNF4A 3'-UTR. Furthermore, the luciferase activity was significantly higher ( $p < .05$ ) in the variant (rs11574744) plasmid when compared to the wildtype HNF4A 3'-UTR (Figure 4.12). Similarly, the luciferase activity of the variant plasmid was significantly higher ( $p < .05$ ) when either the hsa-miR-34a or the hsa-miR-449a were also cotransfected, suggesting that the miRSNP was interfering with the normal mRNA-miRNA activity. Further functional studies will be required to determine the phenotypic consequences of the miRSNPs. Genetic variants in the 3'-UTRs of drug metabolizing genes have not typically been given high priority in functional studies; however, based on these analyses, they may contribute to interindividual variability in the expression of the drug metabolizing enzymes.

Recent studies have shown that other proteins involved in drug metabolism (Tsuchiya et al., 2006; Takagi et al., 2008; To et al., 2008; Pan et al., 2009) are subject to miRNA regulation. The results of this study taken together with previous findings, suggest a complex regulatory mechanism for drug metabolizing enzymes by miRNAs.

In conclusion, these findings provide the first direct evidence that HNF4A and consequently, its downstream targets are regulated by miRNAs. This miRNA regulation of HNF4A can be important for the indirect regulation of CYP2D6. Thus, some of the interindividual variability seen in CYP2D6 mRNA expression and enzyme activity may be due to direct or indirect regulation by miRNAs. Since HNF4A also regulates the expression of a number of genes involved in drug metabolism (including phase II enzymes, drug transporters, transcription factors, etc.) these results have a broader implication. Thus, these miRNAs and miRSNPs may be useful as additional biomarkers



to refine the predicted metabolic phenotype an individual. These novel biomarkers would help improve the efficacy and reduce the side effects of many commonly prescribed drugs.

### **Clinical significance**

In breast cancer therapeutics, tamoxifen remains an important endocrine agent. Especially in many developing countries, including Vietnam and the Philippines, tamoxifen is preferred over aromatase inhibitors (AI) because of its low cost. Hence, it becomes important to understand the factors that lead to interindividual variability in tamoxifen response. A better understanding of this variability would help to improve the effectiveness and compliance of this life saving drug.

In both Vietnamese and Filipino populations, the most frequent (>50%) *CYP2D6* allele is the reduced functional allele, *CYP2D6\*10*. The polymorphisms that lead to reduced enzyme activity interfere with the metabolism of tamoxifen into endoxifen. Thus, in Vietnamese and Filipino women, *CYP2D6* genotyping may be useful in predicting the efficacy of tamoxifen therapy: that is, homozygous carriers of the *CYP2D6\*10* allele may derive reduced benefit from tamoxifen therapy. Prospective genotyping for *CYP2D6* variants may help identify a group of women who may experience greater benefit from tamoxifen therapy and a group that should consider alternatives to tamoxifen therapy. Furthermore, as *CYP2D6* inhibiting drugs create a phenocopy of the *CYP2D6* poor metabolizer phenotypes, they also would likely reduce endoxifen concentration and increase breast cancer recurrence in these women. Although the final conclusions will depend on the prospective outcome studies, based on the

current results, it would be prudent for these women avoid any CYP2D6 inhibitors. Clinical recommendations on *CYP2D6* genotype guided tamoxifen therapy await results from prospective clinical trials that show a direct correlation between endoxifen concentration and clinical outcomes; direct association between endoxifen concentration and clinical outcome has not been demonstrated yet.

These results presented here indicate the importance of conducting a prospective clinical trial to understand the association between endoxifen concentration and clinical outcome. The ongoing clinical trial in Vietnam and the Philippines will be able to bridge the gap. The prospective collection of both endoxifen concentrations and the recurrence outcomes in this study will for the first time, provide a direct assessment of the association of endoxifen with breast cancer recurrence outcomes. The results from this study will likely help guide the personalization of tamoxifen therapy by adjusting doses or choosing other endocrine therapies

Like the *CYP2D6* gene variants, *CYP2D6* gene CNVs can also affect CYP2D6 enzyme activity and consequently, alter drug metabolism and response. The phenotypic consequence of CNVs will depend on the *CYP2D6* allele involved in duplication. In the Vietnamese and Filipino populations, *CYP2D6* gene multiplications are frequent (>50%). Based on tamoxifen metabolism, the CNV alleles [*CYP2D6*(\*36)<sub>n</sub>] are poor metabolizer (null) alleles, suggesting that increased copy numbers of a gene does not always result in a corresponding increase in the phenotype. Thus, in the Vietnamese and Filipino populations, it is necessary to identify the *CYP2D6* alleles involved in the CNV; simply estimating the total number of copies of the *CYP2D6* gene without identifying the *CYP2D6* alleles cannot reliably be used to estimate the phenotype of an individual.

There is interindividual variability in endoxifen concentration even after correcting for both *CYP2D6* polymorphisms and CNVs. Other factors like microRNAs may explain a part of this variability. The results of this study suggest that miRNAs are likely to play an important role in the indirect regulation of *CYP2D6* through the regulation of its transcriptional factor, HNF4A. These endogenous hepatic miRNAs that regulate *CYP2D6* directly or indirectly will help to understand the interindividual variability in therapeutic efficacy and toxicity. Furthermore, polymorphisms that alter the mRNA-miRNA interactions can be used as biomarkers for guiding the clinical use of *CYP2D6* substrate drugs. Ultimately, these novel biomarkers will help improve the efficacy and reduce the side effects of commonly prescribed drugs.

### **Future directions**

Based on the results presented in this thesis, in the future, it would be valuable to further evaluate the following:

(1) *To determine the association between CYP2D6\*10 and clinical outcome in the Vietnamese and Filipino populations.* The *CYP2D6\*10* allele is a very common allele in the Asians (40-60% frequency) and thus perhaps the most common *CYP2D6* variant allele in the world. Hence, demonstrating a clinically relevant relationship between *CYP2D6\*10* allele and clinical outcome will help to explore alternative therapies in the Vietnamese and Filipino patients with poorer outcomes. These may include tamoxifen dose changes, exploration of use of alternative SERMs, and the use of AIs.

(2) *To determine the association between endoxifen concentrations and clinical outcome.* A direct association between endoxifen concentration and clinical outcome has not been

studied yet. In the future, this ongoing clinical trial in Vietnam and the Philippines will be able to bridge the gap by studying the association between these two factors.

Demonstrating a clinically relevant relationship between endoxifen concentration and clinical outcome will help to personalize the tamoxifen therapy in the Vietnamese and Filipino patients with poorer outcomes.

(3) *To evaluate the CYP2D6\*36 structure.* The *CYP2D6\*36* null allele is present in CNVs at a frequency of 30-50% in Japanese, Chinese, Filipino and Vietnamese populations, suggesting that this may be due to a ‘founder effect’. The *CYP2D6\*36* allele should be sequenced to identify the specific break points that create the fusion allele in these populations. This will be useful to understand the origin of these complex *CYP2D6* CNVs and the mechanism of gene duplications.

(4) *To determine the association between HNF4A miRNA regulation and CYP2D6 expression.* The results of the current study indicate that miRNAs regulate *HNF4A* *in vitro*. Further *in vitro* and *in vivo* studies (e.g., nuclear run-on assays, enzyme activity measurement) should be performed to understand the functional significance of these miRNAs on the expression and activity of *CYP2D6*. Also, the effect of the *HNF4A* 3'-UTR miRSNP on the *CYP2D6* enzyme activity needs to be investigated. Identification of miRNAs and miRSNPs that regulate *CYP2D6* will help to explain some of the interindividual variability that is not explained by drug-drug interactions and gene polymorphisms.

(5) *To determine if the miRNAs regulate CYP2D6 expression by targeting the 5'-UTR or coding region.* The results of the bioinformatic analyses using RNA22 algorithm indicate that both 5'-UTR and coding regions are a target of miRNAs. Further laboratory

experiments (e.g., luciferase assays, mRNA and protein expression, enzyme activity measurements) are needed to validate these predictions. Identification of such endogenous hepatic miRNAs that target regions other than the 3'-UTR will explain the differential expression of CYP2D6 mRNA transcripts in the human liver.

