Important Information

***CYP2C19* has been transitioned from the original "Nomenclature Website" (www.cypalleles.ki.se/) to PharmVar on Sept 26, 2017 (the last posted version of the *CYP2C19* page is available through the 'Archive' link). Current gene information can be accessed through the 'genes' tab on the 'menu' bar.**

A number of allelic variants may not have been sequenced for intronic flanking regions, some variants may have very limited information for up and downstream sequences and exons may not have been sequenced, but screened with other methods. Also, for many alleles, the defined haplotype is inferred. Such alleles may carry single nucleotide variants (SNVs) in exons, flanking intronic and/or up- and downstream regions that have not been captured.

Gene Region Mapped

Allele definitions contain variants positioned between -3402 and 90033 of the NG_008384.3 RefSeq. Note that new submissions will need to include sequence covering a minimum of ~2 kb of upstream and 250 bp of downstream region. Intronic variants are not shown as part of haplotype definitions unless the variant impacts function.

Coordinates

Coordinates in the PharmVar database are available in reference to a number of sequences including Human Genome Assemblies GRCh37 and GRCh38, NG_008384.3 and NM_000769.1. The following table summarizes important information regarding differences among respective sequences.

In this ReadMe document coordinates are counted from the "A" in the ATG translation start codon as +1 for NG_008384.3 unless otherwise stated.

PharmVar feature: No more counting. Get coordinates from the sequence start or the ATG start codon with one click! The default setting is NG_008384.3 counting the "A" in the ATG start codon as +1.

Table 1

sequence	notes
CYP2C19	Encoded on the positive strand
NG_008384.3	The *CYP2C19* Reference Sequence (RefSeq) matches the *CYP2C19*1.001* reference allele. PharmVar submissions must be annotated using the RefSeq. This RefSeq represents the 80161A while the vast majority of alleles, including some *1 suballeles have 80161G (I331V). The gene expert panel recommended to keep *1 allele definitions 'as is' (i.e. accept legacy definitions) although this poses an exception to PharmVar allele designation criteria.
LRG_585 (pending)	Locus Reference Genomic (LRG) sequence. LRGs never change once defined.
NM_000769.1	transcript RefSeq
GRCh37	GRCH37 differs from GRCh38 in three positions, 96521422A>G, 96522561T>C and 96602623G>A. Therefore, the number of SNPs displayed in GRCh37 differs from those shown for GRCh38. This sequence corresponds to the *CYP2C19*27* allele definition.
GRCh38	The *CYP2C19* RefSeq sequence matches the sequence in GRCh38; this sequence corresponds to the *CYP2C19*1A* allele definition.

Sequence variants

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

Some SNVs are unique to an allele and only occur on a single haplotype, e.g. 19154G>A is unique to the **2* haplotype. Other SNVs can occur on more than one allele, e.g. -806C>T is found on **17* and **4B*. While it increases **17* expression levels, this SNP has no impact on **4B* which is nonfunctional due to 1A>G, which is only found in the **4* haplotypes.

PharmVar feature: By clicking on a SNV, the variation view window will slide in. This view provides SNV positions across all sequences in both count modes and the link to the NCBI dbSNP identifier (rs number; note that some SNVs may not have been allocated a dbSNP identifier). There is also a bar providing the option to display all haplotypes with the selected variant.

Core allele definitions

For many star alleles there are a growing number of so-called suballeles all sharing one or more 'key' defining sequence variant(s) (see ALLELE DESIGNATION AND EVIDENCE LEVEL CRITERIA document for details). Although suballele information is valuable for e.g. the design (sequence or SNP-panel based) and interpretation of test results, the distinction of suballeles is not necessary for phenotype prediction because **all alleles under a star number are believed to be functionally equal**. Thus, even if a test is capable of distinguishing suballeles, these are generally simply reported as e.g. *CYP2C19*2* or **4, etc*.

