

Article

Splicing Analysis of 16 *PALB2* ClinVar Variants by Minigene Assays: Identification of Six Likely Pathogenic Variants

Alberto Valenzuela-Palomo ¹, Lara Sanoguera-Miralles ¹, Elena Bueno-Martínez ¹, Ada Esteban-Sánchez ², Inés Llinares-Burguet ¹, Alicia García-Álvarez ¹, Pedro Pérez-Segura ², Susana Gómez-Barrero ³, Miguel de la Hoya ² and Eladio A. Velasco-Sampedro ^{1,*}

- ¹ Splicing and Genetic Susceptibility to Cancer, Unidad de Excelencia Instituto de Biología y Genética Molecular, Consejo Superior de Investigaciones Científicas (CSIC-UVA), 47003 Valladolid, Spain
- ² Molecular Oncology Laboratory, Hospital Clínico San Carlos, IdISSC (Instituto de Investigación Sanitaria del Hospital Clínico San Carlos), 28040 Madrid, Spain
- ³ Facultad de Ciencias de la Salud, Universidad Alfonso X “El Sabio”, Avda. de la Universidad 1, Villanueva de la Cañada, 28691 Madrid, Spain
- * Correspondence: eavelsam@ibgm.uva.es



Citation: Valenzuela-Palomo, A.; Sanoguera-Miralles, L.; Bueno-Martínez, E.; Esteban-Sánchez, A.; Llinares-Burguet, I.; García-Álvarez, A.; Pérez-Segura, P.; Gómez-Barrero, S.; de la Hoya, M.; Velasco-Sampedro, E.A. Splicing Analysis of 16 *PALB2* ClinVar Variants by Minigene Assays: Identification of Six Likely Pathogenic Variants. *Cancers* **2022**, *14*, 4541. <https://doi.org/10.3390/cancers14184541>

Academic Editor: Christian Singer

Received: 5 August 2022

Accepted: 15 September 2022

Published: 19 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Simple Summary: *PALB2* pathogenic variants confer high risk of breast cancer. Here, we have analyzed the impact of *PALB2* variants on splicing, a gene expression step that removes introns to form the mature messenger RNA. This process is performed by the splicing machinery through the recognition of specific sequences, namely the 3' and 5' splice sites, which determine the exon ends. Variants at these sequences may trigger anomalous splicing and aberrant transcripts that may be associated with a disease. To test the impact of variants on splicing, we used a biotechnological tool called minigene, which replicates, at small-scale, the human gene of interest. Thus, we checked 16 *PALB2* variants at the intron/exon boundaries using the minigene mgPALB2_ex1-3. We found that twelve variants disrupted splicing, six of which could be classified as likely pathogenic. These results facilitate the clinical management of carrier patients and families since they may benefit from tailored prevention protocols and therapies.

Abstract: *PALB2* loss-of-function variants are associated with significant increased risk of breast cancer as well as other types of tumors. Likewise, splicing disruptions are a common mechanism of disease susceptibility. Indeed, we previously showed, by minigene assays, that 35 out of 42 *PALB2* variants impaired splicing. Taking advantage of one of these constructs (mgPALB2_ex1-3), we proceeded to analyze other variants at exons 1 to 3 reported at the ClinVar database. Thirty-one variants were bioinformatically analyzed with MaxEntScan and SpliceAI. Then, 16 variants were selected for subsequent RNA assays. We identified a total of 12 spliceogenic variants, 11 of which did not produce any trace of the expected minigene full-length transcript. Interestingly, variant c.49-1G > A mimicked previous outcomes in patient RNA (transcript $\Delta(E2p6)$), supporting the reproducibility of the minigene approach. A total of eight variant-induced transcripts were characterized, three of which ($\Delta(E1q17)$, $\Delta(E3p11)$, and $\Delta(E3)$) were predicted to introduce a premature termination codon and to undergo nonsense-mediated decay, and five ($\nabla(E1q9)$, $\Delta(E2p6)$, $\Delta(E2)$, $\nabla(E3q48)$ -a, and $\nabla(E3q48)$ -b) maintained the reading frame. According to an ACMG/AMP (American College of Medical Genetics and Genomics/Association for Molecular Pathology)-based classification scheme, which integrates mgPALB2 data, six *PALB2* variants were classified as pathogenic/likely pathogenic, five as VUS, and five as likely benign. Furthermore, five $\pm 1,2$ variants were catalogued as VUS because they produced significant proportions of in-frame transcripts of unknown impact on protein function.

Keywords: hereditary breast cancer; cancer susceptibility genes; *PALB2*; aberrant splicing; functional assay; minigenes; clinical interpretation

1. Introduction

Hereditary breast cancer (BC) is a highly heterogeneous genetic disease, in which more than 20 genes of the DNA repair pathway have been proposed as breast cancer susceptibility genes [1]. Historically, genetic testing was focused on the main BC genes *BRCA1* and *BRCA2* by different methods [2,3]. The development of next-generation sequencing (NGS) enabled the development of panels of cancer predisposing genes and the simultaneous sequencing of multiple genes, thus boosting efficiency and cost-effectiveness. Recently, two large-scale sequencing studies, which sequenced a panel of breast cancer genes in more than 170,000 women, refined the BC/OC genetic predisposition spectrum [4,5]. At least eight genes were found to be significantly associated with breast cancer risk: *BRCA1* (MIM#113705), *BRCA2* (MIM#600185), *ATM* (MIM#607585), *BARD1* (MIM#601593), *CHEK2* (MIM#604373), *PALB2* (MIM#610355), *RAD51C* (MIM#602774), and *RAD51D* (MIM#602954) [4–6]. Biallelic loss-of-function variants of *PALB2* (also known as *FANCN*) and other BC susceptibility genes, such as *BRCA2*, *RAD51C*, and *BRCA1*, cause Fanconi anemia [7], which is characterized by a high genomic instability and increased cancer predisposition.

The partner and localizer of *BRCA2* (*PALB2*) interacts with *BRCA1* and *BRCA2* and is implicated in repair of double-strand DNA breaks by homologous recombination. While *BRCA1* recruits *PALB2* at the sites of DNA damage, *PALB2* stabilizes *BRCA2* during formation of the *RAD51* nucleoprotein filament [8,9]. *PALB2* loss-of-function variants confer high risk of developing breast cancer (BC) as well as other types of cancers [10–12]. Two recent reports have shown that *PALB2* protein truncating variants are associated with a significantly increased risk of breast cancer (BC relative risk 3.83 and 5.02, respectively) and accounts for 9.5–10.1% of the protein truncating variants of the eight core BC genes mentioned above (0.39–0.56% of all BC cases) [4,5]. Furthermore, associations with estrogen-negative and triple-negative BC are even higher, with relative risks of 7.35 (4.25–12.72) and 10.36 (6.42–16.71), respectively [4]. Hence, *PALB2* belongs to the high-risk category of BC susceptibility genes together with *BRCA1* and *BRCA2*.

On the other hand, a deleterious effect on gene function cannot be assigned for a relevant proportion of variants detected in patients, the so-called variants of uncertain clinical significance (VUS) [13]. In the case of *PALB2*, VUS frequency is approximately four times greater than that of pathogenic variants [5]. Consequently, they represent a challenge in genetic counselling because cancer risk assessment in VUS carriers is only based on cancer family history [14]. Apart from protein translation, there are other gene expression steps that may be targeted by disease-causing variants, such as transcription, splicing, as well as other post-transcriptional mechanisms [15–18]. Functional studies of these processes provide key information for the clinical interpretation of VUS.

The splicing process is controlled by a large collection of splicing factors and cis-acting sequences that include: the 5' or donor (GT) and the 3' or acceptor (AG) splice sites (5'SS and 3'SS, respectively), the polypyrimidine tract and the branch point, as well as exonic and intronic elements that promote (enhancers) or repress (silencers) exon recognition [19]. All these motifs may be targets of splicing-disrupting mutations (spliceogenic variants) so that an unexpectedly large fraction of variants can actually induce splicing anomalies [17, 20,21]. In fact, it was estimated that about 62% of pathogenic variants impair splicing [22]. Interestingly, several cancer susceptibility genes, such as *MLH1*, *MSH2*, and *PMS2*, are enriched in spliceogenic variants [23].

We have focused our efforts on the study of the impact of genetic variants on the splicing of the BC genes by minigene assays, by which we found a large proportion of spliceogenic variants [24]. We comprehensively analyzed by minigene assays several BC susceptibility genes, such as *BRCA2* [25], along with *RAD51C*, *RAD51D*, *PALB2*, and *ATM* within the framework of the European Project BRIDGES (Breast Cancer Risk after Diagnostic Gene Sequencing; <https://bridges-research.eu/>, accessed on 12 July 2022) [26–29].

In a recent work, we studied the *PALB2* gene and tested, in three different minigenes, 42 candidate BRIDGES variants [28], 35 of which disrupted splicing, with 23 of them being

classified as pathogenic or likely pathogenic, demonstrating the usefulness of minigenes for RNA assays and clinical interpretation of variants. Moreover, these constructs are highly valuable since any other potentially spliceogenic variants of the gene of interest can be so assayed.

Taking advantage of the minigene mgPALB2_ex1-3, we selected 31 ClinVar splice-site variants located at exons 1–3 (<https://www.ncbi.nlm.nih.gov/clinvar/>, accessed on 3 August 2021) to carry out splicing assays of the potentially damaging variants.

2. Materials and Methods

2.1. Variant and Transcript Annotations

The analysis of the ClinVar data identified a total of 31 variants at exons 1, 2, and 3 and flanking intronic sequences located at the *PALB2* 5' and 3' splice sites (5'SS and 3'SS, respectively), defined for the purpose of the present study as: (i) intron/exon (IVS –10 to –1/2 nt) boundaries (3'SS) and (ii) exon/intron (2 nt/IVS +1 to +6) boundaries (5'SS). Variants, transcripts and predicted protein products were described according to the Human Genome Variation Society (HGVS) guidelines (<https://varnomen.hgvs.org/>, accessed on 1 June 2022), using the Ensembl reference transcript ID ENSG00000083093 (Genbank NM_024675.4). We also annotated splicing events according to a former shortened description [30,31].