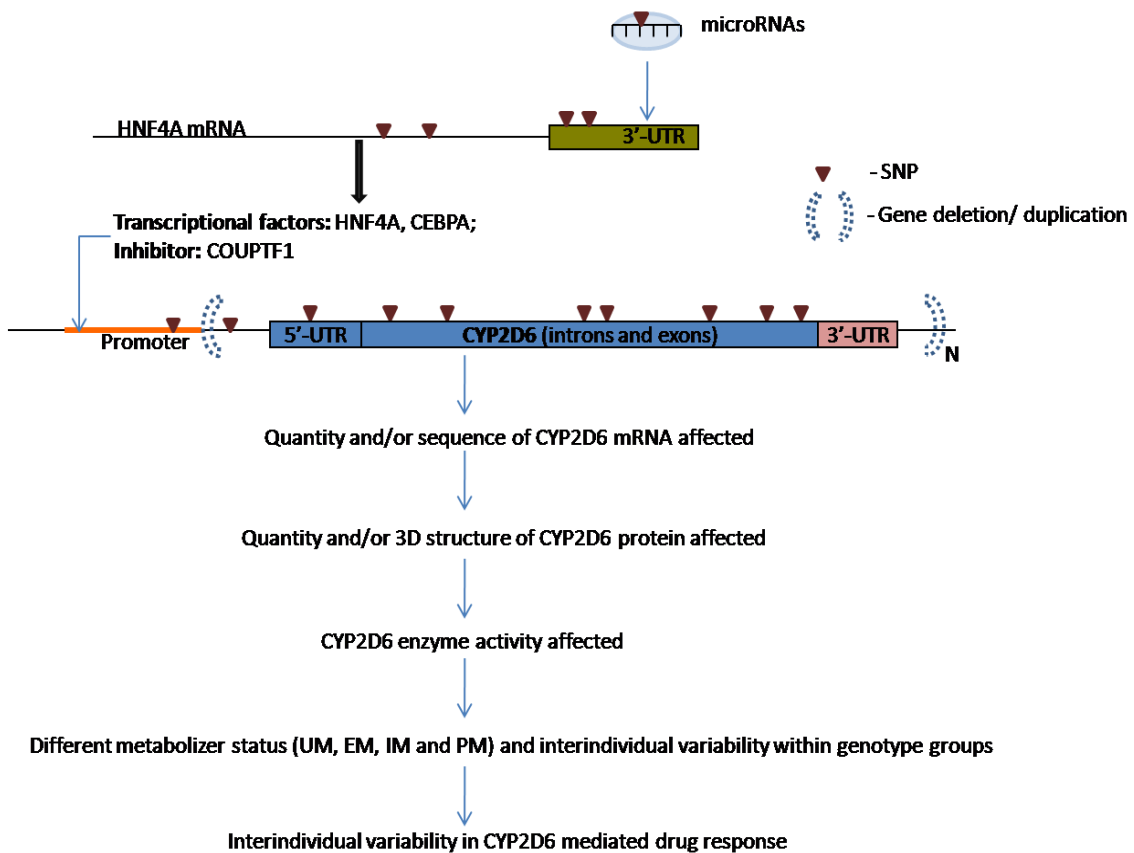
(6) *To identify the microRNAs that are expressed differentially between fetal and adult livers.* A number of CYP450 proteins, including CYP2D6 are only expressed postnatally. Hence, it is possible that post-transcriptional mechanisms like miRNAs may be developmentally regulating the expression of these CYP450s. Hence, miRNAs that are differentially expressed between the fetal and adult livers need to be identified by using methods like quantitative PCR, deep resequencing, and microarrays. Once the differentially expressed miRNAs are identified, the effect of these miRNAs on CYP450 expression can be evaluated by both *in vitro* and *in vivo* experiments (e.g., luciferase assays, mRNA expression, western blot, and enzyme activity measurements). These experiments will help to understand the developmental changes seen in CYP2D6 expression and activity. These results will also likely reveal miRNAs that are important in regulating CYP2D6 activity in adults.

## **Summary**

Interindividual variability in drug metabolism remains a significant contributor to differences in drug efficacy and toxicity. Some of this variability is due to known genetic variations and environmental factors that inhibit the enzymes or alter their expression levels. The studies presented here emphasize on the role of *CYP2D6* gene polymorphisms and copy number variations in the regulation of CYP2D6 enzyme activity and consequently, metabolism of tamoxifen in breast cancer patients.

Furthermore, in addition to the genetic factors, these results also show that miRNAs are involved in regulating HNF4A, a transcriptional regulator of CYP2D6. Ultimately, all these three factors can contribute to the interindividual variability in CYP2D6 mediated drug metabolism and response (Figure 5.5).

**Figure 5.5: Genetic and epigenetic factors regulate the expression and activity of CYP2D6.**



Note: *Cis*-acting genetic (polymorphisms and copy number variations) and *trans*-acting epigenetic factors (microRNA) regulate the expression of CYP2D6 mRNA, protein and consequently interindividual variability in CYP2D6 enzyme activity, drug metabolism and response.

## REFERENCES

- Adams BD, Furneaux H and White BA (2007) The micro-ribonucleic acid (miRNA) miR-206 targets the human estrogen receptor-alpha (ERalpha) and represses ERalpha messenger RNA and protein expression in breast cancer cell lines. *Mol Endocrinol* 21:1132-1147.
- Ahmad A, Ali SM, Ahmad MU, Sheikh S and Ahmad I (2010) Orally administered Endoxifen is a new therapeutic agent for breast cancer. *Breast Cancer Res Treat.*
- Aklillu E, Herrlin K, Gustafsson LL, Bertilsson L and Ingelman-Sundberg M (2002) Evidence for environmental influence on CYP2D6-catalysed debrisoquine hydroxylation as demonstrated by phenotyping and genotyping of Ethiopians living in Ethiopia or in Sweden. *Pharmacogenetics* 12:375-383.
- Aklillu E, Persson I, Bertilsson L, Johansson I, Rodrigues F and Ingelman-Sundberg M (1996) Frequent distribution of ultrarapid metabolizers of debrisoquine in an ethiopian population carrying duplicated and multiduplicated functional CYP2D6 alleles. *J Pharmacol Exp Ther* 278:441-446.
- Al-Dabbagh SG, Idle JR and Smith RL (1981) Animal modelling of human polymorphic drug oxidation--the metabolism of debrisoquine and phenacetin in rat inbred strains. *J Pharm Pharmacol* 33:161-164.
- Ambros V, Bartel B, Bartel DP, Burge CB, Carrington JC, Chen X, Dreyfuss G, Eddy SR, Griffiths-Jones S, Marshall M, Matzke M, Ruvkun G and Tuschl T (2003) A uniform system for microRNA annotation. *RNA* 9:277-279.
- Bao L, Zhou M, Wu L, Lu L, Goldowitz D, Williams RW and Cui Y (2007) PolymiRTS Database: linking polymorphisms in microRNA target sites with complex traits. *Nucleic Acids Res* 35:D51-54.
- Barron TI, Connolly R, Bennett K, Feely J and Kennedy MJ (2007) Early discontinuation of tamoxifen: a lesson for oncologists. *Cancer* 109:832-839.
- Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116:281-297.
- Blevins-Primeau AS, Sun D, Chen G, Sharma AK, Gallagher CJ, Amin S and Lazarus P (2009) Functional significance of UDP-glucuronosyltransferase variants in the metabolism of active tamoxifen metabolites. *Cancer Res* 69:1892-1900.
- Bogaards JJ, Bertrand M, Jackson P, Oudshoorn MJ, Weaver RJ, van Bladeren PJ and Walther B (2000) Determining the best animal model for human cytochrome P450 activities: a comparison of mouse, rat, rabbit, dog, micropig, monkey and man. *Xenobiotica* 30:1131-1152.

Bolt HM and Thier R (2006) Relevance of the deletion polymorphisms of the glutathione S-transferases GSTT1 and GSTM1 in pharmacology and toxicology. *Curr Drug Metab* 7:613-628.

Bonanni B, Lazzeroni M and Veronesi U (2007) Does tamoxifen prophylaxis for breast cancer provide long-term prevention? *Nat Clin Pract Oncol* 4:568-569.

