*Pharmacogenetics* (1996) **6**, 1–42

*Invited review*

# P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature

**David R. Nelson[1], Luc Koymans[2], Tetsuya Kamataki[3], John J. Stegeman[4], René Feyereisen[5], David J. Waxman[6], Michael R. Waterman[7], Osamu Gotoh[8], Minor J. Coon[9], Ron W. Estabrook[10], Irwin C. Gunsalus[11] and Daniel W. Nebert[12]***

[1]*Department of Biochemistry, University of Tennessee, Memphis, TN 38163, USA*
[2]*Janssen Research Foundation, Turnhoutseweg 30, B-2340 Beerse, Belgium*
[3]*Faculty of Pharmaceutical Sciences, Hokkaido University, N12 W6, Sapporo, Hokkaido, 060 Japan*
[4]*Biology Department, Woods Hole Oceanographic Institute, Woods Hole, MA 02543, USA*
[5]*Department of Entomology, University of Arizona, Tucson, AZ 85721, USA*
[6]*Division of Cell and Molecular Biology, Department of Biology, Boston University, Boston, MA 02215, USA*
[7]*Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, TN 37232, USA*
[8]*Saitama Cancer Center, 818 Komuro, Ina, Saitama 362, Japan*
[9]*Department of Biological Chemistry, University of Michigan, Ann Arbor, MI 48109, USA*
[10]*Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX 75235, USA*
[11]*Environmental Research Lab, 1 Sabine Island Drive, Pensacola Beach, FL 32561, USA*
[12]*Center for Environmental Genetics, University of Cincinnati Medical Center, Cincinnati, OH 45267-0056, USA*

Received 20 October 1995

**We provide here a list of 481 P450 genes and 22 pseudogenes, plus all accession numbers that have been reported as of October 18, 1995. These genes have been described in 85 eukaryote (including vertebrates, invertebrates, fungi, and plants) and 20 prokaryote species. Of 74 gene families so far described, 14 families exist in all mammals examined to date. These 14 families comprise 26 mammalian subfamilies, of which 20 and 15 have been mapped in the human genome and the mouse genome, respectively. Each subfamily usually represents a cluster of tightly linked genes widely scattered throughout the genome, but there are exceptions. Interestingly, the *CYP51* family has been found in mammals, filamentous fungi and yeast, and plants – attesting to the fact that this P450 gene family is very ancient. One functional *CYP51* gene and two processed pseudogenes, which are the first examples of intronless pseudogenes within the P450 superfamily, have been mapped to three different human chromosomes.**

**This revision supersedes the four previous updates in which a nomenclature system, based on divergent evolution of the superfamily, has been described. For the gene, we recommend that the italicized root symbol '*CYP*' for human ('*Cyp*' for mouse and Drosophila), representing 'cytochrome P450', be followed by an Arabic number denoting the family, a letter designating the subfamily (when two or more exist), and an Arabic numeral representing the individual gene within the subfamily. A hyphen is no longer recommended in mouse gene nomenclature. '*P*' ('*ps*' in mouse and Drosophila) after the gene number denotes a pseudogene; '*X*' after the gene number means its use has been discontinued. If a gene is the sole member of a family, the subfamily letter and gene number would be helpful but need not be included. The human nomenclature system should be used for all species other than mouse and Drosophila. The cDNAs, mRNAs and enzymes in all species (including mouse) should include all capital letters, and without italics or hyphens. This nomenclature system is similar to that proposed in our previous updates.**

## Introduction

P450 enzymes are important in the oxidative, peroxidative and reductive metabolism of numerous and diverse endogenous compounds such as steroids, bile acids, fatty acids, prostaglandins, leukotrienes, biogenic amines, retinoids, lipid hydroperoxides, and phytoalexins. Many of these enzymes also metabolize a wide range of man-made chemicals including drugs,

*To whom correspondence should be addressed.

environmental chemicals and pollutants. P450 substrates also include natural plant products involved in flavour, odour, flower colour, and the response to wounding. The metabolism of foreign chemicals can frequently produce toxic metabolites, of which some have been implicated as agents responsible for birth defects and other forms of toxicity, as well as tumour initiation and progression.

The original term 'cytochrome P-450' represents a holdover from the time when the protein was given its provisional name (Sato & Omura, 1961). These proteins are not, in fact, 'cytochromes' in the true sense of the word. The Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB) prefers the term 'heme-thiolate protein' instead of 'cytochrome' for P450.

It is now widely accepted that the P450 gene superfamily is very old, the ancestral gene having existed before the time of prokaryote/eukaryote divergence – a time which obviously predates drugs, animal-plant interactions, and combustion of organic matter (*reviewed in* Nelson *et al.*, 1993). Table 1 lists all the gene families thus far identified and their general characteristics/functions. Figure 1 illustrates a phylogenetic tree which includes a single representative of most of the eukaryotic (and selected prokaryotic) families in the P450 superfamily. It has become increasingly clear that the P450 enzymes, as well as other so-called 'drug-metabolizing' enzymes, play important roles in maintaining the steady-state levels of endogenous ligands involved in ligand-modulated transcription of genes effecting growth, differentiation, apoptosis, cellular homeostasis, and neuroendocrine functions (*reviewed in* Nebert, 1994).

The foundations for the P450 gene superfamily nomenclature system were laid with the first report (Nebert *et al.*, 1987) and have been extended in subsequent updates (Nerbert *et al.*, 1989, 1991, Nelson *et al.*, 1993). Each P450 gene almost always produces a single protein. To date, there appear to be only a few exceptions to this rule where differential splicing of the P450 transcript occurs such that entire (translated) exons or portions of exons are exchanged in order to produce an enzyme with a new catalytic activity (*reviewed in* Nelson *et al.*, 1993). Estimates of the number of individual P450 genes in any mammalian species range from 60 to 200 (*reviewed in* Nebert & Nelson, 1991), and this estimate remains reasonable.

## Naming a P450 gene or enzyme

In order to prevent incorrect assignments or duplications of gene names, colleagues are requested to send (at least) the deduced amino acid sequence of any newly discovered P450 cDNA or gene to the P450 Nomenclature Committee. The assignment of names to newly discovered P450 genes is by agreement among members of the P450 Nomenclature Committee, and will be based on decisions about the family or subfamily into which the new gene might best fit, and whether exceptions to the Committee's general rules on P450 nomenclature might be necessary. Normally, the naming of genes will be carried out chronologically, meaning that delays in submitting a new sequence to this

**Table 1.** General overview of P450 families/subfamilies and enzyme functions in various species

| | |
|---|---|
| *CYP1* | Vertebrates: dioxin-inducible; metabolism of polycyclic hydrocarbons, halogenated and heterocyclic hydrocarbons, and aromatic amines |
| *CYP2* | Vertebrates and invertebrates; metabolism of drugs and environmental chemicals |
| *CYP3* | Vertebrates; metabolism of drugs and environmental chemicals |
| *CYP4* | Vertebrates, fatty acid hydroxylases; invertebrates, unknown function(s) |
| *CYP5* | Vertebrates; thromboxane synthase |
| *CYP6* | Insects; metabolism of plant products and pesticides |
| *CYP7A* | Vertebrates; cholesterol 7α-hydroxylase |
| *CYP7B* | Vertebrates; unknown function(s) |
| *CYP8* | Vertebrates; prostacyclin synthase |
| *CYP9* | Insects |
| *CYP10* | Mollusks (mitochondrial enzyme) |
| *CYP11* | Vertebrates; cholesterol side-chain cleavage, steroid 11β-hydroxylase, and aldosterone synthase (mitochondrial enzyme) |
| *CYP12* | Insects (mitochondrial enzyme) |
| *CYP13* | Nematodes |
| *CYP14* | Nematodes |
| *CYP15* | Insects |
| *CYP16* | Nematodes |
| *CYP17* | Vertebrates; steroid 17α-hydroxylase |
| *CYP18* | Insects |
| *CYP19* | Vertebrates; aromatization of androgens |
| *CYP21* | Vertebrates; steroid 21-hydroxylase |
| *CYP24* | Vertebrates; steroid 24-hydroxylase (mitochondrial enzyme) |
| *CYP27* | Vertebrates; steroid 27-hydroxylase (mitochondrial enzyme) |
| *CYP51* | Animals, filamentous fungi, yeast and plants; sterol biosynthesis |
| *CYP52* | Yeast; alkane hydroxylases |
| *CYP53* to *CYP62* | Fungi |
| *CYP71* to *CYP92* | Plants |
| *CYP73* | Plants, cinnamic acid hydroxylase |
| *CYP101* to *CYP118* | Bacteria |

