Title
Genetic variation in membrane transporters

Author
Leabman, Maya Kaushal

Publication Date
2002

Peer reviewed|Thesis/dissertation

Genetic Variation in Membrane Transporters: Implications for Drug Response

by

Maya Kaushal Leabman

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Pharmaceutical Chemistry

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Date University Librarian

Degree Conferred:.....

To my parents and my husband, Michael:

ACKNOWLEDGEMENTS

I would first like this thank my advisor, Dr. Kathleen M. Giacomini, who has been a wonderful teacher, mentor, and role model for me during my graduate career. I greatly appreciate the countless hours Kathy has spent with me over the past five years – preparing me for qualifying exams and helping me with my research. I value the opportunities Kathy has given me and her willingness to allow me to explore new areas of research. Kathy's passion for science and her drive to conduct quality research have inspired me throughout my graduate studies. In addition, Kathy has been an understanding and sensitive advisor, giving me extra words of encouragement or reassurance when I have needed them most.

I thank my thesis committee members, Drs. Deanna Kroetz and Ira Herskowitz, for their help and advice on my research project and their critical review of my dissertation. I also thank Dr. Leslie Benet for being a wonderful chair for my oral qualifying exams. I thank all members of the Pharmacogenetics of Membrane Transporters project – in particular, Conrad Huang, Doug Stryke, Susan Johns, Michiko Kawamoto, and Elaine Carlson – for helping me with many analyses and willingly and happily answering all of my questions. I also thank Dr. Claire Brett for her help and friendship. I thank all members of the Giacomini lab, past and present, for their help in the lab, their input on my research and for making the lab a fun place to work. Moreover, I thank all of them for their friendship and support over the past five years.

I give special thanks to my family for their unconditional support and love during my graduate career. I thank my mother and father for their constant emotional support,

their willingness to listen, and their many words of encouragement. Lastly, I thank my husband Michael – my love, my best friend, and my strongest advocate – for being there through every stage of my graduate career and most importantly, I thank him for his boundless love, patience, and support.

WILEY
Publishers Since 1807

Hoboken, NJ 07030

TEL 201.748.6000
FAX 201.748.6008

October 28, 2002

Dear Maya Leabman:

Thank you for your recent request regarding your contribution to JOURNAL OF PHARMACEUTICAL SCIENCES. We grant you the following rights in your Contribution:

1. The right to share with colleagues print or electronic "preprints" of the unpublished Contribution, in form and content as accepted by Wiley for publication in the Journal. Such preprints may be posted as electronic files on your own website for personal or professional use, or on your internal university or corporate networks/intranet, or secure external website at your institution, but not for commercial sale or for any systematic external distribution by a third party (e.g., a listserve or database connected to a public access server). Prior to publication, you must include the following notice on the preprint: "This is a preprint of an article accepted for publication in [Journal title] © copyright (year) (copyright owner as specified in the Journal)". After publication of the Contribution by Wiley, the preprint notice should be amended to read as follows: "This is a preprint of an article published in [include the complete citation information for the final version of the Contribution as published in the print edition of the Journal]", and should provide an electronic link to the Journal's WWW site, located at the following Wiley URL: http://www.interscience.Wiley.com/. You agree not to update the preprint or replace it with the published version of the Contribution.

2. The right, without charge, to photocopy or to transmit online or to download, print out and distribute to a colleague a copy of the published Contribution in whole or in part, for your personal or professional use, for the advancement of scholarly or scientific research or study, or for corporate informational purposes.

3. The right to republish, without charge, in print format, all or part of the material from the published Contribution in a book written or edited by you.

4. The right to use selected figures and tables, and selected text (up to 250 words, exclusive of the abstract) from the Contribution, for your own teaching purposes, or for incorporation within another work by you that is made part of an edited work published (in print or electronic format) by a third party, or for presentation in electronic format on an internal computer network or external website of yours or your employer.

5. The right to include the Contribution in a compilation for classroom use (course packs) to be distributed to students at your institution free of charge or to be stored in electronic format in datarooms for access by students at your institution as part of their course work (sometimes called "electronic reserve rooms") and for in-house training programs at your employer.

Please include the following notice: [Journal title] © copyright (year) (copyright owner as specified in the Journal).

Please retain this letter for any future uses you may wish to make of your Contribution that are covered by the above grant of rights and for which no additional correspondence is needed. If you have any questions please call 201-748-8607.

Sincerely,

Kathleen Robbins
Permissions Assistant

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO | SANTA BARBARA • SANTA CRUZ

Maya K. Leabman
Department of Biopharmaceutical Sciences
University of California
San Francisco, California 94143-0446

Telephone: (415) 514-0301
FAX: (415) 502-4322
email: kmaya@itsa.ucsf.edu

October 29, 2002

Chaunfayta L. Hightower
Lippincott Williams & Wilkins
Permissions Department
351 West Camden Street
Baltimore, MD 21201
U.S.A.
FAX: (410) 528-4331

To Whom It May Concern:

I am writing to obtain copyright permission to reprint an article that appeared in *Pharmacogenetics* as part of my doctoral dissertation at the University of California, San Francisco. The title of the article is "Polymorphisms in a human kidney xenobiotic transporter, OCT2, exhibit altered function" (Jul; 12(5):395-405, 2002). I would greatly appreciate your prompt attention to this request.

The document can be faxed to 415-502-4322.

Thank you very much for your assistance.

Sincerely,

Maya K. Leabman

Maya K. Leabman
Graduate Student
Department of Biopharmaceutical Sciences
University of California, San Francisco

Permission is granted to reproduce the requested material for use in your academic thesis/dissertation. Permission is granted provided a prominent credit line is placed stating the original source and copyright owner, © Lippincott Williams & Wilkins.

Chaunfayta L. Hightower date: 10-29-02

ABSTRACT

Genetic Variation in Membrane Transporters: Implications for Drug Response

Maya Kaushal Leabman

Membrane transporters are major determinants of drug absorption, distribution, and elimination and thus play a critical role in drug response. The goal of this dissertation research was to identify genetic variants of membrane transporters that are hypothesized to play a role in drug response and to determine the functional significance of variation in membrane transporter genes. In particular, we determined the nature and extent of variation in transporter genes and examined evolutionary, structural, and functional constraints on membrane transporters. First, we identified genetic variants of the human organic cation transporter, OCT2 (*SLC22A2*), which is involved in the renal secretion of organic cations including many clinically used drugs. Twenty-eight variable sites of OCT2 were identified in 247 DNA samples from ethnically diverse populations using denaturing HPLC analysis followed by direct sequencing. Four common non-synonymous variants altered transport function assayed in *Xenopus laevis* oocytes. Comparison of nucleotide diversity (π) and Tajima's *D* at synonymous and non-synonymous sites suggested that selection has acted against amino acid changes in OCT2. To further examine the genetic factors contributing to interindividual variation in renal elimination, we calculated the genetic component (r_{GC}) of renal clearance for five drugs using the RDA method proposed by Kalow *et al.* Heritability of renal clearance of drugs undergoing net secretion was significantly higher than drugs with renal clearance determined by other factors (passive reabsorption and renal blood flow), suggesting that

genetic factors contribute to interindividual variation in net renal secretion. Lastly, we screened for variation in a set of 24 membrane transporter genes and identified 680 SNPs, of which 155 caused amino acid changes. Amino acid diversity (π_{NS}) in transmembrane domains (TMDs) was significantly lower than in loops, suggesting that TMDs have special functional constraints. The allele frequency distribution of non-synonymous changes at evolutionarily conserved (EC) sites had an excess of low frequency alleles, suggesting that changes at EC amino acids are deleterious. These results suggest that analysis of human genetic variation in combination with phylogenic comparisons may help predict structural constraints on proteins.

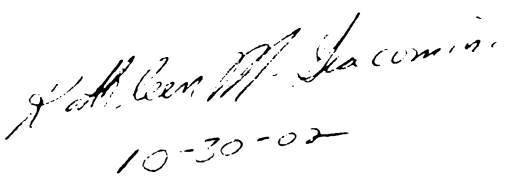


TABLE OF CONTENTS

LIST OF TABLES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 4

LIST OF ABBREVIATIONS

araC	cytosine arabinoside
BBM	brush border membrane
BSEP	bile salt export pump
BLM	basolateral membrane
ClAdo	2-chlorodeoxyadenosine
CNT	concentrative nucleoside transporter
CTR	copper transporter
DAT	dopamine transporter
DMA	dimethylamiloride
DMT	divalent metal transporter
dTub	2'-deoxytubercidin
EC	evolutionarily conserved
ENT	equilibrative nucleoside transporter
EU	evolutionarily unconserved
FIC	familial intrahepatic cholestasis gene
GAT	gamma-aminobutyric acid transporter
IREG	iron-regulated transporter
MDCK	Madin-Darby canine kidney
MPTP	1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine
MPP^+	1-methyl-4-phenylpyridinium
MRP	multidrug resistance-associated protein
NET	norepinephrine transporter

NMN	N^1-methylnicotinamide
NRAMP	natural resistance-associated macrophage protein
OAT	organic anion transporter
OCT	organic cation transporter
OCTN	novel organic cation transporter
PAH	para-aminohippurate
PEPT	peptide transporter
PGE_2	prostaglandin E2
$PGF_{2\alpha}$	prostaglandin F2α
SPNT	sodium-dependent purine nucleoside transporter
TBA	tetrabutylammonium
TBuMA	tributylmethylammonium
TEA	tetraethylammonium
THA	tetrahexylammonium
TMA	tetramethylammonium
TMD	transmembrane domain
TPeA	tetrapentylammonium
TPrA	tetrapropylammonium
VACHT	vesicular acetylcholine transporter
VMAT	vesicular monoamine transporter

CHAPTER 1

ROLE OF ORGANIC CATION TRANSPORTERS IN DRUG ELIMINATION*

INTRODUCTION

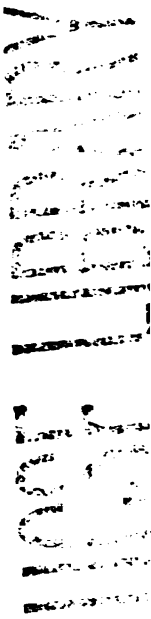
Organic cations are a diverse group of compounds that includes endogenous compounds (*e.g.*, norepinephrine and dopamine), clinically used drugs (*e.g.*, procainamide and cimetidine), as well as toxic substances (*e.g.*, 1-methyl-4-phenylpyridinium [MPP^+]) [1, 2]. The SLC22A family of transport proteins, which includes both the OCTs and OCTNs, mediates the facilitative transport of organic cations in many epithelial tissues throughout the body and is thought to play an important role in the elimination of organic cations – in particular, organic cation drugs – from the systemic circulation. This organic cation transporter family includes six members to date: OCT1-OCT3 and OCTN1-OCTN3. During the past decade, significant progress has been made in the cloning, functional characterization, and localization of these transporters. Furthermore, genomic sequences for these transporters have become increasingly available, facilitating the construction of animal models and providing valuable information about the *in vivo* roles of organic cation transporters.

Over the past couple of years, sequencing of the human genome and large-scale SNP (single nucleotide polymorphism) identification have provided information on human genetic variation which can be applied to transporter research. For example, sequencing of *OCTN2* in discrete populations has allowed the identification and association of polymorphisms in this gene with primary systemic carnitine deficiency (SCD), an autosomal recessive disease [3]. Polymorphisms in other organic cation

*This work was published in *Journal of Pharmaceutical Sciences* 90(4):397-421, 2001. Permission from the publisher is included in "Acknowledgements".

transporter genes have also been identified, and information regarding these polymorphisms is now available in SNP databases. The increase in available data on genetic variation in organic cation transporters will allow us to further elucidate their physiological and pharmacological roles and will ultimately allow us to test hypotheses about the effects of genetic variation in OCTs and OCTNs on drug response.

In this chapter, I summarize and evaluate the progress made in cloning and characterization of organic cation transporters and consider their roles in drug elimination. In particular, three critical areas will be reviewed: 1) the molecular mechanisms of organic cation transport, which includes molecular characteristics, tissue distribution and membrane localization, and functional characteristics such as transport mechanism and substrate specificity; 2) animal models of organic cation transport – the OCT1 knock-out mouse, the OCT3 knock-out mouse, and the juvenile visceral steatosis (*jvs*) mouse; 3) organic cation transporter polymorphism identification studies.



MOLECULAR MECHANISMS OF ORGANIC CATION TRANSPORT

OCT1 (SLC22A1): *1. Molecular Characteristics.* The first organic cation transporter gene, rOCT1, was isolated by Grundemann and co-workers in 1994 from a rat kidney cDNA library using expression cloning techniques in *Xenopus laevis* oocytes [4]. rOCT1 is 554 amino acids in length and is predicted to have 12 transmembrane domains (TMDs). Following the cloning of rOCT1, the rabbit, mouse and human orthologs were cloned by homology [5-8]. These transporters have been termed rbOCT1, mOCT1, and hOCT1. The human and rabbit orthologs are approximately 80% identical to rOCT1,

whereas mOCT1 shares higher sequence identity with rOCT1. In addition, a novel splice variant of rOCT1, termed rOCT1A, was later cloned by Zhang and co-workers [9]. This variant is virtually identical to rOCT1 with the exception of a 104 base pair deletion at the 5-prime end of the cDNA. *2. Tissue Distribution and Localization.* Much work has been done to characterize the tissue expression patterns and membrane localization of OCT1. Northern blot analyses and *in situ* hybridization show that mRNA transcripts of rOCT1 are primarily expressed in the liver (hepatocytes) and kidney (proximal tubules) and are expressed at lower levels in the small intestine (enterocytes) (Table 1) [4, 10]. In the kidney, rOCT1 has been further localized to the basolateral membrane, while in the liver, it has been localized to the sinusoidal membrane of the hepatocyte (Figure 1) [11, 12]. More recent studies have demonstrated that rOCT1 mRNA is present in S1, S2, and S3 segments of proximal tubules in the renal cortex whereas rOCT1 protein is mainly confined to the S1 and S2 segments [13]. Although the rabbit, mouse, and human orthologs of OCT1 have not yet been localized to particular membranes, their tissue expression patterns have been determined and are, in general, similar to those of rOCT1. However, the relative amounts of each OCT1 ortholog may vary from one tissue to another. Both rbOCT1 and hOCT1 are expressed most strongly in the liver and have lower levels of expression in other tissues, including the kidney and intestine (as well as the muscle and heart, for hOCT1) [6-8]. In contrast, the rat ortholog is present in abundance in both the kidney and liver. *3. Functional Characteristics.* All of the OCT1 orthologs have been shown to transport the prototypical organic cations tetraethylammonium (TEA) and/or 1-methyl-4- phenylpyridinium (MPP^+) (Figure 2). Furthermore, studies in *X. laevis* oocytes have shown that transport of these

Table 1: Tissue distribution of cloned organic cation transporters[a]

Transporter	Kidney	Liver	Intestine	Brain	Placenta	Lung	Heart	Ref.
mOCT1	++	++	++	-	-	n/a	-	[5]
rOCT1	++	++	+	-	-	-	-	[10, 14, 15]
rOCT1A	++	++	+	-	n/a	-	-	[9]
rbOCT1	+	++	+	n/a	n/a	-	-	[6]
hOCT1	+	++	+	+	+	-	+	[7, 8]
mOCT2	++	-	n/a	+	n/a	n/a	n/a	[16]
rOCT2	++	-	-	+	n/a	-	-	[15, 17, 18]
hOCT2	++	-	+	+	+	n/a	n/a	[8, 19]
mOCT3	++	n/a	n/a	n/a	n/a	n/a	n/a	[20]
rOCT3	+	-	+	+	++	+	+	[10, 21]
hOCT3	+	++	-	+	+	+	+	[20, 22, 23]
mOCTN1	++	++	n/a	-	n/a	+	-	[24]
rOCTN1	+	++	++	+	+	+	+	[25]
hOCTN1	++	-	-	-	+	+	+	[26]
mOCTN2	++	+	n/a	+	n/a	+	+	[24]
rOCTN2	+	n/a	n/a	+	+	n/a	+	[27]
hOCTN2	++	+	+	+	++	+	++	[28, 29]
mOCTN3	+	-	n/a	-	n/a	-	-	[24]

[a] Tissue distribution determined by Northern blot, RT-PCR, *in situ* hybridization, and/or RNase protection assays. Key: (++) high expression, (+) moderate to low expression, (-) absent, (n/a) not determined.

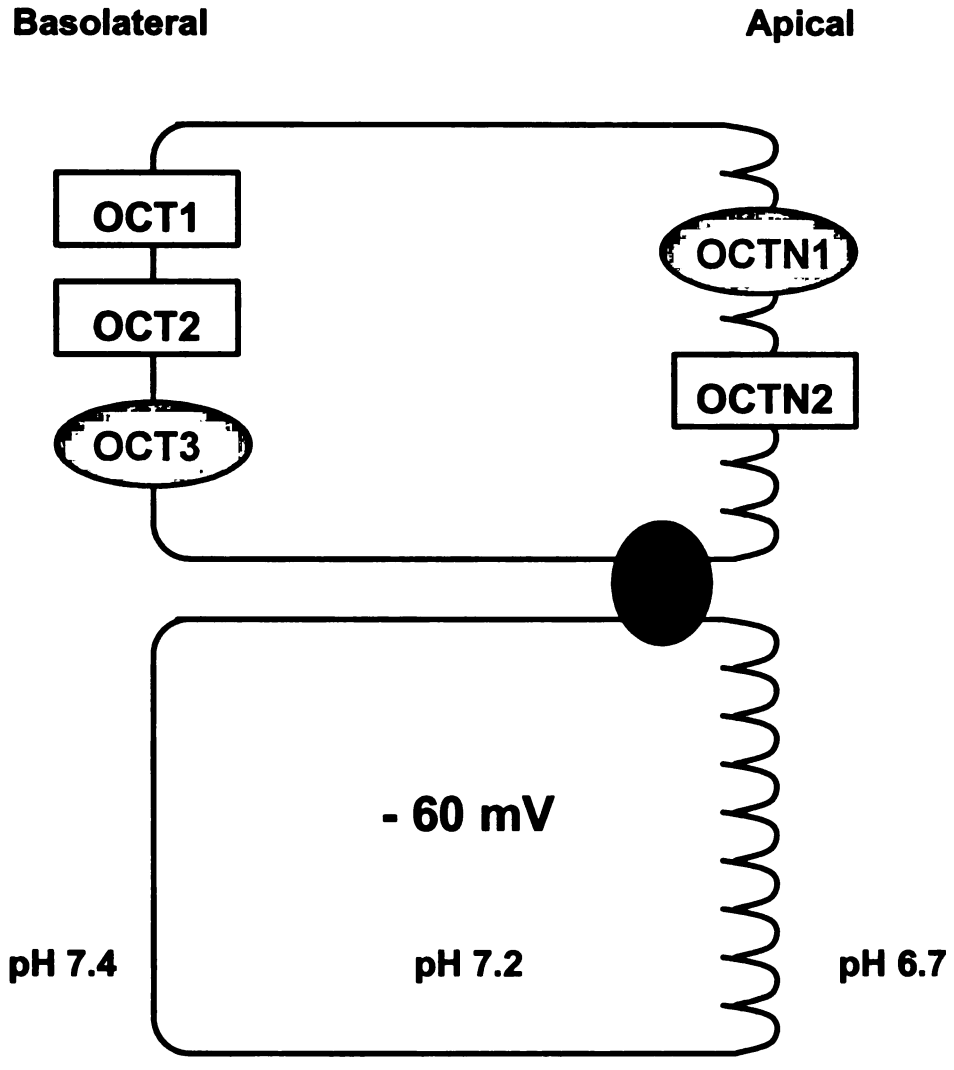
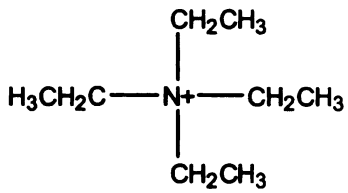
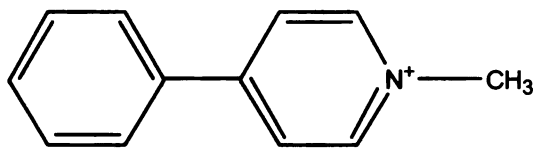


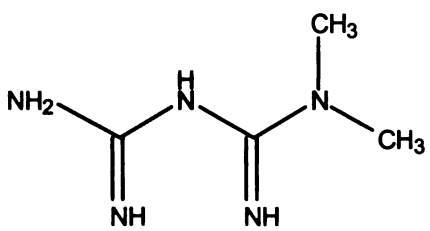
Figure 1. Membrane localization of organic cation transporters in renal epithelial cells. Transporters enclosed in rectangles have been localized by immunohistochemistry; localization of transporters enclosed in shaded ovals is based on functional studies.



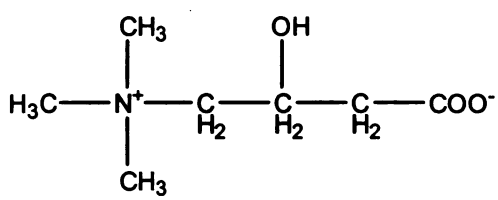
TEA



MPP+



Metformin



Carnitine

Figure 2. Structures of model organic cations that are substrates of organic cation transporters. TEA, tetraethylammonium; MPP^+, 1-methyl-4-phenylpyridinium.

organic cations by rOCT1, rbOCT1, and hOCT1 is sensitive to membrane potential, a characteristic of a basolateral membrane organic cation transporter [2, 4, 6, 7]. Transport of substrates by rOCT1 and hOCT1 is significantly inhibited by numerous drugs, including procainamide, desipramine, clonidine, araC, and AZT. hOCT1 (and hOCT2) also transports the prostaglandins PGE_2 and $PGF_{2\alpha}$, both of which are acidic compounds and metabolites of arachidonic acid in the kidney [30]. Table 2 lists the K_i values of these compounds as well as others for OCT1-OCT3 and OCTN1-OCTN3. It should be noted that some of the compounds in Table 2 have relatively weak interactions with the listed transporter (*e.g.*, high K_i values) and therefore, it is likely that other transporters may play significant *in vivo* roles in the transport of these compounds. Studies have demonstrated that rOCT1 and hOCT1 have the transport properties of the uptake system I in rat hepatocytes; they transport relatively small hydrophilic "type I" organic cations and are inhibited by bulkier "type II" organic cations but not taurocholate and cardiac glycosides [31]. A recent study has shown that hydrophobicity is a major determinant of drug interactions with hOCT1 (Table 2) [32, 33]. In general, the K_i values of most organic cations are higher for hOCT1 than for rOCT1, suggesting that hOCT1 has a lower affinity for organic cations [32, 34]. In a study comparing the kinetics and substrate selectivities among the cloned OCT1 transporters from mouse, rat, rabbit, and human, significant differences were found among these four transporters, which suggests that OCT1 may be responsible, in part, for interspecies differences in the elimination of organic cations [34].

OCT2 (SLC22A2): *1. Molecular Characteristics.* Soon after the cloning of rOCT1, rOCT2 was cloned from a rat kidney cDNA library using RT-PCR [18]. Subsequently, the pig, mouse, and human orthologs were cloned and termed pOCT2, mOCT2, and hOCT2 [8, 15, 16]. The OCT2 cDNAs are predicted to encode proteins of 593, 555, 554 and 555 amino acids in length for the rat, human, pig, and murine homologs, respectively. All of the OCT2 proteins are predicted to consist of 12 TMDs. *2. Tissue Distribution and Localization.* Unlike OCT1, which has a wide tissue distribution, OCT2 is expressed predominantly in the kidney and to a lesser extent in the brain [8, 15-19]. *In situ* hybridization and immunohistochemistry studies demonstrate rOCT2 expression mainly in the S2 and S3 segments of kidney proximal tubules [13, 35]. Furthermore, expression of rOCT2 appears to be higher in the outer stripe of the medulla; this is in contrast to rOCT1, which has higher expression in the renal cortex [13, 35]. The membrane localization of OCT2 has been extensively studied, but many discordant results have been reported [8, 11, 15, 36, 37]. Detailed mechanistic studies in stably transfected MDCK cells suggest that rOCT2 is a potential-sensitive transporter, which is consistent with a renal basolateral membrane organic cation transporter and not a brush border membrane organic cation-proton exchanger [37]. Furthermore, MDCK cells stably transfected with rOCT2-GFP display increased basal and lateral fluorescence [38]. Recent studies have shown that renal basolateral membranes (BLMs), and not brush-border membranes (BBMs), immunoreact with a rOCT2-specific antibody, confirming the basolateral localization of rOCT2 in the kidney. In the brain, rOCT2 appears to be localized to the apical membrane of choroid plexus [39]. *3. Functional Characteristics.* TEA is a substrate of rOCT2, hOCT2, and pOCT2. rOCT2, like OCT1, is a potential-

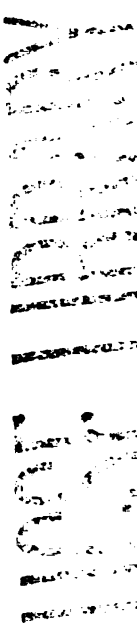
sensitive transporter [37]. The OCT2 class of transporters interacts with many of the same organic cation drugs as OCT1, including procainamide, desipramine, and cimetidine (Table 2). Inhibition of TEA uptake in stably transfected MDCK cells indicates that rOCT1 and rOCT2 have similar affinities for many organic cations [11]. However, more work is needed to determine if the specificities of these two isoforms completely overlap; a recent study has shown that there are differences between rOCT1 and rOCT2 specificities, at least for some compounds [40]. Studies in rOCT2-expressing NIH3T3 cells have shown that the anti-cancer drug cisplatin inhibits TEA uptake with a K_i of 925 μM, but other anti-tumor agents such as daunomycin and vinblastine do not interact with rOCT2 [41]. Several recent studies have also begun to examine the regulation of OCT2 by various steroids and hormones. Urakami *et al.* demonstrated that testosterone increases OCT2 mRNA and protein levels in male and female rats, which is accompanied by a corresponding increase in TEA uptake into renal slices. In contrast, estradiol moderately decreases OCT2 expression and TEA uptake into male rat renal slices [42]. In addition, studies have shown that OCT2 mRNA levels in MDCK cells are significantly increased following exposure of the cells to the steroid hormones, dexamethasone, hydrocortisone, and testosterone, suggesting that steroids may transcriptionally regulate OCT2 [43].

OCT3 (SLC22A3): *1. Molecular Characteristics.* OCT3 was first cloned from a rat placental cDNA library by cross-hybridization [21]. Subsequently, the human and mouse orthologs were cloned [20, 22, 23]. OCT3 from rat and mouse each have 551 amino acid residues, whereas the human has 556; all three transporters are predicted to consist of 12

TMDs. At the protein level, rOCT3 shares 95% and 83% identity with mOCT3 and hOCT3 (EMT), respectively. The OCT3 proteins share ~ 50% identity with the other two OCT isoforms. *2. Tissue Distribution and Localization.* Unlike OCT1 and OCT2, OCT3 is expressed at high levels in the placenta. Northern blot analysis demonstrates that rOCT3 is also expressed in the kidney, intestine, heart, and brain [21]. Studies using *in situ* hybridization have further localized rOCT3 to the hippocampus, cerebellum, and cerebral cortex regions of the rat brain and to the cortex of the mouse kidney [10, 20]. While the expression pattern of mOCT3 is similar to that of rOCT3, hOCT3 has a much broader tissue distribution [23]. Northern blot analysis has shown that in addition to the placenta, hOCT3 is highly expressed in the aorta, liver, prostate, salivary glands, adrenal glands, skeletal muscle, and fetal lung [23]. Intracellular localization of OCT3 has not yet been determined in any species. *3. Functional Studies.* Studies in a transfected human retinal pigment epithelial cell line (HRPE) and *X. laevis* oocytes have demonstrated OCT3-mediated transport of the prototypical organic cations, TEA, guanidine, and MPP^+ [20, 21]. Efflux studies in HeLa cells as well as electrophysiology studies in oocytes have shown that TEA transport via rOCT3 is H^+-independent and membrane-potential dependent, the functional characteristic of a basolateral membrane organic cation transporter [21]. Studies suggest that rOCT3 interacts with dopamine, the neurotoxins amphetamine and methamphetamine, as well as a variety of steroids (Table 2) [10]. TEA uptake in rOCT3-transfected HRPE cells shows marked inhibition by the steroids, beta-estradiol, corticosterone, deoxycorticosterone, papaverine, testosterone, and progesterone. Uptake studies have shown that hOCT3 transports a variety of organic cations, including catecholamines [20, 22].

OCTN1 (SLC22A4): *1. Molecular Characteristics.* The first member of a new subfamily of organic cation transporters (OCTNs), termed hOCTN1, was cloned from human fetal liver by Tamai and co-workers in 1997 [26]. Subsequently, rOCTN1 and mOCTN1, the rat and mouse orthologs, were cloned from placenta and kidney, respectively [24, 25]. hOCTN1, rOCTN1, and mOCTN1 consist of 551, 553, and 553 amino acids, respectively, and are predicted to have 12 TMDs. The rat and mouse OCTN1 orthologs are ~ 85% identical to hOCTN1 at the protein level; in addition, they share ~ 70% identity with OCTN2 and ~ 30% identity with the OCT family. hOCTN1 is unique among the organic cation transporters in that it has a nucleotide binding site sequence motif [26]. *2. Tissue Distribution and Localization.* Initial studies suggest that hOCTN1 has a broad tissue distribution. Northern blot analysis shows strong mRNA expression of hOCTN1 in fetal kidney, lung, and liver as well as adult kidney, trachea, and bone marrow [26]. Weak signals are detected in many other tissues. In addition, hOCTN1 mRNA transcripts are detected in several human cancer cell lines. rOCTN1 is primarily expressed in the intestine, liver, and kidney while mOCTN1 is primarily expressed in the heart, liver, and kidney [24, 25]. Intracellular localization studies have not yet been performed for either transporter. *3. Functional Characteristics.* rOCTN1, mOCTN1, and hOCTN1 accept the prototypical organic cation, TEA, as a substrate [24-26]. However, kinetic studies in transfected cells and *X. laevis* oocytes have shown that their affinity for TEA is lower than that of OCT1, OCT2, and OCT3 (Table 2) [26, 44]. TEA efflux from transfected HEK 293 cells expressing hOCTN1 is stimulated by acidic pH in the external media, suggesting that transport via hOCTN1 may be driven by a

proton gradient, the driving force of the renal brush border membrane transport system [44]. hOCTN1 transports not only TEA, but also other drugs and endogenous compounds including quinidine, verapamil, and carnitine (Table 2) [44]. In addition, hOCTN1 is inhibited by a variety of structurally diverse compounds, including cimetidine, procainamide, pyrilamine, quinine, cephaloridine, and verapamil (Table 2) [44]. rOCTN1 interacts with many of these compounds, but some specificity differences between rOCTN1 and hOCTN1 have been reported (Table 2) [25]. Interestingly, mOCTN1 transports carnitine in a Na^+-dependent manner while transport of carnitine by hOCTN1 is Na^+-independent and rOCTN1 is unable to transport carnitine [24].

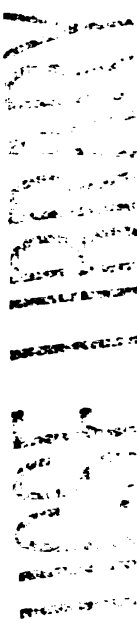


OCTN2 (SLC22A5): *1. Molecular Characteristics.* hOCTN2 was originally cloned from a human placental trophoblast cell line using homology screening. The mouse and rat homologs of OCTN2 have subsequently been isolated [27, 29]. OCTN2 cDNAs are predicted to encode proteins of 557 amino acids and consist of 12 TMDs. The three OCTN2 orthologs share ~ 83% identity in amino acid sequence. OCTN2 is ~ 88% similar to hOCTN1 and, like OCTN1, contains a nucleotide binding site sequence motif. OCTN2 shares ~ 35% identity with members of the OCT family. Two splice variants of rOCTN2 were recently identified [45]. One of the splice variants, found in rat liver RNA, has a deletion of 88 amino acids (all of exon 3 and part of exon 4), which results in a 469 amino acid protein with intact N and C termini but missing four transmembrane domains. The second splice variant is missing all of exon 3 and is predicted to result in a frameshift and a truncated protein of 226 amino acids. *2. Tissue Distribution and Localization.* Northern blot analysis has shown that hOCTN2 is strongly expressed in

fetal as well as adult kidney, skeletal muscle, placenta, heart, prostate, thyroid, and brain [28]. The fetal expression of OCTN2 differs from that of OCTN1 in that hOCTN2 is expressed mainly in the kidney, whereas hOCTN1 is expressed mainly in the liver. However, both hOCTN1 and hOCTN2 have a fairly wide distribution in both adult and fetal tissues. *In situ* hybridization studies demonstrated that rOCTN2 mRNA is expressed in the kidney (cortex), heart, placenta, and brain (Table 1) [27]. The rOCTN2 splice variant transcripts have been detected in the same tissues as the full length transcript, but they are present in much lower abundance (0.1-1%) [45]. Similar to rOCTN2, mOCTN2 is expressed in the kidney, heart, and brain and is also expressed in the liver [24]. OCTN2 has been localized to the apical membrane of renal epithelial cells from both mice and rats [24]. *3. Functional Characteristics.* Studies have demonstrated that OCTN2 functions as both a Na^+-independent organic cation transporter and a Na^+-dependent, high affinity carnitine transporter. However, OCTN2 appears to transport the classical OCT substrate, TEA, to a lesser extent than other organic cation transporters [28, 46]. In addition to carnitine, OCTN2 transports and interacts with a variety of drugs (Table 2). For example, studies with hOCTN2 have demonstrated transport of betaine and the β-lactam antibiotic, cephaloridine (Table 2) [14, 46]. Other drugs that interact with OCTN2 include quinidine, verapamil, cefepime, and emetine (Table 2). The substantial inhibition of hOCTN2 by emetine is of particular note since emetine is known to cause secondary carnitine deficiency in some individuals. This side effect could be due to inhibition of OCTN2 carnitine transport. Recently Wagner and co-workers have shown that carnitine (a zwitterion) transport via hOCTN2 is electrogenic [46]. In addition, they examined whether hOCTN2 is also capable of operating as an organic

cation-proton antiporter using TEA, choline, and carnitine as model organic cations; their results suggest that hOCTN2 does not function in this mode [46].