Bonanni B, Macis D, Maisonneuve P, Johansson HA, Gucciardo G, Oliviero P, Travaglini R, Muraca MG, Rotmensz N, Veronesi U and Decensi AU (2006) Polymorphism in the CYP2D6 tamoxifen-metabolizing gene influences clinical effect but not hot flashes: data from the Italian Tamoxifen Trial. *J Clin Oncol* 24:3708-3709; author reply 3709.

Borges S, Desta Z, Jin Y, Faouzi A, Robarge JD, Philip S, Nguyen A, Stearns V, Hayes D, Rae JM, Skaar TC, Flockhart DA and Li L (2010) Composite Functional Genetic and Comedication CYP2D6 Activity Score in Predicting Tamoxifen Drug Exposure Among Breast Cancer Patients. *J Clin Pharmacol*.

Borges S, Desta Z, Li L, Skaar TC, Ward BA, Nguyen A, Jin Y, Storniolo AM, Nikoloff DM, Wu L, Hillman G, Hayes DF, Stearns V and Flockhart DA (2006) Quantitative effect of CYP2D6 genotype and inhibitors on tamoxifen metabolism: implication for optimization of breast cancer treatment. *Clin Pharmacol Ther* 80:61-74.

Bradford LD (2002) CYP2D6 allele frequency in European Caucasians, Asians, Africans and their descendants. *Pharmacogenomics* 3:229-243.

Cai WM, Chen B and Zhang WX (2007) Frequency of CYP2D6\*10 and \*14 alleles and their influence on the metabolic activity of CYP2D6 in a healthy Chinese population. *Clin Pharmacol Ther* 81:95-98.

Cairns W, Smith CA, McLaren AW and Wolf CR (1996) Characterization of the human cytochrome P4502D6 promoter. A potential role for antagonistic interactions between members of the nuclear receptor family. *J Biol Chem* 271:25269-25276.

Calin GA and Croce CM (2006) MicroRNA-cancer connection: the beginning of a new tale. *Cancer Res* 66:7390-7394.

Carcillo JA, Parise RA, Adedoyin A, Frye R, Branch RA and Romkes M (1996) CYP2D6 mRNA expression in circulating peripheral blood mononuclear cells correlates with in vivo debrisoquine hydroxylase activity in extensive metabolizers. *Res Commun Mol Pathol Pharmacol* 91:149-159.



- Chang J, Powles TJ, Ashley SE, Gregory RK, Tidy VA, Treleaven JG and Singh R (1996) The effect of tamoxifen and hormone replacement therapy on serum cholesterol, bone mineral density and coagulation factors in healthy postmenopausal women participating in a randomised, controlled tamoxifen prevention study. *Ann Oncol* 7:671-675.
- Chen Y and Stallings RL (2007) Differential patterns of microRNA expression in neuroblastoma are correlated with prognosis, differentiation, and apoptosis. *Cancer Res* 67:976-983.
- Chida M, Ariyoshi N, Yokoi T, Nemoto N, Inaba M, Kinoshita M and Kamataki T (2002) New allelic arrangement CYP2D6\*36 x 2 found in a Japanese poor metabolizer of debrisoquine. *Pharmacogenetics* 12:659-662.
- Chou WH, Yan FX, de Leon J, Barnhill J, Rogers T, Cronin M, Pho M, Xiao V, Ryder TB, Liu WW, Teiling C and Wedlund PJ (2000) Extension of a pilot study: impact from the cytochrome P450 2D6 polymorphism on outcome and costs associated with severe mental illness. *J Clin Psychopharmacol* 20:246-251.
- Clarke M (2006) Meta-analyses of adjuvant therapies for women with early breast cancer: the Early Breast Cancer Trialists' Collaborative Group overview. *Ann Oncol* 17 Suppl 10:x59-62.
- Corchero J, Granvil CP, Akiyama TE, Hayhurst GP, Pimprale S, Feigenbaum L, Idle JR and Gonzalez FJ (2001) The CYP2D6 humanized mouse: effect of the human CYP2D6 transgene and HNF4alpha on the disposition of debrisoquine in the mouse. *Mol Pharmacol* 60:1260-1267.
- Craddock N, Hurles ME, Cardin N, Pearson RD, Plagnol V, Robson S, Vukcevic D, Barnes C, Conrad DF, Giannoulatou E, Holmes C, Marchini JL, Stirrups K, Tobin MD, Wain LV, Yau C, Aerts J, Ahmad T, Andrews TD, Arbury H, Attwood A, Auton A, Ball SG, Balmforth AJ, Barrett JC, Barroso I, Barton A, Bennett AJ, Bhaskar S, Blaszczyk K, Bowes J, Brand OJ, Braund PS, Bredin F, Breen G, Brown MJ, Bruce IN, Bull J, Burren OS, Burton J, Byrnes J, Caesar S, Clee CM, Coffey AJ, Connell JM, Cooper JD, Dominiczak AF, Downes K, Drummond HE, Dudakia D, Dunham A, Ebbs B, Eccles D, Edkins S, Edwards C, Elliot A, Emery P, Evans DM, Evans G, Eyre S, Farmer A, Ferrier IN, Feuk L, Fitzgerald T, Flynn E, Forbes A, Forty L, Franklyn JA, Freathy RM, Gibbs P, Gilbert P, Gokumen O, Gordon-Smith K, Gray E, Green E, Groves CJ, Grozeva D, Gwilliam R, Hall A, Hammond N, Hardy M, Harrison P, Hassanali N, Hebaishi H, Hines S, Hinks A, Hitman GA, Hocking L, Howard E, Howard P, Howson JM, Hughes D, Hunt S, Isaacs JD, Jain M, Jewell DP, Johnson T, Jolley JD, Jones IR, Jones LA, et al. (2010) Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. *Nature* 464:713-720.

- Dahl ML, Johansson I, Palmertz MP, Ingelman-Sundberg M and Sjoqvist F (1992) Analysis of the CYP2D6 gene in relation to debrisoquin and desipramine hydroxylation in a Swedish population. *Clin Pharmacol Ther* 51:12-17.
- Desta Z, Ward BA, Soukhova NV and Flockhart DA (2004) Comprehensive evaluation of tamoxifen sequential biotransformation by the human cytochrome P450 system in vitro: prominent roles for CYP3A and CYP2D6. *J Pharmacol Exp Ther* 310:1062-1075.
- Du T and Zamore PD (2005) microPrimer: the biogenesis and function of microRNA. *Development* 132:4645-4652.
- Dunning AM, Dowsett M, Healey CS, Tee L, Luben RN, Folkard E, Novik KL, Kelemen L, Ogata S, Pharoah PD, Easton DF, Day NE and Ponder BA (2004) Polymorphisms associated with circulating sex hormone levels in postmenopausal women. *J Natl Cancer Inst* 96:936-945.
- Duursma AM, Kedde M, Schrier M, le Sage C and Agami R (2008) miR-148 targets human DNMT3b protein coding region. *RNA* 14:872-877.
- Eichelbaum M, Ingelman-Sundberg M and Evans WE (2006) Pharmacogenomics and individualized drug therapy. *Annu Rev Med* 57:119-137.
- Eichelbaum M, Mineshita S, Ohnhaus EE and Zekorn C (1986) The influence of enzyme induction on polymorphic sparteine oxidation. *Br J Clin Pharmacol* 22:49-53.
- Esquela-Kerscher A and Slack FJ (2006) Oncomirs - microRNAs with a role in cancer. *Nat Rev Cancer* 6:259-269.
- Fukuda T, Yamamoto I, Nishida Y, Zhou Q, Ohno M, Takada K and Azuma J (1999) Effect of the CYP2D6\*10 genotype on venlafaxine pharmacokinetics in healthy adult volunteers. *Br J Clin Pharmacol* 47:450-453.
- Fukushima T, Hamada Y, Yamada H and Horii I (2007) Changes of micro-RNA expression in rat liver treated by acetaminophen or carbon tetrachloride--regulating role of micro-RNA for RNA expression. *J Toxicol Sci* 32:401-409.
- Gaedigk A, Bradford LD, Alander SW and Leeder JS (2006) CYP2D6\*36 gene arrangements within the cyp2d6 locus: association of CYP2D6\*36 with poor metabolizer status. *Drug Metab Dispos* 34:563-569.
- Gaedigk A, Gaedigk R and Leeder JS (2005) CYP2D7 splice variants in human liver and brain: does CYP2D7 encode functional protein? *Biochem Biophys Res Commun* 336:1241-1250.