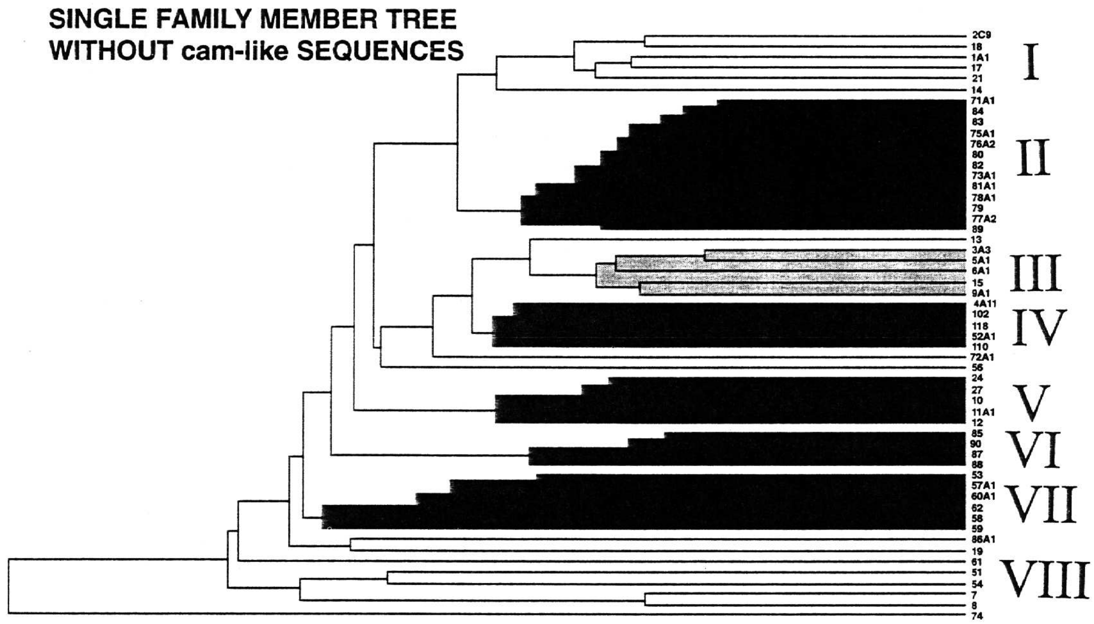


**Fig. 1.** Single-family-member phylogenetic tree, excluding cam (CYP101)-like sequences. The tree is a simple unweighted-pair-group method of analysis (UPGMA) dendrogram, calculated by using identities only and correcting for probable evolutionary multiple hits according to the scales devised by Margaret Dayhoff (1979). The equations used to approximate these tables are given in Nelson & Strobel (1987). The first five gene families (*Group I*) are present in vertebrates, yet *CYP18* has so far only been found in insects. The next 13 clustered gene families (*Group II*) exist in plants. Next, the five families (*Group III*) are found in vertebrates and invertebrates; it is very interesting that the *CYP5* family is most similar to the *CYP3* family, and together they are grouped with the insect families. The next five (*Group IV*) represent, where a catalytic function is known, fatty acid hydroxylases in both eukaryotes and prokaryotes. Genes in the next five families (*Group V*) encode mitochondrial P450 proteins and are associated, in large part, with steroidogenesis. The next four gene families (*Group VI*) exist in plants, and the next six (*Group VII*) families are found in various fungi. Among those genes in *Group VIII*, it is worth noting that the mammalian aromatase (*CYP19*) gene is most homologous to a gene (*CYP86A1*) found in *Arabidopsis thaliana*. Moreover, the fact that the sterol 14-demethylase (*CYP51*) gene – found in animals, plants, filamentous fungi and yeast – is most homologous to a gene (*CYP54*) found in *Neurospora* raises the question as to whether or not the *CYP54* gene might also be found in widely divergent taxa. Lastly, the homology between the cholesterol 7α-hydroxylase (*CYP7A1*) gene and the prostacyclin synthase (*CYP8*) gene is intriguing.

Committee might result in assignment of a higher number (or higher letter for subfamily designation). Prior consultation on the part of prospective authors has, in the past, generally worked well and has eliminated a lot of misunderstandings that characterized the earlier literature. *Prospective authors are therefore reminded that to publish temporary P450-like names for genes/proteins whose sequences have not been examined by the Nomenclature Committee will only add confusion to the literature.*

Recommendations for naming a P450 gene include the italicized root symbol '*CYP*' ('*Cyp*' for the mouse and Drosophila), denoting **cy**tochrome **P**450, an Arabic number designating the P450 family, a letter indicating the subfamily when two or more subfamilies are known to exist within that family, and an Arabic numeral representing the individual gene. With mouse genes, hyphens are no longer being used as they once were. '*P*' ('*ps*' in mouse and Drosophila) after the gene number is used to denote a pseudogene. If no subfamily or second gene exists in a family, the subfamily and gene number would be helpful but need not be included, *e.g. CYP5* or *CYP17*; if a second subfamily in that family is described, then the letter and number must be added (*e.g. Cyp7a1, Cyp7b1*).

The same numbers and letters are recommended for the corresponding gene product (mRNA, cDNA, enzyme): italicized '*CYP1A2*' or '*CYP24*' ('*Cyp1a2*' and

'*Cyp24*', respectively, in mouse) for the gene; non-italicized and all capital letters 'CYP1A2' or 'CYP24' for the cDNA, mRNA and protein in all species – including mouse and Drosophila. Also, there are no spaces between any of the letters or numbers.

In some cases, the gene is known by a trivial name before it is discovered to be a P450. This is perhaps more likely to happen in organisms such as *Arabidopsis thaliana* and *Drosophila melanogaster* (*e.g. Eig17-1* is identical to *Cyp18*). In these cases, the *CYP* name becomes an official *synonym*. If these original gene names continue to be used in publications – the officially assigned name and the species under study should be provided in the title, summary or footnote. Thus, 'olf1' and 'IIC16' are encoded by the rat *CYP2G1* and rabbit *CYP2C16* genes, respectively, and either name might be used as long as the synonym is provided. In the text of the publication one might designate the protein as 'P450 1A1' or 'P450 2G1', or simply '1A1' or '2G1', but one should stay with all capital letters.

Consistent usage of the gene nomenclature is preferred. Authors may continue to use trivial names – as long as the names are compatible with GenBank and other nucleic acid and protein databases. For the protein, this means no hyphen in 'P450', no Greek letters, and no subscripts or superscripts. The past history of naming the P450 steroidogenesis enzymes is confusing. For example, $P450_{7\alpha}$, $P450_{scc}$, $P450_{11\beta1}$, $P450_{11\beta2}$, $P450_{17\alpha}$, $P450_{arom}$, $P450_{c21}$, $P450_{c24}$ and $P450_{c27}$ might be referred to as P450c7, P450scc, P450c11b1, P450c11b2, P450c17, P450arom, P450c21, P450c24 and P450c27, respectively, but **preferably** should be called CYP7, CYP11A1, CYP11B1, CYP11B2, CYP17, CYP19, CYP21, CYP24 and CYP27, respectively.

A P450 protein sequence from one gene family usually is defined as having ≤ 40% amino acid identity to a P450 protein from any other family. This definition of a P450 gene family was originally an arbitrary decision, but unexpectedly has turned out to be very useful. In those instances examined thus far, genes within a defined subfamily have been found to be nonsegregating, *i.e.* to lie within the same 'gene cluster'. For example, the rat *CYP2D1*, *CYP2D2*, *CYP2D3*, *CYP2D4* and *CYP2D5* genes are located adjacent to one another on the same chromosome and form the '*CYP2D* cluster' (Table 2). These clusters most likely have arisen via gene duplication events, mostly during the last 400 million years (Nebert & Gonzalez, 1987; Nelson & Strobel, 1987; Gonzalez & Nebert, 1990; Nebert & Nelson, 1991). Different subfamilies within a gene family, however, are not necessarily genetically segregating distinct entities. For example, *CYP1A* and *CYP1B* are located on different chromosomes, as are *CYP2B*, *CYP2C* and *CYP2D*. By contrast, the *CYP1A*, *CYP11A* and *CYP19* clusters are near one another on the same human and the same mouse chromosome (Table 2). Yeast *CYP52A8* and *CYP52B1* are within 5 kb of one another (Seghezzi *et al.*, 1992), and Drosophila *CYP6A*, *CYP4E*, *CYP4P* and *CYP9B* clusters are within one-tenth the length of the left arm of Chr 2 (R. Feyereisen, personal communication). Also, there is an intermingling of genes in the *CYP2A*, *CYP2B* and *CYP2F* subfamilies along the long arm of human chromosome 19 (Fernandez-Salguero & Gonzalez, 1995). Finally, the *CYP51* family is so ancient – existing in animals, plants and yeast – that the functional gene and two pseudogenes exist on three different human chomosomes (Table 2).

The P450 protein sequences within a given gene family, are > 40% identical. There are more than a half dozen exceptions to this rule, however (*reviewed in* Nelson *et al.*, 1993). Because of such exceptions, which require discussion among members of the Nomenclature Committee, it is imperative that colleagues do not try to make up their own names, but rather submit their new sequences (single-letter amino-acid sequence is sufficient) to the Nomenclature Committee. Although assignment of large partial sequences to a particular subfamily is usually feasible, it may not be possible to assign names to short expressed sequence tag (EST) sequences, PCR fragments or N-terminal sequences. These fragments may have to remain as 'unassigned' until more of the sequence becomes available.

Mammalian sequences within the same subfamily are always > 55% identical. Inclusion of more distant species within the same subfamily (*e.g.* chicken CYP17 with mammalian CYP17 proteins, or trout CYP1A1 with mammalian CYP1A proteins) lowers this value to > 46%, and this value is likely to change as the database continues to expand.

Similarities in enzymic activities and in gene regulation alone *cannot* be used to determine orthologous genes across species or to classify genes within any given family or subfamily (*discussed in* Nelson *et al.*, 1993). It is worth repeating that properties in addition to sequence may provide important information relevant to the identification of orthologous P450 genes across species or phyla. For example, all genes within a given family that have been examined thus far, in vertebrates, contain the same number of exons and similar intron-exon boundaries; when available, this information is an important piece of evidence for P450 gene nomenclature assignments. The *C. elegans CYP13* family is an exception to the rule that intron-exon boundaries are fixed within a single gene family.

