Important Information

CYP2D6 was transitioned from the original 'Human Cytochrome P450 Nomenclature Database' website (www.cypalleles.ki.se/) to PharmVar on Sept 26, 2017. **A PharmVar expert panel curates this gene.**

Several star alleles may not have been sequenced for intronic flanking regions, may have very limited information for up- and downstream sequences, and exons may not have been sequenced but screened with other methods. Also, the haplotype may have been inferred. Such star alleles have been annotated with a limited or moderate evidence level as they may have additional variants that have not been captured when the allele was first defined.

Gene Region Mapped

PharmVar defines haplotypes (star alleles) for *CYP2D6* covering the following regions: 5' flanking sequence, the 5' UTR, all nine exons, all eight introns, the 3' UTR, and the 3' flanking sequence. **Table 1** specifies the positions within the genomic reference sequence NG_008376.4 that delineate the beginning (5' limit) and end (3' limit) of the sequence used to assign *CYP2D6* star alleles.

Single nucleotide variants (SNVs) include single nucleotide changes, as well as small insertions and deletions. All are collectively referred to as SNVs.

Submissions must cover all regions specified in **Table 1** and list all SNVs present within the 5' and 3' limits, with one exception that is referred to as a 'gap' in the upstream region (more details on page 2 below).

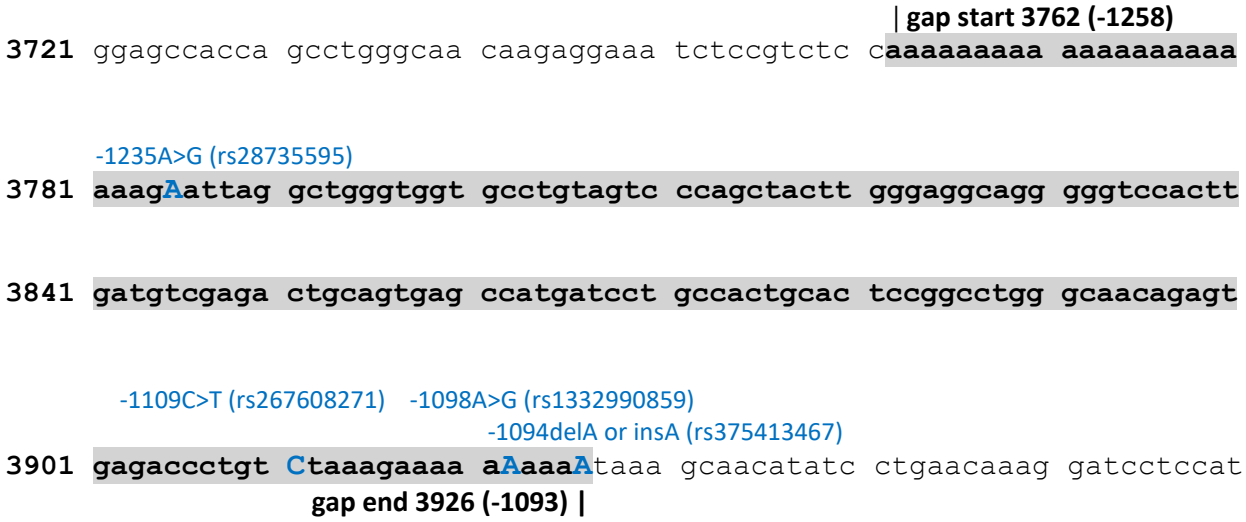
Table 1 Regions required for star allele characterization. The 5' and 3' limits denote the first and last bases for which sequence information must be provided

RefSeq	**5' limit** Counting from the **sequence start**	**5' limit** Counting from the **ATG translation start**	**3' limit** Counting from the **sequence start**	**3' limit** Counting from the **ATG translation start**
NG_008376.4	3436	-1584	9501	4482
	Gap region: 3762 (-1258) – 3926 (-1093) Variants in this region are not part of star allele definitions			
All SNVs within the 5' and 3' limits, with the exception of the gap region, must be submitted for haplotype (star allele) definitions. This includes 5' flanking sequence, the 5' UTR, exons 1-9, introns 1-8, the 3' UTR, and the 3' flanking sequence.				

