Important Information

In the absence of a centralized resource for *NUDT15* nomenclature, some of the allelic variants were assigned star allele numbers when first published (*NUDT15*1-*6* and **10* and **11*), while others were reported without a star allele designation or designations were assigned post-hoc (*NUDT15*7-*9A*). PharmVar provides standardized nomenclature, and an **expert panel curates this gene**.

Some allelic variants have not been fully characterized and may have additional single nucleotide variants (SNVs) and are shown as having a moderate (Mod) evidence level.

Gene Region Mapped

PharmVar defines haplotypes (star alleles) for *NUDT15* covering the following regions: 5' flanking sequence, the 5'UTR, all three exons, exon/intron junctions, and 130 bp of the 3'UTR. **Table 1** specifies the positions within the genomic reference sequence NG_047021.1 that delineate the beginning (5' limit) and end (3' limit) of the sequence used to define *NUDT15* star alleles.

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

Intronic variants are only to be included in star allele definitions if they are within splice recognition sequences, or there is convincing evidence that the variant impacts protein function.

Submissions must cover all regions specified in **Table 1** and list all variants present in the haplotype in the required regions.

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_047021.1	4951	-230	13363	8183
All SNVs within the 5' and 3' limits must be submitted for haplotype (star allele) definitions. This includes 5' flanking sequence, the 5'UTR, exons 1-3 including exon/intron junctions, and a portion of the 3'UTR.				

Coordinates

Variant coordinates in the PharmVar database are available for genomic and transcript reference sequences (RefSeq) and the Human Genome Assemblies GRCh37 and GRCh38.

Table 2 summarizes important information regarding these sequences.

In this document coordinates are counted from the 'A' in the ATG translation start codon as +1 for the NM_018283.4 transcript as reference unless indicated otherwise.

Table 2

Sequence	Notes
NUDT15	Encoded on the positive strand.
NG_047021.1	This genomic RefSeq matches the *NUDT15.001* allele. **NG_047021.1 is used as a reference to define allelic variation (star alleles). PharmVar submissions must be annotated using this RefSeq.**
LRG	There is no Reference Genomic (LRG) for *NUDT15*.
NM_018283.4	The transcript RefSeq matches *NUDT15*1.* NM_018283.4 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 *NUDT15* sequence matches *NUDT15*1.*
GRCh38	The *NUDT15* sequence matches *NUDT15*1.*

PharmVar Feature: Get coordinates from the sequence start or the ATG start codon with one click! The default setting is NG_047021.1 counting the 'A' in the ATG start codon as +1. The ⓘ buttons link out to NCBI to access respective sequences.

Sequence Variants

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

Some SNVs are unique to an allele and only occur on a single haplotype, e.g., c.416G>A (p.139H) is unique to the *NUDT15*4* haplotype. Other variants occur on more than one allele, e.g., the GAGTCG insertion at position c.55 is found in *NUDT15*3* and **6.*

The PharmVar Variant Window

Click on a variant to open this window. As an example, **Figure 1** shows the Variant Window for the *NUDT15*3* variant c.415C>T.

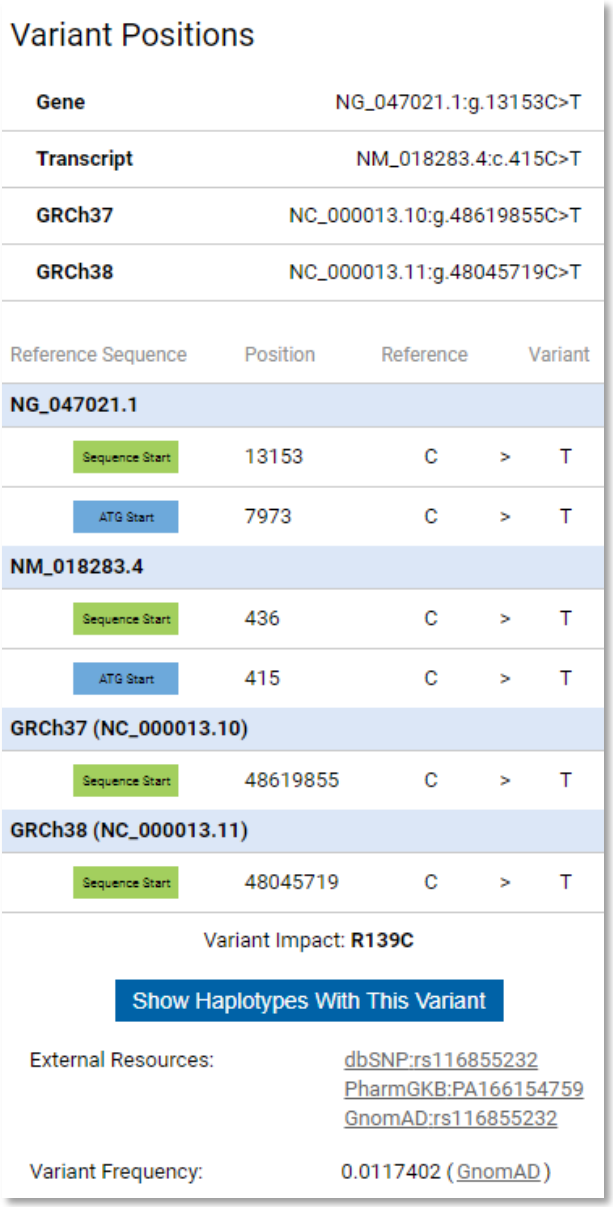
This view provides SNV positions across all reference sequences, links to external resources, including dbSNP (note that some SNVs may not have an rsID), ClinPGx (formerly PharmGKB) and gnomAD, and frequency information. There is also a **button to display all haplotypes (star alleles) that contain the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