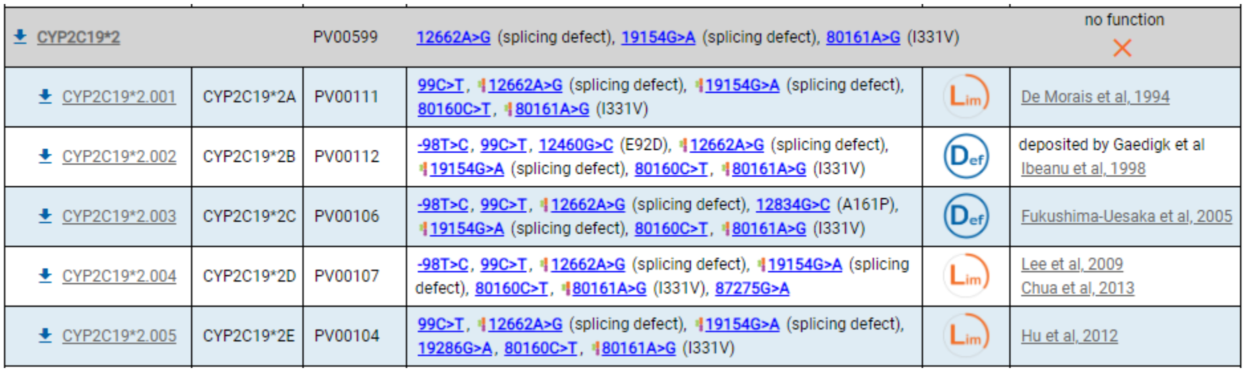
PharmVar has developed core allele definitions in collaboration with the PharmGKB. Only sequence variations which change an amino acid or impact function by changing expression levels or interfere with splicing and are present in **all** suballeles within an allele group, are part of the **core allele definition**. **With this rule-based system suballeles are collapsed into a single 'core' definition representing all suballeles categorized under a star (*) number**.

Core alleles have their own unique PVID.

A core SNV is a sequence variation that is part of a core allele definition.

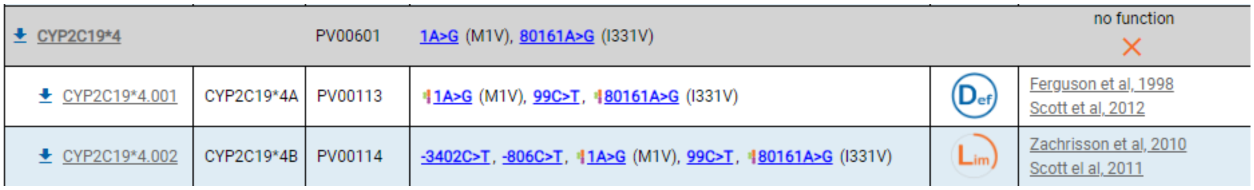
Core allele definitions are highlighted by a gray background in the *CYP2C19* gene page 'Table View'. Suballeles are labeled as e.g. *CYP2C19*2.001*, *CYP2C19*2.002, etc.* and are shown under an allele's core definition with alternating white and blue backgrounds. Function is shown in the core allele bar representing all alleles listed under that core allele definition.

For example, there are numerous *CYP2C19*2* suballeles listed. Of all the SNVs found on these suballeles, only two fulfill the rules and are shared among all suballeles, namely 12662A>G (intron 1) and 19154G>A (c.681G>A). Both render CYP2C19 nonfunctional due to aberrant splicing. Therefore, these two SNVs constitute the *CYP2C19*2* core allele definition. The *CYP2C19*2* core allele definition is shown below along with the first six defined suballeles.