Upstream 'gap region': Two homopolymer strings of 'A' nucleotides frame a 165bp-long segment in the upstream region which is notoriously difficult to resolve by Sanger sequencing and may also cause difficulties for other sequencing methods. It is therefore not only difficult to determine the exact number of A's in the homopolymer strings (which are both reported to have varying numbers of A's) but also to accurately determine the presence of SNVs adjacent to, within, or between the homopolymer strings (i.e., NG_008376.4 positions 3785A>G (-1235A>G), 3911C>T (-1109C>T), 3922A>G (-1098A>G), 3926insA and 3926delA (-1094insA and -1094delA). To the best of current knowledge, variants within this region do not impact CYP2D6 function.

Due to uncertainty around these SNVs, in particular -1235A>G, and the absence of data for many of the star alleles, the consensus among the experts was to exclude this region when defining star alleles. Therefore, the above-mentioned variants were removed from all star alleles on October 23, 2024 (v.6.1.6). See the Change Log document for additional details.

Figure 1 Upstream gap region excluded from all star allele definitions
The gap region is shown in bold with gray background. Start and stop positions are as indicated; SNVs within the region that were removed from allele definitions are shown in blue font



Distant enhancer SNV: There is mixed evidence (PMIDs 23985325, 25381333, and 28745555) that a distant 'enhancer' SNV (rs5758550) located about 115 kb downstream of the *CYP2D6* gene locus impacts expression levels and no convincing evidence that rs5758550 indeed causes a clinically meaningful change in CYP2D6 function (PMIDs 28745555 and 34716917). Although it is possible to experimentally link rs5758550 with the known haplotypes (PMID 32457600). However, star allele definitions do not include rs5758550 due to the lack of evidence supporting its clinical utility.

Coordinates

Variant coordinates in the PharmVar database are available for genomic and transcript reference sequences (RefSeq), as well as the Human Genome Assemblies GRCh37 and GRCh38. **Table 2** summarizes important information regarding these sequences.

In this document, coordinates are counted from the 'A' of the ATG translation start codon using the genomic NG_008376.4 sequence as reference unless indicated otherwise.

Table 2 Reference Sequences Used by PharmVar for Variant Mapping

Sequence	Notes
CYP2D6	Encoded on the negative strand.
NG_008376.4	This genomic RefSeq matches the *CYP2D6*1.001* allele. **NG_008376.4 is used as a reference to define allelic variation (star alleles). PharmVar submissions must be annotated using this RefSeq.**
LRG_303	NG_008376.4 matches the Locus Reference Genomic (LRG) sequence LRG_303.
NM_000106.6	The transcript RefSeq matches *CYP2D6*1.001*. NM_000106.6 represents the MANE Select transcript that best represents gene function, is expressed, highly conserved, and is well-supported by experimental evidence. MANE select transcripts match GRCh38 and are identical to their genomic reference sequences for the 5'-UTR, coding sequence, splicing, and 3'-UTR.
GRCh37	The *CYP2D6* sequence in GRCh37 matches *CYP2D6*2*, which has multiple variants compared to the *CYP2D6*1* reference allele. These differences must be accounted for when using GRCh37 for star allele calling. See page 6 for additional details.
GRCh38	The *CYP2D6* sequence matches the NG_008376.4 RefSeq.
NG_008376.3	PharmVar no longer uses this RefSeq. NG_008376.3 is identical to AY545216 and matches the *CYP2D6*1.001* allele. NG_008376.3 is 819 bp shorter on the 5'end and 1540 bp on the 3' end compared to NG_008376.4; but is otherwise identical.
M33388	M33388 was used as reference sequence by the CYP450 Nomenclature database; it contains sequencing errors (601insC, 1289-90GC>CG, 1328delG and 1440delC) that shift position numbers when compared to NG_008376.4 (e.g., the SNV defining the *CYP2D6*4* allele is at position 1846 according to M33388, but 1847 on NG_008376.4).

PharmVar Feature: Obtain coordinates from the sequence start or the ATG start codon with one click! The default setting is NG_008376.4, counting the 'A' in the ATG start codon as +1. The ⓘ buttons link out to NCBI sequences.