As we go to press, there are 481 P450 genes and 22 pseudogenes that have been described in 85 eukaryotes

**Table 2.** Chromosomal and subchromosomal localization of the human and mouse *CYP* genes. Only references that were not cited in the 1993 update are cited in this Table. If two or more genes exist in a subfamily, only the 'gene cluster' is listed in this Table. For example, *CYP1A1* and *CYP1A2* are located near one another on human chromosome 15 and represent the '*CYP1A* cluster'. *Cyp1a1* and *Cyp1a2* are near one another on mouse Chr 9 and represent the *Cyp1a* cluster. The mouse *Cyp2e1* gene is listed as the *Cyp2e* cluster, although it appears to be the only gene in this mouse subfamily. There are five known *CYP2D* functional genes in the rat *CYP2D* cluster and five known *Cyp2d* functional genes in the mouse *Cyp2d* cluster, and usually one (and at least two pseudogenes) in the human *CYP2D* cluster. Mouse *Cyp3a11*, *Cyp3a13* and *Cyp3a16* represent the *Cyp3a* cluster. Subfamily clusters not yet mapped in the human or mouse genome are also listed. Interestingly, the *CYP51* gene – being so old that it has been found in animals, filamentous fungi and yeast, and plants – has the functional gene and two pseudogenes on three different human chromosomes.

| *P450 gene cluster* | *Chromosomal location* | *References* |
|---|---|---|
| Human | | |
| *CYP1A* | 15q22-qter (near *MPI*) | |
| *CYP1B* | 2 | Sutter *et al.*, 1994 |
| *CYP2A* | 19q13.1-13.2 | |
| *CYP2B* | 19q12-q13.2 | |
| *CYP2C* | 10q24.1-24.3 | |
| *CYP2D* | 22q13.1 | |
| *CYP2E* | 10 | |
| *CYP2F* | 19 | |
| *CYP2G* | | |
| *CYP2J* | | |
| *CYP3* | 7q22.1 | |
| *CYP4A* | 1 | Bell *et al.*, 1993 |
| *CYP4B* | 1p12-p34 | |
| *CYP4F* | | |
| *CYP5* | | |
| *CYP7A* | 8q11-q12 | |
| *CYP7B* | | |
| *CYP8* | | |
| *CYP11A* | 15q23-q24 | |
| *CYP11B* | 8q21-q22 | |
| *CYP17* | 10q24.3 | |
| *CYP19* | 15q21 | |
| *CYP21* | 6p (within *HLA*) | |
| *CYP24* | 20q13.1 | Labuda *et al.*, 1993 |
| *CYP27* | 2q33-qter | |
| *CYP51* | 7 | D. Rozman, M. Stromstedt & M.R. Waterman, personal communication |
| *CYP51P1* | 3 | D. Rozman, M. Stromstedt & M.R. Waterman, personal communication |
| *CYP51P2* | 13 | D. Rozman, M. Stromstedt & M.R. Waterman, personal communication |
| Mouse | | |
| *Cyp1a* | Mid-9 (near *Mpi1*) | |
| *Cyp1b* | | |
| *Cyp2a* | 7 (near *Gpi1*) | |
| *Cyp2b* | Proximal 7 (*Coh*) | |
| *Cyp2c* | 19 | |
| *Cyp2d* | 15 | |
| *Cyp2e* | 7 | |
| *Cyp2f* | 7 | |
| *Cyp2g* | | |
| *Cyp2j* | | |
| *Cyp3* | 5 | |
| *Cyp4a* | 4 (near *Mtv13*) | Bell *et al.*, 1993 |
| *Cyp4b* | | |
| *Cyp4f* | | |

| *P450 gene cluster* | *Chromosomal location* | *References* |
|---|---|---|
| *Cyp5* | | |
| *Cyp7a* | | |
| *Cyp7b* | | |
| *Cyp8* | | |
| *Cyp11a* | 9 | |
| *Cyp11b* | 15 | |
| *Cyp17* | 19 (distal to *Got1*) | |
| *Cyp 19* | 9 | |
| *Cyp21* | 17 (within *H2*) | |
| *Cyp24* | 2 | |
| *Cyp27* | 1 | J.J. Cali & D.W. Russell, personal communication |
| *Cyp51* | | |

(including animals, filamentous fungi and yeast, and plants) and in 20 bacteria (Table 3). Of the 74 gene families so far described, 14 exist in all mammals examined to date. These 14 families comprise 26 mammalian subfamilies (clusters of genes), of which 20 (excluding two human *CYP51* pseudogenes) and 15 have been mapped in the human genome and the mouse genome, respectively (Table 2). Since the 1993 update was as comprehensive as possible, Table 3 includes only those references not included in the 1993 update, plus some older references with new accession numbers.

This P450 nomenclature system affords a convenient medium for colleagues across distantly related fields to converse with one another. What follows is a discussion of several interesting observations or potential problems that have surfaced during the past 3 years.

### *CYP1A* subfamily in nonmammalian vertebrates

On the basis of amino acid similarities and in some cases cDNA-expressed catalytic activities, at present we can be certain about the orthologous *CYP1A1* genes and the orthologous *CYP1A2* genes in all mammals and birds (Nebert & Nelson, 1991). On the other hand, the rainbow trout was found to have the *CYP1A1* gene but not the *CYP1A2* gene, suggesting that *CYP1A2* arose via a gene duplication event in land animals, following the divergence of land animals and sea animals approximately 380 million years ago. Berndtson & Chen (1994) reported two *CYP1A* genes in trout; upon closer inspection (Morrison *et al.*, 1995), what Berndtson & Chen (combined with confusion among Nomenclature Committee members) named 'CYP1A2' is actually the *CYP1A1* described by Heilmann *et al.* (1988) and what Berndtson & Chen named 'CYP1A1' represents a gene duplication event estimated to have occurred in trout about 12–14 million years ago. Although unique to trout (having appeared after divergence from its nearest related neighbours), and since this gene is in the *CYP1A* family, the Nomenclature Committee has agreed upon naming this *CYP1A3*. A similar circumstance arose previously in the *CYP2E* subfamily; all mammals appeared to have only the *CYP2E1* gene, except for rabbit which was found to have a second functional *CYP2E* gene, which was named *CYP2E2* (Nebert *et al.*, 1991).

*CYP1A* genes have now been sequenced from seven fish species (Table 3). Phylogenetic analysis of these genes indicates that the fish *CYP1A* genes are all more closely related to mammalian *CYP1A1* than *CYP1A2*. The P450 Nomenclature Committee has agreed that, if two or more *CYP1A* genes are found in the same nonmammalian vertebrate, the one most closely related to the mammalian *CYP1A1* sequence will be regarded as '*CYP1A1*' and the other will be assigned the next number available in this subfamily.

With the discovery of the mammalian *CYP1B1* gene (Table 3), it is clear that the *CYP1A1* and *CYP1B1* genes diverged at least 400–600 million years ago. With all of the evidence for a second polycyclic hydrocarbon-inducible fish P450 (Morrison *et al.*, 1995), it is extremely likely that a polycyclic hydrocarbon-inducible *CYP1B1* gene will soon be found in nonmammalian vertebrates.

### Explosion of nonmammalian P450 sequences

It is obvious, from one glance at Table 3, that the greatest increase in new P450 genes – since the Nelson *et al.* (1993) update – has come from invertebrates, plants and bacteria. The rate of newly discovered mammalian families (*CYP8*, *CYP51*) and subfamilies (*CYP1B*; and *CYP7B*, in which catalytic function has not yet been determined) has been much slower in the

**Table 3.** Update of all *CYP* genes. Only those references not cited in the 1993 update are listed in this Table; those published before 1993 but found by us since the 1993 update are included. A complete list of the older trivial names for all P450 proteins is also provided in the previous update (Nelson *et al.*, 1993). Small fragments and expressed sequence tags (ESTs) that have not yet been assigned a gene name and number, as well as allelic variants (*v1*, *v2*, etc.), are not included in this Table. All of this information, as well as information about sequence and translation reading errors, can be found on the World Wide Web: http://drnelson.utmem.edu/homepage.html

| *Gene symbol* | *Species* | *References* |
|---|---|---|
| *CYP1A1* | Rat | Omata *et al.*, 1994 |
| | | Chapman *et al.*, 1994 |
| | Human | |
| | Rabbit | |
| | Dog | |
| | Hamster | Ohgiya *et al.*, unpublished |
| | Macaque | Ohmachi *et al.*, unpublished |
| | Guinea pig | Ohgiya *et al.*, 1993 |
| | Trout | Berndtson & Chen, 1994 |
| | Plaice | Leaver *et al.*, 1993 |
| | Red sea bream | Mizukami *et al.*, 1994 |
| | Toadfish | Morrison *et al.*, 1995 |
| | Scup | Morrison *et al.*, 1995 |
| | 4-eye butterfly fish | Vrolijk *et al.*, unpublished |
| | Tomcod | Roy *et al.*, 1995 |
| | Sea bass | X. Stien, J.B. Berge *et al.*, personal communication |
| | Ice cod | J.J. Stegeman, personal communication |
| *Cyp1a1* | Mouse | Jones & Nebert, 1989 |
| *CYP1A2* | Rat | Woelfel *et al.*, 1991 |
| | | Habib *et al.*, 1994 |
| | Human | |
| | Rabbit | |
| | Hamster | |
| | Dog | |
| | Guinea pig | Black, unpublished |
| | | Mori *et al.*, unpublished |
| | Chicken | |
| *Cyp1a2* | Mouse | |
| *CYP1A3* | Trout | Berndtson & Chen, 1994 |
| *CYP1B1* | Human | Sutter *et al.*, 1994 |
| | Rat | Bhattacharyya *et al.*, 1995 |
| *Cyp1b1* | Mouse | Shen, Z. *et al.*, 1993 |
| | | Shen, Z. *et al.*, 1994 |
| | | Savas *et al.*, 1994 |
| *CYP2A1* | Rat | |
| *CYP2A2* | Rat | |
| *CYP2A3* | Rat | |
| *Cyp2a4* | Mouse | |
| *Cyp2a5* | Mouse | |
| *CYP2A6* | Human | |
| *CYP2A7* | Human | |
| *CYP2A8* | Hamster | |
| *CYP2A9* | Hamster | |
| *CYP2A10* | Rabbit | Peng *et al.*, 1993 |
| *CYP2A11* | Rabbit | Peng *et al.*, 1993 |
| *Cyp2a12* | Mouse | Iwasaki *et al.*, 1993 |
| *CYP2A13* | Human | F.J. Gonzalez, personal communication |
| *CYP2B1* | Rat | |