2.2. Bioinformatics: Databases and In Silico Studies

All the 31 *PALB2* ClinVar splice-site variants at exons 1, 2, and 3 were analyzed with MaxEntScan (MES) http://hollywood.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html, accessed on 3 August 2021) to identify potentially spliceogenic variants [32]. Candidate variants were analyzed under the following criteria [28]: (i) splice site disruption at the $\pm 1,2$ (AG/GT) positions and (ii) relevant MES score changes ($\geq 15\%$) [33,34]. All variants were further evaluated with the splice-site predictor SpliceAI (<https://spliceailookup.broadinstitute.org/>, accessed on 3 August 2021) [35]. SpliceAI outputs were helpful to predict putative splicing outcomes based on a “two score” approach (e.g., donor loss + acceptor loss predicts exon skipping, while donor loss + donor gain predicts a donor shift). SpliceAI parameters were as follows: genome version hg38; score type raw; max distance 10,000 nt; Illumina’s pre-computed scores yes. Scores range 0–1 is interpreted as probability of impact on splicing with the following cutoffs: 0.2–0.49 (high recall), 0.5–0.79 (recommended), and >0.8 (high precision). SpliceAI was not herein used to filter out variants. On basis of the MES outcomes, we decided to carry out the subsequent splicing assays for 16 potentially spliceogenic variants.

2.3. Minigene Construction and Mutagenesis

The minigene mgPALB2_ex1–3 was built in the splicing vector pSAD as previously reported [28,36,37] (Figure 1a). In brief, this construct contains a 974 bp insert (final minigene size: 5068 bp) that includes exons 1 to 3. This construct has the special feature of a chimeric exon 1 composed of vector exon 1 and *PALB2* exon 1 so that 5'SS variants of exon 1 can be tested [28].

The wild-type minigene was used as template to generate 16 DNA ClinVar variants by site-directed mutagenesis with the QuikChange Lightning Kit (Agilent, Santa Clara, CA, USA) (Table 1). All mutant constructs were confirmed by sequencing (Macrogen, Madrid, Spain).

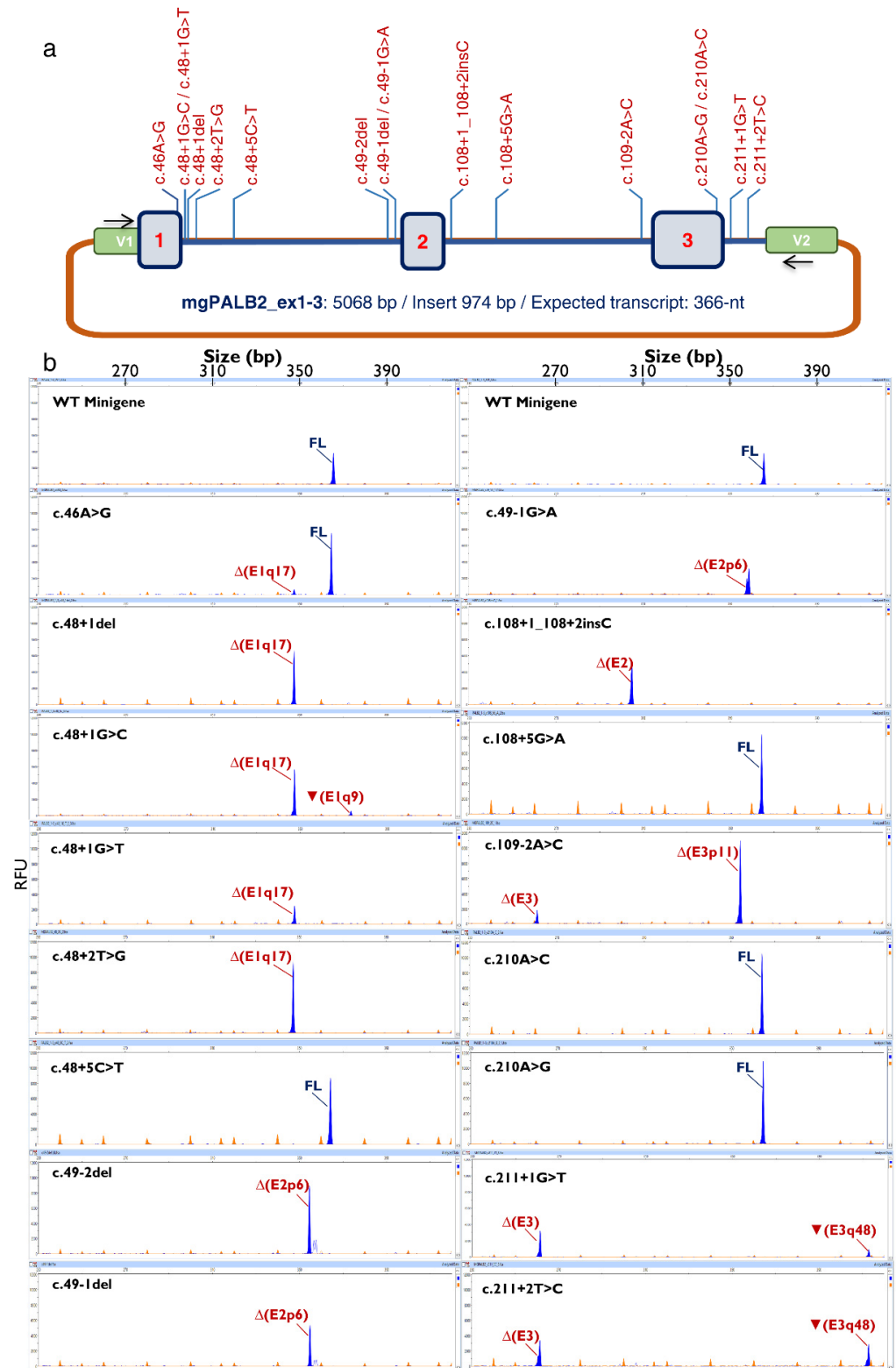


Figure 1. Minigene splicing assays of selected *PALB2* variants. (a) Map of variants in the minigene mgPALB2_ex1–3. V1 and V2 are the vector exons while variants are shown in red above the minigene construct (b) Fluorescent fragment analysis of 16 variants. The electropherogram of the wild-type minigene is shown on the top of each column. FAM-labelled products (blue peaks) were run together with LIZ-600 (orange peaks) as size standard (FL, minigene full-length transcript). The x-axis indicates size in bp (electropherograms on the top) and the y-axis represents relative fluorescence units (RFU).

Table 1. Mutagenesis primers of *PALB2* variants.

Variant	Exon/Intron	Primers (5'→3')
c.46A > G	Ex1	CTGTGAGGAGAAGGAAGAGGTGCCGGGGTGCAGGAAGGG CCCTCCCACACCCCGGCACCTCTTCCTTCTCCTCACAG
c.48 + 1del	IVS1	AGCTGTGAGGAGAAGGAAAAGGGCCGGGGTGCAGGAAG CTTCCCACACCCCGGCCCTTTTCCTTCTCCTCACAGCT
c.48 + 1G > C	IVS1	GCTGTGAGGAGAAGGAAAAGCTGCCGGGGTGCAGGAAG CCTTCCCACACCCCGGCAGCTTTTCCTTCTCCTCACAGC
c.48 + 1G > T	IVS1	TCAGCTGTGAGGAGAAGGAAAAGTTGCCGGGGTGCAGGA TCCCGCACACCCCGCAACTTTTCCTTCTCCTCACAGCTGA
c.48 + 2T > G	IVS1	CAGCTGTGAGGAGAAGGAAAAGTGCCGGGGTGCAGGAAG CTTCCCACACCCCGGCACCTTTTCCTTCTCCTCACAGCTG
c.48 + 5C > T	IVS1	GAGGAGAAGGAAAAGGTGCTGGGGTGCAGGAAGGGCGGA TCCGCCCTTCCCACACCCCGCACCTTTTCCTTCTCCTC
c.49-2del	IVS1	TGCCAGTATTGTTGGTGTTTTTCTTCTTCCGTTAAAGGA TCCTTAAACGGAAGAAGAAAAACACCAACAATACTGGGCA
c.49-1del	IVS1	TGCCAGTATTGTTGGTGTTTTTCTTCTTCCATTAAAGGA TCCTTAAATGGAAGAAGAAAAACACCAACAATACTGGGCA
c.49-1G > A	IVS1	TTCTTCCAATTAAGGAGAAATTAGCATTCTGAAAAGGG CCTTTTCAAGAATGCTAATTCTCCTTTAATTGGAAGAA CCTCAGGCTAAGTGAATCGTATTCTCAAATTAAGGTGTT AACACCTTAATTTGAGAATACGATTCACTTAGCCTGAAGG
c.108 + 1_108 + 2insC	IVS2	TAGCCCGCCTCAGGTAATGAATCGTATTCTCAAATTAA TTAATTTGAGAATACGATTCACTTAGCCTGAAGGCGGCTA
c.108 + 5G > A	IVS2	TTTGTCTCTCTCGCGTCCCAAAGAGCTGAAAAGATTAA TTAATCTTTTTCAGCTCTTTGGGCACGCGAGAGGAGACAAA
c.109-2A > C	IVS2	CCGCAGCTAAAACACTCGGGTAAATCTAGACCATTCACTT AAGTGAATGGTCTAGATTACCGAGTGTTTTAGCTGCGG
c.210A > G	Ex3	CGCAGCTAAAACACTCCGGTAAATCTAGACCATTCACTTA TAAGTGAATGGTCTAGATTACCGAGTGTTTTAGCTGCG
c.210A > C	Ex3	CCGCAGCTAAAACACTCAGTAAATCTAGACCATTCACTT AAGTGAATGGTCTAGATTAACTGAGTGTTTTAGCTGCGG
c.211 + 1G > T	IVS3	ACCGCAGCTAAAACACTCAGGCAAATCTAGACCATTCACT AGTGAATGGTCTAGATTGCTGAGTGTTTTAGCTGCGGT
c.211 + 2T > C	IVS3	ACCGCAGCTAAAACACTCAGGCAAATCTAGACCATTCACT AGTGAATGGTCTAGATTGCTGAGTGTTTTAGCTGCGGT