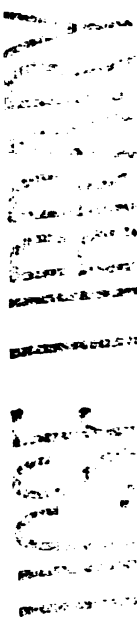
OCTN3 (SLC22A9): *1. Molecular Characteristics.* The most recent member of the OCTN family is mOCTN3. OCTN3 was cloned by Tamai *et al.* from mouse embryo cDNA and encodes a protein of 564 amino acids [24]. It is ~80% identical to OCTN2 protein and ~70% identical to OCTN1 protein. OCTN3 has not yet been cloned from another species. *2. Tissue Distribution and Localization.* RT-PCR and Western blot analysis of adult mouse tissues has shown the mOCTN3 mRNA and protein are expressed in the testis and weakly in the kidney but are not significantly expressed in any other tissue [24]. *3. Functional Characteristics.* Functional studies in transfected HEK293 cells have demonstrated that mOCTN3 transports carnitine in a Na^+-independent manner, in contrast to OCTN1 and OCTN2, which transport carnitine in a Na^+-dependent manner. Furthermore, OCTN3 does not significantly transport the prototypical organic cation TEA. The relative uptake activity ratio of carnitine to TEA was 746 for OCTN3, compared to 1.78 and 11.3 for OCTN1 and OCTN2, respectively [24]. These results suggest OCTN3 has a very high specificity for carnitine compared to the other OCTN transporters.



ANIMAL MODELS

In vitro molecular studies with cloned organic cation transporters have defined their primary functions at the cellular level (*e.g.*, substrate specificities and transport mechanisms), but for many of the OCTs and OCTNs, their role at the level of the whole

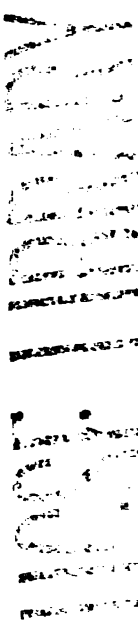
organism is still unknown. Animal models, such as knock-out or transgenic animals, are valuable tools for studying the physiological importance of a particular gene. For example, studies have shown that mice lacking the *Ctr1* gene – a gene encoding for a copper transporter – die *in utero*, suggesting a vital and necessary role of this transporter in development [47, 48]. In the case of xenobiotic transporters, the phenotype of knock-out animals has been shown to be less dramatic. *mdr1a* -/- mice, which lack functional P-glycoprotein, display a phenotype only upon exposure to particular drugs; for example, *mdr1a* -/- mice have heightened sensitivity to ivermectin (a drug used to treat mite infestation) due to increased drug accumulation in the brain [49]. To date, three animals models of organic cation transport have been studied - the OCT1 knock-out mouse, the OCT3 knock-out mouse, and the *jvs* mouse, an inbred strain which has a null OCTN2 gene. Below we summarize these studies with particular emphasis on their physiological and pharmacological relevance.



OCT1: In 2001, Jonker *et al.* generated mice lacking the *Oct1* gene by targeted disruption of exon 7, which encodes for transmembrane domains 7 through 9 of Oct1 [50]. Mice homozygous for this disruption had no detectable transcription of Oct1 mRNA in their liver and kidneys, as shown by Northern analysis. *Oct1* -/- mice were healthy, viable, fertile, and displayed no phenotypic abnormalities, suggesting that Oct1 is not essential for normal health in mice [50]. To determine the pharmacological role of Oct1, the tissue distribution of the model substrate TEA was examined in wild-type and *Oct1* -/- mice. Following an intravenous dose of ^{14}C-TEA, the accumulation of TEA in the liver was reduced six-fold, and direct intestinal excretion of TEA was reduced two-

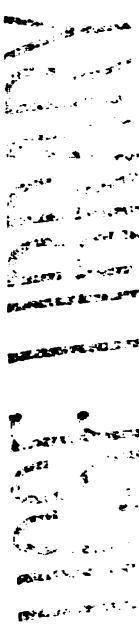
fold in *Oct1* -/- mice compared to wild-type controls [50]. In contrast, cumulative excretion into the urine was increased 1.5-fold, probably due to decreased accumulation in the liver [50]. Although Oct1 is known to be expressed in the kidneys of mice, the increased urinary excretion of TEA in *Oct1* -/- mice suggests that other transporters in the kidney may compensate for its loss of function in the knock-out mice. Oct2 is highly expressed in mice kidneys and has overlapping substrate specificities with Oct1. It is therefore likely to be the transporter compensating for the loss of Oct1 function in the kidney. Two other organic cations, MPP^+ and MIBG, also displayed reduced accumulation in the livers of *Oct1* -/- mice; there were no significant differences in the tissue distributions of cimetidine and choline in *Oct1* -/- and control mice [50].

These results suggest that Oct1 participates in the hepatic and intestinal uptake of organic cations like TEA, MPP^+, and MIBG. These results verify the molecular studies which demonstrated high expression of OCT1 on the BLM of hepatoctyes and enterocytes; in addition, they suggest that OCT1 may be one of the primary organic cation transporters facilitating transport of organic cations from the systemic circulation into hepatocytes. Since many organic cation drugs are eliminated to a significant extent by the liver and since OCT1 transports a variety of clinically used drugs, OCT1 may be an important determinant of the pharmacokinetics of many organic cations.



OCT3: Like the *Oct1* -/- mice, *Oct3* -/- mice are viable and display no apparent phenotype [51]. They have normal life spans and breeding behavior. Furthermore, histology of the placenta and heart, the tissues expressing the highest levels of Oct3, show no cellular alterations compared to wild-type controls [51]. The *in vivo* role of

Oct3 was examined by measuring the tissue accumulation of MPP^+ in both male and female *Oct3* -/- and control mice following an intravenous dose. MPP^+ accumulation was decreased in the hearts of male and female *Oct3* -/- mice while there were no significant differences in MPP^+ levels in any of the other organs examined (*i.e.*, liver, brain, lung, stomach, kidney, adrenal glands, spleen, small intestine, cecum, colon, muscle, uterus, and testis) [51]. Since OCT3 is expressed at high levels in the placenta, MPP^+ concentration was also measured in the embryo, placenta, and amniotic fluid of *Oct3* -/- and control mice. MPP^+ concentrations were reduced only in the embryos of the *Oct3* -/- mice (three-fold lower than wild-type controls), suggesting a role of Oct3 at the fetoplacental interface [51]. Levels of the endogenous monoamines, dopamine and noradrenaline, both known substrates of Oct3, were also measured in the placentas and embryos of *Oct3* -/- and control mice. Although levels were reduced 50% in the *Oct3* -/- mice, the results were not statistically significant (probably due to inherent variability in dopamine and noradrenaline levels) [51]. The combined results from the studies with Oct3 null mice suggest that Oct3 plays an important role in the transport of organic cations in the heart as well as at the fetoplacental interface.



OCTN2: Unlike OCT1 and OCT3, the animal model for OCTN2 is not a genetically engineered knock-out mouse; instead, it is an inbred strain of mice that exhibits systemic carnitine deficiency – the juvenile visceral steatosis (*jvs*) mouse. The *jvs* mouse has been shown to harbor a missense mutation in the *OCTN2* gene (L352R) that completely abolishes carnitine transport in transfected HEK293 cells and *jvs* mouse hepatocytes [3]. Since the identification of the OCTN2 mutation in *jvs* mice, studies have been conducted

in this animal model to examine the *in vivo* role of OCTN2 in organic cation transport. Recently, Ohashi *et al.* examined the pharmacokinetics of TEA in wild-type and *jvs* mice and showed that *jvs* mice have a decreased renal secretory clearance and increased AUC of TEA compared to controls [52]. In addition, tissue-to-plasma concentration ratios (K_p values) were lower in the brain, lung, liver, and spleen and higher in the kidneys of *jvs* mice (no significant difference in K_p values in muscle, heart, gut, testis, and fat) [52]. These studies suggest that OCTN2 plays an important role in the uptake of organic cations into the brain, lung, liver, and spleen. Moreover, the reduced renal secretory clearance and increased K_p value of TEA in the kidneys of *jvs* mice suggests that OCTN2 mediates renal secretion of organic cations. Combined with the *in vitro* studies discussed previously, these studies suggest that OCTN2 functions as an organic cation/carnitine exchanger in which the reabsorption of carnitine is coupled to the secretion of organic cations such as TEA.

Tissue distribution studies of L-carnitine and acetyl-L-carnitine in wild-type and *jvs* mice show decreased K_p values in the muscle, heart, kidney, and gut of *jvs* mice. There were no significant differences in the K_p values of the brain, lung, and liver of wild-type and *jvs* mice. However, uptake of acetyl-L-carnitine into cultured brain capillary endothelial cells (BCECs) of jvs mice was reduced ~ 50% compared to controls, suggesting a potential role of OCTN2 in transport of carnitine across the blood brain barrier.

In addition to the *jvs* mouse, a mouse strain homozygous for a deletion of 450 kb on chromosome 11, which includes the mouse *OCTN2* gene, was recently constructed by Zhu *et al.* [53]. These mice exhibit hypertriglyceridemia as well as other phenotypic

abnormalities [53]. Introduction of a human OCTN2 transgene into this mouse strain corrected the triglyceride abnormalities, suggesting an additional, previously unrecognized role of OCTN2 in the regulation of plasma triglyceride metabolism [53].

GENETIC VARIATION IN HUMAN ORGANIC CATION TRANSPORTERS

The animal models discussed above have provided insight into the physiological roles of OCT1, OCT3, and OCTN2. However, in order to fully understand the *in vivo* consequences of abolished or altered organic cation transport in humans, one must correlate human OCT/OCTN genotypes with a particular disease susceptibility or drug-response phenotype. To date, only one of the organic cation transporters, OCTN2, has been associated with a human phenotype. In 1999, Nezu *et al.* demonstrated that mutations in OCTN2 were associated with the autosomal recessive disease, primary systemic carnitine deficiency [3]. Since this discovery, many mutations in OCTN2 have been documented. In addition to these OCTN2 mutations, mutations in other organic cation transporters are being identified in large scale polymorphism screens, and the data are now available in databases. In this section, we will review the human OCTN2 mutations identified to date and their phenotypic consequences.

Mutations in OCTN2

In the noteworthy study of Nezu *et al.*, four mutations in OCTN2 - a 113 base pair deletion, a frameshift mutation, a nonsense mutation, and a splice mutation - were identified in three pedigrees containing family members affected with systemic carnitine deficiency [3]. Since this discovery, a number of other OCTN2 mutations have been

identified and are summarized in Table 3 and Figure 3. These mutations can be classified into mutations that abolish or decrease transporter function.

Many of the mutations that abolish OCTN2 function result from insertion/deletions or nonsense mutations. Three insertion/deletion mutations (1-176del, 226insC, and IVS8-1G→A) have been identified [3]. The 1-176del mutation results from a deletion of 113 base pairs encompassing the start codon at the 5' end of OCTN2; the next available ATG in frame is located at codon 177. The 226insC mutation contains an insertion of a cytosine just after the start codon and results in a frameshift. The IVS8-1G→A mutation changes the splice acceptor site at the 3' end of intron 8 from AG to AA, most likely resulting in the loss of exon 9 and the creation of a premature stop codon. The nonsense mutations identified to date include W132X, R282X, Y401X, and 458X [54-56]. It is interesting to note that the R282X mutation was found in three separate studies and in individuals of different ethnicities, suggesting that this mutation is a recurrent or ancient founder mutation.

In addition to these truncation mutations, multiple missense mutations that abolish carnitine transport have been identified [54, 56-58]. *In vitro* studies in stably or transiently transfected cell lines have demonstrated that these mutations indeed reduce or abolish function. Three of the six missense mutations that abolish OCTN2 function occur at highly conserved residues. The P478L mutation occurs at a proline residue that is conserved throughout the OCT/OCTN family whereas the R169W and R169Q mutations occur at an arginine residue that is located within a sugar transporter sequence motif and conserved throughout the entire SLC superfamily [54, 57, 58].

A number of missense mutations have been identified that lead to decreased transporter function; these include E452K, S467C, W283C, M179L, W283C, and V446F [57, 59-61]. With the exception of M179L, most of these mutations reduce carnitine transport 90-98%. The E452K mutation decreases the V_{max} of carnitine transport by impairing the Na^+ stimulation of carnitine transport [61]. The exact mechanism by which the other missense mutations decrease the carnitine transport function of OCTN2 is still unknown. Furthermore, correlations between levels of residual carnitine activity and SCD phenotype have not been found.

FUTURE DIRECTIONS

During the past decade significant progress has been made in the functional characterization of organic cation transporters; however, many questions regarding their molecular characteristics remain unanswered. For example, the molecular determinants of substrate specificity and transport mechanism have not yet been identified. Which domains and amino acid residues govern transport function and substrate specificity of organic cation transporters (Figure 4)? Are these residues conserved throughout the SLC22A family (Figure 5)? Such questions can be addressed by conducting *in vitro* site-directed mutagenesis studies to determine the functional importance of particular amino acid residues.

Additionally, studies have demonstrated that many drugs interact with multiple transporter isoforms. Cimetidine is a known substrate of OCT1, OCT2, and OCT3. However, the relative role of each transporter in the systemic elimination of cimetidine remains to be determined. Pharmacokinetic studies in animal models, for example, the

Oct1 -/- mice, will help to further reveal the *in vivo* pharmacological roles of each transporter. In addition, the intracellular localization and driving force(s) of each transporter must be established and much work is needed in this area. Although OCT1, OCT2, and OCTN2 have been convincingly localized in renal epithelial cells, the membrane localization of OCT3, OCTN1, and OCTN3 has not been determined. Such studies may help to define the roles of each transporter within a particular epithelial cell.

In addition to these molecular studies, studies of genetic variation in OCTN2 in human populations have demonstrated that mutations in this transporter are associated with primary carnitine deficiency. Genetic variation in other organic cation transporters could have profound effects on interindividual variability in drug response and pharmacokinetics. For example, studies in *Oct1* -/- mice suggest an important role of OCT1 in the hepatic elimination of organic cation drugs. Individuals carrying deleterious mutations in the *OCT1* gene could have altered hepatic elimination of some organic cation drugs. Future studies correlating organic cation transporter genotypes with *in vitro* and *in vivo* phenotypes will help to further elucidate their physiological and pharmacological roles.

Table 2: Compounds that interact with organic cation transporters

Compound	Transporter[a]	K_i or (K_m)[b] (μM)	Reference
Acebutolol	hOCT1	96	[32]
Acetyl-L-carnitine	hOCTN2	(8.5)	[62]
Acyclovir (ACV)	hOCT1	(151)	[63]
Adrenaline	rOCT2	(1900)	[22]
	hOCT3	ND	[22]
Aldosterone	hOCTN2	ND	[62]
Amantadine	hOCT1	ND	[32]
	hOCT2	(27)	[19]
Amphetamine	rOCT3	42	[10]
araC	rOCT1	ND	[40]
Azidoprocainamide	rOCT1	(54)	[31]
	hOCT1	(101)	[31]
AZT (zidovudine)	rOCT1	ND	[40]
Bamet-R2	hOCT1	(58)	[64]
	hOCT2	(13)	[64]
Bamet-UD2	hOCT1	(15)	[64]
	hOCT2	(8)	[64]
Betaine	hOCTN2	ND	[46]
L-carnitine	hOCTN1	ND	[44]
	mOCTN2	ND	[27]
	rOCTN2	(15)	[27]
	hOCTN2	ND, (4.3), (4.8)	[27, 46, 62]
D-carnitine	hOCTN1	ND	[44]
	hOCTN2	(11), (98)	[46, 62]
Cefepime	rOCTN2	2100	[14]
	hOCTN2	1700	[14]
Cefluprenam	rOCTN2	ND	[14]
	hOCTN2	ND	[14]

Compound	Transporter[a]	K_i or (K_m)[b] (μM)	Reference
Cefoselis	rOCTN2	6400	[14]
	hOCTN2	6400	[14]
Cefsulodin	hOCTN2	ND	[62]
Ceftazidime	hOCTN2	ND	[62]
Cephaloridine	hOCTN1	ND	[44]
	rOCTN2	790	[14]
	hOCTN2	ND, 230	[14, 62]
Choline	rOCT1	(1100), (346), 400	[39, 65, 66]
	hOCT1	ND	[32]
	rOCT2	159, (441)	[39, 66]
	hOCT2	(210), (102)	[8, 39]
	rOCT3	ND	[10, 21]
	rOCTN2	ND	[27]
	hOCTN2	ND	[46]
Cimetidine	rOCT1	5.7, 329, ND	[11, 66, 67]
	rOCT1A	ND	[9]
	rbOCT1	ND	[6]
	hOCT1	166	[32]
	rOCT2	9.4, 198, 373, (21)	[11, 41, 66, 67]
	mOCT3	ND	[20]
	rOCT3	ND	[10, 21]
	hOCT3	ND	[67]
	rOCTN1	1540	[25]
	hOCTN1	ND	[44]
	rOCTN2	ND	[27]
	hOCTN2	ND	[10, 27, 62]
Cisplatin	rOCT2	925	[41]
ClAdo	rOCT1	ND	[40]
Clonidine	rOCT1	1.4	[68]
	hOCT1	0.55	[32]

Compound	Transporter[a]	K_i or (K_m)[b] (μM)	Reference
	mOCT3	ND	[20]
	rOCT3	ND	[10]
	hOCT3	373	[20]
	rOCTN2	ND	[27]
	hOCTN2	ND	[62]
Corticosterone	rOCT1	10, 72	[4, 68]
	hOCT1	7.0	[32]
	pOCT2	0.67	[15]
	rOCT2	4.2, 0.5	[10, 17]
	rOCT3	4.9	[10]
	hOCT3	0.12	[22]
	hOCTN2	ND	[62]
Creatinine	hOCT1	ND	[32]
	rOCT2	ND	[67]
Cyanine-863	rOCT1	0.13, 0.67	[4, 68]
	pOCT2	0.5	[15]
	hOCT2	0.21	[8]
Cysteine	hOCTN2	ND	[46]
Decynium-22	rOCT1	0.36, 22, 0.5	[4, 66, 68]
	hOCT1	4.4, 2.7	[7, 32]
	rOCT2	14	[66]
	pOCT2	0.05	[15]
	hOCT2	0.10	[8]
Deoxycorticosterone	rOCT2	1.9	[10]
	rOCT3	8.4	[10]
Desipramine	rOCT1	2.8	[4]
	hOCT1	5.3	[32]
	hOCT2	16	[8]
	rOCT3	68	[10, 21]
	hOCT3	14	[20]

Compound	Transporter[a]	K_i or (K_m)[b] (μM)	Reference
	rOCTN1	80	[25]
	rOCTN2	ND	[27]
	hOCTN2	ND	[27]
Disopyramide	hOCT1	ND	[32]
Disprocynium24	rOCT1	0.11	[68]
	rOCT2	0.01	[17]
	hOCT3	0.01	[22]
DMA	mOCT3	ND	[20]
	rOCT3	ND	[10, 21]
	rOCTN1	180	[25]
	rOCTN2	ND	[27]
dTub	rOCT1	(10)	[40]
Dopamine	rOCT1	(1100), (51)	[69, 70]
	hOCT1	ND	[32]
	rOCT2	2300, (2100)	[10, 17]
	hOCT2	(390), (330)	[19]
	rOCT3	620	[10]
Emetine	hOCTN2	4.2	[46]
Beta-estradiol	rOCT2	84.8	[10]
	rOCT3	1.1	[10]
Ganciclovir (GCV)	hOCT1	(516)	[63]
Guanidine	rOCT1	724, 4200	[11, 67]
	rOCT1A	ND	[9]
	rOCT2	714, (730)	[11, 67]
	mOCT3	ND	[20]
	rOCT3	ND	[10, 21]
	hOCT3	6200, 13,000	[20, 67]
Histamine	rOCT1	1400	[67]
	rOCT2	(540)	[67]
	hOCT2	(1300)	[19]

Compound	Transporter[a]	K_i or $(K_m)^b$ (μM)	Reference
	hOCT3	(180)	[67]
Imipramine	hOCT3	42	[20]
	rOCTN1	ND	[25]
Indinavir	hOCT1	62	[71]
Levofloxacin	rOCT2	ND	[66]
Lysine	hOCTN2	ND	[46]
Memantine	hOCT2	(34)	[19]
Mepiperphenidol	rOCT1	5.2	[4]
	hOCT2	1.8	[8]
Methamphetamine	rOCT3	247	[10]
Methionine	hOCTN2	ND	[46]
Midazolam	hOCT1	3.7	[32]
MPTP	mOCT3	ND	[20]
	rOCT3	ND	[21]
	rOCTN2	ND	[27]
	hOCTN2	ND	[27, 29]
MPP^+	rOCT1	13, 0.8, (10), 64, (13)	[4, 11, 65, 66, 68]
	rbOCT1	(23)	[6]
	hOCT1	(15), 12	[7, 32]
	rOCT2	1.8, 44	[11, 66]
	hOCT2	(19), 2.4; (16)	[8, 19]
	mOCT3	ND	[20]
	rOCT3	(91), 143; ND	[10, 21]
	hOCT3	ND; 54, (47)	[20, 22]
	rOCTN1	ND	[25]
	rOCTN2	ND	[27]
	hOCTN2	ND	[29, 62]
Nelfinavir	hOCT1	22	[71]
Nicotine	rOCT1	64.3	[11]

Compound	Transporter[a]	K_i or (K_m)[b] (μM)	Reference
	rOCT2	50.5, ND	[11, 66]
	rOCT3	ND	[21]
	hOCT1	ND	[32]
	rOCTN1	ND	[25]
	hOCTN1	ND	[44]
	rOCTN2	ND	[27]
	hOCTN2	ND	[29, 62]
NMN	rOCT1	1000, (340), 2400, 670	[4, 11, 65, 72]
	rbOCT1	ND	[6]
	hOCT1	7700	[32]
	rOCT2	400, 1600	[11, 66]
	hOCT2	(300)	[8]
	mOCT3	ND	[20]
	rOCT3	ND	[10, 21]
N-methyl-quinidine	rOCT1	(7)	[31]
	hOCT1	(12)	[31]
N-methyl-quinine	rOCT1	(17)	[31]
	hOCT1	(20)	[31]
Noradrenaline	rOCT1	(2800)	[69]
	rOCT2	(4400)	[17]
	hOCT3	(510)	[22]
Norepinephrine	rOCT2	11,000	[10]
	hOCT2	(1900)	[19]
	rOCT3	434	[10]
O-methylisoprenaline	rOCT1	43, 25	[4, 68]
	pOCT2	880	[15]
	hOCT2	570	[8]
Pancuronium	hOCT1	ND	[7]
PGE_2	hOCT1	(657)	[30]

Compound	Transporter[a]	K_i or (K_m)[b] (μM)	Reference
	hOCT2	(29)	[30]
$PGF_{2\alpha}$	hOCT1	(477)	[30]
	hOCT2	(334)	[30]
Procainamide	rOCT1	13, 44.4	[4, 66]
	rOCT1A	ND	[9]
	hOCT1	73.9, 107	[32, 33]
	rOCT2	257	[66]
	hOCT2	50	[8]
	hOCT3	738	[20]
	rOCTN1	860	[25]
	hOCTN1	ND	[44]
	rOCTN2	ND	[27]
	hOCTN2	ND	[29, 62]
Progesterone	rOCT2	1.6	[10]
	rOCT3	10.5	[10]
Pyrilamine	hOCTN1	ND	[44]
	hOCTN2	ND	[62]
Quinidine	rOCT1	14.6, 6.0	[11, 68]
	hOCT1	17.5, 23.4	[32, 33]
	rOCT2	19.1, ND	[11, 66]
	hOCTN1	ND	[44]
	hOCTN2	ND	[46, 62]
Quinine	rOCT1	0.93, 4.3	[4, 68]
	hOCT1	22.9, 22.6	[32, 33]
	hOCT2	3.4	[8]
	pOCT2	5.5	[15]
	hOCTN1	ND	[44]
	hOCTN2	ND	[62]
Reserpine	rOCT1	20	[4]
Ritonavir	hOCT1	5.2	[71]

Compound	Transporter[a]	K_i or (K_m)[b] (μM)	Reference
Saquinavir	hOCT1	8.3	[71]
Serotonin	rOCT1	(650)	[69]
	rOCT2	(3600)	[17]
	hOCT2	(80)	[19]
	rOCT3	970, ND	[10, 21]
TBA	mOCT1	7.3	[34]
	rOCT1	17	[34]
	rbOCT1	25	[34]
	hOCT1	52, 30	[33, 34]
TBuMA	hOCT1	66, (53)	[31, 33]
		(34)	[31]
TEA	mOCT1	128	[34]
	rOCT1	(95); 47, (36); 100, 167	[4, 11, 34, 66]
	rbOCT1	ND, 94	[6, 34]
	hOCT1	161, 260, 158	[32-34]
	rOCT2	52, (45); (34), 142	[11, 66]
	pOCT2	(20), 38	[15]
	hOCT2	(76)	[8]
	mOCT3	(1900)	[20]
	rOCT3	(2500)	[21]
	hOCT3	1300	[20]
	rOCTN1	960	[25]
	hOCTN1	(436)	[26]
	mOCTN2	ND	[27]
	rOCTN2	(63)	[27]
	hOCTN2	ND	[27, 29, 46, 62]
Testosterone	rOCT3	ND	[10]
THA	hOCT1	3.0	[33]
	rOCTN2	ND	[27]

Compound	Transporter[a]	K_i or (K_m)[b] (μM)	Reference
TMA	mOCT1	2040	[34]
	rOCT1	1000, 905	[4, 34]
	rbOCT1	5800	[34]
	hOCT1	10,000, 12,000	[33, 34]
	hOCT2	180	[8]
	rOCTN2	ND	[27]
TPeA	mOCT1	ND	[34]
	rOCT1	0.43, ND	[4, 34]
	rbOCT1	ND	[34]
	hOCT1	7.4, 8.6, ND	[32-34]
	hOCT2	1.5	[8]
TPrA	mOCT1	20	[34]
	rOCT1	21	[34]
	rbOCT1	36	[34]
	hOCT1	90, 102	[33, 34]
Tyramine	hOCT3	ND	[22]
	hOCTN2	ND	[62]
Verapamil	hOCT1	2.9	[32]
	rOCTN1	ND	[25]
	hOCTN1	ND	[44]
	hOCTN2	ND	[27, 46, 62]
Vecuronium	hOCT1	120, 232, 237	[7, 32, 33]

[a] Underlining indicates that the compound is a substrate. Otherwise, the compound has only been shown to inhibit the transport of a model compound (*i.e.,* it is not known whether the drug is a substrate or only an inhibitor).

[b] Values in parentheses in the K_i/K_m column indicate a K_m value. Otherwise it is a K_i value. All of the available kinetic constants are reported from the primary literature. ND indicates that a kinetic parameter has not been determined, but the compound has been shown to interact with the transporter.

Table 3: Summary of Human Mutations in *OCTN2*

Polymorphism	Patient Ethnicity[a]	Effect of Polymorphism	Reference
Polymorphisms Abolishing Function			
1-176 del	Japanese (1)	Truncated protein; non-functional	[3]
226insC	Japanese (1)	Truncated protein; non-functional	[3]
W132X	ND (1), Chinese (1)	Truncated protein; non-functional	[3, 58]
IVS8-1G→A	ND (1)	Truncated protein; non-functional	[3]
R282X	East Indian (1), German (3)	Truncated protein; non-functional	[55], [54], [56]
Y401X	ND (1)	Truncated protein; non-functional	[55]
458X	ND (1)	Truncated protein; non-functional	[55]
P478L	Chinese (1)	No carnitine uptake	[58]
R169W	Italian (1) & ND(1)	No carnitine uptake	[57]
G242V	Spanish (1)	No carnitine uptake	[57]
W351R	ND (1)	No carnitine uptake	[57]
R169Q	German (1)	No carnitine uptake	[54]
Y211C	Moroccan (1), Cape Verde (1)	No carnitine uptake	[56]
Polymorphisms with Reduced Function			
E452K	Pakistani (1)	↓ V_{max} due to impaired Na^+ stimulation (3% of control)	[61]
S467C	Japanese (7)	↓ L-carnitine uptake (11% of control)	[59]
W283C	Japanese (1)	↓ L-carnitine uptake (2% of control)	[59]
M179L	Japanese (1)	↓ L-carnitine uptake (74% of control)	[59]
W283R	Caucasian (1)	↓ L-carnitine uptake (2% of control)	[60]
V446F	Caucasian (1)	↓ L-carnitine uptake (2% of control)	[60]
A301D	Italian (1)	↓ L-carnitine uptake (2-3% of control)	[57]

[a] Numbers in parentheses specify the number of individuals in which the polymorphism was identified. ND, not determined.

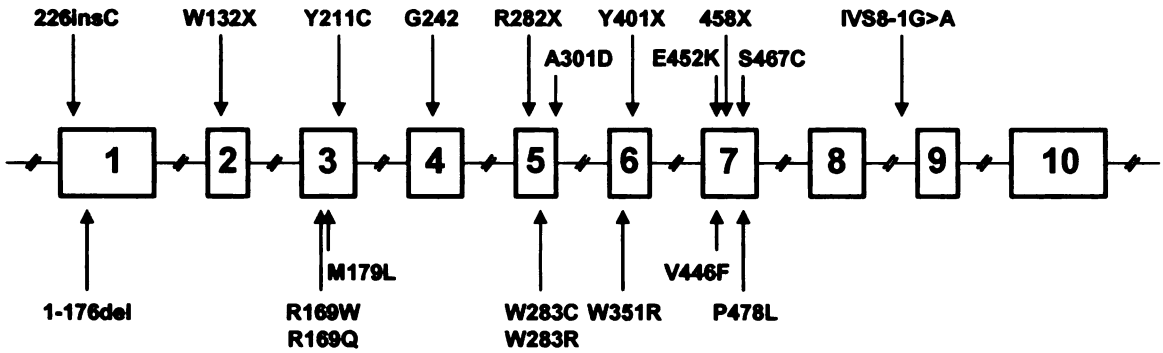


Figure 3. Genomic location of OCTN2 polymorphisms. Numbered rectangles represent exons of OCTN2 and joining lines depict introns. Arrows denote the location of each polymorphism in a particular exon or intron.