| *Gene symbol* | *Species* | *References* |
|---|---|---|
| *CYP2B2* | Rat | Shephard *et al.*, 1994 |
| *CYP2B3* | Rat | Jean *et al.*, 1994 |
| *CYP2B4* | Rabbit | Ryan *et al.*, 1993 |
| *CYP2B4P* | Rabbit | |
| *CYP2B5* | Rabbit | |
| *CYP2B6* | Human | Mimura *et al.*, 1993 |
| *CYP2B7PX* (discontinued number; shown to be alternative splicing product of *CYP2B6*) | | |
| *CYP2B8X* (discontinued number, shown to be promoter region of *CYP2B15*) | | |
| *Cyp2b9* | Mouse | |
| *Cyp2b10* | Mouse | |
| *CYP2B11* | Dog | |
| *CYP2B12* | Rat | |
| *Cyp2b13* | Mouse | |
| *CYP2B14X* (discontinued number; reassigned to *CYP2B16P*) | | |
| *CYP2B14P* | Rat | Jean *et al.*, 1994 |
| *CYP2B15* | Rat | Nakayama *et al.*, 1993 |
| *CYP2B16P* | Rat | Jean *et al.*, 1994<br>Trottier *et al.*, 1996 (coding sequence intact, but pseudogene because of intron splice junction mutation) |
| *CYP2B17* | Macaque | Ohmori *et al.*, 1993 |
| *CYP2B18* | Guinea pig | K. Oguri, personal communication |
| *CYP2C1* | Rabbit | Noshiro *et al.*, unpublished |
| *CYP2C2* | Rabbit | |
| *CYP2C3* | Rabbit | |
| *CYP2C4* | Rabbit | |
| *CYP2C5* | Rabbit | |
| *CYP2C6* | Rat | |
| *CYP2C6P1* | Rat | |
| *CYP2C7* | Rat | |
| *CYP2C8* | Human | |
| *CYP2C9* | Human | de Morais *et al.*, 1993<br>Sandhu *et al.*, 1993 |
| *CYP2C10X* (discontinued number; existence in doubt) | | |
| *CYP2C11* | Rat | Ström *et al.*, 1994 |
| *CYP2C12* | Rat | |
| *CYP2C13* | Rat | Legraverend *et al.*, 1994 |
| *CYP2C14* | Rabbit | |
| *CYP2C15* | Rabbit | |
| *CYP2C16* | Rabbit | |
| *CYP2C17X* (discontinued number; splice variant of *CYP2C18/2C19*) | | |
| *CYP2C18* | Human | de Morais *et al.*, 1993<br>Goldstein *et al.*, 1991 |
| *CYP2C19* | Human | Wrighton *et al.*, 1993<br>de Morais *et al.*, 1994a,b |
| *CYP2C20* | Macaque | |
| *CYP2C21* | Dog | |
| *CYP2C22* | Rat | |
| *CYP2C23* | Rat | Imaoka *et al.*, 1993b<br>Karara *et al.*, 1993<br>Marie *et al.*, 1993 |
| *CYP2C24* | Rat | Zaphiropoulos, 1993 |
| *CYP2C25* | Hamster | Sakuma *et al.*, 1994a |
| *CYP2C26* | Hamster | Sakuma *et al.*, 1994a |
| *CYP2C27* | Hamster | Sakuma *et al.*, 1994a |

| *Gene symbol* | *Species* | *References* |
|---|---|---|
| *CYP2C28* | Hamster | Sakuma *et al.*, 1995 |
| *Cyp2c29* | Mouse | Bornheim & Correia, 1989 |
| | | Matsunaga *et al.*, 1994 |
| *CYP2C30* | Rabbit | Noshiro *et al.*, unpublished |
| *CYP2C31* | Dwarf goat | Zeilmaker *et al.*, 1994 |
| *CYP2C32* | Pig | Zaphiropoulos *et al.*, 1995 |
| *CYP2C33* | Pig | Zaphiropoulos *et al.*, 1995 |
| *CYP2C34* | Pig | Zaphiropoulos *et al.*, 1995 |
| *CYP2C35* | Pig | Zaphiropoulos *et al.*, 1995 |
| *CYP2C36* | Pig | Zaphiropoulos *et al.*, 1995 |
| *CYP2C37* | Macaque | S. Ohmori, personal communication |
| *CYP2D1* | Rat | |
| *CYP2D2* | Rat | |
| *CYP2D3* | Rat | |
| *CYP2D4* | Rat | |
| *CYP2D5* | Rat | |
| *CYP2D6* | Human | |
| *CYP2D7P1* | Human | |
| *CYP2D7P2* | Human | Heim & Meyer, 1992 |
| *CYP2D8P1* | Human | |
| *CYP2D8P2* | Human | Heim & Meyer, 1992 |
| *Cyp2d9* | Mouse | Sueyoshi *et al.*, 1995 |
| *Cyp2d10* | Mouse | |
| *Cyp2d11* | Mouse | |
| *Cyp2d12* | Mouse | M. Negishi, personal communication |
| *Cyp2d13* | Mouse | M. Negishi, personal communication |
| *CYP2D14* | Cow | |
| *CYP2D15* | Dog | Sakamoto *et al.*, 1995a,b |
| | | T. Tasaki *et al.*, unpublished |
| *CYP2D16* | Guinea pig | Jiang *et al.*, 1995 |
| *CYP2D17* | Macaque | M. Lawton, personal communication |
| *CYP2D18* | Rat | Kawashima & Strobel, 1995 |
| *CYP2D19* | Marmoset | T. Igarashi *et al.*, unpublished |
| *CYP2D20* | Hamster | T. Sakuma *et al.*, unpublished |
| *CYP2D21* | Miniature pig | T. Shimojima *et al.*, unpublished |
| *CYP2E1* | Human | Watanabe *et al.*, 1994 |
| | Rat | Richardson *et al.*, 1992 |
| | | Favreau *et al.*, 1992 |
| | Rabbit | |
| | Macaque | |
| | Hamster | Puccini *et al.*, 1992 |
| | | Sakuma *et al.*, 1994b |
| | Marmoset | T. Igarashi *et al.*, unpublished |
| *Cyp2e1* | Mouse | Davis & Felder, 1993 |
| *CYP2E2* | Rabbit | |
| *CYP2F1* | Human | |
| *Cyp2f2* | Mouse | |
| *CYP2G1* | Rat | |
| | Rabbit | |
| *CYP2H1* | Chicken | |
| *CYP2H2* | Chicken | |
| *CYP2J1* | Rabbit | |
| *CYP2J2* | Human | D. Zeldin, personal communication |
| *CYP2J3* | Rat | D. Zeldin, personal communication |

| *Gene symbol* | *Species* | *References* |
|---|---|---|
| *CYP2J3P1* | Rat | D. Zeldin, personal communication |
| *CYP2J3P2* | Rat | D. Zeldin, personal communication |
| *CYP2K1* | Trout | Buhler *et al.*, 1994 |
| *CYP2K2* | Killifish | M. Oleksiak & J. Stegeman, personal communication |
| *CYP2L1* | Spiny lobster | M.O. James, personal communication |
| *CYP2M1* | Trout | Yang *et al.*, 1995 |
| | | J. Stegeman, personal communication |
| *CYP2N1* | Killifish | M. Oleksiak & J. Stegeman, personal communication |
| *CYP2P1* | Killifish | M. Oleksiak & J. Stegeman, personal communication |
| *CYP2P2* | Killifish | M. Oleksiak & J. Stegeman, personal communication |
| *CYP2P3* | Killifish | M. Oleksiak & J. Stegeman, personal communication |
| *CYP2Q1* | Xenopus laevis | H. Ohi, personal communication |
| *CYP3A1* | Rat | Telhada *et al.*, 1992 |
| | | Miyata *et al.*, unpublished |
| | | Kirita & Matsubara, 1993 |
| *CYP3A2* | Rat | Telhada *et al.*, 1992 |
| | | Cooper *et al.*, 1993 |
| | | Yamazoe & Nagata, 1995a |
| | | Miyata *et al.*, unpublished |
| *CYP3A3* | (discontinued number; existence in doubt) | |
| *CYP3A4* | Human | Hashimoto *et al.*, 1993 |
| *CYP3A5* | Human | Jounaidi *et al.*, 1994 |
| *CYP3A5P1* | Human | Schuetz & Guzelian, 1995 |
| *CYP3A5P2* | Human | P. Maurel & Y. Jounaidi, personal communication |
| *CYP3A6* | Rabbit | |
| *CYP3A7* | Human | |
| *CYP3A8* | Macaque | |
| *CYP3A9* | Rat | P. Nef, personal communication |
| *CYP3A10* | Hamster | Subramanian *et al.*, 1995 |
| | | Miyata *et al.*, unpublished |
| *Cyp3a11* | Mouse | |
| *CYP3A12* | Dog | |
| *Cyp3a13* | Mouse | Yanagimoto *et al.*, 1994 |
| *CYP3A14* | Guinea pig | Mori *et al.*, unpublished |
| *CYP3A15* | Guinea pig | Mori *et al.*, unpublished |
| *Cyp3a16* | Mouse | Itoh *et al.*, 1994 |
| *CYP3A17* | Guinea pig | T. Mori, personal communication |
| *CYP3A18* | Rat | Strotkamp *et al.*, 1994 |
| | | Yamazoe & Nagata, 1995b |
| *CYP3A19* | Dwarf goat | Zeilmaker *et al.*, 1994 |
| *CYP3A20* | Guinea pig | T. Mori *et al.*, unpublished |
| *CYP3A21* | Marmoset | T. Igarishi *et al.*, unpublished |
| *CYP3A22* | Miniature pig | T. Shimojima *et al.*, unpublished |
| *CYP3A23* | Rat | Kirita & Matsubara, 1993 |
| | | Komori & Oda, 1994 |
| *CYP4A1* | Rat | Okita & Okita, 1992 |
| | | Aldridge *et al.*, 1995 |
| *CYP4A2* | Rat | |
| *CYP4A3* | Rat | |
| *CYP4A4* | Rabbit | Palmer *et al.*, 1993a |
| *CYP4A5* | Rabbit | |
| *CYP4A6* | Rabbit | Muerhoff *et al.*, 1992 |

