Tamara Schenekar

**Isolation and characterization of the CYP2D6
gene in Felidae
with comparison to other Mammals**

Masterarbeit

zur Erlangung des akademischen Grades
einer Magistra
an der Naturwissenschaftlichen Fakultät der
Karl-Franzens-Universität Graz

Priv-Doz. Dr. Steven Weiss
Institut für Zoologie

2010

TABLE OF CONTENTS

Deutsche Zusammenfassung/German Summary	ii
Summary	iii
Acknowledgements	iv
I. Introduction.....	1
1. Introducing Cytochrome P450s	1
2. Catalytic functions of Cytochrome P450s and their role in the metabolism.....	2
2.1 Cytochrome P450s and their importance in drug metabolism	2
2.2. Induction and inhibition of Cytochrome P450s.....	3
3. Structure of Cytochrome P450s.....	5
4. The P450 gene superfamily.....	7
4.1 Evolution of the P450 genes	8
4.2 CYP Nomenclature.....	11
4.3 Polymorphisms in the P450s and their medical consequences	12
5. CYP2D6.....	14
5.1 Structural properties of the CYP2D6 enzyme.....	14
5.2 Substrates, Inhibitors and Inducers of CYP2D6	15
5.3 The Debrisoquine/Sparteine polymorphism caused by CYP2D6.....	17
5.3.1 Ultrarapid Metabolizers.....	17
5.4 The <i>CYP2D</i> locus	18
5.5 Alleles of <i>CYP2D6</i>	20
5.5.1 Alleles causing normal enzyme function	21
5.5.2 Alleles causing a decreased enzyme function	21
5.5.3 Alleles causing no enzyme function.....	21
5.5.4 Duplications or multiduplications of the <i>CYP2D6</i> -gene as a cause of increase enzyme function	22
5.6 Genetic mechanisms leading to duplication and multiduplication of <i>CYP2D6</i>	23
5.7 CYP2D6 in other mammals	24
5.8 Studies on Cytochromes P450 in domestic cats.....	26
II. Material and Methods.....	29
II.1 Primer Design	29
II.2 RNA Extraction and Reverse Transcription.....	29
II.3 Amplification and Sequencing of <i>CYP2D6</i>	29
II.4 Data Analysis.....	30
III. Results	32
III.1 Sliding Window Analysis	32
III.2 dN/dS Ratios	33
IV. Discussion.....	37
References.....	39
Figure List	45
Appendix.....	46

DEUTSCHE ZUSAMMENFASSUNG /GERMAN SUMMARY

Cytochrome P450 (auch CYPs oder P450s) stellen eine Superfamilie an Hämproteinen dar und sind vor allem dadurch bekannt, dass sie am Metabolismus von etwa 75% der humanmedizinischen Medikamente am Markt beteiligt sind. Weiters spielen eine wichtige Rolle in der Biosynthese von Steroidhormonen, der Oxidation von ungesättigten Fettsäuren und dem Metabolismus von fettlöslichen Proteinen und werden auch mit der Aktivierung von Prokarzinogenen in Verbindung gebracht. Verschiedene Allelvarianten mehrerer P450 Gene verursachen eine hohe Variabilität in Abbau- und Ausscheidungsgeschwindigkeit von bestimmten Medikamenten zwischen einzelnen Personen. Das hochvariable *CYP2D6*-Gen ist das wohl am besten untersuchte Mitglied der P450 Gen-Superfamilie und CYP2D6 ist am Abbau von etwa 25% der handelsüblichen Medikamente beteiligt. Im Gegensatz zu den zahlreichen detaillierten Studien in Menschen ist noch sehr wenig über mögliche Variabilität in Sequenz und Funktion dieses Gens in Tieren bekannt. Orthologe des *CYP2D6*-Gens wurden in mehreren Tieren von veterinärmedizinischem Interesse charakterisiert, wobei das Hauptinteresse der P450-Forschung in Tieren in der Expression und Substratspezifität des Enzyms liegt.

CYP2D6-mRNA von 23 Vertretern der Familie Felidae, sieben Arten umfassend, wurde sequenziert und mit bereits zur Verfügung stehenden Sequenzen zehn weiterer Säugetierarten und mehreren Allelvarianten des Menschen verglichen.

Ein relativ hohes dN/dS-Verhältnis zwischen den Sequenzen wurde beobachtet, vor allem innerhalb der Katzenfamilie. Eine nicht-monotone Verteilung der paarweisen dN/dS-Verhältnisse in Abhängigkeit zur evolutionären Distanz weist auf den Erhalt von leicht schädlichen Mutationen im Zeitraum von drei bis vier Millionen Jahren hin. Weiters wurden Hinweise positiver Selektion auf einzelne Codons gefunden und ein möglicher Zusammenhang zwischen dem Ernährungstyp einzelner Arten und der Aminosäurezusammensetzung an Substratbindestellen wird diskutiert.

SUMMARY

Cytochrome P450s (CYPs, P450s) are a superfamily of hemoproteins best known by the fact that they metabolize about 75% of marketed human drugs. P450s additionally play important roles in the biosynthesis of steroid hormones, oxidation of unsaturated fatty acids and the metabolism of fat-soluble proteins and are also involved in the activation of procarcinogens. Allelic variants of several human P450 genes underlie the high among individual variation in metabolic conversion and excretion rates of drugs and therefore drug response. The highly polymorphic *CYP2D6* gene is the best studied member of the P450 gene superfamily, accounting for about 25% of the metabolism of commonly used drugs. In contrast to detailed studies of this enzyme and its gene in humans, little is known about the sequence variation or function of this gene in non-human animals. Orthologs of the *CYP2D6* gene have been characterized in several animals of interest to veterinary medicine, whereby the main focus of P450 research in animals is expression and substrate specificity.

CYP2D6-mRNA from 23 cats (Felidae) spanning seven species were sequenced and compared to available *CYPD6*-sequences in ten additional mammals including multiple allelic variants in humans.

A relatively high mean dN/dS ratio among the sequences is revealed, especially within Felidae. A non-monotonic distribution in pairwise dN/dS ratios with respect to evolutionary distance suggests the retention of slightly deleterious mutations across three to four MY. Signals of positive selection pressure on specific codons are supported and the possible connection between the dietary type of specific species and specific amino acid changes in substrate binding sites is discussed.

ACKNOWLEDGEMENTS

This thesis would not have been possible without the support of numerous people. I'd like to thank the following people:

Priv-Doz. Dr. Steven Weiss for supervising this master thesis and letting me have my way to work with such "hairy" organisms. Further for his patience with me and his ability to laugh about the numerous problems that occurred in the course of this work. Thank you, thank you, thank you!

Dr. Kathrin Winkler for her extremely helpful guidance and help throughout the whole work in the lab but also with the subsequent steps of this thesis.

Bailey Douglas Kessing for his assistance during my stay in the U.S. and his family for their hospitality and introducing me to the "American way of life".

The **Laboratoy of Genomic Diversity** for providing most of the samples for this study and especially **Dr. Jennifer Troyer**, **Natalie Baggett** and **Dr. Stephen O'Brien** for their support.

My family, especially my parents **Hildegard** and **Willibald Schenekar** for always supporting my dreams and wishes regarding my academic studies.

My study colleagues for sharing my passion for biology, especially **Mag. Maria Luise Spreitzer** and **Mag. Gernot Kunz** for their supporting words, countless, inspiring coffee breaks and all the time spent together.

Christoph Hellig for being an inspiration and motivation for pursuing my dreams in my studies. Rest in peace!

I. INTRODUCTION

1. Introducing Cytochrome P450s

Cytochrome P450s, also called P450s or CYPs, are a superfamily of hemoproteins, known to have an immense impact on the metabolism of a large portion of commonly used drugs (Guengerich 2008). They are involved in the metabolism of numerous additional xenobiotics as well as endogenous substrates (Hasler 1999). P450s are found in prokaryotes, eukaryotes and archaea, and putatively even within viruses (Lamb et al. 2009).

The name 'Cytochrome P450' stems from a spectral characteristic that all P450s have in common: In their reduced form, these proteins have their maximum absorbance at 450 nm. The 'P' stands for 'pigment' (Sato and Omura 1961). Based on these unique spectrophotometric properties, the first P450 was discovered within rat liver microsomes in 1954 by Martin Klingenberg (Klingenberg 1958; Hasler 1999; Porter 2004). The identification of P450s as hemoproteins and further characterization was done by Tsuneo Omura and Ryo Sato (Omura and Sato 1962; Omura and Sato 1964a; Omura and Sato 1964b). Cooper et al. (1965) were the first to note the involvement of P450s in the metabolism of drugs and other xenobiotics. In the 1970s researchers began to ask just how many P450-isoforms existed in the liver. While many P450-isoforms were identified, their enzymatic similarity complicated thorough characterization. Over the years, the plurality of P450s, as well as their differences in induction and inhibition was revealed. These enzymatic studies were followed by molecular genetic studies based on cloning of P450 cDNA. Molecular genetics provided a new way of studying the diversity of P450s. Within the course of these studies the genetic variability of multiple genes within the family was slowly elucidated (Porter 2004).

Because P450s take part in many metabolic pathways, and especially because of their eminent role in drug processing, the polymorphisms and resulting inter-individual variation of the reactions that are catalyzed by these enzymes are a main concern of the burgeoning field of pharmacogenetics (Hasler 1999; van der Weide and Steijns 1999; Weide and Steijns 1999).

2. Catalytic functions of Cytochrome P450s and their role in the metabolism

The reactions catalyzed by Cytochrome P450s are diverse and play crucial parts in:

- 1.) biosynthesis of steroid hormones
- 2.) oxidation of unsaturated fatty acids to intracellular messengers
- 3.) stereo- and regio-specific metabolism of fat-soluble proteins
- 4.) metabolism of drugs, environmental pollutants and other xenobiotics (Hasler 1999)

The catalytic function of P450s is to transfer one oxygen atom into a substrate. Therefore, biochemists classify P450s as monooxygenases (Ingelman-Sundberg 2004). However, there are some additional reactions that P450s catalyze, thus functioning as isomerases, reductases, and dehydrases (Mansuy 1998). The electrons needed for monooxygenation are received either from NADPH or NADH. During the incorporation of one oxygen atom into the product, the P450 undergoes a cyclic series of reactions.

The basic equation for the cytochrome P450-dependent monooxygenation is as follows:

In mammalian tissues there are two types of electron transport systems aiding this reaction. The first transport system is associated with the mitochondria of some cells and involves a FAD-containing reductase and ferredoxin. The other transport system, referred to as microsomal, is located in the endoplasmatic reticulum of many cell types and involves a unique flavoprotein that contains both FAD and FMN (Hasler 1999).

2.1 Cytochrome P450s and their importance in drug metabolism

As mentioned above, P450s metabolize endogenous as well as exogenous substrates. Enzymatic reactions involving exogenous substrates are very often the main focus of research on cytochrome P450 activity. The number of substrates that are metabolized by these enzymes is enormous and estimated to be far beyond 1000 (Hasler 1999).

One of the most extensively studied issues regarding the activity of P450s is their vast impact on the metabolism of drugs. Since the first discovery of CYPs role in drug metabolism (Cooper et al. 1965), our knowledge of such functions has exponentially grown. Today it is known that cytochrome

P450s are responsible for about 75% of the metabolism of marketed drugs. They make up the major part of the phase I enzymes that alter the structure of lipophilic molecules (most commonly used drugs are lipophilic) to increase their water solubility (Ingelman-Sundberg 2004).

One aspect of their role in drug metabolism is the activation of products that need to be converted to their active form. A widely known example for this is the activation of codeine whereby it is converted to morphine (Poulsen et al. 1996). Interindividual variation in the ability to activate drugs, based for example on genetic polymorphisms, can lead to non-responders at one extreme and serious adverse drug reactions at the other extreme (see section 5.3). The problem of adverse drug reactions is one of the main concerns in drug treatment.

Additionally, many therapeutic agents need enzymatic processing before they can be eliminated from the body. If this processing is slow, the xenobiotic accumulates and thus increases its toxicity in the body. Cytochrome P450s are also of great importance in this aspect of drug metabolism.

Another big concern of medical scientists concerning cytochrome P450s doesn't involve drugs directly but rather the activation of procarcinogens. For example, there are several studies on the relationship of tobacco smoke, CYP2D6 and CYP1A1 and lung cancer. Some CYPs are able to activate inhaled procarcinogens like polycyclic aromatic hydrocarbons (PAHs) and aromatic amines contained in tobacco smoke (Hasler 1999).

2.2 Induction and inhibition of Cytochrome P450s

During the last few decades, the ability of certain chemicals to induce or increase CYP-activity has been widely studied. The first environmental chemicals that were found to be able to induce P450 activity were polycyclic aromatic hydrocarbons (Conney, Miller, and Miller 1956) as mentioned above. The expression of an enzyme can be controlled at the level of transcription, which is the most common case, but also at the RNA- and translation-level, as well as post-transcriptional phase where the stabilization of gene products can be affected. The latter case is widely known for members of the CYP subfamily 1A, which can be bound by bulky planar chemicals that cause the enzyme to switch to a DNA-binding state (Hasler 1999).

