

PVS1 Null variant (nonsense, frameshift, canonical +/-1 or 2 splice sites, initiation codon, single or multi-exon deletion) in a gene where loss of function (LOF) is a known mechanism of disease
Caveats:

1. Use caution interpreting LOF variants at the extreme 3' end of a gene
2. Use caution with splice variants that are predicted to lead to exon skipping but leave the remainder of the protein intact

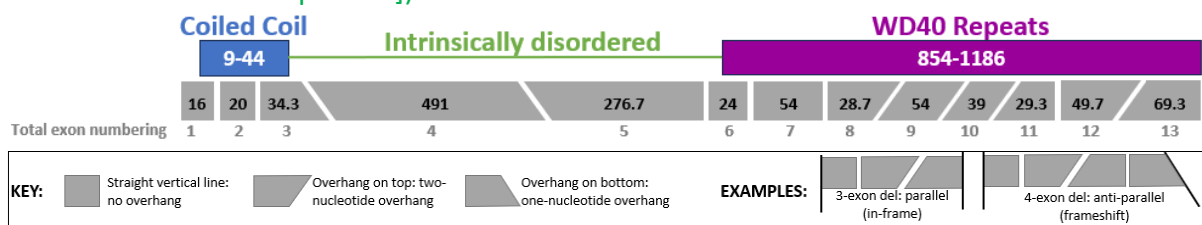
Notes

1. The default RefSeq transcript for nucleotide (c.) annotation is **NM_024675.3/ENST00000261584.8**. Several naturally occurring alternate splicing isoforms have been described¹. Yet, after careful examination, none of them is considered a candidate rescue transcript (very low contribution to overall expression, not coding proteins predicted functional, or both). In keeping with that, we have considered that all presumed LoF events (PVS1 decision tree specifications) occur in biologically relevant transcript(s). (Lopez-Perolio 2019 PMID 30890586).
2. **WD40** beta propeller and Coiled-coil domain (CC) are considered indispensable for PALB2 protein function. PVS1 alterations that are predicted to escape NMD, but that adversely affect the WD40 domain can be granted PVS1 (as opposed to PVS1_Strong as the recommended baseline (Tayoun 2018 PMID 30192042). The following evidence supports this strength change:
 - The WD40 domain interacts with many different protein partners that are involved in the double strand break repair pathway (Reviewed in Ducey 2019 PMID 30638972)
 - Two different C-terminal truncating mutations (c.3549C>A and c.3549C>G) resulting in loss of the last 3 amino acids [p.(Tyr1183Ter)], were identified in *trans* with PALB2 stop-gain variants in three unrelated FA (FA-N) patients (Reid 2007 PMID 17200671)
 - The PALB2 WD40 toroidal structure is “sealed” in the seventh blade by interaction of the C-terminal strand with the incomplete N-terminal blade. The last four residues of PALB2 (Y1183, H1184, Y1185, and S1186) are directly involved in this interaction (molecular Velcro hydrogen bonding) (Oliver 2009 PMID: 28673926). This is the rationale for the clinical relevance of the last 4 amino acids of the protein.
 - Alterations predicted to lead to in frame losses adversely affecting the WD40 structure/function are found in *trans* with LoF PALB2 alterations in Fanconi Anemia patients (Reid 2007 PMID 17200671)
 - Exon 10 donor: c.3113+5G>C (biallelic with c.395delT)
 - Exon 12 donor: c.3350+4A>G (biallelic with c.2393_2394insCT)
 - LoF alterations are rare in GnomAD in all exons
 - GnomAD v2.1 accessed 5/30/2019
 - Total Variants (includes splice acceptor/donor-conservative)
 - 1418 variants

- 336,349 carriers
- LoF Flag (excludes splice acceptor/donor-conservative)
 - 95 variants (6.7%)
 - 239 carriers (.07%)

***PALB2* Exon Map and Domain Map: Use for Single- and Multi-Cassette Exon Losses**

- Number above exon represents the coding exon number
- Overhang on top: a two-nt overhang
- Overhang on bottom: a one-nt overhang
- Parallel lines represent in-frame changes (e.g. EX6_7del is in-frame [5' of exon 6 and 3' of exon 7 are parallel]; however EX6_8del is out-of-frame [5' of exon 6 and 3' of exon 8 are not parallel])



- BRCA1 binds the coiled coil domain
- BRCA2 and RAD51 bind the WD40 domain

PVS1 can be applied as per the below decision tree.

