Alternative splicing and ACMG-AMP-2015-based classification of PALB2 genetic variants: an ENIGMA report

Irene Lopez-Perolio, Raphaël Leman, Raquel Behar, Vanessa Lattimore, John F. Pearson, Laurent Castéra, Alexandra Martins, Dominique Vaur, Nicolas Goardon, Grégoire Davy, et al.

ORIGINAL ARTICLE

Alternative splicing and ACMG-AMP-2015-based classification of PALB2 genetic variants: an ENIGMA report

Irene Lopez-Perolio,[1] Raphaël Leman,[2] Raquel Behar,[1] Vanessa Lattimore,[3] John F Pearson,[3] Laurent Castéra,[2] Alexandra Martins,[4] Dominique Vaur,[2] Nicolas Goardon,[2] Grégoire Davy,[2] Pilar Garre,[1] Vanesa García-Barberán,[1] Patricia Llovet,[1] Pedro Pérez-Segura,[1] Eduardo Díaz-Rubio,[1] Trinidad Caldés,[1] Kathleen S Hruska,[5] Vickie Hsuan,[6] Sitao Wu,[6] Tina Pesaran,[6] Rachid Karam,[6] Johan Vallon-Christersson,[7] Ake Borg,[7] kConFab Investigators,[8,9] Alberto Valenzuela-Palomo,[10] Eladio A Velasco,[10] Melissa Southey,[11] Maaike P G Vreeswijk,[12] Peter Devilee,[12] Anders Kvist,[7] Amanda B Spurdle,[13] Logan C Walker,[3] Sophie Krieger,[2] Miguel de la Hoya[1]

► Additional material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/jmedgenet-2018-105834).

For numbered affiliations see end of article.

Correspondence to
Dr Miguel de la Hoya, Laboratorio OncologíaMolecular, Hospital Clínico San Carlos, Madrid 28040, Spain; mdhoya@hotmail.com

IL-P and RL contributed equally.

SK and MdlH are joint senior authors.

Received 26 October 2018
Revised 10 January 2019
Accepted 6 February 2019
Published Online First 19 March 2019

To cite: Lopez-Perolio I, Leman R, Behar R, *et al*. *J Med Genet* 2019;**56**:453–460.

ABSTRACT

Background *PALB2* monoallelic *loss-of-function* germ-line variants confer a breast cancer risk comparable to the average *BRCA2* pathogenic variant. Recommendations for risk reduction strategies in carriers are similar. Elaborating robust criteria to identify *loss-of-function* variants in *PALB2*—without incurring overprediction—is thus of paramount clinical relevance. Towards this aim, we have performed a comprehensive characterisation of alternative splicing in *PALB2*, analysing its relevance for the classification of truncating and splice site variants according to the 2015 American College of Medical Genetics and Genomics-Association for Molecular Pathology guidelines.

Methods Alternative splicing was characterised in RNAs extracted from blood, breast and *fimbriae*/ovary-related human specimens (n=112). RNAseq, RT-PCR/CE and CloneSeq experiments were performed by five contributing laboratories. Centralised revision/curation was performed to assure high-quality annotations. Additional splicing analyses were performed in *PALB2* c.212–1G>A, c.1684+1G>A, c.2748+2T>G, c.3113+5G>A, c.3350+1G>A, c.3350+4A>C and c.3350+5G>A carriers. The impact of the findings on PVS1 status was evaluated for truncating and splice site variant.

Results We identified 88 naturally occurring alternative splicing events (81 newly described), including 4 in-frame events predicted relevant to evaluate PVS1 status of splice site variants. We did not identify tissue-specific alternate gene transcripts in breast or ovarian-related samples, supporting the clinical relevance of blood-based splicing studies.

Conclusions PVS1 is not necessarily warranted for splice site variants targeting four *PALB2* acceptor sites (exons 2, 5, 7 and 10). As a result, rare variants at these splice sites cannot be assumed *pathogenic/likely pathogenic* without further evidences. Our study puts a warning in up to five *PALB2* genetic variants that are currently reported as *pathogenic/likely pathogenic* in ClinVar.

INTRODUCTION

Monoallelic *loss-of-function* (LoF) germ-line variants in *PALB2* predispose to breast cancer, with estimated absolute risks by age 80 ranging from 33% to 58%, depending on the family history.[1 2] Excess risk for other cancers, such as pancreas, prostate, ovarian and male breast cancer, is still under investigation. Currently, gene panel testing for breast cancer predisposition includes *PALB2*,[2] and LoF germ-line variants in this gene are considered actionable findings in many settings, with proposed actions ranging from increased surveillance to prophylactic surgery.[3–5] Accordingly, classifying *PALB2* LoF variants is of paramount clinical relevance. Yet, the task is not trivial, as proved by the large number of variants of uncertain significance still existing in genes that have been extensively studied, such as *BRCA1* or *BRCA2*.[6]

In the research setting, truncating (nonsense or frameshift) variants predicted to induce nonsense-mediated decay (PTC-NMD variants) and canonical ±1,2 splice site variants (hereafter named splice site variants) at cancer predisposition genes are often assumed pathogenic/likely pathogenic LoF variants.[1 2] However, in the clinical setting a more conservative approach is recommended. According to *the American College of Medical Genetics and Genomics-Association for Molecular Pathology* (ACMG-AMP) interpretation guidelines,[7] a PTC-NMD or splice site variant is a very strong evidence of pathogenicity (PVS1), but not sufficient to classify the variant as pathogenic/likely pathogenic. Additional combinations of strong (PS), moderate (PM) and/or supporting (PP) evidence of pathogenicity are required. Furthermore, PVS1 is not warranted for every PTC-NMD/splice site variant. Indeed, the ACMG-AMP-2015 guidelines specify several caveats, including the possibility of: (i) *rescue* transcripts (alternate gene transcripts that skip the truncating variant, encoding functional or partially functional proteins

and resulting in reduced or no haploinsufficiency), (ii) splice site variants producing transcripts with in-frame deletions/insertions retaining some or all functional capacity and (iii) tissue-specific alternate gene transcripts.[7] Therefore, the accurate interpretation of *PALB2* PTC-NMD and splice site variants according to the ACMG-AMP-2015 guidelines requires reliable information on both protein structure/function and alternative splicing.

To be more precise, *PALB2* PTC-NMD/splice site variants without direct risk estimates and/or functional data (a common scenario in genetic testing) should be classified as likely pathogenic only if PVS1 is warranted. For PTC-NMD variants, PVS1 is warranted if no *rescue* transcripts are predicted. For splice site variants the analysis is more complex. In addition to *rescue* transcripts, the possibility of the variant allele producing transcripts with in-frame alterations retaining coding potential should be considered, although predicting the precise nature of the transcripts produced by a splice site variant is challenging.

In recent years, the Evidence-based Network for the Interpretation of Germ-line Mutant Alleles (ENIGMA consortium) has conducted a comprehensive characterisation of naturally occurring alternate gene transcripts in *BRCA1* and *BRCA2*,[8 9] exploring the impact of the findings for the clinical classification of genetic variants at the two *loci*. Major achievements were the identification of a subset of splice sites variants for which PVS1 was not necessarily warranted, the posterior demonstration that at least one allele containing a splice site variant, *BRCA1* c.[594-2A>C; 641A>G], does not increase breast cancer risk and the observation that splicing assays may lead to erroneous clinical conclusions if alternate gene transcripts are not properly addressed.[8–11] Recommendations based on these studies are documented in the *ENIGMA BRCA1/2 Gene Variant Classification Criteria* (https://enigmaconsortium.org) that support *BRCA1* and *BRCA2* expert panel review interpretation at ClinVar.

