

# The Human Cytochrome P450 (CYP) Allele Nomenclature website: A peer-reviewed database of CYP variants and their associated effects

Sarah C. Sim\* and Magnus Ingelman-Sundberg

Section of Pharmacogenetics, Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden

\*Correspondence to: Tel: +46 8 524 862 39; Fax: +46 8 33 73 27; E-mail: sarah.sim@ki.se

Date received (in revised form): 22nd March 2010

## Abstract

Pharmacogenetics affects both pharmacokinetics and pharmacodynamics, thereby influencing an individual's response to drugs, both in terms of response and adverse reactions. Within the area of pharmacogenetics, findings of genetic variation influencing drug levels have been more prevalent, and variation in the cytochrome P450 (CYP) enzymes is one of the most common causes. Much of the work concerning sequence variations in CYPs aims at finding biomarkers of use for individualised treatment, thereby increasing the treatment response, lowering the number of side effects and decreasing the overall cost of treatment regimens. For over ten years, the Human Cytochrome P450 Allele Nomenclature (CYP-allele) website (<http://www.cypalleles.ki.se/>) has offered a database of genetic information on CYP variants, along with effects at the molecular as well as clinical level. Thus, this database serves as an assembly of past, current and soon-to-be published information on CYP alleles and their outcome effects. The website is used by academic researchers and companies (eg as a tool in drug development and for outlining new research projects). By providing peer-reviewed genetic information on CYP enzymes, the CYP-allele website has become increasingly popular and widely used. Recently, NADPH cytochrome P450 oxidoreductase (POR), the electron donor for CYP enzymes, was included on the website, which already contains 29 CYP genes, hence POR alleles are now also designated using the star allele (POR\*) nomenclature. Although most CYPs on the CYP-allele website are involved in the metabolism of xenobiotics, polymorphic enzymes with endogenous functions are also included. Each gene on the CYP-allele website has its own webpage that lists the different alleles with their nucleotide changes, their functional consequences and links to publications in which the allele has been identified and/or characterised. Thus, the CYP-allele website offers a rapid online publication of new alleles, as well as providing an overview of peer-reviewed data.

**Keywords:** pharmacogenetics, adverse drug reactions, drug response, haplotypes, drug metabolism, cytochrome P450 oxidoreductase (POR)

## Introduction

Only 30–60 per cent of patients respond properly to treatment with antidepressants, beta-blockers, statins and antipsychotic agents.<sup>1</sup> Approximately two days of prolonged hospital visits are caused by

adverse drug reactions (ADRs), and, in the USA, about 100,000 deaths are estimated to be due to ADRs every year.<sup>1</sup> These data emphasise an important problem that can be mostly explained by variable pharmacokinetics due, to a large extent, to

differences in the activity of cytochrome P450 (CYP) enzymes involved in the metabolism of the drugs. Variable metabolism, in turn, can be caused by variation in the *CYP* genes, as is the focus of the current review.

## Cytochromes P450

There are 57 active *CYP* genes in the human genome, which are divided into 18 families. The first three families (*CYP*1–3) are generally involved in the metabolism of exogenous substances such as drugs, whereas *CYP* families with higher numbers are usually involved in the metabolism of endogenous substances. *CYP* enzymes are responsible for 75–80 per cent of all phase I-dependent metabolism and for 65–70 per cent of the clearance of clinically used drugs.<sup>2,3</sup> Variation in *CYP* genes results in phenotypes classically defined as ultrarapid, extensive, intermediate and poor metabolisers. An ultrarapid metaboliser (UM) generally carries duplicated or multi-duplicated gene copies of the same allele, whereas intermediate (IM) and poor metabolisers (PM) characteristically carry one and two defective alleles (eg gene inactivation or deletion), respectively. The term extensive metaboliser (EM) is normally used for subjects carrying two alleles giving normal activity of the *CYP* enzyme (also called the \*1 or consensus allele). The metaboliser phenotypes are mainly used for describing drug metabolism, but genetic variation in *CYP*s with endogenous functions, such as in sterol, steroid, bile acid and fatty acid homeostasis, have also been well characterised, some of which give rise to disease states.