Regulation at the transcriptional level is very important and involves three cytosolic receptors that perceive the amount of environmental xenobiotics. The three receptors are called pregnane X-receptor (PXR), constitutive androgen receptor (CAR) and the Ah-receptor (AHR). The intracellular increase of xenobiotics causes the activation of several P450s, which reduce the concentration of these chemicals through these receptors. Cytosolic receptors thus influence the expression of CYP1A1, CYP1A2 and CYP2S1 (AHR), CYP2C9, CYP3A4 (PXR) and CYP2B6, CYP2C9,

CYP3A4 (CAR) (Ingelman-Sundberg 2004). Most P450 drug metabolism is regulated this way. But various other receptors, like hormone receptors are also participating in these systems. CYP activity can also be decreased or inhibited by many chemicals because of the enzymes' broad substrate specificity. The issue of most concern is the inhibition by drug-drug interactions, especially if a patient receives multiple drug therapy in clinical practice. The inhibition can lead to impaired drug clearance and prolongation of drug action. This effect can be short-lived because the inhibiting drug diffuses away, or, if the inhibiting drug has a drastic effect on the enzyme, the inhibition is irreversible. An example of drugs inhibiting CYP activity are HIV-1 protease inhibitors, such as Ritonavir. They block the activity of CYP3A4, which metabolizes several drugs that are administered simultaneously in therapy (Ernest, Hall, and Jones 2005).

Drugs are not the only substances found to inhibit CYP activity. For example, substances in lime juice and red wine have been related to inhibition of CYP3A4 (Bailey, Dresser, and Bend 2003).

3. Structure of Cytochrome P450s

The most important feature of cytochrome P450s is their heme-thiolate prosthetic group. The heme consists of a protoporphyrin IX that contains iron. The thiolate is the ligand on the polypeptide that binds the iron containing heme (Hasler 1999). Therefore it was suggested by the Nomenclature Committee of the International Union of Biochemistry that cytochrome P450s should be called 'heme thiolate protein P450s', to be more accurate (Mansuy 1998).

The proteins consist of about 500 amino acids (Figure 1). Their sequence identity is extremely low, sometimes less than 20% (Hasemann et al. 1995). All eukaryotic cytochrome P450s are membrane bound, either on microsomal, mitochondrial or nuclear membranes. Their prokaryotic orthologs are soluble proteins. The eukaryotic CYPs are anchored in membranes by a hydrophobic helix serving as the amino-terminal membrane-anchoring segment. This anchor is followed by a hinge, consisting of a cluster of prolines that separates the hydrophobic anchor from globular part of the protein. The highest sequence- and structural conservation is around the enzyme's heme. In this region, there is a completely conserved cysteine. This conservation is necessary to the enzymes' monooxygenase functions. It provides the thiolate ligand to bind the heme (Gotoh et al. 1983). This conserved cysteine is located near the carboxy terminus of the protein, for example at amino acid position 432 (of 487) for CYP2C5 (Williams et al. 2000). The region around this conserved amino acid is widely referred to as the Cys-pocket. The P450 signature motif, a motif of ten amino acids in length, including the heme-binding cysteine, (FXXGXXXCXG, with the Xs representing variable amino acids) is also found within this area (Hasler 1999). The most variable regions are those associated with membrane anchoring, targeting of membrane-bound proteins (Werck-Reichhart and Feyereisen 2000) or substrate-binding (Gotoh 1992). Sites involved in the latter are widely referred to as substrate recognition sites (SRS).

Figure 1: (a) Typical primary structure of a membrane-bound P450 protein (b) Topology diagram of the secondary structures and their arrangements of CYP102; blue boxes indicate α -helices; groups of white arrows outlined with dotted lines form β -sheets; element sizes are not in proportion to their length in the primary sequence (c) folded P450 (CYP2C5) in association with the ER membrane (purple); the heme is colored in orange, the substrate in yellow (from Werck-Reichhart and Feyereisen 2000; p. 3-4).

4. The P450 gene superfamily

Because of the early recognized functional similarity of all P450 enzymes (Guengerich 1992) and the enzymes' corresponding genes, all genes encoding for cytochrome P450 enzymes have been put into the so called P450 gene superfamily (or P450 supergene family). This superfamily is organized based on the genes' evolution and divided into families and subfamilies. Genes encoding for enzymes with amino acid sequences having more than 40% sequence identity belong to the same family and those encoding for sequences possessing more than 55% sequence identity belong to the same subfamily (Gonzalez and Nebert 1990). However, there are some exceptions to this rule (see Nebert and Nelson 1991 for details). It has also been discovered that genes belonging to the same subfamily usually comprise tightly linked gene clusters on the same chromosome (Nelson et al. 1996). Thanks to the completion of the human genome project, the final number of human cytochrome P450 genes is set to 57 and the number of pseudogenes to 58. Those human genes and pseudogenes are put into 18 families and 44 subfamilies (Guengerich 2008).

The number of known CYP genes in general has increased rapidly during the last decades. Today, CYP genes have been discovered in all three domains of life and even within a virus (Table 1).

Table 1: Summary of the numbers of known P450 enzymes in various groups of organisms <http://drnelson.utmem.edu/CytochromeP450.html>; cited Oct. 2009).

	No. of P450 genes	No. of P450 families	No. of P450 subfamilies
Animals	3,282	120	500
Insects	1,675	59	338
Animals without	1,607	69	169
Plants	4,266	126	464
Fungi	2,570	459	1,011
Protists	247	62	119
Bacteria	905	196	409
Archaea	22	12	14
Viruses	2	2	2
Total	11,294	977	2,519

4.1 Evolution of the P450 genes

With the increase of the known P450 sequences, the complexity of their phylogeny also increased. Today there is a widely accepted consensus concerning the evolutionary relationships of the most important CYP families and subfamilies (Figure 2).

The ancestral P450 may have existed as early as 3.5 billion years ago, during the eon of archaeobacteria, preceding the separation of bacteria and eukaryota. The initial branching of the P450 phylogeny is dated to 1.4 billion years ago, coinciding with the prokaryote-eukaryote divergence. Because of the existence of the ancestral P450 before the split of archaea, bacteria and eukaryota, all three domains of life feature P450 genes (Lewis, Watson, and Lake 1998).

P450 Phylogenetic Tree

Figure 2: Abbreviated version of the P450 phylogeny mapped against important developments in the evolution of life. The tree was constructed by combining several P450 phylogenetic trees (from Lewis, Watson, and Lake 1998; p. 254).

Co-evolution is considered to be the main driver for the enormous increase in CYP genes over the last hundreds of millions of years, in the form of so called 'animal-plant-warfare'. Animals and plants diverged about 1.2 billion years ago and by that time animals started to ingest plants. As a form of defense, plants developed phytotoxins. Phytotoxins are stress metabolites produced in order to make the plants less potable or digestible. In return, animals developed new P450s to be able to detoxify these phytotoxins. The intensity of this competition probably reached its peak when animal species evolved that were able to live and feed on land. These species encountered new, terrestrial plant forms, which promoted co-evolution. This might be reflected by the 'explosion' of new genes in the CYP2 family about 400 million years ago. Around this time period, at least 30 gene duplications occurred (Gonzalez and Nebert 1990).

It is also believed that an increase in atmospheric oxygen simultaneous to the colonization of land by animals is linked to the diversification of the P450s (Lewis, Watson, and Lake 1998).

The molecular mechanisms responsible for the evolution of the CYP genes and the resulting species specificities in P450 catalytic functions and regulation were suggested to include:

- gene duplication by unequal crossover
- gene inactivation or terminal mutational events
- various other DNA turnover mechanisms, such as replication slippage, inversions and gene conversions (Gonzalez and Nebert 1990).

CYP genes tend to duplicate by unequal crossover because of the presence of numerous highly similar sequences in their flanking regions. These similarities cause a shifted pairing of the two homologous DNA strands during meiotic pairing. This mechanism can lead to gene duplication as well as gene deletions (Hennig 1998). Such a gene deletion is the cause for the Poor Metabolizer phenotype of the Debrisoquine polymorphism (see section 5.3). Gene conversion also happens during meiosis. It leads to non-Mendelian inheritance of traits or alleles and is caused by a non-reciprocal transfer of information of one chromatid to another.

Two mechanisms are possible: The first begins with a double-strand break (and the loss of a DNA segment). Then a homologous DNA duplex can act as a donor and replace the lost DNA fragment on the other chromatid. The second mechanism is the result of the mismatch repair of heteroduplex DNA. Heteroduplex DNA is created when two single strands of two chromatids of a heterozygous locus combine. A repair system replaces the mismatched base pairs on one strand (Lewin 2000).

4.2 CYP Nomenclature

Because of the rapidly increasing number of described P450 sequences it has been recommended by the Cytochrome P450 Nomenclature Committee to assign names to newly discovered genes or enzymes the way it was first described in Nebert et al. 1991, in order to avoid incorrect assignments or duplications of gene names.

For a gene, the italicized root symbol CYP (in mouse: *Cyp*) is used, denoting **C**ytochrome **P**450. This is followed by an Arabic number designating the P450 family, a letter representing the subfamily and a number indicating the individual gene. If it is assumed that there is no other subfamily and no other functional gene in a particular family, it's not needed to specify subfamily and gene number (e.g. *CYP5* or *CYP17*). If at least a second subfamily in one family exists, the subfamily letter and the genes number must be added. Formerly, in mouse genes, the final number was preceded by a hyphen but that's no longer in use today. In the case of a pseudogene, a 'P' ('ps' in mouse and *Drosophila*) is added to the gene number.

The same nomenclature is used for the genes' products and transcripts. Thereby, the all capital letters rule applies also to mouse and *Drosophila*. The cDNA is written in italicized letters, like the gene itself. The mRNA and the final protein are written in non-italicized letters. Usually, the naming of genes is carried out chronologically, meaning that later submitted sequences will get higher numbers or letters, or higher letters regarding new subfamilies (Nebert and Nelson 1991; Nelson et al. 1993; Nelson et al. 1996).

For instance, the naming of *CYP2D6* (and its orthologous gene in the mouse) and its products is carried out this way:

	DNA/gene		mRNA		cDNA		Protein
Human:	<i>CYP2D6</i>	→	CYP2D6	→	<i>CYP2D6</i>	→	CYP2D6
Mouse:	<i>Cyp2D22</i>	→	CYP2D22	→	<i>CYP2D22</i>	→	CYP2D22

If possible, orthologs are given the same name, even if they don't fit the 55% rule for the same subfamily or the 40% rule for the same family (Nelson 2006).

4.3 Polymorphisms in the P450s and their medical consequences

Several human CYP genes have been found to feature high levels of sequence variation. It has been additionally observed that the metabolic conversion and excretion rates of drugs vary immensely between individuals.

Four major classes of phenotypes, regarding to the ability to metabolize a certain drug have been defined: poor metabolizers (PMs), intermediate metabolizers (IMs), extensive metabolizers (EM) and ultrarapid metabolizers (UMs). A large part of this phenotypic variation is a result of variation in CYP genes (van der Weide and Steijns 1999).

To date, not much is known about sequence polymorphism of the P450 genes in non-human mammals. Even though the genetic diversity in human populations is rather high compared to other mammalian species, it is known that different strains within animal species show significant differences in the metabolism of some xenobiotics (Lewis, Watson, and Lake 1998). Mutation in the CYP genes can result in P450 enzymes with reduced, increased, abolished or altered enzyme activity. The total loss in a certain enzyme activity can be caused by the deletion of the whole gene, but also by mutations that create altered splicing, frame shifting, preterm stop codons, deleted transcriptional start sites and deleterious amino acid changes (Ingelman-Sundberg 2004). Altered substrate specificity can be caused by mutations in the enzymes substrate recognition sites. This has been documented on *CYP2D6**17-allele of the *CYP2D6* gene (Oscarson et al. 1997). Increased enzyme activity is caused by the duplication or multiplication of the CYP genes and has been well-described for *CYP2D6* (Johansson et al. 1993) but was also observed in *CYP2A6* (Rao et al. 2000).

Some of the polymorphisms of most interest for medicine are those found in the genes *CYP2D6*, *CYP1A1*, *CYP2C19* and *CYP2E1*. An allelic variant of *CYP1A1* is believed to be associated with lung cancer. This allele is caused by a point mutation in the heme-binding region. It has been recognized that a considerable high percentage of lung cancer patients carry this allele (Lewis, Watson, and Lake 1998). *CYP2C19* metabolizes a common ulcer drug, Omeprazole. There seems to be a higher frequency of poor metabolizers of this drug in Asians (23%) than in Caucasians (3 – 5%)(Mizutani 2003). *CYP2E1* metabolizes multiple known carcinogens. Some polymorphism has been recognized to be correlated with cancer and allele frequencies differ among ethnic groups (Hirvonen et al. 1993).

However, the clinical significance of an enzyme's polymorphism is not always clear and many other factors, like the environment play a major part in drug metabolism as well as cancer susceptibility (e.g. drug-drug interaction or the exposure to procarcinogens).

The knowledge of the genotype regarding a particular P450 gene is of certain importance in cases when the drug of interest (metabolized by the gene's product) has a relatively low therapeutic

window, meaning that its therapeutic concentration lies just below a concentration that's already toxic for the organism. Poor metabolizers, not able to remove the applied drug will soon suffer from side effects of the drug, while ultra-rapid metabolizers will not experience the therapeutic level at the same given dosage. The inverse effect is observed when the administered drug is actually a prodrug that needs to be activated by a P450 enzyme (van der Weide and Steijns 1999; Ingelman-Sundberg 2004).

Several approaches have been developed to estimate the metabolic capacity of an individual to metabolize a specific drug. Phenotyping is done by the administration of a probe drug that is known to be metabolized only by one certain P450 enzyme. Thus, the so-called metabolic ratio (ratio of the unchanged drug to the metabolized product in blood serum or urine) is measured. This method has the advantage of taking into account all internal and external factors that contribute to the metabolism of the drug. The disadvantages are the rather complicated procedures and the possible risks of adverse drug reactions. Genotyping reveals defined mutations in the CYP gene of interest that cause specific phenotypes. It has the advantage of relatively quick results and non-interference with diseases or administered drugs. Multiple techniques have been developed for genotyping several CYP genes (Ingelman-Sundberg 2004).

