Received: 19 June 2017 | Revised: 3 November 2017 | Accepted: 17 December 2017

DOI: 10.1002/humu.23390

RESEARCH ARTICLE

WILEY HGVS HUMAN GENOME VARIATION SOCIETY

Thorough in silico and in vitro cDNA analysis of 21 putative *BRCA1* and *BRCA2* splice variants and a complex tandem duplication in *BRCA2* allowing the identification of activated cryptic splice donor sites in *BRCA2* exon 11

Annelot Baert[1,2] | Eva Machackova[3] | Ilse Coene[2] | Carol Cremin[4] | Kristin Turner[4] | Cheryl Portigal-Todd[4] | Marie Jill Asrat[4] | Jennifer Nuk[4] | Allison Mindlin[4] | Sean Young[4,5] | Andree MacMillan[6] | Tom Van Maerken[2] | Martin Trbusek[7] | Wendy McKinnon[8] | Marie E. Wood[8] | William D. Foulkes[9] | Marta Santamariña[10] | Miguel de la Hoya[11] | Lenka Foretova[3] | Bruce Poppe[2] | Anne Vral[1] | Toon Rosseel[2] | Kim De Leeneer[2] | Ana Vega[10] | Kathleen B. M. Claes[2]

[1]Department of Basic Medical Sciences, Ghent University, Ghent, Belgium

[2]Center for Medical Genetics, Ghent University Hospital, Ghent, Belgium

[3]Department of Cancer Epidemiology and Genetics, Masaryk Memorial Cancer Institute, Brno, Czech Republic

[4]BC Cancer Agency, Vancouver, British Columbia, Canada

[5]Cancer Genetics and Genomics Laboratory, Department of Pathology and Laboratory Medicine, BC Cancer Agency, Vancouver, British Columbia, Canada

[6]Provincial Medical Genetics Program, Eastern Health, St. John's, Newfoundland and Labrador, Canada

[7]Department of Internal Medicine - Hematology and Oncology, University Hospital Brno, Brno, Czech Republic

[8]Familial Cancer Program, University of Vermont Medical Center, Burlington, Vermont, United States

[9]Cancer Research Program, Research Institute of the McGill University Health Centre, McGill University, Montreal, Quebec, Canada

[10]Fundación Pública Galega de Medicina Xenómica-SERGAS, Grupo de Medicina Xenómica, CIBERER, IDIS, Santiago de Compostela, Spain

[11]Molecular Oncology Laboratory CIBERONC, Hospital Clinico San Carlos, IdISSC (Instituto de Investigación Sanitaria del Hospital Clínico San Carlos), Madrid, Spain

Correspondence
Kathleen B. M. Claes, Center for Medical Genetics, Ghent University Hospital, Corneel Heymanslaan 10, 9000 Gent.
Email: Kathleen.Claes@UGent.be

Funding information
Contract grant sponsors: Stichting tegen Kanker (2012-216); Spanish Ministry of Economy and Innovation; European Regional Development FEDER Funds (PI15/00059); CIBERER grant (ER17P1AC7112/2017); MH CZ - DRO (MMCI, 00209805); Ministry of Education, Youth and Science (NPS I - LO1413). European Union's Horizon 2020 research and innovation programme (No 634935); Spanish Instituto de Salud Carlos III funding (grant PI15/00059); Canadian Institute of Health Research Foundation grant (FDN-148390)

Communicated by Rolf H. Sijmons

Abstract

For 21 putative *BRCA1* and *BRCA2* splice site variants, the concordance between mRNA analysis and predictions by in silico programs was evaluated. Aberrant splicing was confirmed for 12 alterations. In silico prediction tools were helpful to determine for which variants cDNA analysis is warranted, however, predictions for variants in the Cartegni consensus region but outside the canonical sites, were less reliable. Learning algorithms like Adaboost and Random Forest outperformed the classical tools. Further validations are warranted prior to implementation of these novel tools in clinical settings. Additionally, we report here for the first time activated cryptic donor sites in the large exon 11 of *BRCA2* by evaluating the effect at the cDNA level of a novel tandem duplication (5′ breakpoint in intron 4; 3′ breakpoint in exon 11) and of a variant disrupting the splice donor site of exon 11 (c.6841+1G > C). Additional sites were predicted, but not activated. These sites warrant further research to increase our knowledge on *cis* and *trans* acting factors involved in the conservation of correct transcription of this large exon. This may contribute to adequate design of ASOs (antisense oligonucleotides), an emerging therapy to render cancer cells sensitive to PARP inhibitor and platinum therapies.

KEYWORDS
alternative splicing, BRCA1/2, cDNA analysis, in silico prediction tools, learning algorithms

1 | INTRODUCTION

The risk of breast and ovarian cancer increases drastically in individuals carrying a germline *BRCA1* (MIM# 113705) or *BRCA2* (MIM# 600185) mutation. Heterozygous pathogenic mutations are associated with a lifetime risk of 50%–80% for breast cancer and 30%–50% for ovarian cancer (Roy, Chun, & Powell, 2012). However, due to an evolution in sequencing technologies, the number of individuals undergoing screening drastically increased and this led to a significant rise in the number of variants of unknown clinical significance (VUS). The detection of a VUS is a challenge for healthcare providers as the impact on the risk for breast or ovarian cancer risk is unclear (Goldgar et al., 2004). For missense variants, a number of functional tests have been proposed (e.g., Hendriks et al., 2014; Millot, Carvalho, Caputo, & Vreeswijk, 2012), but the implementation on a large scale in a clinical diagnostic setting is not feasible. In contrast, mRNA analyses to investigate intronic and exonic variants that might impair proper RNA splicing are more widely used. The effect of putative splice site variants can be evaluated in silico as a wide variety of tools have been developed (Jian, Boerwinkle, & Liu, 2014b). Indeed, variants in the highly conserved dinucleotides at the splice site donor/acceptor sites at the intron boundaries, in the conserved adjacent sequences or variants in the branchpoints and exonic splicing enhancers (ESE) might lead to aberrant splicing (Houdayer et al., 2012; Spurdle, Couch, Hogervorst, Radice, & Sinilnikova, 2008; Strachan & Read, 2010). Several guidelines have been formulated to assist the decision-making process (Caminsky, Mucaki, & Rogan, 2014; Houdayer et al., 2008; Tang, Prosser, & Love, 2016). Accurate prediction by these tools is dependent on the ability to detect the wild-type (WT) splice site or define regions such as the branchpoints or ESE. Furthermore, it depends on the degree of conservation of the region in which the variant is situated (Caminsky et al., 2014). Several in silico prediction tools have previously been extensively validated for the prediction of aberrant splicing (Caminsky et al., 2014; Houdayer et al., 2008; Tang et al., 2016). The major problem that complicates the use of these tools in diagnostic laboratories is that there is no unified standard to measure how splicing signals change when one allele is substituted by another because most tools only output prediction scores for potential splice sites given an input DNA sequence. Due to complex dependencies existing among the bases around splice sites, none of the frequently used programs perfectly predict the impact on pre-mRNA splicing. Learning algorithms such as AdaBoost (Pashaei, Yilmaz, Ozen, & Aydin, 2016) and Random Forest (Meher, Sahu, & Rao, 2016) are now emerging and need to be validated for their use in research and clinical practice.

For the evaluation of the effect of exonic variants on splicing, ESE-specific in silico prediction tools such as ESEfinder (Cartegni, Wang, Zhu, Zhang, & Krainer, 2003), RESCUE-ESE (Fairbrother, Yeh, & Sharp, 2002), and the quantitative evaluation of hexamers as exonic splicing elements (Ke et al., 2011) can be applied.

In this study we evaluated the concordance between "classic prediction tools" and more recently developed algorithms (AdaBoost and Random Forest) for 11 *BRCA1* and 10 *BRCA2* variants. Especially for variants outside the canonical splice sites (±1, ±2), the frequently applied prediction tools are less adequate (Spurdle et al., 2008). We compared the in silico output with data of in vitro RNA splicing assays in short-term cultured peripheral blood lymphocytes of patients heterozygous for the variants.

In addition, we evaluated the effect at the cDNA level of a novel large tandem duplication spanning exons 5 to a large part of exon 11 in *BRCA2*. We found that besides a skip of the exon, two cryptic donor sites were activated within exon 11. We confirmed that these were also activated in a patient with an interrupted donor site due to a G > C substitution at position +1 of intron 11.

2 | MATERIALS AND METHODS

2.1 | Variant selection

Eleven DNA alterations in *BRCA1* and 10 in *BRCA2* were selected for this study (Supp. Table S1). In addition, we investigated the effect of a large duplication in *BRCA2* at the cDNA level. Some variants have entries in the ClinVar or LOVD databases, but no assays to elucidate splicing effects were previously reported and their pathogenicity remained unclear.

We submitted all variants described in this paper to the LOVD databases for BRCA1 and BRCA2 variants (URL: https://databases.lovd.nl/shared/genes/BRCA1 and https://databases.lovd.nl/shared/genes/BRCA2).