The **top portion** of the variation window **displays SNV coordinates according to Human Gene Variation Society (HGVS)** nomenclature at the gene, transcript and genome levels. Coordinates are displayed 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 variation window **displays SNV positions 'PharmVar-style'** on the gene, transcript and genome levels providing reference and variant nucleotides and positions for both count modes.

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 them as duplications or 'dup' if a base is duplicated. Additional details and examples are provided under Standards.

Figure 1 Variant Window



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

The PharmVar Allele Page

Click on an allele name to open this page. The example in **Figure 2** shows the *NUDT15*3* allele page detailing the definition of the **3* core allele using the genomic and transcript reference sequences and GRCh37 and GRCh38 genome builds. Links to additional resources pertinent to the haplotype are provided here.

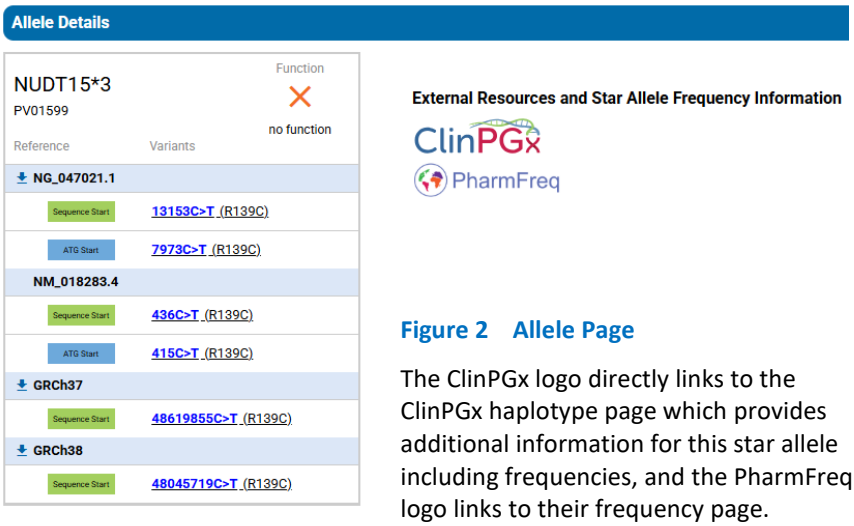


Figure 2 Allele Page

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.

Core Allele Definitions

Some alleles share 'core' defining sequence variants (see Allele Designation and Evidence Levels for details). Therefore, similar alleles are grouped under the same core allele definition (same star number). Although allelic 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 the same star number are assumed to be functionally equivalent based on current literature**. Thus, even if a test can distinguish between suballeles, these can generally simply be reported as e.g., *NUDT15*1, *3,* etc., without the provision of their numeric 'dot' extension.