*CYP*3A4 accounts for about 50 per cent of all *CYP*-dependent drug metabolism, although individuals' capacity for *CYP*3A4-mediated drug metabolism is highly variable. No common genetic variants can account for this variation, despite the fact that 20 different alleles have been described. By contrast, *CYP*2C19 and *CYP*2D6 are highly polymorphic and together account for about 40 per cent of the metabolism of clinically used drugs. In addition, *CYP*1A2, *CYP*2A6 and *CYP*2B6 are polymorphic enzymes that significantly contribute to

xenobiotic metabolism. The characterised genetic polymorphism of these enzymes provides a basis for the possibility to adjust drug dosage and choice of drug therapy according to genotype,<sup>4–8</sup> which includes avoidance of ADRs, since polymorphic *CYP*s are frequently the cause of these, either due to the formation of high levels of metabolites or because of decreased metabolism of the parent drug.

## The Human Cytochrome P450 (CYP) Allele Nomenclature website

### Introduction

Important work in the *CYP* area focused on the identification and characterisation of polymorphic human *CYP* genes, which, in turn, created the need for a unified nomenclature system. Thus, in 1999, a nomenclature committee was formed with the aim of creating a platform for present and future allele nomenclature. Thus, the Human Cytochrome P450 Allele Nomenclature (*CYP*-allele) website (<http://www.cypalleles.ki.se/>) was launched with the purpose of managing allele designations, facilitating rapid online publication and providing a summary of alleles and their associated effects. The nomenclature system chosen for the *CYP*-allele website was based on recognised nomenclature guidelines.<sup>9–13</sup>

Currently, the website covers the nomenclature for polymorphic alleles of 29 *CYP* enzymes and NADPH cytochrome P450 oxidoreductase (POR)<sup>14</sup> (see Table 1). The *CYP*2B6, *CYP*2C9, *CYP*2C19 and *CYP*2D6 genes are particularly polymorphic, all with a high number of functionally different alleles. Each of the genes at the *CYP*-allele website has its own webpage that lists the various alleles with their nucleotide changes, molecular and functional consequences *in vitro* as well as *in vivo*, and also publications identifying or characterising the alleles. In addition, links to the National Center for Biotechnology Information (NCBI) single nucleotide polymorphism database (dbSNP) and papers with allele frequencies are presented. The number of visits is relatively constant over time, about 36,000 per year, and the website is highly cited in publications in the field of pharmacogenomics.

**Inclusion criteria for CYP alleles**

The designation of an allele (such as *CYP2B6\*4*) ideally requires determination of all sequence variations in the gene, although sequencing the intronic regions is generally not necessary. On the website, a gene is considered as the sequence from 5 kilobases upstream from the transcription start site to 500 base pairs downstream of the last exon. If a regulatory element has been characterised at a more distant part of the gene, however, it too is considered to belong to the gene. All known sequence variations within an allele are described on the CYP-allele website, although new allele numbers are currently only designated for alleles that contain at least one functional variation causing consequences such as amino acid substi-

**Table 1.** Polymorphic genes covered on the CYP-allele website. Examples of some important variant alleles are given, as well as representative substrates

Gene	No. alleles	Important variant alleles	Examples of substrates
<i>CYP1A1</i>	11	None	Benzo(a)pyrene
<i>CYP1A2</i>	16	<i>CYP1A2*1F</i> , <i>CYP1A2*1K</i>	Clozapine
<i>CYP1B1</i>	26	Many rare detrimental variants	Oestradiol
<i>CYP2A6</i>	37	<i>CYP2A6*4</i>	Nicotine
<i>CYP2A13</i>	9	<i>CYP2A13*2</i>	N-nitrosomethylphenylamine
<i>CYP2B6</i>	29	<i>CYP2B6*6</i>	Efavirenz
<i>CYP2C8</i>	14	<i>CYP2C8*3</i>	Paclitaxel
<i>CYP2C9</i>	34	<i>CYP2C9*2</i> , *3	Warfarin
<i>CYP2C19</i>	26	<i>CYP2C19*2</i> , *3, *17	Escitalopram
<i>CYP2D6</i>	78	<i>CYP2D6*4</i> , *2 <sub>xn</sub>	Tricyclic antidepressants
<i>CYP2E1</i>	7	None	Ethanol, carbon tetrachloride
<i>CYP2F1</i>	6	<i>CYP2F1*2A</i>	3-methylindole
<i>CYP2J2</i>	10	<i>CYP2J2*2-5</i>	Arachidonic acid, some drugs