| *Gene symbol* | *Species* | *References* |
|---|---|---|
| *CYP4A7* | Rabbit | Sawamura *et al.*, 1993 |
| *CYP4A8* | Rat | |
| *CYP4A9* | Human | J.P. Hardwick, personal communication |
| | | Kawashima *et al.*, 1994 |
| *Cyp4a10* | Mouse | Bell *et al.*, 1993 |
| | | Henderson *et al.*, 1994 |
| *CYP4A11* | Human | Imaoka *et al.*, 1993a |
| | | Palmer *et al.*, 1993b |
| | | Bell *et al.*, 1993 |
| | | Kawashima *et al.*, 1994 |
| | | Y. Kikuta, unpublished |
| *Cyp4a12* | Mouse | Bell *et al.*, 1993 |
| *CYP4A13* | Guinea pig | Bell *et al.*, 1993 |
| *Cyp4a14* | Mouse | D. Bell, personal communication |
| *CYP4B1* | Human | J.P. Hardwick, personal communication |
| | Rabbit | Ryan *et al.*, 1993 |
| | Rat | |
| *Cyp4b1* | Mouse | Imaoka *et al.*, 1995 |
| *CYP4C1* | Cockroach | |
| *CYP4C2* | Mosquito | Scott *et al.*, 1994 |
| *Cyp4c3* | Drosophila | B. Dunkov *et al.*, personal communication |
| *CYP4C4* | Beetle roach | R. Feyereisen, personal communication |
| *CYP4C5* | Beetle roach | R. Feyereisen, personal communication |
| *CYP4C6* | Beetle roach | R. Feyereisen, personal communication |
| *Cyp4d1* | Drosophila | |
| *Cyp4d2* | Drosophila | Frolov & Alatortsev, 1994 |
| *CYP4D3* | House fly | R. Feyereisen, personal communication |
| *CYP4D4* | House fly | R. Feyereisen, personal communication |
| *CYP4D5* | Mosquito | Scott *et al.*, 1994 |
| *CYP4D6* | Mosquito | Scott *et al.*, 1994 |
| *CYP4D7* | Mosquito | Scott *et al.*, 1994 |
| *Cyp4d8* | Drosophila | B. Dunkov *et al.*, personal communication |
| *CYP4D9* | House fly | R. Feyereisen, personal communication |
| *Cyp4e1* | Drosophila | |
| *Cyp4e2* | Drosophila | Amichot *et al.*, 1994 |
| *Cyp4e3* | Drosophila | B. Dunkov *et al.*, personal communication |
| *Cyp4e4* | Drosophila | B. Dunkov *et al.*, personal communication |
| *CYP4F1* | Rat | Chen & Hardwick, 1993 |
| *CYP4F2* | Human | L. Chen & J.P. Hardwick, unpublished |
| | | Kikuta *et al.*, 1994 |
| *CYP4F3* | Human | Kikuta *et al.*, 1993 |
| *CYP4F4* | Rat | H. Kawashima & H.W. Strobel, personal communication |
| *CYP4F5* | Rat | H. Kawashima & H.W. Strobel, personal communication |
| *CYP4F6* | Rat | H. Kawashima & H.W. Strobel, personal communication |
| *Cyp4g1* | Drosophila | L.C. Waters, personal communication |
| | | B. Dunkov *et al.*, personal communication |
| *CYP4G2* | House fly | R. Feyereisen, personal communication |
| *CYP4G3* | House fly | R. Feyereisen, personal communication |
| *CYP4H1* | Mosquito | Scott *et al.*, 1994 |
| *CYP4H2* | Mosquito | Scott *et al.*, 1994 |
| *CYP4H3* | Mosquito | Scott *et al.*, 1994 |

| *Gene symbol* | *Species* | *References* |
|---|---|---|
| *CYP4H4* | Mosquito | Scott *et al.*, 1994 |
| *CYP4H5* | Mosquito | Scott *et al.*, 1994 |
| *CYP4H6* | Mosquito | Scott *et al.*, 1994 |
| *CYP4H7* | Mosquito | Scott *et al.*, 1994 |
| *CYP4H8* | Mosquito | Scott *et al.*, 1994 |
| *CYP4H9* | Mosquito | Scott *et al.*, 1994 |
| *CYP4J1* | Mosquito | Scott *et al.*, 1994 |
| *CYP4J2* | Mosquito | Scott *et al.*, 1994 |
| *CYP4J3* | Mosquito | Scott *et al.*, 1994 |
| *CYP4K1* | Mosquito | Scott *et al.*, 1994 |
| *CYP4L1* | Tobacco hornworm | Snyder *et al.*, 1995 |
| *CYP4L2* | Tobacco hornworm | Snyder *et al.*, 1995 |
| *CYP4M1* | Tobacco hornworm | Snyder *et al.*, 1995 |
| *CYP4M2* | Tobacco hornworm | Snyder *et al.*, 1995 |
| *CYP4M3* | Tobacco hornworm | Snyder *et al.*, 1995 |
| *CYP4N1* | House fly | R. Feyereisen, personal communication |
| *CYP4N2* | House fly | R. Feyereisen, personal communication |
| *Cyp4p1* | Drosophila | B. Dunkov *et al.*, personal communication |
| *CYP5* | Human | Wang *et al.*, 1991 |
| | | Yokoyama *et al.*, 1991 |
| | | Ohashi *et al.*, 1992 |
| | | Lee KD *et al.*, 1994 |
| | | Miyata *et al.*, 1994b |
| | Rat | Tone *et al.*, 1994 |
| | Pig | Shen RF *et al.*, 1994 |
| *Cyp5* | Mouse | Zhang *et al.*, 1993 |
| *CYP6A1* | House fly | Cohen *et al.*, 1994 |
| *Cyp6a2* | Drosophila | |
| *CYP6A3* | House fly | Cohen & Feyereisen, 1995 |
| *CYP6A4* | House fly | Cohen & Feyereisen, 1995 |
| *CYP6A5* | House fly | Cohen & Feyereisen, 1995 |
| *CYP6A6* | House fly | Cohen & Feyereisen, 1995 |
| *CYP6A7* | House fly | R. Feyereisen, personal communication |
| *Cyp6a8* | Drosophila | R. Ganguly, personal communication |
| *Cyp6a9* | Drosophila | R. Ganguly, personal communication |
| *CYP6B1* | Black swallowtail butterfly | Prapaipong *et al.*, 1994 |
| | | Ma *et al.*, 1994 |
| *CYP6B2* | Cotton bollworm | Wang & Hobbs, unpublished |
| *CYP6B3* | Black swallowtail butterfly | Hung *et al.*, 1995 |
| *CYP6B4* | *Papilio glaucus* | M. Schuler, personal communication |
| *CYP6C1* | House fly | Cohen & Feyereisen, 1995 |
| *CYP6C2* | House fly | Cohen & Feyereisen, 1995 |
| *CYP6D1* | House fly | Tomita & Scott, 1995 |
| *CYP7A1* | Human | Nishimoto *et al.*, 1993 |
| | | Thompson *et al.*, 1993 |
| | | Wang & Chiang, 1994 |
| | Rat | Hoekman *et al.*, 1993 |
| | | Chiang & Stroup, 1994 |

| *Gene symbol* | *Species* | *References* |
|---|---|---|
| | | Lee YH *et al.*, 1994 |
| | | Ramirez *et al.*, 1994 |
| | Rabbit | Poorman *et al.*, 1993 |
| | | Kai *et al.*, 1995 |
| | Hamster | Crestani *et al.*, 1993 |
| *Cyp7a1* | Mouse | Tzung *et al.*, 1994 |
| *CYP7B1* | Human | K. Rose, G. Stapleton & R. Lathe, personal communication |
| | Rat | Stapleton *et al.*, 1995 |
| *Cyp7b1* | Mouse | Stapleton *et al.*, 1995 |
| *CYP8* | Human | Miyata *et al.*, 1994a |
| | Cow | Pereira *et al.*, 1993 |
| | | Pereira *et al.*, 1994 |
| | | Hara *et al.*, 1994 |
| *CYP9A1* | Tobacco budworm | Rose *et al.*, unpublished |
| *Cyp9b1* | Drosophila | B. Dunkov *et al.*, personal communication |
| *Cyp9b2* | Drosophila | B. Dunkov *et al.*, personal communication |
| *Cyp9c1* | Drosophila | B. Dunkov *et al.*, personal communication |
| *CYP10* | Pond snail | |
| *CYP11A1* | Human | |
| | Cow | |
| | Pig | Urban *et al.*, 1994 |
| | Rat | |
| | Rabbit | Yang *et al.*, 1993 |
| | Sheep | Okuyama *et al.*, unpublished |
| | Goat | Okuyama *et al.*, unpublished |
| | Chicken | |
| | Trout | Takahashi *et al.*, 1993 |
| *CYP11B1* | Human | Naiki *et al.*, 1993 |
| | Cow | |
| | Rat | Nomura *et al.*, 1993 |
| | | Matsukawa *et al.*, 1993 |
| | | Mukai *et al.*, 1993 |
| | Sheep | Boon *et al.*, 1995 |
| | | Anwar *et al.*, 1994 |
| | Pig | Sun *et al.*, 1995 |
| | Hamster | J.G. LeHoux, personal communication |
| *Cyp11b1* | Mouse | |
| *CYP11B2* | Human | |
| | Rat | Nomura *et al.*, 1993 |
| | | Mukai *et al.*, 1993 |
| | | Zhou *et al.*, 1993 |
| | Cow | |
| | Hamster | LeHoux *et al.*, 1994 |
| | Bullfrog | Nonaka *et al.*, 1995 |
| *Cyp11b2* | Mouse | |
| *CYP11B3* | Rat | Mukai *et al.*, 1993 |
| | | Nomura *et al.*, 1993 |
| | | Zhou *et al.*, unpublished |
| *CYP11B4* | Cow | |
| *CYP11B5P* | | |
| *CYP11B6P* | | |
| *CYP11B7P* | | |