2.4. Splicing Functional Assays

Approximately 2×10^5 MCF-7 and MDA-MB-231 cells were seeded in four-well plates (Nunc, Roskilde, Denmark) to grow up to 90% confluency in 0.5 mL of medium (MEME, 10% fetal bovine serum, 2 mM glutamine, 1% non-essential amino acids, and 1% penicillin/streptomycin). Then, using a standard protocol of transfection, MCF-7 cells were transfected with either the wild-type or the mutant minigenes. To inhibit nonsense-mediated decay (NMD), cells were incubated with cycloheximide 300 µg/mL (Sigma-Aldrich, St. Louis, MO) for 4 h. RNA was extracted after 48 h and purified with the Genematrix Universal RNA purification Kit (EURx, Gdansk, Poland) with on-column DNase I digestion to degrade genomic DNA that could interfere in RT-PCR. Retrotran-

scription was carried out with specific primers of exons V1 and V2 of the pSAD[®] vector as previously described [26,28,38]. The expected size of the minigene mgPALB2_ex1–3 full-length (mgFL) transcript was 366 nt. To estimate the relative abundance of all transcripts, semi-quantitative fluorescent RT-PCRs (26 cycles) were performed with pSPL3_RT-FW and FAM-RTpSAD-RV. FAM-labeled products were run with LIZ-600 Size Standards at the MacroGen facility and analyzed with the Peak Scanner software V1.0 (Life Technologies, Carlsbad, CA, USA). Three independent experiments for each variant were carried out to calculate the average relative proportions of each transcript and the corresponding standard deviations.

2.5. Clinical Classification of PALB2 Variants

We performed a tentative clinical classification of 16 *PALB2* variants according to ACMG/AMP-based guidelines. We used a Bayesian-ACMG/AMP point system that shows higher plasticity in combining different ACMG/AMP criteria and strengths of evidence [39,40]. Point-based variant classification categories are defined as follows: pathogenic (P) $\geq +10$; likely pathogenic (LP) +6 to +9; variant of uncertain significance (VUS) 0 to +5; likely benign (LB) -1 to -6 ; and benign (B) ≤ -7 . The mgPALB2 read-outs were included in the classification system as observable PVS1_O or BP7_O evidence codes of variable strength depending on the splicing outcome (P, supporting (± 1 point); M, moderate (± 2); S, strong (± 4); VS, very strong (± 8)) [29,41]. This score is deduced from the presumed impact of all the transcripts generated by a particular variant. To interpret variants producing ≥ 2 transcripts, we applied the following rules: (i) decode/separate mgPALB2-readouts into individual components (transcripts); (ii) apply ACMG/AMP-based evidence levels to each individual transcript; and (iii) deduce a global PVS1_O (or BP7_O) code strength based on the relative contribution of individual transcripts to the overall expression. Thus, if pathogenic (or benign) supporting transcripts contribute $\geq 90\%$ to the overall expression level, PVS1_O (or BP7_O) codes are applied. If different transcripts support different pathogenic evidence strengths, the lowest strength contributing $>10\%$ to the overall expression is selected as overall evidence strength. At present, $\geq 90\%$ and $\geq 10\%$ cut-offs of the overall mgPALB2 expression are merely operational. Recently, we already used a similar approach to deal with those *PALB2/ATM/RAD51C* minigene readouts that yielded several transcripts per variant [28,29,41].

We considered that functional splicing data (PVS1_O/BP7_O) override predictive splicing codes PVS1 (GT-AG splice site variants) and PP3/BP4 (non-GT-AG variants) so that the latter does not contribute to our variant classification. Otherwise, internal inconsistencies would arise in the ACMG/AMP classification system (e.g., IVS + 1 and IVS + 5 variants with identical splicing impact would score very differently). Furthermore, the ACMG/AMP system implicitly assumes that each piece of evidence contributing to the final classification is independent, which is an assumption barely met by predictive and functional splicing codes, as most splicing analyses (including our mgPALB2 ones) are performed for bioinformatically pre-selected variants. These issues have been extensively discussed elsewhere [28,29,41]. The rarity code PM2 was considered with allele frequency $\leq 0.001\%$ at gnomADv2.1.1 (<https://gnomad.broadinstitute.org>; accessed on 21 June 2022) decreasing, so PM2 evidence strength to “supporting” as previously reported [29]. For *PALB2* variants absent on gnomADv2.1.1, the number of interrogated alleles (allele number) was determined using data of the closest available SNP (≤ 5 nt apart from the variant of interest).

Since no specific *PALB2* recommendations exist for missense variants, we applied general recommendations recently published by ClinGen SVI [42]. Specifically, REVEL ≥ 0.8 supporting pathogenic (moderate strength), REVEL ≤ 0.4 supporting benign (moderate strength), and $0.4 < \text{REVEL} < 0.8$ supporting neither pathogenic nor benign. To obtain REVEL scores, we ran the built-in Ensembl Variant Effect Predictor (www.ensembl.org/Tools/VEP; accessed on 1 June 2022).

3. Results and Discussion

3.1. Bioinformatics Analysis of ClinVar Variants

A total of 31 ClinVar variants comprising 66 submissions to the ClinVar database were chosen from the intron/exon boundaries of *PALB2* exons 1 to 3. All the 31 variants were bioinformatically analyzed with MaxEntScan, 16 of which met the criteria indicated in Materials and Methods, so they were selected for subsequent minigene RNA assays (Table 2). Twelve of these variants targeted splice donor sites, while the remaining four targeted acceptor sites. The 16 MES-selected variants were also analyzed by SpliceAI (Table 3). Changes c.48 + 5C > T and c.108 + 5G > A were not predicted to affect splicing.

3.2. Minigene Splicing Assays of Candidate Variants

The 16 variants were introduced into the wild-type minigene mgPALB2ex1–3 by site-directed mutagenesis and assayed in MCF-7 cells. Twelve variants impaired splicing, eleven of which showed a total impact, as the minigene full-length transcript was absent (Figure 1b, Table 4). These 11 variants affected the AG/GT ($\pm 1,2$) dinucleotides of the 3'SS and 5'SS and showed the strongest impacts on MES scores (Table 2). In contrast, partial splicing anomalies were found in variants at other splice-site positions. Actually, we noticed weak or no splicing effects for those variants involving the antepenultimate nt of exon 1 (c.46A > G), +5 nt of introns 1 and 2 (c.48 + 5C > T and c.108 + 5G > A, respectively), and penultimate nt of exon 3 (c.210A > G and c.210A > C). Splicing disruptions of variants at positions other than $\pm 1,2$ are particularly difficult to predict, as we have pointed out in previous reports [28]. In this study, leaky variants (those that generate non-negligible levels of full-length transcripts) were associated with moderate reductions in the MES score (−19.3% to −29.8%, Table 2). To test the reproducibility of the minigene assay in different cell lines, four variants (c.46A > G, c.48 + 1G > T, c.49-2del, and c.210A > C) were also assayed in the triple-negative breast cancer cell line MDA-MB-231. All the variants mimicked the splicing patterns characterized in MCF-7 cells (Figure 2). Moreover, variant c.49-1G > A replicated the splicing outcomes formerly characterized in patient RNA [43], confirming the reproducibility of minigene assays (Table 3).

A total of eight different anomalous transcripts were characterized (Table 4, Figure 3). Three transcripts ($\Delta(E1q17)$, $\Delta(E3p11)$, $\Delta(E3)$) are predicted to introduce a premature termination codon triggering the NMD surveillance mechanism (PTC-NMD) [28,44], while the remaining five isoforms, including two versions of $\nabla(E3q48)$ (a and b), maintained the reading frame. Minigene assays, together with fluorescent fragment analysis, displayed simplicity, robustness, high resolution, and sensitivity. This strategy allowed us to detect splicing alterations introducing small size changes (i.e., insertion of 9 nt or $\nabla(E1q9)$ or deletion of 6 nt, $\Delta(E2p6)$) as well as some transcripts representing a minor contribution to the overall mgPALB2 expression (i.e., c.46A > G: $\Delta(E1q17)$, 7.5%; c.48 + 1G > C: $\nabla(E1q9)$, 9.2%; Table 4).

Table 2. Bioinformatics analysis of *PALB2* variants with Max Ent Scan.