Gaedigk A, Ndjountche L, Divakaran K, Dianne Bradford L, Zineh I, Oberlander TF, Brousseau DC, McCarver DG, Johnson JA, Alander SW, Wayne Riggs K and Steven Leeder J (2007) Cytochrome P4502D6 (CYP2D6) gene locus heterogeneity: characterization of gene duplication events. *Clin Pharmacol Ther* 81:242-251.

Gaedigk A, Simon SD, Pearce RE, Bradford LD, Kennedy MJ and Leeder JS (2008) The CYP2D6 activity score: translating genotype information into a qualitative measure of phenotype. *Clin Pharmacol Ther* 83:234-242.

Garcia-Barcelo M, Chow LY, Chiu HF, Wing YK, Lee DT, Lam KL and Waye MM (2000) Genetic analysis of the CYP2D6 locus in a Hong Kong Chinese population. *Clin Chem* 46:18-23.

Georgantas RW, 3rd, Hildreth R, Morisot S, Alder J, Liu CG, Heimfeld S, Calin GA, Croce CM and Civin CI (2007) CD34+ hematopoietic stem-progenitor cell microRNA expression and function: a circuit diagram of differentiation control. *Proc Natl Acad Sci U S A* 104:2750-2755.

Georges M, Clop A, Marcq F, Takeda H, Pirottin D, Hiard S, Tordoir X, Caiment F, Meish F, Bibe B, Bouix J, Elsen JM, Eychenne F, Laville E, Larzul C, Milenkovic D, Tobin J and Charlier AC (2006) Polymorphic microRNA-target interactions: a novel source of phenotypic variation. *Cold Spring Harb Symp Quant Biol* 71:343-350.

Gjerde J, Hauglid M, Breilid H, Lundgren S, Varhaug JE, Kisanga ER, Mellgren G, Steen VM and Lien EA (2008) Effects of CYP2D6 and SULT1A1 genotypes including SULT1A1 gene copy number on tamoxifen metabolism. *Ann Oncol* 19:56-61.

Goetz MP, Knox SK, Suman VJ, Rae JM, Safgren SL, Ames MM, Visscher DW, Reynolds C, Couch FJ, Lingle WL, Weinshilboum RM, Fritcher EG, Nibbe AM, Desta Z, Nguyen A, Flockhart DA, Perez EA and Ingle JN (2007) The impact of cytochrome P450 2D6 metabolism in women receiving adjuvant tamoxifen. *Breast Cancer Res Treat* 101:113-121.

Goetz MP, Rae JM, Suman VJ, Safgren SL, Ames MM, Visscher DW, Reynolds C, Couch FJ, Lingle WL, Flockhart DA, Desta Z, Perez EA and Ingle JN (2005) Pharmacogenetics of tamoxifen biotransformation is associated with clinical outcomes of efficacy and hot flashes. *J Clin Oncol* 23:9312-9318.

Gonzalez-Santiago S, Zárata R, Haba-Rodríguez J and al. e (2007) CYP2D6\*4 polymorphism as blood predictive biomarker of breast cancer relapse in patients receiving adjuvant tamoxifen. [abstract]. *J Clin Oncol; ASCO annual meeting proceedings* Part I. No. 18S (June 20 Supplement):590.

Gregory RI, Yan KP, Amuthan G, Chendrimada T, Doratotaj B, Cooch N and Shiekhattar R (2004) The Microprocessor complex mediates the genesis of microRNAs. *Nature* 432:235-240.

- Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A and Enright AJ (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res* 34:D140-144.
- Gruber AR, Lorenz R, Bernhart SH, Neubock R and Hofacker IL (2008) The Vienna RNA websuite. *Nucleic Acids Res* 36:W70-74.
- Guidice JM, Marez D, Sabbagh N, Legrand-Andreoletti M, Spire C, Alcaide E, Lafitte JJ and Broly F (1997) Evidence for CYP2D6 expression in human lung. *Biochem Biophys Res Commun* 241:79-85.
- Hakkola J, Pasanen M, Purkunen R, Saarikoski S, Pelkonen O, Maenpaa J, Rane A and Raunio H (1994) Expression of xenobiotic-metabolizing cytochrome P450 forms in human adult and fetal liver. *Biochem Pharmacol* 48:59-64.
- Han J, Lee Y, Yeom KH, Kim YK, Jin H and Kim VN (2004) The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* 18:3016-3027.
- Hanioka N, Okumura Y, Saito Y, Hichiya H, Soyama A, Saito K, Ueno K, Sawada J and Narimatsu S (2006) Catalytic roles of CYP2D6.10 and CYP2D6.36 enzymes in mexiletine metabolism: in vitro functional analysis of recombinant proteins expressed in *Saccharomyces cerevisiae*. *Biochem Pharmacol* 71:1386-1395.
- Harries LW, Locke JM, Shields B, Hanley NA, Hanley KP, Steele A, Njolstad PR, Ellard S and Hattersley AT (2008) The diabetic phenotype in HNF4A mutation carriers is moderated by the expression of HNF4A isoforms from the P1 promoter during fetal development. *Diabetes* 57:1745-1752.
- Hayhurst GP, Lee YH, Lambert G, Ward JM and Gonzalez FJ (2001) Hepatocyte nuclear factor 4alpha (nuclear receptor 2A1) is essential for maintenance of hepatic gene expression and lipid homeostasis. *Mol Cell Biol* 21:1393-1403.
- Hebbring SJ, Adjei AA, Baer JL, Jenkins GD, Zhang J, Cunningham JM, Schaid DJ, Weinshilboum RM and Thibodeau SN (2007) Human SULT1A1 gene: copy number differences and functional implications. *Hum Mol Genet* 16:463-470.
- Hosono N, Kato M, Kiyotani K, Mushiroda T, Takata S, Sato H, Amitani H, Tsuchiya Y, Yamazaki K, Tsunoda T, Zembutsu H, Nakamura Y and Kubo M (2009) CYP2D6 genotyping for functional-gene dosage analysis by allele copy number detection. *Clin Chem* 55:1546-1554.
- Hosono N, Kubo M, Tsuchiya Y, Sato H, Kitamoto T, Saito S, Ohnishi Y and Nakamura Y (2008) Multiplex PCR-based real-time invader assay (mPCR-RETINA): a novel SNP-based method for detecting allelic asymmetries within copy number variation regions. *Hum Mutat* 29:182-189.

- Iafate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW and Lee C (2004) Detection of large-scale variation in the human genome. *Nat Genet* 36:949-951.
- Ihara A, Yamagata K, Nammo T, Miura A, Yuan M, Tanaka T, Sladek FM, Matsuzawa Y, Miyagawa J and Shimomura I (2005) Functional characterization of the HNF4alpha isoform (HNF4alpha8) expressed in pancreatic beta-cells. *Biochem Biophys Res Commun* 329:984-990.
- Ingelman-Sundberg M (2001) Genetic and environmental causes for interindividual variability in drug pharmacokinetics *International Congress Series* 1220: 175-186.
- Ingelman-Sundberg M (2004) Pharmacogenetics of cytochrome P450 and its applications in drug therapy: the past, present and future. *Trends Pharmacol Sci* 25:193-200.
- Ingelman-Sundberg M (2005) Genetic polymorphisms of cytochrome P450 2D6 (CYP2D6): clinical consequences, evolutionary aspects and functional diversity. *Pharmacogenomics J* 5:6-13.
- Ingelman-Sundberg M and Rodriguez-Antona C (2005) Pharmacogenetics of drug-metabolizing enzymes: implications for a safer and more effective drug therapy. *Philos Trans R Soc Lond B Biol Sci* 360:1563-1570.
- Ingelman-Sundberg M, Sim SC, Gomez A and Rodriguez-Antona C (2007) Influence of cytochrome P450 polymorphisms on drug therapies: pharmacogenetic, pharmacoepigenetic and clinical aspects. *Pharmacol Ther* 116:496-526.
- Ingle JN, Suman VJ, Johnson PA, Krook JE, Mailliard JA, Wheeler RH, Loprinzi CL, Perez EA, Jordan VC and Dowsett M (1999) Evaluation of tamoxifen plus letrozole with assessment of pharmacokinetic interaction in postmenopausal women with metastatic breast cancer. *Clin Cancer Res* 5:1642-1649.
- Iorio MV, Ferracin M, Liu CG, Veronese A, Spizzo R, Sabbioni S, Magri E, Pedriali M, Fabbri M, Campiglio M, Menard S, Palazzo JP, Rosenberg A, Musiani P, Volinia S, Nenci I, Calin GA, Querzoli P, Negrini M and Croce CM (2005) MicroRNA gene expression deregulation in human breast cancer. *Cancer Res* 65:7065-7070.
- Ishiguro A, Kubota T, Ishikawa H and Iga T (2004) Metabolic activity of dextromethorphan O-demethylation in healthy Japanese volunteers carrying duplicated CYP2D6 genes: duplicated allele of CYP2D6\*10 does not increase CYP2D6 metabolic activity. *Clin Chim Acta* 344:201-204.
- Iwazaki N, Kobayashi K, Morimoto K, Hirano M, Kawashima S, Furihata T and Chiba K (2008) Involvement of hepatocyte nuclear factor 4 alpha in transcriptional regulation of the human pregnane X receptor gene in the human liver. *Drug Metab Pharmacokinet* 23:59-66.