| *Gene symbol* | *Species* | *References* |
|---|---|---|
| *CYP11B8P* | Rat | Mukai *et al.*, 1993 |
| *CYP12A1* | House fly | V. Guzov *et al.*, personal communication |
| *CYP13A1* | Caenorhabditis | Wilson *et al.*, 1994 |
| *CYP13A2* | Caenorhabditis | Wilson *et al.*, 1994 |
| *CYP13A3* | Caenorhabditis | Wilson *et al.*, 1994 |
| *CYP13A4* | Caenorhabditis | Wilson *et al.*, 1994 |
| *CYP13A5* | Caenorhabditis | Wilson *et al.*, 1994 |
| *CYP13A6* | Caenorhabditis | Wilson *et al.*, 1994 |
| *CYP13A7* | Caenorhabditis | Wilson *et al.*, 1994 |
| *CYP13A8* | Caenorhabditis | Wilson *et al.*, 1994 |
| *CYP13A9P* | Caenorhabditis | Wilson *et al.*, 1994 |
| *CYP13A10* | Caenorhabditis | Wilson *et al.*, 1994 |
| *CYP14A1* | Caenorhabditis | Wilson *et al.*, 1994 |
| *CYP14A2* | Caenorhabditis | Wilson *et al.*, 1994 |
| *CYP14A3* | Caenorhabditis | Wilson *et al.*, 1994 |
| *CYP14A4* | Caenorhabditis | Wilson *et al.*, 1994 |
| *Cyp15* | Drosophila | R. Feyereisen, personal communication |
| *CYP16* | Caenorhabditis | Wilson *et al.*, 1994 |
| *CYP17* | Human | Biason *et al.*, 1991 |
| | Cow | Hornsby *et al.*, 1992 |
| | Pig | |
| | Chicken | |
| | Rat | Nason *et al.*, 1992 |
| | | Givens *et al.*, 1994 |
| | Trout | |
| | Guinea pig | |
| | Hamster | J.G. LeHoux, personal communication |
| | Sheep | Murray *et al.*, unpublished |
| | Dogfish | J. Trant, personal communication |
| *Cyp17* | Mouse | |
| *Cyp18* | Drosophila | McCarthy & Waterman, personal communication |
| | | Hurban & Thummel, 1993 |
| *CYP19* | Human | Kilgore *et al.*, 1992 |
| | | Harada *et al.*, 1993 |
| | | Mahendroo *et al.*, 1993 |
| | | Toda & Shizuta, 1993 |
| | | Honda *et al.*, 1994 |
| | | Katsumi *et al.*, 1994 |
| | | Toda & Shizuta, 1994 |
| | | Toda *et al.*, 1994 |
| | Chicken | |
| | Rat | Fitzpatrick & Richards, 1993 |
| | Trout | |
| | Goldfish | G. Callard, personal communication |
| | Cow | Hinshelwood *et al.*, 1993 |
| | | Vanselow & Furbass, unpublished |
| | Japanese quail | Harada *et al.*, 1994 |
| | Zebra finch | Shen P *et al.*, 1994 |
| | Channel catfish | J. Trant, personal communication |
| | Medaka | Tanaka *et al.*, 1995 |
| *CYP19* | Mouse | |

| *Gene symbol* | *Species* | *References* |
|---|---|---|
| *CYP21* | Human | Globerman *et al.*, 1988 |
| | | Partanen & Campbell, unpublished |
| | Cow | |
| | Pig | |
| | Sheep | |
| *Cyp21* | Mouse | |
| *CYP21P* | Human | Collier *et al.*, 1993 |
| *Cyp21ps* | Mouse | |
| *CYP24* | Human | Chen *et al.*, 1993 |
| | | Labuda *et al.*, 1993 |
| | Rat | Ohyama *et al.*, 1993 |
| | | Hahn *et al.*, 1994 |
| | | Ohyama *et al.*, 1994 |
| | | Zierold *et al.*, 1994 |
| *CYP27* | Human | Guo *et al.*, 1993 |
| | Rabbit | |
| | Rat | |
| *CYP51* | Human | Y. Aoyama, M. Noshiro & Y. Yoshida, personal communication |
| | | D. Rozman, M. Stromstedt & M.R. Wateman, personal communication |
| | Rat | Aoyama *et al.*, 1994 |
| | | Y. Aoyama, M. Noshiro & Y. Yoshida, personal communication |
| | | Sloane *et al.*, 1995 |
| | *S. cerevisiae* | Johnston *et al.*, 1994 |
| | | Burns *et al.*, 1994 |
| | *S. pombe* | S. Hunt, K. Devlin & C.M. Churcher, unpublished |
| | *C. tropicalis* | |
| | *C. albicans* | |
| | *I. orientalis* | Burgener-Kairuz *et al.*, 1994 |
| | *C. krusei* | Burgener-Kairuz *et al.*, 1994 |
| | *T. glabrata* | Burgener-Kairuz *et al.*, 1994 |
| | *U. maydis* | Hargreaves & Keon, unpublished |
| | *Penicillium italicum* | Nistelrooy *et al.*, unpublished |
| | Wheat | F. Durst, personal communication |
| *CYP51P1* | Human | D. Rozman, M. Stromstedt & M.R. Waterman, personal communication |
| *CYP51P2* | Human | D. Rozman, M. Stromstedt & M.R. Waterman, personal communication |
| *CYP52A1* | *C. tropicalis* | |
| *CYP52A2* | *C. tropicalis* | |
| *CYP52A3* | *C. maltosa* | |
| *CYP52A4* | *C. maltosa* | |
| *CYP52A5* | *C. maltosa* | |
| *CYP52A6* | *C. tropicalis* | |
| *CYP52A7* | *C. tropicalis* | |
| *CYP52A8* | *C. tropicalis* | |
| *CYP52A9* | *C. maltosa* | Ohkuma *et al.*, 1995 |
| *CYP52A10* | *C. maltosa* | M. Takagi, personal communication |
| *CYP52A11* | *C. maltosa* | M. Takagi, personal communication |
| | | M. Ohkuma, unpublished |
| *CYP52B1* | *C. tropicalis* | Seghezzi *et al.*, 1992 |
| *CYP52C1* | *C. tropicalis* | Seghezzi *et al.*, 1992 |
| *CYP52C2* | *C. maltosa* | M. Takagi, unpublished |
| *CYP52D1* | *C. maltosa* | M. Takagi, unpublished |

| Gene symbol | Species | References |
|---|---|---|
| CYP52E1 | C. apicola | K. Lottermoser, unpublished |
| CYP52E2 | C. apicola | K. Lottermoser, personal communication |
| CYP53A1 | Aspergillus niger | |
| CYP53B1 | Rhodotorula minuta | Fukuda et al., unpublished |
| CYP54 | Neurospora crassa | |
| CYP55A1 | Fusarium oxysporum | Tomura et al., 1994 |
| CYP55A2 | Cylindrocarpon tonkinense | Usuda et al., 1995 |
| CYP55A3 | Cylindrocarpon tonkinense | Usuda et al., 1995 |
| CYP56 | S. cerevisiae | |
| CYP57A1 | Nectria haematococca | Maloney & VanEtten, 1994 |
| CYP57A2 | Nectria haematococca | Reimmann & VanEtten, personal communication |
| CYP58 | Fusarium sporotrichioides | T.M. Hohn, personal communication |
| CYP59 | Aspergillus nidulans | Keller et al., 1994<br>Brown et al., personal communication |
| CYP60A1 | Aspergillus parasiticus | Yu et al., 1995 |
| CYP60A2 | Aspergillus nidulans | Brown et al., personal communication |
| CYP60B1 | Aspergillus nidulans | Brown et al., personal communication |
| CYP61 | S. cerevisiae | Lye & Churcher, unpublished |
| CYP62 | Aspergillus nidulans | Brown et al., personal communication |
| CYP71A1[a] | Persea americana | |
| CYP71A2 | Solanum melongena | Umemoto et al., 1993 |
| CYP71A3 | Solanum melongena | Umemoto et al., 1993 |
| CYP71A4 | Solanum melongena | Umemoto et al., 1993 |
| CYP71A5 | Nepeta racemosa | D. Hallahan, personal communication |
| CYP71A6 | Nepeta racemosa | D. Hallahan, personal communication |
| CYP71A7 | Catharanthus roseus | Meijer et al., 1993 |
| CYP71A8 | Mentha piperita | Kang & Choi, unpublished |
| CYP71B1 | Thlaspi arvense | Udvardi et al., 1994 |
| CYP71B2 | Arabidopsis thaliana | M. Mizutani, personal communication |
| CYP71B3 | Arabidopsis thaliana | M. Mizutani, personal communication |
| CYP71B4 | Arabidopsis thaliana | M. Mizutani, personal communication |
| CYP71B5 | Arabidopsis thaliana | M. Mizutani, personal communication |
| CYP71B6 | Arabidopsis thaliana | M. Mizutani, personal communication |
| CYP71C1 | Zea mays | Frey et al., 1995 |
| CYP71C2 | Zea mays | Frey et al., 1995<br>S. Potter, personal communication |
| CYP71C3 | Zea mays | Frey et al., 1995 |
| CYP71C4 | Zea mays | Frey et al., 1995 |
| CYP71C5 | Zea mays | S. Potter, personal communication |
| CYP71D1 | Catharanthus roseus | J. Schroeder, personal communication |
| CYP71D2 | Catharanthus roseus | J. Schroeder, personal communication |
| CYP71D3 | Arabidopsis thaliana | EST Z27299 |
| CYP72A1 | Catharanthus roseus | Vetter et al., 1992<br>Meijer et al., 1993<br>Mangold et al., 1994 |