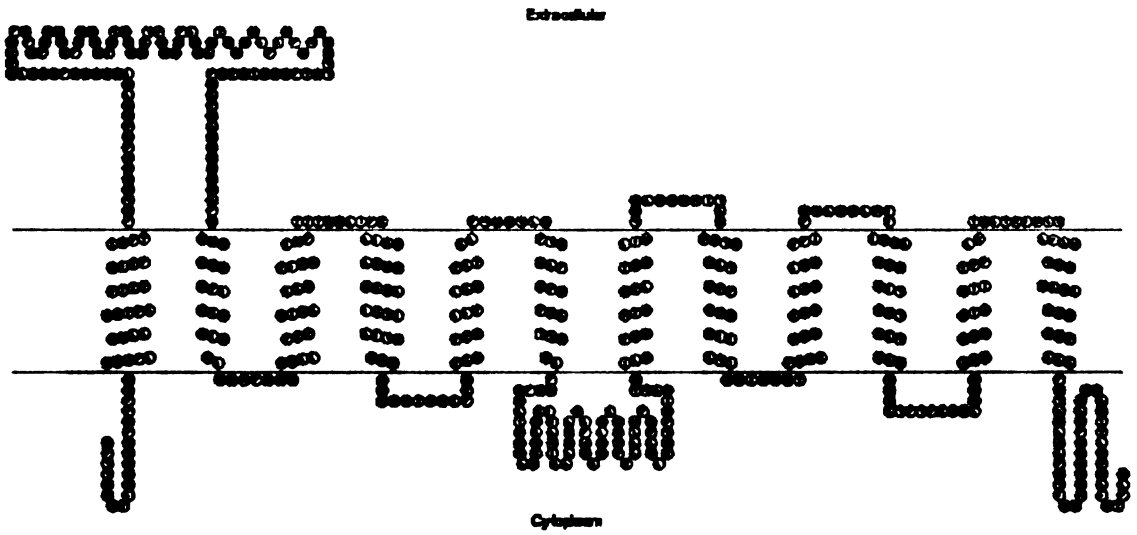


Figure 4. Representative secondary structure of an organic cation transporter (OCT2)

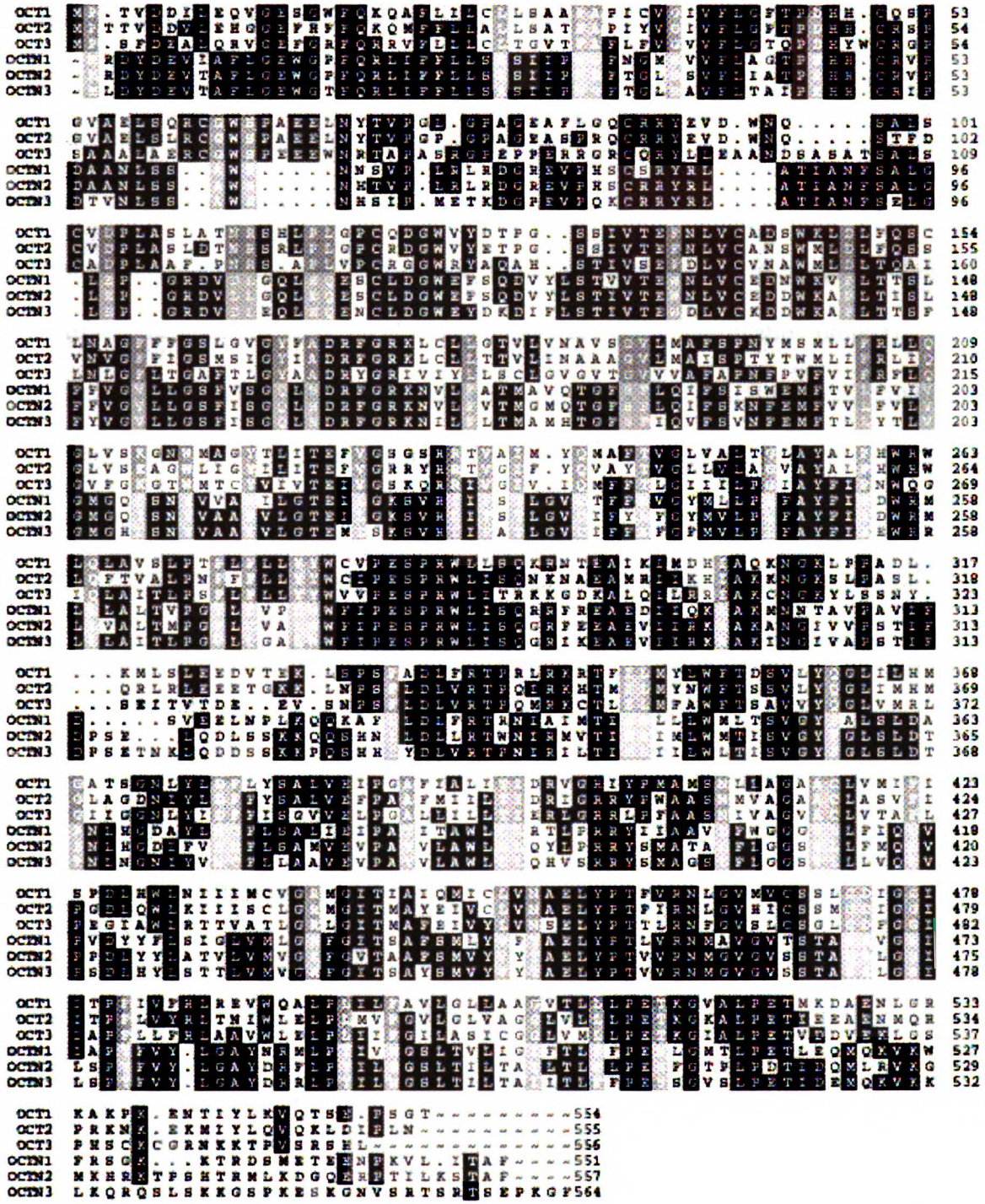
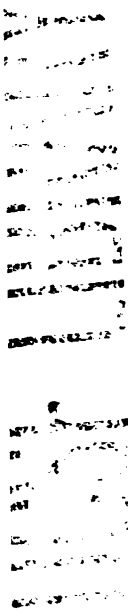


Figure 5. Alignments of OCTs and OCTNs. OCT1-OCT3 and OCTN1 and OCTN2 are human sequences; OCTN3 is the mouse sequence. Black shading indicates residues conserved in at least three of the six sequences. Amino acid residues conserved only in the OCTs or OCTNs are shaded. In addition, residues with similar chemical properties to conserved residues are shaded light gray.

SUMMARY OF DISSERTATION STUDIES

The cDNAs for six organic cation transporter isoforms - OCT1, OCT2, OCT3, OCTN1, OCTN2, and OCTN3 - have been cloned and the transporter proteins have been functionally characterized. *In vitro* molecular studies, and in some cases, studies in animal models and/or humans, have helped define the physiological and pharmacological roles of this transporter family. With many large-scale SNP identification projects underway, we are now learning about the types and levels of genetic variation throughout the human genome. Many studies have associated genetic variation in drug-metabolizing enzymes with drug response phenotypes. In contrast, little is known about genetic variation in membrane transporter genes. The major goal of this dissertation was to identify genetic variants of membrane transporters with roles in drug response and elucidate the consequences of this genetic variation on transporter function.



CHAPTER 2: POLYMORPHISMS IN A HUMAN KIDNEY XENOBIOTIC TRANSPORTER, OCT2, EXHIBIT ALTERED FUNCTION

The goal of this chapter was to identify genetic variants of the human renal organic cation transporter, hOCT2, and determine the significance of common variants to transporter function. This chapter provides a comprehensive genetic analysis and functional characterization of variants of OCT2 *in vitro*. All exons of *OCT2*, as well as 50-100 base pairs of flanking intronic sequence, were screened using denaturing HPLC analysis followed by direct sequencing. Twenty-eight variable sites were identified in a collection of 247 ethnically diverse DNA samples. Eight caused non-synonymous amino acid changes, of which four were present at $\geq 1\%$ in an ethnic population. All four of

these altered transporter function assayed in *Xenopus laevis* oocytes. Analysis of nucleotide diversity (π) revealed a higher prevalence of synonymous ($\pi = 22.4 \times 10^{-4}$) versus non-synonymous ($\pi = 2.1 \times 10^{-4}$) changes in OCT2 than in other genes. In addition, the non-synonymous sites had a significant tendency to exhibit more skewed allele frequencies (more negative Tajima's *D* values) compared to synonymous sites. The population-genetic analysis together with the functional characterization suggests that selection has acted against amino acid changes in OCT2. Such selection may be due to a necessary role of OCT2 in the renal elimination of endogenous amines or xenobiotics, including environmental toxins, neurotoxic amines, and therapeutic drugs.

CHAPTER 3: ESTIMATING THE GENETIC COMPONENT CONTRIBUTING TO VARIATION IN RENAL CLEARANCE

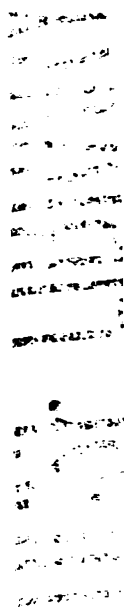
Recent studies in pharmacogenetics have examined the effects of genetic variation in metabolizing enzymes on interindividual variation in drug response. In contrast, little is known about the genetic factors contributing to interindividual variation in renal elimination, the other major pathway of detoxification. Renal excretion is the major pathway for elimination of many clinically used drugs and is the exclusive pathway for eliminating virtually all of the end-products of drug-metabolizing enzymes. In this chapter, we examined the heritability of renal elimination using a method recently proposed by Kalow *et al.* in order to better understand the relative contributions of genetic and environmental factors on renal elimination. We calculated the heritability of renal clearance of five drugs spanning a wide range of renal clearance values; these included terodiline, a drug that undergoes passive reabsorption, metformin, amoxicillin

and ampicillin, which undergo net secretion, and PAH, a drug whose renal clearance is limited by renal blood flow. Using a bootstrap analysis, we showed that heritability of renal clearance of drugs undergoing net secretion is significantly higher than drugs with renal clearance determined by other factors (passive reabsorption and renal blood flow). Furthermore, heritability of renal clearance of net secreted drugs is similar to the heritability of metabolic clearance. These studies suggest that genetic factors may play a significant role in contributing to variation in the net secretion of renally cleared drugs.

CHAPTER 4: NATURAL VARIATION IN MEMBRANE TRANSPORTER GENES REVEALS EVOLUTIONARY CONSTRAINTS

Membrane transporters play an important role in drug response as they provide the targets for many therapeutic drugs and are major determinants of drug absorption, distribution, and elimination. As part of a large pharmacogenetics project, we screened for variation in a set of 24 membrane transporter genes in 247 DNA samples (494 chromosomes) from ethnically diverse populations. We observed that Caucasians had an excess of high frequency alleles in comparison to African Americans, consistent with a historic bottleneck. By screening a functional class of proteins with similar secondary structures, characterized by multiple transmembrane domains (TMDs) joined by loops, we obtained information on variation in distinct regions of these proteins. We observed that amino acid diversity ($\pi_{\text{non-synonymous}}$) in the TMDs was significantly lower than amino acid diversity in the loops. Non-synonymous SNPs at evolutionarily conserved sites were more prevalent at low allele frequencies (< 1%) compared to non-synonymous SNPs at evolutionarily unconserved sites, while the reverse was true at high allele

frequencies (> 10%). These data imply that variation in amino acid sequence at evolutionarily conserved residues and residues within the TMDs are maintained at low allele frequencies due to purifying selection. This study demonstrates the use of human population genetic analysis, in combination with phylogenetic comparisons, to learn about protein structure and function.



REFERENCES

1. Bendayan R. Renal drug transport: a review. *Pharmacother* **16**(6): 971-985, 1996.
2. Zhang L, Brett CM and Giacomini KM. Role of organic cation transporters in drug absorption and elimination. *Annu Rev Pharmacol Toxicol* **38:** 431-460, 1998.
3. Nezu J, Tamai I, Oku A, Ohashi R, Yabuuchi H, Hashimoto N, Nikaido H, Sai Y, Koizumi A, Shoji Y, Takada G, Matsuishi T, Yoshino M, Kato H, Ohura T, Tsujimoto G, Hayakawa J, Shimane M and Tsuji A. Primary systemic carnitine deficiency is caused by mutations in a gene encoding sodium ion-dependent carnitine transporter. *Nat Genet* **21**(1): 91-4, 1999.
4. Grundemann D, Gorboulev V, Gambaryan S, Veyhl M and Koepsell H. Drug excretion mediated by a new prototype of polyspecific transporter. *Nature* **372**(6506): 549-552, 1994.
5. Schweifer N and Barlow DP. The Lx1 gene maps to mouse chromosome 17 and codes for a protein that is homologous to glucose and polyspecific transmembrane transporters. *Mamm Genome* **7**(10): 735-740, 1996.
6. Terashita S, Dresser MJ, Zhang L, Gray AT, Yost SC and Giacomini KM. Molecular cloning and functional expression of a rabbit renal organic cation transporter. *Biochim Biophys Acta* **1369**(1): 1-6, 1998.
7. Zhang L, Dresser MJ, Gray AT, Yost SC, Terashita S and Giacomini KM. Cloning and functional expression of a human liver organic cation transporter. *Mol Pharmacol* **51**(6): 913-921, 1997.

8. Gorboulev V, Ulzheimer JC, Akhoundova A, Ulzheimer-Teuber I, Karbach U, Quester S, Baumann C, Lang F, Busch AE and Koepsell H. Cloning and characterization of two human polyspecific organic cation transporters. *DNA Cell Biol* **16**(7): 871-881, 1997.

9. Zhang L, Dresser MJ, Chun JK, Babbitt PC and Giacomini KM. Cloning and functional characterization of a rat renal organic cation transporter isoform (rOCT1A). *J Biol Chem* **272**(26): 16548-16554, 1997.

10. Wu X, Kekuda R, Huang W, Fei YJ, Leibach FH, Chen J, Conway SJ and Ganapathy V. Identity of the organic cation transporter OCT3 as the extraneuronal monoamine transporter (uptake2) and evidence for the expression of the transporter in the brain. *J Biol Chem* **273**(49): 32776-32786, 1998.

11. Urakami Y, Okuda M, Masuda S, Saito H and Inui KI. Functional characteristics and membrane localization of rat multispecific organic cation transporters, OCT1 and OCT2, mediating tubular secretion of cationic drugs. *J Pharmacol Exp Ther* **287**(2): 800-805, 1998.

12. Meyer-Wentrup F, Karbach U, Gorboulev V, Arndt P and Koepsell H. Membrane localization of the electrogenic cation transporter rOCT1 in rat liver. *Biochem Biophys Res Commun* **248**(3): 673-678, 1998.

13. Karbach U, Kricke J, Meyer-Wentrup F, Gorboulev V, Volk C, Loffing-Cueni D, Kaissling B, Bachmann S and Koepsell H. Localization of organic cation transporters OCT1 and OCT2 in rat kidney. *Am J Physiol Renal Physiol* **279**(4): F679-87, 2000.

14. Ganapathy ME, Huang W, Rajan DP, Carter AL, Sugawara M, Iseki K, Leibach FH and Ganapathy V. beta-lactam antibiotics as substrates for OCTN2, an organic cation/carnitine transporter. *J Biol Chem* **275**(3): 1699-707, 2000.

15. Grundemann D, Babin-Ebell J, Martel F, Ording N, Schmidt A and Schomig E. Primary structure and functional expression of the apical organic cation transporter from kidney epithelial LLC-PK1 cells. *J Biol Chem* **272**(16): 10408-10413, 1997.

16. Mooslehner KA and Allen ND. Cloning of the mouse organic cation transporter 2 gene, Slc22a2, from an enhancer-trap transgene integration locus. *Mamm Genome* **10**(3): 218-224, 1999.

17. Grundemann D, Koster S, Kiefer N, Breidert T, Engelhardt M, Spitzenberger F, Obermuller N and Schomig E. Transport of monoamine transmitters by the organic cation transporter type 2, OCT2. *J Biol Chem* **273**(47): 30915-30920, 1998.

18. Okuda M, Saito H, Urakami Y, Takano M and Inui K. cDNA cloning and functional expression of a novel rat kidney organic cation transporter, OCT2. *Biochem Biophys Res Commun* **224**(2): 500-507, 1996.

19. Busch AE, Karbach U, Miska D, Gorboulev V, Akhoundova A, Volk C, Arndt P, Ulzheimer JC, Sonders MS, Baumann C, Waldegger S, Lang F and Koepsell H. Human neurons express the polyspecific cation transporter hOCT2, which translocates monoamine neurotransmitters, amantadine, and memantine. *Mol Pharmacol* **54**(2): 342-52, 1998.

20. Wu X, Huang W, Ganapathy ME, Wang H, Kekuda R, Conway SJ, Leibach FH and Ganapathy V. Structure, function, and regional distribution of the organic cation transporter OCT3 in the kidney. *Am J Physiol* **279**(3): F449-F458, 2000.

21. Kekuda R, Prasad PD, Wu X, Wang H, Fei YJ, Leibach FH and Ganapathy V. Cloning and functional characterization of a potential-sensitive, polyspecific organic cation transporter (OCT3) most abundantly expressed in placenta. *J Biol Chem* **273**(26): 15971-15979, 1998.

22. Grundemann D, Schechinger B, Rappold GA and Schomig E. Molecular identification of the corticosterone-sensitive extraneuronal catecholamine transporter. *Nat Neurosci* **1**(5): 349-351, 1998.

23. Verhaagh S, Schweifer N, Barlow DP and Zwart R. Cloning of the mouse and human solute carrier 22a3 (Slc22a3/SLC22A3) identifies a conserved cluster of three organic cation transporters on mouse chromosome 17 and human 6q26-q27. *Genomics* **55**(2): 209-218, 1999.

24. Tamai I, Ohashi R, Nezu JI, Sai Y, Kobayashi D, Oku A, Shimane M and Tsuji A. Molecular and functional characterization of organic cation/carnitine transporter family in mice. *J Biol Chem* **275**(51): 40064-72, 2000.

25. Wu X, George RL, Huang W, Wang H, Conway SJ, Leibach FH and Ganapathy V. Structural and functional characteristics and tissue distribution pattern of rat OCTN1, an organic cation transporter, cloned from placenta. *Biochim Biophys Acta* **1466**(1-2): 315-27, 2000.

26. Tamai I, Yabuuchi H, Nezu J, Sai Y, Oku A, Shimane M and Tsuji A. Cloning and characterization of a novel human pH-dependent organic cation transporter, OCTN1. *FEBS Lett* **419**(1): 107-111, 1997.

27. Wu X, Huang W, Prasad PD, Seth P, Rajan DP, Leibach FH, Chen J, Conway SJ and Ganapathy V. Functional characteristics and tissue distribution pattern of organic cation transporter 2 (OCTN2), an organic cation/carnitine transporter. *J Pharmacol Exp Ther* **290**(3): 1482-1492, 1999.

28. Tamai I, Ohashi R, Nezu J, Yabuuchi H, Oku A, Shimane M, Sai Y and Tsuji A. Molecular and functional identification of sodium ion-dependent, high affinity human carnitine transporter OCTN2. *J Biol Chem* **273**(32): 20378-20382, 1998.

29. Wu X, Prasad PD, Leibach FH and Ganapathy V. cDNA sequence, transport function, and genomic organization of human OCTN2, a new member of the organic cation transporter family. *Biochem Biophys Res Commun* **246**(3): 589-595, 1998.

30. Kimura H, Takeda M, Narikawa S, Enomoto A, Ichida K and Endou H. Human organic anion transporters and human organic cation transporters mediate renal transport of prostaglandins. *J Pharmacol Exp Ther* **301**(1): 293-8, 2002.

31. van Montfoort JE, Muller M, Groothuis GM, Meijer DK, Koepsell H and Meier PJ. Comparison of "type I" and "type II" organic cation transport by organic cation transporters and organic anion-transporting polypeptides. *J Pharmacol Exp Ther* **298**(1): 110-5, 2001.

32. Zhang L, Schaner ME and Giacomini KM. Functional characterization of an organic cation transporter (hOCT1) in a transiently transfected human cell line (HeLa). *J Pharmacol Exp Ther* **286**(1): 354-361, 1998.

33. Zhang L, Gorset W, Dresser MJ and Giacomini KM. The interaction of n-tetraalkylammonium compounds with a human organic cation transporter, hOCT1. *J Pharmacol Exp Ther* **288**(3): 1192-8, 1999.

34. Dresser MJ, Gray AT and Giacomini KM. Kinetic and selectivity differences between rodent, rabbit, and human organic cation transporters (OCT1). *J Pharmacol Exp Ther* **292**(3): 1146-52, 2000.

35. Sugawara-Yokoo M, Urakami Y, Koyama H, Fujikura K, Masuda S, Saito H, Naruse T, Inui K and Takata K. Differential localization of organic cation transporters rOCT1 and rOCT2 in the basolateral membrane of rat kidney proximal tubules. *Histochem Cell Biol* **114**(3): 175-80, 2000.

36. Koepsell H, Gorboulev V and Arndt P. Molecular pharmacology of organic cation transporters in kidney. *J Membr Biol* **167**(2): 103-17, 1999.

37. Sweet DH and Pritchard JB. rOCT2 is a basolateral potential-driven carrier, not an organic cation/proton exchanger. *Am J Physiol* **277**(6 Pt 2): F890-F898, 1999.

38. Sweet DH, Miller DS and Pritchard JB. Basolateral localization of organic cation transporter 2 in intact renal proximal tubules. *Am J Physiol Renal Physiol* **279**(5): F826-34, 2000.

39. Sweet DH, Miller DS and Pritchard JB. Ventricular choline transport: a role for organic cation transporter 2 expressed in choroid plexus. *J Biol Chem* **276**(45): 41611-9, 2001.

40. Chen R and Nelson JA. Role of organic cation transporters in the renal secretion of nucleosides. *Biochem Pharmacol* **60**(2): 215-9, 2000.

41. Pan BF, Sweet DH, Pritchard JB, Chen R and Nelson JA. A transfected cell model for the renal toxin transporter, rOCT2. *Toxicol Sci* **47**(2): 181-186, 1999.

42. Urakami Y, Okuda M, Saito H and Inui K. Hormonal regulation of organic cation transporter OCT2 expression in rat kidney. *FEBS Lett* **473**(2): 173-6, 2000.

43. Shu Y, Bello CL, Mangravite LM, Feng B and Giacomini KM. Functional characteristics and steroid hormone-mediated regulation of an organic cation transporter in Madin-Darby canine kidney cells. *J Pharmacol Exp Ther* **299**(1): 392-8, 2001.

44. Yabuuchi H, Tamai I, Nezu J, Sakamoto K, Oku A, Shimane M, Sai Y and Tsuji A. Novel membrane transporter OCTN1 mediates multispecific, bidirectional, and pH-dependent transport of organic cations. *J Pharmacol Exp Ther* **289**(2): 768-773, 1999.

45. Brooks H and Krahenbuhl S. Identification and tissue distribution of two differentially spliced variants of the rat carnitine transporter OCTN2. *FEBS Lett* **508**(2): 175-80, 2001.

46. Wagner CA, Lukewille U, Kaltenbach S, Moschen I, Broer A, Risler T, Broer S and Lang F. Functional and pharmacological characterization of human Na(+)-carnitine cotransporter hOCTN2. *Am J Physiol* **279**(3): F584-F591, 2000.

47. Lee J, Prohaska JR and Thiele DJ. Essential role for mammalian copper transporter Ctr1 in copper homeostasis and embryonic development. *Proc Natl Acad Sci U S A* **98**(12): 6842-7, 2001.

48. Kuo YM, Zhou B, Cosco D and Gitschier J. The copper transporter CTR1 provides an essential function in mammalian embryonic development. *Proc Natl Acad Sci U S A* **98**(12): 6836-41, 2001.

49. Schinkel AH, Smit JJ, van Tellingen O, Beijnen JH, Wagenaar E, van Deemter L, Mol CA, van der Valk MA, Robanus-Maandag EC, te Riele HP and et al. Disruption of the mouse mdr1a P-glycoprotein gene leads to a deficiency in the blood-brain barrier and to increased sensitivity to drugs. *Cell* **77**(4): 491-502, 1994.

50. Jonker JW, Wagenaar E, Mol CA, Buitelaar M, Koepsell H, Smit JW and Schinkel AH. Reduced hepatic uptake and intestinal excretion of organic cations in mice with a targeted disruption of the organic cation transporter 1 (Oct1 [Slc22a1]) gene. *Mol Cell Biol* **21**(16): 5471-7, 2001.

51. Zwart R, Verhaagh S, Buitelaar M, Popp-Snijders C and Barlow DP. Impaired activity of the extraneuronal monoamine transporter system known as uptake-2 in Orct3/Slc22a3-deficient mice. *Mol Cell Biol* **21**(13): 4188-96, 2001.

52. Ohashi R, Tamai I, Nezu Ji J, Nikaido H, Hashimoto N, Oku A, Sai Y, Shimane M and Tsuji A. Molecular and physiological evidence for multifunctionality of carnitine/organic cation transporter OCTN2. *Mol Pharmacol* **59**(2): 358-66, 2001.

53. Zhu Y, Jong MC, Frazer KA, Gong E, Krauss RM, Cheng JF, Boffelli D and Rubin EM. Genomic interval engineering of mice identifies a novel modulator of triglyceride production. *Proc Natl Acad Sci U S A* **97**(3): 1137-42, 2000.

54. Burwinkel B, Kreuder J, Schweitzer S, Vorgerd M, Gempel K, Gerbitz KD and Kilimann MW. Carnitine transporter OCTN2 mutations in systemic primary

carnitine deficiency: a novel Arg169Gln mutation and a recurrent Arg282ter mutation associated with an unconventional splicing abnormality. *Biochem Biophys Res Commun* **261**(2): 484-7, 1999.

55. Wang Y, Ye J, Ganapathy V and Longo N. Mutations in the organic cation/carnitine transporter OCTN2 in primary carnitine deficiency. *Proc Natl Acad Sci U S A* **96**(5): 2356-60, 1999.

56. Vaz FM, Scholte HR, Ruiter J, Hussaarts-Odijk LM, Pereira RR, Schweitzer S, de Klerk JB, Waterham HR and Wanders RJ. Identification of two novel mutations in OCTN2 of three patients with systemic carnitine deficiency. *Hum Genet* **105**(1-2): 157-61, 1999.

57. Wang Y, Taroni F, Garavaglia B and Longo N. Functional analysis of mutations in the OCTN2 transporter causing primary carnitine deficiency: lack of genotype-phenotype correlation. *Hum Mutat* **16**(5): 401-7, 2000.

58. Tang NL, Ganapathy V, Wu X, Hui J, Seth P, Yuen PM, Wanders RJ, Fok TF and Hjelm NM. Mutations of OCTN2, an organic cation/carnitine transporter, lead to deficient cellular carnitine uptake in primary carnitine deficiency. *Hum Mol Genet* **8**(4): 655-60, 1999.

59. Koizumi A, Nozaki J, Ohura T, Kayo T, Wada Y, Nezu J, Ohashi R, Tamai I, Shoji Y, Takada G, Kibira S, Matsuishi T and Tsuji A. Genetic epidemiology of the carnitine transporter OCTN2 gene in a Japanese population and phenotypic characterization in Japanese pedigrees with primary systemic carnitine deficiency. *Hum Mol Genet* **8**(12): 2247-54, 1999.

60. Mayatepek E, Nezu J, Tamai I, Oku A, Katsura M, Shimane M and Tsuji A. Two novel missense mutations of the OCTN2 gene (W283R and V446F) in a patient with primary systemic carnitine deficiency. *Hum Mutat* **15**(1): 118, 2000.

61. Wang Y, Kelly MA, Cowan TM and Longo N. A missense mutation in the OCTN2 gene associated with residual carnitine transport activity. *Hum Mutat* **15**(3): 238-45, 2000.

62. Ohashi R, Tamai I, Yabuuchi H, Nezu JI, Oku A, Sai Y, Shimane M and Tsuji A. Na(+)-dependent carnitine transport by organic cation transporter (OCTN2): its pharmacological and toxicological relevance. *J Pharmacol Exp Ther* **291**(2): 778-84, 1999.

63. Takeda M, Khamdang S, Narikawa S, Kimura H, Kobayashi Y, Yamamoto T, Cha SH, Sekine T and Endou H. Human organic anion transporters and human organic cation transporters mediate renal antiviral transport. *J Pharmacol Exp Ther* **300**(3): 918-24, 2002.

64. Briz O, Serrano MA, Rebollo N, Hagenbuch B, Meier PJ, Koepsell H and Marin JJ. Carriers involved in targeting the cytostatic bile acid-cisplatin derivatives cis-diammine-chloro-cholylglycinate-platinum(II) and cis-diammine-bisursodeoxycholate-platinum(II) toward liver cells. *Mol Pharmacol* **61**(4): 853-60, 2002.

65. Busch AE, Quester S, Ulzheimer JC, Waldegger S, Gorboulev V, Arndt P, Lang F and Koepsell H. Electrogenic properties and substrate specificity of the polyspecific rat cation transporter rOCT1. *J Biol Chem* **271**(51): 32599-604, 1996.

66. Okuda M, Urakami Y, Saito H and Inui K. Molecular mechanisms of organic cation transport in OCT2-expressing Xenopus oocytes. *Biochim Biophys Acta* **1417**(2): 224-231, 1999.

67. Grundemann D, Liebich G, Kiefer N, Koster S and Schomig E. Selective substrates for non-neuronal monoamine transporters. *Mol Pharmacol* **56**(1): 1-10, 1999.

68. Martel F, Vetter T, Russ H, Grundemann D, Azevedo I, Koepsell H and Schomig E. Transport of small organic cations in the rat liver. The role of the organic cation transporter OCT1. *Naunyn Schmiedebergs Arch Pharmacol* **354**(3): 320-6, 1996.

69. Breidert T, Spitzenberger F, Grundemann D and Schomig E. Catecholamine transport by the organic cation transporter type 1 (OCT1). *Br J Pharmacol* **125**(1): 218-24, 1998.

70. Busch AE, Quester S, Ulzheimer JC, Gorboulev V, Akhoundova A, Waldegger S, Lang F and Koepsell H. Monoamine neurotransmitter transport mediated by the polyspecific cation transporter rOCT1. *FEBS Lett* **395**(2-3): 153-6, 1996.

71. Zhang L, Gorset W, Washington CB, Blaschke TF, Kroetz DL and Giacomini KM. Interactions of HIV protease inhibitors with a human organic cation transporter in a mammalian expression system. *Drug Metab Dispos* **28**(3): 329-34, 2000.

72. Okuda M, Urakami Y, Saito H and Inui K. Molecular mechanisms of organic cation transport in OCT2-expressing Xenopus oocytes. *Biochim Biophys Acta* **1417**(2): 224-31, 1999.

CHAPTER 2

POLYMORPHISMS IN A HUMAN KIDNEY XENOBIOTIC TRANSPORTER, OCT2, EXHIBIT ALTERED FUNCTION*

INTRODUCTION

Like disease susceptibility, response to drugs and sensitivity to xenobiotics is a highly variable, complex trait in which genetics plays a role [1]. Enzymes in the liver and transporters in the kidney are critical in the detoxification and elimination of xenobiotics from the systemic circulation and thus are major determinants of drug response and sensitivity. Recent studies have focused on elucidating the consequences of genetic variation in hepatic metabolizing enzymes on inter-individual differences in therapeutic and adverse drug responses and in sensitivity to environmental toxins. In contrast, nothing is known about the role of genetic variation in the transporters that function in renal elimination, the other major pathway of detoxification.

Renal excretion is the major pathway for elimination of many clinically used drugs and is the exclusive pathway for eliminating virtually all of the end-products of drug-metabolizing enzymes. Transporters in the renal tubule epithelium mediate secretion and thus play a critical role in detoxification [2, 3]. Transporters in the kidney also control the exposure of renal cells to nephrotoxic drugs and environmental toxins and thus determine xenobiotic-induced nephrotoxicity. Inter-individual variation in the renal secretion of many drugs, including anti-arrhythmic and anti-diabetic drugs, as well as in drug-induced nephrotoxicity has been documented [4, 5]. It is hypothesized that this variation is due, in part, to genetic variation in transporters in the renal epithelium.

*This work was published in *Pharmacogenetics* 12(5):395-405, 2002. Permission from the publisher is included in the "Acknowledgements".

The superfamily of transporters, solute carrier family 22 (SLC22) is thought to play a major role in renal secretory transport of xenobiotics. SLC22 includes organic anion transporters (OAT: OAT1 - OAT4) and organic cation transporters (OCT: OCT1 - OCT3 and OCTN: OCTN1 - OCTN3). Members of this superfamily of facilitative transporters share sequence homology and a common 12-transmembrane secondary structure. SLC22 transporters interact with structurally diverse chemical compounds and differ in their charge specificity. In particular, OCTs and OCTNs transport positively charged compounds (and some zwitterions) whereas OATs transport anionic compounds [2, 3].

OCT2 is thought to be the major transporter for uptake of many cationic xenobiotics from the bloodstream into renal epithelial cells. OCT2 interacts with a variety of organic cations, including many clinically used drugs (*e.g.*, procainamide and cimetidine), hormones (*e.g.*, norepinephrine and dopamine), as well as toxic substances (*e.g.*, 1-methyl-4-phenylpyridinium [MPP^+]) [6-12]. The importance of OCT2 in renal elimination of xenobiotics is deduced from the findings that its mRNA is abundant in the kidney and that OCT2 protein is located on the basolateral membrane of the proximal tubule epithelium [6, 9, 13, 14]. Functional studies using isolated human kidney tubules demonstrate their ability to take up organic cations with the specificity of OCT2 [14]. These observations suggest that OCT2 governs the entry of organic cations from the blood into the renal tubule, thereby controlling the first step in renal secretion of organic cations.

In this study, we report a comprehensive genetic and functional analysis of a major human kidney xenobiotic transporter in the SLC22 family, OCT2, in ethnically

diverse populations. Our study had two major goals: to identify variants and variant frequencies in OCT2 and to determine the significance of common variants (polymorphisms) to transporter function.

MATERIALS AND METHODS

Gene sequences and PCR primers

Genomic and cDNA sequences of OCT2 were obtained from GenBank (http://www.ncbi.nlm.nih.gov, accession numbers AC024253, AL162582, and NM003058). The genomic structure of OCT2 was deduced by aligning the genomic and cDNA sequences using the Blast program on the National Center for Biotechnology Information website (http://www.ncbi.nlm.nih.gov/BLAST). All intron-exon boundaries conformed to the consensus donor-acceptor sequence for RNA splicing (GT/AG). Primers were designed manually to span the exons and include ~50 base pairs of flanking intronic sequence at the 5' and 3' ends. Primer sequences are available at http://www.pharmgkb.org or http://www.pharmacogenetics.ucsf.edu.

Variant identification

A collection of 247 ethnically identified genomic DNA samples was obtained from the Coriell Institute of Medicine and used to screen for OCT2 variants. PCR, denaturing HPLC, and sequencing were carried out as previously described [15]. Briefly, optimal reaction and cycle conditions were determined for each amplicon. PCR (25 μl reaction volume) was then carried out using a GeneAmp 9700 thermal cycler (Applied Biosystems, Foster City, CA). Following PCR, samples from two to three individuals

were pooled and heteroduplexed by denaturing and then annealing. Heteroduplexed samples were injected into the WAVE DNA Fragment Analysis System (Transgenomic, Omaha, NE) in a 5-10 μl volume. Column temperatures were designed using the DHPLC Melt Program available at the DNA Variation Group website (http://insertion.stanford.edu/melt.html); gradient conditions were determined using the software provided with the WAVE system. Data were analyzed visually. Samples with clear or suggestive variant peaks were then sequenced. Sequencing was performed by using a dye terminator method with the ABI PRISM BigDye system. Samples were sequenced in the forward and reverse directions with the same primers used for denaturing HPLC and were run on an ABI 377 or ABI 3700 automated sequencer.