5. CYP2D6

Of all cytochrome P450s, CYP2D6 (also called debrisoquine-4-hydroxylase), is by far the best studied. Its clinical relevance is due to its involvement of the metabolism of at least 20% of commonly used drug in medical treatment (Bertz and Granneman 1997). It is also well known because of the high phenotypic inter-individual variation in drug metabolism that is caused by genotypic variation of the gene (see section 5.5).

The mRNA of the enzyme is mainly expressed in the liver but is also found in some extrahepatic tissues like lung, brain and gastrointestinal tract. However, the expression levels in these tissues are much lower than in the liver and sometimes the mRNA is hardly detectable. Therefore, some of the findings of extrahepatic mRNA are rather controversial (Zanger, Raimundo, and Eichelbaum 2004).

5.1 Structural properties of the CYP2D6 enzyme

The human CYP2D6 enzyme contains 497 amino acid residues. Its crystal structure was described by Rowland et al. (2006) (Figure 3). Its structure shows the general fold of a cytochrome P450 and it shares many common features with CYP2C9. The active site cavity of the enzyme is rather large with a volume 540 Å³ and resembles a right foot in shape. The big size of the active site cavity is believed to be one reason of the wide diversity in the substrate recognition of CYP2D6. The heme is connected to the protein by hydrogenbonding interactions with the residues of six amino acids (Arg-101, Trp-128, Arg-132, His-376, Ser-441). The conserved cysteine residue is located at position 443 and is penta-coordinated with the heme iron (Smith, Achland, and Jones 1997; Rowland et al. 2006).

Figure 3: Ribbon diagram of the human CYP2D6 protein and its prosthetic group. Red coiled structures represent alpha-helices, green arrows indicate beta-sheets (from Rowland et al. 2006; p. 7618).

5.2 Substrates, Inhibitors and Inducers of CYP2D6

Even though CYP2D6 accounts for less than 2% of the total hepatic cytochrome P450s, it's involved in the metabolism of a large part (25%) of clinically used drugs. The list of CYP2D6 substrates is still steadily growing. All of the substrates are lipophilic bases with a protonable nitrogen atom. Most of the tricyclic antidepressants are metabolized by CYP2D6 and it is also involved in the metabolic pathways of many antiarrhythmics, analgetics and further antidepressants (Table 2).

With CYP2D6 as well, some of the drug substrates are actually prodrugs, that need to be activated by the enzyme, such as Codeine or Hydrocodone (Zanger, Raimundo, and Eichelbaum 2004; Ingelman-Sundberg 2005).

Table 2: Drugs of different therapeutic classes metabolized by CYP2D6 (according to Zanger, Raimundo and Eichelbaum 2004).

Therapeutic class	Drug
Analgesics/antitussives	Codeine
	Dextromethorphan
	Dihydrocodeine
	Ethylmorphine
	Hydrocodone
	Norcodeine
	Oxycodone
Anti-ADHD drug	Atomoxetine
Antiarrhythmics	Aprindine
	Encainide
	Flecainide
	Mexiletine
	N-propylajmaline
	Procainamide
	Propafenone
	Sparteine
Antidementia drugs	Galanthamine
	Nicergoline
Tricyclic antidepressants	Amitriptyline
	Clomipramine
	Desipramine
	Imipramine
	Nortriptyline
	Other antidepressants
	Desmethylcitalopram
	Fluoxetine
	Fluvoxamine
	Maprotiline
	Mianserin
	Minaprine
	Mirtazapine
	Paroxetine D
	Venlafaxine
Antidiabetic	Phenformine

Table 2: (cont.)

Antiestrogen	Tamoxifen
Antihypertensives	Debrisoquine
	Guanoxan
	Indoramin
Antiemetics	Dolasetron
	Ondansetron
	Tropisetron
Antihistamines	Mequitazine
	Promethazine
Antipsychotics	Haloperidol
	Perphenazine
	Risperidone
	Thioridazine
	Zuclopenthixol
Appetite suppressant	Dexfenfluramine
Beta adrenergic blocking agents	Alprenolol
	Bufuralol
	Bunitrolol
	Bupranolol
	Carvedilol
	Metoprolol
	Propranolol
	Timolol
Calcium antagonist	Perhexiline
MAO-inhibitors	Amiflamine
	Brofaromine
Recreational drugs	Methoxyamphetamine
	MDMA, MDME
Vasodilators	Cinnarizine
	Flunarizine

It has also been suggested that CYP2D6 possesses some endogenous substrates. It might play an important role in the metabolism of neurotransmitters. The metabolism of indoethylamines pinoline (6-methoxy-1,2,3,4-tetrahydro-beta-carboline) and 5-methoxy-N,N-dimethyltryptamine (5-MDMT), as well as 5-methoxytryptamine is highly influenced by CYP2D6 (Zanger, Raimundo, and Eichelbaum 2004).

CYP2D6 isn't inducible by common enzyme inducers. More likely, multiple copy number is the cause of increased enzyme function in some individuals (Ingelman-Sundberg 2004). However, a 30% increase in enzyme activity has been observed after treatment with rifampin and also dexamethasone seems to have a positive influence on enzyme activity.

CYP2D6 is inhibited by a number of drugs leading to numerous clinically significant drug interactions. Amongst these drugs acting as inhibitors are quinidine, fluoxetine, flecainidine, paroxetine, mexiletine and propafenone (Zhou 2009).

5.3 The Debrisoquine/Sparteine polymorphism caused by CYP2D6

Long before the cDNA of the CYP2D6-gene was cloned and sequenced and various sequence polymorphisms described, it was known that the reaction to specific drugs like debrisoquine or sparteine varies tremendously between individuals. During the 1970s and 1980s, several research groups showed that some patients receiving drug therapy were unable to oxidize specific drugs and suffered from severe adverse drug reactions (Zanger, Raimundo, and Eichelbaum 2004). Such adverse drug reactions included nausea, blurred vision and diplopia after administration of sparteine, while incapacitating orthostatic hypotension was observed after the intake of debrisoquine. Later it has been shown that poor metabolizers of sparteine are also poor metabolizers of debrisoquine and vice versa and further, that the cause of this phenotype is a monogenetic, autosomal inherited, recessive trait. During the following years, these same poor metabolizers showed difficulty metabolizing a series of additional drugs, such as bufuralol and metoprolol (Distlerath et al. 1985). The enzyme responsible for this phenomenon, today known as CYP2D6, was isolated and purified by Distlerath et al. (1985; Skoda et al. 1988). Several enzymatic studies followed and knowledge of CYP2D6 polymorphism grew. Also, first molecular genetic studies were carried out on the debrisoquine polymorphism and a specific RFLP pattern was linked to the poor metabolizer phenotype (Sachse et al. 1997). Researchers initially believed that one specific genotype caused the PM phenotype but over time more alleles with an impaired or non-functional enzyme were described (see section 5.5).

Soon it became clear that ethnicity plays a role in the frequency of poor metabolizers. Today it is estimated that about 7% of the Caucasian population are poor metabolizers, with *CYP2D6*4* being the main allele causing this phenotype (Mizutani 2003). In Eastern Asians, the frequency of poor metabolizers is 1.5% and the most common allele causing PMs is *CYP2D6*5* (Johansson et al. 1993).

5.3.1 Ultrarapid Metabolizers

The phenotypes regarding the debrisoquine/sparteine polymorphism were identified by measuring the metabolic ratio (MR) between urinary debrisoquine and its metabolized form, 4-hydroxydebrisoquine after the administration of a specific dose of debrisoquine. Poor metabolizers are defined as individuals having a MR of greater than 12.5 and extensive metabolizers as individuals with a ratio lower than this value. In 1992, a small percentage of the Swedish population was found to have a very low metabolic ratio (less than 0.2). Individuals with such a high turnover rate of a drug would need a much higher dosage of it so that they reach therapeutic plasma

levels and experience a therapeutic effect. In 1993, Johansson et al. described the genetic basis behind the ultrarapid metabolism of debrisoquine. Individuals with this extremely low metabolic ratio, today known ultrarapid metabolizers, carry up to 12 extra copies of the CYP2D6 gene (Ingelman-Sundberg 2005; Section 5.5.4)

The frequency of ultrarapid metabolizers is also influenced by people's origin. About 30% of Ethiopians, 10% of people in Spain, Italy and Turkey but only 1-2% of people from Northern Europe and almost no Asian people are of the UM phenotype (Figure 4) (Gonzalez et al. 1988).

Figure 4: Global distribution of the frequency of ultrarapid metabolizers. Arrow indicates the muslim migration which may explain the high percentage of UMs in Spain (from Ingelman-Sundberg 1999; p. 452).

The advent of molecular biology in the middle of the 1980s significantly changed the way in which the CYP2D6-polymorphism was studied. As mentioned above, the first molecular genetic studies were done by RFLP pattern analysis. In 1988, Gonzalez et al. cloned the cDNA of the human *CYP2D6*-gene (known then as *CYPdb1*) for the first time and from then on, the *CYP2D* locus was the focus of numerous molecular genetic studies.

5.4 The CYP2D locus

The human *CYP2D6*-gene and two pseudogenes (*CYP2D7P* and *CYP2D8P*) are located on the long arm of chromosome 22 (region 22.13.1). The two pseudogenes are believed to have arisen via gene duplication events. *CYP2D6*, *CYP2D7P* and *CYP2D8P* are tightly linked and represent the *CYP2D* gene cluster (Figure 5). These three genes are the only members of the *CYP2D* subfamily in humans.

Table 3: Comparisons of the exon and intron length between *CYP2D6*, *CYP2D7* and *CYP2D8P* (according to Kimura et al. 1989).

	<i>CYP2D6</i>	<i>CYP2D7P</i>	<i>CYP2D8P</i>
Exon 1	268	269	265
Intron 1	703	701	1620
Exon 2	172	172	172
Intron 2	550	528	546
Exon 3	153	153	153
Intron 3	88	88	88
Exon 4	161	161	161
Intron 4	433	425	449
Exon 5	177	177	177
Intron 5	190	192	186
Exon 6	142	142	142
Intron 6	207	194	204
Exon 7	188	188	185
Intron 7	454	454	449
Exon 8	142	142	142
Intron 8	98	98	96
Exon 9	252	252	232

The *CYP2D6* gene consists of nine exons and eight introns, like all members of the CYP2 family. It is 4378 bp in length from the transcription start site to the polyadenylation site (Table 3). The TATA box of the *CYP2D6* gene is located -28 to -24 bp of the transcription start site. The 5' untranslated region of the mRNA is 88 base pairs long. The 3' untranslated region from the stop codon to the polyadenylation site is 73 bp in length (Kimura et al. 1989).

CYP2D7P resides 9331 bp upstream of *CYP2D6*. (GenBank Gene ID 1564). It is 4336 bp in length from the transcription start site to the polyadenylation site. Its 5' and 3' untranslated regions are of the same length as those of *CYP2D6*.

It possesses a T-insertion in its first exon (pos. 226) which disrupts the protein reading frame. The shifted frame also causes a premature termination at positions 2587 to 2589 in exon 5. The exonic sequence of *CYP2D7P* shows 97% sequence identity to *CYP2D6*.

CYP2D8P (GenBank Gene ID 1568) lays 4666 bp upstream of *CYP2D7P* and is 5267 bp in length. It contains several insertions and deletions in its exonic sequences that cause a completely disrupted open reading frame. Because of these numerous mutations it is believed that this gene has been a pseudogene for a longer time than *CYP2D7P*. *CYP2D8P* has an exonic sequence identity of 92% to *CYP2D6* (Zhou 2009).

Figure 5: Structural organization of the human *CYP2D* locus on chromosome 22. Numbered black bars indicate the genes' exons.

This arrangement of the *CYP2D* locus and the sequence of the *CYP2D6* gene described above are defined as the reference sequence or wild-type allele because it was the first fully described sequence of *CYP2D6* (Kimura et al. 1989). Since then, the enormous variability of this gene has been uncovered.

5.5 Alleles of *CYP2D6*

By April 2010, at least 78 allelic variants (labeled *CYP2D6**2 to *CYP2D6**74) have been described in humans additional to the wild-type allele (*CYP2D6**1). There are also numerous accessory subvariants of these alleles (indicated by a letter after the alleles' numbers). All of these subvariants share the same key mutations that characterize the specific allelic variant. For instance, all known subvariants of the *CYP2D6**6 allele (*CYP2D6**6A, *CYP2D6**6B, *CYP2D6**6C and *CYP2D6**6D share the same T deletion at position 1707 (<http://www.cypalleles.ki.se>).

The mutations causing these alleles can be point mutations, insertions, deletions, gene rearrangements or deletions or duplications/multiplications of the whole gene. At least 196 single nucleotide polymorphisms (SNPs) have been described. At least 52 of them are non-synonymous SNPs, indicating a potential impact on the protein's function. These mutations can result in alleles with increased, decreased or missing enzyme activity. Some alleles alter the protein's constitution but have no effect on the effectiveness of the enzyme (Table 4). As indicated above, the frequency of the different alleles varies with ethnicity (Zhou 2009).

Table 4: Most common alleles causing different effects on the *CYP2D6* activity (according to (Zhou 2009) and Human Cytochrome P450 (CYP) Nomenclature Committee [cited 2009 Oct 15]; Available from: <http://www.cypalleles.ki.se>).