Jin Y, Desta Z, Stearns V, Ward B, Ho H, Lee KH, Skaar T, Storniolo AM, Li L, Araba A, Blanchard R, Nguyen A, Ullmer L, Hayden J, Lemler S, Weinshilboum RM, Rae JM, Hayes DF and Flockhart DA (2005) CYP2D6 genotype, antidepressant use, and tamoxifen metabolism during adjuvant breast cancer treatment. *J Natl Cancer Inst* 97:30-39.

Johansson I, Oscarson M, Yue QY, Bertilsson L, Sjoqvist F and Ingelman-Sundberg M (1994) Genetic analysis of the Chinese cytochrome P4502D locus: characterization of variant CYP2D6 genes present in subjects with diminished capacity for debrisoquine hydroxylation. *Mol Pharmacol* 46:452-459.

John B, Enright AJ, Aravin A, Tuschl T, Sander C and Marks DS (2004) Human MicroRNA targets. *PLoS Biol* 2:e363.

Johnson MD, Zuo H, Lee KH, Trebley JP, Rae JM, Weatherman RV, Desta Z, Flockhart DA and Skaar TC (2004) Pharmacological characterization of 4-hydroxy-N-desmethyl tamoxifen, a novel active metabolite of tamoxifen. *Breast Cancer Res Treat* 85:151-159.

Kamiyama Y, Matsubara T, Yoshinari K, Nagata K, Kamimura H and Yamazoe Y (2007) Role of human hepatocyte nuclear factor 4alpha in the expression of drug-metabolizing enzymes and transporters in human hepatocytes assessed by use of small interfering RNA. *Drug Metab Pharmacokinet* 22:287-298.

Kato M, Nakamura Y and Tsunoda T (2008) An algorithm for inferring complex haplotypes in a region of copy-number variation. *Am J Hum Genet* 83:157-169.

Kearns GL, Blumer J, Schexnayder S, James LP, Adcock KG, Reed MD, Daniel JF, Gaedigk A and Paul J (2008) Single-dose pharmacokinetics of oral and intravenous pantoprazole in children and adolescents. *J Clin Pharmacol* 48:1356-1365.

Kelly CM, Juurlink DN, Gomes T, Duong-Hua M, Pritchard KI, Austin PC and Paszat LF (2010) Selective serotonin reuptake inhibitors and breast cancer mortality in women receiving tamoxifen: a population based cohort study. *BMJ* 340:c693.

Kertesz M, Iovino N, Unnerstall U, Gaul U and Segal E (2007) The role of site accessibility in microRNA target recognition. *Nat Genet* 39:1278-1284.

Kimura S, Umeno M, Skoda RC, Meyer UA and Gonzalez FJ (1989) The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and identification of the polymorphic CYP2D6 gene, a related gene, and a pseudogene. *Am J Hum Genet* 45:889-904.

Kirchheiner J, Nickchen K, Bauer M, Wong ML, Licinio J, Roots I and Brockmoller J (2004) Pharmacogenetics of antidepressants and antipsychotics: the contribution of allelic variations to the phenotype of drug response. *Mol Psychiatry* 9:442-473.

Kiyotani K, Mushiroda T, Imamura CK, Hosono N, Tsunoda T, Kubo M, Tanigawara Y, Flockhart DA, Desta Z, Skaar TC, Aki F, Hirata K, Takatsuka Y, Okazaki M, Ohsumi S, Yamakawa T, Sasa M, Nakamura Y and Zembutsu H (2010) Significant Effect of Polymorphisms in CYP2D6 and ABCC2 on Clinical Outcomes of Adjuvant Tamoxifen Therapy for Breast Cancer Patients. *J Clin Oncol*.

Kiyotani K, Mushiroda T, Sasa M, Bando Y, Sumitomo I, Hosono N, Kubo M, Nakamura Y and Zembutsu H (2008) Impact of CYP2D6\*10 on recurrence-free survival in breast cancer patients receiving adjuvant tamoxifen therapy. *Cancer Sci* 99:995-999.

Krek A, Grun D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M and Rajewsky N (2005) Combinatorial microRNA target predictions. *Nat Genet* 37:495-500.

Kreuzer KA, Lass U, Bohn A, Landt O and Schmidt CA (1999) LightCycler technology for the quantitation of bcr/abl fusion transcripts. *Cancer Res* 59:3171-3174.

Kubota T, Yamaura Y, Ohkawa N, Hara H and Chiba K (2000) Frequencies of CYP2D6 mutant alleles in a normal Japanese population and metabolic activity of dextromethorphan O-demethylation in different CYP2D6 genotypes. *Br J Clin Pharmacol* 50:31-34.

Laika B, Leucht S, Heres S and Steimer W (2009) Intermediate metabolizer: increased side effects in psychoactive drug therapy. The key to cost-effectiveness of pretreatment CYP2D6 screening? *Pharmacogenomics J* 9:395-403.

Lam YW, Gaedigk A, Ereshefsky L, Alfaro CL and Simpson J (2002) CYP2D6 inhibition by selective serotonin reuptake inhibitors: analysis of achievable steady-state plasma concentrations and the effect of ultrarapid metabolism at CYP2D6. *Pharmacotherapy* 22:1001-1006.

Lee KH, Ward BA, Desta Z, Flockhart DA and Jones DR (2003) Quantification of tamoxifen and three metabolites in plasma by high-performance liquid chromatography with fluorescence detection: application to a clinical trial. *J Chromatogr B Analyt Technol Biomed Life Sci* 791:245-253.

Lee SJ, Lee SS, Jung HJ, Kim HS, Park SJ, Yeo CW and Shin JG (2009) Discovery of novel functional variants and extensive evaluation of CYP2D6 genetic polymorphisms in Koreans. *Drug Metab Dispos* 37:1464-1470.

Lee SS, Cha EY, Jung HJ, Shon JH, Kim EY, Yeo CW and Shin JG (2008) Genetic polymorphism of hepatocyte nuclear factor-4 $\alpha$  influences human cytochrome P450 2D6 activity. *Hepatology* 48:635-645.