2.2 | In silico prediction

Five prediction tools integrated in Alamut® Visual 2.8.1 (Interactive Biosoftware, Rouen, France) were used to assess the impact of the variants on splice sites in silico: MaxEntScan (MES), Human Splicing Finder, Splice Site Finder (SSF), Splice Site Prediction by Neural Network, and GeneSplicer. In Supp. Table S2, the main references and previously determined cut-offs of these programs can be found.

Furthermore, we consulted the dbscSNV v1.1 database for Adaboost and RandomForest scores. A score higher than 0.6 is indicative for aberrant splicing (Jian, Boerwinkle, & Liu, 2014a).

All variants located in the coding regions were also evaluated for their effect on ESE. Disruptions of these 6 bp motifs can result in improper splicing and can be predicted by tools such as ESEFinder (Cartegni et al., 2003) and RESCUE-ESE (Fairbrother et al., 2002), both integrated in Alamut® Visual 2.8.1. The quantitative evaluation of all RNA hexamers as potential exonic splicing elements (Ke et al., 2011) was used to calculate total exonic splicing regulatory sequence (ESRseq) score change (ΔESRseq scores) for every exonic variant. Through analysis of variants in *BRCA2* exon 7, Di Giacomo et al. (2013) experimentally determined a cut-off of −0.663 (Di Giacomo et al., 2013).

Finally, for variants in the coding regions the effect of amino acid substitution itself should also be taken into account. Therefore, Priors (https://priors.hci.utah.edu/PRIORS/), an online tool combining amino acid substitution severity and spliceogenecity-based probability for pathogenicity of *BRCA1* and *BRCA2* point mutations was also consulted (Vallée et al., 2016).

2.3 | Sample collection, cell cultures, RNA isolation, and cDNA preparation

Individuals with variants of interest donated a blood sample in EDTA tubes and signed an informed consent. From these blood samples, phytohemagglutinin stimulated short-term lymphocyte cultures were established (Messiaen & Wimmer, 2008). At day 7, puromycin (200 μg/ml; Sigma–Aldrich, St. Louis, Missouri, USA) was added to the cultures to avoid nonsense-mediated decay (NMD). Four to six hours later, total RNA was extracted using the QIAamp® RNeasy Mini Kit (QIAGEN, Venlo, The Netherlands) according to the manufacturer's instructions. cDNA was synthesized using either the iScript™ cDNA Synthesis Kit (Bio-Rad Laboratories, Hercules, California, USA) or the Superscript® II Reverse Transcriptase Kit (Life Technologies, Carlsbad, California, USA).

2.4 | RT-PCR and sequencing

Splicing aberrations were assessed by means of RT-PCR using either primers in separate exons or primers situated at the exon boundaries (Supp. Table S3) to avoid genomic DNA interference. Fragment sizes were checked on the Labchip GX (Caliper Life Sciences, Waltham, Massachusetts, USA) or on agarose gel. Sanger sequencing was performed using the BigDye® Terminator Cycle Sequencing Kit (Life Technologies, Carlsbad, California, USA). Transcripts from carriers were compared to at least three controls for RT-PCR and at least one control was included for sequencing analysis. For analysis of the *BRCA1* variant c.4675+3A > T, cloning was performed using the TA Cloning™ kit (ThermoFisher Scientific, Waltham, Massachusetts, USA) to characterize the various aberrant transcripts revealed by RT-PCR.

2.5 | Nomenclature

Nucleotide and exon numbering for cDNA is based on NCBI entries NM_007294.3 (*BRCA1*) and NM_000059.3 (*BRCA2*). Nucleotide +1 corresponds to A of AUG translation initiation codon (according to HGVS guidelines). NM_007294.3 omits the historical exon 4 and renumbers the remaining exons sequentially (Ensembl transcript: ENST00000357654.7). In Supp. Table S1, the "legacy numbering" is also provided.

3 | RESULTS

An overview of the *BRCA1* and *BRCA2* variants evaluated can be found in Supp. Table S1, including in silico prediction data as well as the results of cDNA analysis in short-term cultured lymphocytes in the presence of puromycin. Out of 21 tested DNA alterations, 12 DNA alterations (seven *BRCA1* and five *BRCA2* variants) resulted in one or more aberrant mRNA transcripts. Furthermore, the *BRCA2* duplication c.425+415_4780dup{ins GATCGCAGTGA}, spanning exons 5, 6, 7, 8, 9, 10, and a large part of exon 11, was shown to result in a complex splicing pattern.

3.1 | cDNA analysis results for four variants in canonical splice sites

BRCA1 c.5468-1G > A is located in the splice acceptor site of the last coding exon 23. Loss of the WT splice site, unanimously predicted by all prediction tools (Supp. Table S1), results in an out of frame deletion of the first 11 bp of exon 23 as a cryptic splice site acceptor was activated 11 bp downstream of the WT acceptor site. This resulted in an aberrant transcript (r.5468_5478del), leading to a premature termination codon (PTC) (p.(Ala1823Valfs*2)) (Figure 1A). We cannot rule out that this variant also leads to loss of exon 23, the last exon of *BRCA1*, as we have no means to design primers in the next exon.

The substitution *BRCA1* c.4986+1G > C results in the activation of a cryptic splice site at c.4986+65. This leads to an aberrant transcript in which the first 65 intronic nucleotides of intron 15 are retained: r.4986_4987ins4986+1_486+65, p.(1662Phefs*14) (Figure 1B).

BRCA1 c.5152+2dupT inactivates the splice donor site of exon 17, resulting in an aberrant transcript lacking exon 17 (r.5075_5152del; p.(Asp1692_Trp1718delinsGly)) (Figure 1C).

BRCA2 c.6841+1G > C results in the abolishment of the WT splice donor site of exon 11, accurately predicted by all tools. This leads to multiple aberrant transcripts of which a skip of exon 11 (r.1910_6841del, p.(Leu638_Gly2281del)) is the most abundant (Figure 2A–C). However, additional isoforms are obvious from Figure 2A. These are further described under point 3.4.

3.2 | cDNA analysis results for 10 variants outside the canonical splice sites but in the Cartegni consensus region

The Cartegni consensus region encompasses 11 bases of the 5′ splice site (from the three last exonic to the eight first intronic bases) and 14 bases of the 3′ site (from the 12 last intronic to the first two exonic bases) (Cartegni, Chew, & Krainer, 2002).

Two exonic variants in this region both result in aberrant splicing:

- The substitution *BRCA1* c.4674A > G, affecting the second last nucleotide of exon 14, leads to the abolishment of the natural splice donor site as predicted by all in silico tools. The lack of a natural splice donor site resulted in the activation of a cryptic splice site at position c.4664, only predicted by SSF (score = 91.56), leading to a deletion of the last 11 nucleotides of exon 14: r.4665_4675del; p.(Gln1556Glyfs*14). The presence of a small "G" signal at position c.4674, demonstrates some remaining activity from the natural splice site (Supp. Table S1 and Figure 1D).
- *BRCA2* c.517G > C, located in the first nucleotide of exon 7 leads to an aberrant transcript containing an out of frame skip of exon 7 (r.517_631del; p.(Gly173Serfs*19)) due to the inactivation of the WT splice donor (Supp. Table S1 and Figure 4A). MES provided a small reduction in score (−12.1%, lower than the 15% cut-off previously proposed by Houdayer et al. (2012) and Jian et al. (2014a), but higher than the 10% cut-off more recently proposed by Tang et al. (2016)). Priors calculated a greater change of pathogenicity through splicing defects (probability = 0.34) than as to be expected from the