A recent study has identified alternate gene transcripts at the *PALB2* locus, but no inferences in relation to the clinical interpretation of genetic variants were made.[12] Here, we undertake a comprehensive characterisation of *PALB2* alternative splicing, exploring the possible relevance of the findings for the clinical classification of PTC-NMD and splice site variants according to the ACMG-AMP-2015 guidelines.

METHODS

Identification of alternative splicing events

To characterise alternative splicing at the *PALB2* locus, we analysed RNAs isolated from *112* specimens, including lymphoblastic cell lines not treated with the NMD-inhibitor puromycin (*n=68*), matched replicates treated with puromycin (LCLs+Puro, *n=1*), stimulated leucocytes cultures not treated with puromycin (*n=6*), matched replicates treated with puromycin (sLEU+Puro, *n=3*), RNA stabilised peripheral blood samples (PAXgene, QIAGEN, *n=7*; Tempus, Thermo Fisher, *n=10*), non-malignant breast tissue samples from unrelated women (Breast, *n=12*; 10 corresponding to women with a diagnosis of breast cancer, of which 9 are included in SCAN-B, ClinicalTrials.gov identifier: NCT02306096; 2 corresponding to women without a diagnosis of breast cancer included in CASOHAR trial NTC02560818), a human mammary epithelial cell (HMEC, *n=1*, 2 technical replicas included in the analysis), commercially available RNA from non-malignant breast tissue (Clontech *636576*, *n=1*), normal ovarian *fimbriae* tissue samples from prophylactic oophorectomies performed in postmenopausal women without cancer (Fimbriae, *n=2*) and one pool of 3 non-malignant ovarian tissues (Clontech *636555*, *n=1*).

Experiments were performed independently in five ENIGMA laboratories (figure 1). Most samples were analysed by targeted RNAseq (n=72) in laboratory 1 (online supplementary table 1 and 2). Other samples were analysed by whole transcriptome RNAseq (n=13) in laboratories 2 and 3 (online supplementary table 1 and 2), by capillary electrophoresis of RT-PCR products (RT-PCR/CE, n=22) in laboratory 4 (online supplementary table 1, 2, 3 and figures 1A, B), and by whole-gene CloneSeq splicing analysis (n=5) in laboratory 5 (online supplementary figure 1B). We later performed a centralised revision/curation of the data, including the search for putative tissue-specific alternate gene transcripts. To this end, we pooled together all data produced in LCLs±Puro, sLEU±Puro, PAXgene and Tempus samples (hereafter referred collectively as BLOOD), all data produced in non-malignant breast tissues, HMEC and Clontech *636576* (hereafter referred as BREAST) and all data produced in non-malignant ovarian *fimbriae* and Clontech *636555* (hereafter referred as OVARY). The overall workflow is summarised in figure 1 (see online supplementary section 1 for further details).

Annotation of alternative splicing events

We described all alternative splicing events according to HGVS guidelines, using as a reference the Ensembl transcript ENST00000261584.8 (NCBI RefSeq NM_024675.3). For the sake of simplicity, we also identified most events with a code that combines the following symbols: Δ (skipping of reference exonic sequences), ▼ (inclusion of reference intronic sequences), E (exon), I (intron), p (acceptor shift), q (donor shift), AFE (alternative first exon) and IVS± (located at intervening sequence). When necessary, the exact number of nucleotides skipped (or retained) is indicated. Events were annotated as well according to the confidence of the finding (high-confidence vs lower-confidence), predictions on coding potential (LoF vs uncertain) and relative quantification (expression level relative to the corresponding reference transcript) (see online supplementary material section 2 and figures 2-5 for further details).

Analysis of PVS1 status (warranted vs not warranted) for every possible PTC-NMD and splice site variant at the *PALB2* locus

To decide if PVS1 is warranted we used predictions based on: (i) the identification of alternate gene transcripts in control samples, (ii) RNA splicing assays performed previously in carriers of *PALB2* splice site variants (online supplementary table 4) and (iii) novel RNA splicing assays (online supplementary table 4, figures 6A, B and C). In brief, we consider PVS1 warranted for PTC-NMD variants only if no plausible *rescue transcripts* have been detected. Similarly, we consider PVS1 warranted for splice site variants only if all predicted RNA product are bona fide LoF transcripts. To predict possible RNA products, we used splicing assays performed in carriers of splice site variants (assuming that other *PALB2* splice site variants targeting the same splicing site will produce similar transcripts). If no splicing assay was available for a particular splice site, we based predictions on alternate gene transcripts, as previously done for *BRCA1* and *BRCA2*.[9 10] Further details are shown in online supplementary material section 3 and table 4.

RESULTS

We used RNA extracted from different human biological samples (blood-derived, breast and ovary; see 'Methods' section) to characterise naturally occurring alternative splicing at the *PALB2* locus. This study combined targeted RNAseq, whole transcriptome RNAseq, RT-PCR/CE and whole-gene CloneSeq splicing