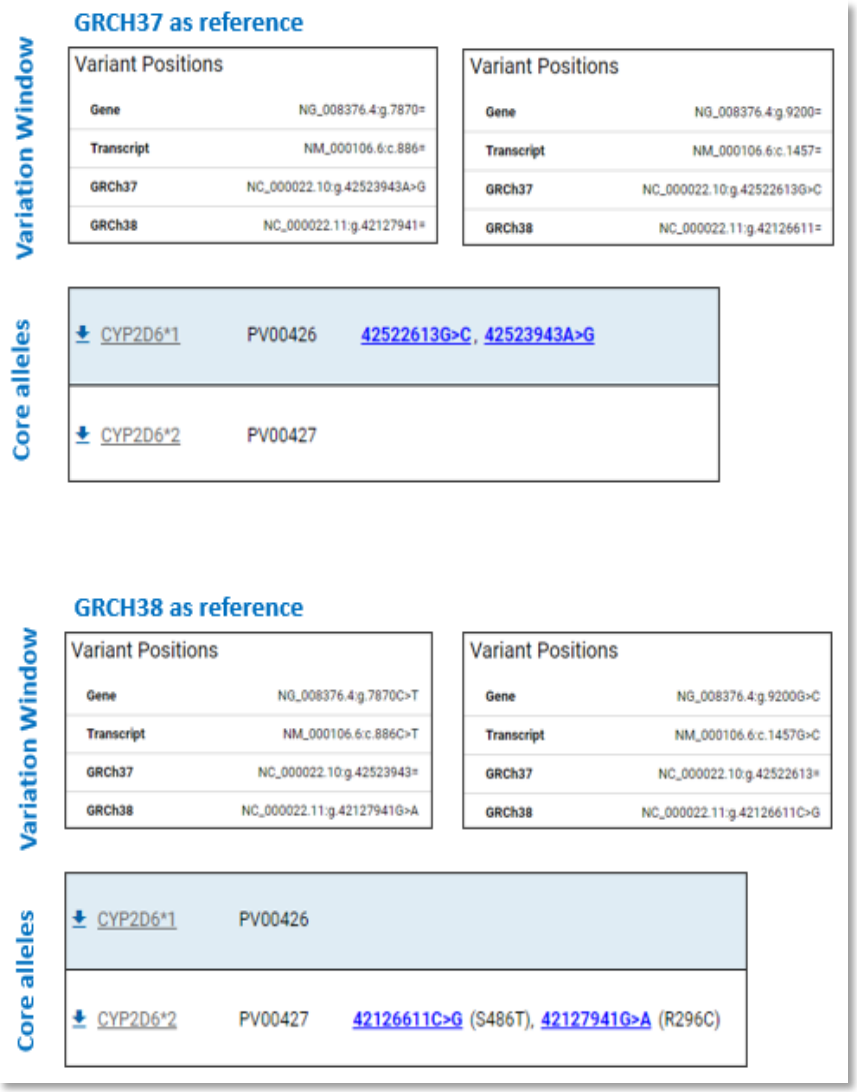
Comparisons with GRCh37

When using GRCh37 as the RefSeq, all allele definitions differ from GRCh37 regarding the number of 'A' nucleotides in the upstream region (rs267608321). The NG_007938.4 RefSeq has 22 A's while GRCh37 has 27 A's. Because *CYP2D6* is encoded on the negative chromosomal strand, this difference is displayed as g.42528030delTTTTT for GRCh37. PharmVar does not capture the 'A' string variation among haplotypes due to difficulties sequencing this region. Therefore, g.42528030delTTTTT is shown for every suballele when using GRCh37 as the reference sequence. The *CYP2D6* sequence in GRCh37 otherwise matches the *CYP2D6*2.001* allele.

Furthermore, alleles having 2851C>T (p.R296C) and 4181G>C (p.S486T) in their core allele definitions in the NG_008376.4 reference do not have these variants in the GRCh37 reference. In contrast, alleles lacking these variants in their core allele definitions when using NG_008376.4 as reference gain both variants when GRCh37 is used as the reference sequence, including *CYP2D6*1*. Since *CYP2D6* is encoded on the negative chromosomal strand, 2851C>T is shown as g.42523943A>G, but 4181G>C is shown as g.42522613G>C on GRCh37. Finally, neither variant is annotated as having an amino acid change on GRCh37, as these 'variants' encode their respective reference amino acids (p.C296 and p.T486, respectively).