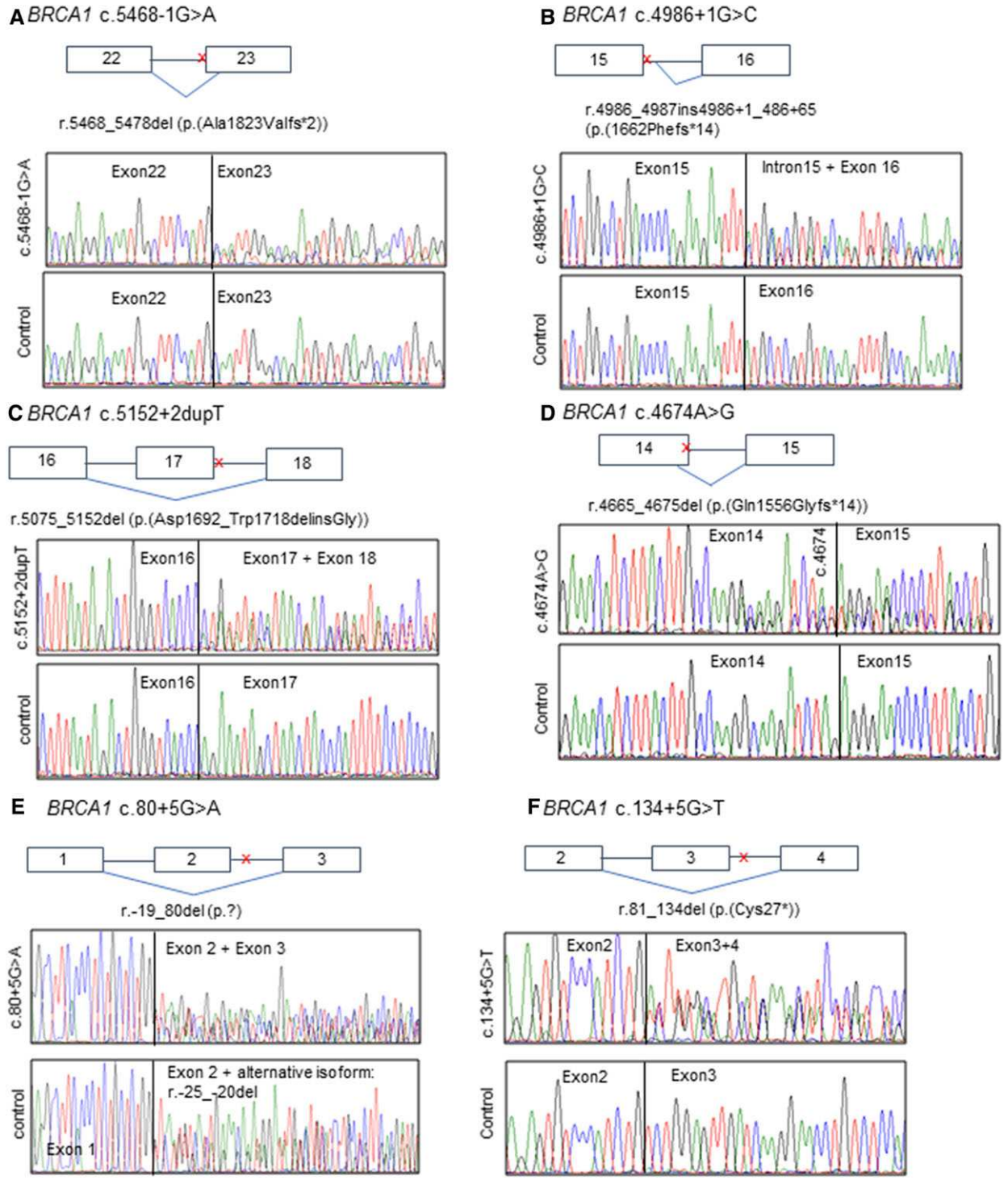

FIGURE 1 Results of in vitro mRNA analysis of six *BRCA1* variants leading to aberrant splicing. Schematic representation of six *BRCA1* variants leading to aberrant splicing with x illustrating the position of the variant. Sequencing results are shown for the patient with the mutation and a negative control. From panel E, it is clear that r.-25_-20del represents a major alternative splice event, observed in both the patient and controls (Colombo et al., 2014; Menéndez et al., 2012). Nucleotide and exon numbering for cDNA is based on NCBI entry NM_007294.3

missense variant itself (probability = 0.02). Only AdaBoost and Random Forest provided high scores indicative for aberrant splicing.

Five out of eight intronic variants in the Cartegni consensus region resulted in aberrant splicing and in silico predictions for these variants were accurate: *BRCA1*c.80+5G > A, c.134+5G > T & c.4675+3A > T, and *BRCA2* c.8488-9T > G & c.8954-5A > G.

- *BRCA1* c.80+5G > A abolishes the WT donor site of exon 2. With primers located in the non-coding exon 1 and in exon 7, we found that this variant leads to out of frame skipping of exon 2 (r.-19_80del; p.?) leading to loss of the translation initiation codon (Supp. Table S1 and Figure 1E). In addition, we found both in the controls and the patient with the mutation an abundant alternative transcript (r.-25_-20del: skip of the last six nucleotides of

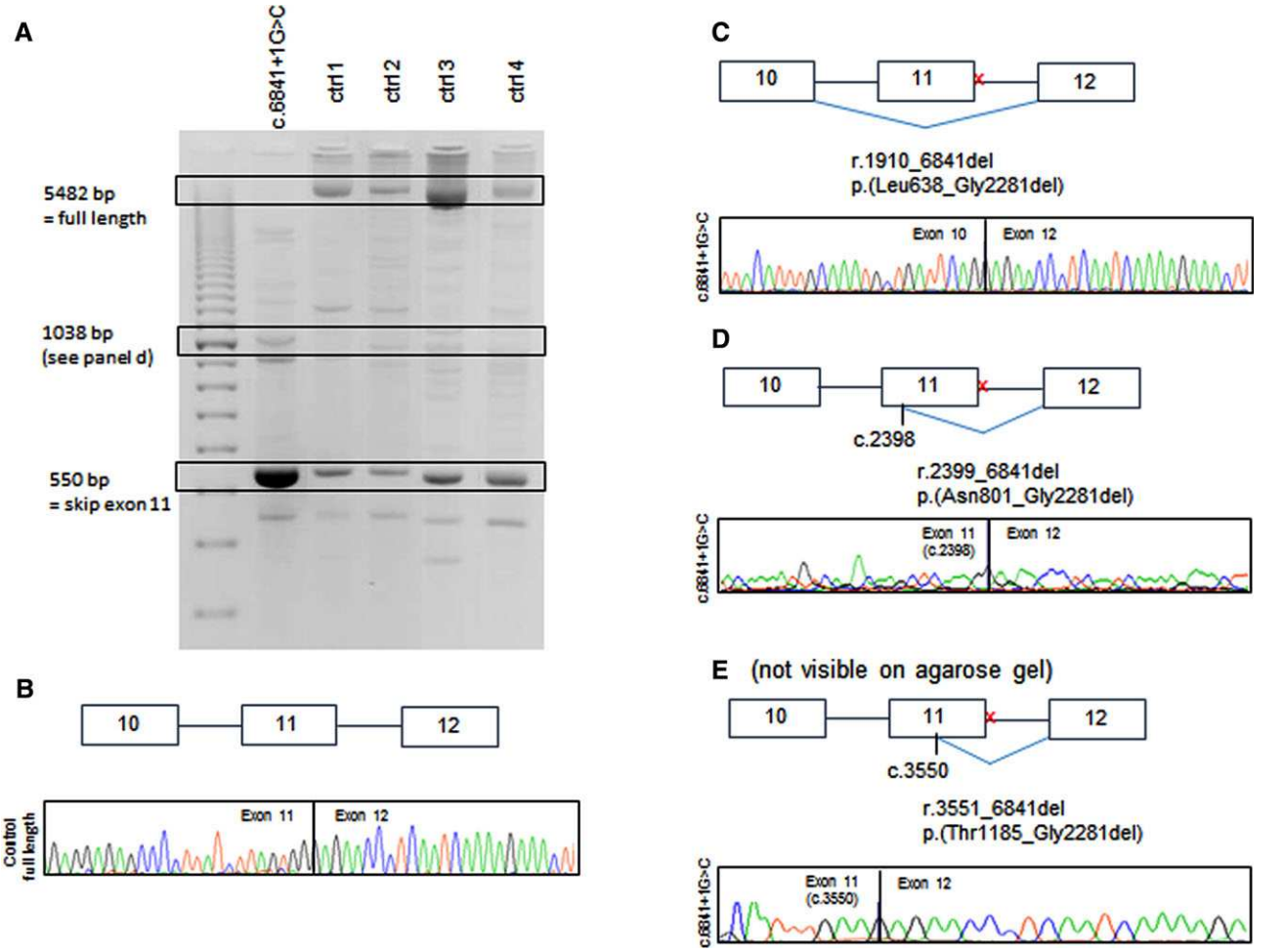

FIGURE 2 Schematic representation of the results of in vitro mRNA analysis for the *BRCA2* variant c.6841+1G > C. **A–C**: RT-PCR products on agarose gel, schematic representation, and sequencing data from PCR products generated with primers spanning exons 10 to 14 for the patient with the c.6841+1G > C variant and four controls. In both the patient with the c.6841+1G > C mutation and the controls, several shorter transcripts were observed (**A**). The full-length transcript (**B**) is present in all control samples, while the skip of exon 11 (c) is highly abundant in the patient but shows lower expression in controls. ddPCR showed that this transcript represented <1% of the transcripts in controls (the "relatively strong" band on the picture may be due to preferential amplification of the shorter transcript). **D and E**: Two additional isoforms were characterized. **D**: A transcript resulting from the activation of a cryptic splice site donor at c.2398, missing nucleotides c.2399_6841 from exon 11 and (**E**) a slightly longer transcript generated by the activation of a cryptic splice site donor at c.3550, lacking nucleotides c.3551_6841 from exon 11. Nucleotide and exon numbering for cDNA is based on NCBI entry NM_000059.3

exon 1), which has previously been described as naturally occurring isoform.

