Humu-2014-0316

Research Article

Functional classification of *BRCA2* DNA variants by splicing assays in a large minigene with 9 exons

Alberto Acedo, Cristina Hernández-Moro, Álvaro Curiel-García, Beatriz Díez-Gómez and Eladio A. Velasco

Splicing and Genetic Susceptibility to Cancer, Instituto de Biología y Genética Molecular (CSIC-UVa), Valladolid, Spain

Contract grant sponsors: Instituto de Salud Carlos III (Spanish Ministry of Economy and Competitiveness) grants PI10/2910 and PI13/1749, Consejería de Sanidad (Junta de Castilla y León) grant BIO/VA08/13. AA was supported by the European Social Fund and Consejeria de Educación de la Junta de Castilla y León under the P.O. Castilla y León 2007-2013. CHM is supported by Fundación Villalar (Castilla y León, Spain)

Corresponding author:

Eladio A. Velasco

Grupo de Splicing y Cáncer

Instituto de Biología y Genética Molecular (IBGM)

Consejo Superior de Investigaciones Científicas (CSIC)-UVa

Sanz y Forés 3.

47003 Valladolid.

Spain

Phone: +34 983184829

Fax: +34 983184800

e-mail: eavelsam@ibgm.uva.es

ABSTRACT

Numerous pathogenic DNA variants impair the splicing mechanism in human genetic diseases. Minigenes are optimal approaches to test variants under the splicing viewpoint without the need of patient samples. We aimed to design a robust minigene construct of the breast cancer gene *BRCA2* in order to investigate the impact of variants on splicing. *BRCA2* exons 19 to 27 (MGBR2_ex19-27) were cloned in the new vector pSAD. It produced a large transcript of the expected size (2,174 nucleotides) and exon structure (V1-ex19-27-V2). Splicing assays showed that 18 (17 splice-site and 1 silencer variants) out of 40 candidate DNA variants induced aberrant patterns. Twenty-four anomalous transcripts were accurately detected by fluorescent-RT-PCR that were generated by exon-skipping, alternative site usage and intron-retention events. Fourteen variants induced major anomalies and were predicted to disrupt protein function so they could be classified as pathogenic. Furthermore, minigene mimicked previously reported patient RNA outcomes of seven variants supporting the reproducibility of minigene assays. Therefore, a relevant fraction of variants are involved in breast cancer through splicing alterations. MGBR2_ex19-

27 is the largest reported *BRCA2* minigene and constitutes a valuable tool for the functional and clinical classification of sequence variations.

Key Words: Hereditary breast and ovarian cancer, BRCA2, splicing, minigenes, Unclassified variants

INTRODUCTION

Recent advances in genomic technologies have allowed the development of new strategies for the molecular diagnostics of human diseases. Actually, hundreds of thousands of DNA variants are detected in massive sequencing projects of genetic disorders. Variations are usually classified attending to their impact on the protein translation, such as synonymous, missense, nonsense and frameshift variants. With the exception of protein-truncating variants, most of the remaining ones are of unknown clinical significance (UV), whose classification as neutral or pathogenic represents a challenge for genetic counselling (Gomez Garcia et al., 2009). Interestingly, recent estimations have shown that an unexpectedly large fraction of genetic diseases are caused by variants that disrupt the splicing process (Wang and Cooper, 2007), ranging from 15% to >60% (Krawczak et al., 1992; Lopez-Bigas et al., 2005). The generation of aberrant splicing events or changes in ratio of isoforms in a specific tissue can be associated with a disorder, its clinical severity and even the penetrance of a particular variant. Intron removal is guided by a large number of splicing factors, ribonucleoproteins and a complex array of sequences including acceptor and donor sites, polypyrimidine tract and branch point, as well as exonic and intronic splicing enhancers and silencers that stimulate or repress exon inclusion into the mature mRNA,

respectively. Consequently, all these elements are potential targets for variants with impact on splicing.