Normal function	Decreased function	No function	Increased function
<i>CYP2D6</i> *1A	<i>CYP2D6</i> *10	<i>CYP2D6</i> *3	<i>CYP2D6</i> *1XN
<i>CYP2D6</i> *2A	<i>CYP2D6</i> *14	<i>CYP2D6</i> *4	<i>CYP2D6</i> *2XN
<i>CYP2D6</i> *17x2	<i>CYP2D6</i> *17	<i>CYP2D6</i> *6	<i>CYP2D6</i> *35X2
<i>CYP2D6</i> *27	<i>CYP2D6</i> *18	<i>CYP2D6</i> *5	<i>CYP2D6</i> *35X2
<i>CYP2D6</i> *35	<i>CYP2D6</i> *36	<i>CYP2D6</i> *7	
<i>CYP2D6</i> *39	<i>CYP2D6</i> *41	<i>CYP2D6</i> *8	
<i>CYP2D6</i> *41x2	<i>CYP2D6</i> *47	<i>CYP2D6</i> *11	
<i>CYP2D6</i> *48	<i>CYP2D6</i> *49	<i>CYP2D6</i> *12	
	<i>CYP2D6</i> *50	<i>CYP2D6</i> *13	
	<i>CYP2D6</i> *51	<i>CYP2D6</i> *14	
	<i>CYP2D6</i> *54	<i>CYP2D6</i> *15	
	<i>CYP2D6</i> *55	<i>CYP2D6</i> *16	
	<i>CYP2D6</i> *57	<i>CYP2D6</i> *18	
		<i>CYP2D6</i> *19	
		<i>CYP2D6</i> *20	
		<i>CYP2D6</i> *21	
		<i>CYP2D6</i> *38	
		<i>CYP2D6</i> *40	
		<i>CYP2D6</i> *42	
		<i>CYP2D6</i> *44	
		<i>CYP2D6</i> *56	
		<i>CYP2D6</i> *62	

5.5.1 Alleles causing normal enzyme function

These alleles have been found to have no influence on the metabolic activity of the enzyme, even if they contain multiple mutations compared to the wild-type allele (Sachse et al. 1997).

The most common allele in Caucasians is the wild-type allele, *CYP2D6*1*, with an allele frequency of 36.4%. In Asians, the allele seems to have a quite similar frequency, (e.g. in Koreans it has been found in 32.2% of the population). It's followed by *CYP2D6*2*, having a frequency of 32.4% in Caucasians. Individuals harboring these alleles are called extensive metabolizers (Lee et al. 2009, Zhou 2009).

5.5.2 Alleles causing a decreased enzyme function

People carrying one of these alleles carry mutations that impair the activity of the CYP2D6 enzyme. The mutations cause changes in the amino acid sequence that result in decreased protein stability, reduced substrate-enzyme affinity or disrupted substrate recognition.

A very common allele in Asians is the *CYP2D6*10* allele (33-43%), whereas it quite rare in Caucasians and other ethnic groups (2-5%). In this allele a SNP causes a proline → serine substitution in the proline-rich region. The *CYP2D6*17* allele is common in populations in Africa (9-34%) but almost absent in Caucasians. It causes several amino acid substitutions resulting in an alteration of the protein structure and substrate recognition. Individuals that carry one or two alleles that cause impaired enzyme activity show either called intermediate or poor metabolizer.

5.5.3 Alleles causing no enzyme function

These alleles do not encode a functional enzyme. A person carrying one of these alleles in the homozygous state or two of them in the heterozygous state displays the PM phenotype.

Mutations causing the total loss of the enzymes activity include:

- deletion of the entire gene (e.g. *CYP2D6*5*)
- SNPs or small deletions or insertions that cause frame shifts or altered splicing (e.g. *CYP2D6*4*)
- mutations that still allow the transcription of a full-length protein but the protein is not functional (e.g. *CYP2D6*12*)
- formations of hybrid genes (e.g. *CYP2D6*16* results in a *CYP2D6-CYP2D7P* hybrid gene)

The deletion of the whole gene is usually caused by unequal recombination after unequal crossover of two homologous but non-allelic sequences or after the formation of a loop structure on

one chromosome. *CYP2D6**5 is the second most common allele causing lack of enzyme function in citizens of Great Britain. Its frequency is about 4-7% and about the same in all ethnic groups. The person whose DNA of chromosome 22 was first sequenced for the Human Genome Project carried the *CYP2D6**5 allele and therefore *CYP2D6* doesn't show up in the Human Genome Project. *CYP2D6**4 represents more than 75% of the mutant alleles in Swedish Caucasians and has an allele frequency of 22% in these people. It's the most common allele of the PM phenotype in Caucasians. It carries several SNPs, one of them causes a splicing defect in exon 3. *CYP2D6**12 carries four SNPs that do not effect splicing or transcription, but the enzyme has almost no activity. *CYP2D6**16 contains a gene that possesses exons 1 to 7 of *CYP2D7P* and exons 8-9 of *CYP2D6*. This gene was generated by large deletions of the pseudogene caused either by unequal crossover or looping-out (Zhou 2009).

5.5.4 Duplications or multiduplications of the *CYP2D6*-gene as a cause of increase enzyme function

In 1993, a Swedish research group reported the occurrence of 0, 1, 2, 3, 4, 5 and 13 copies of a *CYP2D6* allele in Swedish people. Johansson et al. (1993) defined characterized their allele as "*CYP2D6L*-variant", today known as *CYP2D6**2A. Individuals carrying more than two active copies of the gene exhibit the ultrarapid metabolizer (UM) phenotype. The amplification of the gene showed an autosomal dominant inheritance pattern and this was the first time, a stably inherited amplification of a gene was described (Johansson et al. 1993; Zhou 2009). The *CYP2D6**2A allele itself results in a normally functioning enzyme. But when it is duplicated or multi-duplicated, the overall enzymatic activity is increased. Later, further alleles containing duplicated genes have been identified: To date, the alleles *CYP2D6**1, *CYP2D6**4, *CYP2D6**6, *CYP2D6**10, *CYP2D6**17, *CYP2D6**29, *CYP2D6**35, *CYP2D6**43 and *CYP2D6**45 have been found in such tandem repeats.

Alleles carrying a duplication or multi-duplication are indicated with a *XN* after the final variant number or letter with N being replaced with the number of genes in the genotype. For example *CYP2D6**2X4 describes a fourfold amplification of the *CYP2D6**2-allele. Not all duplication events result in an increased enzyme activity. For example, the *CYP2D6**4 allele causes a no enzyme activity at all and therefore the *CYP2D6**4XN variants also result in no enzyme function. The frequency of duplication or multi-duplication of the gene is very dependent on ethnicity, as well.

The most frequent allele containing a duplication of *CYP2D6* is *CYP2D6**2X2 is especially common in Africans (28-56%) whereas it's much rarer in Asians or Caucasians. *CYP2D6**4XN variants have been found in 37% of Asians but show a low frequency in Caucasians (Zhou 2009).

5.6 Genetic mechanisms leading to duplication and multiduplication of *CYP2D6*

It is assumed that the molecular events leading up to five copies of the gene are unequal crossover events during meiosis. This is supported by the fact that a repetitive sequence (CCCACCCTTC) has been found four times in the *CYP2D* gene cluster. Also, Alu sequences (short repetitive sequences found exclusively in primate genomes (Shankar, Kataria, and Mukerji 2009)) were discovered in the 3' regions of *CYP2D6* and *CYP2D7P*. Repetitive sequences like these are known to favor unequal crossover events. There are several models about the exact mechanisms of the duplication and further multi-duplication events of the *CYP2D6*-gene involving these sequences. A very common one suggests that after the first duplication event (caused by unequal crossover of the repetitive sequences), the Alu sequences surrounding the duplicated *CYP2D6* loop out the duplicated sequence. This looped out area can now be duplicated during sister chromatid exchanges (Figure 6) (Lundqvist, Johansson, and Ingelman-Sundberg 1999).

Figure 6: Unequal exchange model as the cause of the duplication and multiduplication of *CYP2D6*. (I) Formation of a duplicated *CYP2D6*-gene via unequal crossover. (II) Duplicated region forms a hairpin loop and is looped out during another unequal crossover event. (III) Formation of additional gene copies with the same mechanisms. Black bars indicate Alu sequences (from Lundqvist, Johansson, and Ingelman-Sundberg 1999; p. 335).

In contrast, it was suggested that alleles carrying 13 copies of the gene have been generated via unequal segregation and extrachromosomal replication of acentric DNA. Thereby, a chromosomal breakage forms an acentric, extrachromosomal DNA structure containing the replicated element. This molecule is then replicated several times (in the course of the cell's cycle) and these multiple copies undergo unequal segregation during meiosis. Finally, the extrachromosomal elements become reintegrated into the native locus in the chromosome (Figure 7) (Kimmel, Axelrod, and Wahl 1992).

Figure 7: Model of the creation of alleles with 13 copies of *CYP2D6*. First, an acentric DNA molecule is formed by a double DNA breakage. This molecule is the replicated multiple times. Multiple copies are unequally segregated during meiosis. At the end, copies are reintroduced into the *CYP2D* locus (from Lundqvist, Johansson, and Ingelman-Sundberg 1999; p. 336).

5.7 *CYP2D6* in other mammals

More recently, numerous sequencing projects of non-human mammal-genomes have been completed leading to the steady increase in the number of P450 genes characterized. The first non-human mammal whose mRNA of its *CYP2D6*-ortholog has been sequenced was *CYP2D4*, the rat ortholog of *CYP2D6* (Matsunaga, Umeno, and Gonzalez 1990).

By February 2010, eight fully characterized *CYP2D6* orthologs of non-human mammals had been described (Table 5 & Figure 8).

Table 5: List of fully characterized *CYP2D6* orthologs in non-human mammals (data from GenBank Homepage [http://www.ncbi.nlm.nih.gov/Genbank]; cited 2009 Oct 15).

Organism	Name of <i>CYP2D6</i> ortholog	Accession Numbers
<i>Pan troglodytes</i> (Common chimpanzee)	<i>CYP2D6</i>	NM_001040622.1
<i>Pan paniscus</i> (Pygmy chimpanzee)	<i>CYP2D6</i>	DQ282163.1
<i>Macaca mulatta</i> (Rhesus macaque)	<i>CYP2D6</i>	XM_001107251.1
<i>Bos taurus</i> (Cattle)	<i>CYP2D6</i>	NM_174529.2
<i>Equus caballus</i> (Horse)	<i>CYP2D50</i>	NM_001111306.1
<i>Mus musculus</i> (House mouse)	<i>CYP2D22</i>	NM_019823.2
<i>Rattus norvegicus</i> (Norwegian rat)	<i>CYP2D4</i>	NM_138515.1
<i>Canis lupus familiaris</i> (Dog)	<i>CYP2D15</i>	NM_001003333.1

Figure 8: Intron & exon organization the nine fully characterized mammalian *CYP2D6*-orthologs.

In addition to these fully characterized genes, the mRNA sequences of two putative orthologs on monkeys, namely *CYP2D17* of Crab-eating macaque (*Macaca fascicularis*) and *CYP2D19* of White-tufted-ear marmoset (*Callithrix jacchus*) have been sequenced. Their GenBank accession numbers are U38218.1 and D29822.1, respectively.

The polymorphism of *CYP2D6* in humans is well known and extensively studied. In contrast, little is known about such polymorphism in other mammals. However, there are some enzymatic studies on the P450 enzymes in general, especially on pets and meat-producing animals because they often receive veterinary treatment. But most of these studies focus on substrate specificity and expression of the enzymes (Garfinkel 1958; Ioannides 2006; Trepanier 2006; Seliskar and Rozman 2007).

In the course these enzymatic studies, some evidence of intra-species polymorphisms regarding enzyme activities have been found. In dogs, significantly different propofol hydroxylase activities (an enzymatic reaction catalyzed by a cytochrome P450) have been observed between beagles and greyhounds (Court et al. 1999). And even within one breed, the beagle, significant

differences in the ability to metabolize certain drugs were found and the phenotypes poor and extensive metabolizers have been defined (Paulson et al. 1999). Also in deer, variation in CYP activity was found between different breeds (Ioannides 2006). Further, horses were also suggested to be classified into extensive and poor metabolizer (Bull et al. 2003). In cats (Fink-Gremmels 2008) and cattle (Ioannides 2006), gender differences in enzymatic P450 activities have been described.

Another important discovery is the fact that even though very high sequence identity between the orthologs in different species is found (Table 6), the enzymes do not have necessarily the same substrate specificity or inhibition pattern. For instance, *CYP2D15* in dogs exhibits enzyme activities like its human ortholog, but shows inhibition patterns after treatment with quinine and quinidine that are more similar to rat *CYP2D1* (Roussel et al. 1998).

Table 6: Pairwise sequence identities *CYP2D6*-orthologs and isozymes in eleven mammals (in %). Upper part represents the pairwise similarities in the nucleotide sequences in the open reading frame of the genes; lower part represents the sequence similarity in the amino acid sequence of the enzymes.