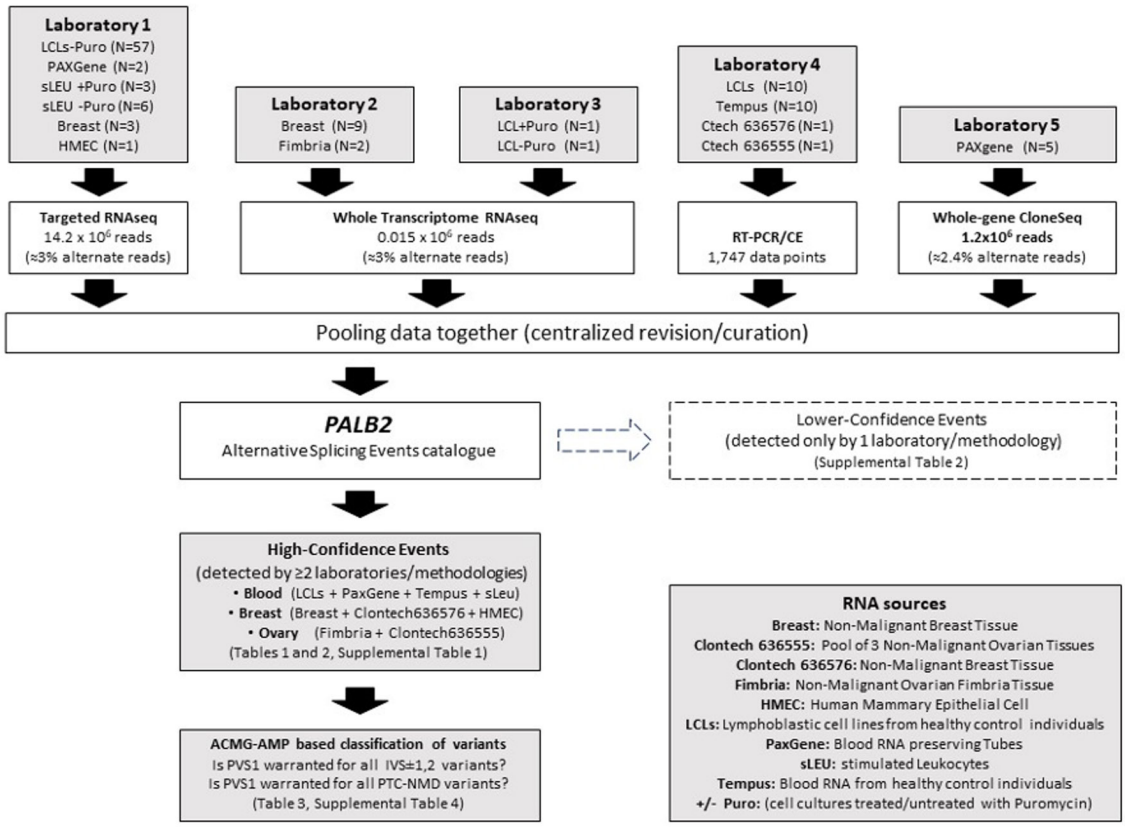

Figure 1 Workflow. The workflow is followed by the Evidence-based Network for the Interpretation of Germ-line Mutant Alleles consortium to characterise the naturally occurring alternative splicing profile at the *PALB2* locus in BLOOD-derived, BREAST-derived and OVARY-derived samples. RNAseq data were produced in five independent laboratories using different methodologies in unrelated samples. Laboratory 1 (Clinical Biology and Oncology Laboratory, Cancer Center François Baclesse, Normandy University Caen, France) performed targeted RNAseq analysis. Laboratories 2 (Division of Oncology and Pathology, Department of Clinical Sciences, Lund University, Sweden) and 3 (Department of Pathology and Biomedical Science, University of Otago Christchurch, New Zealand) performed whole transcriptome RNAseq. Laboratory 4 (Molecular Oncology Laboratory, Academic Hospital San Carlos, Madrid, Spain) performed capillary electrophoresis analysis of real-time PCR products (RT-PCR/CE). Laboratory 5 (Ambry Genetics) performed whole-gene CloneSeq alternative splicing analysis. As indicated, the overall contribution of targeted RNAseq reads to the analysis is roughly 1000× higher than that of whole transcriptome RNAseq. For instance, targeted RNAseq experiments end up with 13 754 118 reads aligned to reference exon-exon junctions, but only 459 186 reads supporting alternative splicing events (≈3%). The same percentage was observed in whole transcriptome RNA experiments, although the total number of reads was much lower (14 933 reads combining data from laboratories 2 and 3). RT-PCR/CE contributed 1747 data points (individual RT-PCR experiments performed with a particular combination of primers in individual samples, including technical replicas). CloneSeq analysis contributed 1.2×10^6 reads (≈2.4% of the reads supporting alternative splicing events). Data were pooled together, reviewed and cross-checked to end up with a list of *high-confidence* naturally occurring alternative splicing events (events detected by different techniques in different samples), and a list of *lower-confidence* splicing events (events not qualifying for higher confidence events). Finally, the possible relevance of high-confidence findings for the initial classification of canonical splicing site and PTC-NMD variants was explored. ACMG-AMP, American College of Medical Genetics and Genomics-Association for Molecular Pathology; HMEC, human mammary epithelial cell; LCL, lymphoblastic cell line; NMD, nonsense-mediated decay.

analysis data that was independently produced at five contributing centres (figure 1). The analysis identified 44 naturally occurring alternative splicing events with high-confidence (online supplementary table 1) and provided evidence for the existence of up to 44 additional (*lower-confidence* events, online supplementary table 2 and supplemental material section 2.2). Most events (37 out of 44 high-confidence and all *lower-confidence* events) have not been described previously in GENCODE (https://www.gencodegenes.org/) or the scientific literature to our knowledge.

Up to 15 high-confidence events preserved a bona fide open reading frame (ie, an ORF spanning from the reference start codon to the reference termination codon, table 1, protein column). Of these, nine were predicted to code for non-functional proteins, and the remaining six for proteins of uncertain functionality (table 1, coding potential column). Twenty-nine high-confidence events did not preserve a bona fide ORF. All of them were predicted to code for non-functional proteins (table 2).

Targeted RNAseq data (online supplemental table 1, laboratory 1) indicated that most *high-confidence* events make on average (n=72 samples) a minor contribution to the expression level (ie, reads supporting the splicing event representing ≤1% of the reads supporting the corresponding reference transcript). The only exceptions were Δ(E1q17), IVS1-463▼(134), Δ(E7p10), Δ(E11), Δ(E11_E12) and Δ(E12), with contributions of ≈2%, ≈5%, ≈1.4%, ≈2%, ≈2% and ≈13%, respectively. In silico analysis suggests that events contributing >1% might be related to the presence of suboptimal splice sites at the

Table 1 High-confidence alternative splicing events at the PALB2 locus (in-frame events)

Designation*	Biotype†	RNA‡	Protein‡	Coding potential§	Rationale§	Blood	Breast	Ovary
▼(AFE600)+ Δ(E1)¶	Terminal modification	r.1_28delins28+805_28+858	p.Asp2_Lys16delins17	Uncertain	Damaging to CC	Yes	Yes	–
▼(E1q9)	Donor shift	r.48_49ins48+1_48+9	(p.Lys16_Leu17ins3)	Uncertain	Uncertain impact on CC	Yes	–	Yes
Δ(E2p6)	Acceptor shift	r.49_54del	(p.Leu17_Lys18del)	Uncertain	Uncertain impact on CC	Yes	Yes	Yes
Δ(E2)	Cassette	r.49_108del	(p.Leu17_Asn36del)	LoF	Damaging to CC	Yes	–	Yes
Δ(E4)	Cassette	r.212_1684del	(p.Glu71_Lys561del)	LoF	Damaging to ChAM	Yes	Yes	Yes
Δ(E5p24)	Acceptor shift	r.1685_1708del	(p.Gly562_Lys569del)	Uncertain	No domain affected	Yes	Yes	Yes
Δ(E6)**	Cassette	r.2515_2586del	(pThr839_Lys862del)	LoF**	Damaging to WD40¥	Yes	Yes	Yes
▼(E7p42)	Acceptor shift	r.2586_2587ins2587-42_2587–1	(p.Lys862_Asn863ins14)	Uncertain	Uncertain impact on WD40	Yes	Yes	Yes
Δ(E7)	Cassette	r.2587_2748del	(p.Arg863_Glu916del)	LoF	Damaging to WD40	Yes	Yes	Yes
Δ(E9p30)	Acceptor shift	r.2835_2864del	(p.Ala946_Glu954del)	LoF	Damaging to WD40	Yes	Yes	Yes
Δ(E9)	Cassette	r.2835_2996del	(p.Ala946_Gly1000del)	LoF	Damaging to WD40	Yes	Yes	Yes
Δ(E9_E10)	Multicassette	r.2835_3113del	(p.Ala946_Trp1038del)	LoF	Damaging to WD40	Yes	Yes	–
Δ(E10p3)	Acceptor shift	r.2997_2999del	(p.Gly1000del)	Uncertain	Uncertain impact on WD40	Yes	Yes	–
Δ(E10)	Cassette	r.2997_3113del	(p.Gly1000_Trp1038del)	LoF	Damaging to WD40	Yes	Yes	Yes
Δ(E11_E12)††	Multicassette	r.3114_3350del	(p.Asn1039_Arg1117del)	LoF	Damaging to WD40	Yes	Yes	Yes

*See supplementary material section 2.1 and figure 2 for details.
†Biotype according to ENCODE.[25]
‡RNA and predicted protein described according to the Human Genome Variation Society guidelines at http://varnomen.hgvs.org/, using Ensembl transcript ENST00000261584.8 as a reference.
§Uncertain coding potential if the transcript encodes a protein predicted to preserve (or partially preserve) functional capacity. See online supplemental material section 2.3 and figure 4 for further details.
¶Only▼(AFE600)+Δ(E1) described in GENCODE (comprehensive gene annotation from GENCODE release 26 retrieved through Ensembl at http://www.ensembl.org/).
**Δ(E6) transcripts code for a hypomorphic protein (instable, but with residual activity).[26]
††Only Δ11_12 described previously in the literature.[12]
CC, N-terminal coiled-coil domain; ChAM, chromatin- associated motif; LOF, loss-of-function; WD40, WD40 β-propeller C-terminal domain.

PALB2 gene (online supplemental figure 7), with Δ(E12) contribution (≈13%) probably explained by the intrinsically weak exon 12 GC donor site.[13] The relatively elevated level of alternative splicing resulting in skipping of exons 11 and/or 12 is supported by targeted and whole transcriptome RNAseq (online supplemental table 1), semi-quantitative RT-PCR/CE analysis (online supplemental figure 1A), whole-gene CloneSeq splicing analysis (online supplemental figure 1B) and quantitative dPCR (online supplemental figure 5B). According to the latter, ≈8%–34% of the *PALB2* transcripts (depending on the sample analysed) may skip exon 11, exon 12 or both.