- Lewis BP, Burge CB and Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 120:15-20.
- Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP and Burge CB (2003) Prediction of mammalian microRNA targets. *Cell* 115:787-798.
- Lim HS, Ju Lee H, Seok Lee K, Sook Lee E, Jang IJ and Ro J (2007) Clinical implications of CYP2D6 genotypes predictive of tamoxifen pharmacokinetics in metastatic breast cancer. *J Clin Oncol* 25:3837-3845.
- Lim YC, Desta Z, Flockhart DA and Skaar TC (2005) Endoxifen (4-hydroxy-N-desmethyl-tamoxifen) has anti-estrogenic effects in breast cancer cells with potency similar to 4-hydroxy-tamoxifen. *Cancer Chemother Pharmacol* 55:471-478.
- Lim YC, Li L, Desta Z, Zhao Q, Rae JM, Flockhart DA and Skaar TC (2006) Endoxifen, a secondary metabolite of tamoxifen, and 4-OH-tamoxifen induce similar changes in global gene expression patterns in MCF-7 breast cancer cells. *J Pharmacol Exp Ther* 318:503-512.
- Lonning PE, Lien EA, Lundgren S and Kvinnsland S (1992) Clinical pharmacokinetics of endocrine agents used in advanced breast cancer. *Clin Pharmacokinet* 22:327-358.
- Lund E, Guttinger S, Calado A, Dahlberg JE and Kutay U (2004) Nuclear export of microRNA precursors. *Science* 303:95-98.
- Masubuchi Y, Iwasa T, Hosokawa S, Suzuki T, Horie T, Imaoka S, Funae Y and Narimatsu S (1997) Selective deficiency of debrisoquine 4-hydroxylase activity in mouse liver microsomes. *J Pharmacol Exp Ther* 282:1435-1441.
- McCarroll SA, Hadnott TN, Perry GH, Sabeti PC, Zody MC, Barrett JC, Dallaire S, Gabriel SB, Lee C, Daly MJ and Altshuler DM (2006) Common deletion polymorphisms in the human genome. *Nat Genet* 38:86-92.
- McFayden MC, Melvin WT and Murray GI (1998) Regional distribution of individual forms of cytochrome P450 mRNA in normal adult human brain. *Biochem Pharmacol* 55:825-830.
- Meijerman I, Sanderson LM, Smits PH, Beijnen JH and Schellens JH (2007) Pharmacogenetic screening of the gene deletion and duplications of CYP2D6. *Drug Metab Rev* 39:45-60.
- Meng F, Henson R, Lang M, Wehbe H, Maheshwari S, Mendell JT, Jiang J, Schmittgen TD and Patel T (2006) Involvement of human micro-RNA in growth and response to chemotherapy in human cholangiocarcinoma cell lines. *Gastroenterology* 130:2113-2129.



- Miksys S, Rao Y, Hoffmann E, Mash DC and Tyndale RF (2002) Regional and cellular expression of CYP2D6 in human brain: higher levels in alcoholics. *J Neurochem* 82:1376-1387.
- Miller TE, Ghoshal K, Ramaswamy B, Roy S, Datta J, Shapiro CL, Jacob S and Majumder S (2008) MicroRNA-221/222 confers tamoxifen resistance in breast cancer by targeting p27Kip1. *J Biol Chem* 283:29897-29903.
- Miranda KC, Huynh T, Tay Y, Ang YS, Tam WL, Thomson AM, Lim B and Rigoutsos I (2006) A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell* 126:1203-1217.
- Mishra PJ, Humeniuk R, Longo-Sorbello GS, Banerjee D and Bertino JR (2007) A miR-24 microRNA binding-site polymorphism in dihydrofolate reductase gene leads to methotrexate resistance. *Proc Natl Acad Sci U S A* 104:13513-13518.
- Nelson DR, Zeldin DC, Hoffman SM, Maltais LJ, Wain HM and Nebert DW (2004) Comparison of cytochrome P450 (CYP) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes and alternative-splice variants. *Pharmacogenetics* 14:1-18.
- Nowell SA, Ahn J, Rae JM, Scheys JO, Trovato A, Sweeney C, MacLeod SL, Kadlubar FF and Ambrosone CB (2005) Association of genetic variation in tamoxifen-metabolizing enzymes with overall survival and recurrence of disease in breast cancer patients. *Breast Cancer Res Treat* 91:249-258.
- Okishiro M, Taguchi T, Jin Kim S, Shimazu K, Tamaki Y and Noguchi S (2009) Genetic polymorphisms of CYP2D6 10 and CYP2C19 2, 3 are not associated with prognosis, endometrial thickness, or bone mineral density in Japanese breast cancer patients treated with adjuvant tamoxifen. *Cancer* 115:952-961.
- Olsen PH and Ambros V (1999) The lin-4 regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation. *Dev Biol* 216:671-680.
- Orom UA, Nielsen FC and Lund AH (2008) MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Mol Cell* 30:460-471.
- Osborne CK (1998) Tamoxifen in the treatment of breast cancer. *N Engl J Med* 339:1609-1618.
- Pan YZ, Gao W and Yu AM (2009) MicroRNAs regulate CYP3A4 expression via direct and indirect targeting. *Drug Metab Dispos* 37:2112-2117.

Place RF, Li LC, Pookot D, Noonan EJ and Dahiya R (2008) MicroRNA-373 induces expression of genes with complementary promoter sequences. *Proc Natl Acad Sci U S A* 105:1608-1613.

Rae JM, Sikora MJ, Henry NL, Li L, Kim S, Oesterreich S, Skaar TC, Nguyen AT, Desta Z, Storniolo AM, Flockhart DA, Hayes DF and Stearns V (2009) Cytochrome P450 2D6 activity predicts discontinuation of tamoxifen therapy in breast cancer patients. *Pharmacogenomics J* 9:258-264.

Ramamoorthy A, Flockhart DA, Hosono N, Kubo M, Nakamura Y and Skaar TC (2010) Differential quantification of CYP2D6 gene copy number by four different quantitative real-time PCR assays. *Pharmacogenet Genomics*.

Rau T, Wohlleben G, Wuttke H, Thuerauf N, Lunkenheimer J, Lanczik M and Eschenhagen T (2004) CYP2D6 genotype: impact on adverse effects and nonresponse during treatment with antidepressants—a pilot study. *Clin Pharmacol Ther* 75:386-393.

Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, Gonzalez JR, Gratacos M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW and Hurles ME (2006) Global variation in copy number in the human genome. *Nature* 444:444-454.

Robarge JD, Li L, Desta Z, Nguyen A and Flockhart DA (2007) The star-allele nomenclature: retooling for translational genomics. *Clin Pharmacol Ther* 82:244-248.

Rodriguez-Antona C, Jover R, Gomez-Lechon MJ and Castell JV (2000) Quantitative RT-PCR measurement of human cytochrome P-450s: application to drug induction studies. *Arch Biochem Biophys* 376:109-116.

Romkes-Sparks M, Mnuskin A, Chern HD, Persad R, Fleming C, Sibley GN, Smith P, Wilkinson GR and Branch RA (1994) Correlation of polymorphic expression of CYP2D6 mRNA in bladder mucosa and tumor tissue to in vivo debrisoquine hydroxylase activity. *Carcinogenesis* 15:1955-1961.

Sachse C, Brockmoller J, Bauer S and Roots I (1997) Cytochrome P450 2D6 variants in a Caucasian population: allele frequencies and phenotypic consequences. *Am J Hum Genet* 60:284-295.

Sakuyama K, Sasaki T, Ujiie S, Obata K, Mizugaki M, Ishikawa M and Hiratsuka M (2008) Functional characterization of 17 CYP2D6 allelic variants (CYP2D6.2, 10, 14A-B, 18, 27, 36, 39, 47-51, 53-55, and 57). *Drug Metab Dispos* 36:2460-2467.