CYP2C19*2		PV00599	12662A>G (splicing defect), 19154G>A (splicing defect), 80161A>G (I331V)		no function ✗
CYP2C19*2.001	CYP2C19*2A	PV00111	99C>T, 12662A>G (splicing defect), 19154G>A (splicing defect), 80160C>T, 80161A>G (I331V)	Lim	De Morais et al, 1994
CYP2C19*2.002	CYP2C19*2B	PV00112	-98T>C, 99C>T, 12460G>C (E92D), 12662A>G (splicing defect), 19154G>A (splicing defect), 80160C>T, 80161A>G (I331V)	Def	deposited by Gaedigk et al Ibeanu et al, 1998
CYP2C19*2.003	CYP2C19*2C	PV00106	-98T>C, 99C>T, 12662A>G (splicing defect), 12834G>C (A161P), 19154G>A (splicing defect), 80160C>T, 80161A>G (I331V)	Def	Fukushima-Uesaka et al, 2005
CYP2C19*2.004	CYP2C19*2D	PV00107	-98T>C, 99C>T, 12662A>G (splicing defect), 19154G>A (splicing defect), 80160C>T, 80161A>G (I331V), 87275G>A	Lim	Lee et al, 2009 Chua et al, 2013
CYP2C19*2.005	CYP2C19*2E	PV00104	99C>T, 12662A>G (splicing defect), 19154G>A (splicing defect), 19286G>A, 80160C>T, 80161A>G (I331V)	Lim	Hu et al, 2012

For *CYP2C19*4*, there is one SNV that is shared among the subvariants, 1A>G (M1V). Therefore, this SNV is the sole variant of the *CYP2C19*4* core allele definition besides80161 (133V) which is present in the vast majority of alleles.



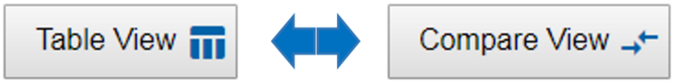
CYP2C19*4		PV00601	1A>G (M1V), 80161A>G (I331V)		no function ✕
CYP2C19*4.001	CYP2C19*4A	PV00113	1A>G (M1V), 99C>T, 80161A>G (I331V)	Def	Ferguson et al, 1998 Scott et al, 2012
CYP2C19*4.002	CYP2C19*4B	PV00114	-3402C>T, -806C>T, 1A>G (M1V), 99C>T, 80161A>G (I331V)	Lim	Zachrisson et al, 2010 Scott el al, 2011

CAVE - the PharmVar Comparative Allele ViewEr

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

To access CAVE switch from Table View to Compare View

Comparative Allele ViewEr (CAVE)



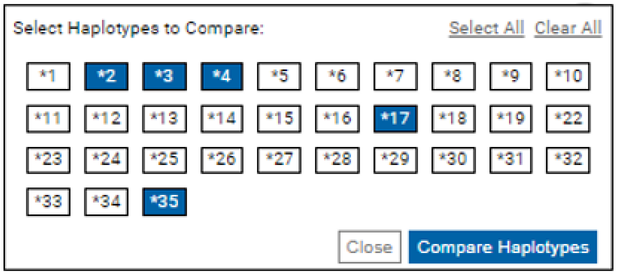
A sequence variation that is part of a core allele definition may be unique or occur in two or more core alleles. It may, however, also be found in some other suballeles. In contrast, some non-core SNVs may be unique and can tentatively be utilized to identify an allele of interest or discriminate it from others. As mentioned above, the CAVE tool uses core allele definitions, and hence unique SNVs that are not part of any core definition are not displayed.

One prime example is -806C>T which is part of the *CYP2C19*17* and core definition, but is also present on one of the two known *CYP2C19*4* suballeles (see screenshot above). Another example is 12662A>G, which is part of two core allele definitions, *CYP2C9*2* and **35.*

Examples:

Select core alleles for comparison in the **selection pad**. To illustrate the example given above and show how these compare with other core alleles we have selected *CYP2C9*2, *3, *4, *17,* and **35*.