- *BRCA1* c.134+5G > T inactivates the donor of exon 3 leading to an out of frame skip of exon 3, resulting in a PTC: r.81_134del; p.(Cys27*) (Figure 1F). Skipping of exon 3 has previously been described as a naturally occurring isoform (Colombo et al., 2014), but from Figure 1F it becomes clear that this isoform is hardly detectable in controls compared to the patient with the c.134+5G > T substitution.
- RT-PCR with primers located in exons 11–16 resulted in a smear of PCR products on agarose gel for the patient with the *BRCA1* variant c.4675+3A > T, wherein the two tested controls only a single band of the expected length was observed (data not shown). To analyze the products of different lengths, cloning followed by sequencing was performed for 57 clones. Two heterozygous SNPs (c.4308T/C and c.4837A/G) allowed to determine if the transcript originated from the WT or the mutant allele. An overview of the aberrant mRNA transcripts is shown in Figure 3 and Table 1; Sanger sequencing results can be found in Supp. Figure S1. The major effect of this mutation is the activation of a cryptic donor site, 11 nucleotides upstream of the WT donor site of exon 14, similar as for the patient with c.4674A > G. In addition, increased expression of the naturally occurring isoform lacking exon 14 (out of frame) was observed. Interestingly, we have no evidence that this represents a major splice event in the patient with c.4674A > G, in the second last nucleotide of exon 14. However, for analysis of this patient a different primer set to generate a shorter fragment was used (Supp. Table S3) and cloning was not performed in that case.
- *BRCA2* c.8954-5A > G (Figure 4B) was found to generate a cryptic splice site, resulting in an out of frame insertion of

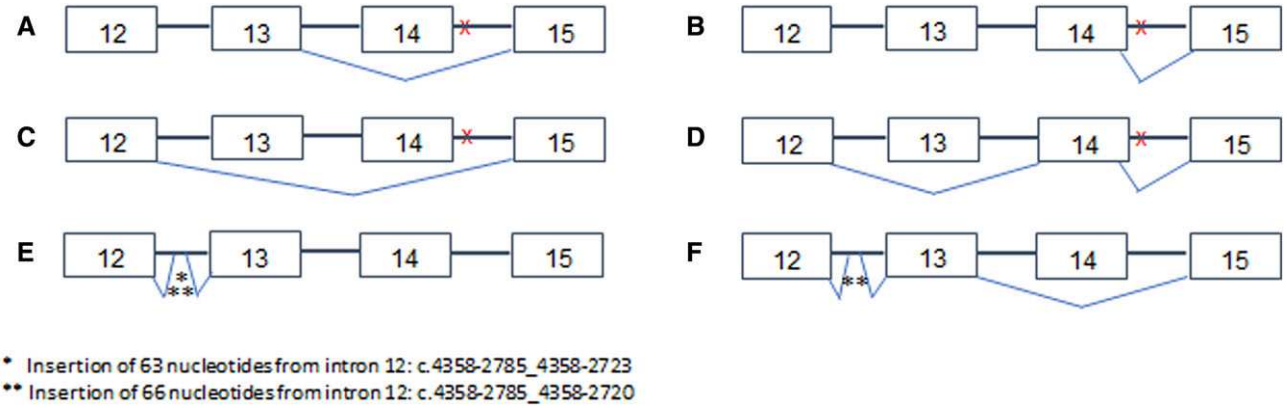

FIGURE 3 Schematic representation of all isoforms obtained by primers located in exons 11 and 16 for *BRCA1* c.4675+3A > T. Cloning of the cDNA fragment revealed several different isoforms transcribed both from the mutant as well as from the WT allele. A schematic representation of the different isoforms is shown (**A–E**). Corresponding Sanger sequencing results can be viewed in Supp. Figure S1. The x indicates the position of the mutation. Panel (e) shows two minor alternative isoforms: an *in frame* insertion of 66 nucleotides (r.4357_4358ins4358-2785_4358-2720, previously reported by our group Claes et al. (2003) and by Colombo et al. (2014)) in addition to an insertion of 63 nucleotides from intron 13 (r.4357_4358ins4358-2785_4358-2723, NCBI RefSeq NM_007300.3) representing less than 2% of the transcripts based on our cloning data

TABLE 1 Overview of aberrant transcripts detected for BRCA1 c.4675+3A > T

Transcript	RNA	Protein	% (Number of clones/total sequenced)	% Originating from the WT allele	Figure 3
Full length	/	/	43.90% (25/57)	96% (24/25)	/
Out of frame skip of exon 14	r.4485_4675del	p.Ser1496Glyfs*14	24.60% (14/57)	7% (1/14)	A
Out of frame deletion of last 11 nucleotides of exon 14	r.4665_4675del	p.Gln1556Glyfs*14	17.50% (10/57)	0% (0/10)	B
In frame skip of exon 13 and 14	r.4358_4675del	p.Ala1453_Leu1558del	3.50% (2/57)	0% (0/2)	C
Out of frame skip of exon 13 and the last 11 nucleotides of exon 14	r.4358_4484del + r.4665_4675del	p.Ala1453Gyfs*10	3.50% (2/57)	0% (0/2)	D
Full length and insertion of 66 nucleotides from intron 12	r.4357_4358ins4358-2785_4358-2720	p.Lys1452_Ala1453ins(22)	3.50% (2/57)	100% (2/2)	E
Full length + insertion of 63 nucleotides from intron 12	r.4357_4358ins4358-2785_4358-2723	p.Lys1452_Ala1453ins(21)	1.70% (1/57)	100% (1/1)	E
Insertion of 66 nucleotides from intron 12 + skip of exon 14	r.4357_4358ins4358-2785_4358-2720 + r.4485_4675del	p.Lys1452_Ala1453ins(22) + p.Ser1496Glyfs*14	1.70% (1/57)	100% (1/1)	F

four nucleotides from intron 22 (r.8953_8954ins8954-4_8954-1; p.(Val2985Aspfs*34)), inducing a premature stop codon, and therefore, prone to NMD.

- The outcome of two variants in the Cartegni site of *BRCA2* intron 19 was different, which was accurately predicted by all in silico prediction tools (Supp. Table S1). *BRCA2* c.8488-9T > G creates a cryptic acceptor site, leading to an aberrant transcript with an out of frame insertion of the last eight nucleotides of intron 19 (r.8487_8488ins8488-8_8488-1; p.(Trp2830Tyrfs*36)) (Supp. Table S1 and Figure 4C), whereas *BRCA2* c.8488-12A > G was not found to induce aberrant splicing.
- Also three other intronic variants in the Cartegni consensus region (*BRCA1* c.4186-10G > A and *BRCA2* c.7618-6G > T; c.8488-12A > G) did not result in aberrant splicing. Based on the MES predictions, aberrant splicing for *BRCA1* c.4186-10G > A (−17.7%) and *BRCA2* c.8488-12A > G (−13.7%) could be expected based on the −10% cut-off proposed by Tang et al. (2016), whereas Houdayer et al. (2012) applied a more stringent cut-off of −15%. For *BRCA2* c.7618-6G > T, the classic splice prediction tools showed increased scores for the natural acceptor site, compatible with normal splicing.

3.3 | cDNA analysis results for seven variants outside the Cartegni consensus region

We evaluated the effect at the mRNA level of two intronic variants *BRCA1* c.5278-22C > G and *BRCA2* c.8488-14A > G.

- *BRCA1* c.5278-22C > G did not result in aberrant splicing, which was correctly predicted by all consulted in silico prediction tools (Supp. Table S1), pointing toward a neutral effect of this variant. This is supported by the fact that more recently a heterozygous *PALB2* mutation c.3362delG was identified in this patient.
- *BRCA2* c.8488-14A > G leads to aberrant splicing. This results from the creation of a cryptic spice site, predicted by all tools consulted, and a reduction in strength of the WT splice site, only predicted by MES. The aberrant transcript contains an out of frame insertion of the last 13 nucleotides of intron 19: r.8432_8433ins8433-13_8433-1; p.(Trp2830Hisfs*19) (Supp. Table S1 and Figure 4D).

None of the five tested exonic variants outside the Cartegni consensus region [two in *BRCA1:* c.1878A > G (p. =) & c.4115G > A; p.(Cys1372Tyr) and three in *BRCA2:* c.3326C > T; p.(Ala1109Val), c.4899C > G; p.(Ile1633Met) & c.6313A > G; p.(Ile2105Val)] affected mRNA splicing (Supp. Table S1). However, the prediction tool Priors, classified three variants [*BRCA1* c.1878A > G (p. =) & c.4115G > A; p.(Cys1372Tyr); *BRCA2* c.3326C > T; p.(Ala1109Val)] as "moderate probability of pathogenicity from the creation of a de novo splice donor" (score 0.3). ΔESRseq scoring predicted a splicing effect for *BRCA2* c.3326C > T and *BRCA1* c.4115G > A (ΔESRseq scores below −0.663). ESEfinder and RESCUE-ESE predicted both disruption and/or creation of a new ESE for every variant except for *BRCA2* c.4899C > G. Furthermore, *BRCA1* c.4115G > A was predicted to activate a cryptic donor site at position c.4112. However, this cryptic donor site was not strong enough to compete with the WT splice site. None of the variants were predicted to have an effect due to an altered protein structure (Prior missense score: 0.02).

3.4 | Identification of multiple cryptic donor sites within the large exon 11 of *BRCA2* in case of interruption of the WT exon 11 donor site

cDNA analysis in lymphocytes of a patient heterozygous for the *BRCA2* tandem duplication c.426+415_4780dup{insGATCGCAGTGA} revealed a complex splicing pattern (Figure 5). The major effects are the inclusion of a cryptic exon in intron 4 (r.425+415_426-209), skipping of exon 11 and aberrant transcripts due to the activation of two cryptic donor sites within exon 11: c.2398 and c.3550. All combinations observed are out of frame.