Hereditary breast and ovarian cancer (HBOC) shows high genetic heterogeneity with more than 1,000 susceptibility loci estimated (Michailidou et al., 2013). Germ line variants in the tumor suppressor genes *BRCA1* (MIM# 113705) and *BRCA2* (MIM# 600185) confer high risk to breast and ovarian cancer as well as other types of tumors (Miki et al., 1994; Wooster et al., 1995; Narod and Foulkes, 2004), accounting for about 16% of familial risk of breast cancer (Stratton and Rahman, 2008). Remarkably, about one half of the >3,800 described variants at the Breast Cancer Information Core database (BIC; https://research.nhgri.nih.gov/projects/bic/Member/index.shtml) are UVs (Frank et al., 2002). We formerly reported that one third of disease-causing variants of 14 exons of *BRCA1/2* disrupted splicing (Sanz et al., 2010; Acedo et al., 2012). Interestingly, it was shown that splicing variants are enriched in cancer-related genes (Sterne-Weiler and Sanford, 2014). The characterization of a deleterious splicing variant is a complex task that requires a functional validation as bioinformatics predictions are not today able to predict the consequence of a splicing motif alteration at the mRNA level. A 5-tier scheme was proposed to improve classification of *spliceogenic* variants that integrates bioinformatic and experimental data (Spurdle et al., 2008) whose purpose was to pre-select candidate variants for functional tests as well as to interpret the splicing outcomes. Splicing assays of patient RNA by RT-PCR is the most straightforward and direct method to assess the clinical relevance of DNA changes (Baralle et al., 2009). However, patient samples are not always available so that functional tests of hybrid minigenes in splicing reporter plasmids, such as pSPL3 (Life Technologies, discontinued), have become valuable tools to check the splicing

profiles induced by a sequence variation. The laborious constructions of minigenes and the low pSPL3 efficiency precluded a more widespread use and the cloning of large minigenes with several exons (Burn et al., 1995). In addition, the natural genomic context is a critical factor for exon recognition (Buratti et al., 2006) that must therefore be considered for minigene design.

Given the growing incidence of splicing alterations in disease-predisposition genes and the need for appropriate tools to identify them, we intended to design a BRCA2 minigene with several exons where candidate *spliceogenic* DNA variants could be easily generated and checked. In a recent work, we comprehensively analyzed 41 DNA variants from *BRCA2* exons 19, 20, 23 and 24 in two minigenes (Acedo et al., 2012) that constituted the starting point of this study. Here we show the construction of the largest *BRCA2* minigene (MGBR2_ex19-27, exons 19 to 27) reported so far in the novel splicing vector pSAD (for Splicing And Disease), which produced a transcript of 2,174 nucleotides. In order to validate it, 40 candidate variants from exon 19 to 27 were assayed, 18 of which disrupted *BRCA2* pre-mRNA processing. We showed that the "maxi-minigene" MGBR2_ex19-27 and the splicing reporter plasmid pSAD are valuable tools to classify sequence variations of uncertain clinical significance under the splicing perspective.

MATERIALS AND METHODS

Ethical approval for this study was obtained from the Ethics Review Committee of Faculty of Medicine at the University of Valladolid (14/04/2010).

Databases and Bioinformatics Tools

Data of *BRCA2* variants detected in breast/ovarian cancer patients were taken from the mutation databases: the BIC and the Universal Mutation Databases (UMD; http://www.umd.be/BRCA2/) (Caputo et al., 2012). To identify potential splicing variants, mutant and normal sequences were analyzed with NNSPLICE version 0.9 for splice sites (http://www.fruitfly.org/seq_tools/splice.html) (Reese et al., 1997), and Human Splicing Finder version 2.4.1 (HSF; http://www.umd.be/HSF/) (Desmet et al., 2009), which includes algorithms for splice sites, silencers and enhancers. Variant descriptions were according to the *BRCA2* GenBank sequence NM_000059.1 and the guidelines of the Human Genome Variation Society (HGVS; http://www.hgvs.org/mutnomen/).

Development of the new splicing vector pSAD

The new splicing plasmid pSAD (Patent P201231427-CSIC, Priority Patent Application filed) has a pSPL3 backbone with the following features (Figure 1A): 1) reduction of intron size by 1.2 kb; 2) introduction of Beta-Galactosidase (*LacZ*) as a second selection marker; 3) new polylinker with 27 restriction sites (pSPL3, 10 sites); 4) strengthening of acceptor site of vector exon V2 (NNSPLICE score 0.99 vs. 0.64) by triple mutagenesis; 5) elimination of old pSPL3 restriction sites (HindIII, XbaI and SalI) that interfere with those of the new multiple cloning site (MCS).

