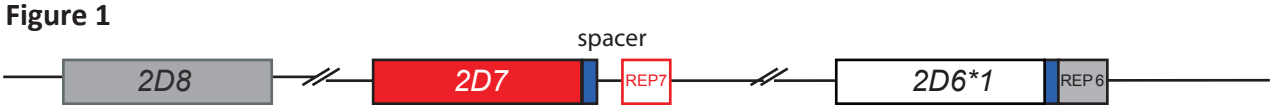
CYP2D6 reference gene locus

The *CYP2D6* gene locus contains three genes, *CYP2D6*, *CYP2D7* and *CYP2D8* (**Figure 1**). *CYP2D7* and *CYP2D8* are considered pseudogenes. All three genes are composed of nine exons and share a high degree of sequence similarity. *CYP2D6* and *CYP2D7* share a common, 0.6 kb long downstream region (blue boxes) and have near-identical repetitive sequences referred to as REP6 and REP7, respectively. A hallmark feature of the 3' region of *CYP2D7* is the presence of a 1.6 kb long additional 'spacer' sequence between the common sequence and REP7 that further distinguishes it from *CYP2D6*.

Figure 1

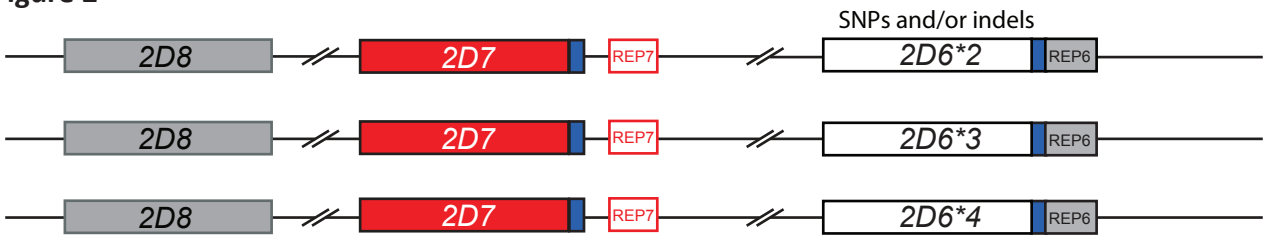


This document provides an overview of genetic variation affecting the *CYP2D6* gene locus and detailed information regarding structural variants including gene deletions and duplication (copy number variation, CNV), conversions and structural rearrangements between *CYP2D6* and its pseudogenes.

Variant *CYP2D6* alleles

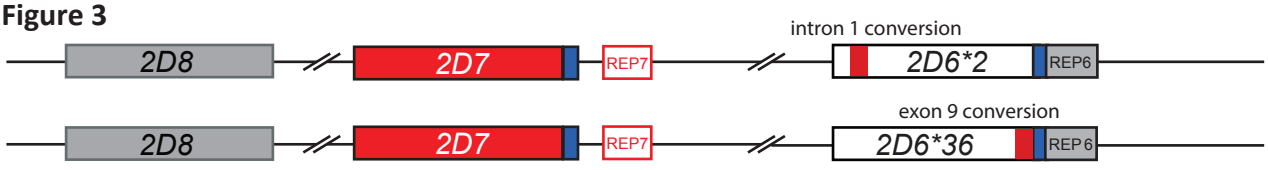
Alleles with SNPs and indels: The majority of *CYP2D6* allelic variants carry SNPs, indels or a combination thereof (**Figure 2**). These alleles may have increased, normal, decreased or no function. All allelic variants with SNPs and/or indels are defined in the PharmVar database.

Figure 2



Alleles with *CYP2D7* conversions: Many *CYP2D6* allelic variants carry a small region that is derived from *CYP2D7* (**Figure 3**). These 'embedded' regions are typically referred to as *CYP2D7* conversions. Note that non-duplicated alleles with embedded *CYP2D7*-derived conversions have a REP6 3' end. Also noteworthy is *CYP2D6*36* which contains an embedded exon 9 conversion. This entity can occur either as a single gene unit (as shown here with a REP6 3'end), or as a hybrid (with a spacer and REP7 3'end as shown in Figures 7 and 8).

Figure 3



Allelic variants reported to have a *CYP2D7* conversion are listed in **Table 1.** All allelic variants in this category are listed in the PharmVar database. Please see the 'change log' document regarding previous and current database annotations of these conversions.