<i>PALB2</i> VARIANTS ¹	# ClinVar Records ²	EXON/INTRON	MES wt	MES m ut	MES Score Change ³	Cryptic/De novo Splice Sites ⁴
c.46A > G	1	Exon 1	5.74	4.05	−29.4%	
c.48 + 1del	1	IVS1	5.74	−12.45	−316.9%	
c.48 + 1G > C	3	IVS1	5.74	−2.53	−144.1%	
c.48 + 1G > T	1	IVS1	5.74	−2.76	−148.1%	
c.48 + 2T > G	1	IVS1	5.74	−1.9	−133.1%	
c.48 + 5C > T	3	IVS1	5.74	4.05	−29.4%	
c.48 + 6G > C	1	IVS1	5.74	5.96	+3.8%	
c.49-2del	1	IVS1	9.28	1.16	−87.5%	3'SS: 5.47 26 nt upstream 3'SS: 8.65, 6 nt downstream
c.49-1del	1	IVS1	9.28	−7.59	−181.8%	3'SS: 5.47 26 nt upstream 3'SS: 8.76, 6 nt downstream
c.49-1G > A	1	IVS1	9.28	0.53	−94.3%	3'SS: 5.47 26 nt upstream 3'SS: 7.53, 6 nt downstream
c.50T > G	2	Exon 2	9.28	9.02	−2.8%	3'SS: 5.47 26 nt upstream
c.50dup	1	Exon 2	9.28	9.25	−0.3%	3'SS: 5.47 26 nt upstream
c.106C > T	4	Exon 2	10.86	9.66	−11%	de novo 5'SS: 3.81 4 nt downstream
c.108G > A	1	Exon 2	10.86	10.08	−7.2%	
c.108 + 1_108 + 2insC	1	IVS2	10.86	−4.16	−138.3%	
c.108 + 4A > G	3	IVS2	10.86	10.28	−5.3%	
c.108 + 5G > A	1	IVS2	10.86	8.76	−19.3%	
c.108 + 6T > C	1	IVS2	10.86	9.88	−9%	
c.109-2A > C	3	IVS2	10.06	2.02	−79.9%	3'SS: 5.47, 11 nt downstream
c.109C > G	2	Exon 3	10.06	11.86	+11.8%	
c.109C > T	5	Exon 3	10.06	9.82	−2.4%	
c.109C > A	7	Exon3	10.06	10.18	+1.2%	
c.110G > T	2	Exon 3	10.06	9.69	−3.7%	
c.110G > A	9	Exon 3	10.06	9.64	−4.2%	
c.111T > C	3	Exon 3	10.06	10.3	+2.4%	
c.210A > G	2	Exon 3	8.76	6.15	−29.8%	5'SS: 7.88, 48 nt downstream
c.210A > C	1	Exon 3	8.76	6.89	−21.3%	5'SS: 7.88, 48 nt downstream
c.211 + 1G > T	1	IVS3	8.76	0.26	−97.0%	5'SS: 7.88, 48 nt downstream
c.211 + 2T > C	1	IVS3	8.76	1.01	−88.5%	5'SS: 7.88, 48 nt downstream
c.211 + 4A > G	1	IVS3	8.76	7.25	−17.2%	5'SS: 7.88, 48 nt downstream
c.211 + 6T > A	1	IVS3	8.76	8.59	−1.9%	5'SS: 7.88, 48 nt downstream

¹ Selected variants are shown in red. ² #, number of ClinVar Records; ³ MES score changes ($\Delta\%$): mutant (mut) vs. wild type (wt). ⁴ Positions of cryptic/de novo splice sites are relative to the corresponding canonical splice site.

Table 3. SpliceAI predictions, mgPALB2 read-outs, and experimental splicing data in carriers.

PALB2 Variants ¹	SpliceAI ²				Predicted Splicing Outcome	mgPALB2 Read-out (>10%)	Experimental Data in RNA from Carriers
	AL (>20%)	DL (>20%)	AG (>20%)	DG (>20%)			
c.46A > G	-	-	-	0.27 (+11)	-	mgFL (92.5%); Δ(E1q17) (7.5%)	Normal [45]
c.48 + 1del	-	0.94 (−2)	-	0.25 (−171)	Δ(E1q169)	Δ(E1q17) (100%)	
c.48 + 1G > C	-	0.94 (−1)	-	0.45 (+8)	▼(E1q9)	Δ(E1q17) (90.8%)/▼(E1q9) (9.2%)	
c.48 + 1G > T	-	0.94 (−1)	-	0.30 (+8)	▼(E1q9)	Δ(E1q17) (100%)	
c.48 + 2 T > G	-	0.94 (−2)	-	-	-	Δ(E1q17) (100%)	
c.48 + 5C > T	-	-	-	-	-	FL (100%)	
c.49-2del	0.98 (+1)	-	0.74 (+7)	-	Δ(E2p6)	Δ(E2p6) (100%)	
c.49-1del	0.98 (+1)	-	0.79 (+6)	-	Δ(E2p6)	Δ(E2p6) (100%)	
c.49-1G > A	0.98 (+1)	-	0.65 (+7)	-	Δ(E2p6)	Δ(E2p6) (100%)	Δ(E2p6) [43]
c.108 + 1_108 + 2insC	0.96 (−61)	0.99 (−2)	-	-	Δ(E2)	Δ(E2) (100%)	
c.108 + 5G > A	-	-	-	-	-	FL (100%)	
c.109-2A > C	1 (+2)	0.33 (+104)	0.55(+13)	-	Δ(E3p11)	Δ(E3)	Δ(E3p11) (85%)/Δ(E3) (15%)
c.210A > G	-	-	-	0.25 (+49)	-	FL (100%)	
c.210A > C	-	-	-	0.28(+49)	-	FL (100%)	
c.211 + 1G > T	0.42 (−103)	1 (−1)	-	0.47 (+47)	▼(E3q48)	Δ(E3)	Δ(E3) (73.3%)/▼(E3q48a) (26.7%)
c.211 + 2T > C	0.32 (−104)	0.99 (−2)	-	0.59 (+46)	▼(E3q48)	Δ(E3)	Δ(E3) (48.1%)/▼(E3q48b) (51.9%)

¹ Bold-highlighted variants for which SpliceAI predictions are, in our opinion, accurate, rightly predicting the exact experimental read-out. ² SpliceAI parameters were as follows (genome version hg38; score type raw; max distance 10,000 nt; Illumina's pre-computed scores yes). Acceptor loss (AL), donor loss (DL), acceptor gain (AG), and donor gain (DG) scores (and positions) are shown. Color codes indicate scores in the 20–49 (high recall), 50–79 (recommended), and 80–100 (high precision) ranges, as per Illumina's specifications. Scores < 20% are not shown. SpliceAI positions are annotated as (−) if upstream of the variant or (+) if downstream. Yet, the SpliceAI annotation (relative to the forward strand) becomes confusing for genes located in the antisense strand, such as *PALB2*. For that reason, in the present table, upstream (+) and downstream (−) positions are not shown as per SpliceAI but relative to *PALB2* coding strand. A minimum of two scores above the threshold are required to predict a specific aberrant outcome (e.g., for a variant damaging a donor site, an acceptor loss scoring at the right position predicts exon skipping, while a donor gain will predict use of a cryptic/de novo site). Since SpliceAI predictions for c.46A >G, c.210A >G, and c.210A >C do not fulfill the two score approach, they were considered negative (no splicing alteration) and therefore accurate.

Table 4. Splicing outcomes of *PALB2* variants.

Variant (HGVS) ¹	Bioinformatics Summary (MES) ²	Canonical Transcript	PTC-Transcripts ³	In-Frame Transcripts ⁴
Wild type mgPB2_ex1–3		100%		
c.46A > G	(↓) 5'SS (5.74→4.05)	92.5% ± 0.1%	Δ(E1q17) (7.5% ± 0.1%)	
c.48 + 1del	(−) 5'SS (5.74→−12.45)	-	Δ(E1q17) (100%)	
c.48 + 1G > C	(−) 5'SS (5.74→−2.53)	-	Δ(E1q17) (90.8% ± 0.6%)	▼(E1q9) (9.2% ± 0.6%)
c.48 + 1G > T	(−) 5'SS (5.74→−2.76)	-	Δ(E1q17) (100%)	
c.48 + 2T > G	(−) 5'SS (5.74→−1.9)	-	Δ(E1q17) (100%)	
c.48 + 5C > T	(↓) 5'SS (5.74→4.05)	100%		
c.49-2del	(−) 3'SS (9.28→1.16) (+) 3'SS (8.65) 6 nt downstream	-		Δ(E2p6) (100%)
c.49-1del	(−) 3'SS (9.28→−7.59) (+) 3'SS (8.76) 6 nt downstream	-		Δ(E2p6) (100%)
c.49-1G > A	(−) 3'SS (9.28→0.53) (+) 3'SS (7.53) 6 nt downstream	-		Δ(E2p6) (100%)
c.108 + 1_108 + 2insC	(−) 5'SS (10.86→−4.16)	-		Δ(E2) (100%)
c.108 + 5G > A	(↓) 5'SS (10.86→8.76)	100%		
c.109-2A > C	(−) 5'SS (10.06→2.02) (+) 3'SS (5.47) 11 nt downstream	-	Δ(E3p11) (85% ± 0.5%) Δ(E3) (15% ± 0.5%)	
c.210A > G	(↓) 5'SS (8.76→6.15)	100%		
c.210A > C	(↓) 5'SS (8.76→6.89)	100%		
c.211 + 1G > T	(↓) 3'SS (8.76→0.26)	-	Δ(E3) (73.3% ± 0.6%)	▼(E3q48a) (26.7% ± 0.6%)
c.211 + 2T > C	(−) 3'SS (8.76→1.01)	-	Δ(E3) (48.1% ± 7.4%)	▼(E3q48b) (51.9% ± 7.4%)

¹ Bold font: No traces or <5% of the full-length transcript. ² (−) site disruption; (+) New site; (↓) Reduction of MES score; Cr., Cryptic. ³ PTC, Premature Termination Codon; ^{3,4} Δ, loss of exonic sequences; ▼, inclusion of intronic sequences; E (exon), p (acceptor shift), q (donor shift). When necessary, the exact number of nt inserted or deleted is indicated. For example, transcript ▼(E1q9) denotes the use of an alternative donor site that is located nine nucleotides downstream of exon 1, causing the addition of 9 nt to the mature mRNA.