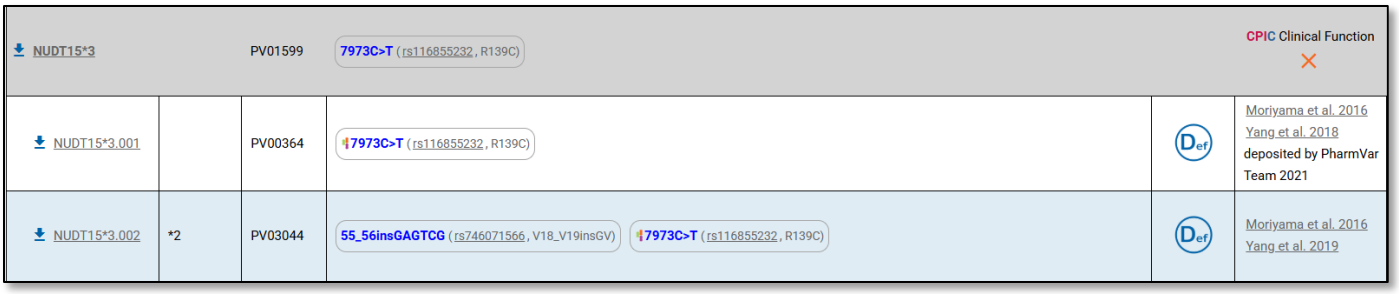
PharmVar has developed core allele definitions in collaboration with ClinPGx. **Only sequence variations that change an amino acid, impact function by altering expression levels, or interfere with splicing and are present in *all* alleles within an allele group are part of the core allele definition.** A core variant is a sequence variation that is part of a core allele definition. Core SNVs are highlighted 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 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 in the *NUDT15* gene page 'Table View'. Alleles of an allele group are labeled as e.g., *NUDT15*1.001*, *NUDT15*1.002,* etc. and are shown under an allele's core 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.

As shown in **Figure 3**, there are two suballeles for **3.* Both have c.415C>T (p.R139C) which eliminates function. The **3.002* suballele (formerly **2.001*) additionally has c.55_56insGAGTCG.

Figure 3 The *NUDT15*3* core allele and suballeles



NUDT15*3		PV01599	7973C>T (rs116855232, R139C)		CPIC Clinical Function ✕
NUDT15*3.001		PV00364	7973C>T (rs116855232, R139C)	Def	Moriyama et al. 2016 Yang et al. 2018 deposited by PharmVar Team 2021
NUDT15*3.002	*2	PV03044	55_56insGAGTCG (rs746071566, V18_V19insGV) 7973C>T (rs116855232, R139C)	Def	Moriyama et al. 2016 Yang et al. 2019

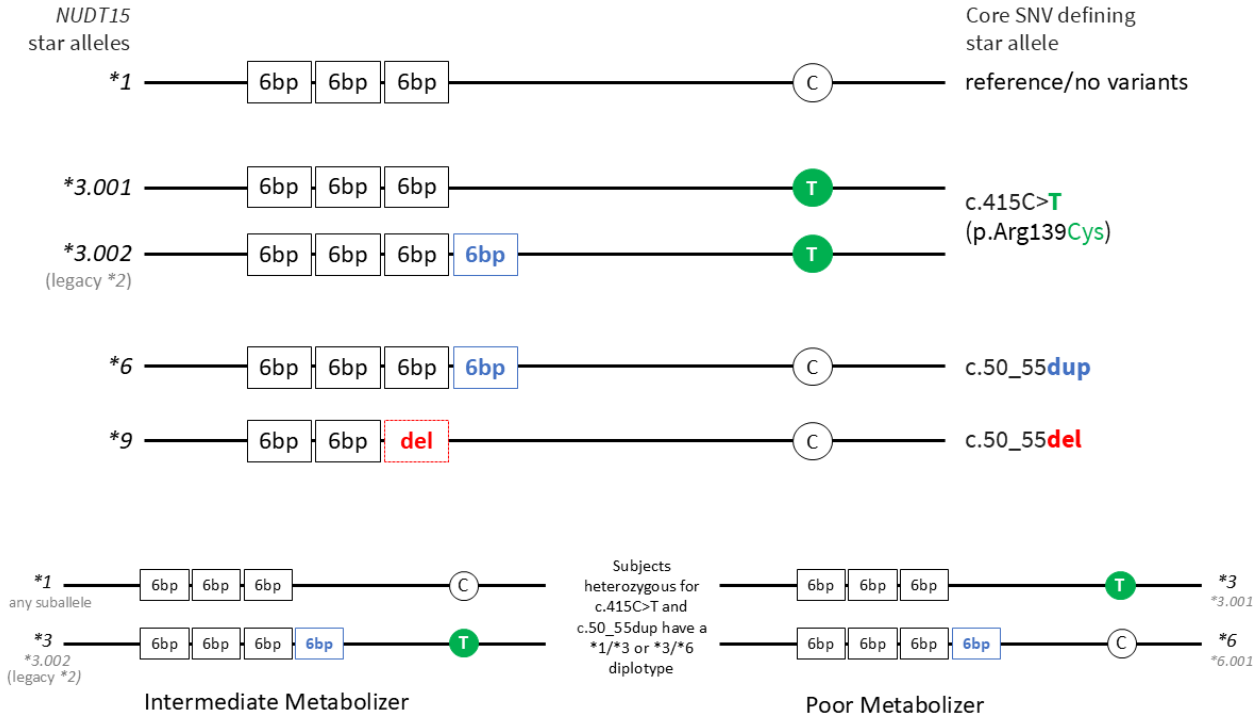
The *NUDT15*2* merger with **3* and challenges of phasing c.415C>T and c.55_56insGAGTCG