Construction of Minigene MGBR2_EX19-27

MGBR2_EX19-27 was constructed in four cloning steps in this order: exons 19 to 24+exon 25+exon 26+exon 27 (Table 1; Figure 1B). Two distinct cloning strategies were used: classical restriction digestion/ligation and overlap extension PCR cloning (Bryksin and Matsumura, 2010). All the inserts were amplified with Phusion High Fidelity polymerase and primers indicated in Table 1. Exons 19-20 and 21-24 were subcloned into the pSAD vector with between the *Xho*I and *Kpn*I sites (MGBR2i_ex19-24). Exon 25 was inserted 3' to exon 24 in the appropriate orientation by overlap extension PCR (MGBR2i_ex19-25). Exon 26 was then cloned downstream exon 25 into the *Kpn*I site (MGBR2i_ex19-26). Finally, we designed a strategy where 752 bp of vector pSAD, including part of the MCS, the vector intron, and 50 bp of exon V2 with its acceptor site, were replaced by 221 bp of intron 26 and 723 bp of exon 27 (MGBR2_ex19-27), so that the final construct contains a chimeric exon 27-V2 with the acceptor site of BRCA2 exon 27 and the annealing site of specific vector primer RT-PSAD-RV.

Site directed mutagenesis. Generation of exonic microdeletions

Mutagenesis was carried out according to the PCR mutagenesis protocol (http://www.methodbook.net/pcr/pcrmut.html) with Pfu UltraHF polymerase (Agilent, Santa Clara, CA). Wild type (wt) minigene MGBR2_ex19-27 was used as template to generate 40 DNA variants reported in the BIC/UMD databases (Supp. Table S1). Thirty-four exonic microdeletions of 30 bp were also generated to map splicing regulatory elements. The first two and the last three nucleotides of each exon were always preserved to avoid disruptions of the natural acceptor and donor sites, respectively. Deletions were introduced

by PCR-mutagenesis with chimeric 50-mer primers (Supp. Table S1) containing 25 nucleotides of each border of the deletion.

Transfection of eukaryotic cells

Approximately 10^5 HeLa (human cervical carcinoma) or MCF-7 (human breast adenocarcinoma) cells were grown to 90% confluency in 0.5 mL of medium (DMEM, 10% Fetal Bovine Serum, 1% glucose and 1% Penicillin/Streptomycin) in 4-well plates (Nunc, Roskilde, Denmark). Cells were transiently transfected with 1 µg of each minigene and 2 µL of Lipofectamine 2000 (Life Technologies, Carlsbad, CA). To inhibit nonsense mediated decay (NMD), cells were incubated with cycloheximide (Sigma-Aldrich, St. Louis, MO) 300 µg/mL (4 hours). RNA was purified with Nucleospin-RNA-II (Macherey-Nagel, Düren, Germany) with on-column rDNAse treatment.

RT-PCR of minigenes

Retrotranscription was carried out with 200 ng of RNA and the Transcriptor cDNA synthesis kit (Roche Applied Science, Penzberg, Germany). Two µL of cDNA were amplified with GoTaq polymerase (Promega, Madison, WI, USA) and specific primers of the vector exons RTpSAD-FW and RTpSAD-RV (Patent P201231427,CSIC), or one of both combined with one of the following *BRCA2* exonic primers: RTBR2_ex24-FW (5'-TTTTTAGATCCAGACTTTCAGC-3'), RTBR2_ex25-FW (5'-TTTGCTGGAGATTTTTCTGTG-3'), RTBR2_ex22-RV (5'-TGGATCTGAGCTTGTTTCTT-3'), RTBR2_ex23-RV (5'-ATTCTGTATCTCTTTCCTTCTGTT-3') and RTBR2_ex25-RV (5'-CACAGAAAAATCTCCAGCAAA-3'). RT-PCR primers from pSPL3, SD6-PSPL3_RTFW (5'-TCACCTGGACAACCTCAAAG-3') and SA2-PSPL3_RTREV (5'TGAGGAGTGAATTGGTCGAA3'), were also used. Samples were denatured at

94ºC for 5 min, followed by 35 cycles consisting of 94ºC for 20 sec, 58ºC for 20 sec, and 72ºC (1 min/kb), and a final extension step at 72ºC for 5 min.