We evaluated if the cryptic donor sites at c.2398 and c.3550 were also activated in a patient with the substitution *BRCA2* c.6841+1G > C, inactivating the WT donor site of exon 11. Besides skipping of exon 11, several other aberrant transcripts were observed. Also in this patient, the cryptic splice donor sites at c.2398 and c.3550 are activated, resulting in the skip of respectively 4,443 and 3,291 nucleotides from the 3′part of exon 11: r.2399_6841del; p.(Asn801_Gly2281del) (Figure 2D) and r.3551_6841del; p.(Thr1185_Gly2281del) (Figure 2E), respectively. The activation of c.2398 leads to a band of 1,038 bp, visible on the gel (Figure 2A), which was confirmed by direct sequencing. No band in concordance with activation of c.3550 was observed in the initial RT-PCR. However, by means of allele specific primers (Supp. Table S1), we could confirm the activation of this particular cryptic splice site.

4 | DISCUSSION

We evaluated the effect at the cDNA level of 21 VUS in *BRCA1* ($n = 11$) and *BRCA2* ($n = 10$) and demonstrated aberrant splicing for 12 variants (seven in *BRCA1* and five in *BRCA2*). In addition, we evaluated the effect of a large tandem duplication in *BRCA2*, with breakpoints in intron 4 and exon 11.

By analyzing controls in all assays, we verified that none of the detected aberrant *BRCA2* transcripts were reported as naturally occurring alternative transcripts (Fackenthal et al., 2016). A skip of *BRCA2* exon 11 was also observed in controls, though be it at very low level (Figure 2). Some of the variants studied in *BRCA1*, lead to a strong upregulation of low abundant naturally occurring transcripts in our patients: skip of exon 2 (r.-19_80), exon 3 (r.81_134del), exon 13 and 14 (r.4358_4675del), exon 14 (r.4485_4675del) and exon 17 (r.5075_5152del) but none of these are considered as naturally occurring *in frame* RNA isoforms that may rescue gene functionality (cfr. ENIGMA BRCA1/2 Gene Variant Classification Criteria v2.4, Table 6—downloaded from: https://enigmaconsortium.org/wp-content/uploads/2016/01/ENIGMA_Rules_2017-05-09.pdf).

mRNA splicing assays yield insights in the exact aberrant transcript, which may in turn shed light on the clinical importance of the variant. The majority of the tested splicing variants lead to a frameshift and the creation of a PTC and are therefore expected to be degraded by NMD (Baert et al., 2016; Perrin-vidoz, Sinilnikova, Stoppa-lyonnet, Lenoir, & Mazoyer, 2002; Ware et al., 2006). Aberrant *BRCA1* transcripts can escape NMD when the PTC is located near the C-terminus or near the translation initiation codon (Perrin-vidoz et al., 2002). Based on this rule, *BRCA1* c.134+5G > T and *BRCA1* c.5468-1G > A may escape NMD. Translation of the transcript generated from the *BRCA1* c.134+5G > T allele would result in only a very small, most likely non-functional, protein (p.(Cys27*)). *BRCA1* c.5468-1G > A leads to an aberrant transcript with a stop codon at amino acid position p.1825 in the last exon. Unfortunately, cultured lymphocytes without puromycin were not available to check the expression of the aberrant transcript in the absence of NMD inhibition. But even if this transcript avoids NMD, two nonsense mutations resulting in a stop codon downstream of amino acid 1825 (c.5503C > T; p.(Arg1835*) and c.5559C > A; p.(Tyr1853*)) are also considered as pathogenic mutations (see ClinVar database entries 55601 and 55629) and evasion of NMD was demonstrated for both variants (Perrin-vidoz et al., 2002).

A pathogenic effect for *BRCA1* c.80+5G > A is expected as the AUG translation initiation codon is lost due to the skip of exon 2. Interruption of the translation initiation codon by a substitution has

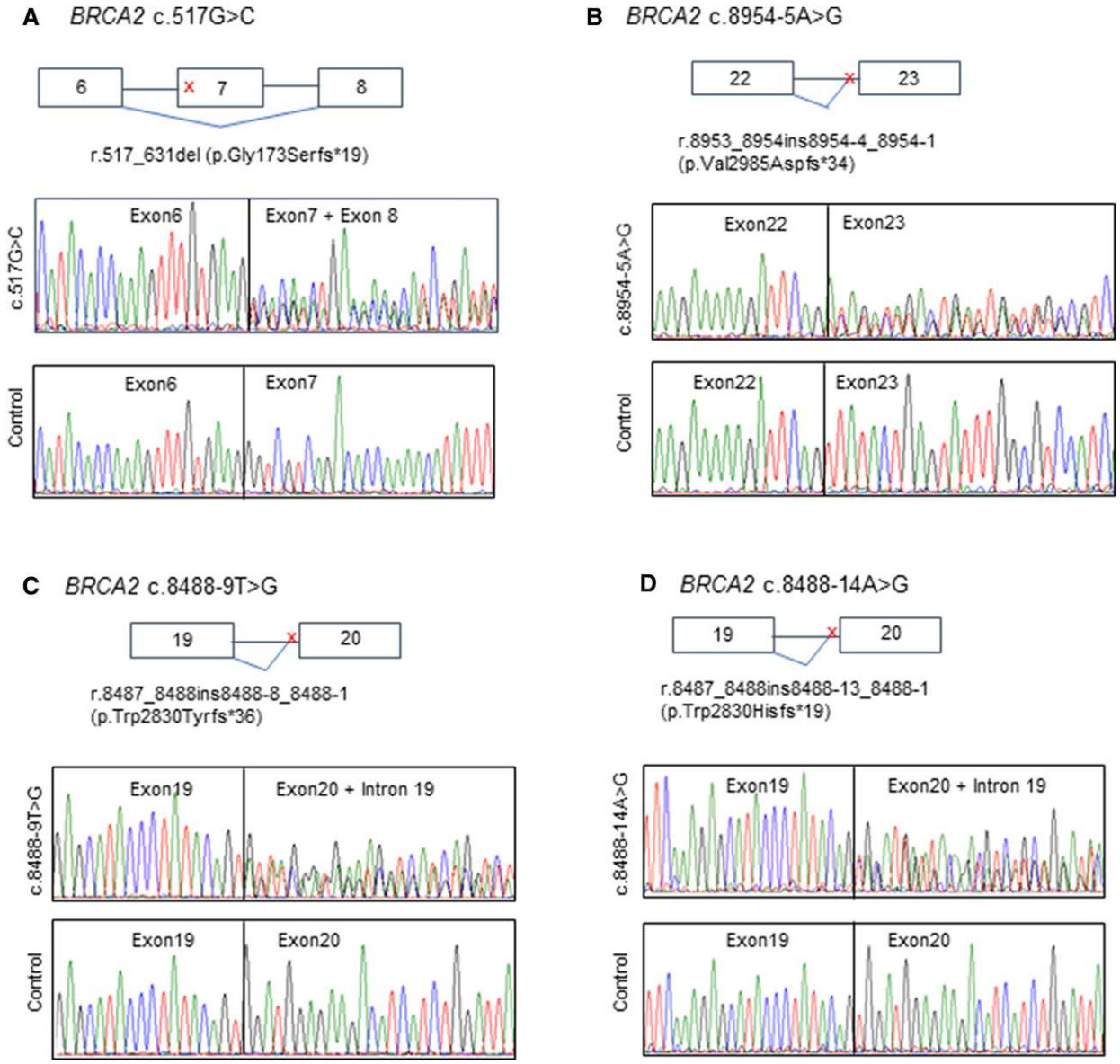

FIGURE 4 Results of in vitro mRNA analysis of four *BRCA2* variants. Schematic representation and sequencing results of four *BRCA2* variants leading to aberrant splicing with the x illustrating the position of the variant and controls. Nucleotide and exon numbering for cDNA is based on NCBI entry NM_000059.3

previously been classified as deleterious (ClinVar ID: 54432, 54745, 54746, 55072, and 267523).

BRCA1 c.5152+2dupT, leads to an in frame skip of exon 17 (r.5075_5152del; p.(Arg1692_Trp1717delinsGly)). Several naturally occurring transcripts lack exon 17 (Colombo et al., 2014), but in our sequencing data the skip of exon 17 was only very obvious in the patient heterozygous for c.5152+2dupT (see Figure 1C). A transcript lacking an in frame exon will avoid NMD. But a skip of exon 17 (amino acids: 1692_1718) leads to disruption of a large part of the first BRCT domain (amino acids 1671_1745), required for several functions of the BRCA1 protein (Roy et al., 2012). Several missense variants in these BRCT repeats lead to impaired folding (Glover, 2006; Lovelock et al., 2006, 2007) and shown to be deleterious based on multifactorial likelihood analysis (Lindor, Goldgar, Tavtigian, Plon, & Couch, 2013). Similarly, *BRCA1* c.5468-1G > A; p.(Ala1823Valfs*2) results in loss of the last 31 amino acids of the second BRCT repeat (amino acids 1,779–1,854) and is therefore expected to cause improper folding of this motif. The p.Val1838Glu variant located in this domain is a proven class 5 variant (Lindor et al., 2013). Binding of several interaction partners might thus be impaired by these two variants, and a deleterious effect may be further confirmed by functional assays.