Continued

**Table 1.** Continued

Gene	No. alleles	Important variant alleles	Examples of substrates
<i>CYP2R1</i>	2	<i>CYP2R1*2</i>	Vitamin D25
<i>CYP2S1</i>	5	<i>CYP2S1*3</i>	Orphan enzyme
<i>CYP2W1</i>	6	<i>CYP2W1*6</i>	Indoles
<i>CYP3A4</i>	20	<i>CYP3A4*1B</i>	Clinically used drugs
<i>CYP3A5</i>	11	<i>CYP3A5*3</i>	Cyclophosphamide
<i>CYP3A7</i>	3	<i>CYP3A7*2</i>	Clinically used drugs
<i>CYP3A43</i>	3	None	None
<i>CYP4A11</i>	None	Only single SNPs	Arachidonic acid
<i>CYP4A22</i>	15	Not known	Orphan enzyme
<i>CYP4B1</i>	7	Not known	Lauric acid
<i>CYP4F2</i>	3	<i>CYP4F2*3</i>	Vitamin K1
<i>CYP5A1</i>	9	Not known	Prostaglandin H2
<i>CYP8A1</i>	4	Not known	Prostaglandin H2
<i>CYP19A1</i>	5	<i>CYP19A1*4</i> , *5	Androstenedione
<i>CYP21A2</i>	119	Several rare variants	Progesterone
<i>CYP26A1</i>	4	<i>CYP26A1*3</i> , *4	Retinoic acid
<i>POR</i>	41	<i>POR*2</i> , *3	Electron donor to CYPs

tutions, translation terminations, splice defects, differential transcription rates etc. Nevertheless, the allele is required to be well characterised regarding linkage or lack of linkage of the consequential SNP with other nucleotide variations, including those in exons, intron–exon junctions and flanking regions. Inferring haplotypes by program analyses is generally avoided, although such alleles have occasionally been included on the CYP-allele website. When the characterised sequence variant is found in different constellations with non-causative (eg silent) ones, the different combinations are defined as sub-alleles and receive letters in addition to the number (eg *CYP2B6\*4A*, *CYP2B6\*4B*). When several effective polymorphisms are present on the same allele, however, the allelic number given is based on the polymorphism that causes the

most severe consequence, such as a splice defect (eg *CYP2C19\*2A*), so alleles that additionally contain sequence variants with less severe effects will share the same allele number, together with an additional letter (eg *CYP2C19\*2B*). Combinations of variants that are also present alone and that are considered similarly effective (eg different amino acid substitutions) are given unique allele numbers (eg *CYP2B6\*6*). Notably, the earliest described alleles on the website do not follow the nomenclature system, but the allelic designations have remained, for practical reasons.

### Submission of new alleles

When new alleles are identified, relevant information is sent to the Webmaster. Inclusion criteria (<http://www.cypalleles.ki.se/criteria.htm>) involve complete characterisation of the gene sequence, covering exons and exon–intron junctions at the minimum, investigation of linkage with other sequence variants and potential *in vitro* or *in vivo* findings. It is advised that the authors of a manuscript that describes a novel allele contact the Webmaster before submission, in order to review the data and assign a new allele name to be used in the manuscript. Usage of star allele designations that have not been approved by the nomenclature committee is strongly discouraged, because of the apparent risk of confusion and of using the same allele name for different variants. All information sent to the Webmaster is kept strictly confidential until publication of the manuscript or until the authors request it to be released. Thus, there are likely to be allele names designated by the CYP-allele website that have not yet been published, further emphasising the importance of refraining from using unauthorised allele names.