As shown in **Figure 2,** 2851C>T and 4181G>C are called variants when using GRCh38 as reference but not when using GRCh37.

Figure 2 Comparisons with GRCh37



Sequence Variants

Sequence variants detailed in the PharmVar database comprise single nucleotide polymorphisms (SNPs) and insertions and deletions of single or multiple nucleotides, which are collectively referred to single nucleotide variants (SNVs) or simply 'variants'.

Some variants only occur on a single star allele, e.g., 2550delA is only found on *CYP2D6*3*. Others may occur on two or more star alleles, e.g., 2851C>T is found on **2, *8, *11, *17*, and over 60 other alleles, and 1022C>T is found on **17, *40, *58, *64, *141,* and **154*.

The PharmVar Variant Window

Click on a variant to open this window. The example in **Figure 3** shows the Variant Window for the *CYP2D6*3* core variant 2550delA (c.775del).

This view provides SNV positions across all reference sequences, links to external resources including dbSNP (note that some SNVs do not have an rsID), ClinPGx (formerly PharmGKB) and gnomAD, and frequency information. There is also a **button to display all haplotypes (star alleles) that contain the selected variant**.

The **top portion** of the variant window **displays variant coordinates according to the Human Gene Variation Society (HGVS)** nomenclature at the gene, transcript, and genome levels. Coordinates are shown as obtained through the NCBI Variation Services and only manually curated if no results are returned. Additional details on HGVS nomenclature can be found here.

The **middle portion** of the variant window **displays SNV positions PharmVar-style** at the gene, transcript, and genome levels, providing reference and variant nucleotides and positions for both count modes.

Figure 3 Variant Window



PharmVar consistently uses the 3' rule (left alignment) to determine the positions of deleted or inserted nucleotide(s). HGVS and PharmVar-style positions may differ for such variants which is most likely due to differences in sequence alignments. Also, PharmVar displays single nucleotide insertions as 'ins' while HGVS displays 'dup' if a base is duplicated. Additional details and examples are provided under Standards.

The **bottom portion** of the variant window provides a filter option to display all the star alleles with the selected SNV, a link to dbSNP (if the variant has an rsID), as well as links to ClinPGx and gnomAD that provide additional external resources for the variant, including frequency.

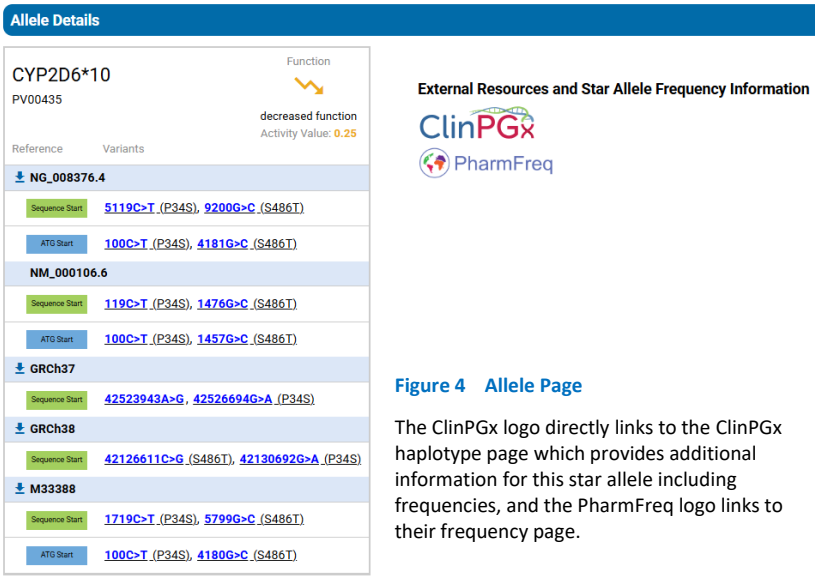


Figure 4 Allele Page

The ClinPGx logo directly links to the ClinPGx haplotype page which provides additional information for this star allele including frequencies, and the PharmFreq logo links to their frequency page.

