This document provides an overview of genetic variation affecting the *CYP2B6* gene locus and detailed information regarding structural variants.

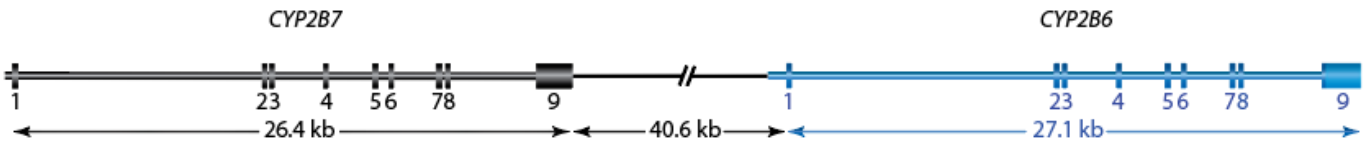
CYP2B6 reference gene locus

The *CYP2B6* gene locus contains two genes, *CYP2B6* and *CYP2B7* (**Figure 1**) of which the latter is considered to be a pseudogene (also referred to as *CYP2B7P*). The distance between the two genes is 41.6 kb. Both genes are composed of nine exons and share a high degree of sequence similarity. An alignment of *CYP2B6* and *CYP2B7* gene sequences revealed few differences among the two genes within exon 1 (9 nucleotide positions differ), exon 2 (9), exon 3 (6), exon 4 (13), exon 5 (2), exon 6 (4), exon 7 (10), exon 8 (3) and exon 9 (5) **(Table 1).**

Table 1 Differences among *CYP2B6* and *CYP2B7* exons

***CYP2B6* position GRCh38**	***CYP2B6* exon**	***CYP2B6* position GRCh38**	***CYP2B6* exon**	***CYP2B6* position GRCh38**	***CYP2B6* exon**
40991311	Exon 1	41004406	Exon 3	41010131	Exon 6
40991347	Exon 1	41004430	Exon 3	41012310	Exon 7
40991371	Exon 1	41004298	Exon 3	41012327	Exon 7
40991381	Exon 1	41006906	Exon 4	41012349	Exon 7
40991388	Exon 1	41006913	Exon 4	41012354	Exon 7
40991390	Exon 1	41006936	Exon 4	41012393	Exon 7
40991391	Exon 1	41006967	Exon 4	41012394	Exon 7
40991428	Exon 1	41006981	Exon 4	41012411	Exon 7
40991446	Exon 1	41007013	Exon 4	41012413	Exon 7
41004093	Exon 2	41007028	Exon 4	41012465	Exon 7
41004102	Exon 2	41007034	Exon 4	41012475	Exon 7
41004134	Exon 2	41007038	Exon 4	41012679	Exon 8
41004138	Exon 2	41007040	Exon 4	41012711	Exon 8
41004139	Exon 2	41007054	Exon 4	41012789	Exon 8
41004141	Exon 2	41007058	Exon 4	41016674	Exon 9
41004148	Exon 2	41007060	Exon 4	41016726	Exon 9
41004152	Exon 2	41009287	Exon 5	41016743	Exon 9
41004155	Exon 2	41009358	Exon 5	41016746	Exon 9
41004299	Exon 3	41010006	Exon 6	41016810	Exon 9
41004302	Exon 3	41010037	Exon 6		
41004355	Exon 3	41010039	Exon 6		

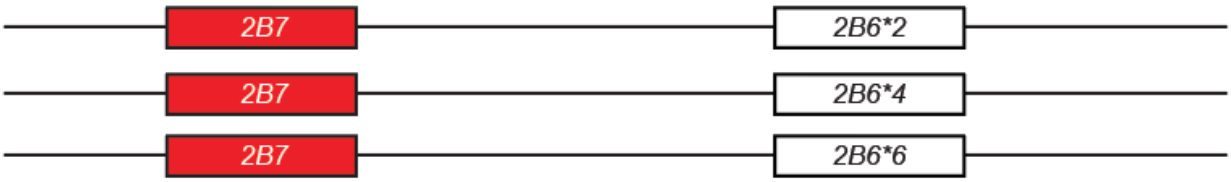
Figure 1 Overview of the *CYP2B* gene locus. Both genes are encoded by the positive strand. Boxes denote exons along exon number. The length of each gene and the distance between the two genes are shown in base pairs (bp).



Variant *CYP2B6* alleles

Alleles with SNVs and indels: The majority of *CYP2B6* allelic variants carry SNVs, indels or a combination thereof (**Figure 2**). These alleles may have increased, normal, decreased or no function. All allelic variants with SNVs and/or indels are defined in the PharmVar database.

Figure 2



Copy number variation (CNV)

CYP2B6 deletion and duplication events have been reported particularly for Africans at frequencies of 0.5-0.9%% and 0.4%, respectively (PMIDs 23164804 and 29261188). However, it remains unknown whether these are affecting the entire gene locus, or only a segment of the gene within the locus, as their precise breakpoints have not been mapped. It is likely that deletion and duplication events crossing between *CYP2B6* and *CYP2B7* are the underlying mechanisms of forming hybrids between these two genes. For *CYP2B6*29* and **30* breakpoints were narrowed to a 529-bp interval within introns 4 and exons 5 of *CYP2B6* and *CYP2B7P1* where the two genes are nearly identical (PMID 23164804).