Table 1

allele designation	*CYP2D7*
many alleles	intron 1 conversion: 214G>C, 221C>A, 223C>G, 227T>C, 232G>C, 233A>C, 245A>G (NG_008376.3 and M33388) For alleles defined based on exon sequence only, it remains unknown whether the intron 1 conversion is present. The number of alleles with the intron 1 conversion may therefore be underestimated.
***4.013** (with REP6 or REP7)	exon 9 conversion: 4125G>C, 4129C>G, 4132A>G, 4134T>C, 4156C>T, 4157A>C, 4159G>C, 4165T>G, 4167T>C, 4168G>A, 4169C>G, 4170T>C, 4173C>T (NG_008376.3; M33388 not shown)
***35.002**	conversion upstream of exon 1: -431C>T, -354A>G, -334G>C, -331T>G, -328C>T, -327A>G, -321C>G, -320A>G, -276C>T, -275C>T, -272C>T, -268G>A, -267G>C, -232G>C, -225A>G (NG_008376.3 and M33388)
***36** (with REP6, REP7 or REPdel)	exon 9 conversion (see above)
***57** (nature of REP unknown)	exon 9 conversion (see above)
***82**	exon 2 conversion: 973C>A, 983A>G, 996C>G, 1013T>C, 1021A>T, 1022C>A, 1027A>G, 1035T>C (NG_008376.3; M33388 not shown)
***83** (with REP6)	exon 9 conversion (see above)

Copy number variation (CNV)

Gene deletion: The allelic variant defined as *CYP2D6*5* is characterized by a deletion of the entire *CYP2D6* gene as shown in **Figure 4**; the REP region for this allele is referred to as REPdel. It remains unknown, however, whether all deletion alleles identified as *CYP2D6*5* have identical breakpoints. *CYP2D6*5* is described in the PharmVar database as 'deletion of the entire gene'.

The gene deletion is typically reported as *CYP2D6*5* and diplotypes described as e.g. *CYP2D6*1/*5*. Some investigators or laboratories may describe a diplotype as *(*1/*1) x1, (*1/*1) 1N or (*1/*5) 1N* or similar (there are currently no standards describing how to textually display *CYP2D6* CNVs).

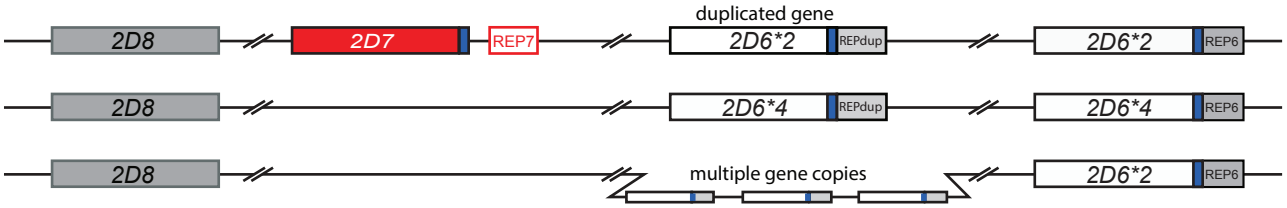
Figure 4

Gene duplications and multiplications: A number of *CYP2D6* alleles are known to occur as duplications or multiplications (**Figure 5**). Duplicated gene copies have a *CYP2D6*-like downstream

region without the spacer, but the REP element is not identical to REP6 and thus, is referred to as REPdup. Figure 5 provides selected examples of alleles carrying two presumed identical gene copies. The majority of these copy number gains are gene duplications. The gene copy between the *CYP2D7* gene and the most 3' gene copy is typically referred to as the 'duplicated' gene copy.

Numerous allelic variants carrying two or more gene copies have been described for *CYP2D6*. Note that duplications/multiplications have been described for normal (e.g. *CYP2D6*1, *2*), decreased (e.g. *CYP2D6*41*) and nonfunctional (e.g. *CYP2D6*4*) allelic variants. PharmVar does not list gene duplications as separate entities in the database. The user is referred to **Table 2** that summarizes the gene duplications that have been reported in the literature to date.

Figure 5



In very few instances have both, or all, gene copies been sequenced. Gene copies are generally deemed identical, an assumption based on limited data (i.e. when both gene copies were genotyped and shown to carry the same set of key identifying SNPs or indels, or in rare occasions, sequenced). This may, however, not always be the case.

If the number of gene copies is known, duplications are commonly annotated as x2, x3, etc. and if the number is unknown as "xN". "Unknown to exist" in **Table 2** indicates that a multiplication has not been described in the literature or has not been submitted to PharmVar so far. If it is unknown which of the two chromosomes carries the duplication/multiplication and/or the number of gene copies is unknown, diplotypes described on reports as e.g. **1/*2* dup, (**1/*2*) x2, (**1/*2*) xN, (**1/*2*) N3 or similar (as noted above there are currently no standards of how to display *CYP2D6* CNVs).