Overall coverage in whole transcriptome RNAseq was substantially lower than in targeted RNAseq experiments (figure 1). As a result, several events representing ≤1% of the targeted RNAseq reads were not detected by this approach. Only one major discrepancy was observed related to *PALB2* Δ(E4_E5), which represented ≤1% of the corresponding reference signal in targeted RNAseq and whole-exon GenClone experiments, but >5% in RNAseq data generated by laboratory 3. However, subsequent digital PCR quantification in BLOOD, BREAST and OVARY confirmed that Δ(E4_E5) represents, on average, ≤1% of the corresponding reference signal (online supplemental figure 5).

Despite the lower coverage, whole transcriptome RNAseq and/or RT-PCR/CE experiments allowed us to detect 50 splicing events in BREAST, and 29 in OVARY. Of these, 24 splicing events—among them Δ(E1q17), IVS-463▼(134), Δ(E7p10), Δ(E11), Δ(E11_E12) and Δ(E12)—were detected in both tissues (table 1 and online supplemental table 1). Equally relevant, we did not identify tissue-specific *PALB2* alternate gene transcripts (neither in BREAST nor in OVARY), suggesting that if they exist, they are expressed at very low levels—supporting the clinical relevance of BLOOD-based *PALB2* splicing studies.

Finally, we used data on alternate gene transcripts to analyse if PVS1 is warranted for all possible PTC-NMD/splice site variants at the *PALB2* gene. In brief, we concluded that PVS1 is warranted for every possible PTC-NMD variant, regardless of the location, that is, we have not identified any plausible *rescue transcript* (see 'Discussion' section). By contrast, we conclude that PVS1 is not necessarily warranted for every possible splice site variant. To be more precise, we propose that PVS1 may not be warranted for splice site variants located at the acceptor sites of exons 2, 5, 7 and 10. For this subset of splice site variants, the production of RNA transcripts retaining some or all functional capacity is plausible (see table 3 for further details). If splicing assays and/or clinical data supporting pathogenicity are lacking, we recommend caution when classifying splice site variants at these specific sites, that is, such variants should not be assumed *pathogenic*/*likely pathogenic*.

DISCUSSION

Alternative splicing probably occurs in all metazoan organisms, and increasing prevalence has been linked to phenotypic complexity.[14] Virtually all human multiexon *loci* produce alternate gene transcripts.[15] Apart from a presumed role in expanding protein diversity[16] that is currently under dispute,[17 18] some authors have suggested that alternative splicing may buffer mutational consequences.[19] The latter possibility has obvious implications for the clinical interpretation of genetic testing results. The ACMG-AMP-2015 guidelines acknowledge this by recommending caution about overinterpreting the impact of PTC-NMD and splice site variants if multiple transcripts are present.[7] Here, we have addressed this relevant aspect of alternative splicing for the particular case of classifying genetic variants at the breast cancer predisposition gene *PALB2*.

Alternative splicing analysis might be influenced by many factors, including collection of RNA samples, experimental design and detection sensitivity. For instance, one study characterising alternative splicing at breast cancer susceptibility genes by RNAseq noticed the poor performance of PAXgene if compared with LCL samples,[12] and a previous ENIGMA collaborative study comparing RT-PCR splicing protocols across different laboratories concluded that primers design and detection sensitivity (rather than RNA extraction and/or cDNA synthesis protocols) had an impact on the

Table 2 High-confidence alternative splicing events at the PALB2 locus (PTC-NMD events)

Designation*	Biotype†	RNA‡	Protein	Coding potential	Blood	Breast	Ovary
Δ(E1q169)	Donor shift	r.-121_48del	Non-coding	LoF	Yes	Yes	Yes
Δ(E1q17)§¶	Donor shift	r.32_48del	p.Cys11Phefs*25	LoF	Yes	Yes	Yes
▼(E1q337)	Donor shift	r.48_49ins48+1_48+337	p.Leu17Valfs*19	LoF	Yes	–	–
IVS1-463▼(134)§¶	Cassette	r.48_49ins49-463_49–330	p.Leu17Valfs*11	LoF	Yes	Yes	–
▼(E2p26)	Acceptor shift	r.48_49ins49-26_49–1	p.Leu17Tyrfs*9	LoF	Yes	–	Yes
▼(I2)	Intron retention	r.108_109ins108+1_109–1	p.R37_S1186delins11	LoF	Yes	–	–
▼(E3p36)**	Acceptor shift	r.108_109ins109-36_109–1	p.Arg37_Ser1186deldelins11	LoF	Yes	Yes	Yes
▼(E4p25)	Acceptor shift	r.211_212ins212-25_212–1	p.Glu71Valfs*10	LoF	Yes	–	–
Δ(E4_E5)§¶	Multicassette	r.212_2514del	p.Glu71Aspfs*1	LoF	Yes	Yes	–
Δ(E5p139)	Acceptor shift	r.1685_1823del	p.Gly562Valfs*19	LoF	Yes	Yes	–
Δ(E5)	Cassette	r.1685_2514del	p.Gly562Aspfs*1	LoF	Yes	–	–
▼(E6p28)	Acceptor shift	r.2514_2515ins2515-28_2515–1	p.Glu840Asnfs*9	LoF	Yes	Yes	Yes
▼(E7p20)	Acceptor shift	r.2586_2587ins2587-20_2587–1	p.Pro864Cysfs*13	LoF	Yes	–	Yes
Δ(E7p2)	Acceptor shift	r.2587_2588del	p.Asn863Serfs*20	LoF	Yes	Yes	–
Δ(E7p10)	Acceptor shift	r.2587_2596del	p.Asn863Valfs*4	LoF	Yes	Yes	Yes
Δ(E7p25)	Acceptor shift	r.2587_2611del	p.Asn863Metfs*1	LoF	Yes	Yes	Yes
▼(E8p30)††	Acceptor shift	r.2748_2749ins2749-30_2749–1	p.Val917_Ser1186delins9	LoF	Yes	–	Yes
Δ(E8)	Cassette	r.2749_2834del	p.Val917Glyfs*6	LoF	Yes	Yes	Yes
Δ(E8_E9)	Multicassette	r.2749_2996del	p.Val917Argfs*10	LoF	Yes	Yes	–
Δ(E10p2)	Acceptor shift	r.2997_2998del	p.Gly1000Glnfs*9	LoF	Yes	–	–
Δ(E10q31)	Donor shift	r.3083_3113del	p.Thr1029Ilefs*1	LoF	Yes	Yes	–
▼(E11p23)	Acceptor shift	r.113_3114ins3111-23_3114–1	p.Trp1038Cysfs*7	LoF	Yes	Yes	Yes
Δ(E11p2)	Acceptor shift	r.3114_3115del	p.Trp1038Ter	LoF	Yes	Yes	Yes
Δ(E11)§	Cassette	r.3114_3201del	p.Asn1039Glyfs*5	LoF	Yes	Yes	Yes
Δ(E11)+▼(E12p446)	Mixed	r.3114_3201del+r0.3201_3202ins3202-446_3202–1	p.Trp1038Cysfs*3	LoF	Yes	–	–
Δ(E11)+▼(E12p65)	Mixed	r.3114_3201del+r0.3201_3202ins3202-65_3202–1	p.Trp1038Ter	LoF	Yes	–	–
▼(E12p65)	Acceptor shift	r.3201_3202ins3202-65_3202–1	p.Gly1068Ilefs*28	LoF	Yes	Yes	Yes
Δ(E12p136)	Acceptor shift	r.3202_3337del	p.Leu1069Argfs*9	LoF	Yes	–	–
Δ(E12)§¶	Cassette	r.3202_3350del	(p.Gly1068_Ser1186delins4)	LoF	Yes	Yes	Yes