*NUDT15*2* and **3* were consolidated under the same star number as both have c.415C>T (p.R139C) which eliminates enzymatic activity. The allele formerly listed as **2.001* was re-designated as **3.002*.

Figure 4 provides a graphical overview of the *NUDT15* star alleles having a GAGTCG insertion or deletion. As illustrated in the figure, the c.415C>T variant can occur without (**3.001*) and with (**3.002*) the insertion, and the insertion can also occur without c.415C>T (**6*). Therefore, as further detailed in the bottom panel of the figure, a patient with a heterozygous test result for both c.415C>T and c.50_55dup without any information regarding variant phase (i.e., whether the variants are *in cis* or *in trans*), the genotype is ambiguous. This means that the patient has either a *NUDT15*1/*3* (intermediate metabolizer) or **3/*6* (poor metabolizer) diplotype with the former having a **3.002* suballele and the latter having a **3.001* suballele. As the predicted phenotypes for *NUDT15*1/*3* and **3/*6* differ, it is important to be aware of this ambiguity. Currently, there is no standardized approach to reporting ambiguous diplotypes when phase is unknown.

Lastly, if a pharmacogenetic test does not interrogate the GAGTCG repeat, *NUDT15*6* and **9* alleles are not detected and 'defaulted' to **1*, while **3.002* will be identified as **3* due to having c.415C>T.

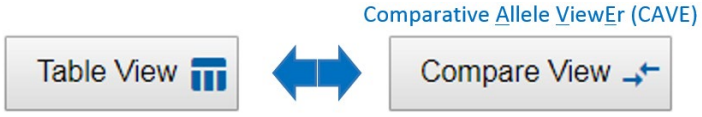
Figure 4 Overview of *NUDT15* star alleles having a 6-bp GAGTCG insertion or deletion and diplotype ambiguity



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

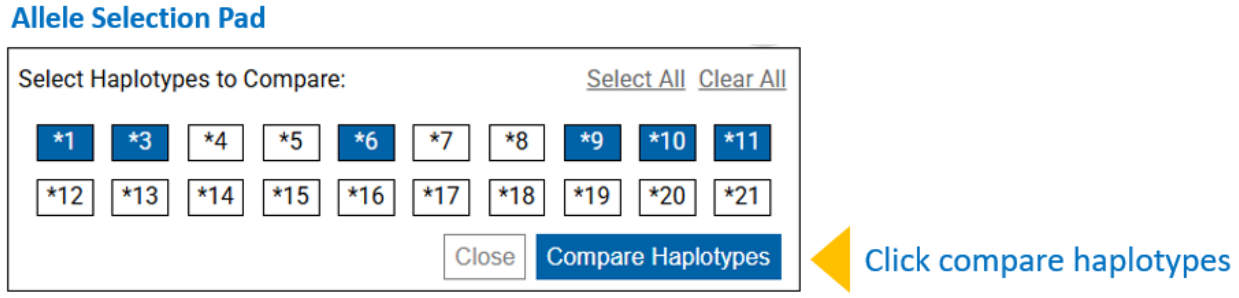


A sequence variation that is part of a core allele definition may be unique or occur in two or more core alleles. It may, however, also be found in some other suballeles. In contrast, some non-core SNVs may be unique and can tentatively be utilized to identify an allele of interest or discriminate it 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.