Aberrant transcripts induced by *BRCA2* c.6841+1G > C are all in frame (p.(Thr1185_Gly2281del), p.(Asn801_Gly2281del), and p.(Leu638_Gly2281del)). However, the size of the deletions and the location in the BRC repeat region, necessary for binding of RAD51 (Gudmundsdottir & Ashworth, 2006), suggest a pathogenic effect.

mRNA analysis of three variants situated in intron 19 of *BRCA2* revealed different effects. Variants c.8488-9T > G and c.8488-14A > G resulted in aberrant splicing, whereas a variant situated in between those two (c.8488-12A > G) did not impair proper mRNA splicing.

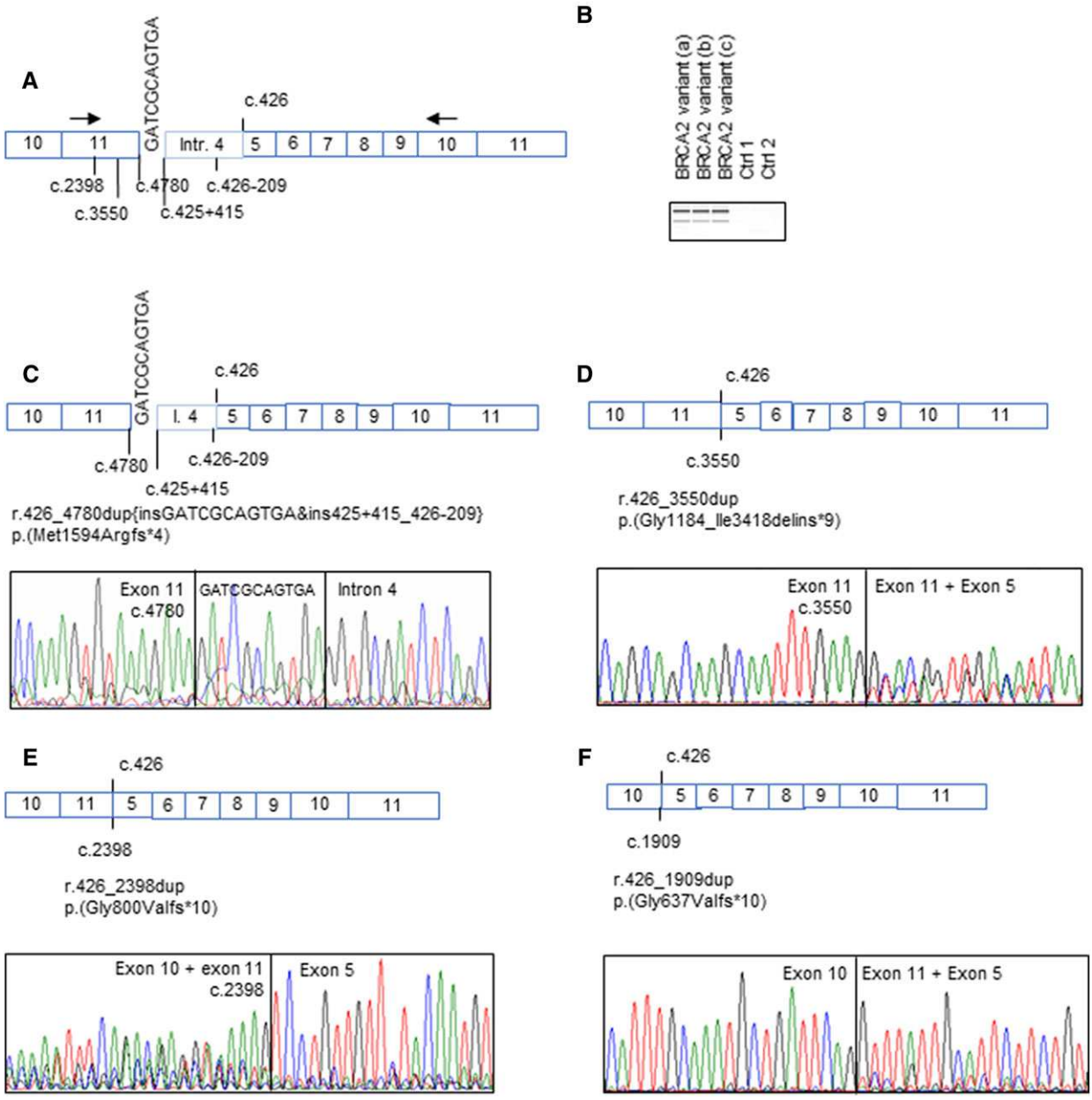

FIGURE 5 RT-PCR results for the patient heterozygous for *BRCA2* c.426+415_4780dup{ins GATCGCAGTGA}. **A**: Schematic representation of the complex tandem duplication. **B–D**: RT-PCR combining a forward primer in exon 11 (c.3183_ 3202) and a reverse primer in exon 10 (c.1261_1242) (as indicated in panel a), did not result in a PCR product in two controls but resulted in two fragments in the patient with the duplication (panel **B**: three replicates shown for the patient with the duplication in BRCA2). Sequencing results for these two fragments are shown in panels **C and D**. **C**: The largest PCR product contained all nucleotides from the start of the primer till c.4780 (the presumed breakpoint in exon 11), an insertion of 11 nucleotides (GATCGCAGTGA), followed by 293 nucleotides from intron 4 (from c.425+415 to c.426-209) and continued by all nucleotides from exon 5 to position c.4780 in exon 11; **D**: a smaller fragment was produced by the activation of a cryptic splice site donor at nucleotide position c.3550. This fragment does not contain nucleotides c.3551 to c.4780, the 11 bp sequence GATCGCAGTGA nor the 293 nucleotides between c.425+415 to c.426-209 from intron 4. **E and F**: RT-PCR with a forward primer in exon 10 (c.1544_1565) and a reverse primer in exon 7 (c.1078_1099) allowed identification of two additional transcripts: (**E**) sequencing revealed the activation of a cryptic splice site donor at nucleotide position c.2398, followed by the insertion of 11 nucleotides (GATCGCAGTGA) and 293 nucleotides of intron 4 (c.425+415_426-209); (**F**) in addition, a shorter aberrant fragment resulting from the complete skip of exon 11 was obtained

Both c.8488-9T > G and c.8488-14A > G create a novel acceptor site, whereas the natural splice acceptor site is relatively weak; a naturally occurring skip of exon 20 has been demonstrated (Fackenthal et al., 2016). The variant c.8488-12A > G does not result in a de novo "AG" splice acceptor site, and did induce aberrant splicing. Only variants in the canonical splice acceptor site of *BRCA2* exon 20 (c.8488-1/2) have previously been reported to result in exon skipping (ClinVar and LOVD) (Fackenthal et al., 2016). We assume that the aberrant transcripts induced by c.8488-9T > G and c.8488-14A > G will be degraded by NMD as NMD was demonstrated for all tested PTC-introducing mutations in *BRCA2*, even when the stop codon is situated towards the C-terminal end (Ware et al., 2006). Even in the absence of NMD how-

ever, the predicted proteins with a frameshift starting from amino acid 2830, will lack most of the BRCA2 DNA binding domain (amino acids 2481_3186), critical for BRCA2 functionality.

Out of the seven tested exonic variants studied, only two (*BRCA1* c.4674A > G & *BRCA2* c.517G > C) resulted in aberrant splicing. *BRCA1* c.4674A > G, located in the second last nucleotide of exon 14, resulted in aberrant splicing due to an abolished WT splice donor site and an activated a cryptic (GC) donor site upstream at position c.4664, only predicted by SSF (Supp. Table S1). *BRCA2* c.517G > C is located in the first nucleotide of exon 7 and induces aberrant splicing due to loss of the natural splice site. However, splice site prediction tools demonstrate only a marginal decrease in 3′ splice site strength (Supp. Table S1). The skip of exon 7 (r.517_631del; p.(Gly173Serfs*19)) is likely to undergo degradation of the transcript by NMD. A variant at the same position (c.517G > T) was previously tested using both minigene and RNA splicing analysis and also resulted in skipping of exon 7 (Gaildrat et al., 2012). A higher than expected presence of splicing mutations is detected in exon 7 of *BRCA2*. This exon is characterized by the presence of ESE's and the combination with relatively weak natural splice sites explains the high number of splicing variants in this exon (Di Giacomo et al., 2013).

In contrast, no such variants have been reported in the large exons 10/11 of *BRCA1/2*. None of the four variants in the large exons studied here, showed improper splicing. ΔESRseq scoring was used to evaluate regulatory elements in these exons. Based on the experimentally determined cut-off of −0.6331 for variants in *BRCA2* exon 7 (Di Giacomo et al., 2013), aberrant splicing was a priori not expected for two out of three studied *BRCA2* exon 11 variants (ΔESRseq score for c.4899C > G = 0.1009 and for c.6313A > G = 0.8822; Supp. Table S1). However, for *BRCA2* c.3326C > T and *BRCA1* c.4115G > A the ΔESRseq scores of respectively −2.9402 and −1.3814 would imply an effect on exonic splice regulation which could not be demonstrated by mRNA analysis. Therefore, the cut-off, experimentally determined for variants in *BRCA2* exon 7, containing a large number of ESE's in combination with weak splice donor and acceptor sites, may not be applicable for variants in other exons. Furthermore, neither the results of ESEfinder nor the results of RESCUE-ESE were concordant with the mRNA results for exonic variants included in this study. This illustrates the limited predictive value of these tools, although some may be useful for the evaluation of variants in exons enriched for exonic splicing regulators (like *BRCA2* exon 7, *MLH1* exon 10, *BRCA*1 exon 6, *CFTR* exon 12, and *NF1* exon 46 (historical numbering: 37) (Soukarieh et al., 2016)).