*See 'Methods' section.
†Biotype according to ENCODE.[25]
‡RNA described according to the Human Genome Variation Society rules at http://varnomen.hgvs.org/, using Ensembl transcript ENST00000261584.8 as a reference.
§described previously in the literature.[12]
¶ described in comprenesive gene annotation from GENCODE realese 26 retrieved through Ensembl at http://www.ensembl.org/
**The predicted 36 nucleotides insertion includes an in-frame PTC (p.Arg37_Ser1186delinsKTYFWGCFCLL).
††The predicted 30 nucleotides insertion includes an in-frame PTC (p.Val917_Ser1186delinsHNFWLLCFI).

analytical outcome.[20] A strength of our study design was the application of different assay designs, RNA samples and subsequent levels of sensitivity and/or filtering, by five independent laboratories to identify *PALB2* alternative splicing events (see online supplementary material section 1 for further details). We elected to define high-confidence splicing events as those found in at least two different data sets (the rationale being that events detected by a minimum of two laboratories, two sample types and two methodologies are very unlikely to represent technical artefacts and/or biological outliers), but acknowledge that such definition may lead to exclusion of real events found by a single laboratory. A higher stringency of high-confidence splicing events found by more than two laboratories was not used due to differences in the level of sensitivity between assays.

Overall, we identified 44 high-confidence alternative splicing events at the *PALB2 locus*, and we provide evidence for 44 additional events (although we cannot discard the possibility that some of the latter represent technical artefacts and/or biological outliers). Interestingly, all *PALB2* reference exons are affected by one or more *high-confidence* alternative splicing events, suggesting that no *PALB2* exon should be annotated as constitutive. Despite the considerable number of alternative splicing events identified, our data suggest that their contribution to the overall *PALB2* expression is low in all three tissues investigated. Splice site and PTC-NMD variants in cancer susceptibility genes can be overinterpreted (misinterpreted as pathogenic), if alternate gene transcripts are not properly considered.[7 10 11 21–23] In the past, this has led to errors in the clinical management of families carrying the *BRCA1* allele c.[594-2A>C; 641A>G].[23] The low level of alternative splicing observed for *PALB2* in BLOOD, BREAST and OVARY suggests that overinterpreting genetic variants at this locus is less likely to occur. However, some of the alternative splicing events we report can be relevant for the clinical interpretation of *PALB2* PTC-NMD and splice site variants, in particular to decide if PVS1 is warranted.

PTC-NMD variants: the existence of *rescue* transcripts reducing or eliminating the functional and clinical impact of certain PTC-NMD variants in cancer susceptibility genes has been confirmed for *APC*[22] and *BRCA1*.[11] More specifically, the alternate gene transcript APC Δ(E9p303) explains the association of PTC-NMD variants located at codons 312–412 with mild disease,[22] and the alternate gene transcript *BRCA1* Δ(E9_E10) explains the low breast cancer risk observed in carriers of the splice site variant *BRCA1* c.594-2A>C.[11] However, we

Table 3 Proposed classification of *PALB2* splice site variants according to the ACMG-AMP-2015 guidelines (based solely on location and MAF)

Splice site variant		Predicted RNA products/coding potential*		PVS1*	gnomAD†	PM2†	Classification‡
		*LoF**	Uncertain*				
E1 donor	c.48+1,2	Δ(E1q17)§	–	Warranted	–	Yes	Likely pathogenic
E2 acceptor	c.49–1,2	–	Δ(E2p6)§	Not warranted	NFE (1allele)	Yes	**Uncertain significance**
E2 donor	c.108+1,2	Δ(E2)/▼(I2)	–	Warranted	–	Yes	Likely pathogenic
E3 acceptor	c.109–1,2	▼(E3p36)/Δ(E3)	–	Warranted	–	Yes	Likely pathogenic
E3 donor	c.211+1,2	Δ(E3)	–	Warranted	–	Yes	Likely pathogenic
E4 acceptor	c.212–1,2	Δ(E4_E5)†	–	Warranted	NFE (1 allele)	Yes	Likely pathogenic
E4 donor	c.1684+1,2	Δ(E4_E5)†	–	Warranted	–	Yes	Likely pathogenic
E5 acceptor	c.1685–1,2	Δ(E5)	Δ(E5p24)	Not warranted	NFE (1 allele)	Yes	**Uncertain significance**
E5 donor	c.2514+1,2	Δ(E5)	–	Warranted	SAS (1 allele)	Yes	Likely pathogenic
E6 acceptor	c.2515–1,2	Δ(E6)†	–	Warranted	AMR (1 allele)	Yes	Likely pathogenic
E6 donor	c.2586+1,2	Δ(E6)†	–	Warranted	SAS (1 allele)	Yes	Likely pathogenic
E7 acceptor	c.2587–1,2	▼(E7p20)/Δ(E7p2)/Δ(E7p10)/Δ(E7p25)/Δ(E7)	▼(E7p42)	Not warranted	SAS (1allele)	Yes	**Uncertain significance**
E7 donor	c.2748+1,2	Δ(E7)§	–	Warranted	NFE (1 allele)	Yes	Likely pathogenic
E8 acceptor	c.2749–1,2	▼(E8p30)/Δ(E8)	–	Warranted	–	Yes	Likely pathogenic
E8 donor	c.2834+1,2	Δ(E8)	–	Warranted	–	Yes	Likely pathogenic
E9 acceptor	c.2835–1,2	Δ(E9p30)§/Δ(E9)§	–	Warranted	–	Yes	Likely pathogenic
E9 donor	c.2996+1,2	Δ(E9)/Δ(E9_E10)	–	Warranted	–	Yes	Likely pathogenic
E10 acceptor	c.2997–1,2	Δ(E10p2)/Δ(E9_E10)/Δ(E10)	Δ(E10p3)	Not warranted	SAS (1 allele)	Yes	**Uncertain significance**
E10 donor	c.3113+1,2	Δ(E10q31)§/Δ(E9_E10)§/Δ(E10)§	–	Warranted	–	Yes	Likely pathogenic
E11 acceptor	c.3114–1,2	Δ(E11)/Δ(E11p2)/Δ(E11p23)/Δ(E11_E12)	–	Warranted	–	Yes	Likely pathogenic
E11 donor	c.3201+1,2	Δ(E11)/Δ(E11_E12)	–	Warranted	–	Yes	Likely pathogenic
E12 acceptor	c.3202–1,2	▼(E12p65)/Δ(E12p136)/Δ(E11_E12)/Δ(E12)	–	Warranted	–	Yes	Likely pathogenic
E12 donor	c.3350+1,2	Δ(E11_E12)§/Δ(E12)§-		Warranted	–	Yes	Likely pathogenic
E13 acceptor	c.3351–1,2	–	–	Warranted	–	Yes	Likely pathogenic

*If available (§), predictions on possible RNA products are based on splicing assays performed in representative examples of splice site variants (see online supplementary table 4). If not, predictions are based on the possible upregulation of naturally occurring alternate gene transcripts. Predicted RNA products are classified according to their coding potential as loss-of-function (LoF) or uncertain (the possibility of coding for a functional or partially functional protein cannot be disregarded). If only LoF transcripts are predicted, we assume that PVS1 is warranted. If ≥1 transcript with uncertain coding potential is predicted, we propose that PVS1 (based solely on variant location) is not warranted.

†After reviewing gnomAD, we conclude that PM2 is met for all possible splice site variants.