PVS1_Variable[RNA] shall be used for observed splice defects, whether from canonical +/-1,2 positions or other spliceogenic regions (including mid-exonic missense/synonymous variants that cause splice defects) with baseline weight as per the below decision tree. Weight can be further modified based on the quality of the RNA study including consideration of concepts such as:

- Starting material (where patient material is preferable to in vitro minigene)
- Use of NMD inhibitors where translation does occur such as cell lines (Thermann 1998 PMID 9628884, Cater 1995 PMID 7499432)
- Primer design (to make sure it's comprehensive to capture possible multicassette events)
- Method of quantification
 - where e.g. capillary electrophoresis is preferable to estimation by gel band density
 - where SNP analysis is most preferred (where analysis of exonic SNPs and their relative presence in aberrant and WT transcripts is informative)
- Quantification (where complete effects should have increased weight over incomplete effects)
- Specific guidance on the use of RNA evidence in variant assessment is not a gene-specific consideration for *PALB2* at this time, therefore discretion is left to assessors

until further guidance is provided for this general concept from the Sequence Variant Interpretation group.

***PALB2* PVS1/ PVS1_Variable[RNA] Guide (Adapted from PMID 30192042)**

1. PVS1 decision tree, based on ACMG/AMP rationale (Tayoun et al, 2018), introducing some code strength modifications (**upgrades** and **downgrades** as indicated), and a few instances not addressed by Tayoun et al. Specifically: (i) GT-AG variants not affecting coding exons, (ii) +2T>C variants predicted to create de novo functional GC sites, (iii) +2C>T variants targeting native GC sites and, (iv) considering the relevance of the protein domain targeted by ≥1 exon duplications in deciding the PVS1 evidence strength
2. We have considered NM_024675.3 the clinically relevant reference transcript (13 exons, 13 coding exons, start codon located in exon 1, encoding a 1186aa protein-UniProtKB Q86Yc2)
3. We are not aware of any potential rescue transcripts (i.e. for the sake of simplicity, in the decision tree we will not refer to “exon is absent from biologically-relevant transcripts”) Lopez-Perolio et al, 2019.
4. We define two clinically relevant domains (Figure 1): (i) a Coiled-coil (CC) domain spanning residues 10-40 (Song et al, 2018) and (ii) a WD40 domain spanning residues 856-1186 (Oliver et al, 2009).
5. Based on clinical, functional and structural data, we have considered in-frame alterations targeting the CC and WD40 domains as PVS1 (see footnote G for an exception)
6. Based on clinical and structural data, we have considered in-frame alterations targeting FATKIN as PVS1, the only exception being very small in-frame alterations with PROVEAN score suggesting pathogenic, that were considered PVS1_Supporting
7. As far as we know, p.Tyr1183Ter is the last PTC variant known to be pathogenic (Reid et al, 2007)
8. p.Tyr1183Ter pathogenicity suggests that the last four residues (YHYS) of *PALB2*, or at least some of them, as critical for *PALB2* function. This is supported by structural considerations (Oliver et al, 2009).
9. The existence of experimental data (literature and/or personal communication from HBOP VCEP members) supporting the PVS1 weight are denoted by red-underline in the PVS1 decision tree.

Initiation Codon	Closest potential in-frame start codon p.Met296 in Exon 4. If used, critical CC domain will be lost (Song et al, 2018). p.Val132fs, p.Leu253fs clinically relevant FA-N (Reid et al, 2007)		PVS1 (upgraded from PVS1_Moderate)
Nonsense or Frameshift	Predicted to undergo NMD (Introduction of a premature termination codon between p.Asp2 and Leu1101/upstream of c.3301)		PVS1
	Not predicted to undergo NMD (p.Ser1102_Ser1186)	Truncated/alterd region is critical to protein function CC (AA10-40) WD40 (AA856-1183) p.(Tyr1183Ter) in exon 13 is the most C-terminal variant known to be pathogenic	PVS1^A (upgraded from PVS1_Strong)
		Role of region in protein function is unknown nonsense encoding p.His1184X, p.Tyr1185X, or p.Ser1186X	PVS1_Moderate^B
		Frameshift impacting p.Tyr1183 and coding for an alternative/extended C-terminal end	PVS1_Strong^C
		Frameshift NOT impacting p.Tyr1183 and coding for an alternative/extended C-terminal end	PVS1_Supporting^D (downgraded from Moderate)