	<i>Homo sapiens</i>	<i>Pan troglodytes</i>	<i>Pan paniscus</i>	<i>Macaca mulatta</i>	<i>Macaca fascicularis</i>	<i>Callithrix jacchus</i>	<i>Equus caballus</i>	<i>Bos taurus</i>	<i>Rattus norvegicus</i>	<i>Mus musculus</i>	<i>Canis lupus familiaris</i>
<i>Homo sapiens</i>	/	98,4	99	94,5	94,2	91,6	83,9	83,3	78,6	79,2	80,9
<i>Pan troglodytes</i>	98,5	/	99,2	94,6	94,3	91,5	83,8	83	78,7	79,3	80,9
<i>Pan paniscus</i>	99,1	99,3	/	94,8	94,4	91,7	84,1	83,2	78,7	79,3	80,7
<i>Macaca mulatta</i>	93,7	93,5	94,1	/	98,7	91	83,7	82,3	77,9	78,9	80,5
<i>Macaca fascicularis</i>	93,3	93,5	93,7	98,7	/	91	83,7	82,5	77,8	78,8	80,7
<i>Callithrix jacchus</i>	90,9	90,7	90,9	90,5	91,1	/	82,7	81,4	77,9	78,3	79,7
<i>Equus caballus</i>	77,4	77,2	77,4	77,8	78	78,4	/	85,2	78,5	77,8	82,9
<i>Bos taurus</i>	77,6	77	77,6	77,2	77,2	77,6	80	/	76,3	76,1	82,7
<i>Rattus norvegicus</i>	76,4	76,2	76,6	76,2	76,2	75,2	74,4	73	/	90,8	74,9
<i>Mus musculus</i>	75	74,8	75,2	75,6	75,6	73,4	71,8	70	87,2	/	75
<i>Canis lupus familiaris</i>	75,2	74,6	74,8	75,8	76,4	75,4	75	75,4	70	69	/

5.8 Studies on Cytochrome P450s in domestic cats

The role of CYPs in domestic cats has gained more and more attention during the last ten years. There are two major reasons for this: First, cats as a popular companion animal have become frequent veterinary patients and therefore receive a lot of medical treatment (Fink-Gremmels 2008). Second, cats are at high risk of acetaminophen intoxication. This is because of their low levels of

high-affinity acetaminophen-UDP-glucuronosyltransferase (acetaminophen-UGT). This enzyme is responsible of the metabolism of the commonly used drug acetaminophen via glucuronidation and therefore cats exhibit a deficiency in this glucuronidation (Court and Greenblatt 1997). Cats are therefore at high risks of adverse drug reactions from many other common drugs and toxins (Tanaka et al. 2005), giving feline CYPs an even greater importance with respect to their impact on drug metabolism.

The first studies on cytochrome P450 enzyme activities were comparative involving several species. Thereby, significant interspecific differences regarding CYP orthologs were observed. Despite the relative close systematic proximity of domestic cat and domestic dog, their enzymatic activities of various cytochrome P450s differed significantly (Table 7)(Chauret et al. 1997).

Table 7: Comparisons of enzyme activities of various cytochrome P450s in cats and dogs. Percentages represent enzyme activity in cats relative to those in dogs (according to Chauret et al. 1997; Fink-Gremmels 2008).

Human CYP ortholog	Enzyme activity	% enzyme activity in cats compared to dogs
CYP1A1/2	Phenacetin-O-deethylase	61,7
CYP2A6	Coumarin-7-hydroxylase	69,9
CYP2C8/9	Tolbutamide	28,6
CYP2C19	S-methenytol-4-hydroxylase	62,9
CYP2D6	Dextrometorphan-O-demethylase	276,9
CYP2E1	Chlorzoxazone-6-hydroxylase	77,4
CYP3A4	Testosterone-6 β -hydroxylase	96,9

Phipps et al. (2003) reported strong similarities of the enzyme activity of CYP2D6 in cats and humans. The feline version of this enzyme was inhibited by the same chemicals quinidine and quinine and had the same substrate specificity. Another feline enzyme, CYP2A4 showed some differences in its enzyme kinetics to its human counterpart. Also, the absence of CYP2C9 and CYP2C19-activity was observed (Phipps, Collins, and Blanchflower 2003).

Shah et al. (2007) further characterized some enzyme activities in cats and related them to specific cytochrome P450s. They also discovered that with some of these enzymes, there are significant sexual differences in their activity (Table 8). CYP2D-activities were three times higher in female cats but CYP2A activities were only one-fifth compared to male cats. It was suggested that substrates these enzymes should be prescribed in different dosages for male and female cats (Shah et al. 2007).

Table 8: Comparison of enzyme activities of four CYP-subfamilies in male and female cats (according to Shah et al. 2007).

CYP subfamily	Enzyme activity	Sexual differences in enzyme activity
CYP1A	Ethoxyresorufin-O-deethylation	no
CYP2C	Tolbutamide hydroxylation	no (total enzyme activity negligible)
CYP2D	Bufuralol 1'-hydroxylation	yes
CYP3A	Midazolam 1'- & 4-hydroxylation	yes

To date, the mRNAs of the following feline *CYPs* have been sequenced:

Gilep et al. (2000) submitted the sequence of the mRNA of *CYP17A1* of the domestic cat to Genbank, but unfortunately, no further information of this feline gene is available. In 2005, the cDNAs of three forms of *CYP2E* – subfamily members have been cloned and sequenced. Apparently, cats express three isoforms of *CYP2E* (*CYP2E-a*, *CYP2E-b* and *CYP2E-c*), in contrast to humans who only express one. The different forms of cDNAs also showed a different expression pattern. mRNA of *CYP2E-a* was found in liver-, lung-, kidney-, stomach- and intestines-tissue and mononuclear cells, whereas *CYP2E-b*- and *CYP2E-c*- mRNA was only found in liver and mononuclear cells. It was suggested that cats have at least two *CYP2E*-genes: *CYP2E1* and *CYP2E2* (Tanaka et al. 2005). One year later, the same research group cloned and sequenced cDNA of the cat *CYP1A1* and *CYP1A2*. The mRNA of *CYP1A1* was found in liver, lung, pancreas, stomach and small intestines. In contrast to that, mRNA of *CYP1A2* was only observed in liver tissue (Tanaka et al. 2006).

Despite the knowledge of the eminent role of *CYP2D6* in drug metabolism, and the probable interspecific variation of this enzymes' activity in domestic cat, the *CYP2D6*-gene hasn't been sequenced in this species. Additionally, little is known about the evolution and possible selection pressures on this gene. Therefore, the aim of this study was, to sequence and characterize the *CYP2D6*-gene in several new species and, by comparing the sequence differences of this gene between closely related species (namely seven members of the family Felidae), as well as with more distantly related species (members of several orders of mammals), to gain first insights into the evolution and eventual selection pressures on it.

II. MATERIAL AND METHODS

Liver samples of three domestic cats (*Felis catus*), four cheetahs (*Acinonyx jubatus*), three bobcats (*Lynx rufus*), three leopard cats (*Prionailurus bengalensis*), three cougars (*Puma concolor*), three lions (*Panthera leo*), and three Siberian tigers (*Panthera tigris altaica*), were stored at -70°C or in liquid nitrogen. One additional liver sample from a domestic cat was received from a local veterinarian and stored at 4°C in RNAlater (Ambion, Austin, TX).

II.1 Primer Design

Primers for Polymerase Chain Reactions were designed against the predicted feline CYP2D6-mRNA sequence received from GARFIELD online browser (Pontius and O'Brien 2007) and placed in conserved regions based on an alignment of mRNA sequences from *Homo sapiens*, *Pan troglodytes*, *Macaca mulatta*, *Bos taurus*, *Canis lupus familiaris* and *Felis catus*), resulting in selected primers FcCYP2D6-190fb (5'-AGCCCAGCACGCACTGAG-3'), FcCYP2D6-4191r (5'-AGGAGCACAGGGATGGAGTTCAGG-3'), FcCYP2D6-6125f (5'-ATAAAGCTGTGAGCAACG-3') and FcCYP2D6-8179r (5'-ACTGGTTTATTGCATATCAGG-3').

II.2 RNA Extraction & Reverse Transcription

Total RNA was extracted using RNeasy Mini Kit (Qiagen) or SV Total RNA Isolation System (Promega) according to manufacturer's guidelines and stored at -60°C. cDNA was prepared using the ImProm-II™ Reverse Transcription System (Promega) and reverse transcription reactions were carried out with both random and oligodT primers.

II.3 Amplification and Sequencing of CYP2D6

The 5'-fragment was amplified using primer pair FcCYP2D6-190fb/FcCYP2D6-4191r and cDNA of reverse transcription reactions with random primers. The 3'-fragment was amplified using primer pair FcCYP2D6-6125f/FcCYP2D6-8179r 4191r and cDNA of reverse transcription reactions with oligodT primers. Each 25 ul reaction mix consisted of 4 µl reverse transcription reaction, 25 pmol of both forward and reverse primer, 5 µl GC-Buffer, 0.75 µl DMSO, 0.25 mM dNTP mix, 0.5 U Phusion™ DNA Polymerase (New England Biolabs) and H2O or 2 µl AmpliTaq® 360 Buffer, 10×, 1 mM MgCl₂, 1 mM dNTP mix, 0.75 U AmpliTaq® 360 DNA Polymerase (Applied Biosystems) and H2O. PCR-conditions were as follows: denaturation of 98°C for 30 sec, followed by 35 cycles at 98°C for 10 sec, 60°C for 30 sec and 72°C for 35 sec and final extension at 72°C for 10 min for the 5'-fragment and identical conditions with an annealing temperature of 55°C for the 3'-fragment. If necessary, PCR

products were cut out from an agarose gel and cleaned using the Wizard® SV Gel and PCR Clean-Up System (Promega). PCR product (20 to 50 ng) was mixed with 1.25 pmol of forward or reverse primer, 2 µl buffer (5X) and 0.2 enzyme from the BigDye® Terminator Reaction Mix (Applied Biosystems). PCR products were sequenced with forward and reverse primers, visualized on an ABI 3130xl Genetic Analyzer (Applied Biosystems) and manually checked for heterozygote positions.

Divergence patterns of *CYP2D6* sequences were compared with one mtDNA (NADH-5) and three nuclear protein coding genes; the recombination activating gene 2 (RAG2), guanine nucleotide binding protein alpha z subunit (GNAZ) and ATPase, copper-transporting, alpha polypeptide (ATP7A) (Table A1; Johnson et al. 2006). These genes were chosen based on availability of a relatively long exonic sequence for most studied species.

II.4 Data Analysis

Pairwise p-distances were calculated in MEGA 4.1 software (Tamura et al. 2007) and Maximum Likelihood (ML) distances in PAUP 4.0 (Swofford 2002) using a substitution models chosen from the program MODELTEST 3.7 (Posada and Crandall 1998). Sliding windows analysis was carried out using DnaSP 5.10 (Librado and Rozas 2009) with 100 bp windows and 25 bp steps. A consensus secondary structure was predicted using the online tool GOR IV (Garnier, Gibrat, and Robson 1996). The numbers of synonymous and non-synonymous substitutions as well as dN/dS ratios were calculated in 99 bp-step windows using the method of Nei and Gojobori (1986). Pairwise dN/dS ratios of nuclear genes were calculated in DnaSP and plotted against pairwise NADH-5 distances.

Codon-specific dN/dS ratios were calculated using a ML approach and multiple models (M0, M3, M7, M8) in the program CODEML of the PAML software package (Yang 2007). Positive selection was tested for using likelihood ratio tests of two site-model pairs: M0 vs. M3 and M7 vs. M8 (Yang, Swanson, and Vacquier 2000).

Changes in AA properties were evaluated using the software TreeSAAP (Woolley et al. 2003). Of 31 AA properties tested, those putatively indicating positive selection were selected and a sliding window analysis was conducted to detect regions under positive selection. To reduce false positives, magnitudes of property changes were divided into eight categories and only highly significant ($p < 0.001$) changes within the three highest categories (i.e. 6-8) were considered (McClellan et al. 2005).

III. RESULTS

A total of 1506 base pairs were aligned in 67 sequences (Figure S1¹). There were 676 segregating sites, 487 of which were parsimony informative resulting from 931 substitutions and 15 insertion/deletion polymorphisms. All primates carried a 9 bp deletion from position 10 to 18 resulting in three AAs less than other species. Two Siberian tigers (PTI-72 & PTI-73) showed a 3 bp insertion (TAC) at positions 331-333 and therefore an additional tyrosine in the protein. The 3 bp deletion in the human allele *CYP2D6*9* occurred at position 853-855.

The AA alignment revealed 257 of the 501 sites to be variable, of which 187 were parsimony informative. The P450 signature motif was located at sites 440-449. Of the four AA residues identified to be involved in substrate binding (Rowland et al. 2006) only one was conserved (Figure S2¹). Little saturation was seen at first and third codon sites, whereas no saturation was observed with respect to transversions and substitutions at second codon positions.

III.1 Sliding Window Analysis

The *CYPD6* gene revealed the highest variability between positions 526 and 875 and a peak around position 650 (Figure 9), whereby felid sequences displayed a low anomaly at this position. The most conserved region was between positions 1251 and 1450 encompassing the cysteine-pocket, whereby felids showed additional variation in this region. Primate sequences displayed an additional signal of variation around position 352. Relative to interspecific comparisons, human alleles showed little variation.

¹ Supplementary material available at
<http://www.kfunigraz.ac.at/zoowww/onlydownloading/Schenekarmastherthesissupplementary.pdf>

Figure 9: Results of sliding window analysis for various data subsets (window size: 100 bp; step size 25 bp): Black bars on top of both panels indicate substrate recognition sites described by Gotoh (1992). The membrane anchor (striped) and the proline-rich hinge region (checkered) described by Yamazaki et al. (1993) as well as the cysteine pocket (Modi et al. 1996) are also shown. Asterisks indicate amino acids involved in substrate binding (Rowland et al. 2006).