‡According to the ACMG-AMP-2015 guidelines, if PVS1 and PM2 are warranted, splice site variants should be classified as likely pathogenic. Otherwise, splice site variants should be classified as uncertain significance. This analysis has highlighted seven splice site variants in ClinVar needing additional justification for assertion as pathogenic/likely pathogenic (see online supplementary table 5 for further details).

ACMG-AMP, American College of Medical Genetics and Genomics-Association for Molecular Pathology; AMR, American; NFE, non-finish Europeans; SAS, South Asia.

Table 4 Known *PALB2* splice site variants for which we put a warning

Splicing site	Variant reported	dbSNP	ClinVar			Proposed ACMG-AMP-2015 classification
			Classification	Review status	Assertion method	
E2 acceptor	c.49-2A>T	rs786203245	Likely pathogenic	**	Ambry autosomal dominant Invitae Variant Classification Sherlock	Uncertain significance
E5 acceptor	c.1685-2A>G	rs754660432	Likely pathogenic	**	GeneDx variant classification Ambry autosomal dominant	
	c.1685–1G>C	rs1057520645	Pathogenic	*	GeneDx variant classification	
E7 acceptor	c.2587-2A>C	rs1060502787	Likely pathogenic	*	Invitae Variant Classification Sherlock	
E10 acceptor	c.2997-2A>C	–	Likely pathogenic	*	Ambry autosomal dominant	

These five *PALB2* variants are classified as pathogenic/likely pathogenic based on assertion criteria defined by the submitters. Ambry Genetics and/or GeneDx classify the indicated variants as pathogenic based on the fact that these are very rare variants located at canonical splice sites, predicted to abolish or significantly reduce native site using in silico predictors and identified in affected/+family history cohort. Invitae classifies the indicated variants as likely pathogenic based on the fact that donor and acceptor splice site variants are typically loss-of-function and loss-of-function variants in *PALB2* are known to be pathogenic. Remarkably, for any of these variants classification is based on splicing assays, and/or in segregation information supporting pathogenicity (Tina Pesaran, unpublished data; Kathleen S Hruska, unpublished data, Inviate ClinVar summary evidences). These are splice site variants targeting acceptor sites for which, in our opinion (table 3), PVS1 is not necessarily warranted. For that reason, we propose that, in absence of functional and/or genetic data, these variants should be classified according to the ACMG-AMP-2015 guidelines as uncertain significance.

ACMG-AMP, American College of Medical Genetics and Genomics-Association for Molecular Pathology.

have not identified plausible rescue transcripts for *PALB2*. Alternate gene transcripts Δ(E2p6), Δ(E6), Δ(E5p24) and Δ(E10p3) might code for functional or partially functional proteins, but their respective contribution to the overall *PALB2* expression (<1%) is too low to be plausible *rescue* transcripts. By contrast, the combined expression of Δ(E11_E12) and Δ(E12) might represent 8%–34% of the overall gene expression (depending on samples and methodologies), but the predicted proteins encoded by these two transcripts (table 1) are unlike to be functional, as they lack part of the C-terminal WD40 β-propeller domain (online supplementary material section 2.3) that mediates PALB2 interaction with several key homologous recombination proteins, including BRCA2 and RAD51.[24] For that reason, we do not consider Δ(E11_E12) and Δ(E12) plausible rescue transcripts, although we cannot rule out the possibility of truncating variants in exons 11 and/or 12 conferring lower cancer risk than truncating variants in other *PALB2* exons.

Canonical ±1,2 splice site variants: we propose that naturally occurring alternate gene transcripts provide predictive information identifying seven *PALB2* canonical splice sites for which, in absence of splicing assays, PVS1 is not warranted (variants targeting exons 2, 5, 7 and 10 acceptor sites). For exon 2 acceptor site, the proposal is based on experimental data obtained in a *PALB2* c.49–1G>A (IVS1-1G>A) carrier indicating upregulation of Δ(E2p6) (Dr Georgios Tsaousis, Genekor Medical, personal communication, June 2018). The possibility that Δ(E2p6) code for a functional/partially functional protein cannot be discarded (see online supplementary material section 2.3), supporting our conservative stance. For the remaining splice sites, we hypothesise that naturally occurring alternate gene transcripts (even if lowly expressed in control samples) may become upregulated if splice site variants impair the expression of reference transcripts. The hypothesis is supported by several observations made in carriers of *PALB2* (among them, the upregulation of Δ(E2p6) in c.49–1G>A carriers), *BRCA1* and *BRCA2* splice site variants (see online supplementary table 4). Note that we propose that PVS1 is not warranted for splice site variants if at least one RNA product with uncertain coding potential is predicted, regardless of other predictions. For instance, we propose that PVS1 is not warranted for variants targeting the *PALB2* exon 7 acceptor site because one RNA product of uncertain coding potential, ▼(E7p42), is predicted (table 3), despite the fact that up to five bona fide LoF transcripts are also predicted (▼(E7p20), Δ(E7p2), Δ(E7p10), Δ(E7p25) and Δ(E7)). When classifying splice site variants in high-risk breast cancer genes as pathogenic/likely pathogenic without functional or genetic data, we favour a very conservative approach. We have identified 43 different *PALB2* splice site variants in ClinVar (last accessed 13 April 2018), all of them reported as pathogenic/likely pathogenic. For four of these variants, we think that the pathogenic/likely pathogenic classification may not be justified without considering additional clinical and/or splicing data (table 4).

In short, we highlight the fact that, where alternate gene transcripts exist, assertions of pathogenicity are warranted only with the support of additional quantitative splicing assays, and preferably clinical evidence.

Author affiliations
[1]Molecular Oncology Laboratory CIBERONC, Hospital Clínico San Carlos, IdISSC (Instituto de Investigación Sanitaria del Hospital Clínico San Carlos), Madrid, Spain
[2]Laboratory of Clinical Biology and Oncology, Centre François Baclesse, Inserm U1245 Genomics and Personalized Medicine in Cancer and Neurological Disorders, Normandy University, Caen, France
[3]Department of Pathology and Biomedical Science, University of Otago, Christchurch, New Zealand
[4]Inserm U1245 Genomics and Personalized Medecine in Cancer and Neurological Disorders, UNIROUEN, Normandie Université, Normandy Centre for Genomic and Personalized Medicine, Rouen, France
[5]GeneDx, Gaithersburg, Maryland, USA
[6]Ambry Genetics, Aliso Viejo, CA, USA
[7]Division of Oncology and Pathology, Department of Clinical Sciences Lund, Lund University, Lund, Sweden
[8]Peter MacCallum Cancer Centre, Melbourne, VIC, Australia
[9]The Sir Peter MacCallum Department of Oncology, University of Melbourne, Parkville, Australia
[10]Splicing and genetic susceptibility to cancer, Instituto de Biología y Genética Molecular (CSIC-UVa), Valladolid, Spain
[11]Genetic Epidemiology Laboratory, Department of Clinical Pathology, The University of Melbourne, Melbourne, VIC, Australia
[12]Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands
[13]Molecular Cancer Epidemiology Laboratory, QIMR Berghofer Medical Research Institute, Brisbane, Australia

Correction notice This article has been corrected since it was published Online First. The affiliations have been corrected.

Acknowledgements The authors would like to thank A Leconte at cancer center, F Baclesse and to Dr C Baudouin at Polyclinique du Parc, Caen (France) for their participation to CASOHAR clinical trial to obtain breast tissue from healthy volunteers. The authors would like to thank Heather Thorne, Eveline Niedermayr, all the kConFab research nurses and staff, the heads and staff of the Family Cancer Clinics and the Clinical Follow-Up Study (which has received funding from the NHMRC, the National Breast Cancer Foundation, Cancer Australia and the National Institute of Health [USA]) for their contributions to this resource, and the many families who contribute to kConFab. The authors would like to thank the SCAN-B collaborators at participating hospitals for support of SCAN-B. The authors would also like to thank Ingrid Hedenfalk and staff at the collaborating gynaecology and pathology clinics in Lund for providing the fimbriae samples.