Table 2

duplications	multiplications (N≥3)	function for x2	references	PMID
****1x2***	**1xN*	increased	Dahl et al. 1995 Sachse et al. 1997 Gaedigk et al. 2007 Hosono et al. 2009 Gaedigk et al. 2012 Del Tredici et al.	7616439 9012401 17259947 19541866 22111604 29674966
****2x2***	**2xN*	increased	Johansson et al. 1993 Dahl et al. 1995 Aklillu et al. 1996 Gaedigk et al. 2007	7903454 7616439 8764380 17259947

duplications	multiplications (N≥3)	function for x2	references	PMID
			Hosono et al. 2009 Gaedigk et al. 2012 Del Tredici et al. 2018	19541866 22111604 29674966
***3x2**	unknown to exist	none	Del Tredici et al. 2018	29674966
***4x2**	*4xN	none	Løvlie et al. 1997 Sachse et al. 1998 Gaedigk et al. 2007 Gaedigk et al. 2012 Del Tredici et al. 2018	9170153 10022755 17259947 22111604 29674966
***6x2**	unknown to exist	none	Gaedigk et al. 2007 Del Tredici et al. 2018	17259947 29674966
***9x2**	unknown to exist	normal	Gaedigk et al. 2011 Del Tredici et al. 2018	22044417 29674966
***10x2**	*10xN	decreased	Garcia-Barceló M, 2000 Ji et al. 2002 Mitsunaga et al. 2002 Ishiguro et al. 2004 Gaedigk et al. 2007 Hosono et al. 2009 Del Tredici et al. 2018	10973875 12089164 12175908 15149890 17259947 19541866 29674966
***17x2**	unknown to exist	normal	Cai et al. 2006 Gaedigk et al. 2007 Gaedigk et al. 2012 Del Tredici et al. 2018	16550211 17259947 22111604 29674966
***29x2**	unknown to exist	normal	Gaedigk et al. 2007 Gaedigk et al. 2012 Del Tredici et al. 2018	17259947 22111604 29674966
***35x2**	unknown to exist	increased	Griese et al. 1998 Gaedigk et al. 2007 Gaedigk et al. 2012 Del Tredici et al. 2018	9511177 17259947 22111604 29674966
***41x2**	*41xN	normal	Gaedigk et al. 2007 Gaedigk et al. 2012 Del Tredici et al. 2018	17259947 22111604 29674966
***43x2**	unknown to exist	uncertain	Gaedigk et al. 2007	17259947
***45x2**	unknown to exist	increased	Gaedigk et al. 2007	17259947

Structural variants

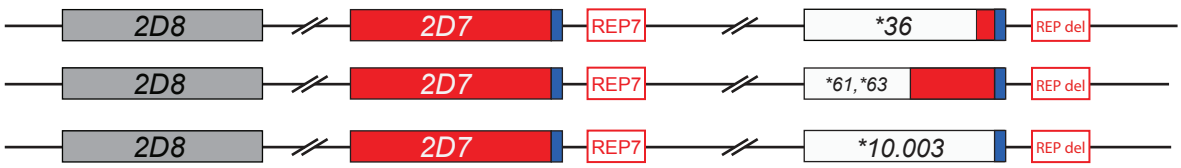
CYP2D7-2D6 hybrid genes: The 5'-portion of these structural variants is derived from *CYP2D7* and the 3'-portion is derived from *CYP2D6* (**Figure 6**). These hybrids are believed to be the product of a large deletion between *CYP2D7* and *CYP2D6*. A series of hybrid genes following this structure has been reported; three representative structures are shown in the graph below. A hallmark feature of these hybrids is a T-insertion in exon 1, which renders these hybrids non-functional. Hybrid genes with this structure have been consolidated under the *CYP2D6*13* allele designation (**Table 3**). *CYP2D6*13* hybrids are described in the PharmVar database as *CYP2D7-2D6* hybrid genes (see 'Structural Variation Document' for *CYP2D6'*).

Figure 6

CYP2D6-2D7 hybrid genes (singletons): The 5'-portion of these structural variants is derived from *CYP2D6* and the 3'-portion is derived from *CYP2D7* (**Figure 7**). Of note, such hybrids have a *CYP2D7*-like downstream region containing a 'spacer' sequence, which is a hallmark feature of *CYP2D7*. The REP sequence, however, is most often matching REPdel (seen in the *CYP2D6*5* gene deletion). Known *CYP2D6-2D7* hybrid genes include **10.003, *61* and **63*. Of importance, **36* has been described as having a *CYP2D6*-derived (rare cases) or *CYP2D7*-like downstream region with the spacer. The latter is a true hybrid, i.e. *2D6* switches over to *2D7*, while the former has an embedded *2D7*-derived sequence (also referred to as conversion). *CYP2D6-2D7* hybrid genes are listed in **Table 3** and are described in the PharmVar database as *CYP2D6-2D7* hybrid genes (see 'Structural Variation Document' for *CYP2D6)*.