Population-genetic analysis

The neutral parameter, θ, was estimated based on the number of segregating sites (S) and was normalized to the number of chromosomes studied (*n*) and the total sequence length (L) as shown in the following formula [16]:

$$\theta = \frac{S}{a_1 L}$$

where

$$a_1 = \sum_{i=1}^{n-1} \frac{1}{i}$$

Nucleotide diversity, π, was estimated by:

$$\pi = [\sum 2p_j(1-p_j)]/(1-1/n),$$

as j increases from 1 to S (*e.g.*, the number of segregating sites) and where p_j is the observed frequency of the j^{th} SNP and n is the number of chromosomes [16]. Tajima's D statistic was calculated as described by Tajima [16]. These three statistics were also estimated separately for synonymous and non-synonymous sites in OCT2 [17].

Haplotypes were determined using an iterative procedure [18]. In short, the algorithm identifies unambiguous haplotypes by searching for complete homozygosity across a region or alternatively, heterozygosity at a single site. Once a set of unambiguous haplotypes is identified, other candidate haplotypes can be called in individuals having more than one heterozygous site by removing the unambiguous haplotype. Once the set of potential haplotype pairs is established for each individual, a maximum likelihood step is applied to determine likelihoods of alternative phases. This approach routinely yields results that differ from more recent Bayesian methods only for the rarest haplotypes [19]. All analysis applied here concerns the unambiguous haplotypes, for which phase identification has high confidence.

Variant construction

Variants were constructed by site-directed mutagenesis. PCR with mutagenic sense and anti-sense primers was carried out using the high fidelity DNA polymerase *PfuTurbo* (Stratagene, La Jolla, CA). PCR cycle conditions were as follows: 95°C for 10 min followed by 14 cycles of 95°C for 30 s, 5°C for 1 min and 68°C for 12 min, and a final hold at 68°C for 12 min. After digesting the parental DNA template with *Dpn*I, the PCR product was transformed into DH5α bacteria, and plasmids containing the variant OCT2

inserts were isolated. The variants were fully sequenced to verify that they had only the desired mutation. Each variant was then subcloned into the pOX expression vector for expression in oocytes.

Expression of OCT2 reference and variants in *Xenopus laevis* oocytes

Transporter cRNA was transcribed *in vitro* using T3 RNA polymerase from templates containing the OCT2 reference or variants inserted into the pOX mammalian expression vector. Oocytes were harvested from *Xenopus laevis* (Xenopus One, Dexter, MI). Oocytes were manually dissected and treated with collagenase D in Ca^{++}-free ORII buffer as previously described [20]. Healthy stage IV and stage V oocytes were sorted and injected with capped cRNA. For dose-response studies, oocytes were injected with increasing amounts of RNA, ranging from 5 to 75 ng. Following injection, the remaining RNA from each dose was run on an RNA gel to verify that none of the RNA was degraded. For all other studies, oocytes were injected with 30-50 ng cRNA. Oocytes were stored in modified Barth's medium at 18°C.

Uptake studies

Transport of [^{3}H]-1-methyl-4-phenylpyridinium (MPP^+) was measured in oocytes 2-4 days after injection as described previously [20]. Briefly, groups of 6 to 9 oocytes were incubated in room-temperature Na^+ buffer containing 1 μM MPP^+ (0.1 μM ^{3}H-MPP^+ and 0.9 μM unlabeled MPP^+). 0.2 μM ^{3}H-MPP^+ and 0.8 μM unlabeled MPP^+ were used in the Michaelis-Menten studies. For kinetic and inhibition studies, unlabeled substrate or inhibitor was added to the reaction mixes as needed. Uptake was stopped after 1 h by

washing the oocytes five times with ice-cold Na^+ buffer. Oocytes were then lysed with 10% SDS and the radioactivity associated with each oocyte counted. To determine apparent K_m and K_i values for MPP^+ and tetrabutylammonium (TBA), respectively, data from the kinetic studies were fit using non-linear regression analysis as previously described [21]. Michaelis-Menten studies with MPP^+ and kinetic inhibition studies with TBA in OCT2 reference, M165I and A270S expressing oocytes were repeated four to five times in different batches of oocytes and data in Table 3 are shown as mean ± standard error from these four to five experiments. All other studies were repeated two to three times in different batches of oocytes.

RESULTS

Variant Identification

To identify variants of OCT2, we screened all 11 exons, as well as 50-100 base pairs of flanking intronic sequence, in a collection of 247 ethnically diverse, ethnically identified genomic DNA samples (see Methods). This data set is available at the PharmGKB and UCSF Pharmacogenetics websites (http://www.pharmgkb.org and http://pharmacogenetics.ucsf.edu). Twenty-eight variable sites were identified (Table 1) and 15 of these were singletons (*e.g.*, present on only one chromosome). Twelve of the 28 were in non-coding or intronic regions of the gene and 16 were in coding regions. Seven of the 16 coding-region single-nucleotide polymorphisms (cSNPs) are synonymous, eight are non-synonymous, and one is a single-nucleotide insertion that leads to a prematurely terminated protein. These 16 cSNPs are distributed throughout the

loops and transmembrane domains (TMDs) of the protein (nine cSNPs in the loops and seven cSNPs in the TMDs) (Figure 1).

Four non-synonymous cSNPs (M165I, A270S, R400C, and K432Q) had ethnic-specific allele frequencies ≥1%. As shown in Figure 2, M165I and R400C were present only in the African American population sample, and K432Q was present in both the African American and Mexican American population samples. A270S was present in all of the populations screened and had a particularly high allele frequency over all populations, 12.7%. The four other non-synonymous cSNPs as well as the insertion variant were present in only one of 494 chromosomes.

Table 2 summarizes estimates of the neutral parameter (θ) and nucleotide diversity (π) calculated for the entire sample as well as for each ethnic group. It should be noted that the number of individuals included in the Mexican American and Pacific Islander populations are small (ten and seven, respectively), and thus the parameters for these groups may not reflect actual population values. For the entire sample, θ ($\times 10^4$) and π ($\times 10^4$) were calculated to be 11.76 ± 3.14 and 5.96 ± 3.69, respectively. A comparison across ethnic groups shows that θ was highest for the African American population (9.69 ± 3.03). Furthermore, θ for the total population analyzed was greater than that of the individual subpopulations, consistent with the existence of population substructure in this sample set. Values of nucleotide diversity (π) ranged from 3.06 in Asian Americans to 6.46 in African Americans. The ratio of nucleotide diversity at synonymous versus non-synonymous sites was greater than 7.5 for all populations. Population substructure was quantified with Wright's F_{ST}, a measure of the proportion of the total variance in nucleotide frequency that occurs

Table 1: Summary of SNPs

exon	SNP #	cDNA position [1]	NT change	AA position	AA Change
1	1.1	-47	C to T	non-coding	
	1.2	57	G to A	19	syn.
	1.3 [2]	**134**	**Insertion (A)**	**45**	
	1.4	**160**	**C to T**	**54**	**Pro to Ser**
	1.5	390	G to T	130	syn.
2	**2.1**	**481**	**T to C**	**161**	**Phe to Leu**
	2.2	**493**	**A to G**	**165**	**Met to Val**
	2.3*	**495**	**G to A**	**165**	**Met to Ile**
	2.4	(+32)	G to C	intron 2	
3	3.1	(-18)	T to C	intron 2	
	3.2	669	C to T	223	syn.
4	**4.1***	**808**	**G to T**	**270**	**Ala to Ser**
5	5.1	(-19)	C to T	intron 4	
	5.2	**890**	**C to G**	**297**	**Ala to Gly**
	5.3	(+28)	deletion	intron 5	
	5.4	(+32)	deletion	intron 5	
6		No SNPs			
7	**7.1***	**1198**	**C to T**	**400**	**Arg to Cys**
	7.2	1203	C to T	401	syn.
8	8.1	(-35)	T to C	intron 7	
	8.2*	**1294**	**A to C**	**432**	**Lys to Gln**
9	9.1	1398	C to T	466	syn.
	9.2	(+25)	T to C	intron 9	
10	10.1	1506	A to G	502	syn.
	10.2	1587	C to T	529	syn.
11	11.1	+107	C to A	non-coding	
	11.2	+201	G to A	non-coding	
	11.3	+254	insertion (T)	non-coding	
	11.4	+272	T to C	non-coding	

Table 1: Summary of SNPs (continued)

SNP #	# chromosomes	Total [3] (n=494)	CA (n=200)	AA (n=200)	AS (n=60)	ME (n=20)	PA (n=14)
1.1	1	0.002	0.000	0.000	0.017	0.000	0.000
1.2	1	0.002	0.000	0.000	0.017	0.000	0.000
1.3 [2]	**1**	**0.002**	**0.005**	**0.000**	**0.000**	**0.000**	**0.000**
1.4	**1**	**0.002**	**0.000**	**0.005**	**0.000**	**0.000**	**0.000**
1.5	137	0.278	0.394	0.205	0.183	0.250	0.143
2.1	**1**	**0.002**	**0.005**	**0.000**	**0.000**	**0.000**	**0.000**
2.2	**1**	**0.002**	**0.000**	**0.005**	**0.000**	**0.000**	**0.000**
2.3*	**2**	**0.004**	**0.000**	**0.010**	**0.000**	**0.000**	**0.000**
2.4	150	0.307	0.296	0.419	0.033	0.300	0.071
3.1	6	0.012	0.000	0.025	0.000	0.050	0.000
3.2	1	0.002	0.000	0.000	0.017	0.000	0.000
4.1*	**62**	**0.127**	**0.157**	**0.110**	**0.086**	**0.150**	**0.071**
5.1	1	0.002	0.000	0.005	0.000	0.000	0.000
5.2	**1**	**0.002**	**0.005**	**0.000**	**0.000**	**0.000**	**0.000**
5.3	58	0.118	0.131	0.145	0.017	0.050	0.071
5.4	58	0.118	0.131	0.145	0.017	0.050	0.071
7.1*	**3**	**0.006**	**0.000**	**0.015**	**0.000**	**0.000**	**0.000**
7.2	1	0.002	0.005	0.000	0.000	0.000	0.000
8.1	6	0.012	0.000	0.005	0.067	0.000	0.071
8.2*	**3**	**0.006**	**0.000**	**0.010**	**0.000**	**0.050**	**0.000**
9.1	1	0.002	0.000	0.005	0.000	0.000	0.000
9.2	1	0.002	0.000	0.005	0.000	0.000	0.000
10.1	173	0.350	0.290	0.500	0.117	0.300	0.143
10.2	1	0.002	0.000	0.000	0.017	0.000	0.000
11.1	5	0.010	0.020	0.005	0.000	0.000	0.000
11.2	1	0.002	0.000	0.005	0.000	0.000	0.000
11.3	6	0.012	0.000	0.030	0.000	0.000	0.000
11.4	1	0.002	0.005	0.005	0.000	0.000	0.000

[1] cDNA numbers are relative to the ATG start site and based on the cDNA sequence from GenBank accession number NM003058.

[2] Non-synonymous changes are shown in boldface.

[3] Columns refer to the frequencies of each variable site in each ethnic group. Some samples contained amplicons that could not be sequenced. Therefore, allele frequencies are based on non-missing data. Abbreviations are as follows: Total = entire sample, CA = Caucasian, AA = African American, AS = Asian American, ME = Mexican American, PA = Pacific Islander. n is the number of chromosomes in each ethnic group.

* Non-synonymous variants with ethnic-specific allele frequencies ≥ 1%.

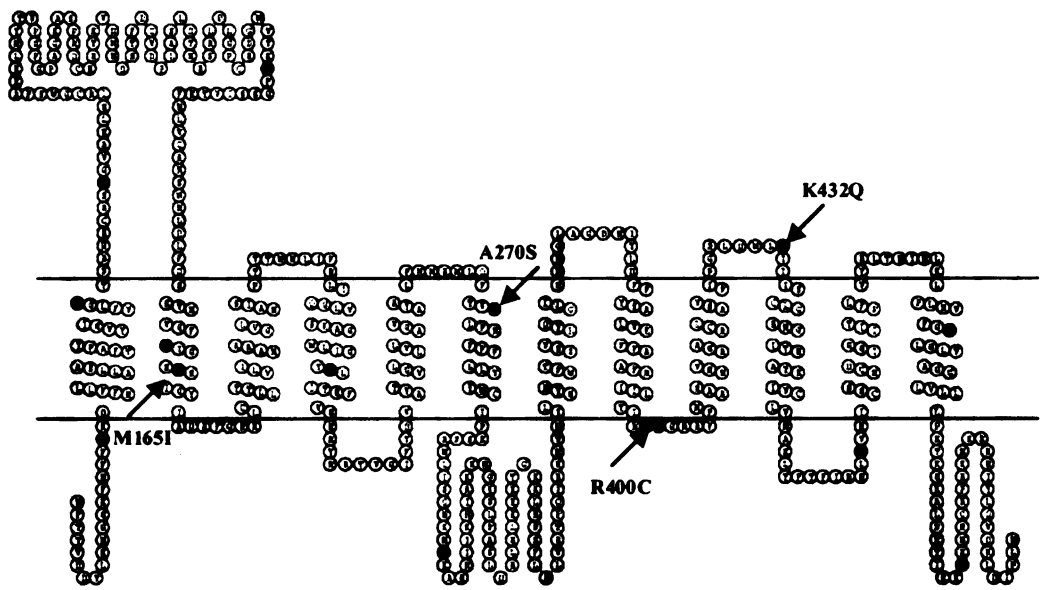


Figure 1. Secondary structure of OCT2 with coding-region SNPs. The transmembrane topology schematic was rendered using TOPO (S.J. Johns [UCSF, San Francisco] and R.C. Speth [Washington State University, Pullman], transmembrane protein display software, available at the UCSF Sequence Analysis Consulting Group website, http://www.sacs.ucsf.edu/TOPO/topo.html). Non-synonymous amino acid changes are shown in red; synonymous changes in green. The one single nucleotide insertion is shown in blue. The four non-synonymous variants that were functionally characterized are indicated by arrows.

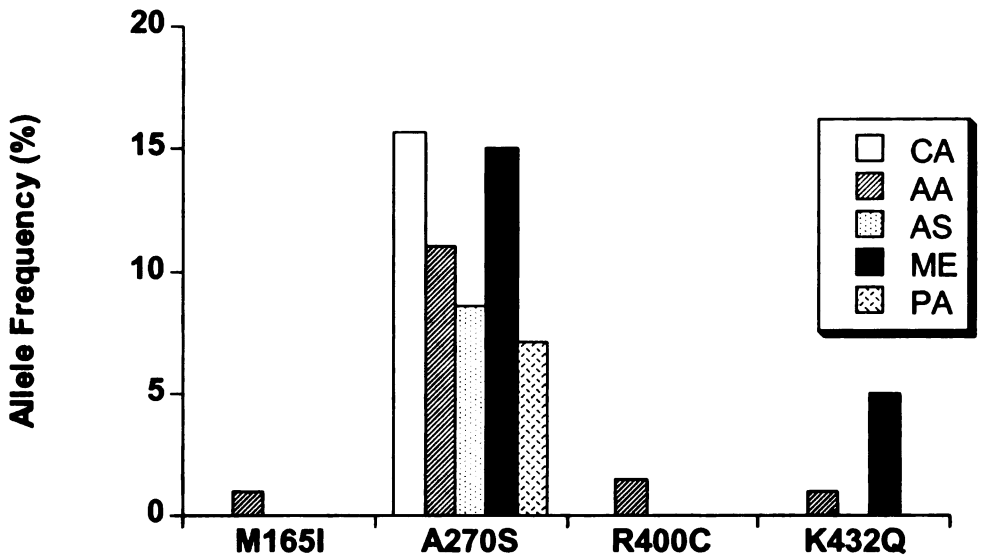


Figure 2. Ethnic-specific frequencies of four non-synonymous cSNPs. Frequencies of M165I, A270S, R400C, and K432Q variants in each ethnic population. CA, Caucasian; AA, African American; AS, Asian American; ME, Mexican; PA, Pacific Islander.

Table 2: Estimates of the neutral parameter (θ), nucleotide diversity (π) and Tajima's D [2]

Population	SNP Type	bp Screened	N [3]	# Variable Sites	θ (x 10^{+4})	π (x 10^{+4})	D
Total [1]	All	3514	494	28	11.76 ± 3.14	5.96 ± 3.69	-1.28
	non-coding	1468	494	12	9.59 ± 3.31	5.13 ± 3.89	-1.03
	Coding	2046	494	16	14.15 ± 4.44	6.88 ± 4.94	-1.21
	synonymous	582	494	7	26.41 ± 11.16	22.44 ± 17.45	-0.29
	non-syn.	1464	494	9	10.40 ± 3.98	2.11 ± 2.56	-1.64
CA	All	3514	200	12	5.82 ± 2.10	5.93 ± 3.69	0.05
	non-coding	1468	200	5	4.61 ± 2.30	4.88 ± 3.78	0.11
	Coding	2046	200	7	7.15 ± 3.12	7.10 ± 5.06	-0.01
	synonymous	582	200	3	13.06 ± 8.06	22.86 ± 17.73	1.20
	non-syn.	1464	200	4	5.33 ± 2.91	2.27 ± 2.67	-1.02
AA	All	3514	200	20	9.69 ± 3.03	6.46 ± 3.95	-0.89
	non-coding	1468	200	11	10.15 ± 3.78	6.21 ± 4.47	-0.93
	Coding	2046	200	9	9.19 ± 3.66	6.75 ± 4.88	-0.60
	synonymous	582	200	3	13.06 ± 8.06	21.48 ± 16.99	1.03
	non-syn.	1464	200	6	8.00 ± 3.70	2.24 ± 2.65	-1.46
AS	All	3514	60	11	6.71 ± 2.73	3.06 ± 2.27	-1.54
	non-coding	1468	60	5	5.81 ± 3.05	1.58 ± 1.88	-1.70
	Coding	2046	60	6	7.71 ± 3.79	4.70 ± 3.84	-0.96
	synonymous	582	60	5	27.42 ± 14.38	16.21 ± 14.26	-0.95
	non-syn.	1464	60	1	1.68 ± 1.74	1.17 ± 1.87	-0.39
ME	All	3514	20	8	6.42 ± 3.22	5.54 ± 3.66	-0.46
	non-coding	1468	20	4	6.11 ± 3.75	4.02 ± 3.47	-0.99
	Coding	2046	20	4	6.76 ± 4.15	7.23 ± 5.37	0.20
	synonymous	582	20	2	14.42 ± 11.41	21.40 ± 17.75	1.14
	non-syn.	1464	20	2	4.41 ± 3.49	2.89 ± 3.22	-0.81
PA	All	3514	14	7	6.27 ± 3.42	3.53 ± 2.67	-1.61
	non-coding	1468	14	4	6.82 ± 4.34	3.10 ± 2.99	-1.80
	Coding	2046	14	3	5.66 ± 3.95	4.02 ± 3.67	-0.89
	synonymous	582	14	2	16.08 ± 13.02	13.49 ± 13.42	-0.44
	non-syn.	1464	14	1	2.46 ± 2.65	1.12 ± 1.92	-1.16

[1] Total = entire sample, CA = Caucasian, AA = African American, AS = Asian American, ME = Mexican American, PA = Pacific Islander.
[2] θ, π and Tajima's D were calculated as described in the Methods.
[3] N = number of chromosomes

among populations [17]. For these OCT2 data, we obtained F_{ST} = 0.178, which is consistent with population substructure and similar to what has been reported in other genes [22].

The occurrence of θ consistently higher than π results in a consistently negative Tajima's *D* statistic (Table 2). This trend occurs when θ is inflated by a large number of rare SNPs that do not contribute significantly to nucleotide diversity. Although these estimates of Tajima's *D* statistic were not significantly different from zero, the trend toward negative values is commonly seen in human variation and may be due to a genome-wide excess of rare variants that occurs with population expansion [17]. Another possible explanation for negative Tajima's *D* values is that natural selection is driving down the frequency of slightly deleterious variants, a possibility that is especially plausible for the non-synonymous changes.

Haplotype Analysis

Haplotypes were determined using an inferential procedure [18]. This method identified 49 distinct haplotypes, of which 13 were unambiguous (*i.e.*, present in individuals who were homozygous or heterozygous at a single site) (Figure 3). Ten variable sites were represented in the 13 unambiguous haplotypes. Because the combined frequency of the unambiguous haplotypes was greater than 85%, we focused on this group. These haplotypes were named based on two criteria; first, their evolutionary relationship and, second, their frequency in the sample set. Figures 4A and 4B show the ethnic-specific distribution of these haplotypes. Haplotypes *1 and *2A accounted for approximately 60% of the total sample and were present in all ethnic

groups. Haplotype *2A was the most frequent among Caucasians, whereas haplotype *1 was the most frequent in the African American population sample. With the exception of haplotypes *1, *2A, *3C, and *3D, all other unambiguous haplotypes were present in only one or two ethnic groups, consistent with the existence of population substructure in this sample set. Only one of the common non-synonymous variants (A270S) was present in an unambiguous haplotype (haplotypes *3D, *3E, and *6).

When the remaining 36 haplotypes were included in this analysis, the data were again consistent with the existence of population substructure. Thirty-four of the 49 haplotypes (69%) were found in only one ethnic group in our sample. Twenty were present only in the African American sample, eight only in the Caucasian sample, five only in the Asian American sample, and one only in the Mexican American sample. In contrast, only three of the 49 haplotypes (haplotypes *1, *2A, and *3C; 6%) were present in all ethnic population samples.

A cladogram showing the relationships among the 13 unambiguous haplotypes is shown in Figure 5. This cladogram suggests that site 1.5, specifying the fifth SNP in exon 1, has recurrently appeared, as it occurs in three independent positions in the cladogram. Alternatively, haplotypes *2B and *3E may have arisen from recombination of *3A and *3D with haplotype *2A. In addition, site 10.1 occurs twice. Thus, haplotypes *1, *2A, *2B, and *3A are identical apart from these two SNPs.

Functional Activity of Common Variants

To examine the functional consequences of genetic variation in OCT2 *in vitro*, the four common non-synonymous coding region variants (M165I, A270S, R400C, and

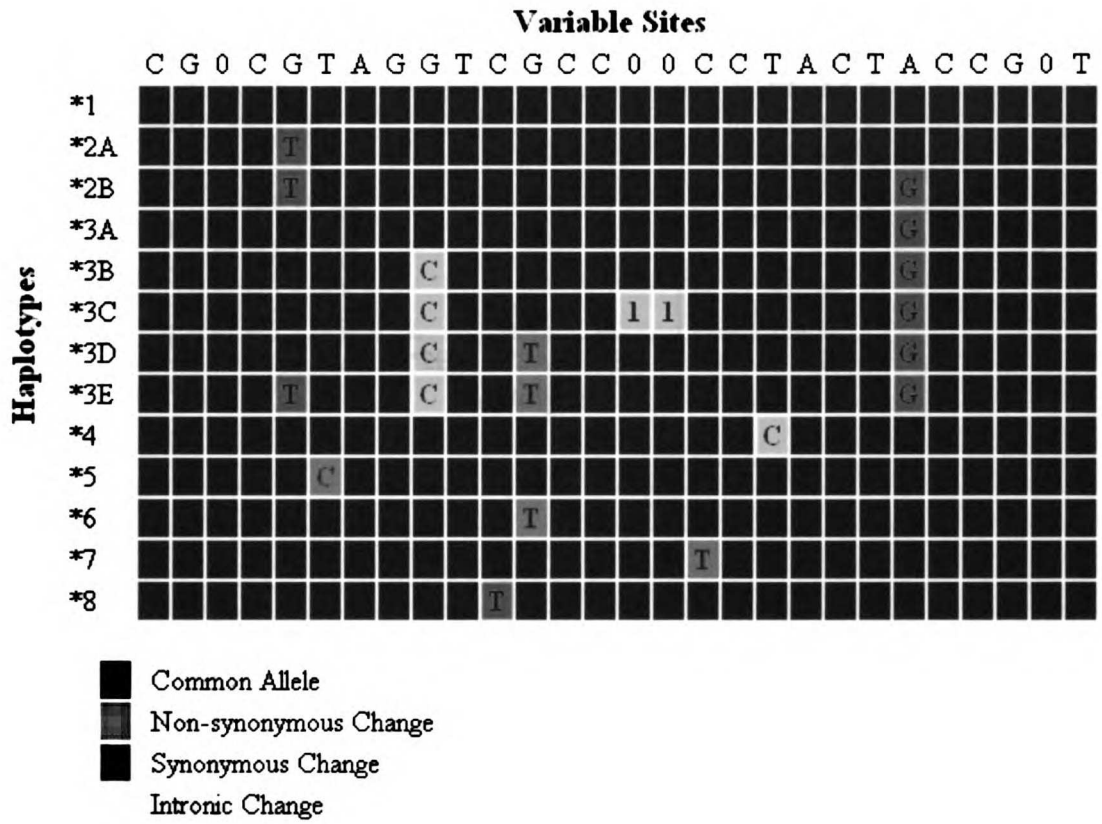
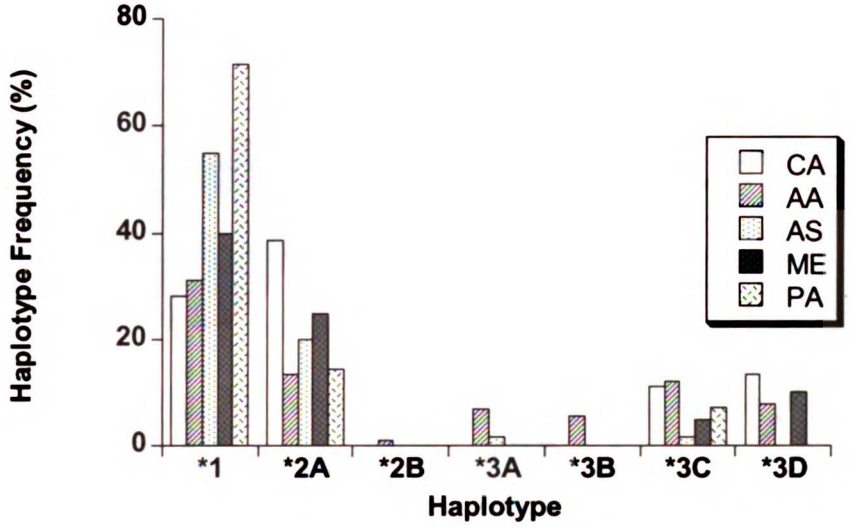


Figure 3. Sequences of thirteen unambiguous haplotypes. Sequences of the unambiguous haplotypes at each of the 28 variable sites. The first row denotes the sequence of the common allele at each site. 0 signifies no insertion or deletion; 1 signifies the presence of an insertion or deletion of a single nucleotide. The ten sites with changes in this group of haplotypes are (from the left): 1.5, 2.1, 2.4, 3.2, 4.1, 5.3, 5.4, 7.1, 8.1, and 10.1.

A



B

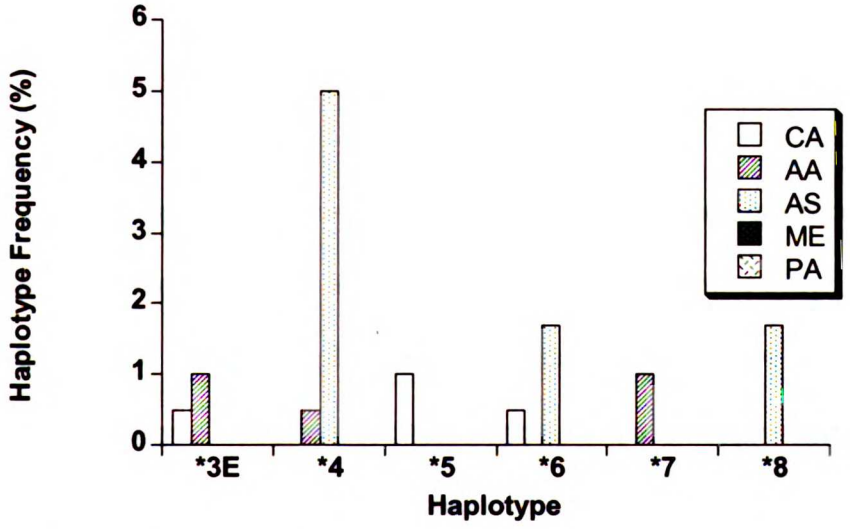


Figure 4. Ethnic-specific frequencies of 13 unambiguous haplotypes. A) Frequencies of haplotypes *1-*3D. B) Frequencies of haplotypes *3E-*8. CA, Caucasian; AA, African American; AS, Asian American; ME, Mexican; PA, Pacific Islander.

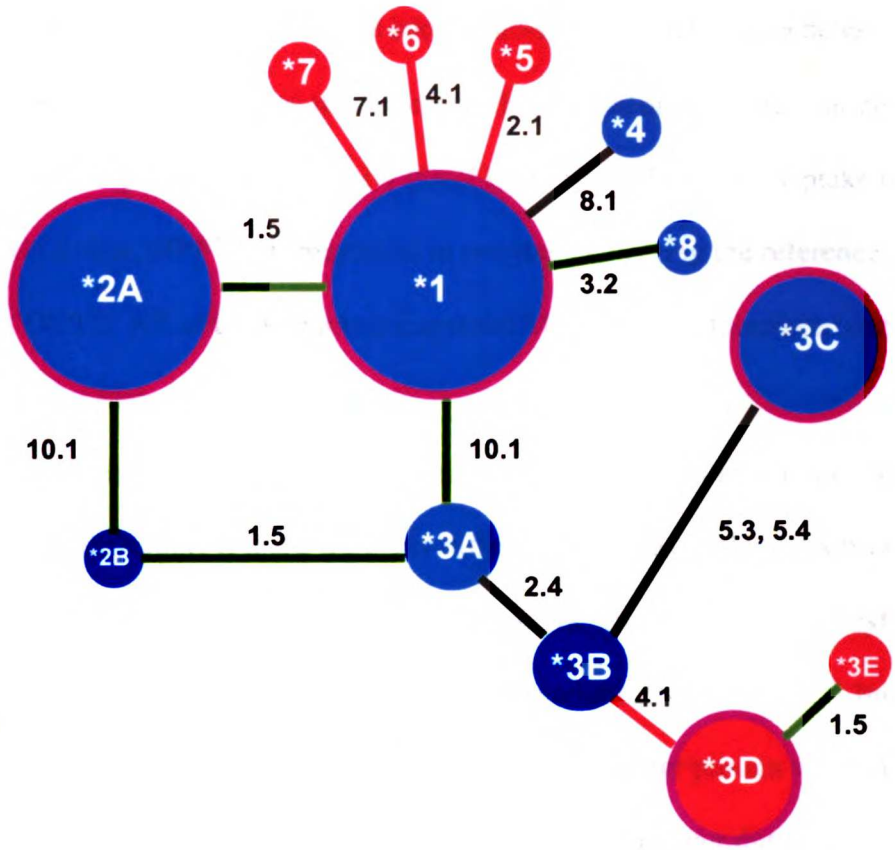


Figure 5. Cladogram showing the relationships between haplotypes. Haplotypes with red-filled circles contain a non-synonymous variant; those with blue-filled circles do not. A pink halo around a haplotype indicates that this haplotype is present in at least three ethnic groups. Lines between haplotypes indicate the type of change (black, non-coding; green, synonymous; red, non-synonymous). The relative areas of the circles are proportional to the relative frequencies of each haplotype.

K432Q) were constructed using site-directed mutagenesis and expressed in *Xenopus laevis* oocytes: each of these variants had ethnic-specific allele frequencies ≥ 1%. The R400C variant is particularly notable because it changed an arginine residue that is conserved throughout the OCT family (OCT1, OCT2, and OCT3). Uptake of the neurotoxic amine, MPP^+, was measured in oocytes expressing the reference and variant forms of OCT2. All of the variants transported MPP^+ and are therefore functional in *X. laevis* oocytes (data not shown).

Following this initial test for function, RNA dose-response studies were conducted to determine if there were any differences in maximal transport rates (*i.e.*, V_{max}) between the reference and variant transporters. In these studies, oocytes were injected with increasing amounts of reference or variant RNA (5-75 ng). Injecting the oocytes with high doses of RNA ensured that the transporter protein was maximally expressed and controlled for RNA degradation that might occur during injection. Uptake of a saturating amount of MPP^+ (100 μM) was then measured. This procedure measures, in effect, an apparent V_{max} for the reference and variant transporters and can therefore give insight into the effect of the amino acid substitutions on transporter expression level or turnover rate. Figures 6 and 7 show the RNA dose-response curves for M165I and R400C compared to the reference. Both variants had dose-response curves significantly lower than that of the reference. Of the four variants tested, R400C had the lowest maximal activity (3- to 4-fold lower than the reference). Both the high frequency variant, A270S, and K432Q had dose-response curves similar to the OCT2 reference.

We next determined whether the variants exhibited alterations in their affinities for the prototypical organic cation MPP^+. The uptake of increasing concentrations of

MPP^+ was measured in oocytes expressing the reference and variant transporters, and the data were fit as previously described [21]. The K_m values of MPP^+ are listed in Table 3 for the OCT2 reference and all four variants. A representative curve of MPP^+ uptake for the reference OCT2 and variant K432Q is shown in Figure 8. Although the average K_m values for M165I and A270S were slightly larger than that of the OCT2 reference, these differences were not statistically significant. In three out of four experiments, the M165I and A270S variants had K_m values very close to that of the OCT2 reference. The K_m was significantly lower for the K432Q variant, suggesting that it had an increased affinity for MPP^+.