As above mentioned, total splicing disruptions were exclusively due to changes in the canonical AG/GT dinucleotides (Table 4). There are only a few exceptions of non-spliceogenic $\pm 1,2$ variants, basically consisting of the change of the consensus AG or GT dinucleotides into atypical splice sites, such as the GC 5' splice sites that account for about 1% of human 5'SS [46,47]. Thus, it has been assessed that about 15–18% of + 2T > C variants, which generate an atypical GC donor site, are able to produce the full-length transcript [48], as is the case of the *PALB2* c.108 + 2T > C variant [28]. However, in this study, we have shown that c.211 + 2T > C produced just aberrant transcripts similarly to variant c.48 + 2T > C [28]. SpliceAI analysis of c.211 + 2T > C (donor loss = 0.99) correctly predicted a total impact on splicing. Curiously, we found another atypical splice-site recognition in a previous study [29]. ATM variant c.1898 + 2T > G creates an intronic GG dinucleotide that might represent an extremely rare 5'SS (~0.01% of human exons) [47]. In fact, we found that this GG 5'SS was used in 13% of minigene transcripts producing the full-length isoform [29]. Therefore, the splicing outcomes of any variant should be carefully analyzed since the generation and use of active atypical sites may rescue the production of the full-length transcript, and thus, these data may modify the clinical classification of variants. Most importantly, up to now, the use of uncommon splice sites cannot be predicted, so they can only be detected by splicing assays. In this regard, minigenes provide a substantial advantage over RNA assays in carriers since variant read-out is not mistaken with wt allele expression. Then, any residual full-length transcript produced by the variant can be tracked by the highly sensitive fluorescent fragment electrophoresis. Conversely, partial splicing outcomes producing the full-length transcript are not simply identified in patient RNA assays unless a coding heterozygous SNP was also present so that the wild-type and variant alleles can be distinguished.

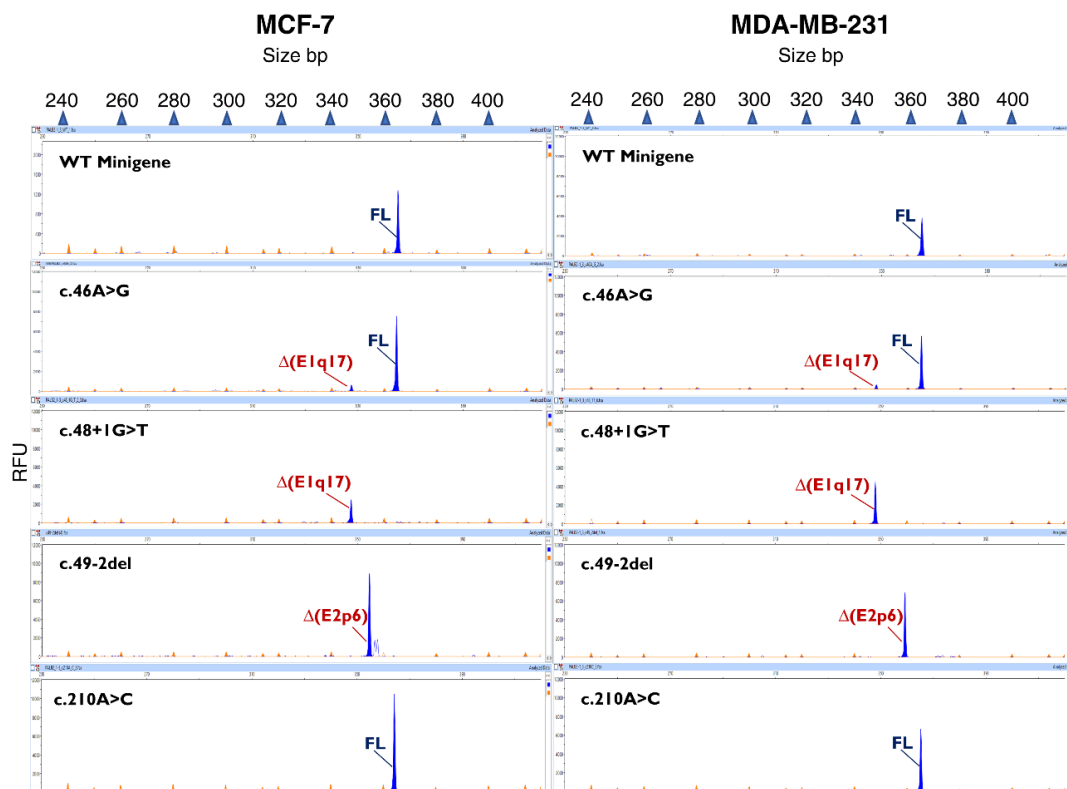


Figure 2. Reproducibility of *PALB2* RNA assays in MDA-MB-231 (left) and MCF-7 (right) cells. The wild-type and mutant minigenes of c.46A > G, c.48 + 1G > T, c.49-2del, and c.210A > C were tested in MCF-7 and MDA-MB-231 cells. RT-PCR products were run by fluorescent fragment electrophoresis using LIZ-600 as size standard. The x-axis indicates size in bp (electropherograms on the top) and the y-axis represents relative fluorescence units (RFU).

Concerning the splicing output, SpliceAI produced reliable predictions for 12 out of the 16 assayed variants (Table 3). Interestingly, two false-positive variants selected on MES score (c.48 + 5C > T and c.108 + 5G > A) were ruled out by SpliceAI. MES accurately predicted splice-site disruptions or their weakening for the twelve spliceogenic variants although estimations failed in the case of four variants. By increasing the MES threshold to -30% , the specificity of the selection procedure would have considerably improved. We firmly believe that bioinformatics predictions are only useful to filter out variants and select those potentially spliceogenic, but at present, RNA assays are critical at validating a splicing effect.

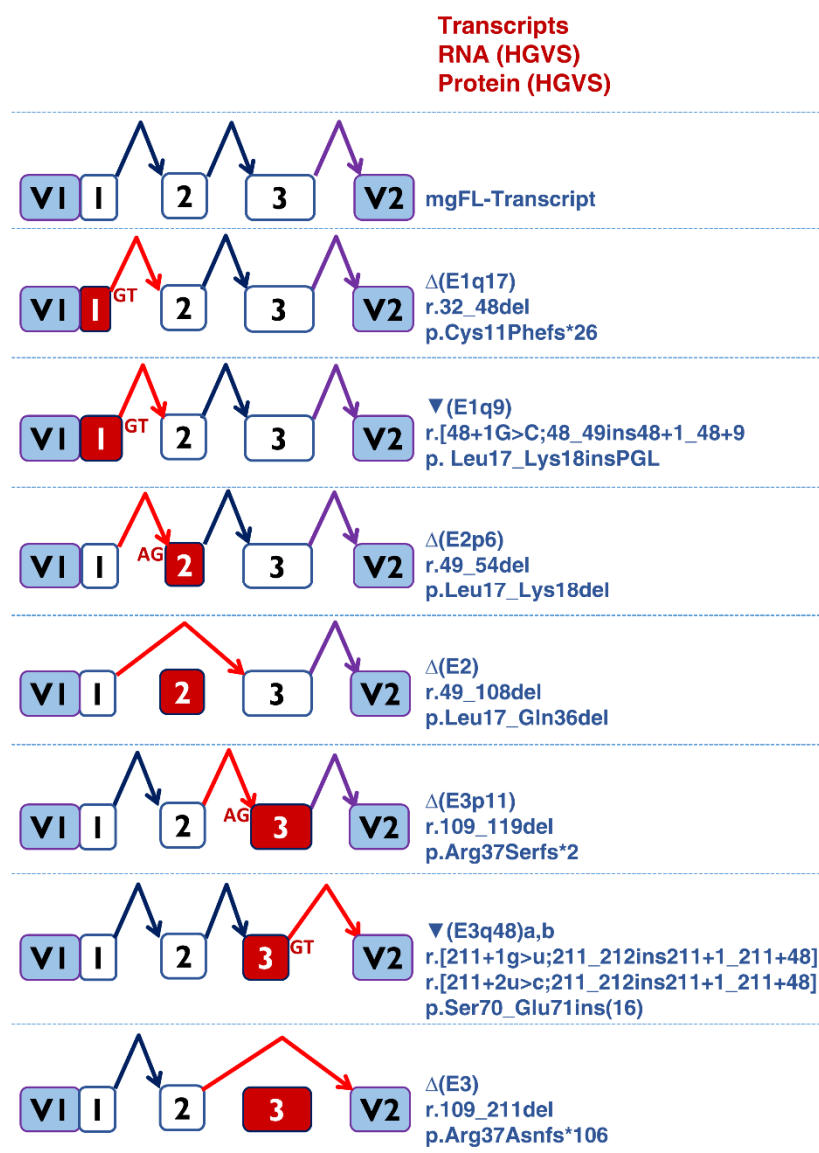


Figure 3. Transcripts produced by *PALB2* variants. Diagrams of the splicing reactions. Exons and the splicing reactions are indicated by boxes and elbow arrows, respectively. Anomalous events, exon skipping or alternative site usage (AG or GT sites) and exons are indicated in red. The impact of each transcript at the RNA and protein levels are described following the Human Genome Variation Society (HGVS) recommendations (right).

3.3. ACMG/AMP-Based Interpretation of Variants

PALB2 expert panel specifications of the ACMG/AMP guidelines are not yet available (<https://clinicalgenome.org/>, last accessed on 07 July 2022); so, as indicated in Materials and Methods, we classified 16 *PALB2* variants according to generic ACMG/AMP-based

classification guidelines combined with some *PALB2* specifications previously developed by our group [28]. This approach integrates mg*PALB2* readouts as observable PVS1_O/BP7_O evidence codes (Table 5). Thus, the three PTC-NMD transcripts (Δ (E1q17), Δ (E3p11), and Δ (E3)) are considered a very strong evidence of pathogenicity (P_VS). Likewise, the in-frame transcript Δ (E2) deletes a key *PALB2* domain (CC domain), where residues Leu17, Leu21, Leu24, Tyr28, Thr31, and Leu35 mediate important interactions in the *PALB2* homodimer and/or the *PALB2*/BRCA1 heterodimer [28,49]. Then, Δ (E2) was deemed a very strong evidence of pathogenicity (P_VS). In addition, the in-frame isoforms ∇ (E1q9) and Δ (E2p6) are predicted to disrupt critical regions for *PALB2*, inserting three or deleting two amino acids at the CC domain, respectively. However, in both cases, a functional impact on protein function cannot be predicted. Therefore, as we had previously pointed out [28], we think that both transcripts provide a moderate evidence of pathogenicity (P_M).