- Schaeffeler E, Schwab M, Eichelbaum M and Zanger UM (2003) CYP2D6 genotyping strategy based on gene copy number determination by TaqMan real-time PCR. *Hum Mutat* 22:476-485.
- Schroth W, Antoniadou L, Fritz P, Schwab M, Muerdter T, Zanger UM, Simon W, Eichelbaum M and Brauch H (2007) Breast cancer treatment outcome with adjuvant tamoxifen relative to patient CYP2D6 and CYP2C19 genotypes. *J Clin Oncol* 25:5187-5193.
- Schroth W, Goetz MP, Hamann U, Fasching PA, Schmidt M, Winter S, Fritz P, Simon W, Suman VJ, Ames MM, Safgren SL, Kuffel MJ, Ulmer HU, Bolander J, Strick R, Beckmann MW, Koelbl H, Weinshilboum RM, Ingle JN, Eichelbaum M, Schwab M and Brauch H (2009) Association between CYP2D6 polymorphisms and outcomes among women with early stage breast cancer treated with tamoxifen. *JAMA* 302:1429-1436.
- Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Maner S, Massa H, Walker M, Chi M, Navin N, Lucito R, Healy J, Hicks J, Ye K, Reiner A, Gilliam TC, Trask B, Patterson N, Zetterberg A and Wigler M (2004) Large-scale copy number polymorphism in the human genome. *Science* 305:525-528.
- Shimada T, Tsumura F, Yamazaki H, Guengerich FP and Inoue K (2001) Characterization of (+/-)-bufuralol hydroxylation activities in liver microsomes of Japanese and Caucasian subjects genotyped for CYP2D6. *Pharmacogenetics* 11:143-156.
- Shrestha S, Tang J and Kaslow RA (2009) Gene copy number: learning to count past two. *Nat Med* 15:1127-1129.
- Siegle I, Fritz P, Eckhardt K, Zanger UM and Eichelbaum M (2001) Cellular localization and regional distribution of CYP2D6 mRNA and protein expression in human brain. *Pharmacogenetics* 11:237-245.
- Silva JM, Li MZ, Chang K, Ge W, Golding MC, Rickles RJ, Siolas D, Hu G, Paddison PJ, Schlabach MR, Sheth N, Bradshaw J, Burchard J, Kulkarni A, Cavet G, Sachidanandam R, McCombie WR, Cleary MA, Elledge SJ and Hannon GJ (2005) Second-generation shRNA libraries covering the mouse and human genomes. *Nat Genet* 37:1281-1288.
- Sistonen J, Sajantila A, Lao O, Corander J, Barbujani G and Fuselli S (2007) CYP2D6 worldwide genetic variation shows high frequency of altered activity variants and no continental structure. *Pharmacogenet Genomics* 17:93-101.
- Snider NT, Sikora MJ, Sridar C, Feuerstein TJ, Rae JM and Hollenberg PF (2008) The endocannabinoid anandamide is a substrate for the human polymorphic cytochrome P450 2D6. *J Pharmacol Exp Ther* 327:538-545.

- Soyama A, Saito Y, Kubo T, Miyajima A, Ohno Y, Komamura K, Ueno K, Kamakura S, Kitakaze M, Tomoike H, Ozawa S and Sawada J (2006) Sequence-based analysis of the CYP2D6\*36-CYP2D6\*10 tandem-type arrangement, a major CYP2D6\*10 haplotype in the Japanese population. *Drug Metab Pharmacokinet* 21:208-216.
- Stearns V, Johnson MD, Rae JM, Morocho A, Novielli A, Bhargava P, Hayes DF, Desta Z and Flockhart DA (2003) Active tamoxifen metabolite plasma concentrations after coadministration of tamoxifen and the selective serotonin reuptake inhibitor paroxetine. *J Natl Cancer Inst* 95:1758-1764.
- Stoffel M and Duncan SA (1997) The maturity-onset diabetes of the young (MODY1) transcription factor HNF4alpha regulates expression of genes required for glucose transport and metabolism. *Proc Natl Acad Sci U S A* 94:13209-13214.
- Sun D, Sharma AK, Dellinger RW, Blevins-Primeau AS, Balliet RM, Chen G, Boyiri T, Amin S and Lazarus P (2007) Glucuronidation of active tamoxifen metabolites by the human UDP glucuronosyltransferases. *Drug Metab Dispos* 35:2006-2014.
- Takagi S, Nakajima M, Mohri T and Yokoi T (2008) Post-transcriptional regulation of human pregnane X receptor by micro-RNA affects the expression of cytochrome P450 3A4. *J Biol Chem* 283:9674-9680.
- To KK, Zhan Z, Litman T and Bates SE (2008) Regulation of ABCG2 expression at the 3' untranslated region of its mRNA through modulation of transcript stability and protein translation by a putative microRNA in the S1 colon cancer cell line. *Mol Cell Biol* 28:5147-5161.
- Treluyer JM, Jacqz-Aigrain E, Alvarez F and Cresteil T (1991) Expression of CYP2D6 in developing human liver. *Eur J Biochem* 202:583-588.
- Tsuchiya Y, Nakajima M, Takagi S, Taniya T and Yokoi T (2006) MicroRNA regulates the expression of human cytochrome P450 1B1. *Cancer Res* 66:9090-9098.
- Veiga MI, Asimus S, Ferreira PE, Martins JP, Cavaco I, Ribeiro V, Hai TN, Petzold MG, Bjorkman A, Ashton M and Gil JP (2009) Pharmacogenomics of CYP2A6, CYP2B6, CYP2C19, CYP2D6, CYP3A4, CYP3A5 and MDR1 in Vietnam. *Eur J Clin Pharmacol* 65:355-363.
- Wadelius M, Darj E, Frenne G and Rane A (1997) Induction of CYP2D6 in pregnancy. *Clin Pharmacol Ther* 62:400-407.
- Wain LV, Armour JA and Tobin MD (2009) Genomic copy number variation, human health, and disease. *Lancet* 374:340-350.

Wang J, Pitarque M and Ingelman-Sundberg M (2006) 3'-UTR polymorphism in the human CYP2A6 gene affects mRNA stability and enzyme expression. *Biochem Biophys Res Commun* 340:491-497.

Wegman P, Elingarami S, Carstensen J, Stal O, Nordenskjold B and Wingren S (2007) Genetic variants of CYP3A5, CYP2D6, SULT1A1, UGT2B15 and tamoxifen response in postmenopausal patients with breast cancer. *Breast Cancer Res* 9:R7.

Wegman P, Vainikka L, Stal O, Nordenskjold B, Skoog L, Rutqvist LE and Wingren S (2005) Genotype of metabolic enzymes and the benefit of tamoxifen in postmenopausal breast cancer patients. *Breast Cancer Res* 7:R284-290.

Weinshilboum R (2003) Inheritance and drug response. *N Engl J Med* 348:529-537.

Wennerholm A, Johansson I, Massele AY, Lande M, Alm C, Aden-Abdi Y, Dahl ML, Ingelman-Sundberg M, Bertilsson L and Gustafsson LL (1999) Decreased capacity for debrisoquine metabolism among black Tanzanians: analyses of the CYP2D6 genotype and phenotype. *Pharmacogenetics* 9:707-714.

Winer EP, Hudis C, Burstein HJ, Wolff AC, Pritchard KI, Ingle JN, Chlebowski RT, Gelber R, Edge SB, Gralow J, Cobleigh MA, Mamounas EP, Goldstein LJ, Whelan TJ, Powles TJ, Bryant J, Perkins C, Perotti J, Braun S, Langer AS, Browman GP and Somerfield MR (2005) American Society of Clinical Oncology technology assessment on the use of aromatase inhibitors as adjuvant therapy for postmenopausal women with hormone receptor-positive breast cancer: status report 2004. *J Clin Oncol* 23:619-629.

Xia L, Zhang D, Du R, Pan Y, Zhao L, Sun S, Hong L, Liu J and Fan D (2008) miR-15b and miR-16 modulate multidrug resistance by targeting BCL2 in human gastric cancer cells. *Int J Cancer* 123:372-379.

Xie X, Lu J, Kulbokas EJ, Golub TR, Mootha V, Lindblad-Toh K, Lander ES and Kellis M (2005) Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature* 434:338-345.

Xin F, Li M, Balch C, Thomson M, Fan M, Liu Y, Hammond SM, Kim S and Nephew KP (2009) Computational analysis of microRNA profiles and their target genes suggests significant involvement in breast cancer antiestrogen resistance. *Bioinformatics* 25:430-434.

Xu Y, Sun Y, Yao L, Shi L, Wu Y, Ouyang T, Li J, Wang T, Fan Z, Fan T, Lin B, He L, Li P and Xie Y (2008) Association between CYP2D6 \*10 genotype and survival of breast cancer patients receiving tamoxifen treatment. *Ann Oncol* 19:1423-1429.

Yekta S, Shih IH and Bartel DP (2004) MicroRNA-directed cleavage of HOXB8 mRNA. *Science* 304:594-596.

Yu A, Kneller BM, Rettie AE and Haining RL (2002) Expression, purification, biochemical characterization, and comparative function of human cytochrome P450 2D6.1, 2D6.2, 2D6.10, and 2D6.17 allelic isoforms. *J Pharmacol Exp Ther* 303:1291-1300.

Yu KD and Shao ZM (2009) Genetic matters of CYP2D6 in breast cancer: copy number variations and nucleotide polymorphisms. *Nat Rev Cancer* 9:842.