The PharmVar Allele Page

Click on an allele name to open this page. **Figure 4** shows the *CYP2D6*10* Allele Page detailing the definition of this core allele using the genomic and transcript reference sequences and GRCh37 and GRCh38 genome builds.

Core Allele Definitions

Some alleles share 'core' defining sequence variants (see Allele designation Criteria and Evidence Level for details). Therefore, similar alleles are grouped under the same core allele definition (same star number). Although allele information may be valuable for test design (sequence or SNV panel-based) or the interpretation of test results, the distinction of alleles with the same star number (also referred to as suballeles) is not considered necessary for phenotype prediction because **all alleles under a star number are assumed to be functionally equivalent based on current literature**. Thus, even if a test can distinguish between alleles of the same star number (suballeles), these can generally be reported as, for example, *CYP2D6*2* or **4* without the provision of their numeric 'dot' extension.

PharmVar has developed core allele definitions in collaboration with ClinPGx. **Only sequence variations that change an amino acid, impact function by altering expression levels, or interfere with splicing and are present in *all* alleles within an allele group are part of the core allele**

definition. A core variant is a sequence variation that is part of a core allele definition. Core SNVs are highlighted in suballeles with the PharmVar logo for easy identification. Core alleles may contain a single core variant or a combination of two or more core variants. All alleles, including their core variant definitions, have unique PharmVar identifiers (PVID).

This rule-based system collapses all alleles of a group into a single 'core' allele definition. Thus, the core allele represents all alleles categorized under a star (*) number.

Core allele definitions are highlighted by a grey background on the *CYP2D6* gene page 'Table View'. Suballeles are labeled as, e.g., *CYP2D6*2.001*, *CYP2D6*2.002,* etc*.,* and are shown under their core allele definition with alternating white and blue backgrounds.

CPIC Clinical Allele Function is shown in the core allele bar, representing all alleles listed under that core allele definition.

Figure 5 shows the core allele definition of *CYP2D6*2* and selected sub alleles. Of all the SNVs found in this allele group, only two fulfill the abovementioned rules and are shared among all the alleles, namely 2851C>T (p.R296C) and 4181G>C (p.S486T). Therefore, these two core variants constitute the *CYP2D6*2* core allele.

Figure 5 The *CYP2D6*2* core allele and selected suballeles

CYP2D6*2		PV00427	2851C>T (rs16947, R296C) 4181G>C (rs1135840, S486T)		CPIC Clinical Function →
CYP2D6*2.001	CYP2D6*2A	PV02536	-1584C>G (rs1080985) -740C>T (rs28624811) -678G>A (rs28633410) 214G>C (rs1080995) 221C>A (rs1080996) 223C>G (rs1080997) 227T>C (rs1080998) 232G>C (rs29001518) 233A>C (rs28695233) 245A>G (rs1081000) 310G>T (rs28371699) 745C>G (rs28371701) 842T>G (rs28371702) 1662G>C (rs1058164) 2851C>T (rs16947, R296C) 3385A>C (rs1985842) 3585G>A (rs28371730) 3791C>T (rs4987144) 4181G>C (rs1135840, S486T) 4482G>A (rs12169962)	Def	Johansson et al. 1993 Panserat et al. 1994 Raimundo et al. 2000 Sakuyama et al. 2008 deposited by Gaedigk et al 2018 deposited by Nofziger 2018
CYP2D6*2.002	CYP2D6*2B	PV00150	1038C>T (rs1081003) 1662G>C (rs1058164) 2851C>T (rs16947, R296C) 4181G>C (rs1135840, S486T)	Lim	Marez et al. 1997
CYP2D6*2.003	CYP2D6*2C	PV00149	1662G>C (rs1058164) 2471T>C (rs17002852) 2851C>T (rs16947, R296C) 4181G>C (rs1135840, S486T)	Lim	Marez et al. 1997
CYP2D6*2.004	CYP2D6*2D	PV03050	-1584C>G (rs1080985) -740C>T (rs28624811) -678G>A (rs28633410) 214G>C (rs1080995) 221C>A (rs1080996) 223C>G (rs1080997) 227T>C (rs1080998) 232G>C (rs29001518) 233A>C (rs28695233) 245A>G (rs1081000) 310G>T (rs28371699) 745C>G (rs28371701) 842T>G (rs28371702) 2851C>T (rs16947, R296C) 3385A>C (rs1985842) 3585G>A (rs28371730) 3791C>T (rs4987144) 4181G>C (rs1135840, S486T) 4482G>A (rs12169962)	Mod	Marez et al. 1997 Hitchman et al. 2026

As shown in **Figure 6** only the 2550delA frameshift is shared among the *CYP2D6*3* suballeles. Therefore, this SNV is only one core variant defining the **3* allele.