Of note, *CYP2D6-2D7* hybrids with a REP7-like downstream region may support amplification with certain XL-PCR-based *CYP2D6*5* assays leading to false-positive **CYP2D6*5* calls.

Figure 7



CYP2D6-2D7 hybrid genes (in tandems): The downstream region of *CYP2D6-2D7* hybrid genes in tandems is typically *CYP2D7*-derived, i.e. containing the 'spacer' sequence followed by REP7. See below for more information on 'tandem' arrangements.

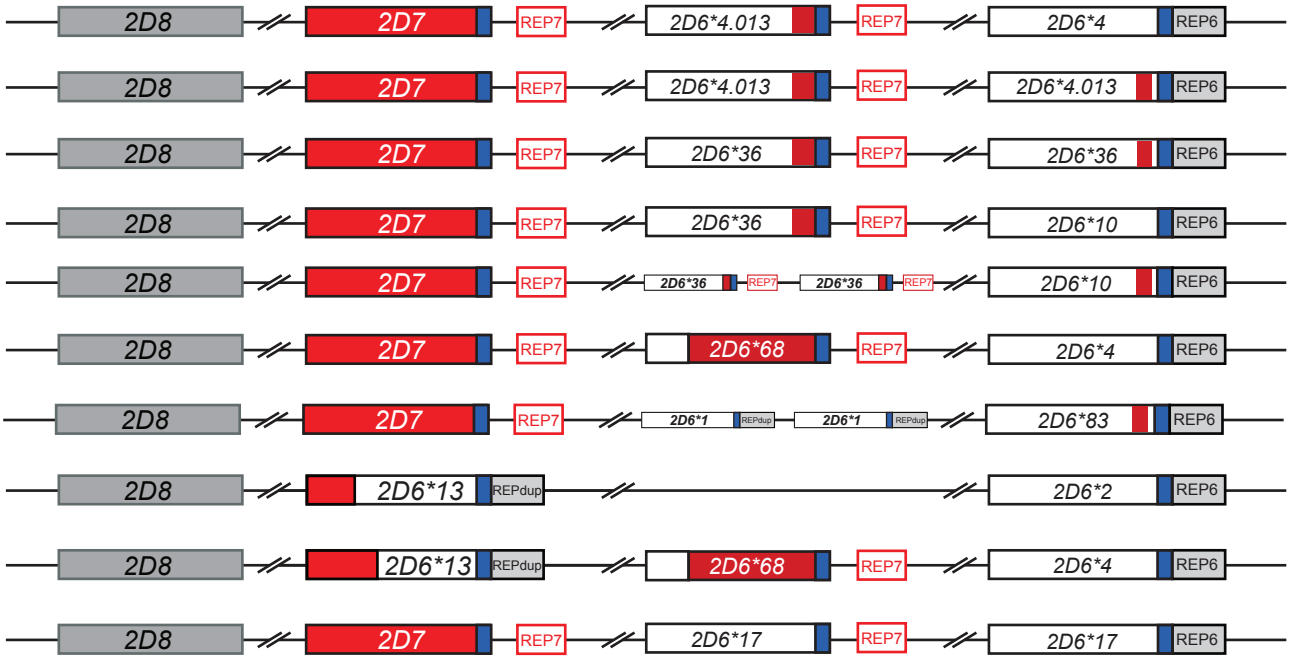
Table 3

allele designation	hybrid structure	switch region	function	references	PMID
***4.013**	*2D6-2D7*	Within or upstream of exon 9, *2D7*-derived downstream region when duplicated	none	Gaedigk et al. 2006	16415111
***10.003**	*2D6-2D7*	downstream exon 9; *2D7-derived* downstream region	decreased	Ishiguro et al. 2004	15313161
***13**	*2D7-2D6*	The following switch regions have been described: intron 1, exon 2, intron 2-exon 3, intron 4, exon 5, exon 7-intron 8, intron 7, exon 9	none	Panserat et al. 1995 Gaedigk et al. 2010 Gaedigk et al, 2010 Black et al, 2012	8554938 20017671 21833166 22004686
***36**	*2D6-2D7*	Within or upstream of exon 9; *2D6* or *2D7* derived downstream region	none	Gaedigk et al. 2006 Hosono et al. 2009 Kramer et al. 2009 Del Tredici et al. 2018	16415111 19541866 19741566 29674966
***61**	*2D6-2D7*	intron 7	uncertain	Kramer et al. 2009	19741566
***63**	*2D6-2D7*	exon 8	uncertain	Kramer et al. 2009	19741566
***68**	*2D6-2D7*	intron 1	none	Kramer et al. 2009 Gaedigk et al. 2012	19741566 22111604