III.2 dN/dS Ratios

CYP2D6 sequences displayed the highest mean dN/dS ratio (0.565 ± 0.484) among the four nuclear genes analyzed (i.e. *RAG2*, 0.127 ± 0.085 ; *GNAZ*, 0.008 ± 0.112 ; *ATP7A*, 0.317 ± 0.298). Intraspecific *CYP2D6* dN/dS ratios were significantly higher (0.730 ± 0.663) than interspecific ratios (0.440 ± 0.253) (Mann-Whitney U; $p < 0.0001$). The distribution of dN/dS ratios showed a marked peak within Felidae, followed by a sharp reduction and gradual negative correlation with increasing evolutionary distance (Figure 10). The highest interspecific dN/dS ratios were all within Felidae whereby ratios > 1 involved tigers and lions, as well as puma and leopard cats, reflecting evolutionary distances of 3.72 Ma and 6.7 Ma, respectively (Johnson et al. 2006). There were no significant differences in the number of segregating sites with respect to the three specified secondary structures of the gene.

Figure 10: Bi-variate plot of pairwise dN/dS ratios of *CYP2D6* (o), *RAG2* (△), *GNAZ* (+) and *ATP7A* (x) against pairwise p-distances of the mtDNA NADH-5 gene, used as a surrogate for evolutionary distance.

Likelihood ratio tests indicated both varying dN/dS ratios among sites and positive selection in felids (M0 vs. M3: $p < 0.001$; M7 vs. M8: $p < 0.001$), as well as for the large dataset (M0 vs. M3: $p < 0.001$; M7 vs. M8: $p < 0.005$) (Figure 11a). Substrate recognition sites, as well as the trans-membrane anchor included multiple codons with high dN/dS ratios, whereas the conserved cysteine-pocket as well as proline-rich hinge region revealed only one ratio > 1 . PAML detected a signal of positive selection for 16 codon sites within Felidae, and 15 sites overall (Table 9). For the large subset, three of 31 analyzed AA properties revealed a signal of positive selection ($p < 0.001$): alpha-helical tendencies, equilibrium constant and power to be at the C-terminal. Codon sites with high dN/dS ratios showed relatively high numbers of radical AA property changes (Table 9, Figure 11b) compared to the mean across all 501 sites (0.46 ± 0.94).

Figure 11: Results of PAML and TreeSAAP analyses: (a) Mean dN/dS ratio for each codon site. Dotted line indicates a ratio of 1. (b) Sliding window analysis of the three AA properties detected to be under positive selection. The dotted line represents the cut-off value for significant ($p < 0.001$) positive selection pressure. Black bars on top of both panels indicate substrate recognition sites described by Gotoh (1992). The membrane anchor (striped) and the proline-rich hinge region (checkered) described by Yamazaki et al. (1993) as well as the cysteine pocket (Modi et al. 1996) are also shown. Asterisks indicate AAs involved in substrate binding.

Table 9: Results of PAML and TreeSAAP analyses. Suggested codons under positive selection (codon), mean dN/dS ratio (dN/dS), posterior probability, and number of radical changes in selected AA properties (# rad. changes).

PAML						TreeSAAP	
felid subset			large subset			large subset	
codon	dN/dS	post. prob.	codon	dN/dS	post. prob.	codon	# rad. changes
8	6.845	0.992**	8	1.561	0.945	8	2
165	3.987	0.545	58	1.177	0.530	58	5
191	6.847	0.992**	59	1.206	0.552	59	3
212	6.517	0.936	62	1.364	0.721	62	2
224	5.414	0.771	96	1.226	0.575	96	0
236	5.263	0.749	110	1.236	0.587	110	1
241	6.514	0.936	213	1.357	0.708	213	2
256	3.790	0.517	216	1.552	0.933	216	5
273	6.525	0.937	241	1.430	0.786	241	2
312	6.167	0.881	247	1.452	0.811	247	2
346	5.525	0.787	308	1.407	0.761	308	3
438	5.051	0.718	361	1.313	0.663	361	4
446	6.255	0.896	446	1.349	0.700	446	4
460	4.678	0.664	486	1.297	0.646	486	3
493	4.880	0.692	499	1.499	0.866	499	2
499	6.704	0.969*					

*: post. prob. > 0.95 ; **: post. prob. > 0.99

IV. DISCUSSION

The mean dN/dS ratio across *CYP2D6* (0.565) is substantially higher than averages reported in multiple studies of mammalian genes. Wang et al. (2007) reported a mean dN/dS ratio between human and mouse of 0.138 (across 12,052 genes) and between human and old world monkeys of 0.177 (across 14,204 genes). The non-monotonic distribution of dN/dS values with respect to evolutionary distance (Figure 10) appears indicative of the initial retention and eventual loss of slightly deleterious mutations (Rocha et al. 2006). Such a pattern across evolutionary time frames of several million years does not correspond with the explanation of this phenomena occurring within species or across closely related species (Yang, Swanson, and Vacquier 2000; Rocha et al. 2006). An alternate explanation may relate to the fact that slightly deleterious mutations are not removed as effectively in small populations (Hughes and Friedman 2009). Three of the studied felids are listed as vulnerable or endangered by the IUCN and felids in general have relatively low effective population sizes. But our comparisons are generally interspecific so it is not clear how large an effect small effective population sizes should have on our data. Interestingly, there was no statistical support for radical physiochemical property changes within Felidae despite the support for positive selection across a number of specific codons.

Several additional characteristics suggest adaptive significance. Some regions containing predicted substrate recognition or membrane anchor sites indeed displayed dN/dS ratios > 1 (Figure 11b) as well as radical changes in AA physiochemical properties. Two of the polymorphic AA substrate binding sites revealed a correlation with diet. Position 124 revealed a common phenylalanine across all carnivores and primates, whereas herbivores had either a valine or isoleucine; position 220 revealed a common aspartic acid across all carnivores whereas herbivores and omnivores had either glutamine or glutamine acid. While it has been hypothesized that P450 evolution is promoted by the presence of plant toxins in the diet, there has been little concrete evidence of a dietary link with major feeding groups (Fink-Gremmels 2008).

Surprisingly, sequence variation of human alleles was very low despite the gene's high expression variability. While in our data only putatively functional alleles were used, several of them are described to result in decreased enzyme function. Nonetheless, variable copy number clearly plays a significant role in the high expression variability in humans. The prevalence of variable copy number or enzyme dysfunction of *CYP2D6* in non-humans is not known.

We conclude that the *CYP2D6* gene in non-human mammals exhibits considerable variation that most likely reflects at least some adaptive significance and should be considered a candidate study target in understanding dietary adaptation. Whether based on copy number variation or

sequence variation, a more thorough consideration of *CYP2D6* variation in domestic or veterinary animals may uncover breed-specific differences that can support more effective drug therapy.

REFERENCES

- Bailey, D. G., G. K. Dresser, and J. R. Bend. 2003 Bergamottin, lime juice, and red wine as inhibitors of cytochrome P450 3A4 activity: comparison with grapefruit juice. *Clin Pharmacol Ther* **73**, 529-537.
- Bertz, R. J., and G. R. Granneman. 1997 Use of in vitro and in vivo data to estimate the likelihood of metabolic pharmacokinetic interactions. *Clin Pharmacokinet* **32**, 210-258.
- Bull, S., S. Nijmeier, A. Werners, and J. Fink-Gremmels. 2003 Cytochrome P450 2D6 polymorphism in horses and ponies. *J. vet. Pharmacol. Therap* **26**, 85-86.
- Chauret, N., A. Gauthier, J. Martin, and D. A. Nicoll-Griffith. 1997 In vitro comparison of cytochrome P450-mediated metabolic activities in human, dog, cat, and horse. *Drug Metab Dispos* **25**, 1130-1136.
- Conney, A. H., E. C. Miller, and J. A. Miller. 1956 The metabolism of methylated aminoazo dyes. V. Evidence for induction of enzyme synthesis in the rat by 3-methylcholanthrene. *Cancer Res* **16**, 450-459.
- Cooper, D. Y., S. Levin, S. Narasimhulu, and O. Rosenthal. 1965 Photochemical Action Spectrum of the Terminal Oxidase of Mixed Function Oxidase Systems. *Science* **147**, 400-402.
- Court, M. H., and D. J. Greenblatt. 1997 Molecular basis for deficient acetaminophen glucuronidation in cats. An interspecies comparison of enzyme kinetics in liver microsomes. *Biochem Pharmacol* **53**, 1041-1047.
- Court, M. H., B. L. Hay-Kraus, D. W. Hill, A. J. Kind, and D. J. Greenblatt. 1999 Propofol hydroxylation by dog liver microsomes: assay development and dog breed differences. *Drug Metab Dispos* **27**, 1293-1299.
- Distlerath, L. M., P. E. Reilly, M. V. Martin, G. G. Davis, G. R. Wilkinson, and F. P. Guengerich. 1985 Purification and characterization of the human liver cytochromes P-450 involved in debrisoquine 4-hydroxylation and phenacetin O-deethylation, two prototypes for genetic polymorphism in oxidative drug metabolism. *J Biol Chem* **260**, 9057-9067.
- Ernest, C. S., 2nd, S. D. Hall, and D. R. Jones. 2005 Mechanism-based inactivation of CYP3A by HIV protease inhibitors. *J Pharmacol Exp Ther* **312**, 583-591.
- Fink-Gremmels, J. 2008 Implications of hepatic cytochrome P450-related biotransformation processes in veterinary sciences. *Eur J Pharmacol* **585**, 502-509.
- Garfinkel, D. 1958 Studies on pig liver microsomes. I. Enzymic and pigment composition of different microsomal fractions. *Arch Biochem Biophys* **77**, 493-509.
- Garnier, J., J.-F. Gibrat, and B. Robson. 1996 GOR secondary structure prediction method version IV. *Methods Enzymol* **266**, 14. (doi:10.1016/S0076-6879(96)66034-0)
- Gilep, A. A., I. V. Huryeva, O. L. Guryev, S. A. Usanov, and R. W. Estabrook. 2000. Cloning and heterologous expression of the cDNA encoding the cytochrome P450 steroid 17alpha-hydroxylase/17,20 lyase from cat.
- Gonzalez, F. J., and D. W. Nebert. 1990 Evolution of the P450 gene superfamily: animal-plant 'warfare', molecular drive and human genetic differences in drug oxidation. *Trends Genet* **6**, 182-186.

- Gonzalez, F. J., F. Vilbois, J. P. Hardwick, O. W. McBride, D. W. Nebert, H. V. Gelboin, and U. A. Meyer. 1988 Human debrisoquine 4-hydroxylase (P450IID1): cDNA and deduced amino acid sequence and assignment of the CYP2D locus to chromosome 22. *Genomics* **2**, 174-179.
- Gotoh, O. 1992 Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. *J Biol Chem* **267**, 83-90.
- Gotoh, O., Y. Tagashira, T. Iizuka, and Y. Fujii-Kuriyama. 1983 Structural characteristics of cytochrome P-450. Possible location of the heme-binding cysteine in determined amino-acid sequences. *J Biochem* **93**, 807-817.
- Guengerich, F. P. 2008 Cytochrome p450 and chemical toxicology. *Chem Res Toxicol* **21**, 70-83.
- Guengerich, F. P. 1992 Characterization of human cytochrome P450 enzymes. *Faseb J* **6**, 745-748.
- Hasemann, C. A., R. G. Kurumbail, S. S. Boddupalli, J. A. Peterson, and J. Deisenhofer. 1995 Structure and function of cytochrome P450s: a comparative analysis of three crystal structures. *Structure* **3**, 41-62.
- Hasler, J. A. 1999 Pharmacogenetics of cytochrome P450s. *Mol Aspects Med* **20**, 12-24, 25-137.
- Hennig, W. 1998. *Genetik*. Springer- Verlag, Berlin Heidelberg.
- Hirvonen, A., K. Husgafvel-Pursiainen, S. Anttila, A. Karjalainen, and H. Vainio. 1993 The human CYP2E1 gene and lung cancer: DraI and RsaI restriction fragment length polymorphisms in a Finnish study population. *Carcinogenesis* **14**, 85-88.
- Hughes, A. L., and R. Friedman. 2009 More radical AA replacements in primates than in rodents: support for the evolutionary role of effective population size. *Gene* **440**, 50-56. (doi:10.1016/j.gene.2009.03.012)
- Ingelman-Sundberg, M. 1999 Duplication, multiduplication, and amplification of genes encoding drug-metabolizing enzymes: evolutionary, toxicological, and clinical pharmacological aspects. *Drug Metab Rev* **31**, 449-459.
- Ingelman-Sundberg, M. 2005 Genetic polymorphisms of cytochrome P450 2D6 (CYP2D6): clinical consequences, evolutionary aspects and functional diversity. *Pharmacogenomics J* **5**, 6-13.
- Ingelman-Sundberg, M. 2004 Human drug metabolising cytochrome P450 enzymes: properties and polymorphisms. *Naunyn Schmiedebergs Arch Pharmacol* **369**, 89-104.
- Ioannides, C. 2006 Cytochrome p450 expression in the liver of food-producing animals. *Curr Drug Metab* **7**, 335-348.
- Johansson, I., E. Lundqvist, L. Bertilsson, M. L. Dahl, F. Sjoqvist, and M. Ingelman-Sundberg. 1993 Inherited amplification of an active gene in the cytochrome P450 CYP2D locus as a cause of ultrarapid metabolism of debrisoquine. *Proc Natl Acad Sci U S A* **90**, 11825-11829.
- Johnson, W. E., E. Eizirik, J. Pecon-Slatery, W. J. Murphy, A. Antunes, E. Teeling, and S. J. O'Brien. 2006 The late Miocene radiation of modern Felidae: a genetic assessment. *Science* **311**, 73-77. (doi:10.1126/science.1122277)
- Kimmel, M., D. E. Axelrod, and G. M. Wahl. 1992 A branching process model of gene amplification following chromosome breakage. *Mutat Res* **276**, 225-239.