To determine whether the OCT2 variants altered recognition by other molecules, in particular, inhibitors, we carried out inhibition studies with a series of n-tetraalkylammonium (nTAA) compounds of increasing hydrophobicity. In the first study, inhibition of ^3H-MPP^+ uptake by tetramethylammonium (TMA), tetraethylammonium (TEA), tetrapropylammonium (TPrA), and tetrabutylammonium (TBA) was measured in oocytes expressing the OCT2 reference and variants. TMA, TEA, and TPrA inhibited the reference form of OCT2 and all OCT2 variants similarly (data not shown). In contrast, the most hydrophobic compound, TBA, inhibited some of the variants differently from the OCT2 reference and was studied further. Inhibition of ^3H-MPP^+ uptake by increasing concentrations of TBA was measured and the data fit as previously described to determine apparent K_i values [21]. Figure 9 shows a representative curve for inhibition of MPP^+ uptake by TBA in oocytes expressing the K432Q variant or OCT2 reference. A summary of the K_i values of TBA for the four variants and the OCT2 reference is shown in Table 3. These studies suggest that

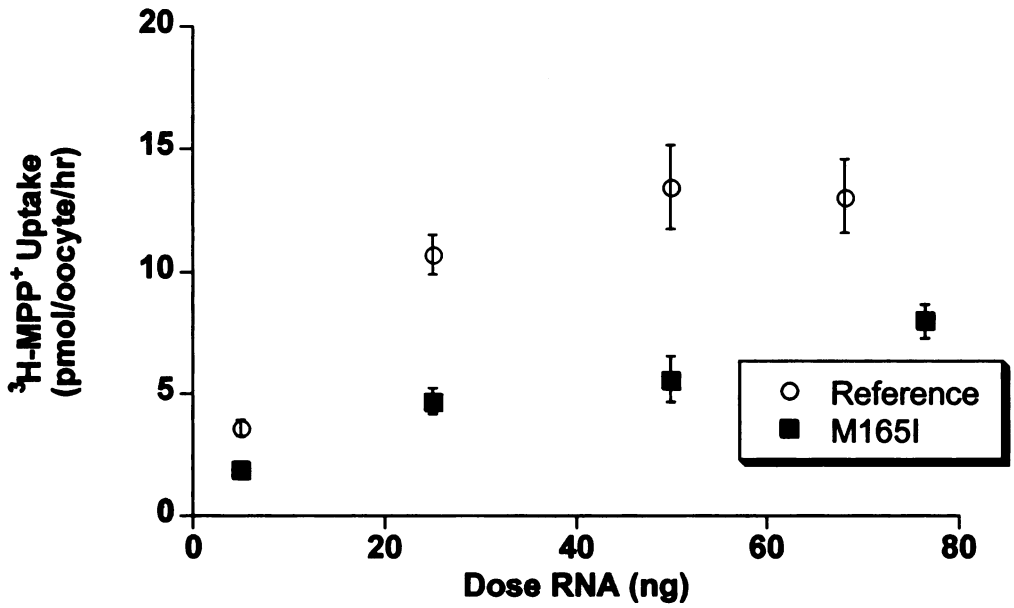


Figure 6. cRNA dose-dependent uptake of MPP^+ for OCT2 reference and variant M165I. *Xenopus laevis* oocytes were injected with approximately 5, 25, 50, and 75 ng of reference (O) or variant (■) OCT2 RNA. Uptake of MPP^+ (100 μM unlabeled MPP^+, 0.1 μM ^{3}H-MPP^+) was measured at room temperature after a 1 hour incubation. RNA remaining after injection of oocytes was run on a gel to verify that RNA was still intact. Data are representative of experiments carried out with three different batches of oocytes. Each value represents mean ± standard error from 6-9 oocytes.

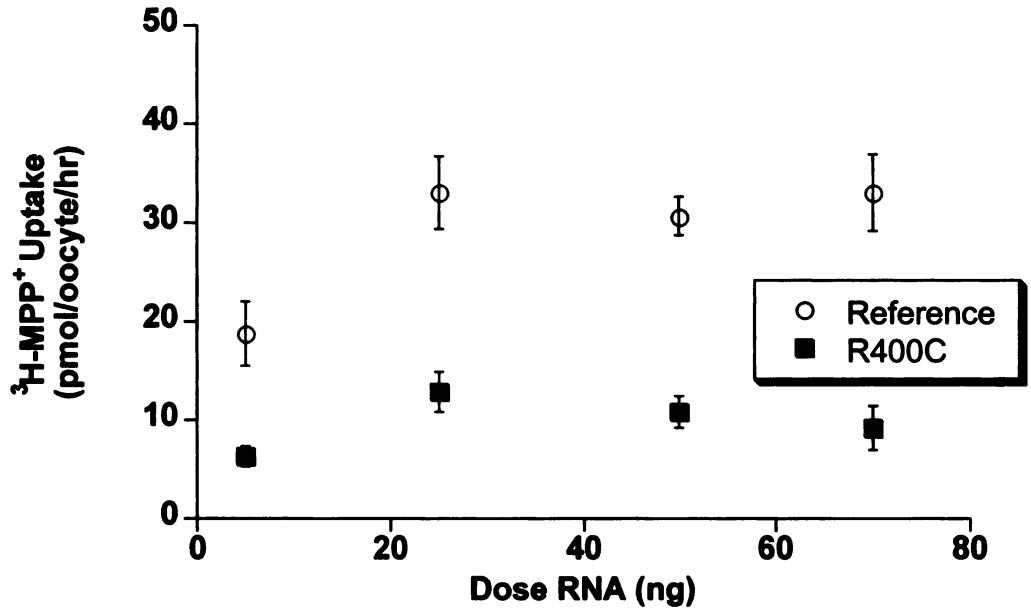


Figure 7. cRNA dose-dependent uptake of MPP^+ for OCT2 reference and variant R400C. *Xenopus laevis* oocytes were injected with approximately 5, 25, 50, and 75 ng of reference (O) or variant (■) OCT2 RNA. Uptake of MPP^+ (100 μM unlabeled MPP^+, 0.1 μM ^{3}H-MPP^+) was measured at room temperature after a 1 hour incubation. RNA remaining after injection of oocytes was run on a gel to verify that RNA was still intact. Data are representative of experiments carried out with three different batches of oocytes. Each value represents mean ± standard error from 6-9 oocytes.

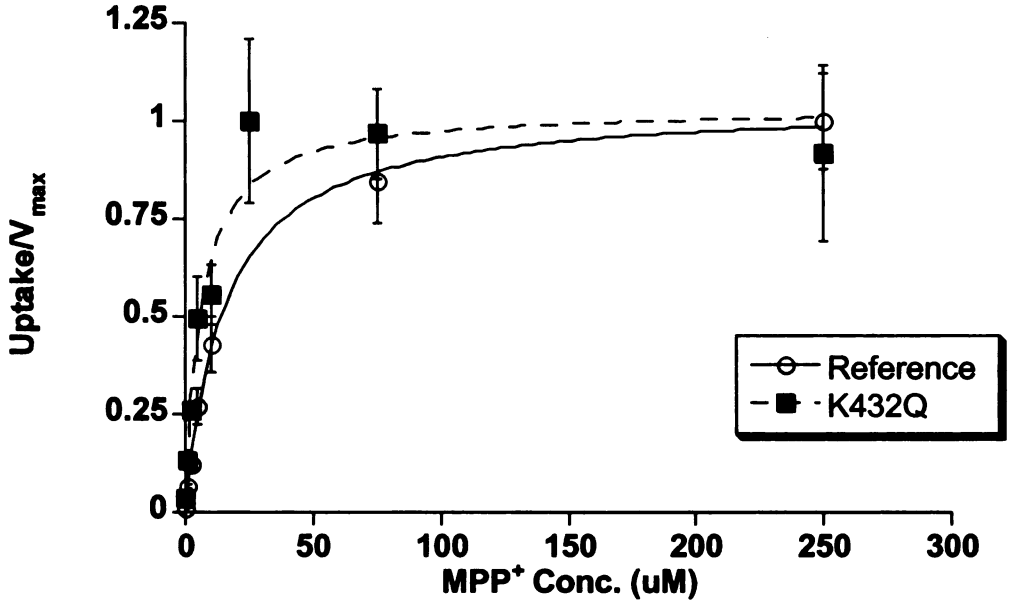


Figure 8. Representative kinetics of MPP^+ uptake. Michaelis-Menten curve for the OCT2 reference (O) and K432Q variant (■). Oocytes were injected with 30-50 ng of OCT2 reference or variant RNA, and uptake of ^3H-MPP^+ in the presence of increasing concentrations of MPP^+ was measured after a 75-minute incubation. Data are representative of experiments carried out with 2-3 different batches of oocytes. Each value represents mean ± standard error from 6-9 oocytes.

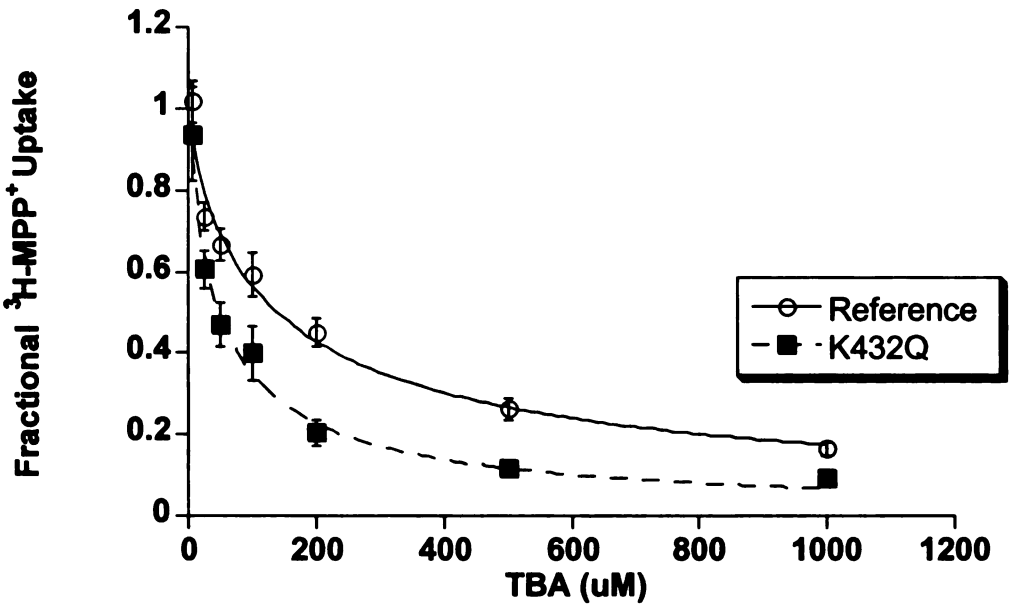


Figure 9. Representative kinetics of inhibition of MPP^+ by tetrabutylammonium. Kinetics of MPP^+ inhibition by TBA for the OCT2 reference (O) and K432Q variant (■). Oocytes were injected with 30-50 ng of OCT2 reference or variant RNA, and uptake of ^3H-MPP^+ in the presence of increasing concentrations of TBA was measured after a 75-minute incubation. Data are representative of experiments carried out with 2-3 different batches of oocytes. Each value represents mean ± standard error from 6-9 oocytes.

Table 3: Kinetic characteristics and interactions with various organic cations[1]

	$K_{m,\ MPP+}$ (μM)	$K_{i,\ TBA}$ (μM)	Inhibition by: Metformin[1] (% control)	Phenformin (% control)	Procainamide (% control)	Quinidine (% control)
Reference	11.8 ± 1.3	148 ± 33	28.3 ± 3.7	28.9 ± 5.5	53.9 ± 5.6	52.0 ± 11.8
M165I	18.2 ± 8.2	146 ± 67	29.2 ± 4.2	28.5 ± 2.1	41.1 ± 1.2	69.2 ± 13.5
A270S	20.8 ± 8.1	274 ± 16*	32.5 ± 0.4	35.1 ± 7.1	56.3 ± 5.6	47.1 ± 1.9
R400C	11.2 ± 5.7	99 ± 24*	19.8 ± 7.7	21.7 ± 1.9	74.5 ±12.6	41.1 ± 24.3
K432Q	6.6 ± 0.6*	77 ± 13*	28.0 ± 2.2	16.2 ± 5.7	66.1 ±10.7	62.5 ± 1.4

[1] Data is listed as mean ± standard error
[2] Concentrations of metformin, phenformin, procainamide, and quinidine were 2.5 mM, 100 μM, 100μM, and 50 μM, respectively.
* $P < 0.05$ by analysis of variance

TBA was a more potent inhibitor of the R400C and K432Q variants ($p<0.05$) than of the OCT2 reference. In addition, the data suggest a decreased inhibition of the A270S variant by TBA ($p<0.05$). There was no significant difference in the inhibition potency of TBA between the OCT2 reference and M165I variant.

Since OCT2 interacts with a variety of drugs [3], inhibition of MPP^+ uptake by various organic cation drugs was also measured in oocytes expressing the reference and variant forms of OCT2. These studies (Table 3) showed that all of the variants behaved similarly to the reference form of OCT2 with respect to inhibition by metformin, phenformin, procainamide, and quinidine when assayed at a single concentration of each organic cation. Whether the variants exhibit differences in inhibition at other concentrations remains to be determined.

DISCUSSION

We report here a comprehensive genetic and functional analysis of variation in the human kidney xenobiotic transporter, OCT2, a major transporter involved in the renal secretion of organic cations. OCT2, one of six organic cation transporters that have been biochemically characterized (out of a total of seven known from the human genome sequence), interacts with a variety of small hydrophilic organic cations, including endogenous compounds (*e.g.*, dopamine and norepinephrine), toxic substances (*e.g.*, MPP^+), and clinically used drugs (*e.g.*, metformin, procainamide, and cimetidine) [6-12]. To understand better the potential effects of genetic variation in a xenobiotic transporter on renal elimination, we screened a large, ethnically diverse collection of genomic DNA samples to identify variants of OCT2. Population genetic analysis of the variants and the

biochemical characterization of four common non-synonymous variants suggests that OCT2 plays a significant role in renal elimination.

Genetic Variation

We estimated the neutral parameter (θ), nucleotide diversity (π), and Tajima's *D* for the entire coding sequence of OCT2 and 50-100 base pairs of flanking intronic sequence. Our values of θ and π (11.76 ± 3.14 and 5.96 ± 3.69, respectively) fall within the range of those reported for other genes [22-25]. The most striking result from our analysis of genetic variation in OCT2 was the difference in nucleotide diversity (π) at synonymous versus non-synonymous sites in the coding region. The ratio of $\pi_{\text{synonymous}}$ to $\pi_{\text{non-synonymous}}$ for OCT2 was more than two-fold higher than that reported by Cargill *et al.* and Halushka *et al.*, who examined genetic variation in over 75 genes each [23-25]. We have analyzed the same set of DNA samples used for OCT2 to identify variants of another xenobiotic transporter, MDR1, and also observed that the ratio of $\pi_{\text{synonymous}}$ to $\pi_{\text{non-synonymous}}$ for MDR1 is like that of other genes [22-25]. The higher ratio for OCT2 was seen in the total population as well as in each ethnic group. In addition, Tajima's *D* for non-synonymous sites was consistently lower and more negative compared to the corresponding *D* for synonymous sites. This trend occurred in the entire sample as well as within each ethnic group, with the exception of the Asian American group. Although none of the *D* values reached statistical significance alone, the probability of observing five out of six non-synonymous *D* values less than the synonymous *D* values is less than 10% (using non-parametric analysis). In addition, in a set of 1000 coalescence simulations of 494 chromosomes having 7 synonymous changes in 582 sites and 9 non-

synonymous changes in 1464 sites, the distribution of $D_{syn} - D_{nonsyn}$ did not even span the observed value of 1.35. This implies that the non-synonymous sites had a significantly depressed Tajima's *D* compared to the synonymous sites. The prevalence of lower and more negative *D* values for non-synonymous sites combined with a high ratio of $\pi_{synonymous}$ to $\pi_{non\text{-}synonymous}$ suggests that selection is acting to preserve the amino acid sequence of OCT2. Such selection may be due to an important role of this transporter in elimination and detoxification of organic cations by the kidney or by the choroid plexus. Recent studies indicate that OCT2 may play an important role in brain choline homeostasis by mediating choline transport across the ventricular membrane of the choroid plexus [26]. Therefore, selection against amino acid changes may be due to an important endogenous role of this transporter as well. Another explanation is that OCT2 contains more functionally critical domains than other proteins and is less tolerant of amino acid substitutions. The amino acid sequence of OCT2 does not appear to be more conserved in different mammalian species than other xenobiotic transporters, for example, MDR1: both human MDR1 and OCT2 are approximately 80-85% identical to their respective rat and mouse isoforms. Thus, there does not appear to be special conservation of amino acid sequence in OCT2.

Genetic variation has been examined not only in MDR1 but also in one other xenobiotic transporter, OATP-C [27-30]. In a recent report, 69 individuals were screened for variation in the coding sequence of *SLC21A6*, which codes for an organic anion transporter (OATP-C) [31]. Of fourteen non-synonymous SNPs, five had ethnic-specific allele frequencies of 9% to 74%. The occurrence of non-synonymous SNPs at such high frequencies is in sharp contrast to OCT2, in which only one of eight non-synonymous

variants had an allele frequency greater than 10%. Although the data for OATP-C are not sufficient to calculate nucleotide diversities, it appears that non-synonymous variants of OATP-C occur at a greater frequency that those of OCT2. The comparisons with MDR1 and OATP-C reinforce the suggestion that OCT2 is relatively intolerant of non-synonymous changes.

There are several notable features of the cladogram that bear comment. First of all, the variant site 1.5 appears to have arisen three independent times. One possibility is that these occurrences reflect *de novo* mutation. Because variant site 1.5 is a G → C change, this would be an unconventional mutational hotspot. Another possibility is that some of the appearances of 1.5 resulted from recombination/gene conversion, for example, between haplotypes *2A and *3A or *2A and *3D. It is striking that haplotypes *1 and *2A differ by only a single variant site but that *1 has six descendent haplotypes whereas *2A has none or only one. This pattern could be explained if the variant 1.5 arose repeatedly. Another possibility is that *2A arose recently and underwent rapid expansion.

Functional Analysis

In vitro phenotypic data on the four most common non-synonymous variants revealed functional differences between the OCT2 reference and the variants. Variants M165I and R400C had significantly lower V_{max} values than the reference form of OCT2 (~50% and 70% lower, respectively). This decreased V_{max} may be due to reduced expression of the transporter or lower catalytic activity. Although it is possible that these two variants function normally in renal epithelial cells, the *in vitro* studies were

conducted under conditions in which the transporter proteins were over-expressed; such conditions are not likely to be achieved in renal epithelial cells. It is therefore likely that these variants will result in an overall decrease in transporter function *in vivo*. Studies in human kidney tissue may help to elucidate the *in vivo* functional activity and expression levels of these variants. R400C and K432Q showed increased sensitivity to inhibition by tetrabutylammonium compared to the reference form of OCT2. The most frequent non-synonymous variant, A270S, was less sensitive to inhibition by tetrabutylammonium (increased $K_{i,\,TBA}$). In addition, K432Q had a stronger affinity for the prototypical organic cation MPP^+. In general, the less frequent non-synonymous variants resulted in more significant and deleterious functional changes (*e.g.* decreased V_{max} for M165I and R400C and altered interactions with organic cations for R400C and K432Q). In contrast, the most frequent non-synonymous variant exhibited more subtle functional differences from the reference form of OCT2. The observation that variants with more significant functional changes occur only at low frequency supports the conclusions drawn from the ratio of non-synonymous to synonymous nucleotide changes as well as the *D* values and suggests that there is selection against changes in the amino acid sequence of OCT2.

We expect that some of the non-synonymous and insertion variants found on only one chromosome in our sample set would also result in significant functional changes. For example, the single-nucleotide insertion at position 134 results in a prematurely truncated protein of 47 amino acids and would almost certainly abolish transporter function. In addition, the P54S variant changes a proline residue that is conserved throughout the OCT family and thus may exhibit decreased function. Collectively, variants that are expected to affect function (P54S, M165I, R400C, K432Q, and insertion

allele, 1.3) were found at a frequency of 0.02 (10/494). Individuals with such variants could exhibit poor renal elimination of xenobiotics.

Implications

The allele frequencies and functional characteristics of the four non-synonymous variants give insight into the reason for the low ratio of nucleotide diversity at non-synonymous versus synonymous sites and the lower, more negative *D* values at non-synonymous sites. Two of the four variants (M165I and R400C) had significantly reduced activity compared to the OCT2 reference and were present in the population at low allele frequencies overall (although they were present at a frequency of approximately 1% in at least one ethnic population). In contrast, the non-synonymous variant with the highest allele frequency, A270S, had more subtle effects on transporter function. These observations suggest that many non-synonymous mutations in OCT2 adversely affect transporter function. Moreover, the data suggest that common non-synonymous mutations may have subtle functional differences that could be overlooked if a detailed functional characterization is not conducted. When summed across multiple genes, these common variations with minor functional differences could lead to an altered drug response phenotype *in vivo*. The population-genetic analysis suggests that during the history of the human species, selection has acted against amino acid changes in OCT2. The significant skew of non-synonymous SNPs toward rarer frequencies than synonymous SNPs implies that extant amino acid variants are detected by selection acting in the population at large. Such selection may be due to a necessary role of OCT2 in the renal elimination of endogenous or xenobiotic organic cations. Selection for OCT2

function in the context of evolutionary time may have acted on environmental xenobiotics such as plant toxins related to food supply.

In modern times, variation in OCT2 may influence drug response by altering the renal elimination of xenobiotics. All of the four variants affected transporter function in our heterologous expressions system. These *in vitro* data are difficult to extrapolate to *in vivo* renal clearance; however, they suggest that altered renal elimination of drugs may occur in individuals who carry these variants. Future studies in humans with OCT2 variants will elucidate the relationship between genetic variation in OCT2 and renal drug elimination and toxicities.

REFERENCES

1. Weber WW. The legacy of pharmacogenetics and potential applications. *Mutation Research* **479:** 1-18, 2001.

2. Burckhardt G and Wolff NA. Structure of renal organic anion and cation transporters. *Am J Physiol Renal Physiol* **278**(6)**:** F853-66, 2000.

3. Dresser MJ, Leabman Kaushal M and Giacomini KM. Transporters involved in the elimination of drugs in the kidney: organic anion transporters and organic cation transporters. *J Pharm Sci* **90**(4)**:** 397-421, 2001.

4. Sirtori CR, Franceschini G, Galli-Kienle M, Cighetti G, Galli G, Bondioli A and Conti F. Disposition of metformin (N,N-dimethylbiguanide) in man. *Clin Pharmacol Ther* **24**(6)**:** 683-93, 1978.

5. Reidenberg MM, Camacho M, Kluger J and Drayer DE. Aging and renal clearance of procainamide and acetylprocainamide. *Clin Pharmacol Ther* **28**(6)**:** 732-5, 1980.

6. Okuda M, Saito H, Urakami Y, Takano M and Inui K. cDNA cloning and functional expression of a novel rat kidney organic cation transporter, OCT2. *Biochem Biophys Res Commun* **224**(2)**:** 500-507, 1996.

7. Gorboulev V, Ulzheimer JC, Akhoundova A, Ulzheimer-Teuber I, Karbach U, Quester S, Baumann C, Lang F, Busch AE and Koepsell H. Cloning and characterization of two human polyspecific organic cation transporters. *DNA Cell Biol* **16**(7)**:** 871-881, 1997.

8. Busch AE, Karbach U, Miska D, Gorboulev V, Akhoundova A, Volk C, Arndt P, Ulzheimer JC, Sonders MS, Baumann C, Waldegger S, Lang F and Koepsell H.

Human neurons express the polyspecific cation transporter hOCT2, which translocates monoamine neurotransmitters, amantadine, and memantine. *Mol Pharmacol* **54**(2): 342-52, 1998.

9. Urakami Y, Okuda M, Masuda S, Saito H and Inui KI. Functional characteristics and membrane localization of rat multispecific organic cation transporters, OCT1 and OCT2, mediating tubular secretion of cationic drugs. *J Pharmacol Exp Ther* **287**(2): 800-805, 1998.

10. Okuda M, Urakami Y, Saito H and Inui K. Molecular mechanisms of organic cation transport in OCT2-expressing Xenopus oocytes. *Biochim Biophys Acta* **1417**(2): 224-231, 1999.

11. Pan BF, Sweet DH, Pritchard JB, Chen R and Nelson JA. A transfected cell model for the renal toxin transporter, rOCT2. *Toxicol Sci* **47**(2): 181-186, 1999.

12. Arndt P, Volk C, Gorboulev V, Budiman T, Popp C, Ulzheimer-Teuber I, Akhoundova A, Koppatz S, Bamberg E, Nagel G and Koepsell H. Interaction of cations, anions, and weak base quinine with rat renal cation transporter rOCT2 compared with rOCT1. *Am J Physiol Renal Physiol* **281**(3): F454-68, 2001.

13. Sugawara-Yokoo M, Urakami Y, Koyama H, Fujikura K, Masuda S, Saito H, Naruse T, Inui K and Takata K. Differential localization of organic cation transporters rOCT1 and rOCT2 in the basolateral membrane of rat kidney proximal tubules. *Histochem Cell Biol* **114**(3): 175-80, 2000.

14. Pietig G, Mehrens T, Hirsch JR, Cetinkaya I, Piechota H and Schlatter E. Properties and regulation of organic cation transport in freshly isolated human proximal tubules. *J Biol Chem* **276**(36): 33741-6, 2001.

15. Glatt CE, DeYoung JA, Delgado S, Service SK, Giacomini KM, Edwards RH, Risch N and Freimer NB. Screening a large reference sample to identify very low frequency sequence variants: comparisons between two genes. *Nat Genet* **27**(4): 435-8, 2001.

16. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**(3): 585-95, 1989.

17. Hartl DL and Clark AG. *Principles of Population Genetics*. Sinauer Associates, Inc., Sunderland, 1997.

18. Clark AG. Inference of haplotypes from PCR-amplified samples of diploid populations. *Mol Biol Evol* **7**(2): 111-22, 1990.

19. Stephens M, Smith NJ and Donnelly P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* **68**(4): 978-89, 2001.

20. Giacomini KM, Markovich D, Werner A, Biber J, Wu X and Murer H. Expression of a renal Na(+)-nucleoside cotransport system (N2) in Xenopus laevis oocytes. *Pflugers Arch* **427**(3-4): 381-3, 1994.

21. Zhang L, Schaner ME and Giacomini KM. Functional characterization of an organic cation transporter (hOCT1) in a transiently transfected human cell line (HeLa). *J Pharmacol Exp Ther* **286**(1): 354-361, 1998.

22. Stephens JC, Schneider JA, Tanguay DA, Choi J, Acharya T, Stanley SE, Jiang R, Messer CJ, Chew A, Han JH, Duan J, Carr JL, Lee MS, Koshy B, Kumar AM, Zhang G, Newell WR, Windemuth A, Xu C, Kalbfleisch TS, Shaner SL, Arnold K, Schulz V, Drysdale CM, Nandabalan K, Judson RS, Ruano G and Vovis GF.

Haplotype variation and linkage disequilibrium in 313 human genes. [Erratum In: Science 2001 Aug 10;293(5532):1048]. *Science* **293**(5529): 489-93, 2001.

23. Nickerson DA, Taylor SL, Weiss KM, Clark AG, Hutchinson RG, Stengard J, Salomaa V, Vartiainen E, Boerwinkle E and Sing CF. DNA sequence diversity in a 9.7-kb region of the human lipoprotein lipase gene [see comments]. *Nat Genet* **19**(3): 233-40, 1998.

24. Halushka MK, Fan JB, Bentley K, Hsie L, Shen N, Weder A, Cooper R, Lipshutz R and Chakravarti A. Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. *Nat Genet* **22**(3): 239-47, 1999.

25. Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, Patil N, Lane CR, Lim EP, Kalayanaraman N, Nemesh J, Ziaugra L, Friedland L, Rolfe A, Warrington J, Lipshutz R, Daley GQ and Lander ES. Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat Genet* **22**(3): 231-8, 1999.

26. Sweet DH, Miller DS and Pritchard JB. Ventricular choline transport: a role for organic cation transporter 2 expressed in choroid plexus. *J Biol Chem* **276**(45): 41611-9, 2001.

27. Konig J, Cui Y, Nies AT and Keppler D. A novel human organic anion transporting polypeptide localized to the basolateral hepatocyte membrane. *Am J Physiol Gastrointest Liver Physiol* **278**(1): G156-64, 2000.

28. Tamai I, Nezu J, Uchino H, Sai Y, Oku A, Shimane M and Tsuji A. Molecular identification and characterization of novel members of the human organic anion transporter (OATP) family. *Biochem Biophys Res Commun* **273**(1): 251-60, 2000.

29. Silverman JA. Multidrug-resistance transporters. *Pharm Biotechnol* **12:** 353-86, 1999.

30. Ambudkar SV, Dey S, Hrycyna CA, Ramachandra M, Pastan I and Gottesman MM. Biochemical, cellular, and pharmacological aspects of the multidrug transporter. *Annu Rev Pharmacol Toxicol* **39:** 361-98, 1999.

31. Tirona RG, Leake BF, Merino G and Kim RB. Polymorphisms in *OATP-C*: Identification of Multiple Allelic Variants Associated with Altered Transport Activity Among European- and African-Americans. *Journal of Biological Chemistry*, in press.

CHAPTER 3

ESTIMATING THE GENETIC COMPONENT CONTRIBUTING TO VARIATION IN RENAL CLEARANCE

INTRODUCTION

Twin studies have long been a valuable tool for measuring the contribution of heredity to a particular trait, such as drug response. In some of the earliest pharmacogenetic studies, monozygotic and dizygotic twin pairs were used to show that variation in the elimination of many drugs was largely influenced by heredity [1]. Monozygotic twin pairs showed little variation in the elimination of various drugs whereas dizygotic twin pairs, sharing only about half of their genes, showed much greater variability. It is now known that some of the variability in drug elimination observed in dizygotic twins is due to genetic differences in drug metabolizing enzymes in the liver. Recent studies in pharmacogenetics have continued to focus on the genes encoding metabolizing enzymes as the heritable factor in drug response. In contrast, no study to our knowledge has examined the heritability of renal elimination, the other major pathway of detoxification. In this chapter, we have estimated the heritability of renal clearance using a method recently proposed by Kalow *et al.* [2, 3]

Renal excretion is the major pathway for elimination of many clinically used drugs and is the exclusive pathway for eliminating virtually all of the end-products of drug-metabolizing enzymes [4]. Excretion of a particular drug or metabolite is the net result of three major processes: filtration at the glomerulus, secretion, and/or reabsorption [5]. Filtration is a passive process depending mainly on the size of the compound being

filtered. In contrast, secretion is an active process. Transporters in the proximal tubule epithelium, for example, organic cation transporters (OCTs) and organic anion transporters (OATs), mediate the active secretion of a variety of small hydrophilic compounds [6]. Reabsorption can occur by either passive or active processes; however, most exogenous compounds undergo only passive reabsorption [5].

The new methodology proposed by Kalow *et al.*, termed Repeated Drug Application (RDA), can be used to estimate heritability without using the twin study design. This method relies on data obtained from studies in which a group of individuals receives an identical dose of drug on two or more occasions and allows for the calculation of variance in a pharmacokinetic parameter within an individual and between individuals. Variance between individuals, $SD_{between}^2$, is influenced mainly by variation in environmental and genetic factors as well as measurement errors ($SD_{between}^2 = SD_E^2 + SD_G^2 + SD_M^2$) whereas variance within an individual, SD_{within}^2, is influenced only by environmental factors and measurement errors ($SD_{within}^2 = SD_E^2 + SD_M^2$). Thus, the genetic component (r_{GC}) contributing to variation in a pharmacokinetic parameter can be estimated as,

$$(1) \qquad r_{GC} = (SD_{between}^2 - SD_{within}^2) / SD_{between}^2$$

Kalow *et al.* applied the above formula to calculate the r_{GC} of metabolic clearance of various drugs using published data and showed that the equation correctly predicts the genetic component contributing to variation in metabolic clearance (*e.g.*, r_{GC} values calculated from RDA studies agreed with independent estimates of heritability from twin studies). We have used this formula to calculate the r_{GC} of renal clearance to better understand the contribution of genetics to variation in renal clearance.

METHODS

Drugs chosen for analysis: Rationale

According to the well-stirred model of clearance, renal clearance is dependent upon renal blood flow (Q_k), plasma protein binding (fu), and intrinsic renal clearance ($CL_{i,\,renal}$) as shown in equation 2 [5].

$$(2) \qquad CL_{renal} = Q_k*(fu*CL_{i,renal})/(Q_k + fu*CL_{i,renal})$$

Intrinsic renal clearance, $CL_{i,\,renal}$, is the sum of clearance by filtration and secretion, minus clearance by reabsorption.

$$(3) \qquad CL_{i,\,renal} = CL_{GFR} + CL_{secretion} - CL_{reabsorption}$$

When intrinsic renal clearance ($CL_{i,\,renal}$) is very large, renal clearance will approximate renal blood flow. Alternatively, when intrinsic renal clearance is very small ($fu*CL_{i,renal} << Q_k$), renal clearance is proportional to $CL_{i,\,renal}$ (*e.g.* $CL_{renal} = fu*CL_{i,\,renal}$).

We estimated the r_{GC} of renal clearance for five drugs spanning a wide range of renal clearance values: 1) terodiline, a hydrophobic drug that has a renal clearance much less than that of GFR and undergoes passive reabsorption; 2&3) amoxicillin and ampicillin, organic anions that undergo net secretion via a transporter-mediated mechanism; 4) metformin, an organic cation that undergoes extensive secretion and is also likely secreted via a transporter-mediated mechanism; and 5) PAH (para-aminohippurate), a compound whose renal clearance is limited by blood flow. Variance data for overall renal clearance, between and within individuals, was obtained from four independent studies and r_{GC} was calculated according to equation 1 [7-10]. When

individual data for glomerular filtration rate (*e.g.* creatinine clearance) were available, net secretion clearance values were obtained by subtracting creatinine clearance from the total renal clearance, and r_{GC} of secretion clearance was calculated (Table 1).