Structural variants

CYP2B7-2B6 hybrid genes:

The 5'-portion of these structural variants is derived from *CYP2B7* and the 3'-portion is derived from *CYP2B6* (**Figure 3**). This hybrid is believed to be the product of a large deletion between *CYP2B6* and *CYP2B7* that was facilitated by repetitive Alu elements in intron 4 (**Table 2**). *CYP2B6*29* shares complete sequence identity with exons 1 to 4 of *CYP2B7* and exons 5 to 9 of *CYP2B6* and has been annotated as having decreased function. The *CYP2D6*29* hybrid is described in the PharmVar database as *CYP2B7-2B6* hybrid gene (see 'Structural Variation Document' for CYP2B6').

Figure 3

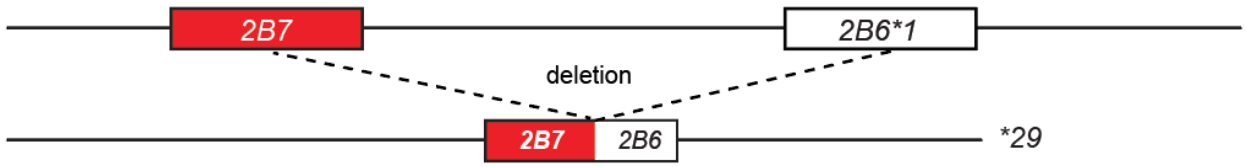


Table 2 ***CYP2B7-2B6* Hybrid Genes**

allele designation	hybrid structure	switch region	function	references	PMID
****29***	*2B7-2B6*	intron 4	decreased	Rotger et al, 2007; Martis et al 2012	17885627 23164804

CYP2B6-2B7 hybrid genes:

This structural variant is believed to be the product of an unequal cross-over and duplication event that was also facilitated by intron 4 Alu elements. As shown in **Figure 4**, the *CYP2B6-2B7* hybrid cataloged as **30* (**Table 3**) represents an additional gene copy that is located between *CYP2B7* and *CYP2B6*. Such arrangements are referred to as tandem arrangements to distinguish allelic variants with two or more gene units that are not identical from allelic variants with two or more identical duplicated gene copies. Such arrangements may, however, also be simply be referred to as duplications in the literature reflecting the gain of a gene copy. The **30* hybrid is likely nonfunctional due to the presence of multiple SNVs including R378X that causes a premature stop codon. The function of the **30*-containing tandem described by Martis et al

depends on the other gene copy within the tandem. Since Martis el al could not resolve the phase of SNPs causing Q172H and K262R, the downstream *CYP2B6* copy could be a **1, *4, *6 or *9*.

Figure 4

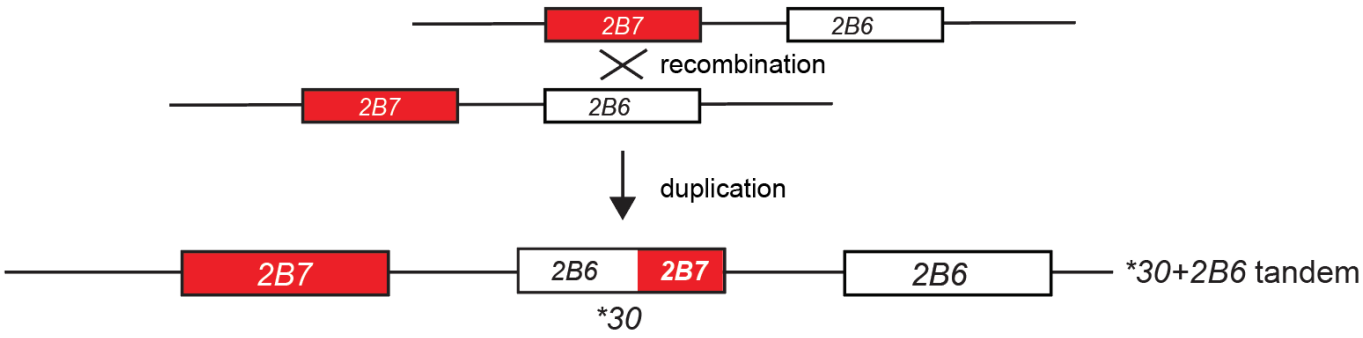


Table 3 *CYP2B6-2B76* Hybrid Genes

allele designation	hybrid structure	switch region	function	references	PMID
30	*2B6-2B7*	intron 4	none	Martis et al 2012	23164804

References

The references provided in this Structural Variation document include the citation in which an allele was first published and additional reference(s) describing important updates and/or information regarding function. The reference list is not intended to provide a complete bibliography for an allele.

Users are encouraged to share their research with and/or bring important literature that might have inadvertently been missed to the attention of PharmVar.

Allele frequencies

CYP2B6 allele frequency tables have been developed for CPIC guidelines and are available through the PharmGKB at https://www.pharmgkb.org/page/cyp2b6RefMaterials. A comprehensive list of frequencies including population-specific information and references can be found in the *CYP2B6* allele frequency table in the 'references' tab. These tables are periodically updated.

Changes and Edits

Suballeles are designated using the revised nomenclature system that replaced letters (e.g. A, B, etc.) with .001, .002, etc.).

A number of changes and edits have been made compared to the original allele annotations to standardize annotations across genes and correct errors. Please see the **Change Log** document for specific details.