- Kimura, S., M. Umeno, R. C. Skoda, U. A. Meyer, and F. J. Gonzalez. 1989 The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and identification of the polymorphic CYP2D6 gene, a related gene, and a pseudogene. *Am J Hum Genet* **45**, 889-904.
- Klingenberg, M. 1958 Pigments of rat liver microsomes. *Arch Biochem Biophys* **75**, 376-386.
- Lamb, D. C., L. Lei, A. G. Warrilow, G. I. Lepesheva, J. G. Mullins, M. R. Waterman, and S. L. Kelly. 2009 The first virally encoded cytochrome p450. *J Virol* **83**, 8266-8269.
- Lee, S. J., S. S. Lee, H. J. Jung, H. S. Kim, S. J. Park, C. W. Yeo, and J. G. Shin. 2009 Discovery of novel functional variants and extensive evaluation of CYP2D6 genetic polymorphisms in Koreans. *Drug Metab Dispos* **37**, 1464-1470.
- Lewin, B. 2000. *Genes VII*. Oxford University Press, New York.
- Lewis, D. F., E. Watson, and B. G. Lake. 1998 Evolution of the cytochrome P450 superfamily: sequence alignments and pharmacogenetics. *Mutat Res* **410**, 245-270.
- Librado, P., and J. Rozas. 2009 DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451-1452. (doi:10.1093/bioinformatics/btp187)
- Lundqvist, E., I. Johansson, and M. Ingelman-Sundberg. 1999 Genetic mechanisms for duplication and multiduplication of the human CYP2D6 gene and methods for detection of duplicated CYP2D6 genes. *Gene* **226**, 327-338.
- Mansuy, D. 1998 The great diversity of reactions catalyzed by cytochrome P450s. *Comp Biochem Physiol C Pharmacol Toxicol Endocrinol* **121**, 5-14.
- Matsunaga, E., M. Umeno, and F. J. Gonzalez. 1990 The rat P450 IID subfamily: complete sequences of four closely linked genes and evidence that gene conversions maintained sequence homogeneity at the heme-binding region of the cytochrome P450 active site. *J Mol Evol* **30**, 155-169.
- McClellan, D. A., E. J. Palfreyman, M. J. Smith, J. L. Moss, R. G. Christensen, and J. K. Sailsbery. 2005 Physicochemical Evolution and Molecular Adaptation of the Cetacean and Artiodactyl Cytochrome b Proteins. *Mol Biol Evol* **22**, 19. (doi:10.1093/molbev/msi028)
- Mizutani, T. 2003 PM frequencies of major CYPs in Asians and Caucasians. *Drug Metab Rev* **35**, 99-106.
- Modi, S., M. J. Paine, M. J. Sutcliffe, L. Y. Lian, W. U. Primrose, C. R. Wolf, and G. C. Roberts. 1996 A model for human cytochrome P450 2D6 based on homology modeling and NMR studies of substrate binding. *Biochemistry* **35**, 4540-4550. (doi:10.1021/bi952742o)
- Nebert, D. W., and D. R. Nelson. 1991 P450 gene nomenclature based on evolution. *Methods Enzymol* **206**, 3-11.
- Nei, M., and T. Gojobori. 1986 Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* **3**, 418-426.
- Nelson, D. R. 2006 Cytochrome P450 nomenclature, 2004. *Methods Mol Biol* **320**, 1-10.
- Nelson, D. R., T. Kamataki, D. J. Waxman, F. P. Guengerich, R. W. Estabrook, R. Feyereisen, F. J. Gonzalez, M. J. Coon, I. C. Gunsalus, O. Gotoh, and et al. 1993 The P450 superfamily: update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature. *DNA Cell Biol* **12**, 1-51.

- Nelson, D. R., L. Koymans, T. Kamataki, J. J. Stegeman, R. Feyereisen, D. J. Waxman, M. R. Waterman, O. Gotoh, M. J. Coon, R. W. Estabrook, I. C. Gunsalus, and D. W. Nebert. 1996 P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics* **6**, 1-42.
- Omura, T., and R. Sato. 1962 A new cytochrome in liver microsomes. *J Biol Chem* **237**, 1375-1376.
- Omura, T., and R. Sato. 1964a The Carbon Monoxide-Binding Pigment of Liver Microsomes. I. Evidence for Its Hemoprotein Nature. *J Biol Chem* **239**, 2370-2378.
- Omura, T., and R. Sato. 1964b The Carbon Monoxide-Binding Pigment of Liver Microsomes. II. Solubilization, Purification, and Properties. *J Biol Chem* **239**, 2379-2385.
- Oscarson, M., M. Hidestrand, I. Johansson, and M. Ingelman-Sundberg. 1997 A combination of mutations in the CYP2D6*17 (CYP2D6Z) allele causes alterations in enzyme function. *Mol Pharmacol* **52**, 1034-1040.
- Paulson, S. K., L. Engel, B. Reitz, S. Bolten, E. G. Burton, T. J. Maziasz, B. Yan, and G. L. Schoenhard. 1999 Evidence for polymorphism in the canine metabolism of the cyclooxygenase 2 inhibitor, celecoxib. *Drug Metab Dispos* **27**, 1133-1142.
- Phipps, A. N., C. Collins, and S. E. Blanchflower. 2003. In vitro characterization of feline P450 enzymes. 9th International COngress of the European Association of Veterinary Pharmacology and Toxicology, Lisbon.
- Pontius, J. U., and S. J. O'Brien. 2007 Genome Annotation Resource Fields--GARFIELD: a genome browser for *Felis catus*. *J Hered* **98**, 386-389.
- Porter, T. D. 2004 Jud Coon: 35 years of p450 research, a synopsis of p450 history. *Drug Metab Dispos* **32**, 1-6.
- Posada, D., and K. A. Crandall. 1998 MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817-818. (doi:10.1093/bioinformatics/14.9.817)
- Poulsen, L., K. Brosen, L. Arendt-Nielsen, L. F. Gram, K. Elbaek, and S. H. Sindrup. 1996 Codeine and morphine in extensive and poor metabolizers of sparteine: pharmacokinetics, analgesic effect and side effects. *Eur J Clin Pharmacol* **51**, 289-295.
- Rao, Y., E. Hoffmann, M. Zia, L. Bodin, M. Zeman, E. M. Sellers, and R. F. Tyndale. 2000 Duplications and defects in the CYP2A6 gene: identification, genotyping, and in vivo effects on smoking. *Mol Pharmacol* **58**, 747-755.
- Rocha, E. P., J. M. Smith, L. D. Hurst, M. T. Holden, J. E. Cooper, N. H. Smith, and E. J. Feil. 2006 Comparisons of dN/dS are time dependent for closely related bacterial genomes. *J Theor Biol* **239**, 226-235. (doi:10.1016/j.jtbi.2005.08.037)
- Roussel, F., D. B. Duignan, M. P. Lawton, R. S. Obach, C. A. Strick, and D. J. Tweedie. 1998 Expression and characterization of canine cytochrome P450 2D15. *Arch Biochem Biophys* **357**, 27-36.
- Rowland, P., F. E. Blaney, M. G. Smyth, J. J. Jones, V. R. Leydon, A. K. Oxbrow, C. J. Lewis, M. G. Tennant, S. Modi, D. S. Eggleston, R. J. Chenery, and A. M. Bridges. 2006 Crystal structure of human cytochrome P450 2D6. *J Biol Chem* **281**, 7614-7622.
- Sachse, C., J. Brockmoller, S. Bauer, and I. Roots. 1997 Cytochrome P450 2D6 variants in a Caucasian population: allele frequencies and phenotypic consequences. *Am J Hum Genet* **60**, 284-295.

- Sato, R., and T. Omura. 1961. A carbon monoxide-binding pigment of liver microsomes. Pp. 529. Proc 5th Intern Congr Biochem, Vienna.
- Seliskar, M., and D. Rozman. 2007 Mammalian cytochrome P450s--importance of tissue specificity. *Biochim Biophys Acta* **1770**, 458-466.
- Shah, S. S., S. Sanda, N. L. Regmi, K. Sasaki, and M. Shimoda. 2007 Characterization of cytochrome P450-mediated drug metabolism in cats. *J Vet Pharmacol Ther* **30**, 422-428.
- Shankar, R., B. Kataria, and M. Mukerji. 2009 Finding Alu in primate genomes with AF-1. *Bioinformatics* **3**, 287-288.
- Skoda, R. C., F. J. Gonzalez, A. Demierre, and U. A. Meyer. 1988 Two mutant alleles of the human cytochrome P-450db1 gene (P450C2D1) associated with genetically deficient metabolism of debrisoquine and other drugs. *Proc Natl Acad Sci U S A* **85**, 5240-5243.
- Smith, D. A., M. J. Achland, and B. C. Jones. 1997 Properties of cytochrome P450 isozymes and their substrates. I. Active site characteristics. *Drug Discovery Today* **2**, 406-414.
- Swofford, D. L. 2002. *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4. Sinauer Associates, Sunderland, Massachusetts.
- Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007 MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**, 1596-1599. (doi:10.1093/molbev/msm092)
- Tanaka, N., T. Miyasho, R. Shinkyo, T. Sakaki, and H. Yokota. 2006 cDNA cloning and characterization of feline CYP1A1 and CYP1A2. *Life Sci* **79**, 2463-2473.
- Tanaka, N., R. Shinkyo, T. Sakaki, M. Kasamatsu, S. Imaoka, Y. Funae, and H. Yokota. 2005 Cytochrome P450 2E polymorphism in feline liver. *Biochim Biophys Acta* **1726**, 194-205.
- Trepanier, L. A. 2006 Cytochrome P450 and its role in veterinary drug interactions. *Vet Clin North Am Small Anim Pract* **36**, 975-985, v.
- van der Weide, J., and L. S. Steijns. 1999 Cytochrome P450 enzyme system: genetic polymorphisms and impact on clinical pharmacology. *Ann Clin Biochem* **36 (Pt 6)**, 722-729.
- Wang, H. Y., H. C. Chien, N. Osada, K. Hashimoto, S. Sugano, T. Gojobori, C. K. Chou, S. F. Tsai, C. I. Wu, and C. K. Shen. 2007 Rate of evolution in brain-expressed genes in humans and other primates. *PLoS Biol* **5**, e13. (doi:10.1371/journal.pbio.0050013)
- Weide, v. d., J., and L. S. Steijns. 1999 Cytochrome P450 enzyme system: genetic polymorphisms and impact on clinical pharmacology. *Ann Clin Biochem* **36 (Pt 6)**, 722-729.
- Werck-Reichhart, D., and R. Feyereisen. 2000 Cytochrome P450s: a success story. *Genome Biol* **1**, REVIEWS3003.
- Williams, P. A., J. Cosme, V. Sridhar, E. F. Johnson, and D. E. McRee. 2000 Mammalian microsomal cytochrome P450 monooxygenase: structural adaptations for membrane binding and functional diversity. *Mol Cell* **5**, 121-131.
- Woolley, S., J. Johnson, M. J. Smith, K. A. Crandall, and D. A. McClellan. 2003 TreeSAAP: selection on AA properties using phylogenetic trees. *Bioinformatics* **19**, 671-672. (doi:10.1093/bioinformatics/btg043)

- Yamazaki, S., K. Sato, K. Suhara, M. Sakaguchi, K. Mihara, and T. Omura. 1993 Importance of the proline-rich region following signal-anchor sequence in the formation of correct conformation of microsomal cytochrome P-450s. *J Biochem* **114**, 652-657.
- Yang, Z. 2007 PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* **24**, 1586-1591. (doi:10.1093/molbev/msm088)
- Yang, Z., W. J. Swanson, and V. D. Vacquier. 2000 Maximum-likelihood analysis of molecular adaptation in abalone sperm lysin reveals variable selective pressures among lineages and sites. *Mol Biol Evol* **17**, 1446-1455.
- Zanger, U. M., S. Raimundo, and M. Eichelbaum. 2004 Cytochrome P450 2D6: overview and update on pharmacology, genetics, biochemistry. *Naunyn Schmiedebergs Arch Pharmacol* 369, 23-37.
- Zhou, S. F. 2009 Polymorphism of human cytochrome P450 2D6 and its clinical significance: Part I. *Clin Pharmacokinet* 48, 689-723.