On the other hand, the contribution of ∇ (E3q48) a and b (insertion of 16 new amino acids: VKSRPFTYACFIIHFP and GKSRRPFTYACFIIHFP, respectively) is unclear. As we previously reported, the 16-aminoacids insertion was classified as a supporting evidence of pathogenicity (P_P) based on bioinformatics predictions (PROVEAN score of -15.84 , deleterious) [28]. Finally, the FL-transcript with the missense variant c.46A > G (p.Lys16Glu) was considered a supporting benign evidence BP4 (-1) since the REVEL score (0.075) suggests no impact on protein function.

All the 16 variants are absent in the gnomAD database, so they meet the PM2 rarity code (Table 5) that we have considered a supporting evidence of pathogenicity (PM2_P; +1 point) as previously mentioned [29]. As indicated above (see Section 2.5), once we incorporate minigene readouts into the classification scheme, predictive splicing codes PVS1 (GT-AG variants) or PP3/BP4 (non GT-AG variants) are no longer taken into consideration.

Finally, we considered that some pathogenic (PS2, PM1, PM6, PP2, PP4, PP5) and benign (BS2, BP1, BP3, BP5, BP6) codes are not applicable to the classification of any of the herein described *PALB2* variants. In addition, the PM3 evidence (in trans with a pathogenic variant in a recessive disorder) was not applied to any of the variants because they were not found in Fanconi Anemia patients (based on ClinVar database, Leiden Open Variation Database, <https://databases.lovd.nl/shared/genes/PALB2>, accessed 8 July 2022, and literature searches).

Taking these considerations altogether, six variants were classified as likely pathogenic (+9 points of the Bayesian scale), five as VUS (+2 or +3 points), and five as likely benign (-1 or -3 points) (Table 5). Remarkably, five $\pm 1,2$ variants (c.49-2del, c.49-1del, c.49-1G >A, c.211 + 1G >T, and c.211 + 2T >C) were catalogued as VUS because they produced the in-frame transcripts Δ (E2p6) (100% of the overall expression) or ∇ (E3q48) (27%–52% of the overall expression), whose impact on *PALB2* function is uncertain. Therefore, it is essential to elucidate if these transcripts retain the DNA repair activity to ascertain the pathogenicity of these five variants. Hence, the PVS1 splicing predictive evidence of $\pm 1,2$ variants may lead to their clinical misinterpretation [50].

Table 5. ACMG/AMP-based classification of 16 *PALB2* variants at exons 1 to 3.

Variants	ClinVar ¹	ACMG-AMP ² Classification	Splicing Predictive PVS1/PP3 ³	PVS1_O/BP7_O (mgPALB2 Readouts) ⁴	PM2 ⁵
c.46A > G (p.Lys16Glu)	B	LB (−1)	PP3	BP7_O_M (−2): 93% [BP7_O_M, FL] ⁶ + 7% [PVS1_O, Δ(E1q17)]	PM2_P (+1)
c.48 + 1del	LP	LP (+9)	PVS1	PVS1_O (+8): 100% [PVS1_O, Δ(E1q17)]	PM2_P (+1)
c.48 + 1G > C	Conflicting	LP (+9)	PVS1	PVS1_O (+8): 91% [PVS1_O, Δ(E1q17)] + 9% [PVS1_O_M, ▼(E1q9)]	PM2_P (+1)
c.48 + 1G > T	P/LP	LP (+9)	PVS1	PVS1_O (+8): 100% [PVS1_O, Δ(E1q17)]	PM2_P (+1)
c.48 + 2T > G	LP	LP (+9)	PVS1	PVS1_O (+8): 100% [PVS1_O, Δ(E1q17)]	PM2_P (+1)
c.48 + 5C > T	Conflicting	LB (−3)	PP3	BP7_O_S (−4): 100% [BP7_O_S, FL] ⁷	PM2_P (+1)
c.49-2del	VUS	VUS (+3)	PVS1	PVS1_O_M (+2): 100% [PVS1_O_M, Δ(E2p6)]	PM2_P (+1)
c.49-1del	P	VUS (+3)	PVS1	PVS1_O_M (+2): 100% [PVS1_O_M, Δ(E2p6)]	PM2_P (+1)
c.49-1G > A	LP	VUS (+3)	PVS1	PVS1_O_M (+2): 100% [PVS1_O_M, Δ(E2p6)]	PM2_P (+1)
c.108 + 1_108 + 2insC	LP	LP (+9)	PVS1	PVS1_O (+8): 100% [PVS1_O, Δ(E2)]	PM2_P (+1)
c.108 + 5G > A	VUS	LB (−3)	PP3	BP7_O_S (−4): 100% [BP7_O_S, FL] ⁷	PM2_P (+1)
c.109-2A > C	LP	LP (+9)	PVS1	PVS1_O (+8): 85% [PVS1_O, Δ(E3p11)] + 15% [PVS1_O, Δ(E3)]	PM2_P (+1)
c.210A > G (p.Ser70=)	Conflicting	LB (−3)	PP3	BP7_O_S (−4): 100% [BP7_O_S, FL] ⁸	PM2_P (+1)
c.210A > C (p.Ser70=)	LB	LB (−3)	PP3	BP7_O_S (−4): 100% [BP7_O_S, FL] ⁸	PM2_P (+1)
c.211 + 1G > T	P/LP	VUS (+2)	PVS1	PVS1_O_P (+1): 73% [PVS1_O, Δ(E3)] + 27% [PVS1_O_P, ▼(E3q48a)]	PM2_P (+1)
c.211 + 2T > C	LP	VUS (+2)	PVS1	PVS1_O_P (+1): 48% PVS1_O, Δ(E3) + 52% [PVS1_O_P, ▼(E3q48b)]	PM2_P (+1)

¹ Clinical interpretation at the ClinVar database (accessed on 9 September 2022). ² Point-based variant classification categories are defined as follows: pathogenic (P) $\geq +10$; likely pathogenic (LP) +6 to +9; variant of uncertain significance (VUS) 0 to +5; likely benign (LB) −1 to −6; and benign (B) ≤ -7 . ³ The predictive splicing codes were not taken into account in this study since they were considered redundant when splicing assay data are available. ⁴ Deconvolution of minigene readouts and assigned score according to the rules indicated in Materials and Methods. Note that transcripts representing <10% of the overall expression (e.g., Δ(E1q17) in c.46A > G) do not contribute to the final PVS1_O/BP7_O evidence strength assignment. If two transcripts representing >10% of the overall expression each support different evidence strengths, the most conservative strength is assigned (e.g., c.211 + 1G > T minigene readout ends up as PVS1_O_P even if the major signal supports PVS1_O). ⁵ Rarity code PM2 with allele frequency $\leq 0.01\%$ at gnomADv2.1.1. ⁶ The FL transcript carries a missense variant r.46A > G (p.Lys16Glu) that qualifies for protein predictive evidence BP4 (multiple lines of computational evidence suggest no impact) with moderate strength (REVEL score = 0.075). Based on that, FL expression as observed in the assay qualifies for BP7_O_M. ⁷ The FL transcripts has a wt sequence that qualifies for BP7_S. ⁸ The FL transcripts carries a synonymous variant r.210A > G or r.210A > C that, once an impact on splicing has been excluded, qualifies for BP7_O.

4. Conclusions

We tested 16 variants at *PALB2* exons 1 to 3 by hybrid minigenes. Twelve variants impaired splicing, and eleven produced negligible levels of the mgFL-transcript. Integrating our previous results for *PALB2* [28], we analyzed a total 58 potential spliceogenic variants, 47 of which (81%) induced splicing anomalies, supporting the high sensitivity and specificity of our selection criteria as well as the efficacy of our minigene approach. By an ACMG/AMP-based strategy, a total of 29 variants were classified as pathogenic/likely pathogenic and, equally relevant, 13 variants as likely benign, whereas 16 variants were kept as VUS. Interestingly, another 56 ClinVar variants at exons 4 to 12 (accessed on 3 August 2021) would be potentially spliceogenic as per MES scores (data not shown), so in future projects, they could be assayed in our two previously reported *PALB2* minigenes: mgPALB2_ex4-6 and mgPALB2_ex5-12 [28]. Moreover, the ACMG/AMP-based guidelines provide a useful framework for the clinical interpretation of variants when splicing data are available. Finally, minigene assays allowed assessing more than 600 variants of the main breast cancer susceptibility genes up to now, demonstrating their high simplicity and robustness. Furthermore, this tool has been used to successfully assay variants at other disease genes, such as *UGT1A1* (Crigler–Najjar syndrome) [51], *CHD7* (Charge syndrome) [52], or *TRPM4* (colorectal cancer) [53], among others (<http://www.ibgm.med.uva.es/servicios/servicio-de-splicing-minigenes/>, accessed on 13 July 2022).

Author Contributions: Conceptualization, E.A.V.-S.; data curation, A.V.-P., formal analysis, A.V.-P., E.B.-M., L.S.-M., A.E.-S., M.d.l.H. and E.A.V.-S.; funding acquisition, M.d.l.H. and E.A.V.-S.; investigation, A.V.-P., E.B.-M., L.S.-M., A.E.-S., I.L.-B., P.P.-S., A.G.-Á., S.G.-B., M.d.l.H. and E.A.V.-S.; methodology, A.V.-P., L.S.-M., E.B.-M., I.L.-B., A.G.-Á., M.d.l.H. and E.A.V.-S.; supervision, E.A.V.-S.; writing—original draft, A.V.-P. and E.A.V.-S.; writing—review and editing, A.V.-P., S.G.-B., L.S.-M., M.d.l.H. and E.A.V.-S. All authors have read and agreed to the published version of the manuscript.