To date, no known DNA variations inducing aberrant splicing of *BRCA2* exon 11 have been reported. Here, we detected the activation of two cryptic splice donor sites at c.2398 and c.3550 in the absence of the WT splice site by evaluating the impact on pre-mRNA splicing of a large tandem duplication with a breakpoint 2061 bp upstream of the WT donor site. Activation of both cryptic splice donor sites was confirmed in a patient with a substitution abolishing the WT splice donor site of exon 11 (c.6841+1G > C). Scanning exon 11 with in silico prediction tools available through Alamut 2.8.1 revealed several alternative cryptic splice site donors, but none of them were activated in the patient with c.6841+1G > C. The activation of the cryptic splice donor sites at positions c.2398 and c.3550 might be due to ESE motifs, lying in close vicinity to the activated splice donor sites (Caceres, Hurst, Cáceres, & Hurst, 2013; Ke et al., 2011). The influence of possible ESE's on the cryptic splice site c.3550 and c.2398 should be further experimentally validated (Di Giacomo et al., 2013; Raponi et al., 2014). Knowledge on potentially activated splice sites in case of inactivation of the WT site, may be important to design adequate *BRCA2* ASOs (antisense oligonucleotides), which have been proposed as a promising avenue to prevent resistance to PARP inhibitor therapy in several tumor types (Rytelewski et al., 2016) or as potential therapeutic anti-cancer agent in combination with cisplatin (Rytelewski et al., 2014).

Another aim of this study was to test the performance of in silico prediction tools. For variants outside the canonical splice sites but within the Cartegni consensus region, Adaboost and Random Forest outperformed the tools integrated in Alamut (cfr. *BRCA2* c.7618-6G > T & c.8488-12A > G). The data from Adaboost and Random Forest are extracted from dbscSNV (Liu, Wu, Li, & Boerwinkle, 2016), which currently only include data on substitutions within the Cartegni consensus regions. For all variants studied and located in these regions, the predictions were accurate. However, these programs provide only one score. If this score is higher than 0.6, aberrant splicing is expected but they do not provide an indication on the potential activation/creation of cryptic splice sites; effects need to be studied by cDNA analysis.

The classic in silico prediction tools remain very useful for the evaluation of intronic variants outside the Cartegni consensus region: aberrant scores were obtained for all variants leading to mis-splicing, in agreement with previous studies suggesting that cDNA analysis is not required for intronic variants if no effect on splicing is predicted by multiple in silico prediction tools (Théry et al., 2011; Vreeswijk et al., 2009).

Priors assesses whether pathogenicity of a missense variant is more likely to originate from the amino acid substitution or a splicing defect (Vallée et al., 2016). Of all exonic variants included in this study Priors gave the highest score for impaired splicing (0.34) for *BRCA1* c.4674A > G & *BRCA2* c.517G > C, and both variants lead to aberrant splicing. However, this score is only marginally higher compared with those for other variants [*BRCA1* c.1878A > G (p. =) & c.4115G > A; p.(Cys1372Tyr) and *BRCA2* c.3326C > T; p.(Ala1109Val)] not leading to aberrant splicing (all with a Priors score of 0.3). A validation study on a large number of exonic variants would allow to determine a threshold for the Priors tool (currently not integrated in Alamut) above which a strong correlation between in silico prediction and aberrant mRNA splicing exists.

To conclude, we confirmed aberrant splicing leading to disruption of functionally important domains for 12 out of 21 tested DNA alterations, suggesting that these variants may be associated with a pathogenic effect. One limitation of the study is that we were not able to assess allele-specific transcript expression for all variants, which is a prerequisite of ENIGMA to definitively classify a variant as (likely) pathogenic (https://enigmaconsortium.org/wp-content/uploads/2016/01/ENIGMA_Rules_2017-05-09.pdf). Two novel in silico prediction tools, Adaboost and Random Forest, performed very well in our cohort, indicate which substitutions in the Cartegni consensus region are expected to lead to aberrant splicing,

but can currently not be used for variants outside these regions and are only suitable for the evaluation of substitutions. Our study is the first to reveal actively used cryptic splice donor sites within the large exon 11 of *BRCA2*. This finding opens perspectives for new research on the identification of *cis* and *trans* acting factors involved in correct splicing of this large exon.

ACKNOWLEDGMENTS

The authors wish to thank the patients for donating a blood sample.

DISCLOSURE STATEMENT

The authors declare no conflict of interest.

ORCID

William D. Foulkes http://orcid.org/0000-0001-7427-4651
Kathleen B. M. Claes http://orcid.org/0000-0003-0841-7372

REFERENCES

Baert, A., Depuydt, J., Van Maerken, T., Poppe, B., Malfait, F., Storm, K., … Vral, A. (2016). Increased chromosomal radiosensitivity in asymptomatic carriers of a heterozygous BRCA1 mutation. *Breast Cancer Research*, *18*(1), 52. https://doi.org/10.1186/s13058-016-0709-1

Caceres, E. F., Hurst, L. D., Cáceres, E. F., & Hurst, L. D. (2013). The evolution, impact and properties of exonic splice enhancers. *Genome Biology*, *14*(1465–6914 (Electronic)), R143. https://doi.org/10.1186/gb-2013-14-12-r143

Caminsky, N. G., Mucaki, E. J., & Rogan, P. K. (2014). Interpretation of mRNA splicing mutations in genetic disease: Review of the literature and guidelines for information-theoretical analysis. *F1000Research*, *3*(May 2016), 282. https://doi.org/10.12688/f1000research.5654.1

Cartegni, L., Chew, S. L., & Krainer, A. R. (2002). Listening to silence and understanding nonsense: Exonic mutations that affect splicing. *Nature Reviews Genetics*, *3*(4), 285–298. https://doi.org/10.1038/nrg775

Cartegni, L., Wang, J., Zhu, Z., Zhang, M. Q., & Krainer, A. R. (2003). ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Research*, *31*(13), 3568–3571. https://doi.org/10.1093/nar/gkg616

Claes, K., Poppe, B., Machackova, E., Coene, I., Foretova, L., De Paepe, A., & Messiaen, L. (2003). Differentiating pathogenic mutations from polymorphic alterations in the splice sites of BRCA1 and BRCA2. *Genes Chromosomes and Cancer*, *37*(3), 314–320. https://doi.org/10.1002/gcc.10221

Colombo, M., Blok, M. J., Whiley, P., Santamariña, M., Gutiérrez-Enríquez, S., Romero, A., … de la Hoya, M. (2014). Comprehensive annotation of splice junctions supports pervasive alternative splicing at the BRCA1 locus: A report from the ENIGMA consortium. *Human Molecular Genetics*, *23*(14), 3666–3680. https://doi.org/10.1093/hmg/ddu075

Di Giacomo, D., Gaildrat, P., Abuli, A., Abdat, J., Frébourg, T., Tosi, M., & Martins, A. (2013). Functional analysis of a large set of brca2 exon 7 variants highlights the predictive value of hexamer scores in detecting alterations of exonic splicing regulatory elements. *Human Mutation*, *34*(11), 1547–1557. https://doi.org/10.1002/humu.22428

Fackenthal, J. D., Yoshimatsu, T., Zhang, B., de Garibay, G. R., Colombo, M., De Vecchi, G., … de la Hoya, M. (2016). Naturally occurring BRCA2 alternative mRNA splicing events in clinically relevant samples. *Journal of Medical Genetics*, *53*, 1–11. https://doi.org/10.1136/jmedgenet-2015-103570

Fairbrother, W. G., Yeh, R., & Sharp, P. A. (2002). Predictive identification of exonic splicing enhancers in human genes. *Science*, *297*, 1007–1013.