| *Gene symbol* | *Species* | *References* |
|---|---|---|
| *CYP72A2* | *Arabidopsis thaliana* | EST T13009 |
| *CYP73A1* | *Helianthus tuberosus* | Teutsch *et al.*, 1993 |
| | | Werck-Reichhart *et al.*, 1993 |
| *CYP73A2* | *Phaseolus aureus* | Mizutani *et al.*, 1993 |
| *CYP73A3* | *Medicago sativa* | Fahrendorf & Dixon, 1993 |
| *CYP73A4* | *Catharanthus roseus* | Meijer *et al.*, 1993 |
| | | Hotze & Schroeder, unpublished |
| *CYP73A5* | *Arabidopsis thaliana* | M. Mizutani, personal communication |
| | | P. Urban & M. Kazmaier, personal communication |
| *CYP73A6* | *Zea mays* | S. Potter, personal communication |
| *CYP73A7* | *Zea mays* | S. Potter, personal communication |
| *CYP73A8* | *Zea Mays* | S. Potter, personal communication |
| *CYP73A9* | *Pisum sativum* | Frank *et al.*, 1995 |
| *CYP74* | *Linum usitatissimum* | Song *et al.*, 1993 |
| | *Parthenium argentatum* | Pan *et al.*, 1995 |
| | *Arabidopsis thaliana* | T. Newman, unpublished |
| *CYP75A1* | *Petunia hybrida* | Toguri *et al.*, 1993a |
| | | Holton *et al.*, 1993 |
| | | Ohbayashi *et al.*, unpublished |
| *CYP75A2* | *Solanum melongena* | Toguri *et al.*, 1993b,c |
| *CYP75A3* | *Petunia hybrida* | Holton *et al.*, 1993 |
| *CYP76A1* | *Solanum melongena* | Toguri *et al.*, 1993b,c |
| *CYP76A2* | *Solanum melongena* | Toguri *et al.*, 1993b,c |
| *CYP76B1* | *Helianthus tuberosus* | F. Durst, personal communication |
| *CYP76C1* | *Arabidopsis thaliana* | M. Mizutani, personal communication |
| *CYP77A1* | *Solanum melongena* | Toguri & Tokugawa, 1994 |
| *CYP77A2* | *Solanum melongena* | Toguri & Tokugawa, 1994 |
| *CYP78A1* | *Zea mays* | Larkin, 1994 |
| *CYP78A2* | *Phalaenopsis sp.* (orchid) | J. Nadeau & S. O'Neill, personal communication |
| *CYP79* | *Sorghum bicolor* | Sibbesen *et al.*, 1994 |
| | | Koch *et al.*, 1995 |
| *CYP80* | *Berberis stolonifera* | Kraus & Kutchan, 1995 |
| *CYP81A1* | *Zea mays* | S. Potter, personal communication |
| *CYP81A2* | *Zea mays* | S. Potter, personal communication |
| *CYP81A3* | *Zea mays* | S. Potter, personal communication |
| *CYP81A4* | *Zea mays* | S. Potter, personal communication |
| *CYP81B1* | *Helianthus tuberosus* | F. Durst, personal communication |
| *CYP82* | *Pisum sativum* | Frank *et al.*, 1995 |
| *CYP83A1* | *Arabidopsis thaliana* | C. Chappel, unpublished |
| *CYP83B1* | *Arabidopsis thaliana* | M. Mizutani, personal communication |
| *CYP84* | *Arabidopsis thaliana* | C. Chappel, personal communication |
| *CYP85* | *Lycopersicon esculentum* | G. Bishop, personal communication |
| *CYP86A1* | *Arabidopsis thaliana* | F. Durst, personal communication |
| | | GenEMBL Z26357, Z26358 |

| *Gene symbol* | *Species* | *References* |
|---|---|---|
| *CYP86A2* | *Arabidopsis thaliana* | GenEMBL T04172 |
| *CYP86A3* | *Brassica campestris* | GenEMBL L35832 |
| *CYP87* | *Helianthus annuus* | A. Steinmetz & F. Durst, personal communication |
| *CYP88* | *Zea mays* | R. Winkler, personal communication |
| *CYP89* | *Vicia sativa* | F. Durst, personal communication |
| *CYP90* | *Arabidopsis thaliana* | M. Szekeres, unpublished |
| *CYP91A1* | *Arabidopsis thaliana* | M. Mizutani, personal communication |
| *CYP91A2* | *Arabidopsis thaliana* | M. Mizutani, personal communication |
| *CYP92* | *Zea mays* | M. Persans, personal communication |
| *CYP101* | *Pseudomonas putida* | Koga *et al.*, 1989 |
| *CYP102* | *Bacillus megaterium* | Ravichandran *et al.*, 1993 |
| *CYP103* | *Agrobacterium tumefaciens* | |
| *CYP104* | *Agrobacterium tumefaciens* | |
| *CYP105A1* | *Streptomyces griseolus* | |
| *CYP105A2* | *Amycolata autotrophica* | Kawauchi *et al.*, 1994 |
| *CYP105B1* | *Streptomyces griseolus* | |
| *CYP105C1* | *Streptomyces sp.* | |
| *CYP105D1* | *Streptomyces griseus* | Trower *et al.*, 1993 |
| *CYP105E1* | *Rhodococcus fascians* | Crespi *et al.*, 1994 |
| *CYP106A1* | *Bacillus megaterium* | |
| *CYP106A2* | *Bacillus megaterium* | Rauschenbach *et al.*, 1993 |
| *CYP107A1* | *Saccharopolyspora erythraea* | |
| *CYP107B1* | *Saccharopolyspora erythraea* | |
| *CYP107C1* | *Streptomyces thermotolerans* | Arisawa *et al.*, 1994<br>Schoner *et al.*, 1992 |
| *CYP107D1* | *Streptomyces antibioticus* | Rodriguez *et al.*, unpublished |
| *CYP107E1* | *Micromosospora griseorubida* | Inouye *et al.*, 1994 |
| *CYP107F1* | *Streptomyces griseus* | Ueda & Horinouchi, unpublished |
| *CYP107G1* | *Streptomyces hygroscopicus* | Schwecke *et al.*, 1995 |
| *CYP108* | *Pseudomonas spp.* | Hasemann *et al.*, 1994 |
| *CYP109* | *Bacillus subtilis* | |
| *CYP110* | *Anabaena sp.* | |
| *CYP111* | *Pseudomonas incognita* | Ropp *et al.*, 1993 |
| *CYP112* | *Bradyrhizobium japonicum* | Tully & Keister, 1993 |

| *Gene symbol* | *Species* | *References* |
|---|---|---|
| *CYP113A1* | *Saccharopolyspora erythraea* | Stassi *et al.*, 1993 |
| *CYP113B1* | *Streptomyces fradiae* | Merson-Davies & Cundliffe, 1994 |
| *CYP114* | *Bradyrhizobium japonicum* | Tully & Keister, 1993 |
| *CYP115P* | *Bradyrhizobium japonicum* | Tully & Keister, 1993 |
| *CYP116* | *Rhodococcus sp.* | Eltis *et al.*, 1993<br>Nagy *et al.*, 1995 |
| *CYP117* | *Bradyrhizobium japonicum* | Tully & Keister, 1993 |
| *CYP118* | *Mycobacterium leprae* | D.R. Smith, unpublished |

[a]The common names for most of the plant species can be found on Nelson's homepage in the World Wide Web and in Table 4.

past 3 years. This may reflect the possibility that there are fewer mammalian families/subfamilies yet to be discovered. The answer as to whether or not this possibility is correct can be found in the EST databases for human sequences that are already available (*vide infra*).

## Accession numbers for the P450 genes and gene products

We have attempted to amass all available P450 accession numbers from SwissProt, NBRF-PIR, and GenBank/EMBL (Table 4). The means of obtaining and maintaining this database have been described in Nelson *et al.* (1993). Some accession numbers appear more than once in Table 4. Some accession numbers were found as cross-references to other databases, and some numbers were obtained from published papers. This information can also be found on the World Wide Web site at (http://drnelson.utmem.edu/homepage.html). If you know of additional accession numbers not listed in Table 4, please let us know.