Statistical Analysis

To determine whether there were statistically significant differences in the $r_{GC, renal}$ values of metformin, amoxicillin, and ampicillin (drugs whose renal clearance is dependent on active secretion) compared to PAH and terodiline (drugs whose renal clearance is independent of active secretion), a bootstrap analysis was conducted. In each iteration of the bootstrap analysis, heritability ($r_{GC, renal}$) was calculated according to equation 1 for each of the five drugs by randomly choosing n subjects from each study population of n individuals (*e.g.*, n = 12 for the amoxicillin and ampicillin studies). Following this, the $r_{GC, renal}$ values were averaged over each of the two groups. Thus, with each iteration of the bootstrap calculation, an average $r_{GC, renal}$ was obtained for group 1 drugs (metformin, amoxicillin, and ampicillin) and group 2 drugs (PAH and terodiline). One thousand iterations were performed. The Mann-Whitney rank sum test was then used to determine whether these two populations of values were significantly different. z_T values were computed for each group and compared to the normal distribution with an infinite number of degrees of freedom.

RESULTS AND DISCUSSION

The data in Table 1 show that metformin, amoxicillin, and ampicillin have r_{GC} renal clearance values ranging from 0.6 to 0.9, suggesting that there is a strong genetic component contributing to variation in the total renal clearance of these compounds. Since both glomerular filtration and net secretion play significant roles in the renal elimination of these drugs, these high r_{GC} values could be due to the genetic contribution of glomerular filtration and/or net secretion to variation in renal clearance. For both metformin and amoxicillin, r_{GC} values of net secretion are virtually unchanged compared to those of total renal clearance (decrease by less than 5%), suggesting that both total renal clearance and net secretory clearance of these drugs are largely influenced by genetic variation. In contrast, the r_{GC} of renal clearance of ampicillin is lower than that of both metformin and amoxicillin, suggesting that genetic factors contribute less to its renal clearance than to those of metformin and amoxicillin. This observation also supports the notion that genetic variation plays a smaller role in the net secretory clearance of ampicillin compared to amoxicillin and metformin.

Recent studies in heterologous expression systems have shown that β-lactam antibiotics, such as ampicillin and amoxicillin, interact with an organic anion transporter (OAT1) on the basolateral membrane of renal proximal tubule cells as well as a secretory inorganic phosphate transporter (Npt1) and an absorptive peptide transporter (PEPT2) on the luminal membrane [6, 11, 12]. These observations suggest that ampicillin and amoxicillin have components of both active secretion and active reabsorption, with the secretory component being larger than that of the reabsorptive component. Both drugs have very similar chemical structures and *in vivo* pharmacokinetics ($CL_{renal,\ amoxicillin} \sim CL$

$_{renal,ampicillin}$) [8]. However, amoxicillin has been shown to have a stronger affinity for the peptide transporters, PEPT1 and PEPT2; thus, its renal clearance may be more sensitive to alterations in renal transporter function [12]. Such alterations in function could be caused by variation in transporter genes and may explain the higher r_{GC} renal secretion value of amoxicillin.

In contrast to metformin, amoxicillin, and ampicillin, the r_{GC} renal clearance values for PAH and terodiline were considerably lower than 0.5, suggesting that genetic factors do not substantially contribute to variation in the renal clearance of these drugs. PAH undergoes extensive renal secretion; however, unlike metformin, amoxicillin and ampicillin, the renal clearance of PAH is thought to be limited by renal blood flow [7]. Therefore, the low r_{GC} value of PAH suggests that variation in renal blood flow may not have a strong genetic component. Terodiline undergoes extensive passive renal reabsorption. The low r_{GC} value of terodiline (0.37) suggests that variation in passive renal reabsorption does not have a strong genetic component. Environmental factors which affect urine pH and urine flow rate may produce substantial variation in passive reabsorption. Hallen *et al.* demonstrated that there was indeed a strong correlation between urine pH and renal reabsorption ($r^2 = 0.83$) which contributed to variation in terodiline renal clearance within and between individuals. A less significant correlation between urine flow rate and renal clearance was also observed [9].

We were interested in whether the genetic contribution to renal clearance ($r_{GC,renal}$) was significantly greater for drugs whose renal clearance was dependent upon active secretion (*i.e.*, metformin, amoxicillin, and ampicillin) in comparison to drugs whose renal clearance was independent of secretion (*i.e.*, dependent on reabsorption or blood

flow). Therefore, we compared $r_{GC,\,renal}$ of metformin, amoxicillin, and ampicillin to $r_{GC,renal}$ of PAH and terodiline using a bootstrap analysis followed by the Mann-Whitney rank sum test (see Methods). We found that $r_{GC,\,renal}$ of metformin, amoxicillin, and ampicillin was significantly larger than the $r_{GC,\,renal}$ of PAH and terodiline ($T_{group\,1}$ = 1,481,993, $T_{group\,2}$ = 517,224, z_T = 37.3, and $p<0.001$). This finding suggests that there is a difference in the genetic component contributing to the variation in renal clearance of these two groups of drugs. That is, renal clearance by secretion may be substantially more heritable than renal clearance controlled by passive reabsorption or limited by renal blood flow.

In addition to the above comparison, we were also interested in comparing the heritability of renal clearance to the heritability of metabolic clearance. Table 2 lists the genetic components (r_{GC}) contributing to the metabolic ratios and rates of metabolism for various drugs as tabulated and reported by Kalow *et al.* With the exception of the r_{GC} value for xanthine oxidase, all of the r_{GC} values for metabolism are greater than 0.50 and fall within the range calculated for metformin, amoxicillin, and ampicillin (0.64-0.94), suggesting that the heritability of renal clearance by filtration and secretion is similar to the heritability of metabolism. Moreover, like metabolic clearance, variation in net renal secretion may be due to genetic variation in the proteins controlling this process.

The RDA method was shown to predict the genetic component of metabolic clearance of various drugs. However, Kalow *et al.* pointed out some assumptions and limitations associated with this method which need to be discussed in regards to the studies presented here. First, a given genetic contribution (r_{GC}) applies only to the population in which it was established; for example, r_{GC} will likely be much higher in a

Table 1: Genetic component contributing to variation in net renal clearance and renal secretion of a variety of drugs

Drug	N[a]	CL_{renal}[b]	$SD^2_{between}$[c] (Total)	$SD^2_{between}$[d] (Secretion)	SD^2_{within}	r_{GC}[e] (Total)	r_{GC}[f] (Secretion)	Ref.
Metformin	3	450	5343	3953	299	0.94	0.92	[10]
Amoxicillin	12	154	476	317	44	0.91	0.86	[8]
Ampicillin	12	165	919	598	334	0.64	0.44	[8]
PAH	8	578	8117	---	6582	0.19	N/A	[7]
Terodiline	8	11.3	6.4	---	4	0.37	N/A	[9]

[a] N is the number of individuals included in each study.
[b] Mean renal clearance values in each study. Units of renal clearance are ml/min.
[c] Standard deviation in total renal clearance between individuals
[d] Standard deviation in net secretion clearance between individuals
[e] Genetic component contributing to variation in total renal clearance:
$(SD_{between}^2{}_{,total} - SD_{within}^2) / SD_{between}^2{}_{,total}$
[f] Genetic component contributing to variation in net secretion clearance:
$(SD_{between}^2{}_{,secretion} - SD_{within}^2) / SD_{between}^2{}_{,secretion}$

Table 2: Genetic components contributing to variation in metabolic ratios and clearance of various drugs as reported by Kalow *et al.*

Drug	Metabolic Ratio or Rate of Metabolism	N	Enzyme	r_{GC}	Reference
Caffeine[a]	(AAMU+1U+1X)/17X	12	CYP1A2	0.69	[13]
Caffeine	AAMU/(AAMU+1U+1X)	12	NAT2	0.95	[13]
Caffeine	1U/1X	12	XO[b]	0.24	[13]
Oxazepam	S-oxazepam/R-oxazepam	11	UGT2B7	0.98	[2]
Ethanol	Rate of metabolism (mg/kg/hr)	11	ADH	0.57	[14]
Dex[c]	Dex/Dextrorphan	7	CYP2D6	0.97	[15]

[a] Caffeine metabolite abbreviations are as follows: AAMU, 5-acetylamino-6-amino-3-methyluracil; 1U, 1-methylurate; 1X, 1-methylxanthine; 17X, 1,7-dimethylxanthine
[b] XO, xanthine oxidase
[c] Dex, dextromethorphan

healthy population compared to that of a diseased population [2, 3]. All of the studies discussed here included only healthy individuals, as judged by physical examinations and laboratory tests. The amoxicillin/ampicillin study was unique because it used a healthy, elderly population (age 69-83) with a ~30% decrease in GFR. Nevertheless, the pharmacokinetics of both amoxicillin and ampicillin in this elderly population were similar to those of a younger population [8].

Another limitation of the RDA methodology is that it relies on uniformity of environmental factors within and between individuals. This limitation can be avoided by controlling environmental factors such as food and drug intake [2, 3]. With the exception of the metformin study, no medications or alcohol were allowed at least 72 hours before and during the studies presented here; additionally, volunteers were subject to an overnight fast and given controlled meals and water during the study. It is possible that the high r_{GC} value obtained from the metformin study could have been caused by an inflated interindividual variance due to environmental factors. However, even if V_{inter} for metformin was inflated by 200%, the r_{GC} would remain greater than 0.80, suggesting a significant contribution of genetics to metformin renal clearance. Another variable that may inflate interindividual variation is the time interval between the two doses [2, 3]. Larger time intervals could lead to a greater possibility of an increased V_{inter}. The time intervals between dosing in the studies discussed here ranged from one to four weeks. It is not expected that such small intervals would substantially affect interindividual variance (V_{inter}).

CONCLUSIONS

Using published RDA data, we have demonstrated that genetic factors contribute to variation in renal clearance, just as they contribute to variation in metabolic clearance. Although the analysis presented here was compiled from only four studies, the trends are strong and suggest that variation in secretion clearance has a large genetic component. In contrast, variation in the renal clearance of drugs limited by renal blood flow or of drugs undergoing passive reabsorption may be primarily due to environmental factors. Active secretion is mediated by transporters in the proximal tubules of the kidney, for example, organic cation transporter (OCTs), organic anion transporters (OATs), and multidrug resistance-associated proteins (MRPs) [6]. These transporters have diverse substrate specificities and interact with many clinically used drugs. Metformin, an organic cation drug, is a known substrate of the kidney-specific organic cation transporter, OCT2. Similarly, anionic β-lactam antibiotics, such as amoxicillin and ampicillin, are thought to interact with multiple transporters in the kidney (OAT1, OAT3, PEPT2, and Npt1) [6, 11, 12]. It is likely that genetic variation in these transporters contributes to the high r_{GC} values observed for metformin, amoxicillin, and ampicillin. Identifying genetic variants of renal drug transporter genes, for example, OCT2, OAT1 and OAT3, may ultimately reveal the molecular mechanisms by which heritability of renal secretion is determined.

REFERENCES

1. Vesell ES. Genetic and environmental factors causing variation in drug response. *Mutat Res* **247**(2): 241-57, 1991.
2. Kalow W, Tang BK and Endrenyi L. Hypothesis: comparisons of inter- and intra-individual variations can substitute for twin studies in drug research. *Pharmacogenetics* **8**(4): 283-9, 1998.
3. Kalow W, Endrenyi L and Tang B. Repeat administration of drugs as a means to assess the genetic component in pharmacological variability. *Pharmacology* **58**(6): 281-4, 1999.
4. Bendayan R. Renal drug transport: a review. *Pharmacotherapy* **16**(6): 971-85, 1996.
5. Roland M and Tozer TN. *Clinical Pharmacokinetics: Concepts and Applications*. Williams & Wilkins, Media, 1995.
6. Dresser MJ, Leabman MK and Giacomini KM. Transporters involved in the elimination of drugs in the kidney: organic anion transporters and organic cation transporters. *J Pharm Sci* **90**(4): 397-421, 2001.
7. Hirata-Dulas CA, Awni WM, Matzke GR, Halstenson CE and Guay DR. Evaluation of two intravenous single-bolus methods for measuring effective renal plasma flow. *Am J Kidney Dis* **23**(3): 374-81, 1994.
8. Sjovall J, Alvan G and Huitfeldt B. Intra- and inter-individual variation in pharmacokinetics of intravenously infused amoxycillin and ampicillin to elderly volunteers. *Br J Clin Pharmacol* **21**(2): 171-81, 1986.

9. Hallen B, Guilbaud O, Stromberg S and Lindeke B. Single-dose pharmacokinetics of terodiline, including a stable isotope technique for improvement of statistical evaluations. *Biopharm Drug Dispos* **9**(3): 229-50, 1988.

10. Pentikainen PJ, Neuvonen PJ and Penttila A. Pharmacokinetics of metformin after intravenous and oral administration to man. *Eur J Clin Pharmacol* **16**(3): 195-202, 1979.

11. Uchino H, Tamai I, Yabuuchi H, China K, Miyamoto K, Takeda E and Tsuji A. Faropenem transport across the renal epithelial luminal membrane via inorganic phosphate transporter Npt1. *Antimicrob Agents Chemother* **44**(3): 574-7, 2000.

12. Terada T, Saito H, Mukai M and Inui K. Recognition of beta-lactam antibiotics by rat peptide transporters, PEPT1 and PEPT2, in LLC-PK1 cells. *Am J Physiol* **273**(5 Pt 2): F706-11, 1997.

13. Denaro CP, Wilson M, Jacob P, 3rd and Benowitz NL. Validation of urine caffeine metabolite ratios with use of stable isotope-labeled caffeine clearance. *Clin Pharmacol Ther* **59**(3): 284-96, 1996.

14. Kopun M and Propping P. The kinetics of ethanol absorption and elimination in twins and supplementary repetitive experiments in singleton subjects. *Eur J Clin Pharmacol* **11**(5): 337-44, 1977.

15. Ereshefsky L, Riesenman C and Lam YW. Antidepressant drug interactions and the cytochrome P450 system. The role of cytochrome P450 2D6. *Clin Pharmacokinet* **29 Suppl 1:** 10-8; discussion 18-9, 1995.

CHAPTER 4

NATURAL VARIATION IN MEMBRANE TRANSPORTER GENES REVEALS EVOLUTIONARY CONSTRAINTS

INTRODUCTION

With the completion of the draft sequence from the Human Genome Project and the development of high-throughput sequencing technology, several large-scale investigations of human sequence variation have been carried out [1-4]. These investigations have provided valuable information about the nature and frequency of sequence variation in the human genome. For example, studies by Cargill *et al.* and Halushka *et al.* identified differences in the level of genetic diversity among SNP types, such as coding and non-coding SNPs as well as synonymous and non-synonymous SNPs [2, 3]. More recently, patterns of haplotype diversity across the human genome have been characterized [1, 4]. To identify genetic variants, these studies have typically screened 24 to 40 chromosomes within an ethnic population; and therefore have identified common variants (frequencies greater than or equal to 5%) with high accuracy. However, the studies have not had the power to identify less common variants which may have larger deleterious effects. Further, the studies, to date, have sequenced genes from a wide variety of structural and functional classes; therefore, little is known about the relative levels of genetic diversity within classes of genes.

Membrane transporters play a critical role in a variety of physiological processes. They are responsible for maintaining cellular and organismal homeostasis by importing nutrients essential for cellular metabolism and exporting cellular waste products and

potentially toxic compounds. Furthermore, membrane transporters are important in drug response as they provide the targets for approximately 15-20% of commonly used drugs and are major determinants of drug absorption, distribution, and elimination. Membrane transport proteins share a similar secondary structure, characterized by multiple membrane-spanning domains joined by alternating intracellular and extracellular loops. Two of the major superfamilies of membrane transport proteins are the ABC (ATP-Binding Cassette) transporters, which include P-glycoprotein, a protein that pumps xenobiotics from cells, and the SLC (Solute Carrier) family, which includes neurotransmitter, nutrient, and heavy metal transporters.

In this study, we examined genetic diversity by screening for variation in a set of 24 genes encoding membrane transporters as part of a pharmacogenetics project that seeks to identify genes that determine drug response. We identified polymorphisms in this set of membrane transporter genes by screening an ethnically diverse collection of genomic DNA, 494 chromosomes in total. Sequencing this functionally and structurally similar class of proteins allowed us to determine the levels and patterns of genetic diversity across different structural regions of membrane transporters and in different transporter families. By sequencing an unusually large sample of 494 chromosomes, we obtained information about rare variants and greater statistical power to detect departures from the expected neutral frequency distribution of polymorphisms.

MATERIALS AND METHODS

Obtaining genomic sequences of 24 transporter genes

Starting with the mRNA accession number for a given gene, two versions of the data were retrieved from NCBI. The GenBank version provided information on the organism and chromosome. The FASTA version was used to extract the coding sequence of the gene. A BLAST search was done using the coding sequence on either the HTG or the NR databases at NCBI to find possible genomic sequences. Hits were checked to insure that they were from humans and were located on the proper chromosome. The hits were further checked to insure that they had splice site boundaries and to determine if exon coverage of the gene was complete. The final results related the exons to their location in the genomic sequence(s).

Variant identification

PCR amplification: PCR primers were designed using the Virtual Genome PCR primer selection website (http://alces.med.umn.edu/websub.html). Primers were made to specifically amplify each exon and a minimum of 35 bases of 5' and 3' flanking intronic sequence. For optimal DHPLC and DNA sequence analyses, amplicons were designed to be between 200 and 500 bases in length. Small, closely spaced exons were combined and analyzed in a single amplicon; large exons (>500 bases) were divided into smaller, overlapping amplicons. PCR conditions were optimized by amplifying each amplicon in six random DNA samples. AmpliTaq Gold® DNA Polymerase (Applied Biosystems) and a GeneAmp PCR System 9700 Thermal Cycler (Applied Biosystems) were used to carry out PCR. Standard PCR cycle conditions were: 94°C for 10 min, 35 cycles of 94°C

for 30 sec, 55°C for 30 sec and 72°C for 1 min, and a final extension step at 72°C for 10 min. Touchdown PCR cycle conditions were: 95°C for 10 min, 14 cycles of denaturating at 94°C for 20 sec, annealing at 63-56°C for 20 sec, decreasing by 0.5°C per cycle, and extending at 72°C for 45 sec., 16 standard cycles at 94°C for 20 sec, 56°C for 20 sec and 72°C for 45 sec, and a final extension at 72°C for 10 min. PCR fidelity and common variants were determined by DNA sequence analysis of the six random DNA samples.

DHPLC: Denaturing high-performance liquid chromatography (DHPLC) was used to survey the experimental sample population when an amplicon demonstrated fewer than 4 variant chromosomes in the six random test samples. PCR amplicons from 3 individuals were pooled and heteroduplexed by denaturing at 95°C for 8 min followed by annealing cycles from 94-25°C stepping down 0.5°C for 19 sec each cycle. Five microliters of each heteroduplexed sample pool were analyzed using a HELIX DHPLC System (Varian Inc). The column temperature for analysis was estimated for each amplicon sequence using the Stanford DHPLC melt program [http://insertion.stanford.edu/melt.html]. The gradient conditions were estimated by amplicon size. The final analysis conditions were determined empirically by varying conditions within the estimated temperature and gradient range on a standard control pool sample. The experimental sample pools were then analyzed at two selected temperatures, the elution profiles were scored by visual inspection, and the three samples in a pool with variant peaks were sequenced. Homoduplex pools were inferred to be reference sequence.

Sequencing: Both the forward and reverse DNA sequences of purified PCR products were determined using ABI PRISM BigDye terminator cycle sequencing Version 2.0 and an ABI Prism 3700 DNA analyzer. DNA sequence files (.ab1) from all samples were imported into Sequencher (Gene Codes Corp.), and aligned with the amplicon reference sequence. Heterozygous variants were identified in aligned sequences and scored in Sequencher using IUPAC/IUB ambiguity codes.

Population genetic statistics

The neutral parameter (θ), nucleotide diversity (π), and Tajima's *D* statistic were calculated as described by Tajima and as described in Chapter 2 [5]. Each parameter was calculated for various gene regions (*e.g.*, coding, non-coding, and intron-exon boundaries) as well as for various sites within the coding region (*e.g.*, synonymous and non-synonymous sites, evolutionarily conserved and unconserved sites, and sites within predicted transmembrane domains and loops). Synonymous and non-synonymous sites were calculated as described previously [6]. Evolutionarily conserved and unconserved amino acid residues were classified based on sequence alignments with mammalian homologs. Protein sequences of mammalian homologs were obtained through a series of database searches. Results were checked by hand to assure that only family members and not similar proteins were included in the dataset. At least three protein sequences were necessary for the alignments. Alignments were generated using the GCG program Pileup. Transmembrane domain and loop regions were classified based on topology data in the literature. If such data was not available for a particular transporter, topology data

from the SwissProtein database was used. These alignments were used in calculation of population genetic parameters for different regions of the proteins.

In the analysis comparing allele frequency distributions of evolutionarily conserved (EC) and evolutionarily unconserved (EU) non-synonymous SNPs, we assigned non-synonymous SNPs as EC or EU based on alignments with 2 orthologs (rat, mouse or rabbit). We were able to assign 118 of the 155 non-synonymous SNPs as EC or EU using this procedure.

Data Analysis

The observed allele frequency distribution of non-coding, synonymous, and non-synonymous SNPs was compared to the distribution expected under the neutral mutation model. According to the neutral mutation model, the expected number of SNPs (G_n) with a frequency of i/n can be obtained from the following equation:

$$G_n(i) = \theta[1/i + 1/(n-i)]$$

where θ is the neutral parameter, n is the number of chromosomes sequenced, and i/n is the expected allele frequency. The expected and observed allele frequency distributions were compared using a chi-square test of binned data. Several different binnings were used, and qualitative results (*i.e.*, our assessment of significance) did not depend on binning.

RESULTS

Sequence variation in 24 membrane transporter genes

To identify polymorphisms in the set of 24 membrane transporter genes, we screened all exons as well as 50-100 base pairs of flanking intronic sequence in an ethnically diverse collection of genomic DNA using denaturing HPLC analysis followed by direct sequencing. The collection of DNA used for screening included samples from 247 unrelated individuals: 100 Caucasians, 100 African Americans, 30 Asians, 14 Mexicans, and 7 Pacific Islanders. We identified 680 SNPs in almost 96 kb of genomic sequence (Table 1*a*). Of these, 330 were located in the 55 kb of coding sequence and 350 were located in the 41 kb of non-coding sequence. Within the coding region, 175 of the SNPs were synonymous and 155 of the SNPs were non-synonymous. One of the 155 non-synonymous SNPs changed an amino acid to a termination codon. Only 12 of the non-coding SNPs were located within the intron-exon boundaries (the first 8 bases and last 3 bases at the 5' and 3' ends, respectively, of the introns). In addition to these biallelic SNPs, two tri-allelic sites were identified in the coding region.

Due to the large amount of genomic sequence screened, we were also able to identify a significant number of insertion/deletion polymorphisms, which occurred at a frequency of ca. 1/25 that of SNPs (29/680) (Table 2). Eight of the insertion/deletion mutations were in the coding regions and 21 were in the non-coding regions. Of the eight coding region insertions/deletions, three resulted in a frameshift; one added an amino acid; one deleted an amino acid; and three added or deleted two or more amino acids. The deficit of insertion/deletions in the coding region compared to the non-coding region suggests selection against insertions and deletions within the coding region. For 15 of the

Table 1. Summary statistics of variation in transporter genes

a.

Sequence Section	bp	SNPs	θ^1	π	D
Total Population:					
Total	96074	680	10.44 ± 2.01	5.09 ± 2.44	-1.56
Coding	54700	330	8.90 ± 1.75	3.96 ± 1.94	-1.67
Non-Coding	41374	350	12.48 ± 2.45	6.57 ± 3.19	-1.43
Intron-Exon Boundary	4437	12	3.99 ± 1.38	2.20 ± 1.66	-0.99
Synonymous	12820	175	20.14 ± 4.10	9.73 ± 4.86	-1.53
Non-synonymous	41880	155	5.46 ± 1.12	2.20 ± 1.12	-1.77
AA Population:					
Total	96074	527	9.34 ± 2.08	5.47 ± 2.63	-1.33
Coding	54700	260	8.09 ± 1.83	4.38 ± 2.14	-1.46
Non-Coding	41374	267	10.99 ± 2.49	6.90 ± 3.36	-1.18
Intron-Exon Boundary	4437	9	3.45 ± 1.38	2.03 ± 1.57	-0.94
Synonymous	12820	142	18.86 ± 4.41	11.00 ± 5.48	-1.31
Non-synonymous	41880	118	4.80 ± 1.14	2.35 ± 1.20	-1.59
CA Population:					
Total	96074	299	5.30 ± 1.20	4.46 ± 2.15	-0.50
Coding	54700	135	4.20 ± 0.99	3.29 ± 1.62	-0.68
Non-Coding	41374	164	6.75 ± 1.56	6.02 ± 2.93	-0.34
Intron-Exon Boundary	4437	9	3.45 ± 1.38	2.31 ± 1.72	-0.75
Synonymous	12820	71	9.43 ± 2.34	7.71 ± 3.92	-0.55
Non-synonymous	41880	64	2.60 ± 0.65	1.93 ± 1.00	-0.78
AS Population:					
Total	96074	212	4.73 ± 1.33	3.77 ± 1.84	-0.72
Coding	54700	93	3.65 ± 1.07	3.09 ± 1.55	-0.53
Non-Coding	41374	119	6.17 ± 1.78	4.66 ± 2.32	-0.85
Intron-Exon Boundary	4437	2	0.97 ± 0.73	1.63 ± 1.37	1.17
Synonymous	12820	53	8.87 ± 2.71	7.64 ± 3.93	-0.47
Non-synonymous	41880	40	2.05 ± 0.65	1.70 ± 0.90	-0.56

Table 1. Summary statistics of variation in transporter genes (continued)

b.

Gene Region	bp	SNPs	θ^1	π	*D*
Total Population:					
Total	96074	680	10.44 ± 2.01	5.09 ± 2.44	-1.56
Coding	54700	330	8.90 ± 1.75	3.96 ± 1.94	-1.67
Conserved[2] (total)	33835	160	6.98 ± 1.43	3.39 ± 1.70	-1.52
Conserved (synonymous)	8052	96	17.59 ± 3.78	11.39 ± 5.80	-1.02
Conserved (non-synonymous)	25783	64	3.66 ± 0.83	0.90 ± 0.54	-2.13
Unconserved (total)	10983	101	13.57 ± 2.90	6.71 ± 3.48	-1.47
Unconserved (synonymous)	2497	41	24.23 ± 5.94	6.50 ± 4.26	-1.99
Unconserved (non-synonymous)	8486	60	10.43 ± 2.39	6.77 ± 3.59	-0.99
Loop (total)	36904	235	9.39 ± 1.88	4.33 ± 2.14	-1.61
Loop (synonymous)	8291	119	21.17 ± 4.45	10.04 ± 5.15	-1.54
Loop (non-synonymous)	28613	116	5.98 ± 1.26	2.68 ± 1.38	-1.62
Loop (non-synonymous, conserved)	16554	43	3.83 ± 0.93	1.16 ± 0.73	-1.91
Loop (non-synonymous, unconserved)	6329	47	10.95 ± 2.62	7.32 ± 3.98	-0.91
TMD (total)	17796	95	7.88 ± 1.69	3.20 ± 1.70	-1.72
TMD (synonymous)	4530	56	18.24 ± 4.22	9.17 ± 5.05	-1.39
TMD (non-synonymous)	13266	39	4.34 ± 1.07	1.16 ± 0.77	-1.98
TMD (non-synonymous, conserved)	9229	21	3.36 ± 0.97	0.43 ± 0.44	-2.17
TMD (non-synonymous, unconserved)	2157	13	8.89 ± 2.98	5.16 ± 3.73	-0.95

[1] Values of θ, π, and Tajima's *D* are listed as mean x 10^4 ± standard deviation.

[2] Conserved and unconserved regions are defined based on protein sequence alignments with at least two other mammalian orthologs and therefore, refer to evolutionary conservation.

Table 2. Insertion and deletion mutations in transporter genes

Gene	Exon[1]	Freq.	# bp	Sequence Containing Ins/del[2]	AA Change[3]	Basis for Ins/del
Coding insertions:						
OCT2	1.2	0.002	1	GGGCT**A**TCACCC	frameshift	
NRAMP1	13.1	0.002	6	CC**GTTCGC**GTTCGCC	+VR	$(GTTCGC)_1 \rightarrow (GTTCGC)_2$
CNT1	4.4	0.291	3	TTGGGC**TTG**TCAAGCC	+V	
Non-coding insertions:						
FIC1	2.1	0.002	1	CCC(**T**)$_8$AAAGAA	-	$T_8 \rightarrow T_9$
GAT1	9.3	0.002	1	(**C**)$_8$GCACCCC	-	$C_8 \rightarrow C_9$
IREG1	1.2	0.002	1	TGACCG**G**AAAGCAT	-	
MRP1	6.2	0.004	1	AGG(**C**)$_6$AGACCTC	-	$C_6 \rightarrow C_7$
OCT2	11.3	0.012	1	AGGG(**T**)$_8$CCATCTT	-	$T_8 \rightarrow T_9$
BSEP	19.1	0.039	3	CTATTT(ATA)$_2$**ATA**AAG	-	$(ATA)_2 \rightarrow (ATA)_3$
FIC1	19.2	0.089	5	(ATAAA)$_3$**ATAAA**AACAGCA	-	$(ATAAA)_3 \rightarrow (ATAAA)_4$
VMAT1	3.1	0.16	1	TTTCAA**A**TTCCCTT	-	$A_2 \rightarrow A_3$ (?)
SPNT1	1.1	0.19	1	CTTGGCC**C**TGAATT	-	
Coding deletions:						
ENT2	6.1	0.002	6	CCGCAG**GTGGCG**TGGACG	-SGV, +M	
ENT2	8.3	0.002	2	AACCTT**CA**GTCTTC	frameshift	
NRAMP1	2.3	0.002	9	CAGCCC**GACCAGCCC**AGGG	-TSP	Recombination between repeats
CNT1	11.3	0.012	1	AGCTG**G**TCTAC	frameshift	
OCT1	7.5	0.105	3	CTCGTC**ATG**ATT	-M	
Non-coding deletions:						
IREG1	3.1	0.002	3	AGTGT**TGT**TATATAATTA	-	
MDR3	25.1	0.002	1	AGTT(**A**)$_3$GAGTAG	-	$A_3 \rightarrow A_2$
MRP1	27.1	0.002	11	TCCTTA**CTCTCTCCCTT**CACT	-	
NET	13.4	0.004	6	GGGGGA**CAGGGA**GG	-	Recombination between repeats
BSEP	2.1	0.012	1	GCGTTG**A**TTTTTT	-	
MRP1	19.1	0.033	2	TCACAC**AT**GTGCAC	-	
PEPT1	4.4	0.085	1	CATAAC**C**AGTCCT	-	
GAT1	9.2	0.098	1	AGGGCG(**C**)$_8$GCAC4	-	$C_8 \rightarrow C_7$
OCT2	5.4	0.11	1	CAAATCA(**G**)$_4$ATG	-	$G_4 \rightarrow G_3$
OCT2	5.3	0.118	1	CAAATC**A**GGGGAT	-	
GAT1	9.1	0.155	1	AGGGCG(**C**)$_8$GCAC4	-	$C_8 \rightarrow C_7$
OCT1	7.8	0.706	8	GTAAGTTG**GTAAGTTG**TCTG	-	Recombination between repeats

[1]Numbers in the exon column refer to the exon containing the indel and the variant number in that particular exon (*i.e.*, 1.3 is the third variant in exon 1).
[2]Inserted and deleted nucleotides are shown in bold. Underlined sequence indicates repeat units that may have been subject to recombination.
[3]Insertions of an amino acid are denoted by a "+" followed by the inserted amino acid(s); similarly deletions are denoted by a "-" followed by the deleted amino acid(s).

29 insertion/deletion mutations, plausible proposals for the origin (occurrence at single nucleotide repeats or at short tandem duplications) can be made (Table 2).

To determine the false positive rate of our polymorphism identification method (denaturing HPLC followed by direct sequencing), we resequenced all 314 singleton SNPs. Approximately 95% (297/314) of the singletons were verified upon resequencing, indicating a 5% false positive rate. To determine the false negative rate of our two-step screening procedure, we resequenced 14 amplicons (4923 bp) in all 247 DNA samples. We identified 11 new SNPs in 1.2×10^6 bp of resequenced DNA, leading to an average of 9.0×10^{-6} SNPs missed per base pair screened. All of the missed SNPs were very rare: eight were singletons, two were doubletons, and one was a tripleton. Our false negative rate is approximatey two-fold lower than that reported by Cargill *et al.* who resequenced 10 genes (20475 bp) in 20 DNA samples and on average missed 17.1×10^{-6} SNPs per base pair screened.

Deviations for a particular allele from Hardy-Weinberg equilibrium in an ethnic population for which many alleles conform to Hardy-Weinberg equilibrium may result from sequencing errors. Of the 680 SNPs, we estimate that 73 (32, 22, and 19 in the African American, Caucasian, and Asian populations, respectively) failed to conform to Hardy-Weinberg equilibrium within an ethnic population ($p < 0.000046$, corrected for 1080 tests performed).