Tandem arrangements: To distinguish allelic variants with two or more gene units that are not identical from allelic variants with identical units that are duplicated or multiplied, they are often referred to as 'tandems'. As shown in **Figure 8**, tandems can harbor two or more gene copies. In the majority of tandems, at least one copy constitutes a *CYP2D7-2D6* or *CYP2D6-2D7* hybrid. The latter are characterized by having a *CYP2D7*-derived downstream region including the spacer.

The PharmVar database does not list tandem arrangements. Please see **Table 4** for a listing of tandem arrangements that have been described in the literature and/or submitted to PharmVar.

Figure 8



As described earlier, *CYP2D6*36* can occur as a 'singleton' with an embedded *CYP2D7*-derived exon 9 conversion or as a true hybrid with a switch to *CYP2D7*. The latter configuation is part of the *CYP2D6*36+*10* tandem that is foremost found in individuals of East Asian ancestry. Interestingly, the most distal gene copies in all tandems described to date have *REP6* downstream regions as shown in **Figure 8**.

Regarding *CYP2D6*68*, this hybrid has only been described to occur in tandem as shown above (i.e. *CYP2D6*68+*4*). It can, however, not be excluded that the **68 CYP2D6-2D7* hybrid gene (with the same or similar switch region) also exists in other tandem arrangements or as a 'singleton'.

The duplicated gene copy of *CYP2D6*17x2* most often has a REPdup without the spacer, like other duplicated/multiplied alleles described earlier. In rare cases, however, the duplicated gene copy has the spacer sequence and tentative a REP7 element as shown in **Figure 8**.

Lastly, the *CYP2D6* portion of hybrid genes found in tandem tend to have the same sequence as the downstream *CYP2D6* gene.

Expert panel members involved with clinical testing have observed test results that suggest the existence of numerous additional CNV structures with and without hybrid genes that have not been further characterized, published or submitted to PharmVar.

Table 4 summarizes the tandem arrangements that have been reported to date.

The genes are displayed in the order they are located on the chromosome, i.e. **36+*10* indicates that the **36* hybrid is located upstream of the **10* gene copy. This tandem may be displayed by others as e.g. **36-*10* or **10-*36* (there are currently no standards describing how to textually display *CYP2D6* CNVs).

Table 4

allele designation	function	references	PMID
1x2+*83	increased	Gaedigk et al. 2012	22111604
1+*90	unknown	Gaedigk et al.	submitted
4.013+*4	none	Gaedigk et al. 2012 Del Tredici et al. 2018	22111604 29674966
4.013xN+*4	none	Del Tredici et al. 2018	29674966
4.013+*4xN	none	Del Tredici et al. 2018	29674966
13*+*1	normal	Gaedigk et al. 2010	20017671
13+*1x2	increased	Black et al. 2012	22004686
13+*2	normal	Gaedigk et al. 2010	20017671
13+*68+*4	none	Black et al. 2012	22004686
17x2	normal	Gaedigk et al. 2007 Gaedigk et al. 2012	17259947 22111604
36x2	none	Gaedigk et al. 2006 Hosono et al. 2009 Del Tredici et al. 2018	16415111 19541866 29674966
36x2+*10	decreased	Hosono et al. 2009 Gaedigk et al. 2012 Del Tredici et al. 2018	19541866 22111604 29674966
36+*10	decreased	Johansson et al. 1994 Leathart et al. 1998 Gaedigk et al. 2006 Hosono et al. 2009 Del Tredici et al. 2018	7935325 9918137 16415111 19541866 29674966
36+*10x2	decreased-normal	Hosono et al. 2009 Gaedigk et al. 2012 Del Tredici et al. 2018	19541866 22111604 29674966
57+*10	decreased	Soyama et al. 2006 Sakuyama et al. 2008	16858124 18784265
68+*4	none	Kramer et al. 2009 Gaedigk et al. 2012	19741566 22111604

References

The references provided in the PharmVar database and the ReadMe document include the citation in which an allele was first published. For some alleles additional reference(s) describe 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

CYP2D6 allele frequency tables have been developed for CPIC guidelines and are available through the PharmGKB *Gene-specific Information Table*.

A comprehensive list of allele frequencies and references can be found in the *CYP2D6* allele frequency table in the source 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.