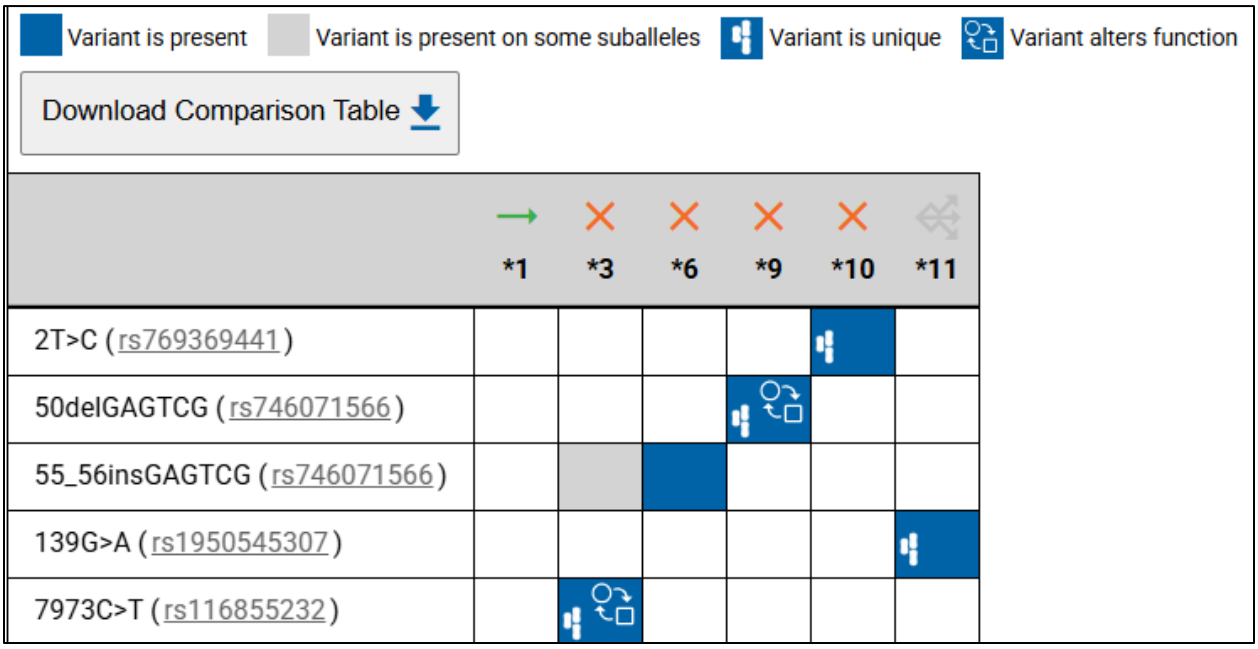
For example, c.55_56insGAGTCG (p.V18_V19insGV) is part of one core allele definition, *NUDT15*6*, but also found on the **3.002* suballele.

Example:

Select core alleles for comparison in the **selection pad**. For this example, we have selected six star alleles, i.e., **1, *3, *6, *9*, **10* and **11*.



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



- c.415C>T is shown in **blue** for *NUDT15*3* indicating that this variant is part of its core definition. This variant also has the function symbol indicating that it has been shown to alter function.
- The GAGTCG insertion and deletion are also shown in **blue** and annotated as having a functional impact.

  The GAGTCG insertion is shown in blue for *NUDT15*6* because it is part of its core allele defitnition and in gray for *NUDT15*3* because it is present in only one of its subalelles. It is also shown in blue for *NUDT15*9,* the haplotype with the GATTCG deletion.

Annotations for rs746071566 (GAGTCG del/dup)

As shown in **Figures 5 and 6** below, *NUDT15*3.002* (formerly **2*) and **6* contain an additional copy of the GAGTCG repeat (4 vs 3 copies of the repeat motif, rs746071566) while *NUDT15*9* has a GAGTCG deletion and carries only 2 copies of the repeat motif.

Due to the nature of this insertion/deletion variation which is referred to as 'del/dup' by dbSNP, annotations may differ among databases. Each Figure shows the reference sequence and different annotations including those used by PharmVar and dbSNP. Note that PharmVar annotates these as an insertion or deletion following the 3' alignment rule while dbSNP provides the number of repeat motives present. Note that **Figure 5** also visualizes the annotations used by Moriyama et al when they first described the insertion in 2016 using the 5' alignment mode; this explains why the motif was initially described as GGAGTC opposed to GAGTCG. Finally, HGVS (also using the 3'alignment rule) annotates the insertion and deletion as c.50_55dup (p.Gly_Val18dup) and c.50_55del (p.Gly17_Val 18del), respectively.

Repeat motifs in **Figures 5** and **6** are shown in color with those in red annotated as being inserted or deleted. Nucleotide positions are counted from the ATG start codon; amino acids are shown above the nucleotide sequence.