Semiquantitative fluorescent RT-PCRs were done in triplicate with one FAM-labelled vector primer. One µL of a dilution 1/10-1/20 of the RT-PCR products was mixed with 18 µL of Hi-Di Formamide (Life Technologies) and 0.2 µL of Genescan 500 Rox Size Standard (Life Technologies) or 1.0 µL LIZ-1200. Samples were run on an ABI3130 sequencer and analyzed with Peak Scanner (Life Technologies). Mean peak areas of each transcript and standard deviations were calculated.

RESULTS

Construction and validation of the minigene MGBR2_ex19-27

With the purpose to maintain the natural genomic context, a 9-exon minigene was constructed (MGBR2_ex19-27) that contained an insert of 5,620 bp with 9 exons of *BRCA2* (19 to 27) (Figure 1B) representing a 5-fold reduction of the original genomic region (28,717 bp). It constitutes the largest *BRCA2* minigene ever reported. To check it, constructs with *BRCA2* exons 19-24 (MGBR2i_ex19-24), 19-25 (MGBR2i_ex19-25), 19-26 (MGBR2i_ex19-26) and 19-27 (MGBR2_ex19-27) as well as the plasmid pSAD were transfected and RNA retrotranscribed and amplified with specific vector exon primers (Figure 1C). An increasing ladder of transcripts of the expected sizes from minigene MGBR2i_ex19-24 to 19-27 could be observed, as exons 25, 26 and 27 were successively incorporated in the transcript, respectively. The final construct MGBR2_ex19-27 produced a large transcript of 2,174 nt of the appropriate size and exon composition (Figure 1) that was completely sequenced and

consisted of V1, *BRCA2*-19-20-21-22-23-24-25-26-27 and V2 exons (canonical transcript), without any anomaly.

Splicing assays of DNA variants from exons 19 to 27

To validate minigene MGBR2_ex19-27 for functional classification of DNA variants, 40 variants were tested, four of which from exons 19, 20, 23 and 24 were previously assayed in pSPL3 (c.8486A>T, c.8488-1G>A, c.9117G>A, c.9256+1G>A, minigenes 19-20 and 23-24) (Acedo et al., 2012). A total of 166 BIC and UMD variants of exons 21, 22, 25, 26 and 27 were analyzed with Human Splicing Finder and NNSPLICE. Thirty six out of them were selected according to the following criteria: disruption of canonical acceptor and donor sites (only one variant per splice site was usually selected), creation of new splice sites, ESE elimination or ESS creation (Table 2; Supp. Tables S1 and S2). Fourteen were intronic (mostly surrounding the splice sites) and 22 exonic that comprised 11 predicted missense, 4 frameshift, 3 synonymous and 2 nonsense changes, as well as 2 splice site deletions.

All variants were generated in MGBR2_ex19-27 and assayed in HeLa cells. Eighteen of them (45%) showed anomalous splicing patterns by capillary electrophoresis (CE) (Figure 2), which were 12 intronic single nucleotide substitutions, two splice site deletions, two predicted missense and two synonymous changes (Table 2). Regarding the bioinformatics output, 17 disrupted the acceptor or donor sites and one created two silencers.

A total of 25 different isoforms, including the canonical transcript, were detected and characterized by CE except for two, designated as ins23 (exon 20) and 436-nt (exon 26). The anomalous events were exon skipping, intron retention and usage of alternative donor or acceptor splice sites. Most of the characterized RNA isoforms (18/22) were predicted to

truncate the BRCA2 protein, including intron 19 retention, exon 20 skipping, exon 21 skipping, ivs21-ins46, exon 22-del1, exon 22 skipping, exon 22-del7, intron 22 retention, ivs22-ins4, ivs22del11- retention, exon 23 skipping, exon 24 skipping, exon 24-del43, ivs25-ins70nt, exon 25 skipping, exon 26-del28, exon 27-del424 and exon 27-del62 (Figure 2). Only four isoforms kept the ORF that included exon 19 skipping and exon 20-del12, both of which were discussed elsewhere (Acedo et al., 2012), and exon 25-del27 and exon 26 skipping whose roles in the disease are arguable. Interestingly, most variants induced more than one splicing isoform, but remarkably, c.8488-1G>A triggered 5 different transcripts. To compare the splicing outputs in different cell lines, seven variants with impact on splicing (c.8486A>C, c.8488-1G>A, c.9117G>A, c.9256G>T, c.8948_8953+5del, c.8953+1G>T and c.9649-2A>G) and the wt minigene were also checked in MCF-7 but no significant differences were observed with HeLa cells as we previously demonstrated for MCF-10A cells (Sanz et al., 2010) (data not shown).

