Changes and Edits

Several changes and edits have been made compared to the original allele annotations in the Human Cytochrome P450 (*CYP*) Allele Nomenclature Database (see archive) to standardize annotations across genes and correct errors. A summary of all changes and edits is provided in **Tables 1, 2 and 3**.

As shown in **Table 1**, numerous changes have been made during the curation process. Limiting the upstream and downstream regions used for allele definitions, as well as the removal of introns of unknown functional consequence, caused the retirement of several suballeles or merging of suballeles as detailed in the Table.

***CYP3A4*1G*:** The defining variant of this allele is **c.1026+12G>A (rs2242480)** in intron 10 (this variant is also part of several other star alleles). The *CYP3A4*1G* allele has been associated with increased CYP3A4 activity, while others reported associations with decreased activity. It is unclear though from published evidence, if c.1026+12G>A is indeed the underlying cause. Observed altered activity might be caused by another single nucleotide variant(s) (SNVs) that is in linkage with c.1026+12G>A or be confounded by activity encoded by the other genes in the *CYP3A* locus. Although all other intronic SNVs of unknown function have been removed from allele definitions, **PharmVar continues to display *CYP3A4*1G* under its 'legacy' name, as well as displays c.1026+12G>A in other haplotypes, until additional evidence is available**. If new information supports c.1026+12G>A being a function altering SNV, it will receive core SNV status, and consequently, **1G* will receive its own star number. Otherwise, *CYP3A4*1G* will be retired and c.1026+12G>A removed from all allele definitions.

c.-392G>A (rs2740574): The SNV at c.-392 was defined as A>G in the past (using the now discontinued reference sequence AF280107.01). Since this SNV is defined as G>A on the current RefSeq (NG_008421.1), all allele definitions were updated accordingly. Specifically, all alleles with the c.-392A>G SNV now match the RefSeq and are thus no longer showing the variant, while all other alleles gained c.-392G>A. Note that the 'legacy nomenclature' showed *CYP3A4*1A* without a SNV and *CYP3A4*1B* having c.-392A>G. Now, *CYP3A4*1B* corresponds to *CYP3A4*1.001* (having a G at c.-392) and *CYP3A4*1A* allele corresponds to *CYP3A4*1.002* (having an A at c.-392). Legacy names are cross-referenced with their new PharmVar names on the gene page.

Missing information for c. 522-191C>T (rs35599367, core SNV of *CYP3A4*22*) and/or c.1026+12G>A (rs2242480, defining *CYP3A4*1G*): For many alleles there is no published information for these intronic SNVs; such alleles received an evidence level of LIM. An allele received an evidence level of MOD, if linkage data obtained from the 1000 Genomes Project data strongly supported the absence or presence of these intronic SNV(s). For example, *CYP3A4*4* received an evidence level of MOD because the two intronic SNVs were not found on any sample in the 1000 (1K) Genomes Project that was positive for the **4* core SNV. Alleles not represented in the 1K Genomes Project received an evidence level of LIM.

All future submissions must provide information for c.522-191C>T and c.1026+12G>A. PharmVar also encourages submissions for existing allele definitions to confirm the status of these SNVs.

Table 1 Changes made for transition into the PharmVar database

Archived legacy page	change/edit	Transition into the PharmVar db
The reference sequence was updated to NG_008421.1 which caused a shift in positions compared to AF280107.01 the reference sequence used on the legacy page, i.e. all SNVs with the exception of those in exon 1 are shifted when counting from the A of the ATG start codon.		
1J, *1N, *1P, *1Q, *1R, *1S	removed	These suballeles were defined by intronic SNVs only; these intronic SNVs are of unknown functional consequence
1F, *1K, *1L	removed	These suballeles were defined only by upstream SNVs outside the region used for star allele definition by PharmVar (1 kb upstream of the ATG start codon)
1C, *1D	removed	These suballeles were defined by upstream region only; no information for their respective coding regions
1H	merged	**1H* merged with **1G* after removing a SNV in the downstream region that is outside the region used for star allele definition by PharmVar (250 bp downstream of the translation stop codon)
c.-392G>A	revised	See note above. This SNV was revised from c.-392A>G to c.-392G>A; thus *CYP3A4*1B, *23, *24* are no longer showing the c.-392A>G variant
15B	merged	**15B* merged with **15A*. After an upstream SNV was removed this suballele no longer differed from **15A*
16B	merged	**16B* merged with **16A.* There was no information supporting the absence of c.1026+12G>A (rs2242480) for **16A* and status of c.-392G>A was inconclusive for both
18B	merged	**18B* merged with **18A.* There was no information supporting the absence of c.1026+12G>A (rs2242480) for **18A*