Funding: The EAV lab is supported by grants from the Spanish Ministry of Science and Innovation, Plan Nacional de I+D+I 2013–2016, ISCIII (PI20/00225), co-funded by FEDER (the European Regional Development Fund, European Union) and the Consejería de Educación, Junta de Castilla y León, ref. CSI242P18 (actuación cofinanciada P.O. FEDER 2014–2020 de Castilla y León) and Programa Estratégico Instituto de Biología y Genética Molecular (IBGM), Escalera de Excelencia, Junta de Castilla y León (Ref. CLU-2019-02). The MdlH lab is supported by a grant from the Spanish Ministry of Science and Innovation, Plan Nacional de I+D+I 2013–2016, ISCIII (PI20/00110), co-funded by FEDER. L.S.-M. is supported by a predoctoral fellowship from the AECC Scientific Foundation, Sede Provincial de Valladolid (2019–2023). E.B.-M. is a postdoctoral researcher funded by the University of Valladolid (POSTDOC-UVA05, 2022–2025). I.L.-B. is supported by a predoctoral fellowship from the Consejería de Educación, Junta de Castilla y León (2022–2025). A.E.-S. is supported through the Operational Program for Youth Employment and the Youth Employment Initiative (YEI), set up by the Community of Madrid in 2020 and co-financed by the European Social Fund.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All sequencing and fragment analysis data will be available after publication at Digital. CSIC.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nielsen, F.C.; van Overeem Hansen, T.; Sørensen, C.S. Hereditary breast and ovarian cancer: New genes in confined pathways. *Nat. Rev. Cancer* **2016**, *16*, 599–612. [[CrossRef](#)] [[PubMed](#)]
2. Gerhardus, A.; Schleberger, H.; Schlegelberger, B.; Gadzicki, D. Diagnostic accuracy of methods for the detection of BRCA1 and BRCA2 mutations: A systematic review. *Eur. J. Hum. Genet.* **2007**, *15*, 619–627.
3. Velasco, E.; Infante, M.; Durán, M.; Esteban-Cardenosa, E.; Lastra, E.; García-Girón, C.; Miner, C.; Duran, M.; Esteban-Cardenosa, E.; Lastra, E.; et al. Rapid mutation detection in complex genes by heteroduplex analysis with capillary array electrophoresis. *Electrophoresis* **2005**, *26*, 2539–2552. [[CrossRef](#)] [[PubMed](#)]

4. Dorling, L.; Carvalho, S.; Allen, J.; González-Neira, A.; Luccarini, C.; Wahlström, C.; Pooley, K.A.; Parsons, M.T.; Fortunato, C.; Wang, Q.; et al. Breast Cancer Risk Genes — Association Analysis in More than 113,000 Women. *N. Engl. J. Med.* **2021**, *384*, 428–439.
5. Hu, C.; Hart, S.N.; Gnanaolivu, R.; Huang, H.; Lee, K.Y.; Na, J.; Gao, C.; Lilyquist, J.; Yadav, S.; Boddicker, N.J.; et al. A Population-Based Study of Genes Previously Implicated in Breast Cancer. *N. Engl. J. Med.* **2021**, *384*, 440–451. [[CrossRef](#)]
6. Narod, S.A. Which Genes for Hereditary Breast Cancer? *N. Engl. J. Med.* **2021**, *384*, 471–473.
7. Reid, S.; Schindler, D.; Hanenberg, H.; Barker, K.; Hanks, S.; Kalb, R.; Neveling, K.; Kelly, P.; Seal, S.; Freund, M.; et al. Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer. *Nat. Genet.* **2007**, *39*, 162–164. [[CrossRef](#)]
8. Zhao, W.; Wiese, C.; Kwon, Y.; Hromas, R.; Sung, P. The BRCA tumor suppressor network in chromosome damage repair by homologous recombination. *Annu. Rev. Biochem.* **2019**, *88*, 221–245. [[CrossRef](#)]
9. Nepomuceno, T.C.; De Gregoriis, G.; de Oliveira, F.M.B.; Suarez-Kurtz, G.; Monteiro, A.N.; Carvalho, M.A. The role of PALB2 in the DNA damage response and cancer predisposition. *Int. J. Mol. Sci.* **2017**, *18*, 1886. [[CrossRef](#)]
10. Rahman, N.; Seal, S.; Thompson, D.; Kelly, P.; Renwick, A.; Elliott, A.; Reid, S.; Spanova, K.; Barfoot, R.; Chagtai, T.; et al. PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat. Genet.* **2007**, *39*, 165–167. [[CrossRef](#)]
11. Antoniou, A.C.; Foulkes, W.D.; Tischkowitz, M. Breast cancer risk in women with PALB2 mutations in different populations. *Lancet Oncol.* **2015**, *16*, e375. [[CrossRef](#)]
12. Yang, X.; Leslie, G.; Doroszuk, A.; Schneider, S.; Allen, J.; Decker, B.; Dunning, A.M.; Redman, J.; Scarth, J.; Plaskocinska, I.; et al. Cancer risks associated with germline PALB2 pathogenic variants: An international study of 524 families. *J. Clin. Oncol.* **2020**, *38*, 674–685. [[CrossRef](#)]
13. Eccles, D.M.; Mitchell, G.; Monteiro, A.N.A.; Schmutzler, R.; Couch, F.J.; Spurdle, A.B.; Gómez-García, E.B. ENIGMA Clinical Working Group BRCA1 and BRCA2 genetic testing—pitfalls and recommendations for managing variants of uncertain clinical significance. *Ann. Oncol.* **2015**, *26*, 2057–2065. [[CrossRef](#)]
14. Radice, P.; de Summa, S.; Caleca, L.; Tommasi, S. Unclassified variants in BRCA genes: Guidelines for interpretation. *Ann. Oncol.* **2011**, *22*, i18–i23. [[CrossRef](#)]
15. De Vooght, K.M.K.; Van Wijk, R.; Van Solinge, W.W. Management of gene promoter mutations in molecular diagnostics. *Clin. Chem.* **2009**, *55*, 698–708. [[CrossRef](#)]
16. Fraile-Bethencourt, E.; Valenzuela-Palomo, A.; Díez-Gómez, B.; Infante, M.; Durán, M.; Marcos, G.; Lastra, E.; Gómez-Barrero, S.; Velasco, E.A. Genetic dissection of the BRCA2 promoter and transcriptional impact of DNA variants. *Breast Cancer Res. Treat.* **2018**, *171*, 53–63. [[CrossRef](#)]
17. Buratti, E.; Baralle, M.; Baralle, F.E. Defective splicing, disease and therapy: Searching for master checkpoints in exon definition. *Nucleic Acids Res.* **2006**, *34*, 3494–3510. [[CrossRef](#)]
18. Manning, K.S.; Cooper, T.A. The roles of RNA processing in translating genotype to phenotype. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 102–114. [[CrossRef](#)]
19. Wang, G.-S.S.; Cooper, T.A. Splicing in disease: Disruption of the splicing code and the decoding machinery. *Nat. Rev. Genet.* **2007**, *8*, 749–761. [[CrossRef](#)]
20. Baralle, D.; Lucassen, A.; Buratti, E. Missed threads. The impact of pre-mRNA splicing defects on clinical practice. *EMBO Rep.* **2009**, *10*, 810–816. [[CrossRef](#)]
21. Scotti, M.M.; Swanson, M.S. RNA mis-splicing in disease. *Nat. Rev. Genet.* **2016**, *17*, 19–32. [[CrossRef](#)]
22. Lopez-Bigas, N.; Audit, B.; Ouzounis, C.; Parra, G.; Guigo, R.; López-Bigas, N.; Audit, B.; Ouzounis, C.; Parra, G.; Guigó, R. Are splicing mutations the most frequent cause of hereditary disease? *FEBS Lett.* **2005**, *579*, 1900–1903. [[CrossRef](#)]
23. Rhine, C.L.; Cygan, K.J.; Soemedi, R.; Maguire, S.; Murray, M.F.; Monaghan, S.F.; Fairbrother, W.G. Hereditary cancer genes are highly susceptible to splicing mutations. *PLoS Genet.* **2018**, *14*, e1007231. [[CrossRef](#)]
24. Sanz, D.J.; Acedo, A.; Infante, M.; Durán, M.; Pérez-Cabornero, L.; Esteban-Cardenosa, E.; Lastra, E.; Pagani, F.; Miner, C.; Velasco, E.A. A high proportion of DNA variants of BRCA1 and BRCA2 is associated with aberrant splicing in breast/ovarian cancer patients. *Clin. Cancer Res.* **2010**, *16*, 1957–1967. [[CrossRef](#)]
25. Fraile-Bethencourt, E.; Díez-Gómez, B.; Velásquez-Zapata, V.; Acedo, A.; Sanz, D.J.; Velasco, E.A. Functional classification of DNA variants by hybrid minigenes: Identification of 30 spliceogenic variants of BRCA2 exons 17 and 18. *PLoS Genet.* **2017**, *13*, e1006691. [[CrossRef](#)]
26. Sanoguera-Miralles, L.; Valenzuela-Palomo, A.; Bueno-Martínez, E.; Llovet, P.; Díez-Gómez, B.; Caloca, M.J.; Pérez-Segura, P.; Fraile-Bethencourt, E.; Colmena, M.; Carvalho, S.; et al. Comprehensive Functional Characterization and Clinical Interpretation of 20 Splice-Site Variants of the RAD51C Gene. *Cancers* **2020**, *12*, 3771. [[CrossRef](#)]
27. Bueno-Martínez, E.; Sanoguera-Miralles, L.; Valenzuela-Palomo, A.; Lorca, V.; Gómez-Sanz, A.; Carvalho, S.; Allen, J.; Infante, M.; Pérez-Segura, P.; Lázaro, C.; et al. Rad51d aberrant splicing in breast cancer: Identification of splicing regulatory elements and minigene-based evaluation of 53 dna variants. *Cancers* **2021**, *13*, 2845. [[CrossRef](#)]
28. Valenzuela-Palomo, A.; Bueno-Martínez, E.; Sanoguera-Miralles, L.; Lorca, V.; Fraile-Bethencourt, E.; Esteban-Sánchez, A.; Gómez-Barrero, S.; Carvalho, S.; Allen, J.; García-Álvarez, A.; et al. Splicing predictions, minigene analyses, and ACMG-AMP clinical classification of 42 germline PALB2 splice-site variants. *J. Pathol.* **2022**, *256*, 321–334. [[CrossRef](#)]