Impact of NMD on splicing

The NMD process can imbalance the ratio of isoforms that could be essential for the clinical classification of variants. We checked the effect of the NMD inhibitor cycloheximide on the splicing outcome of four variants (c.8486A>T, c.8754+4A>G, c.8948_8953+5del and c.9649-2A>G). With the exception of c.8754+4A>G that produced 100% of the aberrant transcript Ivs21-ins46 in both experimental conditions, we observed remarkable differences in peak areas or band intensity between presence and absence of cycloheximide (data not shown). Figure 3 shows the effect of cycloheximide on the splicing outcome of variant c.8948_8953+5del. Changes in band intensities (transcripts) and the presence of a new one were already visualized in agarose gels (Figure 3A, left), but these were even more

pronounced in CE electropherograms (Figure 3A, right). Thus, in the absence of cycloheximide c.8948_8953+5del generated four isoforms: intron 22-del11 retention, exon 22del7, exon 22 skipping and exon 22del31, whereas NMD inhibition enabled the detection of a new transcript (exon 22del41) and modified their relative proportions (Figure 3B).

Mapping of splicing enhancers

Efficient exon recognition depends on additional regulatory sequences besides the canonical splice sites. Density of active ESEs is highest near splice sites (~50 nt at both exon ends) with a maximum between 10 and 20 nucleotides from the natural donor and acceptor sites within each exon (Fairbrother et al., 2004) so DNA variants in these regions have higher likelihood of disrupting ESEs. Theoretically, disruptions of active enhancers would trigger anomalous splicing reactions. To map functional ESEs we introduced 34 30-nt microdeletions of the 5' and 3' ends of each exon by mutagenesis, two on each exon end that overlapped 5 nucleotides, except for the final exon 27 where only two deletions were made on its 5' end. All these microdeletions would cover the critical ESE-containing regions of the nine exons. Two microdeletions, c.8334_8363del (exon 19), c.8490_8519del (exon 20), clearly affected the splicing process (Figure 4). Faint abnormal bands were observed in deletions c.8359_8388del (exon 19), c.8575_8604del (exon 20), c.8697_8726del and c.8722_8751del (exon 21), c.8956_8985del and c.8981_9010del (exon 23) that revealed weak effects (data not shown). These results suggest that the deleted sequences contain supplementary regulatory elements that would promote exon recognition.

DISCUSSION

The clinical classification of DNA variants in disease genes poses a challenge for medical geneticists since it provides essential information for clinical management of patients and asymptomatic carriers. Despite most of the disease-causing variants are assumed to impair the protein function there has been found that a high fraction of them can actually disrupt the splicing pattern of the responsible gene (Acedo et al., 2012). However, the identification of genomic variants that trigger aberrant splicing is not as simple. The ideal manner to study it is on patient RNA but this is often difficult to obtain. Hence it is required a reliable and straightforward substitute methodology to assess splicing. Precisely, *ex vivo* assays of DNA variants with splicing reporter minigenes have emerged to circumvent this problem. Indeed, pSAD is an efficient splicing vector, since it allowed the construction of the largest minigene with nine exons reported so far. Remarkably, the new minigene was genetically engineered to construct a chimeric last exon 27-V2 preserving the necessary sequences to anneal the specific V2 RT-PCR primer that only amplifies RNA synthesized by the minigene.

The wt minigene MGBR2_ex19-27 mostly expressed the full-length transcript (2,174 nucleotides) confirming its stability and robustness. This minigene also allowed us to identify 18 out of 40 variants that induce splicing anomalies. Until now we have functionally assayed 139 reported variants from 22 *BRCA1/2* exons, 63 of which disrupted splicing (45.3%), accounting for 7% (63/905) of all variants of these exons recorded at the BIC database. However, this may be even underrated taking into account a previous comprehensive estimation (Mucaki et al., 2011) and that silencer and enhancer alterations have not been included, supporting that variants with impact on splicing constitute a

prevalent etiopathogenic mechanism in HBOC. In any case, *BRCA* minigenes, such as MGBR2_ex19-27, are powerful tools to establish the exact proportion of *spliceogenic* variants. Indeed, current ongoing projects of our laboratory are developing large pSAD-minigenes to cover all the *BRCA1-BRCA2* exons.