**Table 4.** Accession numbers for cytochrome P450 sequence

| *Gene* | *Species* | *PIR Database* | | *SwissProt* | *GenEMBL* | *DDBJ* | | |
|---|---|---|---|---|---|---|---|---|
| *CYP1A1* | hum | B23585 | A24797 | P04798 | M12079 | M31664 → | M31667 | X04300 |
| | | S15803 | S16714 | | X02612 | K03191 | D12525 | D01198 |
| | | S19336 | | | D10855 | D01150 | | |
| | mon | S21761 | | P33616 | D17575 | | | |
| | rat | A00185 | A24406* | P00185 | K02246 | M12170 | M14633 | M26126 |
| | | S09617* | S15584 | | M26129 | X00469 | I00732 | |
| | | A93513 | D60822* | | | | | |
| | | S45716 | | | | | | |
| | ham | JX0189 | JS0746 | Q00557 | D10251 | D12977 | D10913 | |
| | gpi | F24406* | S43414 | | D11043 | | | |
| | | S43414 | | | | | | |
| | rab | A27821 | B27821 | P05176 | D00212 | M11727 | M19917 | X05685 |
| | | A25143 | | | | | | |
| | dog | C37222 | | | | | | |
| *CYP1A1* | tro | A28789 | | P10609 | M21310 | S69277 | | |
| *CYP1A1* | Ppl | S34184 | | | X73631 | | | |
| *CYP1A1* | Ota | | | | U14161 | | | |
| *CYP1A1* | Sch | | | | U14162 | | | |

| Gene | Species | PIR Database | | SwissProt | GenEMBL | DDBJ | | |
|---|---|---|---|---|---|---|---|---|
| *CYP1A1* | Mto | | | | L41886 | L41917 | | |
| *CYP1A1* | Cca | | | | U19855 | | | |
| *CYP1A1* | Pma | | | | | | | |
| *Cyp1a1* | mou | A00184 | S15588 | P00184 | Y00071 | M10021 | K02588 | X01681 |
| | | A23923 | A24953* | | M11515 | M33935 | M25623 | |
| | | C24406* | | | | | | |
| *CYP1A2* | hum | A23585 | S07373 | P05177 | L00383 → | L00389 | M12078 | M14337 |
| | | S16718 | S22433 | | M31664 → | M31667 | Z00036 | M55053 |
| | | A25892 | A60881* | | M38504 | U02993 | | |
| | mon | | | | | | | |
| | rat | A22562 | B24406* | P04799 | K02422 | K03241 | M26127 | X01031 |
| | | A20963 | S16875 | | | | | |
| | | S26822* | D60822* | | | | | |
| | | A61400 | A44612* | | | | | |
| | ham | S13885 | JX0190 | P24453 | M34446 | M63787 | D10252 | D10914 |
| | | PX0036* | | | | | | |
| | gpi | | | | D50457 | U23501 | | |
| | rab | A00187 | S02038 | P00187 | D00213 | M11728 | M36538 | X05686 |
| | | B25143 | A00188 | | X13853 | | | |
| | | B27821 | | | | | | |
| | dog | B37222 | A60463* | | | | | |
| | chi | JT0575 | | Q01741 | M64537 | | | |
| *Cyp1a2* | mou | A00186 | B23923 | P00186 | X00479 | X04283 | K02589 | X01682 |
| | | A26373 | A93512 | | M10022 | M25624 | | |
| | | B92495 | B24953* | | | | | |
| | | B45955* | D24406* | | | | | |
| | | E24406* | A45955* | | | | | |
| *CYP1A3* | tro | | | | S69278 | | | |
| *CYP1B1* | hum | | | | U03688 | | | |
| *CYP1B1* | rat | | | | X83867 | | | |
| *Cyp1b1* | mou | | | | U02479 | X78445 | U03283 | |
| *CYP2A1* | rat | A29560 | A34272 | P11711 | J02669 | M33312 | | |
| | | S03981* | C41425* | | | | | |
| *CYP2A2* | rat | A31887 | B34272 | P15149 | J04187 | M33313 | M33325 | M34392 |
| | | S03982* | S26821* | | | | | |
| *CYP2A3* | rat | S15056 | A32030 | P20812 | M33190 | J02852 | | |
| | | S12708 | | | | | | |
| *Cyp2a4* | mou | A33531 | S16067 | P15392 | M25146 | M25147 | M26202 | M26203 |
| | | A30499 | | | M26205 → | M26208 | J03549 | J04631 |
| *Cyp2a5* | mou | B33531 | S16068 | P20852 | M25204 → | M25211 | M26204 | J04631 |
| | | B30499 | | | M19319 | | | |
| *Cyp2a4 or 5* | | S03979 | | | | | | |
| *CYP2A6* | hum | A00190 | S05946 | P11509 | K03192 | M33316 | M33318 | X06401 |
| | | S04581 | S09329 | P00190 | X13897 | X13929 | X13930 | |
| | | A34271 | B34271 | P10890 | | | | |
| | | S04698 | A61272* | | | | | |
| | | S17220* | | | | | | |
| *CYP2A7* | hum | C34271 | | P20853 | M33317 | | | |
| *CYP2A8* | ham | S13884 | A33293 | P24454 | M27906 | M34446 | M34447 | M63788 |
| | | PX0037* | | | | | | |
| *CYP2A9* | ham | | P24455 | M34446 | M34448 | M63789 | | |
| *CYP2A10* | rab | | Q05555 | L10236 | | | | |
| *CYP2A11* | rab | | | Q05556 | L10237 | | | |
| *CYP2A10/11* | rab | A31944* | | | | | | |
| *Cyp2a12* | mou | S32491 | | | L06463 | | | |
| *CYP2A13* | hum | | | | | | | |

| *Gene* | *Species* | *PIR Database* | | *SwissProt* | *GenEMBL* | *DDBJ* | | |
|---|---|---|---|---|---|---|---|---|
| *CYP2A* | bov | A35704* | | P22779 | | | | |
| *CYP2A* | mon | S36874* | | | | | | |
| *CYP2A* | bab | S21737* | | P80055* | | | | |
| *CYP2B1* | rat | A00176 | S03854 | P00176 | D00250 | K01721 | L00313 → | L00320 |
| | | A22363 | S19172 | | M11251 | M26854* | M26855* | M37134 |
| | | | | | J00719 | M19972 | M21412 | |
| *CYP2B2* | rat | A00177 | A21872 | P04167 | J00720 → | J00728 | K00996 | K01626 |
| | | S15589 | B00176 | | K01721 | K02427 | K02428 | M13234 |
| | | S03855* | A32736 | | M15947 | M26853* | M27076 | I00525 |
| | | A21162 | | | M34452 | Y00410 | M21403 | M13650 |
| | | | | | S51970 | L28169 | | |
| *CYP2B1* or *2B2* | | A34259* | A92255* | | | | | |
| | | B92255* | A29298* | | | | | |
| | | A60822* | | | | | | |
| *CYP2B3* | rat | A25459 | A29818 | P13107 | M20406 | M19973 | U16209 → | U16214 |
| *CYP2B4* | rab | A00178 | A00179 | P00177 | M20856 | M20857 | L10912 | S64259 |
| | | B27717 | C27717 | P00178 | | | | |
| | | E27717 | S31279 | P00179 | | | | |
| | | S31278 | S35666 | | | | | |
| | | S31277 | | | | | | |
| *CYP2B4P* | rab | | | | | | | |
| *CYP2B5* | rab | D27717 | A27717 | P12789 | M18820 | M20855 | | |
| *CYP2B6* | hum | S04578 | A32969 | P20813 | J02864 | M29874 | X06399 | X06400 |
| | | S07598 | S04579 | | X13494 | X16864 | | |
| | | S04580 | | | | | | |
| *CYP2B7PX* | hum | | | | J02864 | M29873 | | |
| *CYP2B8* | rat | | | | J04808 | | | |
| *Cyp2b9* | mou | A31047 | | P12790 | M21855 | M60267 → | M60273 | |
| *Cyp2b10* | mou | B31047 | A60559* | P12791 | M21856 | | | |
| *CYP2B11* | dog | S11305 | | P24460 | M33575 | M92447 | | |
| *CYP2B12* | rat | S18907 | S27160 | P33272 | X63545 | S48369 | X63545 | |
| *Cyp2b13* | mou | | | | K02409* | M60352 → | M60358 | |
| *CYP2B14X* | rat | | | | | | | |
| *CYP2B15* | rat | | | | D17343 → | D17349 | | |
| *CYP2B16P* | rat | | | | | | | |
| *CYP2B17* | mon | JT0676 | | | | | | |
| *CYP2B18* | gpi | | | | | | | |
| *CYP2B* | gpi | A36154* | S15135* | P34033* | | | | |
| | | S28205* | | | | | | |
| *Cyp2b* | mou | A21630* | | | | | | |
| *Cyp2b* | mou | | | | M60359* | | | |
| *CYP2C1* | rab | A00181 | A34257* | P00180 | K01522 | M74199 | M76597 | D26152 |
| *CYP2C2* | rab | A00182 | A27718 | P00181 | K01521 | M14955 | M19137 | M76596 |
| | | S15587* | | | | | | |
| *CYP2C3* | rab | A00183 | A22606 | P00182 | J02901 | M31245 → | M31254 | K01523 |
| | | A34534 | | | | | | |
| *CYP2C4* | rab | A26731 | B34257* | P11371 | J02716 | M74200 → | M74203 | |
| *CYP2C5* | rab | A00180 | A37828 | P00179 | J05575 | M11299 | M74204 → | M74206 |
| | | S20227 | S16715 | | M17026 | M55664 | | |
| *CYP2C6* | rat | A25954 | B25585 | P05178 | J04466 | K03501 | M13711 | |
| | | A28516 | S00955* | | M24237 → | M24239 | | |
| | | A41425* | | | M36848 | X06712* | M14776 | |
| *CYP2C6P* | rat | | | | J03509 | M18774 | M18336 | |
| *CYP2C7* | rat | A25585 | B25954 | P05179 | M14775 | M18774 | M31031 | N00038 |
| | | B28516 | A60563* | | M18335 | X12595 | J03509 | |
| | | S24582* | A27425* | | | | | |