Contributors IL-P, RL, RB, VL, JFP, LC, AM, DV, NG, GD, PG, VG-B, PLI, PP-S, ED-R, TC, KSH, VH, SW, TP, RK, JV-C, AV-P and MS, contributed to data acquisition, revised the manuscript for important intellectual content and approved the final version. kConFAB provided research resources used in this study. ABS coordinated the ENIGMA consortium. AB, EAV, MPV, PD, AK, ABS, LW and SK contributed to the conception and design of the study, contributed to obtain all necessary approvals and clearances to conduct the research, contributed to data acquisition, contributed to data analysis, contributed to grant funding, revised the manuscript for important intellectual content and approved the final version. MdH contributed to obtain all necessary approvals and clearances to conduct the research, contributed to the conception and design of the study, contributed to data acquisition, contributed to data analysis, contributed to grant funding, wrote the manuscript and approved the final version.

Funding PD, MPGW, EAV, AB, AK and MH have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 634935. RL is supported by a Normandy-University, Federation-Hospitalo-Universitaire (FHU) grant. VL is supported by a Mackenzie Family Cancer Postdoctoral Fellowship. RB is supported by funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 634935. AM is supported by a French Cancéropôle Nord-Ouest (CNO) grant. AB is supported by Mrs Berta Kamprad Foundation. EAV and MH are supported by Spanish Instituto de Salud Carlos III (ISCIII) funding (grants PI17/00227 to EAV and PI15/00059 to MH), an initiative of the Spanish Ministry of Economy and Innovation partially supported by European Regional Development FEDER Funds. ABS is supported by an NHMRC Senior Research Fellowship (ID1061779). LCW is supported by the Rutherford Discovery Fellowship. SK is supported by Ligue Contre le Cancer, Normandie. kConFab is supported by a grant from the National Breast Cancer Foundation, and previously by the National Health and Medical Research Council (NHMRC), the Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania and South Australia and the Cancer Foundation of Western Australia.

Competing interests VH, SW, PT, RK were employees of Ambry Genetics when they were engaged with this project. KSH was employee of GeneDx when she was engaged with this project. EDR has consulting or advisory roles in Amgen, Bayer, Genómica, Servier and Merck. EDR has got research funding from: Roche, Merck-Serono, Amgen, AstraZeneca and Sysmex.

Patient consent for publication Not required.

Ethics approval Ethics approval Academic Hospital San Carlos ethics committee (reference numbers 15/139 E and 16/505 E). The SCAN-B study has been approved by the Lund Regional Ethical Review Board, Sweden (approval number 2009/658). The fimbriae tissue samples were obtained and analysed with approval by the Lund Regional Ethical Review Board, Sweden (approval number 2014/717). French Biomedicine Agency. CASOHAR trial ethic committee (NTC NTC02560818). Ambry

Genetics' patient's information has been de-identified, and this study has been approved and carried out in accordance with the recommendations of the Western Institutional Review Board (WIRB; IRB Tracking Number:20171324). Whole-transcriptome RNAseq study was approved by the New Zealand Southern Health and Disability Ethics Committee (12/STH/44).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Targeted RNAseq data contributed by laboratory 1 is available from Dr Sophie Krieger on reasonable request. Whole-transcriptome RNAseq data generated by laboratory 3 is available from Dr Logan Walker on reasonable request. Targeted RNAseq data contributed by laboratory 2 is available from SCAN-B and Ingrid Hedenfalk, respectively, but restrictions apply to the availability of these data, which were used under licence for the current study, and so are not publicly available. Data are however available from the authors on reasonable request and with permission of SCAN-B or Ingrid Hedenfalk.

REFERENCES

1 Antoniou AC, Casadei S, Heikkinen T, Barrowdale D, Pylkäs K, Roberts J, Lee A, Subramanian D, De Leeneer K, Fostira F, Tomiak E, Neuhausen SL, Teo ZL, Khan S, Aittomäki K, Moilanen JS, Turnbull C, Seal S, Mannermaa A, Kallioniemi A, Lindeman GJ, Buys SS, Andrulis IL, Radice P, Tondini C, Manoukian S, Toland AE, Miron P, Weitzel JN, Domchek SM, Poppe B, Claes KB, Yannoukakos D, Concannon P, Bernstein JL, James PA, Easton DF, Goldgar DE, Hopper JL, Rahman N, Peterlongo P, Nevanlinna H, King MC, Couch FJ, Southey MC, Winqvist R, Foulkes WD, Tischkowitz M. Breast-cancer risk in families with mutations in PALB2. *N Engl J Med* 2014;371:497–506.

2 Easton DF, Pharoah PD, Antoniou AC, Tischkowitz M, Tavtigian SV, Nathanson KL, Devilee P, Meindl A, Couch FJ, Southey M, Goldgar DE, Evans DG, Chenevix-Trench G, Rahman N, Robson M, Domchek SM, Foulkes WD. Gene-panel sequencing and the prediction of breast-cancer risk. *N Engl J Med* 2015;372:2243–57.

3 Desmond A, Kurian AW, Gabree M, Mills MA, Anderson MJ, Kobayashi Y, Horick N, Yang S, Shannon KM, Tung N, Ford JM, Lincoln SE, Ellisen LW. Clinical Actionability of Multigene Panel Testing for Hereditary Breast and Ovarian Cancer Risk Assessment. *JAMA Oncol* 2015;1:943–51.

4 Graffeo R, Livraghi L, Pagani O, Goldhirsch A, Partridge AH, Garber JE. Time to incorporate germline multigene panel testing into breast and ovarian cancer patient care. *Breast Cancer Res Treat* 2016;160:393–410.

5 Daly MB, Pilarski R, Berry M, Buys SS, Farmer M, Friedman S, Garber JE, Kauff ND, Khan S, Klein C, Kohlmann W, Kurian A, Litton JK, Madlensky L, Merajver SD, Offit K, Pal T, Reiser G, Shannon KM, Swisher E, Vinayak S, Voian NC, Weitzel JN, Wick MJ, Wiesner GL, Dwyer M, Darlow S. NCCN Guidelines Insights: Genetic/Familial High-Risk Assessment: Breast and Ovarian, Version 2.2017. *Journal of the National Comprehensive Cancer Network* 2017;15:9–20.

6 Eccles DM, Mitchell G, Monteiro AN, Schmutzler R, Couch FJ, Spurdle AB, Gómez-García EB. ENIGMA Clinical Working Group. BRCA1 and BRCA2 genetic testing-pitfalls and recommendations for managing variants of uncertain clinical significance. *Ann Oncol* 2015;26:2057–65.

7 Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL. ACMG Laboratory Quality Assurance Committee. 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;17:405–23.

8 Colombo M, Blok MJ, Whiley P, Santamariña M, Gutiérrez-Enríquez S, Romero A, Garre P, Becker A, Smith LD, De Vecchi G, Brandão RD, Tserpelis D, Brown M, Blanco A, Bonache S, Menéndez M, Houdayer C, Foglia C, Fackenthal JD, Baralle D, Wappenschmidt B, Díaz-Rubio E, Caldés T, Walker L, Díez O, Vega A, Spurdle AB, Radice P, De La Hoya M. kConFaB Investigators. Comprehensive annotation of splice junctions supports pervasive alternative splicing at the BRCA1 locus: a report from the ENIGMA consortium. *Hum Mol Genet* 2014;23:3666–80.