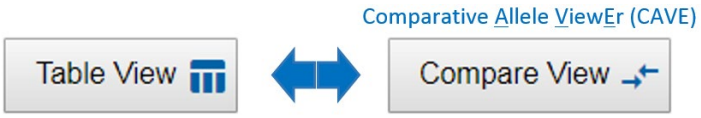
Figure 6 The *CYP2D6*3* core allele and selected suballeles

CYP2D6*3		PV00428	2550delA (rs35742686, R259fs)		CPIC Clinical Function ×
CYP2D6*3.001	CYP2D6*3A	PV00221	2550delA (rs35742686, R259fs)	Def	Kagimoto et al 1990 Scantamburlo et al. 2017 Liau et al. 2019
CYP2D6*3.002	CYP2D6*3B	PV00220	1750A>G (rs1135824, N166D) 2550delA (rs35742686, R259fs)	Lim	Marez et al. 1997
CYP2D6*3.003		PV02310	198G>A (rs201475960) 2550delA (rs35742686, R259fs)	Def	deposited by Nofziger et al

CAVE - the PharmVar Comparative Allele ViewEr

PharmVar has developed the CAVE tool to compare core alleles, visualize which core SNVs are present in alleles of interest and to identify SNVs that are unique to the selected alleles.

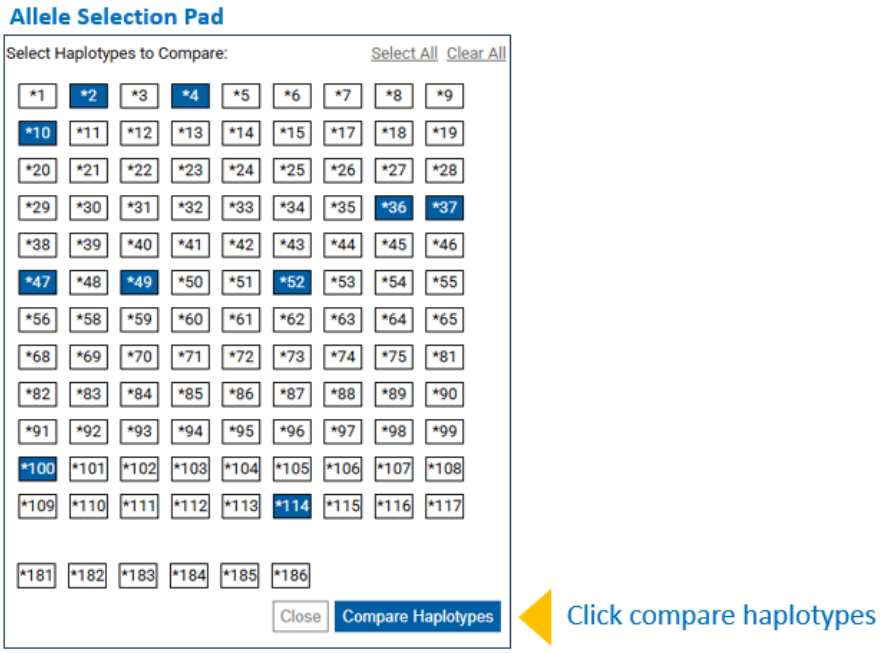
To access CAVE, switch from Table View to Compare View



A sequence variation that is part of a core allele may be unique or occur in two or more core alleles. It may, however, also be found in some allele(s) grouped under another star number. In some instances, a non-core SNV may be unique and therefore useful for identifying an allele of interest or discriminating it from others. Because the CAVE tool uses core allele definitions, unique SNVs that are not part of any core alleles are not displayed.