Gaildrat, P., Krieger, S., Di Giacomo, D., Abdat, J., Revillion, F., Caputo, S., … Martins, A. (2012). Multiple sequence variants of BRCA2 exon 7 alter splicing regulation. *Journal of Medical Genetics*, *49*(10), 609–617. https://doi.org/10.1136/jmedgenet-2012-100965

Glover, J. N. M. (2006). Insights into the molecular basis of human hereditary breast cancer from studies of the BRCA1 BRCT domain. *Familial Cancer*, *5*(1), 89–93. https://doi.org/10.1007/s10689-005-2579-z

Goldgar, D. E., Easton, D. F., Deffenbaugh, A. M., Monteiro, A. N., Tavtigian, S. V., & Couch, F. J. (2004). Integrated evaluation of DNA sequence variants of unknown clinical significance: Application to BRCA1 and BRCA2. *American Journal of Human Genetics*, *75*(4), 535–544. https://doi.org/10.1086/424388S0002-9297(07)62706-2 [pii]

Gudmundsdottir, K., & Ashworth, A. (2006). The roles of BRCA1 and BRCA2 and associated proteins in the maintenance of genomic stability. *Oncogene*, *25*(43), 5864–5874. https://doi.org/10.1038/sj.onc.1209874

Hendriks, G., Morolli, B., Calléja, F. M. G. R., Plomp, A., Mesman, R. L. S., Meijers, M., … Vrieling, H. (2014). An efficient pipeline for the generation and functional analysis of human BRCA2 variants of uncertain significance. *Human Mutation*, *35*(11), 1382–1391. https://doi.org/10.1002/humu.22678

Houdayer, C., Caux-Moncoutier, V., Krieger, S., Barrois, M., Bonnet, F., Bourdon, V., … Stoppa-Lyonnet, D. (2012). Guidelines for splicing analysis in molecular diagnosis derived from a set of 327 combined in silico/in vitro studies on BRCA1 and BRCA2 variants. *Human Mutation*, *33*(8), 1228–1238. https://doi.org/10.1002/humu.22101

Houdayer, C., Dehainault, C., Mattler, C., Michaux, D., Caux-Moncoutier, V., Pagès-Berhouet, S., … Stoppa-Lyonnet, D. (2008). Evaluation of in silico splice tools for decision-making in molecular diagnosis. *Human Mutation*, *29*(7), 975–982. https://doi.org/10.1002/humu.20765

Jian, X., Boerwinkle, E., & Liu, X. (2014a). In silico prediction of splice-altering single nucleotide variants in the human genome. *Nucleic Acids Research*, *42*(22), 13534–13544. https://doi.org/10.1093/nar/gku1206

Jian, X., Boerwinkle, E., & Liu, X. (2014b). In silico tools for splicing defect prediction: A survey from the viewpoint of end users. *Genetics in Medicine*, *16*(7), 497–503. https://doi.org/10.1038/gim.2013.176

Ke, S., Shang, S., Kalachikov, S. M., Morozova, I., Yu, L., Russo, J. J., … Chasin, L. A. (2011). Quantitative evaluation of all hexamers as exonic splicing elements. *Genome Research*, *21*(8), 1360–1374. https://doi.org/10.1101/gr.119628.110

Lindor, N. M., Goldgar, D. E., Tavtigian, S. V., Plon, S. E., & Couch, F. J. (2013). BRCA1/2 sequence variants of uncertain significance: A primer for providers to assist in discussions and in medical management. *The Oncologist*, *18*(5), 518–524. https://doi.org/10.1634/theoncologist.2012-0452

Liu, X., Wu, C., Li, C., & Boerwinkle, E. (2016). dbNSFP v3.0: A one-stop database of functional predictions and annotations for human nonsynonymous and splice-site SNVs. *Human Mutation*, *37*(3), 235–241. https://doi.org/10.1002/humu.22932

Lovelock, P. K., Healey, S., Au, W., Sum, E. Y., Tesoriero, A., Wong, E. M., … Brown, M. A. (2006). Genetic, functional, and histopathological evaluation of two C-terminal BRCA1 missense variants. *Journal of Medical Genetics*, *43*(1), 74–83. https://doi.org/10.1136/jmg.2005.033258

Lovelock, P. K., Spurdle, A. B., Mok, M. T., Farrugia, D. J., Lakhani, S. R., Healey, S., … Brown, M. A. (2007). Identification of BRCA1 missense substitutions that confer partial functional activity:

Potential moderate risk variants? *Breast Cancer Research*, *9*(6), R82. https://doi.org/10.1186/bcr1826

Meher, P. K., Sahu, T. K., & Rao, A. R. (2016). Prediction of donor splice sites using random forest with a new sequence encoding approach. *BioData Mining*, *9*, 4. https://doi.org/10.1186/s13040-016-0086-4

Menéndez, M., Castellsagué, J., Mirete, M., Pros, E., Feliubadaló, L., Osorio, A. … Lázaro, C. (2012). Assessing the RNA effect of 26 DNA variants in the BRCA1 and BRCA2 genes. *Breast Cancer Research and Treatment*, *132*(3), 979–992.

Messiaen, L. M., & Wimmer, K. (2008). Neurofibromatoses. In D. Kaufmann (Ed.), *Monographs in Human Genetics* (Vol. *16*, pp. 63–77). Basel: Karger.

Millot, G., Carvalho, M. A., Caputo, S. M., & Vreeswijk, M. P. G. (2012). A guide for functional analysis of BRCA1 variants of uncertain significance (VUS). *Human Mutation*, *33*(11), 1526–1537. https://doi.org/10.1002/humu.22150.A

Pashaei, E., Yilmaz, A., Ozen, M., & Aydin, N. (2016). A novel method for splice sites prediction using sequence component and hidden Markov model. *Conference Proceedings: IEEE Engineering in Medicine and Biology Society*, 2016, 3076–3079.

Perrin-vidoz, L., Sinilnikova, O. M., Stoppa-lyonnet, D., Lenoir, G. M., & Mazoyer, S. (2002). The nonsense-mediated mRNA decay pathway triggers degradation of most BRCA1 mRNAs bearing premature termination codons. *Human Molecular Genetics*, *11*(23), 2805–2814.

Raponi, M., Smith, L. D., Silipo, M., Stuani, C., Buratti, E., & Baralle, D. (2014). BRCA1 exon 11 a model of long exon splicing regulation © 2014 Landes Bioscience. *RNA Biology*, *11*(4), 351–359.

Roy, R., Chun, J., & Powell, S. N. (2012). BRCA1 and BRCA2: Different roles in a common pathway of genome protection. *Nature Reviews. Cancer*, *12*(1), 68–78. https://doi.org/10.1038/nrc3181

Rytelewski, M., Tong, J. G., Buensuceso, A., Leong, H. S., Maleki Vareki, S., Figueredo, R., … Koropatnick, J., (2014). BRCA2 inhibition enhances cisplatin-mediated alterations in tumor cell proliferation, metabolism, and metastasis. *Molecular Oncology*, *8*(8), 1429–1440. https://doi.org/10.1016/j.molonc.2014.05.017

Rytelewski, M., Vareki, S. M., Mangala, L. S., Jiang, D., Pradeep, S., Rodriguez-aguayo, C., & Koropatnick, J. D. (2016). Reciprocal positive selection for weakness - preventing olaparib resistance by inhibiting BRCA2. *Oncotarget*, *7*(15), 20825–20839. https://doi.org/10.18632/oncotarget.7883

Soukarieh, O., Gaildrat, P., Hamieh, M., Drouet, A., Baert-Desurmont, S., Frébourg, T., … Martins, A. (2016). Exonic splicing mutations are more prevalent than currently estimated and can be predicted by using in silico tools. *PLoS Genetics*, *12*(1), 1–26. https://doi.org/10.1371/journal.pgen.1005756

Spurdle, A. B., Couch, F. J., Hogervorst, F. B. L., Radice, P., & Sinilnikova, O. M. (2008). Prediction and assessment of splicing alterations: Implications for clinical testing. *Human Mutation*, *29*(11), 1304–1313. https://doi.org/10.1002/humu.20901

Strachan, T., & Read, A. (2010). Nucleic acid structure and gene expression. *Human* Molecular Genetics (4th ed., pp. 1–27). Taylor & Francis Group: New York, New York, USA.

Tang, R., Prosser, D. O., & Love, D. R. (2016). Evaluation of bioinformatic programmes for the analysis of variants within splice site consensus regions. *Advances in Bloinformatics*, *2016*, https://doi.org/10.1155/2016/5614058

Théry, J. C., Krieger, S., Gaildrat, P., Révillion, F., Buisine, M.-P., Killian, A., … Tosi, M. (2011). Contribution of bioinformatics predictions and functional splicing assays to the interpretation of unclassified variants of the BRCA genes. *European Journal of Human Genetics*, *19*(10), 1052–1058. https://doi.org/10.1038/ejhg.2011.100

Vallée, M. P., Di Sera, T. L., Nix, D. A., Paquette, A. M., Parsons, M. T., Bell, R., … Tavtigian, S. V. (2016). Adding in silico assessment of potential splice aberration to the integrated evaluation of BRCA gene unclassified variants. *Human Mutation*, *37*(7), 627–639. https://doi.org/10.1002/humu.22973

Vreeswijk, M. P. G., Kraan, J. N., Van Der Klift, H. M., Vink, G. R., Cornelisse, C. J., Wijnen, J. T., … Devilee, P. (2009). Intronic variants in BRCA1 and BRCA2 that affect RNA splicing can be reliably selected by splice-site prediction programs. *Human Mutation*, *30*(1), 107–114. https://doi.org/10.1002/humu.20811

Ware, M. D., DeSilva, D., Sinilnikova, O. M., Stoppa-Lyonnet, D., Tavtigian, S. V., & Mazoyer, S. (2006). Does nonsense-mediated mRNA decay explain the ovarian cancer cluster region of the BRCA2 gene? *Nature*, *25*(2), 323–328. https://doi.org/10.1038/sj.onc.1209033

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Baert A, Machackova E, Coene I, et al. Thorough in silico and in vitro cDNA analysis of 21 putative *BRCA1* and *BRCA2* splice variants and a complex tandem duplication in *BRCA2*, allowing the identification of activated cryptic splice donor sites in *BRCA2* exon 11. *Human Mutation*. 2017;1–12. https://doi.org/10.1002/humu.23390