Figure 5 ***NUDT15*2*** **and** ****6*** **(rs746071566, GAGTCG del/dup | variable number of repeats)**

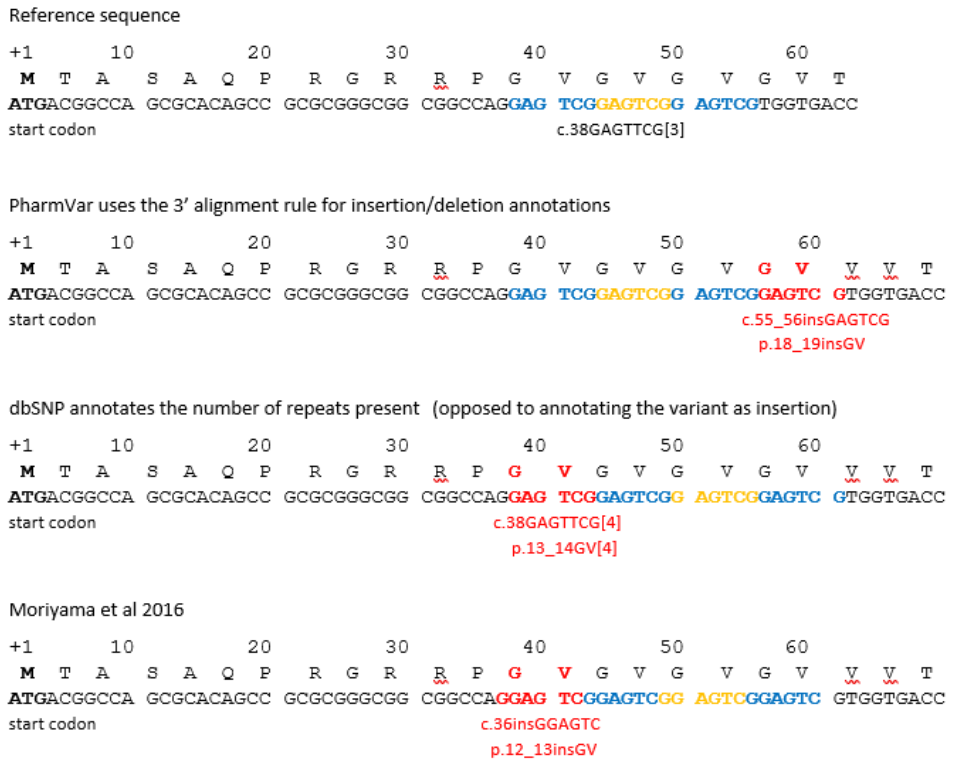


Figure 6 ***NUDT15*9*** **(rs746071566, GAGTCG del/dup | variable number of repeats)**

Function

PharmVar displays 'CPIC clinical allele function' as determined by the Clinical Pharmacogenetic Implementation Consortium (CPIC). The *NUDT15* allele functionality table is available at ClinPGx https://www.clinpgx.org/page/nudt15RefMaterials. This table has been created as part of the CPIC guideline process following the Consortium's consensus-based framework for assigning allele function (Tibben et al, Am J Hum Genet 2025, PMID 41174864). Information about the process can also be found here.

Additional information may have become available since the functionality table has been created. Extrapolation of the activity of an allele to a particular substrate should be done with caution due to potential unknown substrate-specific activities of the allele.

Variant Groupings

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

There are no examples for this feature for *NUDT15.*

Allele Frequencies

NUDT15 allele frequency tables have been developed for CPIC guidelines and are available through ClinPGx at https://www.clinpgx.org/page/nudt15RefMaterials. A list of frequencies including population-specific information and references can be found in the *NUDT15* allele frequency Table in the 'references' tab. These tables are 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.

Frequencies for the SNV can also be obtained by the external links to dbSNP, ClinPGx and gnomAD (see **Figure 1**).

Importantly, **the frequency of a variant does not reflect the frequency of a haplotype (or star allele) if it is part of two or more haplotypes**. Thus, the frequency of c.55_56insGAGTCG represents all haplotypes with this variant.

Allele frequencies: Available allele frequency data are summarized by ClinPGx and PharmFreq for several sources and populations. Both can be easily accessed from PharmVar through the **Allele Page** as described on page 4 and shown in **Figure 2**. Data may not be available for each star allele for each of the data resources.

References

The references provided in the PharmVar database include the citation in which an allele was first published. Additional reference(s) describe important updates supporting the haplotype. 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'.

Changes and Corrections

Changes are recorded in the **Change Log** document.