In this sense, it is also worth mentioning the simplicity of the minigene test where insert cloning can be considered the rate-limiting step. This process might be accelerated through gene synthesis of the complete construct (vector + exons) although this would significantly increase the final costs. Once the minigene is assembled, it can be used as a template to introduce any change. Moreover, many candidate splice variant from these nine exons (MGBR2_ex19-27) can be tested in parallel by a simple and rapid assay consisting of site-directed mutagenesis–transfection–RT-PCR–electrophoresis/sequencing in only 2-3 weeks. In addition, fluorescent-CE is a highly sensitive technique that is able to detect very rare RT-PCR products and to discriminate isoforms that differ in size by a few nucleotides (Acedo et al., 2012; Whiley et al., 2014). CE has been proposed by the ENIGMA (Evidence-based Network for the Interpretation of Germline Mutant Alleles) consortium to detect low-abundance transcripts (Whiley et al., 2014). The use of NMD inhibitors still improves the detection and quantification of transcripts that are critical to evaluate their contribution to the disease.

Genomic context and splicing outcomes of DNA variants

The ultimate goal of constructing of large minigenes such as MGBR2_ex19-27 was to imitate utmost the natural genomic context where the splicing reactions take place since it constitutes a critical factor influencing pre-mRNA processing. This effect is well illustrated by exon 37 processing of the *NF1* gene, which is only efficiently recognized in minigenes with at least exons 34 to 38 (Buratti et al., 2006; Baralle et al., 2006). Interestingly, one variant of exon 37 (c.6792C>G) only replicated the *in vivo* splicing pattern of patient when it was assayed in a minigene with exons 31 to 38. Similarly, we previously observed that variant c.8488-1G>A principally induced exon 20 skipping in a two-exon minigene (19 and 20) whereas in MGBR2_ex19-27 produced five different aberrant transcripts, being intron 19 retention the most abundant isoform like in patient RNA (Acedo et al., 2012). Moreover, study of the impact of variant c.7806-14T>C on two naturally occurring splicing events (Δ17,18 and Δ18) that were upregulated in carriers of allele C showed almost identical results in lymphocytes and a pSAD-minigene with exons 16 to 20 (de Garibay et al., 2014). Therefore, the restoration of the genomic background is a key issue to replicate the *in vivo* pattern, so any tested exon of a minigene should be flanked by their natural 5' and 3' exons. In this regard, large genes, such as TTN with 364 exons and a mRNA over 100 kb long (Chauveau et al., 2014), represent an interesting technical challenge for the construction and validation of reliable minigenes to assay potential *spliceogenic* variants.

We further compared our results with other previous studies with minigenes or patient RNA. Variants c.8754+4A>G, c.8754+5G>A (46-nt insertion of intron 21, also c.8754+1G>A and c.8754+4A>G not tested here), c.9117G>A (exon 23 skipping), c.9256+1G>A (exon 24 skipping) and c.9501+3A>T (partial exon 25 skipping) were previously

reported to display the same splicing patterns in lymphocytes (Claes et al., 2003; Bonatti et al., 2006; Bonnet et al., 2008; Hansen et al., 2008; Vreeswijk et al., 2009; Thomassen et al., 2012). Also, variant c.8486A>G, in the same position that c.8486A>T tested by us (Table 2), was reported to have the same outcome (exon 19 skipping) (Houdayer et al., 2012). Therefore, a total of seven variants from exons 19 to 27 yield the same or similar results in patient and minigene assays, supporting the reproducibility of MGBR2_ex19-27 results. On the other hand, genetic variants c.8755-1G>A, c.8953+1G>T and c.9257-1G>C were reported with different splicing profiles (Machackova et al., 2008; Whiley et al., 2011). In the two latter cases this may be due to the use of NMD inhibitors that could modify the number and amount of RNA isoforms so both results are not comparable.