FIGURE LIST

- Figure 1:** (a) Typical primary structure of a membrane-bound P450 protein (b) Topology diagram of the secondary structures and their arrangements of CYP102; blue boxes indicate α -helices; groups of white arrows outlined with dotted lines form β -sheets; element sizes are not in proportion to their length in the primary sequence (c) folded P450 (CYP2C5) in association with the ER membrane (purple); the heme is colored in orange, the substrate in yellow (from Werck-Reichhart and Feyereisen 2000; p. 3-4).....p. 6
- Figure 2:** Abbreviated version of the P450 phylogeny mapped against important developments in the evolution of life. The tree was constructed by combining several P450 phylogenetic trees (from Lewis, Watson, and Lake 1998; p. 254).....p. 9
- Figure 3:** Ribbon diagram of the human CYP2D6 protein and its prosthetic group. Red coiled structures represent alpha-helices, green arrows indicate beta-sheets (from Rowland et al. 2006; p. 7618).....p. 14
- Figure 4:** Global distribution of the frequency of ultrarapid metabolizers. Arrow indicates the muslim migration which may explain the high percentage of UMs in Spain (from Ingelman-Sundberg 1999; p. 452).....p. 18
- Figure 5:** Structural organization of the human *CYP2D* locus on chromosome 22. Numbered black bars indicate the genes' exons.P. 19
- Figure 6:** Unequal exchange model as the cause of the duplication and multiduplication of *CYP2D6*. (I) Formation of a duplicated *CYP2D6*-gene via unequal crossover. (II) Duplicated region forms a hairpin loop and is looped out during another unequal crossover event. (III) Formation of additional gene copies with the same mechanisms. Black bars indicate Alu sequences (from Lundqvist, Johansson, and Ingelman-Sundberg 1999; p. 335).....p. 23
- Figure 7:** Model of the creation of alleles with 13 copies of *CYP2D6*. First, an acentric DNA molecule is formed by a double DNA breakage. This molecule is replicated multiple times. Multiple copies are unequally segregated during meiosis. At the end, copies are reintroduced into the *CYP2D* locus (from Lundqvist, Johansson, and Ingelman-Sundberg 1999; p. 336).....p. 24
- Figure 8:** Intron & exon organization the nine fully characterized mammalian *CYP2D6*-orthologs..p. 25
- Figure 9.** Results of sliding window analysis for various data subsets (window size: 100 bp; step size 25 bp): Black bars on top of both panels indicate substrate recognition sites described by Gotoh (1992). The membrane anchor (striped) and the proline-rich hinge region (checkered) described by Yamazaki et al. (1993) as well as the cysteine pocket (Modi et al. 1996) are also shown. Asterisks indicate amino acids involved in substrate binding (Rowland et al. 2006).....p. 33
- Figure 10:** Bi-variate plot of pairwise dN/dS ratios of *CYP2D6* (o), RAG2 (Δ), GNAZ (+) and ATP7A (x) against pairwise p-distances of the mtDNA NADH-5 gene, used as a surrogate for evolutionary distance.p. 34
- Figure 11:** Results of PAML and TreeSAAP analyses: (a) Mean dN/dS ratio for each codon site. Dotted line indicates a ratio of 1. (b) Sliding window analysis of the three AA properties detected to be under positive selection. The dotted line represents the cut-off value for significant ($p < 0.001$) positive selection pressure. Black bars on top of both panels indicate substrate recognition sites described by Gotoh (1992). The membrane anchor (striped) and the proline-rich hinge region (checkered) described by Yamazaki et al. (1993) as well as the cysteine pocket (Modi et al. 1996) are also shown. Asterisks indicate AAs involved in substrate binding.....p. 35

APPENDIX

Table A1: GenBank accession numbers of used sequence data

Species	Sample code ¹	Gene	Accession No.
<i>Acinonyx jubatus jubatus</i> (Cheetah)	AJU_173	CYP2D6	### ²
<i>Acinonyx jubatus</i> (Cheetah)	AJU_287	CYP2D6	###
<i>Acinonyx jubatus</i> (Cheetah)	AJU_430	CYP2D6	###
<i>Acinonyx jubatus</i> (Cheetah)	AJU_580	CYP2D6	###
<i>Felis catus</i> (Domestic cat)	Fc_CYP2D6_mRNA_complete	CYP2D6	###
<i>Felis catus</i> (Domestic cat)	FCA_6840	CYP2D6	###
<i>Felis catus</i> (Domestic cat)	FCA_6974	CYP2D6	###
<i>Felis catus</i> (Domestic cat)	FCA_7318	CYP2D6	###
<i>Lynx rufus</i> (Bobcat)	LRU_50	CYP2D6	###
<i>Lynx rufus</i> (Bobcat)	LRU_61	CYP2D6	###
<i>Lynx rufus</i> (Bobcat)	LRU_64	CYP2D6	###
<i>Prionailurus bengalensis</i> (Leopard cat)	PBE_15	CYP2D6	###
<i>Prionailurus bengalensis euptilura</i> (Leopard cat)	PBE_31	CYP2D6	###
<i>Prionailurus bengalensis euptilura</i> (Leopard cat)	PBE_55	CYP2D6	###
<i>Puma concolor</i> (Cougar)	PCO_1249	CYP2D6	###
<i>Puma concolor coryi</i> (Cougar)	PCO_1302	CYP2D6	###
<i>Puma concolor</i> (Cougar)	PCO_1362	CYP2D6	###
<i>Panthera leo</i> (Lion)	PLE_1774	CYP2D6	###
<i>Panthera leo</i> (Lion)	PLE_1780	CYP2D6	###
<i>Panthera leo</i> (Lion)	PLE_1796	CYP2D6	###
<i>Panthera tigris altaica</i> (Siberian Tiger)	PTI_66	CYP2D6	###
<i>Panthera tigris altaica</i> (Siberian Tiger)	PTI_72	CYP2D6	###
<i>Panthera tigris altaica</i> (Siberian Tiger)	PTI_73	CYP2D6	###
<i>Acinonyx jubatus</i> (Cheetah)	---	ATP7A	DQ082604
<i>Bos taurus</i> (Cattle)	---	ATP7A	XM_615430.4
<i>Callithrix jacchus</i> (White-tufted-ear marmoset)	---	ATP7A	n.a. ³
<i>Canis lupus familiaris</i> (Dog)	---	ATP7A	AY011436.1
<i>Equus caballus</i> (Horse)	---	ATP7A	AY011430.1
<i>Felis catus</i> (Domestic cat)	---	ATP7A	DQ082590
<i>Homo sapiens</i> (Human)	---	ATP7A	NM_000052.4
<i>Lynx rufus</i> (Bobcat)	---	ATP7A	DQ082608
<i>Macaca fascicularis</i> (Crab-eating macaque)	---	ATP7A	n.a.
<i>Macaca mulatta</i> (Rhesus monkey)	---	ATP7A	AY011415.1
<i>Mus musculus</i> (House mouse)	---	ATP7A	AY011397.1
<i>Pan paniscus</i> (Pygmy chimpanzee)	---	ATP7A	n.a.
<i>Pan troglodytes</i> (Common chimpanzee)	---	ATP7A	AF535046.1
<i>Panthera leo</i> (Lion)	---	ATP7A	DQ082622
<i>Panthera tigris</i> (Tiger)	---	ATP7A	DQ082624
<i>Prionailurus bengalensis</i> (Leopard cat)	---	ATP7A	DQ082599
<i>Puma concolor</i> (Cougar)	---	ATP7A	DQ082602
<i>Rattus norvegicus</i> (Norway rat)	---	ATP7A	AY011398.1

Table A1 (cont.)

<i>Bos taurus</i> (Catte)	---	CYP2D6	NM_174529.2
<i>Callithrix jacchus</i> (White-tufted-ear marmoset)	---	CYP2D19	D29822.1
<i>Canis lupus familiaris</i> (Dog)	---	CYP2D15	NM_001003333.1
<i>Equus caballus</i> (Horse)	---	CYP2D50	NM_001111306.1
<i>Felis catus</i> (Domestic cat)	---	CYP2D6	chrB4:154039538..154047755 ⁴
<i>Homo sapiens</i> (Human)	---	CYP2D6	NM_000106.4
<i>Homo sapiens</i> (Human)	---	CYP2D6	AY545216.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	EU530602.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	DQ282145.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	DQ282144.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	DQ282146.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	EU530603.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	DQ282156.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	DQ282147.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	DQ282151.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	DQ282152.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	DQ282157.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	DQ282158.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	DQ282153.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	DQ282159.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	DQ282154.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	DQ282155.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	M33388.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	BC067432.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	BC075023.2
<i>Homo sapiens</i> (Human)	---	CYP2D6	BC075024.2
<i>Homo sapiens</i> (Human)	---	CYP2D6	BC106758.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	DD182331.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	DD182332.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	DD182333.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	DD182334.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	BD318197.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	BD318198.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	BD318199.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	BD318200.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	BC126858.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	NG_008376.1
<i>Homo sapiens</i> (Human)	---	CYP2D6	GQ374179.1
<i>Macaca fascicularis</i> (Crab-eating macaque)	---	CYP2D17	U38218.1
<i>Macaca mulatta</i> (Rhesus monkey)	---	CYP2D6	XM_001107251.1
<i>Mus musculus</i> (House mouse)	---	CYP2D22	NM_019823.2
<i>Pan paniscus</i> (Pygmy chimpanzee)	---	CYP2D6	DQ282163.1
<i>Pan troglodytes</i> (Common chimpanzee)	---	CYP2D6	NM_001040622.1
<i>Rattus norvegicus</i> (Norway rat)	---	CYP2D4	NM_138515.1
<i>Acinonyx jubatus</i> (Cheetah)	---	GNAZ	DQ082128

Table A1 (cont.)

<i>Bos taurus</i> (Catte)	---	GNAZ	n.a.
<i>Callithrix jacchus</i> (White-tufted-ear marmoset)	---	GNAZ	n.a.
<i>Canis lupus familiaris</i> (Dog)	---	GNAZ	XM_843512.1
<i>Equus caballus</i> (Horse)	---	GNAZ	XM_001488441.2
<i>Felis catus</i> (Domestic cat)	---	GNAZ	DQ082114
<i>Homo sapiens</i> (Human)	---	GNAZ	NM_002073.2
<i>Lynx rufus</i> (Bobcat)	---	GNAZ	DQ082132
<i>Macaca fascicularis</i> (Crab-eating macaque)	---	GNAZ	n.a.
<i>Macaca mulatta</i> (Rhesus monkey)	---	GNAZ	XM_001092302.1
<i>Mus musculus</i> (House mouse)	---	GNAZ	NM_010311.3
<i>Pan paniscus</i> (Pygmy chimpanzee)	---	GNAZ	n.a.
<i>Pan troglodytes</i> (Common chimpanzee)	---	GNAZ	XM_515016.2
<i>Panthera leo</i> (Lion)	---	GNAZ	DQ082146
<i>Panthera tigris</i> (Tiger)	---	GNAZ	DQ082148
<i>Prionailurus bengalensis</i> (Leopard cat)	---	GNAZ	DQ082123
<i>Puma concolor</i> (Cougar)	---	GNAZ	DQ082126
<i>Rattus norvegicus</i> (Norway rat)	---	GNAZ	NM_013189.1
<i>Acinonyx jubatus</i> (Cheetah)	---	NADH-5	n.s. ⁵
<i>Bos taurus</i> (Catte)	---	NADH-5	NC_006853.1
<i>Callithrix jacchus</i> (White-tufted-ear marmoset)	---	NADH-5	n.a.
<i>Canis lupus familiaris</i> (Dog)	---	NADH-5	NC_002008.4
<i>Equus caballus</i> (Horse)	---	NADH-5	NC_001640.1
<i>Felis catus</i> (Domestic cat)	---	NADH-5	n.s.
<i>Homo sapiens</i> (Human)	---	NADH-5	NC_012920.1
<i>Lynx rufus</i> (Bobcat)	---	NADH-5	n.s.
<i>Macaca fascicularis</i> (Crab-eating macaque)	---	NADH-5	n.a.
<i>Macaca mulatta</i> (Rhesus monkey)	---	NADH-5	NC_005943.1
<i>Mus musculus</i> (House mouse)	---	NADH-5	NC_010339.1
<i>Pan paniscus</i> (Pygmy chimpanzee)	---	NADH-5	n.a.
<i>Pan troglodytes</i> (Common chimpanzee)	---	NADH-5	NC_001643.1
<i>Panthera leo</i> (Lion)	---	NADH-5	n.s.
<i>Panthera tigris</i> (Tiger)	---	NADH-5	n.s.
<i>Prionailurus bengalensis</i> (Leopard cat)	---	NADH-5	n.s.
<i>Puma concolor</i> (Cougar)	---	NADH-5	n.s.
<i>Rattus norvegicus</i> (Norway rat)	---	NADH-5	NC_001665.2
<i>Acinonyx jubatus</i> (Cheetah)	---	RAG2	DQ082344
<i>Bos taurus</i> (Catte)	---	RAG2	NM_001082469.1
<i>Callithrix jacchus</i> (White-tufted-ear marmoset)	---	RAG2	n.a.
<i>Canis lupus familiaris</i> (Dog)	---	RAG2	XM_540537.2
<i>Equus caballus</i> (Horse)	---	RAG2	XM_001487973.2
<i>Felis catus</i> (Domestic cat)	---	RAG2	DQ082330
<i>Homo sapiens</i> (Human)	---	RAG2	NM_000536.2
<i>Lynx rufus</i> (Bobcat)	---	RAG2	DQ082348
<i>Macaca fascicularis</i> (Crab-eating macaque)	---	RAG2	n.a.
<i>Macaca mulatta</i> (Rhesus monkey)	---	RAG2	XM_001114748.1

Table A1 (cont.)

<i>Mus musculus</i> (House mouse)	---	RAG2	NM_009020.3
<i>Pan paniscus</i> (Pygmy chimpanzee)	---	RAG2	n.a.
<i>Pan troglodytes</i> (Common chimpanzee)	---	RAG2	XM_001155150.1
<i>Panthera leo</i> (Lion)	---	RAG2	DQ082362
<i>Panthera tigris</i> (Tiger)	---	RAG2	DQ082364
<i>Prionailurus bengalensis</i> (Leopard cat)	---	RAG2	DQ082339
<i>Puma concolor</i> (Cougar)	---	RAG2	DQ082342
<i>Rattus norvegicus</i> (Norway rat)	---	RAG2	XM_230374.4

¹ Only available for sequences obtained in this study

² ###: Will be submitted

³ n.a.: Sequence data not available. Species was excluded from analysis involving this gene

⁴ Positional information of the putative *CYP2D6*-gene. Data from GARFIELD-Web-Browser (<http://lgd.abcc.ncifcrf.gov>)

⁵ n.s.: Sequence data not submitted to Genbank