The Webmaster (and in rare cases also the Editorial and/or Advisory Board) reviews the submission to evaluate whether there are enough data to support a new allele designation. Only peer-reviewed data are thus published on the CYP-allele website. Papers describing additional characterisation of a known allele—with respect to, for example, *in vitro* or *in vivo* activity—are also peer reviewed and can be linked to the respective allele

on the webpage. Suggestions of papers that should be included with respect to further characterisation of alleles are appreciated.

## Conclusions

The CYP-allele website is widely used and well acknowledged within the scientific community. It serves the purpose of a unified and easily accessible nomenclature system for CYP enzymes, as well as for the CYP electron donor POR. The purpose of the CYP-allele website is to facilitate rapid online publication as well as providing a summary of the characteristics of specific alleles. The database has proven to be a useful resource in the area of pharmacogenomics.

## References

- Eichelbaum, M., Ingelman-Sundberg, M. and Evans, W.E. (2006), 'Pharmacogenomics and individualized drug therapy', *Annu. Rev. Med.* Vol. 57, pp. 119–137.
- Bertz, R.J. and Granneman, G.R. (1997), 'Use of *in vitro* and *in vivo* data to estimate the likelihood of metabolic pharmacokinetic interactions', *Clin. Pharmacokinet.* Vol. 32, pp. 210–258.
- Evans, W.E. and Relling, M.V. (1999), 'Pharmacogenomics: Translating functional genomics into rational therapeutics', *Science* Vol. 286, pp. 487–491.
- Weinshilboum, R. (2003), 'Inheritance and drug response', *N. Engl. J. Med.* Vol. 348, pp. 529–537.
- Ingelman-Sundberg, M. (2004), 'Pharmacogenetics of cytochrome P450 and its applications in drug therapy: The past, present and future', *Trends Pharmacol. Sci.* Vol. 25, pp. 193–200.
- Kirchheiner, J., Nickchen, K., Bauer, M., Wong, M.L. *et al.* (2004), 'Pharmacogenetics of antidepressants and antipsychotics: the contribution of allelic variations to the phenotype of drug response', *Mol. Psychiatry* Vol. 9, pp. 442–473.
- Pirmohamed, M. and Park, B.K. (2001), 'Genetic susceptibility to adverse drug reactions', *Trends Pharmacol. Sci.* Vol. 22, pp. 298–305.
- Ingelman-Sundberg, M., Oscarson, M. and McLellan, R.A. (1999), 'Polymorphic human cytochrome P450 enzymes: An opportunity for individualized drug treatment', *Trends Pharmacol. Sci.* Vol. 20, pp. 342–349.
- Shows, T.B., McAlpine, P.J., Boucheix, C., Collins, F.S. *et al.* (1987), 'Guidelines for human gene nomenclature. An international system for human gene nomenclature (ISGN, 1987)', *Cytogenet. Cell Genet.* Vol. 46, pp. 11–28.
- Daly, A.K., Brockmoller, J., Broly, F., Eichelbaum, M. *et al.* (1996), 'Nomenclature for human CYP2D6 alleles', *Pharmacogenetics* Vol. 6, pp. 193–201.
- Antonarakis, S.E. (1998), 'Recommendations for a nomenclature system for human gene mutations. Nomenclature Working Group', *Hum. Mutat.* Vol. 11, pp. 1–3.
- den Dunnen, J.T. and Antonarakis, S.E. (2001), 'Nomenclature for the description of human sequence variations', *Hum. Genet.* Vol. 109, pp. 121–124.
- Nelson, D.R., Zeldin, D.C., Hoffman, S.M., Maltais, L.J. *et al.* (2004), 'Comparison of cytochrome P450 (CYP) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes and alternative-splice variants', *Pharmacogenetics* Vol. 14, pp. 1–18.
- Sim, S.C., Miller, W.L., Zhong, X.B., Arlt, W. *et al.* (2009), 'Nomenclature for alleles of the cytochrome P450 oxidoreductase gene', *Pharmacogenet. Genomics* Vol. 19, pp. 565–566.