Zanger UM, Raimundo S and Eichelbaum M (2004) Cytochrome P450 2D6: overview and update on pharmacology, genetics, biochemistry. *Naunyn Schmiedebergs Arch Pharmacol* 369:23-37.

Zhao JJ, Lin J, Yang H, Kong W, He L, Ma X, Coppola D and Cheng JQ (2008) MicroRNA-221/222 negatively regulates estrogen receptor alpha and is associated with tamoxifen resistance in breast cancer. *J Biol Chem* 283:31079-31086.

Zineh I, Beitelshees AL, Gaedigk A, Walker JR, Pauly DF, Eberst K, Leeder JS, Phillips MS, Gelfand CA and Johnson JA (2004) Pharmacokinetics and CYP2D6 genotypes do not predict metoprolol adverse events or efficacy in hypertension. *Clin Pharmacol Ther* 76:536-544.

## CURRICULUM VITAE

Anuradha Ramamoorthy

### Education

2005-2010 Ph.D. Medical and Molecular Genetics,  
Indiana University, USA.  
2003-2004 M.Phil. Genetics  
University of Madras, India.  
2000-2002 M.Sc. Environmental Toxicology,  
University of Madras, India.  
1997-2000 B.Sc. Microbiology,  
Dr. MGR Janaki College, University of Madras, India.

### Honors and Awards

2005-2006 Indiana University Purdue University Indianapolis (IUPUI)  
Fellowship  
2002 Dr. C.R. Krishnamurthy Memorial Medal for Environmental  
Toxicology  
I place in M.Sc. Environmental Toxicology degree examinations  
2000 University of Madras III place in French in B.Sc. degree  
examinations  
University of Madras XIV place in English in B.Sc. degree  
examinations  
Dr. MGR Janaki College I place in Microbiology in B.Sc. degree  
examinations

### Fellowships

11/2008-Present US Department of Defense (DoD) Breast Cancer  
Predoctoral Fellowship. Title: Mechanisms of Variability in  
Breast Cancer Therapies: Role of MicroRNAs in Hepatic Drug  
Metabolism

### Research and Training Experience

2005-2010 Indiana University School of Medicine, Department of  
Medical and Molecular Genetics.  
Mentor: Dr. David Flockhart, M.D., Ph.D.  
The specific aim is to determine the contributions of CYP2D6  
genotype, copy number polymorphism and microRNAs to the  
variability in CYP2D6 metabolism. Since hepatic CYP2D6 activity  
determines the therapeutic exposure of the breast cancer drug

tamoxifen, a better understanding of the regulation of this enzyme will help explain the variable efficacy and side effects of this and related drugs.

- Participation certification in the Principles of Clinical Pharmacology Course conducted by NIH.
- Participation in *NIH T32* Clinical Pharmacology Training Course.
- Certificate in the Business of Life Sciences (Graduate Student Associate).

2006 Indiana University School of Medicine, Department of Medical and Molecular Genetics. Position held: Part-time Research Assistant (Supervisor: Elliot Rosen, Ph.D.).

2005-2006 Indiana University School of Medicine, Department of Medical and Molecular Genetics. Research rotations (Mentors: Brittney-Shea Herbert, Ph.D., Sean Mooney, Ph.D. Elliot Rosen, Ph.D., Howard Edenberg, Ph.D., and Dr. David Flockhart, M.D. Ph.D.). Performed four different research projects involving various molecular biology techniques and bioinformatic analysis.

2004- 2005 University of Madras, Department of Genetics. Position held: Honorary Research Assistant (Supervisor: Srikumari Sri Sailapathy, Ph.D.).

2003-2004 University of Madras, Department of Genetics (Advisor: Srikumari Sri Sailapathy, Ph.D.).  
M.Phil. dissertation: Evaluation of multiplex families with stuttering – a genetic perspective. The overall aim was to evaluate the association between Dopamine D2 Receptor (DRD2) genotype and the stuttering phenotype.

### **Skills**

DNA and RNA isolation, genotyping, reverse transcription, Real Time and conventional PCR, primer design, restriction fragment length polymorphism (RFLP), gel extraction, PCR purification, DNA and RNA electrophoresis, polyacrylamide gel electrophoresis (PAGE), bacterial cloning, plasmid preparation, bacterial transformation, spectrophotometry, western blotting, immunoprecipitation, cell culture, mammalian cell transfection, stable gene expression, microarray basic data analysis, Primer Express, SPSS, DNA sequencing data analysis, pedigree analysis.



## **Publications**

1. Ramamoorthy, A., Flockhart, D.A., Hosono, N., Kubo, M., Nakamura, Y., and Skaar, T.C. CYP2D6 copy number assays: Differential quantification of CYP2D6 gene copy number by four different quantitative real-time PCR assays. *Pharmacogenet Genomics*. 2010 Jul. 20 (7):451-4.
2. Ramamoorthy, A., Desta, Z., Lang, L., Love, R., Flockhart, D.A., Skaar, T.C., et al., CYP2D6 copy number variations in Vietnamese and Filipino breast cancer subjects: effect of \*36 on tamoxifen metabolites. In preparation.
3. Desta, Z., Ramamoorthy, A., Lang, L., Flockhart, D.A., Skaar, T.C., Love, R., et al., Cytochrome P450 (CYP) 2D6 and UDP-Glucuronosyl transferase (UGT) 2B7 genetic variants predict tamoxifen active metabolite (endoxifen) in Vietnamese and Filipino breast cancer patients. In preparation.
4. Ramamoorthy, A., Flockhart, D.A., and Skaar, T.C. *In silico* and *in vitro* identification of microRNAs regulating hepatic drug metabolism: post-transcriptional regulation of HNF4A by microRNAs. In preparation.
5. Ramamoorthy, A., Flockhart, D.A., and Skaar, T.C. Rifampin regulates hepatic miRNA expression. In preparation.

## **Abstracts and Presentations**

1. Ramamoorthy, A., Flockhart, D.A., and Skaar, T.C. *In Silico* analysis of miRNAs that target cytochrome P450s. PGRN, 2007.
2. Ramamoorthy, A., Flockhart, D.A., and Skaar, T.C. Identification of microRNAs that target CYP2D6. ASCPT, 2008.
3. Ramamoorthy, A., Flockhart, D.A., and Skaar, T.C. *In silico* analysis of miRNAs that target cytochrome P450s that are involved in the metabolism of breast cancer drugs. Amelia Project, 2008.
4. Ramamoorthy, A., Flockhart, D.A., and Skaar, T.C. Identification of microRNAs that target HNF4A. ASCPT, 2009.
5. Ramamoorthy, A., Hosono, N., Kubo, M., Nakamura, Y., Flockhart, D.A., and Skaar, T.C. CYP2D6 copy number assays: Differentiation between multiple copies of functional and nonfunctional (\*36) alleles in Asian populations. For ASCPT, 2010.
6. Ramamoorthy, A., Desta, Z., Lang, L., Love, R., Flockhart, D.A., Skaar, T.C., et al., CYP2D6 copy number variations in Vietnamese and Filipino breast cancer subjects: effect of \*36 on tamoxifen metabolites. For ASCPT, 2010.
7. Desta, Z., Ramamoorthy, A., Lang, L., Flockhart, D.A., Skaar, T.C., Love, R., et al., Cytochrome P450 (CYP) 2D6 and UDP-Glucuronosyl transferase (UGT) 2B7 genetic variants predict tamoxifen active metabolite (endoxifen) in Vietnamese and Filipino breast cancer patients. For ASCPT, 2010.
8. Ramamoorthy, A., Flockhart, D.A., and Skaar, T.C. Rifampin regulates hepatic miRNA expression. For ASCPT, 2010.

**Teaching Experience**

2008 and 2009      Teaching Assistant, Q580 Basic Human Genetics  
Lecture: "Recombinant DNA technology"  
(Supervisor: Terry Reed, Ph.D.)

**Professional Activities**

2010                      International Society for the Study of Xenobiotics (ISSX)  
2006-Present          American Society of Clinical Pharmacology and  
Therapeutics (ASCPT)  
2009                      American College of Clinical Pharmacology (ACCP)  
2006-2007              American Academy of Arts and Sciences (AAAS)/ Science:  
Program for Excellence in Science