To further quantify the amount of variation in our set of genes, we calculated two measures of nucleotide diversity, the average heterozygosity (π) and the population mutation parameter (θ). In addition, Tajima's *D* was calculated to detect deviations from the neutral mutation model. These parameters were calculated for the entire region

sequenced as well as for each SNP type (*i.e.*, coding versus non-coding and synonymous versus non-synonymous). For the entire set of genes, θ ($\times 10^4$) and π ($\times 10^4$) were 10.4 ± 2.0 and 5.1 ± 2.4, respectively (Table 1*a*), which are similar to values previously reported for other genes (θ and π ranging from 5.0 to 8.3 $\times 10^{-4}$), suggesting that, on average, genetic variation in membrane transporters is similar to that in other genes [2-4]. As in previous reports, we found that non-coding regions have greater levels of nucleotide diversity (π) than coding regions (6.57 ± 3.19 and 3.96 ± 1.94; $p<0.05$), and within the coding region, synonymous sites have greater levels of nucleotide diversity than non-synonymous sites (9.73 ± 4.86 and 2.20 ± 1.12; $p<0.05$). In Table 3, we show the values of θ and π for each gene. θ (x 10^4) ranged from 1.62 to 21.65 and π (x 10^4) ranged from 0.65 to 13.95. All transporter genes had negative Tajima's *D* values with the exception of PEPT2.

The ratio of π for non-synonymous sites (π_{NS}) to π for synonymous sites (π_S) provides a measure of selection for function of a given gene [7]. We have calculated π_{NS}/π_S for each of the transporter genes and compared them with the ratios previously determined for VMAT2, SERT and TLR4 (Table 3) [8, 9]. Eleven genes exhibited low π_{NS}/π_S ratios (<0.1), indicating that mutations that change amino acids are rapidly eliminated from the population relative to synonymous changes. Eleven genes had ratios between 0.1 and 0.5, suggesting that they are under some selection. Strikingly, VMAT1 had a ratio of 1.7, suggesting that it may be under positive selection, implying that amino acid diversity is advantageous for VMAT1. It is notable that some genes (ENT1, FIC1, and SERT) exhibit low π_S values (<1), suggesting that either these genes have intrinsically reduced mutability or there was some other cause for elimination of

Table 3. Population genetic parameters of transporter genes

Gene	HGNC Symbol	θ_{Total} [1]	π_{Total}	Tajima's D_{Total}	π_S	π_{NS}	π_{NS}/π_S
DAT	SLC6A3	12.11	7.61	-1.19	14.29	0.00	0.0000
NRAMP1	SLC11A1	9.27	4.04	-1.47	8.47	0.07	0.0077
MRP1	ABCC1	9.91	4.07	-2.63	13.61	0.15	0.0110
OCT3	SLC22A3	7.40	5.04	-0.69	19.92	0.23	0.0115
NET	SLC6A2	10.75	4.27	-1.89	8.58	0.11	0.0130
GAT1	SLC6A1	8.31	1.06	-2.21	2.21	0.03	0.0132
VMAT2	SLC18A2	n/a	n/a	n/a	5.58	0.17	0.0305
DMT1	SLC11A2	11.19	4.52	-1.74	4.97	0.19	0.0381
ENT2	SLC29A2	6.34	1.44	-1.37	2.00	0.13	0.0663
IREG1	SLC11A3	6.12	4.40	-0.50	7.67	0.76	0.0989
OCT2	SLC22A2	11.15	5.74	-1.42	22.54	2.23	0.0990
TLR4	TLR4	n/a	n/a	n/a	10.10 [2]	1.30	0.1287
MDR3	ABCB4	8.03	3.55	-2.00	9.60	1.43	0.1486
BSEP	ABCB11	8.93	6.99	-0.85	12.62	2.30	0.1821
MDR1	ABCB1	9.39	5.64	-1.75	10.68	2.37	0.2218
SERT	SLC6A4	n/a	n/a	n/a	0.77	0.18	0.2338
MRP2	ABCC2	10.14	3.26	-3.23	8.81	2.17	0.2466
CNT1	SLC28A1	21.65	10.90	-2.39	22.52	6.20	0.2753
PEPT1	SLC15A1	12.13	4.09	-2.54	9.74	3.66	0.3762
PEPT2	SLC15A2	9.65	13.95	1.43	19.66	8.73	0.4443
OCT1	SLC22A1	21.42	8.56	-2.43	11.20	5.11	0.4563
ENT1	SLC29A1	7.69	0.89	-1.93	0.48	0.23	0.4813
VACHT	SLC18A3	9.01	3.73	-1.10	4.46	2.21	0.4963
FIC1	ATP8B1	10.19	2.99	-2.98	0.86	0.84	0.9732
SPNT1	SLC28A2	8.30	6.37	-0.66	7.61	7.64	1.0044
VMAT1	SLC18A1	18.98	6.54	-2.69	4.25	7.08	1.6650
CTR1	SLC31A1	1.62	0.65	-0.22	0.00	1.31	n/a
Superfamilies:							
ABC		9.35	4.66	-1.50	11.08	1.66	0.1495
SLC		11.17	5.61	-1.50	9.84	2.75	0.2794

[1] Values of π and θ are listed (x 10^{-4}).
[2] π value for TLR4 is based on synonymous and intronic SNPs.

silent variation, such as a recent selective sweep. Differences in mutability are plausible, perhaps caused by different chromosomal domains or replicons that are somehow protected from mutagenesis. We note that genes that are adjacent to each other (MDR1 and MDR3) have similar π_S values (π_S x 10^4 is 10.68 and 9.60, respectively) and that the members of the OCT1-OCT2-OCT3 cluster also have similar π_S values (π_S x 10^4 is 11.2, 22.5, and 19.9, respectively). Values of θ were greater than π for all gene regions, resulting in negative Tajima's *D* values for 23 of 24 genes and consistent with population expansion or negative (purifying) selection.

Within the membrane transport proteins that were analyzed, there were two major superfamilies, ABC transporters (5 members) and SLC transporters (18 members). Average nucleotide diversity of the ABC transporters was similar to that of the SLC transporters (Table 3). All of the ABC transporters had similar levels of overall diversity (as indicated by π_{Total}). Four of these transporters exhibited intermediate π_{NS}/π_S values, indicating similar levels of selection, whereas one (MRP1) had a very low ratio. Nucleotide diversity of SLC superfamily members exhibited wide variation, as did π_{NS}/π_S (Table 3). The three neurotransmitter transporters in the SLC6 family (GAT, NET, and DAT) exhibited very low π_{NS}/π_S ratios consistent with strong purifying selection. The heavy metal transporters in the SLC11 family (NRAMP1, DMT1, and IREG) exhibited typical π values and low π_{NS}/π_S. In contrast, the organic cation transporters in the SLC22 family (OCT1, OCT2, and OCT3), all of which exhibit typical π values have widely different π_{NS}/π_S ratios (0.46, 0.099, and 0.012, respectively). The two members of the vesicular monoamine transporter family (SLC18; VAChT and VMAT1) exhibited high

π_{NS}/π_S (0.5, 1.0, respectively), in contrast to the low π_{NS}/π_S previously reported for VMAT2 (0.03) [9].

We identified a total of 421 population-specific SNPs, of which 248 were singletons (Table 4). Of the 259 SNPs that were not population-specific, 83 were present in all five populations and 176 were present in two, three, or four populations. An examination of SNP sharing between pairs of populations revealed that each population – Caucasian, Asian, Mexican, and Pacific Islander – shared the most SNPs with the African American population. Because we screened an equal number of chromosomes in the African American and Caucasian populations, we could compare the levels of variation in these two population samples. θ and π were higher (and Tajima's *D* was more negative) over all gene regions in the African American sample compared to the Caucasian sample, consistent with African Americans descending from a larger, and/or more ancient population. In agreement with these results, the African American sample had a substantially larger number of population-specific SNPs – 278 compared to only 80 for Caucasians. One hundred thirty-three (48%) of the African American-specific SNPs were singletons (occurring on only one of 494 chromosomes) while 71 (89%) of the Caucasian-specific SNPs were singletons. These singletons reflect a mixture of new mutations that have not been subjected to selection and old mutations that have been subjected to selection. The ratio of the number of singletons in African Americans to that in Caucasians provides a measure of the distribution of new mutations (similar in both populations) plus old mutations, some of which are shared by both populations and others which are present only in African Americans. We observe a ratio of approximately 2:1, in contrast to Stephens *et al.*, who reported a ratio of 3:1 [4]. Because we identified rarer

Table 4. Population specific SNPs and shared SNPs

a.

Population Specific SNPs	# SNPs	# Singletons
African Americans	278	133
Caucasians	80	71
Other	63	44
Total	421	248

b.

Shared SNPs	# SNPs
All 5 Populations	83
2-4 Populations	176
Total	259

singletons than Stephens *et al.*, our ratio should lead to a better estimation of new mutations.

Nucleotide diversity across structural regions

Because membrane transporters have two distinct types of secondary structure, transmembrane domains (TMDs) and loops (see Fig. 1), we compared nucleotide diversity of these structural regions. Amino acid diversity (π_{NS}) in the TMDs was significantly lower than amino acid diversity in the loops ($\pi_{TMD\text{-}NS} = 1.16 \times 10^{-4}$ versus $\pi_{loop\text{-}NS} = 2.68 \times 10^{-4}$; $p<0.05$) (see Table 1*b* and Fig. 1). If one makes the generally accepted assumption that evolutionarily conserved (EC) residues have lower nucleotide diversity than evolutionarily unconserved (EU) residues, such a difference could result if TMD regions contain a greater proportion of evolutionarily conserved residues than do loops. For 20 of 24 transporters, we have identified at least two other mammalian species and used these comparisons for defining evolutionarily conserved and unconserved positions (see Materials and Methods). We observed that 83% of the residues in the TMDs and 74% of the residues in the loops are evolutionarily conserved. As noted above, in general, π_{NS} for evolutionarily unconserved (EU) residues is greater than π_{NS} for evolutionarily conserved (EC) residues. We observed this trend in the loops ($\pi_{NS\text{-}EU}$ and $\pi_{NS\text{-}EC}$ are 7.32 and 1.16×10^{-4}, respectively, in the loops) and in the TMDs ($\pi_{NS\text{-}EU}$ and $\pi_{NS\text{-}EC}$ are 5.16 and 0.43×10^{-4}, respectively, in the TMDs). These results suggest that population genetic analyses (in particular, amino acid diversity) parallel phylogenetic comparisons.

We next analyzed the nature of the changes in the TMDs and loop regions, on the basis of our vertebrate ortholog alignments and block analysis to define relationships

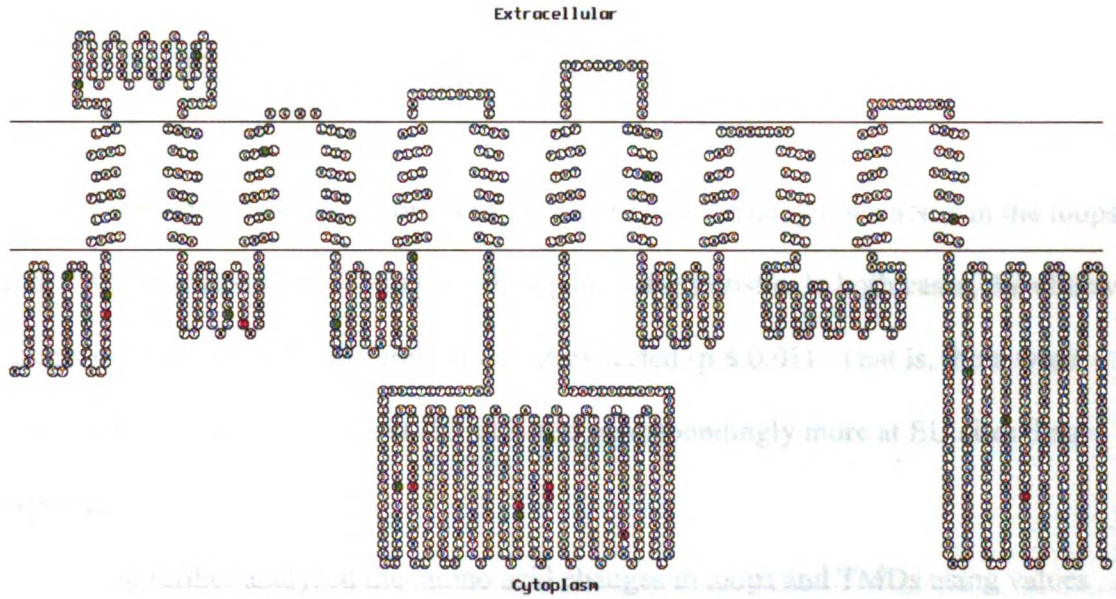


Figure 1. Predicted secondary structure of representative membrane transporter, BSEP, with coding region SNPs. The transmembrane topology schematic was rendered using TOPO (S.J. Johns [UCSF, San Francisco] and R.C. Speth [Washington State University, Pullman], transmembrane protein display software, available at the UCSF Sequence Analysis Consulting Group website, http://www.sacs.ucsf.edu/TOPO/topo.html). Non-synonymous amino acid changes are shown in red; synonymous changes in green.

among amino acids [10, 11]. We observed 86 non-synonymous changes in loops, of which 45 were at EC positions and 41 at EU positions, and 32 non-synonymous changes in the TMDs, of which 20 were at EC positions and 12 at EU positions. Expected counts of non-synonymous changes in the loops and TMDs were calculated based on the fraction of conserved residues in each region (74% and 83%, respectively). Differences in the observed and expected counts of EC and EU non-synonymous SNPs in the loops and TMDs were quantified using the chi-squared test statistic. In both cases, the observed counts were significantly different from the expected ($p \leq 0.01$). That is, there were fewer non-synonymous SNPs at EC sites and correspondingly more at EU sites than expected.

We further analyzed the amino acid changes in loops and TMDs using values generated by protein block matrices, the SLIM matrix for amino acid changes in the transmembrane domains and BLOSUM62 for the loops. The SLIM matrix is based on amino acid changes of membrane proteins: SLIM values less than 0 indicate changes that are considered evolutionarily unfavorable; SLIM values greater than or equal to 0 are considered evolutionarily favorable [10]. Within the TMDs, 28% of the amino acid changes had SLIM values less than 0 whereas 72% had SLIM values greater than or equal to 0, indicating that the majority of changes in the TMDs (72%) are evolutionarily favorable or tolerated. Thus, the SLIM values indicate that amino acid changes in TMDs occur at positions that are evolutionarily unconstrained. The BLOSUM62 matrix is based on amino acid changes in a structurally diverse protein set, most of which are globular [11]. BLOSUM62 values less than 0 are considered to be evolutionarily unfavorable whereas values greater or equal to 0 are considered favorable [2]. Within the

loops, 61% of amino acid changes had BLOSUM62 values less than 0 and 39% had values greater than or equal to 0, indicating that the majority of changes are evolutionarily unfavorable based on analysis of globular protein blocks. We suggest that the apparent high fraction of unfavorable amino acid changes that occur in loops predicted by BLOSUM62 reflects different structural and functional demands on loops in comparison to globular proteins. The EC versus EU criterion is thus more effective in predicting favorable and unfavorable amino acid changes in loops than BLOSUM62 values. It may be worth developing block matrices analogous to BLOSUM62, but tuned specifically for loop domains of membrane proteins.

Frequency distributions and evolutionary constraints of minor alleles

Because we screened 494 chromosomes, binomial sampling theory suggests that we have a ~99% chance of identifying SNPs that occur at a frequency of 1% in our total sample and an ~86% chance of identifying SNPs that occur at a frequency of 1% in our two largest ethnic samples (Caucasians and African Americans, each of 200 chromosomes). Of the 680 SNPs, 248 (36 %) were found only once in 494 chromosomes. Of these, 133 were specific to the African American population, 71 to the Caucasian sample, and 44 in the other samples (Table 4). Few population-specific alleles were found at high frequency. Only four of the 278 African American-specific alleles have a relative frequency greater than 0.10. The highest frequency population-specific allele was 0.287 in the African American population sample, 0.107 in the Asian sample, and 0.04 in the Caucasian sample. Eighty-three (12 %) SNPs were present in all populations and 176 (26 %) were found in two to four populations.

In Table 5*a*, we present the observed minor allele frequency distributions of non-coding, synonymous, and non-synonymous SNPs. Similar to previous reports, we observed a higher percentage of SNPs in the lowest allele frequency class and a lower percentage in the higher allele frequency classes [2, 3]. This trend was somewhat more pronounced for non-synonymous SNPs compared to synonymous and non-coding SNPs, consistent with the idea that selection acts more strongly on non-synonymous SNPs. However, when we tested the heterogeneity of the frequency distributions of synonymous SNPs versus non-synonymous SNPs, we found them to be homogeneous ($X_{7df}^2 = 7.24$, $p = 0.40$). Similarly, coding versus non-coding SNPs appeared to have similar frequency spectra in this sample ($X_{7df}^2 = 6.85$, $p = 0.44$). The overall trends in the allele frequency distributions of the African American and Caucasian samples (Table 5 *b, c*) were similar to those seen in the total sample. The allele frequency distribution of coding region SNPs in the Caucasian sample was significantly different from that of the African American sample ($X_{7df}^2 = 16.21$, $p = 0.0127$). In particular, the Caucasian distribution had an increased percentage of high-frequency SNPs and a decreased percentage of intermediate-frequency SNPs.

We compared the observed minor allele frequency distributions with that predicted under the infinite-sites, neutral model, which is based on the assumptions that all sites are mutable, alleles are lost through genetic drift but not by selection, and population size is fixed (Table 5*a*). Relative to this model, we observed a higher percentage of low frequency alleles (14.8% versus 34.3% for alleles with frequencies less than or equal to 0.2%) for the non-coding and synonymous sites, which are not expected to be under selection. We also observed a reduced percentage of high frequency alleles

Table 5. Observed and Expected Minor Allele Frequency Distributions

a. Total Population

Minor Allele Freq.	Expected[1]	NS[2]	NS	S	S	NC	NC	NC&S	NC&S
	%	*n*	%	*n*	%	*n*	%	*n*	%
p ≤0.002	**14.8**	64	**41.3**	62	**35.4**	118	**33.7**	180	**34.3**
0.002<p<0.005	**7.4**	14	**9.0**	21	**12.0**	38	**10.9**	59	**11.2**
0.005≤p<0.010	**8.6**	23	**14.8**	17	**9.7**	34	**9.7**	51	**9.7**
0.010≤p<0.020	**11.2**	18	**11.6**	18	**10.3**	34	**9.7**	52	**9.9**
0.020≤p<0.050	**14.4**	11	**7.1**	23	**13.1**	43	**12.3**	66	**12.6**
0.050≤p<0.100	**11.2**	6	**3.9**	10	**5.7**	18	**5.1**	28	**5.3**
0.100≤p<0.200	**11.9**	7	**4.5**	7	**4.0**	27	**7.7**	34	**6.5**
0.200≤p	**20.4**	12	**7.7**	17	**9.7**	38	**10.9**	55	**10.5**

b. African American Population

Minor Allele Freq.	Expected	NS	NS	S	S	NC	NC	NC&S	NC&S
	%	*n*	%	*n*	%	*n*	%	*n*	%
p ≤0.005	**17.1**	45	**38.1**	41	**28.9**	84	**31.5**	125	**30.6**
0.005<p≤0.010	**8.6**	12	**10.2**	16	**11.3**	25	**9.4**	41	**10.0**
0.010<p≤0.020	**10.1**	16	**13.6**	13	**9.2**	22	**8.2**	35	**8.6**
0.020<p≤0.050	**14.9**	16	**13.6**	31	**21.8**	39	**14.6**	70	**17.1**
0.050<p≤0.100	**12.3**	8	**6.8**	12	**8.5**	37	**13.9**	49	**12.0**
0.100<p≤0.200	**13.6**	12	**10.2**	12	**8.5**	23	**8.6**	35	**8.6**
0.200<p≤0.500	**23.2**	9	**7.6**	17	**12.0**	37	**13.9**	54	**13.2**

c. Caucasian Population

Minor Allele Freq.	Expected	NS	NS	S	S	NC	NC	NC&S	NC&S
	%	*n*	%	*n*	%	*n*	%	*n*	%
p ≤0.005	**17.1**	24	**38.7**	25	**35.7**	44	**27.2**	69	**29.7**
0.005<p≤0.010	**8.6**	7	**11.3**	7	**10.0**	12	**7.4**	19	**8.2**
0.010<p≤0.020	**10.1**	3	**4.8**	6	**8.6**	10	**6.2**	16	**6.9**
0.020<p≤0.050	**14.9**	4	**6.5**	5	**7.1**	18	**11.1**	23	**9.9**
0.050<p≤0.100	**12.3**	7	**11.3**	5	**7.1**	18	**11.1**	23	**9.9**
0.100<p≤0.200	**13.6**	7	**11.3**	7	**10.0**	20	**12.3**	27	**11.6**
0.200<p≤0.500	**23.2**	10	**16.1**	15	**21.4**	40	**24.7**	55	**23.7**

[1] The expected percentage of SNPs in each frequency class was predicted based on the infinite sites neutral mutation model. The remaining columns list the observed numbers (*n*) and observed percentages (%) of each SNP type for non-synonymous, synonymous, and non-coding SNPs as well as for non-coding and synonymous SNPs combined.
[2] Abbreviations are as follows: NS, non-synonymous; S, synonymous; NC, non-coding.

(43.5% versus 22.3% for alleles with frequencies greater than or equal to 5%). These observations are similar to those of Glatt *et al.* and can be explained by population expansion [9].

We identified 155 non-synonymous SNPs, which were distributed across various frequencies, and would ultimately like to know their effect on function. We first calculated the ratio of non-synonymous versus synonymous changes as a function of allele frequency (Table 6*a*). As observed by Fay *et al.*, we found that less common alleles (frequencies of 0.2% or less than 1%) had NS/S ratios of 1.03 and 1.01, respectively, which were greater than NS/S ratios for more common alleles (>1%, 0.72; >5%, 0.74) [7]. The differences in these ratios suggest that less common alleles are more likely to result in non-synonymous changes than common alleles and are therefore more likely to affect function. Other potential measures of functional consequences of non-synonymous alleles are whether they affect an evolutionarily conserved residue and whether they cause a large chemical change. Of the 155 non-synonymous SNPs, we were able to assign 118 as affecting evolutionarily conserved (EC) or evolutionarily unconserved (EU) amino acid residues by identifying at least two vertebrate sequences in addition to human. Striking differences in the allele frequency distributions of non-synonymous SNPs at EU and EC sites were observed (Table 6*a* and Fig. 2). For example, for the rarest alleles, occurring at 0.2%, EU/EC was 13/34 whereas for the common alleles, occurring at greater than 5%, EU/EC was 16/5. This difference can also be observed by comparing EU/EC for alleles occurring at less than 1% (26/47) with those occurring at greater than 1% (27/18). This difference is highly significant ($X^2_{3df} = 18.13$, $p < 0.0005$). We also observed that non-synonymous SNPs at EC sites were more likely

Table 6. Nature of non-synonymous SNPs as a function of frequency

a.

SNP Frequency	NS/S	EC SNPs (#)	EU SNPs (#)	EC SNPs (%)	EU SNPs (%)
≤ 0.2%	1.03	34	13	52.3	24.5
< 1%	1.01	47	26	72.3	49.1
≥ 1%	0.72	18	27	27.7	50.9
≥ 5%	0.74	5	16	7.7	30.2

b.

	BLOSUM62 Score[1]						SLIM Score	
SNP Frequency	*Total* < 0	≥ 0	*Loops* < 0	≥ 0	*TMDs* < 0	≥ 0	*TMDs* < 0	≥ 0
≤ 0.2%	36 (57.1)	27 (42.9)	27 (58.7)	19 (41.3)	9 (52.9)	8 (47.1)	5 (29.4)	12 (70.6)
≤ 1%	64 (60.4)	42 (39.6)	47 (62.7)	28 (37.3)	17 (54.8)	14 (45.2)	10 (32.3)	21 (67.7)
> 1%	27 (56.3)	21 (43.8)	24 (60.0)	16 (40.0)	3 (37.5)	5 (62.5)	1 (12.5)	7 (87.5)
≥ 5%	13 (52.0)	12 (48.0)	12 (57.1)	9 (42.9)	1 (25.0)	3 (75.0)	0 (0.0)	4 (100.0)
All	91 (59.1)	63 (40.9)	71 (61.7)	44 (38.3)	20 (51.3)	19 (38.7)	11 (28.2)	28 (71.8)

[1]BLOSUM62 and SLIM values for non-synonymous SNPs in the loops, TMDs, and/or total protein were assigned according to the amino acid matrices of Henikoff *et al.* and Muller *et al.* (as described in Material and Methods). Values were assigned for different allele frequency classes (≤ 0.2%, ≤ 1%, > 1%, ≥ 5%, and all frequencies).

*Values are listed as the number of SNPs in a particular frequency class. Percentage of SNPs is listed in parentheses below.

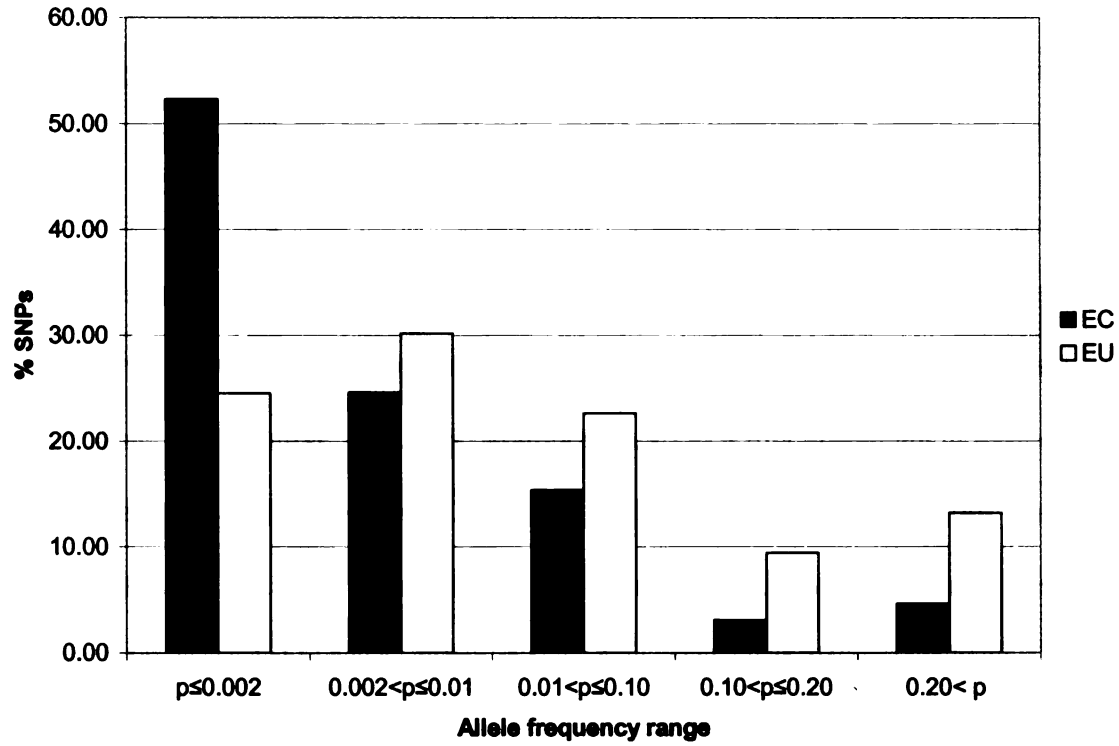


Figure 2. Allele frequency distribution of evolutionarily conserved (EC) and evolutionarily unconserved (EU) non-synonymous SNPs. Percentage of non-synonymous SNPs at EC (black columns) sites and EU (white columns) sites are shown for various allele frequency ranges. Amino acid residues were classified as EC or EU based on sequence alignments of each gene with two mammalian orthologs (rat, mouse, or rabbit).

to be ethnic-specific than non-synonymous SNPs at EU sites. In particular, 50 out of 65 (77%) non-synonymous SNPs at EC sites were present in only 1 population while for non-synonymous SNPs at EU sites, 28 out of 53 (53%) SNPs were present in only 1 population ($X_{1df}^2 = 7.56$, $p = 0.006$).

A similar analysis of all 155 nonsynonymous SNPs was carried out using Grantham's values, which provide a measure of chemical similarity[12]. No differences were seen when amino acid changes were characterized according to Grantham values (for example, conservative versus moderately radical) as a function of allele frequency. We used the same criteria as Cargill *et al.* to characterize the non-synonymous changes using BLOSUM62, scoring values less than 0 or greater than or equal to 0 as a function of allele frequency (Table 6*b*) [2]. We observed no significant differences in analyzing total non-synonymous changes or changes in loops or TMDs at different allele frequencies. Analysis of TMDs by the SLIM matrix revealed an interesting trend: For alleles occurring at less than 1% frequency, 10 exhibited a score less than 0 and 21 greater or equal to 0 whereas for alleles occurring at greater than 1%, 1 exhibited a score less than 0 and 7 exhibited a score greater than 0 (Table 6*b* and Fig. 2*b*). These data are consistent with rare alleles being more deleterious, but the data failed to attain significance by a Fisher exact test ($p = 0.212$).

DISCUSSION

We report here a comprehensive description of genetic variation in a set of 24 membrane transporter genes in 494 chromosomes from ethnically diverse populations. Previous large-scale screening studies have included genes encoding a variety of different

protein types such as globular, cytoplasmic as well as membrane proteins (receptors, channels and transporters). By screening for variation in a functionally and structurally similar class of proteins, we were able to learn about the relative levels of nucleotide diversity across different structural regions of membrane transporters and to compare this set of genes with unselected gene sets. Furthermore, sequencing a large number of chromosomes allowed us to identify rare variants and thereby better understand the selective forces acting on the genes.

Variation in membrane transporters genes

Nucleotide Diversity

The study presented here focused specifically on membrane transporters with known and potential roles in response to drugs and other xenobiotics. Some differences are apparent between our study and previous SNP identification studies. For example, our θ values (Table 1*a*) were somewhat higher than those reported by Halushka *et al.* and substantially higher than those of Cargill *et al.* and Stephens *et al.* An explanation is that the value of θ is related to the number and (proportion) of African Americans screened (100 (0.40); 40 (0.54); 21 (0.25), and 17 (0.30), in our study and Halushka *et al.*, Stephens *et al*, and Cargill *et al.*, respectively), because African American populations are known to have high levels of nucleotide diversity [2-4]. Our π values were similar to those reported by Cargill *et al.* in all gene classes (4.9 versus 5.1 x 10^{-4} for π_{total}) but lower than those of Stephens *et al.* (9 x 10^{-4}). In our set of membrane transporters, the average nucleotide diversity over all regions (π_{total}) was greater than the amino acid diversity (π_{NS}) (4.9 versus 2.3 x 10^{-4}), consistent with the findings of Cargill *et al.*, and

indicating that the amino acid changes are subject to purifying selection in our set of genes.

Of the 680 SNPs that we identified, 62% were population specific in comparison with approximately 50% observed by Stephens *et al.* [4]. This difference may reflect the fact that we screened a larger number of individuals and therefore detected rare variants that are necessarily population specific. We observed that the number of singletons in African Americans was twice that in Caucasians in contrast to Stephens *et al.* who observed a threefold difference. The ratio of singletons in African Americans to those in Caucasians reflects an excess of rare ancestral alleles in the African American population sample, assuming that the rate of mutation is similar in both populations. Our ratio indicates that new mutations account for no more than 50% of rare alleles in the African American population. We anticipate that deeper screening of these populations will reveal a further reduction in this ratio and an apparent increase in the fraction of rare mutations that are new.

In general, population-specific alleles were found at low frequency. Only 4/278 African American-specific alleles and 1/50 Asian-specific alleles had frequencies greater than 0.1. Strikingly, the Caucasian population sample had no population-specific alleles (0/80) even at moderately high frequency ($\geq$ 5%) in contrast to the African American population sample which had 31/278. The relatively high incidence of moderately frequent population-specific alleles in African Americans may facilitate identification of ethnic-specific disease loci in this population.

We noted that the minor allele frequency distribution of coding region SNPs differed between African American and Caucasian population samples (Table 5). This

difference was notable for intermediate and high frequency alleles and reflects population demography. These observations are consistent with the view that the Caucasian population derived from an ancestral African population and that low frequency alleles were lost in the derived population while high frequency alleles were maintained.

In our study, the African American sample had the largest θ in comparison with the Caucasian and Asian samples (Table 1). Our coding region θ (x 10^4) values for Caucasians were similar to that of Halushka *et al.* (4.2 versus 4.5); our θ (x 10^4) for African Americans was larger than that of Halushka *et al.* (8.1 versus 6.3) [3]. This difference presumably reflects the increased number of rare SNPs identified in sequencing a sample of 100 African Americans versus 40 Africans analyzed by Halushka *et al.*

For 11 of the 24 genes (OCT1, PEPT1, CNT1, ENT1, FIC1, MDR3, MRP1, MRP2, GAT1, NET, VMAT1), values of Tajima's *D* were significantly negative, ranging from -1.89 to -2.98; values were also negative for 12/13 of the other genes (see Table 3), but did not reach significance. Negative values of Tajima's *D* reflect an increase in the number of rare variants compared to that predicted under the neutral mutation model, which could be due to population expansion or negative selection. Since the same population was used to screen for genetic variation in all transporter genes, and hence would be expected to show the same degree of distortion due to demography, it may be that genes with significantly negative Tajima's *D* values are subject to purifying selection, but sampling properties of Tajima's *D* in an expanding population would need to be determined to assess statistical significance. We observed that Tajima's *D* is considerably more negative in the African American population sample than in the

Caucasian sample (-1.33 versus -0.5). For most genes, the reverse is true, which is interpreted to indicate that the African population has had a more stable, large long term population size (less growth) compared to Caucasians. The more negative Tajima's *D* in African Americans relative to Caucasians in our study might arise if there were weak purifying selection such that the larger effective size in Africans allows selection to discriminate with greater precision among the transporter alleles. It is also possible that the African populations have been subjected to greater levels of selection for tolerance of xenobiotics.