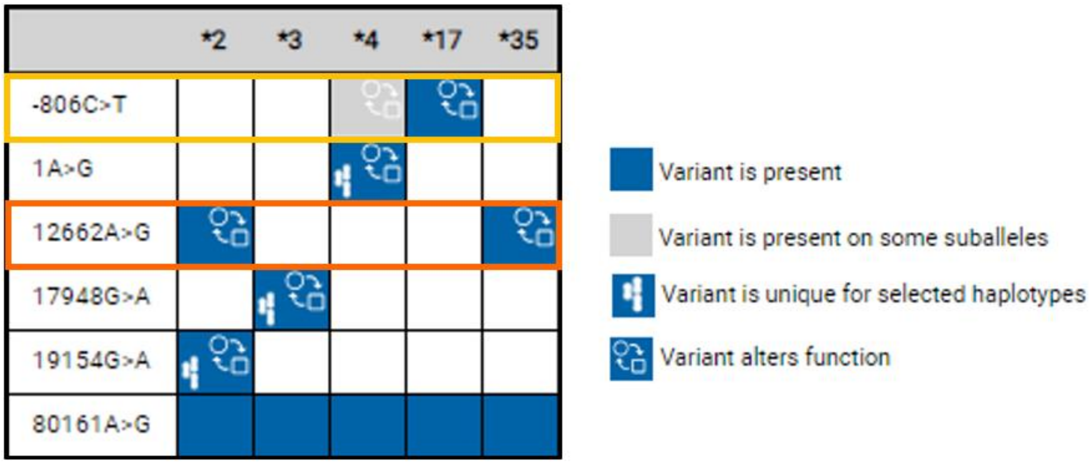
Allele Selection Pad


Select Haplotypes to Compare:

Click compare haplotypes

CAVE graphically displays the core SNVs that are found in the selected alleles

- -806C>T is shown in **blue** for **17* indicating that it is part of its core allele definition (highlighted by yellow frame). This SNV is not part of the **4* core allele defitnion, because is has not been found on all **4* suballeles; it is therefore shown in **gray.** Thus, -806C>T is not a unique SNV and should always be tested along the **4*-defining SNV (1A>G).
- Similarly, 12662A>G is shown in **blue** for **3* and **35* indicating that it is part of their respective core allele definitions (highlighted by orange frame). In order to discriminate these two alleles, 19154G>A, a SNV unique for **2* must be interogated.
- 80161A>G (I331V) is present on all selected alleles and therefore cannot be utilized to discriminate any of the alleles shown.



	*2	*3	*4	*17	*35
-806C>T			▨ ⟳	■ ⟳	
1A>G			■ ▮ ⟳		
12662A>G	■ ⟳				■ ⟳
17948G>A		■ ▮ ⟳			
19154G>A	■ ▮ ⟳				
80161A>G	■	■	■	■	■

■ Variant is present
▨ Variant is present on some suballeles
▮ Variant is unique for selected haplotypes
⟳ Variant alters function

Function

The function of an allele is shown according to the information curated by the PharmGKB and Clinical Pharmacogenetic Implementation Consortium (CPIC) available at https://www.pharmgkb.org/page/cyp2c19RefMaterials. Allele functionality tables are created as part of a more formalized CPIC guideline process and are periodically updated. Additional information may have become available since the functionality table has been created. For some alleles, the assigned activity may differ from that shown on the archived Nomenclature pages. Extrapolation of the activity of a particular allele to particular substrates should be done with caution due to potential unknown substrate-specific activities of the allele.

References

The references provided in the PharmVar database include the citation in which an allele was first published. For some alleles additional reference(s) describe important updates and/or function. The reference list is not intended to provide a complete bibliography for an allele. Haplotypes not published elsewhere are listed as "deposited by".

Allele frequencies

CYP2C19 allele frequency tables have been developed for CPIC guidelines and are available through the PharmGKB at https://www.pharmgkb.org/page/cyp2c19RefMaterials. A comprehensive list of frequencies and references can be found in the *CYP2C19* allele frequency Table in the source tab. These tables are periodically updated.

Changes and Edits

A number of changes and edits have been made to the annotations on the original P450 Nomenclature site to standardize annotations across genes and correct errors. Please see the **CHANGE Log** document for details.