9 Fackenthal JD, Yoshimatsu T, Zhang B, de Garibay GR, Colombo M, De Vecchi G, Ayoub SC, Lal K, Olopade OI, Vega A, Santamariña M, Blanco A, Wappenschmidt B, Becker A, Houdayer C, Walker LC, López-Perolio I, Thomassen M, Parsons M, Whiley P, Blok MJ, Brandão RD, Tserpelis D, Baralle D, Montalban G, Gutiérrez-Enríquez S, Díez O, Lazaro C, Spurdle AB, Radice P, de la Hoya M. kConFaB Investigators. Naturally occurring BRCA2 alternative mRNA splicing events in clinically relevant samples. *J Med Genet* 2016;53:548–58.

10 Dosil V, Tosar A, Cañadas C, Pérez-Segura P, Díaz-Rubio E, Caldés T, de la Hoya M. Alternative splicing and molecular characterization of splice site variants: BRCA1 c.591C>T as a case study. *Clin Chem* 2010;56:53–61.

11 de la Hoya M, Soukarieh O, López-Perolio I, Vega A, Walker LC, van Ierland Y, Baralle D, Santamariña M, Lattimore V, Wijnen J, Whiley P, Blanco A, Raponi M, Hauke J, Wappenschmidt B, Becker A, Hansen TV, Behar R, Investigators K, Niederacher D, Arnold N, Dworniczak B, Steinemann D, Faust U, Rubinstein W, Hulick PJ, Houdayer C, Caputo SM, Castera L, Pesaran T, Chao E, Brewer C, Southey MC, van Asperen CJ, Singer CF, Sullivan J, Poplawski N, Mai P, Peto J, Johnson N, Burwinkel B, Surowy H, Bojesen SE, Flyger H, Lindblom A, Margolin S, Chang-Claude J, Rudolph A, Radice P, Galastri L, Olson JE, Hallberg E, Giles GG, Milne RL, Andrulis IL, Glendon G, Hall P, Czene K, Blows F, Shah M, Wang Q, Dennis J, Michailidou K, McGuffog L, Bolla MK, Antoniou AC, Easton DF, Couch FJ, Tavtigian S, Vreeswijk MP, Parsons M, Meeks HD, Martins A, Goldgar DE, Spurdle AB. Combined genetic and splicing analysis of BRCA1 c.[594-2A>C; 641A>G] highlights the relevance of naturally occurring in-frame transcripts for developing disease gene variant classification algorithms. *Hum Mol Genet* 2016;25:2256–68.

12 Davy G, Rousselin A, Goardon N, Castéra L, Harter V, Legros A, Muller E, Fouillet R, Brault B, Smirnova AS, Lemoine F, de la Grange P, Guillaud-Bataille M, Caux-Moncoutier V, Houdayer C, Bonnet F, Blanc-Fournier C, Gaildrat P, Frebourg T, Martins A, Vaur D, Krieger S. Detecting splicing patterns in genes involved in hereditary breast and ovarian cancer. *Eur J Hum Genet* 2017;25:1147–54.

13 Thanaraj TA, Clark F. Human GC-AG alternative intron isoforms with weak donor sites show enhanced consensus at acceptor exon positions. *Nucleic Acids Res* 2001;29:2581–93.

14 Kim E, Magen A, Ast G. Different levels of alternative splicing among eukaryotes. *Nucleic Acids Res* 2007;35:125–31.

15 Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F, Xue C, Marinov GK, Khatun J, Williams BA, Zaleski C, Rozowsky J, Röder M, Kokocinski F, Abdelhamid RF, Alioto T, Antoshechkin I, Baer MT, Bar NS, Batut P, Bell K, Bell I, Chakrabortty S, Chen X, Chrast J, Curado J, Derrien T, Drenkow J, Dumais E, Dumais J, Duttagupta R, Falconnet E, Fastuca M, Fejes-Toth K, Ferreira P, Foissac S, Fullwood MJ, Gao H, Gonzalez D, Gordon A, Gunawardena H, Howald C, Jha S, Johnson R, Kapranov P, King B, Kingswood C, Luo OJ, Park E, Persaud K, Preall JB, Ribeca P, Risk B, Robyr D, Sammeth M, Schaffer L, See LH, Shahab A, Skancke J, Suzuki AM, Takahashi H, Tilgner H, Trout D, Walters N, Wang H, Wrobel J, Yu Y, Ruan X, Hayashizaki Y, Harrow J, Gerstein M, Hubbard T, Reymond A, Antonarakis SE, Hannon G, Giddings MC, Ruan Y, Wold B, Carninci P, Guigó R, Gingeras TR. Landscape of transcription in human cells. *Nature* 2012;489:101–8.

16 Blencowe BJ. Alternative splicing: new insights from global analyses. *Cell* 2006;126:37–47.

17 Tress ML, Abascal F, Valencia A. Most Alternative Isoforms Are Not Functionally Important. *Trends Biochem Sci* 2017;42:408–10.

18 Blencowe BJ. The Relationship between Alternative Splicing and Proteomic Complexity. *Trends Biochem Sci* 2017;42:407–8.

19 Niklas KJ, Bondos SE, Dunker AK, Newman SA. Rethinking gene regulatory networks in light of alternative splicing, intrinsically disordered protein domains, and post-translational modifications. *Front Cell Dev Biol* 2015;3:8.

20 Whiley PJ, de la Hoya M, Thomassen M, Becker A, Brandão R, Pedersen IS, Montagna M, Menéndez M, Quiles F, Gutiérrez-Enríquez S, De Leeneer K, Tenés A, Montalban G, Tserpelis D, Yoshimatsu T, Tirapo C, Raponi M, Caldes T, Blanco A, Santamariña M, Guidugli L, de Garibay GR, Wong M, Tancredi M, Fachal L, Ding YC, Kruse T, Lattimore V, Kwong A, Chan TL, Colombo M, De Vecchi G, Caligo M, Baralle D, Lázaro C, Couch F, Radice P, Southey MC, Neuhausen S, Houdayer C, Fackenthal J, Hansen TV, Vega A, Diez O, Blok R, Claes K, Wappenschmidt B, Walker L, Spurdle AB, Brown MA. ENIGMA consortium. Comparison of mRNA splicing assay protocols across multiple laboratories: recommendations for best practice in standardized clinical testing. *Clin Chem* 2014;60:341–52.

21 Tesoriero AA, Wong EM, Jenkins MA, Hopper JL, Brown MA, Chenevix-Trench G, Spurdle AB, Southey MC. kConFab. Molecular characterization and cancer risk associated with BRCA1 and BRCA2 splice site variants identified in multiple-case breast cancer families. *Hum Mutat* 2005;26:495.

22 Nieuwenhuis MH, Vasen HF. Correlations between mutation site in APC and phenotype of familial adenomatous polyposis (FAP): a review of the literature. *Crit Rev Oncol Hematol* 2007;61:153–61.

23 Rosenthal ET, Bowles KR, Pruss D, van Kan A, Vail PJ, McElroy H, Wenstrup RJ. Exceptions to the rule: case studies in the prediction of pathogenicity for genetic variants in hereditary cancer genes. *Clin Genet* 2015;88:533–41.

24 Park JY, Zhang F, Andreassen PR. PALB2: the hub of a network of tumor suppressors involved in DNA damage responses. *Biochim Biophys Acta* 2014;1846:263–75.

25 Mudge JM, Frankish A, Fernandez-Banet J, Alioto T, Derrien T, Howald C, Reymond A, Guigó R, Hubbard T, Harrow J. The origins, evolution, and functional potential of alternative splicing in vertebrates. *Mol Biol Evol* 2011;28:2949–59.

26 Byrd PJ, Stewart GS, Smith A, Eaton C, Taylor AJ, Guy C, Eringyte I, Fooks P, Last JI, Horsley R, Oliver AW, Janic D, Dokmanovic L, Stankovic T, Taylor AM. A Hypomorphic PALB2 Allele Gives Rise to an Unusual Form of FA-N Associated with Lymphoid Tumour Development. *PLoS Genet* 2016;12:e1005945.