To ensure consistent display across all genes PharmVar uses a number of conventions for storing and displaying allelic data. See the **Standards Document** for more information. Note that PharmVar is using the **3' Rule** (or left-alignment). Both alleles with indels were aligned according to this rule.

Changes and edits made after the gene has been transitioned into the PharmVar databases are listed in **Table 2**.

Table 2 Changes made after the transition into the PharmVar database

Allele	change/edit	After transition to the PharmVar db

Alleles accepted and/or released by PharmVar

Table 3 provides a summary of allele definitions that have been accepted and/or released by PharmVar and the submitting author.

Table 3 Alleles accepted and/or released by PharmVar

CYP3A4 allele	PharmVar submission ID	acceptance date	release date	db version	submitted by	PharmVar ID	notes
***12.001**	01_seq1 02_seq11	01-19-2021 02-16-2021	04-30-2021	4.2.7	Drögemöller	PV01516	confirmatory
***15.001**	01_seq2 02_seq12	01-19-2021 02-16-2021	04-30-2021	4.2.7	Drögemöller	PV01527	confirmatory
***1G** (*X.001)[1]	01_seq3	01-19-2021	hold	tbd	Drögemöller	tbd	novel suballele
***1G** (*X.002)[1]	01_seq4	01-19-2021	hold	tbd	Drögemöller	tbd	confirmatory
Note that submission 01 utilized sequence data published earlier (Drögemöller et al 2013, PMID 23423246) to complement published data; however, sequence data did not cover c. 522-191C>T, therefore the evidence level is Moderate (Mod)							
***1.002**	02_seq1	02-16-2021	04-30-2021	4.2.7	Gaedigk et al	PV01515	confirmatory
***2.001**	02_seq2	02-16-2021	04-30-2021	4.2.7	Gaedigk et al	PV01500	confirmatory
***1G** (*X.002)[1]	02_seq3	02-16-2021	hold	tbd	Gaedigk et al	PV01498	confirmatory
***3.001**	02_seq4	02-16-2021	04-30-2021	4.2.7	Gaedigk et al	PV01523	confirmatory
***35.001**	02_seq5	02-16-2021	04-30-2021	4.2.7	Gaedigk et al	PV01492	novel allele
***14.001**	02_seq6	02-16-2021	04-30-2021	4.2.7	Gaedigk et al	PV01512	confirmatory
***1.006**	02_seq7	02-16-2021	04-30-2021	4.2.7	Gaedigk et al	PV01524	novel suballele
***1G** (*X.001)[1]	02_seq8	02-16-2021	hold	tbd	Gaedigk et al	tbd	confirmatory
***1G** (*X.003)[1]	02_seq9	02-16-2021	hold	tbd	Gaedigk et al	tbd	novel suballele
***16.001**	02_seq10	02-16-2021	04-30-2021	4.2.7	Gaedigk et al	PV01693	confirmatory
***22.001**	02_seq13	02-16-2021	04-30-2021	4.2.7	Gaedigk et al	PV01521	confirmatory

CYP3A4 allele	PharmVar submission ID	acceptance date	release date	db version	submitted by	PharmVar ID	notes
***1.001**	02_seq14	02-16-2021	04-30-2021	4.2.7	Gaedigk et al	PV01494	confirmatory
***10.001**	02_seq15	02-16-2021	04-30-2021	4.2.7	Gaedigk et al	PV01519	confirmatory
***25**	pending				Rodriguez-Antona et al		
***27**	pending				Rodriguez-Antona et al		

[1] **X* denotes variants of the **1G* allele; these alleles will be released if **1G* receives its own star allele designation