Deletion (Single exon to full gene)	Full gene deletion		PVS1_Stand Alone (upgraded from PVS1)
	Single to multi exon deletion – Disrupts reading frame and is or is-not predicted to undergo NMD		PVS1
	Single to multi exon deletion preserves reading frame	Altered region is relevant for protein function Targeting exon 2 (CC domain) -and/or- targeting ≥1 WD40 coding exons (exons 6, 7, 8, 9, 10, 11, 12, 13)E	PVS1 (upgraded from PVS1_Strong)
		Functional relevance of the altered region is unknown: Targeting neither CC (exon2), nor WD40 (exon 6-13) Removes ≥10% (>356nt) of coding sequence	PVS1_Strong
		Functional relevance of the altered region is unknown: Targeting neither CC (exon2), nor WD40 (exon 6-13) Removes <10% (>356nt) of coding sequence	PVS1_N/A
Duplication (≥1 exon in size and must be completely contained within gene)	Reading frame disrupted and NMD predicted to occur		PVS1 (if proven in tandem) -or- PVS1_Strong (if presumed in tandem)
	Preserves reading frame, but disrupts the WD40 domain (both breakpoints contained within the domain)		PVS1 (if proven in tandem) -or- PVS1_Strong (if presumed in tandem)
	Preserves reading frame, but not completely contained within the WD40 domain (one breakpoint contained within the domain and one breakpoint outside the domain)		PVS1_Supporting (if proven in tandem) -or- PVS1_N/A (if presumed in tandem)
	Proven not in tandem		PVS1_N/A

Canonical GT-AG +/-1,2 Splice Sites G to non-G at last nucleotide of exon when adjacent intronic sequence is not gtrrgt (where r is a purine) can provide same weight as PVS1 indicates but notched one-level-down in strength	Disrupts reading frame and is predicted to undergo NMD		PVS1 (variants listed in A)
	Disrupts reading frame and is NOT predicted to undergo NMD: Truncated/altered region is critical to protein function: Splicing alteration introduces a PTC upstream of p.His1184		PVS1 (upgraded from PVS1_Strong) (variants listed in B)
	In-frame splicing alteration predicted	Target region critical to protein function (predicted damaging to CC or WD40 domains) ^F	PVS1 (upgraded from PVS1_Strong) (variants listed in C)
		Role of target region unknown: removes ≥10% of coding sequence	PVS1_Strong (variants listed in D)
		Role of target region unknown: removes <10% of coding sequence	PVS1_Moderate ^G c.49-2A> C, T c.49-1G> A , C, T
		Role of target region unknown: inserts <10% coding sequence	PVS1_Supporting ^H c.211+1G> A , C, T c.211+2T> A, C, G
	No splicing alteration predicted	Variant creates a GC site predicted functional	PVS1_N/A c.108+2T> C
		Non-canonical GC improved to GT	PVS1_N/A c.3350+2C> T

A) Two different C-terminal truncating mutations (c.3549C>A, c.3549C>G) introducing the same stop codon [p.(Tyr1183Ter)] were identified in trans with *PALB2* stop-gain variants in three unrelated FA (FA-N) patients (Reid 2007). For that reason, variants introducing stop gains upstream of His1184 are considered PVS1, regardless of NMD status. Pathogenicity of p.(Tyr1183Ter) points to last four residues (YHYS), or at least some of them, as critical for *PALB2* function. Probably, this critical role has a structural basis, as WD40 C-terminal residues interacting with WD40 N-terminal residues seal the toroidal structure in a “molecular Velcro” interaction, and lack of residues YHYS (is predicted to) prevent the closure of the WD40 ring (Oliver et al, 2009). Note that, at present, we do not know if all four residues or only a subset are structurally critical.

B) Depending on the location, these variant will preserve one, two, or three residues at the C-terminal YHYS end. The role of deleted residues is not necessarily critical (i.e. the preserved residues might be sufficient to seal the toroidal structure), and <10% protein removed, qualifying for moderate.

C) These variants will encode proteins with alternative C-terminal ends. We expect these alternative C-terminal ends being unable to recapitulate the “molecular Velcro” interaction with WD40 N-terminal residues, but we cannot discard the possibility of an alternative C-terminal end (partially) mimicking the critical role of the native C-terminal end in WD40 folding. A similar mechanism has been proposed at the N-terminal end of the WD40 domain (upstream sequence mimicking 7D β -strand) to explain the apparent hypomorphic nature of exon 6 skipping (Byrd et al, 2016)

D) Depending on the location, these variants will preserve one, two or three residues at the C-terminal YHYS end (as per B). Compare with B, these variant will code additional residues. We cannot discard the possibility of an alternative C-terminal end mimicking the critical role of YHYS in WD40 folding

E) Exon 6 codes for WD40 β -strand 7D that is critical for WD-40 toroidal folding (Oliver et al, 2009). Yet, Byrd et al 2016 suggests that exon 6 deletion might be to some extent hypomorphic due to some ability of exon 5 residues to mimic WD40 β -strand 7D.