Insertions and Deletions

Previous large-scale screening studies have provided relatively little information on insertion and deletion mutations presumably for technical reasons and because of their rarity [2-4]. Clark *et al.* (1998) reported the identification of 9 indel sites and 79 SNPs in screening the entire lipoprotein lipase gene in 71 individuals [13, 14]. All of these indels were intronic. Similarly, Saito *et al.* identified 29 non-coding indels and 297 SNPs in screening nine ABC transporter genes [15]. In contrast, we observed 29 indel mutations and 680 SNPs; eight indels affected the coding region. Normalizing for the percentage of intronic sequence screened, our frequency of intronic indels was similar to that reported for lipoprotein lipase (0.07 versus 1.3, respectively) [13, 14]. The low frequency of coding indels (approximately 1.2%) presumably reflects their severe consequences on protein function. Of the eight coding indels, five added or deleted amino acids but otherwise conserved the reading frame. It is interesting to note that two of these, one of which adds an amino acid and one of which deletes an amino acid, occurred at high

frequency (0.29 for CNT1 4.4 and 0.105 for OCT1 7.5). Both of these variants exhibit transport function although there may be differences in specificity (unpublished observations). As anticipated from studies of indel formation in bacteria and yeast, indels occurred at sites of repetitive nucleotides in 16 of 29 instances (e.g., G4 → G5, $[ATAAA]_3$ → $[ATAAA]_4$).

Selection on Membrane Transporters and Domains

The data in this study suggest that most of the 24 membrane transporters are under negative selection, as indicated by the negative Tajima's *D* values and the low ratios of π_{NS} to π_S. We found that 23/24 genes had negative Tajima's *D* values, consistent with Stephens *et al.*, who found that 90% of 313 genes had negative Tajima's *D* values. Although these negative *D* values may be explained by population expansion, they are also consistent with negative selection. The mean ratio of π_{NS}/π_S in this study (0.23) was considerably lower than that of Cargill *et al.* (0.64), who analyzed 106 genes from diverse classes [2]. The lower ratio in our study suggests that membrane transporter genes may be under more negative selection than the diverse set analyzed by Cargill *et al.* The heavy metal transporters (NRAMP1, DMT1 and IREG1) as well as the plasma membrane neurotransmitter transporters in the SLC6 family (DAT, NET, and GAT1) appear to be under particularly strong negative selection as indicated by both negative Tajima's *D* values and very low π_{NS}/π_S values. The π_{NS}/π_S reported for SERT (0.23), another member of the SLC6 family, was higher than observed for DAT, NET and GAT1 [9].

Selection acts to preserve function of a protein. Within membrane transporter proteins, there are two secondary structure domains, TMDs and loops, which have

distinct amino acid compositions consistent with their cellular environment (aqueous versus lipid) and function. We asked the question of whether there are differences in selective pressures acting on these two domains. We first noted that amino acid diversity (π_{NS}) was lower in the TMDs than in the loops (1.16×10^{-4} versus 2.68×10^{-4}) and lower than in the largely globular proteins described by Cargill *et al.* (2.75×10^{-4}) [2]. The restricted variation in TMDs suggests that these domains have greater structural constraints than loops and globular domains. The changes in TMDs that we observed were generally well-predicted to be tolerated using the SLIM matrix (72% were considered evolutionarily favorable) and occurred preferentially at evolutionarily unconserved sites versus conserved sites (Table 6). In contrast the majority of changes in loops were predicted to be poorly tolerated using BLOSUM62, which suggests that it may be useful to construct a blocks matrix using loop domains.

Our studies have identified a large number of alleles that occur at frequencies of 1% or less (Table 5). Because greater than 50% of all known disease-causing mutations are due to amino acid changes, it is of interest to predict the functional consequences of non-synonymous variants, both at low frequency and at moderate to high frequencies [16]. We evaluated the possible functional consequences of the amino acid changes occurring at different allele frequencies by a variety of methods. Of these, the NS/S and EU/EC ratios exhibited significant differences as a function of allele frequency. The NS/S ratio has been used previously to evaluate possible functional consequences from previous data sets of Cargill *et al.* and Halushka *et al.* [2, 3, 7]. Our observations are qualitatively similar to this prior work in that the less common alleles exhibited a greater NS/S ratio in comparison to the more common alleles, indicating that rarer alleles are

more likely to change amino acids and alter protein function. Because we analyzed more chromosomes for each allele, we anticipated that our NS/S ratio would be greater than that of Cargill *et al.* and Halushka *et al.* at the lowest allele frequencies. The fact that this was not observed may indicate that rare alleles of transporter genes are not as likely to affect function as rare alleles of other classes of genes.

By taking advantage of multiple alignments of closely related transporters, we were able to characterize non-synonymous positions as either EU or EC. The EU/EC ratio revealed a great difference between alleles occurring at different frequencies, with the rarer alleles having an increased fraction of alterations at EC sites (Table 6). The EU/EC ratio provides a simple method of incorporating phylogenetic conservation for evaluation of amino acid changes and should be increasingly valuable as more phylogenetic sequence information becomes available. More complex methods for predicting deleterious amino acid changes using phylogenetic information are under development [17].

A substantial number of Mendelian disorders are associated with non-functional genetic variants of membrane transport proteins. These include glucose-galactose malabsorption syndrome, Menke's syndrome and Tangier's disease. Whether the variants that we have identified contribute to diseases or to alterations in drug response is under study. In a recent study by Miller *et al.*, examination of known disease-causing mutations in seven genes demonstrated that disease-causing mutations were more prevalent at evolutionarily conserved sites [16]. The results from this study support our analysis and suggest that amino acid changes at evolutionarily conserved sites have

deleterious effects on protein function and are therefore maintained at low allele frequencies.

In conclusion, we identified genetic variants in a structurally and functionally similar class of proteins (*e.g.*, membrane transporters) by screening a large collection of DNA (247 samples) from ethnically diverse populations. Because we screened such a large sample of DNA, we identified lower frequency variants than in previous studies. This enabled us to analyze the characteristics of amino acid changes as a function of allele frequency and identify predictors of the allele frequency distributions of non-synonymous SNPs. We observed that the allele frequency distribution of non-synonymous SNPs at evolutionarily conserved (EC) sites contained an excess of rare alleles, suggesting EC SNPs may be deleterious to protein function. In addition, we found that transmembrane domains have significantly less amino acid variation than loops, demonstrating that there are segmental differences in variation within membrane transport proteins and suggesting that amino acid variation in TMDs is constrained within humans. This may be due to an important functional and structural role of TMDs. The lower amino acid diversity in TMDs was consistent with our phylogenetic comparisons which indicated that TMDs contain a larger fraction of evolutionarily conserved amino acids than loops. This study demonstrates the use of human population genetic analysis, in combination with phylogenetic comparisons, to learn about protein structure and function. Our results suggest that analysis of nucleotide diversity in different protein segments and in different categories of non-synonymous variants (*i.e.*, EC and EU) may ultimately aid in predicting deleterious alleles which may be implicated in disease or drug-response phenotypes.

REFERENCES

1. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ and Altshuler D. The structure of haplotype blocks in the human genome. *Science* **296**(5576): 2225-9, 2002.
2. Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, Patil N, Lane CR, Lim EP, Kalayanaraman N, Nemesh J, Ziaugra L, Friedland L, Rolfe A, Warrington J, Lipshutz R, Daley GQ and Lander ES. Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat Genet* **22**(3): 231-8, 1999.
3. Halushka MK, Fan JB, Bentley K, Hsie L, Shen N, Weder A, Cooper R, Lipshutz R and Chakravarti A. Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. *Nat Genet* **22**(3): 239-47, 1999.
4. Stephens JC, Schneider JA, Tanguay DA, Choi J, Acharya T, Stanley SE, Jiang R, Messer CJ, Chew A, Han JH, Duan J, Carr JL, Lee MS, Koshy B, Kumar AM, Zhang G, Newell WR, Windemuth A, Xu C, Kalbfleisch TS, Shaner SL, Arnold K, Schulz V, Drysdale CM, Nandabalan K, Judson RS, Ruano G and Vovis GF. Haplotype variation and linkage disequilibrium in 313 human genes. [Erratum In: Science 2001 Aug 10;293(5532):1048]. *Science* **293**(5529): 489-93, 2001.
5. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**(3): 585-95, 1989.
6. Hartl DL and Clark AG. *Principles of Population Genetics*. Sinauer Associates, Inc., Sunderland, MA, 1997.

7. Fay JC, Wyckoff GJ and Wu CI. Positive and negative selection on the human genome. *Genetics* **158**(3): 1227-34, 2001.

8. Smirnova I, Hamblin MT, McBride C, Beutler B and Di Rienzo A. Excess of rare amino acid polymorphisms in the Toll-like receptor 4 in humans. *Genetics* **158**(4): 1657-64, 2001.

9. Glatt CE, DeYoung JA, Delgado S, Service SK, Giacomini KM, Edwards RH, Risch N and Freimer NB. Screening a large reference sample to identify very low frequency sequence variants: comparisons between two genes. *Nat Genet* **27**(4): 435-8, 2001.

10. Muller T, Rahmann S and Rehmsmeier M. Non-symmetric score matrices and the detection of homologous transmembrane proteins. *Bioinformatics* **17 Suppl 1:** S182-9, 2001.

11. Henikoff S and Henikoff JG. Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci U S A* **89**(22): 10915-9, 1992.

12. Grantham R. Amino acid difference formula to help explain protein evolution. *Science* **185**(4154): 862-4, 1974.

13. Clark AG, Weiss KM, Nickerson DA, Taylor SL, Buchanan A, Stengard J, Salomaa V, Vartiainen E, Perola M, Boerwinkle E and Sing CF. Haplotype structure and population genetic inferences from nucleotide-sequence variation in human lipoprotein lipase. *Am J Hum Genet* **63**(2): 595-612, 1998.

14. Nickerson DA, Taylor SL, Weiss KM, Clark AG, Hutchinson RG, Stengard J, Salomaa V, Vartiainen E, Boerwinkle E and Sing CF. DNA sequence diversity in

a 9.7-kb region of the human lipoprotein lipase gene [see comments]. *Nat Genet* **19**(3): 233-40, 1998.

15. Saito S, Iida A, Sekine A, Miura Y, Ogawa C, Kawauchi S, Higuchi S and Nakamura Y. Three hundred twenty-six genetic variations in genes encoding nine members of ATP-binding cassette, subfamily B (ABCB/MDR/TAP), in the Japanese population. *J Hum Genet* **47**(1): 38-50, 2002.

16. Miller MP and Kumar S. Understanding human disease mutations through the use of interspecific genetic variation. *Hum Mol Genet* **10**(21): 2319-28, 2001.

17. Ng PC and Henikoff S. Predicting deleterious amino acid substitutions. *Genome Res* **11**(5): 863-74, 2001.

CHAPTER 5

SUMMARY AND CONCLUSIONS

Over the past decade, significant progress has been made in understanding the molecular mechanisms and biological significance of organic cation transporters. Genes for six organic cation transporters have been cloned: OCT1-OCT3 and OCTN1-OCTN3 [1-15]. Molecular studies with these organic cation transporters have elucidated their transport mechanisms, substrate specificities, tissue distributions, and membrane localizations [16, 17]. Studies in animal models have begun to reveal their *in vivo* physiological and pharmacological roles. Recently, genetic variants in human OCTN2 have been associated with an autosomal recessive disease, primary carnitine deficiency, demonstrating its physiological importance in maintaining cellular carnitine levels [18]. To date, OCTN2 is the only organic cation transporter for which genetic variants have been associated with a human phenotype. However, studies in heterologous expression systems and studies in animal models have demonstrated that many OCTs and OCTNs transport a variety of therapeutic drugs and play important roles in the disposition of organic cation drugs in the systemic circulation. It is likely that genetic variation in these transporters could lead to interindividual variation in drug response. The major goals of the work presented here were to identify genetic variants of organic cation transporters, examine the effects of this genetic variation on transporter function, and determine the significance of genetic variation in membrane transporters in relation to global variation in the human genome.

It is well established that there is significant interindividual variation in drug response [19]. Early pharmacogenetic studies demonstrated that genetic variation in drug metabolizing enzymes contributes significantly to variation in drug response, namely, to hepatic elimination. The genetic factors contributing to variation in renal elimination, the other major pathway of elimination, are, however, not yet well understood. Renal secretion is mediated by transporters in the proximal tubule epithelium of the kidney [16, 17, 20]. Studies suggest that the human organic cation transporter, OCT2, is the major renal transporter governing the first step in renal secretion - transport of organic cations from the blood into the renal tubule [21]. Therefore, it is possible that genetic variation in OCT2 contributes to interindividual variation in the renal clearance of organic cation drugs.

To better understand the functional significance of variation in OCT2, we identified SNPs of OCT2 in a large collection of DNA from ethnically diverse populations and functionally characterized four common, non-synonymous SNPs (Chapter 2). SNPs were identified in the coding regions and flanking intronic regions of the OCT2 gene using the high-throughput technique of denaturing HPLC analysis followed by direct DNA sequencing. In total, we identified 28 variants of OCT2, 12 in the non-coding or intronic regions and 18 in the coding region. Only four of eight non-synonymous coding-region SNPs had allele frequencies $\geq$1% in an ethnic population. Each of these four variants affected transporter function assayed in *Xenopus laevis* oocytes.

The most common variant, A270S, with an overall allele frequency of 12%, displayed slight alterations in its interaction with organic cations, but exhibited functional

characteristics most similar to that of the reference OCT2. In contrast, the less common variants, M165I, R400C, and K432Q with overall allele frequencies less than 1% displayed significantly reduced transport function (*e.g.*, decreased V_{max}) and/or altered interactions with organic cations. These results suggest that less common non-synonymous variants have more significant effects on transporter function. More detailed functional analyses of rare and common non-synonymous variants of other transporters will be needed to determine whether this trend can be generalized to other transporters.

We used the SNP data in this study to predict haplotypes of OCT2 and identified 13 unambiguous haplotypes. Only one of the four common non-synonymous variants, A270S, was present in unambiguous haplotypes (haplotypes *3D and *3E). Although this variant had minimal effects on transporter function, it may be interesting to examine the *in vitro* and *in vivo* functions of the *3D and *3E haplotypes.

In addition to our functional analysis, we also examined nucleotide diversity in OCT2 and compared it to that reported for other genes. Diversity at synonymous sites ($\pi_{synonymous}$) in OCT2 was much greater than that at non-synonymous sites ($\pi_{non\text{-}synonymous}$), and the ratio of $\pi_{synonymous}$ to $\pi_{non\text{-}synonymous}$ was much greater than that reported for other genes. In combination with our functional analysis showing that non-synonymous changes in OCT2 alter transporter function, these results suggest that selection acts against amino acid changes in OCT2. Such selection may be due to an important role of OCT2 in renal elimination of endogenous compounds and xenobiotics.

In Chapter 3, we further investigated the genetic contribution to variation in renal clearance by calculating the heritability of renal clearance using the "Repeat Drug

Application" methodology (RDA) proposed by Kalow *et al.* (1998) [22]. This method can be used as a substitute for twin studies to dissect the environmental and genetic components contributing to a particular trait. We calculated the genetic component contributing to variation in renal clearance ($r_{GC, renal}$) for five drugs – terodiline, amoxicillin, ampicillin, metformin, and para-aminohippurate (PAH). These drugs were chosen because they span a range of renal clearance values and because renal clearance data within and between individuals were available in the literature.

We demonstrated that the heritability of renal clearance of drugs undergoing net renal secretion (*e.g.*, metformin, amoxicillin, and ampicillin) is significantly greater than the heritability of renal clearance of drugs that are passively reabsorbed or drugs whose renal clearance is controlled by renal blood flow. Active secretion is controlled by transporters in the proximal tubule epithelium such as organic cation transporters (OCTs), organic anion transporters (OATs), and multidrug resistance associated proteins (MRPs) [16, 17, 20]. Studies have demonstrated that the renal organic cation transporter, OCT2, transports metformin and that the organic anion transporter, OAT1, transports both amoxicillin and ampicillin [23, 24]. It is likely that the high heritability of net secretion is due to genetic variation in these renal transporters.

It is interesting to consider the molecular basis for the high heritability of metformin renal clearance given our knowledge of genetic variation in the human *OCT2* gene (Chapter 2). In Chapter 2, we demonstrated that four non-synonymous variants of OCT2 had altered functional characteristics, such as decreased maximal transport rates and altered interactions with organic cations. These variants may contribute to interindividual variation in metformin renal clearance and the high heritability of

metformin renal clearance calculated in Chapter 3. One possible candidate SNP that may contribute to the high heritability of metformin renal clearance is the A270S variant, with an overall population allele frequency of 12%. Although the effect of this variant on transporter function assayed in *X. laevis* oocytes was subtle, its effect on transporter function in proximal tubule cells *in vivo* is unknown. Moreover, we identified two haplotypes of OCT2 that contain the A270S variant which may also exhibit functional differences from the reference haplotype. Variation in the promoter region of OCT2, or in other transporters that work in concert with OCT2 to mediate secretion, may also contribute to variation in renal secretion. Future studies correlating an individual's transporter genotype or haplotype with a renal clearance phenotype may help to elucidate the molecular mechanism controlling heritability of renal secretion.

In Chapter 4, we expanded upon the genetic analysis of OCT2 discussed in Chapter 2 and examined genetic variation in 24 membrane transporter genes with roles in drug response in 247 DNA samples from ethnically diverse populations. By examining variation in nucleotide diversity over different genomic and structural units of membrane transporters, we were able to identify predictors of transporter function. The studies conducted in this chapter differ in several ways from previous large-scale genetic screening studies [25-27]. First, we screened a larger number of chromosomes from Caucasian and African American populations (200 chromosomes in each population) which allowed us to identify rarer variants than in previous studies, learn about differences in genetic variation between Caucasians and African Americans, and examine frequency distributions of variants. Secondly, we screened a structurally similar class of

proteins – membrane transporters – which allowed us to examine genetic variation in two different structural units of proteins, transmembrane domains and loops.

One of the major findings in this chapter was that amino acid diversity (quantified by $\pi_{\text{non-synonymous}}$) was significantly lower in the transmembrane domains compared to the loops. In other words, there appear to be segmental differences in variation within membrane transporter proteins. In agreement with these results, we observed that the majority of amino acid changes within the transmembrane domains were predicted to be well-tolerated using the SLIM matrix (*i.e.*, 72% of the amino acid changes within transmembrane domains have positive SLIM scores) [28]. This finding suggests that transmembrane domains may have specialized functions that restrict variation.

A second major finding in this chapter was that the degree of evolutionary conservation in the coding region is a strong predictor of allele frequencies. We observed that amino acid changes at evolutionarily conserved (EC) sites occur preferentially at low allele frequencies. The ratio of non-synonymous changes at evolutionarily conserved sites to evolutionarily unconserved sites (EC/EU) was much greater at low allele frequencies (< 0.2%) than at high allele frequencies (> 5%), suggesting that changes at evolutionarily conserved sites may be deleterious. Our findings also have important implications in the current debate over whether common or rare variants contribute to complex phenotypes and suggest that rare alleles may have more severe effects on protein function and thereby contribute to interindividual differences in drug response. Our studies also revealed common alleles that may have subtle effects on transporter function. It is interesting to note that the majority of potentially deleterious alleles (*e.g.*, non-synonymous SNPs at evolutionarily conserved sites) that we identified were found in

only one population. This suggests that drug-response phenotypes with similar characteristics across multiple populations may have different underlying genetic mechanisms which are specific to a particular ethnic population.

Our genetic analysis suggests that evolutionarily conserved amino acid residues in the coding regions are likely to play important roles in protein function. The importance of evolutionarily conserved nucleotide sequence in non-coding regions, such as promoter or intronic regions, is still unknown. It will be interesting in future studies to compare levels of nucleotide diversity at evolutionarily conserved and unconserved sites in non-coding regions. Examining levels of nucleotide diversity across various regions of non-coding sequence may help to define regions which are important in regulating transcription or RNA splicing.

Other future studies which stem from the analyses presented in Chapter 4 include the development of specialized amino acid block matrices for various structural domains within proteins. Such matrices may aid in predicting favorable and unfavorable amino acid changes and ultimately in predicting the effects of amino acid changes on protein function. For example, it may be useful to generate an amino acid matrix for the loop regions of membrane spanning proteins. Although the SLIM matrix was a fairly good predictor of the allele frequency distribution of non-synonymous changes in the transmembrane domains (with rare amino acid changes having more negative SLIM values than more common amino acid changes), BLOSUM62 was not as effective at predicting rare variants within the loop regions. In addition, block matrices for the carboxy- and N-termini as well as block matrices specifically for small loops and larger intra- or extra-cellular domains, may be helpful for predicting favorable and unfavorable

amino acid changes. It may also be useful to examine non-synonymous changes in the context of a three-dimensional protein structure.

The field of pharmacogenetics has evolved from initial studies examining genetic variation in drug metabolizing enzymes. Much progress has been made over the past decade in identifying drug metabolizing enzyme variants and relating these variants to interindividual differences in drug response [19]. The research presented in this dissertation has extended the field of pharmacogenetics by examining genetic variation in another group of genes with important roles in drug response - membrane transporter genes. In Chapter 2, we examined genetic variation in the human organic cation transporter, OCT2, which is a major secretory transporter in the renal epithelium of proximal tubule cells and interacts with a variety of clinically used drugs. In Chapter 3, we demonstrated that interindividual variation in renal secretion of select organic cation and organic anion drugs has a large genetic component which may be due, in part, to genetic variation in renal transporters, such as OCT2. Lastly, in Chapter 4, we examined genetic variation in 24 membrane transporter genes and used analysis of nucleotide and amino acid diversity to learn about the selective forces acting on this set of genes. Future studies are needed to elucidate further the *in vivo* effects of genetic variation in membrane transporter genes, like OCT2, on drug response. Such studies may include candidate SNP- or haplotype-based association studies, and genotype to phenotype or phenotype to genotype studies. Furthermore, specialized bioinformatics tools, such as block matrices based on evolutionary conservation of secondary and tertiary protein structures, may become increasing useful to help predict the effects of genetic variation on protein function.

REFERENCES

1. Gorboulev V, Ulzheimer JC, Akhoundova A, Ulzheimer-Teuber I, Karbach U, Quester S, Baumann C, Lang F, Busch AE and Koepsell H. Cloning and characterization of two human polyspecific organic cation transporters. *DNA Cell Biol* **16**(7): 871-881, 1997.
2. Grundemann D, Gorboulev V, Gambaryan S, Veyhl M and Koepsell H. Drug excretion mediated by a new prototype of polyspecific transporter. *Nature* **372**(6506): 549-552, 1994.
3. Grundemann D, Schechinger B, Rappold GA and Schomig E. Molecular identification of the corticosterone-sensitive extraneuronal catecholamine transporter. *Nat Neurosci* **1**(5): 349-351, 1998.
4. Kekuda R, Prasad PD, Wu X, Wang H, Fei YJ, Leibach FH and Ganapathy V. Cloning and functional characterization of a potential-sensitive, polyspecific organic cation transporter (OCT3) most abundantly expressed in placenta. *J Biol Chem* **273**(26): 15971-15979, 1998.
5. Mooslehner KA and Allen ND. Cloning of the mouse organic cation transporter 2 gene, Slc22a2, from an enhancer-trap transgene integration locus. *Mamm Genome* **10**(3): 218-224, 1999.
6. Okuda M, Saito H, Urakami Y, Takano M and Inui K. cDNA cloning and functional expression of a novel rat kidney organic cation transporter, OCT2. *Biochem Biophys Res Commun* **224**(2): 500-507, 1996.

7. Tamai I, Yabuuchi H, Nezu J, Sai Y, Oku A, Shimane M and Tsuji A. Cloning and characterization of a novel human pH-dependent organic cation transporter, OCTN1. *FEBS Lett* **419**(1): 107-111, 1997.

8. Tamai I, Ohashi R, Nezu JI, Sai Y, Kobayashi D, Oku A, Shimane M and Tsuji A. Molecular and functional characterization of organic cation/carnitine transporter family in mice. *J Biol Chem* **275**(51): 40064-72, 2000.

9. Tamai I, Ohashi R, Nezu J, Yabuuchi H, Oku A, Shimane M, Sai Y and Tsuji A. Molecular and functional identification of sodium ion-dependent, high affinity human carnitine transporter OCTN2. *J Biol Chem* **273**(32): 20378-20382, 1998.

10. Terashita S, Dresser MJ, Zhang L, Gray AT, Yost SC and Giacomini KM. Molecular cloning and functional expression of a rabbit renal organic cation transporter. *Biochim Biophys Acta* **1369**(1): 1-6, 1998.

11. Verhaagh S, Schweifer N, Barlow DP and Zwart R. Cloning of the mouse and human solute carrier 22a3 (Slc22a3/SLC22A3) identifies a conserved cluster of three organic cation transporters on mouse chromosome 17 and human 6q26-q27. *Genomics* **55**(2): 209-218, 1999.

12. Wu X, Kekuda R, Huang W, Fei YJ, Leibach FH, Chen J, Conway SJ and Ganapathy V. Identity of the organic cation transporter OCT3 as the extraneuronal monoamine transporter (uptake2) and evidence for the expression of the transporter in the brain. *J Biol Chem* **273**(49): 32776-32786, 1998.

13. Wu X, Prasad PD, Leibach FH and Ganapathy V. cDNA sequence, transport function, and genomic organization of human OCTN2, a new member of the

organic cation transporter family. *Biochem Biophys Res Commun* **246**(3): 589-595, 1998.

14. Zhang L, Dresser MJ, Chun JK, Babbitt PC and Giacomini KM. Cloning and functional characterization of a rat renal organic cation transporter isoform (rOCT1A). *J Biol Chem* **272**(26): 16548-16554, 1997.

15. Zhang L, Dresser MJ, Gray AT, Yost SC, Terashita S and Giacomini KM. Cloning and functional expression of a human liver organic cation transporter. *Mol Pharmacol* **51**(6): 913-921, 1997.

16. Dresser MJ, Leabman MK and Giacomini KM. Transporters involved in the elimination of drugs in the kidney: organic anion transporters and organic cation transporters. *J Pharm Sci* **90**(4): 397-421, 2001.

17. Koepsell H, Gorboulev V and Arndt P. Molecular pharmacology of organic cation transporters in kidney. *J Membr Biol* **167**(2): 103-17, 1999.

18. Nezu J, Tamai I, Oku A, Ohashi R, Yabuuchi H, Hashimoto N, Nikaido H, Sai Y, Koizumi A, Shoji Y, Takada G, Matsuishi T, Yoshino M, Kato H, Ohura T, Tsujimoto G, Hayakawa J, Shimane M and Tsuji A. Primary systemic carnitine deficiency is caused by mutations in a gene encoding sodium ion-dependent carnitine transporter. *Nat Genet* **21**(1): 91-4, 1999.

19. Evans WE and Relling MV. Pharmacogenomics: translating functional genomics into rational therapeutics. *Science* **286**(5439): 487-91, 1999.

20. Bendayan R. Renal drug transport: a review. *Pharmacother* **16**(6): 971-985, 1996.

21. Pietig G, Mehrens T, Hirsch JR, Cetinkaya I, Piechota H and Schlatter E. Properties and regulation of organic cation transport in freshly isolated human proximal tubules. *J Biol Chem* **276**(36): 33741-6, 2001.

22. Kalow W, Tang BK and Endrenyi L. Hypothesis: comparisons of inter- and intra-individual variations can substitute for twin studies in drug research. *Pharmacogenetics* **8**(4): 283-9, 1998.

23. Uchino H, Tamai I, Yabuuchi H, China K, Miyamoto K, Takeda E and Tsuji A. Faropenem transport across the renal epithelial luminal membrane via inorganic phosphate transporter Npt1. *Antimicrob Agents Chemother* **44**(3): 574-7, 2000.

24. Terada T, Saito H, Mukai M and Inui K. Recognition of beta-lactam antibiotics by rat peptide transporters, PEPT1 and PEPT2, in LLC-PK1 cells. *Am J Physiol* **273**(5 Pt 2): F706-11, 1997.

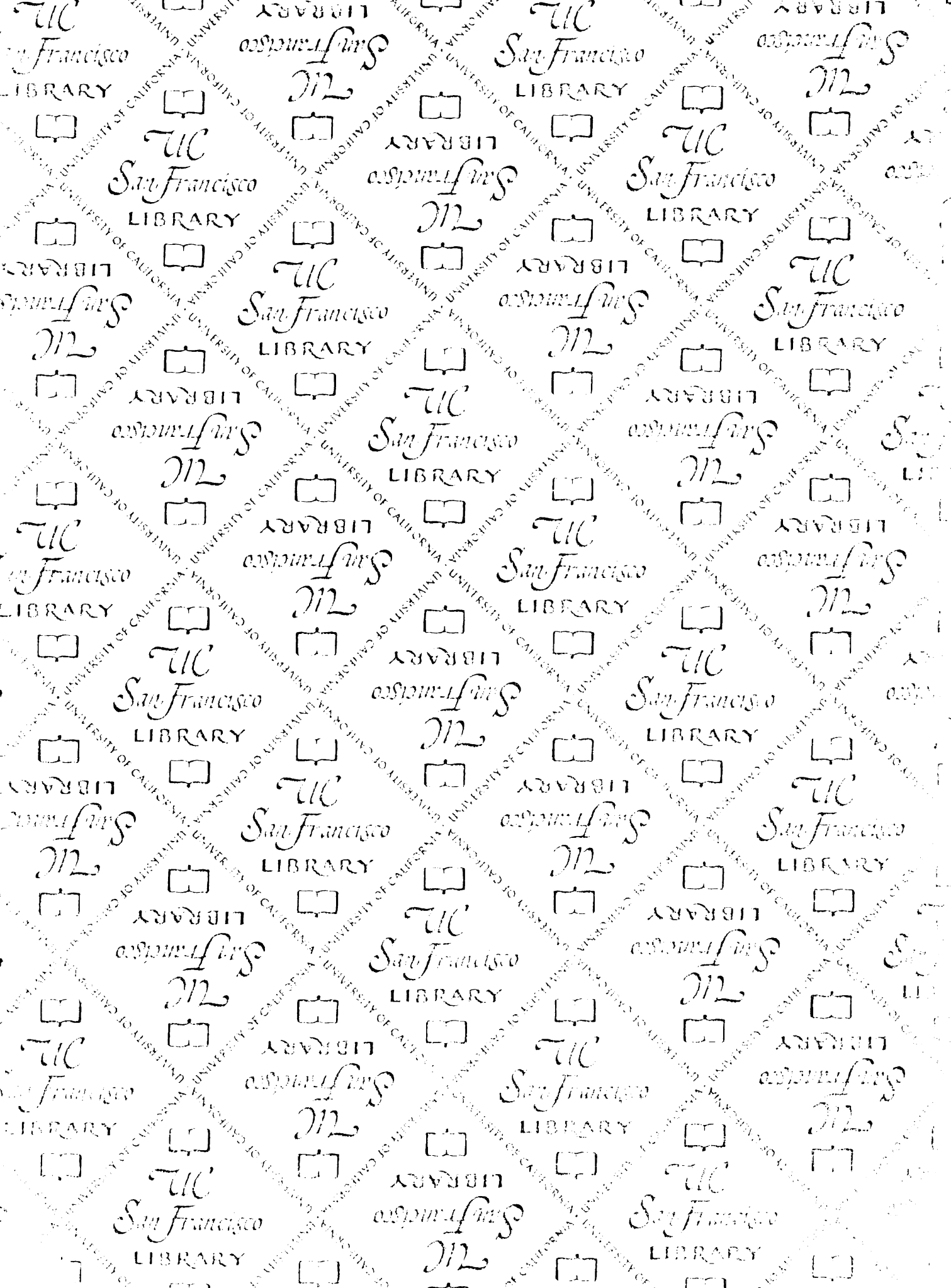
25. Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, Patil N, Lane CR, Lim EP, Kalayanaraman N, Nemesh J, Ziaugra L, Friedland L, Rolfe A, Warrington J, Lipshutz R, Daley GQ and Lander ES. Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat Genet* **22**(3): 231-8, 1999.

26. Halushka MK, Fan JB, Bentley K, Hsie L, Shen N, Weder A, Cooper R, Lipshutz R and Chakravarti A. Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. *Nat Genet* **22**(3): 239-47, 1999.

27. Stephens JC, Schneider JA, Tanguay DA, Choi J, Acharya T, Stanley SE, Jiang R, Messer CJ, Chew A, Han JH, Duan J, Carr JL, Lee MS, Koshy B, Kumar AM, Zhang G, Newell WR, Windemuth A, Xu C, Kalbfleisch TS, Shaner SL, Arnold K, Schulz V, Drysdale CM, Nandabalan K, Judson RS, Ruano G and Vovis GF.

Haplotype variation and linkage disequilibrium in 313 human genes. [Erratum In: Science 2001 Aug 10;293(5532):1048]. *Science* **293**(5529): 489-93, 2001.

28. Muller T, Rahmann S and Rehmsmeier M. Non-symmetric score matrices and the detection of homologous transmembrane proteins. *Bioinformatics* **17 Suppl 1:** S182-9, 2001.


UC
San Francisco
LIBRARY