CAVE comparisons can be downloaded into an Excel spreadsheet using the 'Download Comparision table' button.

The 100C>T SNV is a classic example of an SNV being part of multiple core allele definitions, including *CYP2D6*10*, **36*, **49*, **100*, **114*, and many others. Also, as detailed below, it is present on most but not all *CYP2D6*4* suballeles.



Select core alleles for comparison in the **selection pad**. To illustrate the example given above, ten alleles were selected, of which eight have 100C>T in their core allele definitions, i.e., **10*, **36*, **37*, **47*, **49*, **52*, **100*, and **114*.

CAVE graphically displays the core SNVs that are found in the selected alleles along their rsIDs and impact (amino acid change, splice defect or expression).

PharmVar no longer provides variant-level function annotations in the CAVE tool.
(The 'variant alters function' symbol has been removed from all genes on July 30, 2026).
Instead, CPIC clinical allele function is now provided.

Comparative Allele ViewEr (CAVE) using NG_008376.4 as reference

Variant is present | Variant is present on some suballeles | Variant is unique

CPIC Clinical Function

Normal Function | No Function | Increased Function | Decreased Function | N/A Function Not Assigned | Uncertain Function | Unknown Function

Download Comparison Table

Variant Impact		*2	*4	*10	*36	*37	*47	*49	*52	*100	*114
73C>T (rs267608313)	R25W						■				
100C>T (rs1065852)	P34S		▒	■	■	■	■	■	■	■	■
1612T>A (rs1135822)	F120I				▒			■			
1759G>A (rs5030865)	G169R										■
1847G>A (rs3892097)	splice defect		■								
1944G>A (rs745365204)	R201H					■					
2829delC (rs267608279)	S288fs									■	
2851C>T (rs16947)	R296C	■	▒								■
3878G>A (rs28371733)	E418K								■		
Ex9-conv	P469A, T470A, H478S, G479R, F481V, A482S		▒		■						
4181G>C (rs1135840)	S486T	■	▒	■	■	■	■	■	■	■	■

- 100C>T (highlighted by orange frame) is shown by **blue** boxes for **10*, **36*, **37, *49, *52, *100* and **114*, indicating that it is part of their core allele definitions. This SNV is shown by a **gray** box for **4* to indicate that it is not found on all **4* suballeles. Since 100C>T is not unique to **10*, this allele cannot be identified unequivocally by genotyping solely for 100C>T; the presence of other alleles with 100C>T must be ruled out by testing for their respective core SNVs.
- Similarly, 4181G>C (highlighted by yellow frame) is also present on **10*, **36*, **37, *49, *52, *100* and **114* but not all **4* suballeles. Other variants are also found on two or more star alleles.
- As the graph illustrates, some variants are unique to a star allele as indicated by the PharmVar symbol while other are found on two or more star alelles.

Variant Groupings

This feature allows PharmVar to display amino acid changes that are caused by two SNVs or a series of SNVs within a defined region that cause multiple amino acid changes.

For example, 1660G>A causes a valine to methionine change at amino acid position 136 in *CYP2D6*107* and is displayed as '1660G>A (V136M)' on the *CYP2D6* page. However, if both 1660G>A and 1662G>C are present, the amino acid is changed to an isoleucine. The variant grouping displays the combined effect as '1660G>A+1662G>C (V136I)', which is part of the *CYP2D6*29* and **70* alleles.