F) Predicted damaging to CC (exon 2) -or- WD40 (exons 6 to 13) domains if ≥ 1 coding exon skipped

G) For variants targeting the exon 2 acceptor site, the predicted/observed outcome Δ (E2p6) skips two residues located at the CC domain p.(Leu17_Lys18del). p.Lys18Arg has no impact on HR efficiency (Boonen et al, 2020). p.Leu17Pro shows HR efficiency reduced >50%. Yet, we do not consider the latter result a formal proof that Leu 17 is critical for HR activity, as proline is a well establish α -helix breaker (Li et al, 1996)

H) For variants targeting the exon 3 donor site, the predicted/observed outcome ∇ (E3q48) introduces 17 novel residues in the protein p.E71delinsGKSRPFTYACFIHFPE. Supporting strength is based on bioinformatics (PROVEAN score -15.84)

Splicing predictions based on SpliceAI Delta Scores >0.2. Individual genetic variants with experimental splicing data (PVS1_O_variable strength) are <u>underlined red</u>							
PVS1 (list A) NMD-Prone		PVS1 (list B) NMD-Escaping Frameshift		PVS1 (list C) NMD-Escaping: In-Frame Function Critical		PVS1_Strong (List D)** NMD-Escaping: In-Frame Function Unknown ≥10%	
c.48+1G>	<u>A</u> , C, T	c.3202-2A>	C, G, T	c.108+1G>	<u>A</u> , C, T	c.212-2A>	C, <u>G</u> , T
c.48+2T>	A, <u>C</u> , G	c.3202-1G>	A, C, T	c.108+2T>	A, G	c.212-1G>	<u>A</u> , C, T
c.49-2A>	G	c.3350+1G>	<u>A</u> , C, T	c.2515-2A>	C, <u>G</u> , T	c.1684+1G>	<u>A</u> , C, T
c.109-2A>	C, <u>G</u> , T	c.3350+2C>	A, G	c.2515-1G>	A, C, <u>I</u> *	c.1684+2T>	A, C, G
c.109-1G>	A, C, T	c.3351-2A>	C, G, T	c.2586+1G>	<u>A</u> , C, T*	c.1685-2A>	<u>C</u> , <u>G</u> , T
c.1685-1G>	A	c.3351-1G>	A, C, T	c.2586+2T>	A, C, G*	c.1685-1G>	C, T
c.2514+1G>	A, C, T			c.2748+1G>	A, C, <u>I</u> *		
c.2514+2T>	A, <u>C</u> , G			c.2748+2T>	A, C, <u>G</u>		
c.2587-2A>	<u>C</u> , <u>G</u> , T			c.2835-2A>	C, G, T		
c.2587-1G>	A, C, T			c.2835-1G>	<u>C</u> , T		
c.2749-2A>	C, G, T			c.2996+2T>	A, C, G		
c.2749-1G>	A, C, <u>I</u>			c.2996+1G>	A, C, T		
c.2834+1G>	<u>A</u> , C, T			c.2997-2A>	G, T		
c.2834+2T>	A, <u>C</u> , G			c.3113+1G>	A, C, T		
c.2835-1G>	A			c.3113+2T>	A, C, G		
c.2997-2A>	<u>C</u>						
c.2997-1G>	<u>A</u> , C, T						
c.3114-2A>	C, G, T						
c.3114-1G>	A, C, T						
c.3201+1G>	<u>A</u> , C, T						
c.3201+2T>	A, C, G						

* Variants targeting exon 6 donor and acceptor sites are predicted to cause Δ(E6) p.(Thr839_Lys862del). Exon 6 codes for WD40 β-strand 7D that is critical for WD-40 toroidal folding (Oliver et al, 2009). Yet, Byrd et al 2016 suggests that exon 6 deletion might be to some extent hypomorphic due to some ability of exon 5 residues to mimic WD40 β-strand 7D.

**Due to its size, splicing analyses involving *PALB2* exons 4 (1437nt) and/or exon 5 (830nt) are challenging. Indeed, we are aware of conflicting data regarding the impact of exon 4 acceptor/donor site variants on splicing. In agreement with ACMG/AMP recommendations (lowest PVS1 strength among the different scenarios, Abou Tayoun et al, 2018), we have assigned a strong evidence strength to exon 4 acceptor/donor site variants based on Δ(E4) (No-FS, not targeting critical domains, >10% of the coding sequence). We have assigned a strong evidence strength to exon 5 acceptor site variants based on Δ(5Ep510) (No-FS, not targeting critical domains, >10% of the coding sequence)