29. Bueno-Martínez, E.; Sanoguera-Miralles, L.; Valenzuela-Palomo, A.; Esteban-Sánchez, A.; Lorca, V.; Llinares-Burguet, I.; Allen, J.; García-Álvarez, A.; Pérez-Segura, P.; Durán, M.; et al. Minigene-based splicing analysis and ACMG/AMP-based tentative classification of 56 ATM variants. *J. Pathol.* **2022**, *258*, 83–101. [[CrossRef](#)]
30. Lopez-Perolio, I.; Leman, R.; Behar, R.; Lattimore, V.; Pearson, J.F.; Castéra, L.; Martins, A.; Vaur, D.; Goardon, N.; Davy, G.; et al. Alternative splicing and ACMG-AMP-2015-based classification of PALB2 genetic variants: An ENIGMA report. *J. Med. Genet.* **2019**, *56*, 453–460. [[CrossRef](#)]
31. Fraile-Bethencourt, E.; Valenzuela-Palomo, A.; Díez-Gómez, B.; Caloca, M.J.; Gómez-Barrero, S.; Velasco, E.A. Minigene Splicing Assays Identify 12 Spliceogenic Variants of BRCA2 Exons 14 and 15. *Front. Genet.* **2019**, *10*, 503. [[CrossRef](#)] [[PubMed](#)]
32. Yeo, G.; Burge, C.B. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J. Comput. Biol.* **2004**, *11*, 377–394. [[CrossRef](#)] [[PubMed](#)]
33. Houdayer, C.; Caux-Moncoutier, V.; Krieger, S.; Barrois, M.; Bonnet, F.; Bourdon, V.; Bronner, M.; Buisson, M.; Coulet, F.; Gaildrat, P.; et al. Guidelines for splicing analysis in molecular diagnosis derived from a set of 327 combined in silico/in vitro studies on BRCA1 and BRCA2 variants. *Hum. Mutat.* **2012**, *33*, 1228–1238. [[CrossRef](#)] [[PubMed](#)]
34. Moles-Fernández, A.; Duran-Lozano, L.; Montalban, G.; Bonache, S.; López-Perolio, I.; Menéndez, M.; Santamariña, M.; Behar, R.; Blanco, A.; Carrasco, E.; et al. Computational Tools for Splicing Defect Prediction in Breast/Ovarian Cancer Genes: How Efficient Are They at Predicting RNA Alterations? *Front. Genet.* **2018**, *9*, 366. [[CrossRef](#)]
35. Jaganathan, K.; Kyriazopoulou Panagiotopoulou, S.; McRae, J.F.; Darbandi, S.F.; Knowles, D.; Li, Y.I.; Kosmicki, J.A.; Arbelaez, J.; Cui, W.; Schwartz, G.B.; et al. Predicting Splicing from Primary Sequence with Deep Learning. *Cell* **2019**, *176*, 535–548.e24. [[CrossRef](#)]
36. De Garibay, G.R.; Acedo, A.; García-Casado, Z.; Gutiérrez-Enríquez, S.; Tosar, A.; Romero, A.; Garre, P.; Lloret, G.; Thomassen, M.; Díez, O.; et al. Capillary Electrophoresis Analysis of Conventional Splicing Assays: IARC Analytical and Clinical Classification of 31 BRCA2 Genetic Variants. *Hum. Mutat.* **2014**, *35*, 53–57. [[CrossRef](#)]
37. Acedo, A.; Hernández-Moro, C.; Curiel-García, Á.; Díez-Gómez, B.; Velasco, E.A. Functional classification of BRCA2 DNA variants by splicing assays in a large minigene with 9 exons. *Hum. Mutat.* **2015**, *36*, 210–221. [[CrossRef](#)]
38. Fraile-Bethencourt, E.; Valenzuela-Palomo, A.; Díez-Gómez, B.; Goína, E.; Acedo, A.; Buratti, E.; Velasco, E.A. Mis-splicing in breast cancer: Identification of pathogenic BRCA2 variants by systematic minigene assays. *J. Pathol.* **2019**, *248*, 409–420. [[CrossRef](#)]
39. Tavtigian, S.V.; Greenblatt, M.S.; Harrison, S.M.; Nussbaum, R.L.; Prabhu, S.A.; Boucher, K.M.; Biesecker, L.G. ClinGen Sequence Variant Interpretation Working Group (ClinGen SVI) Modeling the ACMG/AMP variant classification guidelines as a Bayesian classification framework. *Genet. Med.* **2018**, *20*, 1054–1060. [[CrossRef](#)]
40. Tavtigian, S.V.; Harrison, S.M.; Boucher, K.M.; Biesecker, L.G. Fitting a naturally scaled point system to the ACMG/AMP variant classification guidelines. *Hum. Mutat.* **2020**, *41*, 1734–1737. [[CrossRef](#)]
41. Sanoguera-Miralles, L.; Bueno-Martínez, E.; Valenzuela-Palomo, A.; Esteban-Sánchez, A.; Llinares-Burguet, I.; Pérez-Segura, P.; García-Álvarez, A.; de la Hoya, M.; Velasco-Sampedro, E.A. Minigene Splicing Assays Identify 20 Spliceogenic Variants of the Breast/Ovarian Cancer Susceptibility Gene RAD51C. *Cancers* **2022**, *14*, 2960. [[CrossRef](#)]
42. Wilcox, E.H.; Sarmady, M.; Wulf, B.; Wright, M.W.; Rehm, H.L.; Biesecker, L.G.; Abou Tayoun, A.N. Evaluating the impact of in silico predictors on clinical variant classification. *Genet. Med.* **2022**, *24*, 924–930. [[CrossRef](#)]
43. Tsaousis, G.N.; Papadopoulou, E.; Apeessos, A.; Agiannitopoulos, K.; Pepe, G.; Kampouri, S.; Diamantopoulos, N.; Floros, T.; Iosifidou, R.; Katopodi, O.; et al. Analysis of hereditary cancer syndromes by using a panel of genes: Novel and multiple pathogenic mutations. *BMC Cancer* **2019**, *19*, 535. [[CrossRef](#)]
44. Lykke-Andersen, S.; Jensen, T.H. Nonsense-mediated mRNA decay: An intricate machinery that shapes transcriptomes. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 665–677. [[CrossRef](#)]
45. Casadei, S.; Gulsuner, S.; Shirts, B.H.; Mandell, J.B.; Kortbawi, H.M.; Norquist, B.S.; Swisher, E.M.; Lee, M.K.; Goldberg, Y.; O'Connor, R.; et al. Characterization of splice-altering mutations in inherited predisposition to cancer. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 26798–26807. [[CrossRef](#)]
46. Thanaraj, T.A.; Clark, F. Human GC-AG alternative intron isoforms with weak donor sites show enhanced consensus at acceptor exon positions. *Nucleic Acids Res.* **2001**, *29*, 2581–2593. [[CrossRef](#)]
47. Parada, G.E.; Munita, R.; Cerda, C.A.; Gysling, K. A comprehensive survey of non-canonical splice sites in the human transcriptome. *Nucleic Acids Res.* **2014**, *42*, 10564–10578. [[CrossRef](#)]
48. Lin, J.H.; Tang, X.Y.; Boulling, A.; Zou, W.B.; Masson, E.; Fichou, Y.; Raud, L.; Le Tertre, M.; Deng, S.J.; Berlivet, I.; et al. First estimate of the scale of canonical 5' splice site GT>GC variants capable of generating wild-type transcripts. *Hum. Mutat.* **2019**, *40*, 1856–1873. [[CrossRef](#)]
49. Song, F.; Li, M.; Liu, G.; Swapna, G.V.T.; Daigham, N.S.; Xia, B.; Montelione, G.T.; Bunting, S.F. Antiparallel Coiled-Coil Interactions Mediate the Homodimerization of the DNA Damage-Repair Protein PALB2. *Biochemistry* **2018**, *57*, 6581–6591. [[CrossRef](#)]
50. Richards, S.; Aziz, N.; Bale, S.; Bick, D.; Das, S.; Gastier-Foster, J.; Grody, W.W.; Hegde, M.; Lyon, E.; Spector, E.; et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **2015**, *17*, 405–424. [[CrossRef](#)]
51. Gailite, L.; Valenzuela-Palomo, A.; Sanoguera-Miralles, L.; Rots, D.; Kreile, M.; Velasco, E.A. UGT1A1 Variants c.864+5G>T and c.996+2_996+5del of a Crigler-Najjar Patient Induce Aberrant Splicing in Minigene Assays. *Front. Genet.* **2020**, *11*, 169. [[CrossRef](#)]

52. Villate, O.; Ibarluzea, N.; Fraile-Bethencourt, E.; Valenzuela, A.; Velasco, E.A.; Grozeva, D.; Raymond, F.L.; Botella, M.P.; Tejada, M.-I. Functional Analyses of a Novel Splice Variant in the CHD7 Gene, Found by Next Generation Sequencing, Confirm Its Pathogenicity in a Spanish Patient and Diagnose Him with CHARGE Syndrome. *Front. Genet.* **2018**, *9*, 26–31. [[CrossRef](#)]
53. Zhu, L.; Miao, B.; Dymerska, D.; Kuswik, M.; Bueno-Martínez, E.; Sanoguera-Miralles, L.; Velasco, E.A.; Paramasivam, N.; Schlesner, M.; Kumar, A.; et al. Germline Variants of CYBA and TRPM4 Predispose to Familial Colorectal Cancer. *Cancers* **2022**, *14*, 670. [[CrossRef](#)]