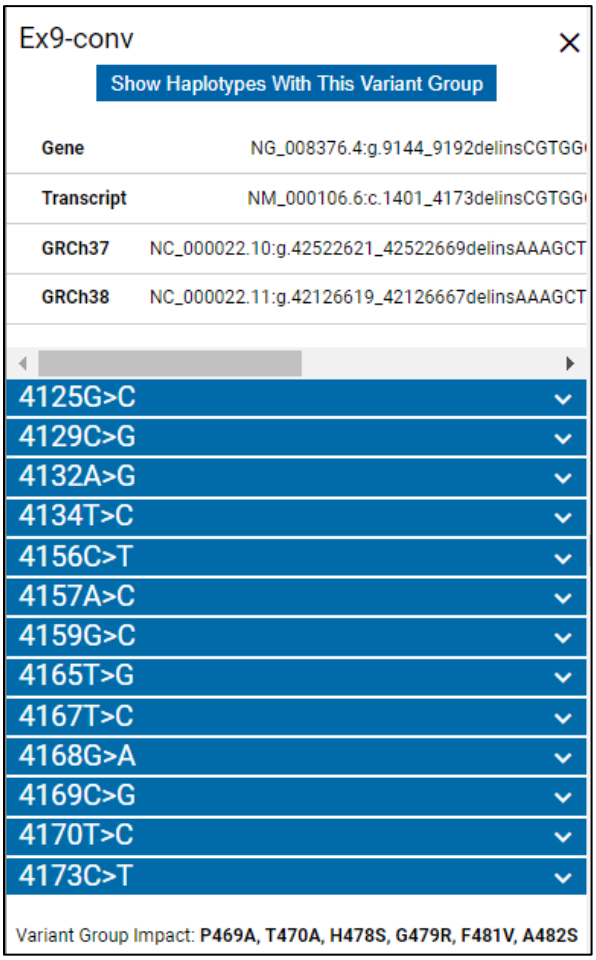
Another example is a *CYP2D7*-derived sequence in exon 9 that affects 13 nucleotides and leads to six amino acid changes which is known as the **exon 9 conversion or 'E9-conv'** for short as depicted in the following graphical representation.



Differences between *CYP2D6* and the *CYP2D7*-drived sequences are shown in red. The first and last positions of the E9-conv correspond to 4125 and 4173 when counting from the ATG start codon, and 9144 and 9192 when counting from the sequence start of NG_008376.4.

The Ex9-conv is displayed for core alleles and suballeles as:





In the **Variant Window**, the Ex9-conv grouping expands to list all SNVs, which can be expanded further for additional information.

Note that not every variant in the Ex9-conv contributes to an amino acid change: two SNVs must be present to cause the p.H478S and p.A482S amino acid changes, respectively.

Allele Function

PharmVar displays 'CPIC clinical allele function' as determined by the Clinical Pharmacogenetic Implementation Consortium (CPIC) in the Table view and in the graphical CAVE displays.

The *CYP2D6* allele functionality table is available at ClinPGx (https://www.clinpgx.org/page/cyp2d6RefMaterials). This table has been created as part of the CPIC guideline process (Pharmacogene Curation SOP) and is periodically updated.

Variant and Allele Frequencies

Variant frequencies: PharmVar displays the frequency of a sequence variation in the **Variant Window** using gnomAD as resource. The frequency provided represents the global frequency of the SNV which may considerably vary across populations.

The frequency of a variant can also be obtained through the external links to dbSNP, ClinPGx, and gnomAD (see **Figure 1**). The ClinPGx link directs the user to population specific SNV frequency

information with 'one click'. This is a convenient way to access other *CYP2D6* information within ClinPGx.

Importantly, **the frequency of a variant does not reflect the frequency of a haplotype (or star allele) if the variant is part of two or more haplotypes**. Thus, the frequency of 100C>T, as shown in **Figure 3**, represents the aggregate frequency of all haplotypes (star alleles) with this variant.

Allele frequencies: Available allele frequency data are summarized on the ClinPGx haplotype page for several sources and populations. The ClinPGx haplotype page can be accessed from PharmVar through its **Allele Page** as described on page 6 and shown in **Figure 4**. Allele frequency data provided by PharmFreq can now also be accessed through the **Allele Page.**

Data may not be available for each star allele for each of the data resources.

CYP2D6 Structural Variation Document

This document provides detailed information regarding structural variation, including gene deletions and duplications, as well as conversions and structural rearrangements between *CYP2D6* and its pseudogenes.

Important: Structural variants, with the exception of gene duplications, are listed on the PharmVar *CYP2D6* gene page and are described in this document.

References

The references provided in the PharmVar database include the citation in which an allele was first published. Additional reference(s) support the haplotype or updates/revisions. The reference list is not intended to provide a complete bibliography for an allele.

Haplotypes submitted to PharmVar but not published elsewhere are listed as 'deposited by'. Additional submitter information is provided in the Change Log document.

Changes and Corrections

Changes and edits are logged in the **CHANGE Log** document.