Clinical classification of DNA variants

The clinical interpretation of the splicing results of a genetic variant is a particularly complex task. Identification of *spliceogenic* variants will aid in breast cancer prediction, prevention and surveillance, and will contribute to elucidate the BC predisposition spectrum. It is accepted that a variant would be considered pathogenic when it causes major or total splicing aberrations and generates a PTC or an in-frame deletion of a known functional domain. This matches class C5 of the previously reported 5-tiered classification scheme of variants (Spurdle et al., 2008). Fifteen variants induced major aberrant transcripts of which 14 fulfilled these criteria of pathogenicity: c.8486A>T, c.8488-1G>A, c.8633-24_8634del, c.8754G>A, c.8754+4A>G, c.8754+5G>A, c.8754+5G>T, c.8755-1G>A, c.8948_8953+5del, c.8953+1G>T, c.9117G>A, c.9256+1G>A, and c.9502-2A>C.

In contrast, variant c.9257-1G>C and c.9502-2A>C induced isoforms that were predicted to cause in-frame deletions of 9 (exon 25-del27) and 49 codons (exon 26

skipping), respectively. The first one was predicted to delete amino acids 3086 to 3094 of which Tyr3092 is conserved across species and Gly3086 in vertebrates. Exon 26 skipping constitutes 33% of aberrant transcripts induced by c.9502-2A>C (Table 2). It is expected to delete 49 amino acids from Asn3168 to Leu3216 and involves 35 amino acids (3168-3192) of the C-terminal end of the oligonucleotide/oligosaccharide-binding fold (OB3; amino acids 3051-3192) that also binds p53 (Rajagopalan et al., 2010) where only Leu3180 is conserved among vertebrates. Therefore, both variants require further studies to classify them as neutral or deleterious.

Only three DNA changes produced weak effects (<15% of abnormal isoforms): c.9502-12T>G (polypyrimidine tract of intron 25), c.9501+3A>T (intron 25) and c.9698G>T (exon 27). Thus, their role in BC is questionable but they might constitute low-penetrance or disease-modifier alleles as it was demonstrated in *CFTR* related disorders or Menkes disease (Steiner et al., 2004; Douglas and Wood, 2011). These variants might provide low-penetrance alleles that might interact synergistically with other protector and risk alleles to modify the overall BC risk. The integration of all these data into a single model to calculate BC risk would improve disease prediction and prevention.

Splicing Regulatory Sequences

Seventeen variants affected the splice junctions and only one created splicing silencers (c.9698G>T). The splice site software was accurate to determine variants with impact on splicing, as previously reported (Houdayer et al., 2008), but the splicing outcomes (exon skipping, use of cryptic sites, etc) were unpredictable reinforcing the need for a confirmatory functional test. Interestingly, c.9698G>T had a weak/partial effect such as other enhancer/silencer alterations (*BRCA1* c.5123C>A or *BRCA2* c.8257_8259del) (Bonnet

et al., 2008; Millevoi et al., 2010), in contrast to *BRCA1*-c.5434C>G (Gaildrat et al., 2010), or *BRCA2*-c.145G>T (Sanz et al., 2010) associated with total disruptions. ESE/ESS predictions programs have shown low accuracy as these elements are constituted by short-degenerate motifs. Combined computational and experimental approaches are expected to ultimately yield a full map of functional cis-acting elements operating at the RNA level that guide physiological or pathological alternative splicing events. Indeed, the functional ESE mapping by exonic microdeletions has been revealed as a helpful approach to specifically delimitate enhancer-containing sequences (Figure 4) where only predicted ESE-disrupting variants would be chosen and assayed. Thus, three previously characterized SRE variants, c.8378G>A (exon 19), c.8969G>A (exon 23) and c.9006A>T (exon 23) (Acedo et al., 2012), are placed within the positive microdeletions c.8359_8388del, c.8956_8985del and c.8981_9010del, respectively, which strengthens the utility of this strategy to locate putative ESE-variants. The ESE motifs involved in exon recognition can be finely mapped with subsequent smaller deletions of the 30-nt positive segment. To identify the specific splicing factors that bind such motifs, novel experiments of inhibition with siRNAs or *in vitro* binding experiments would be required (Goina et al., 2008).

We can therefore conclude that splicing impairment significantly contributes to the inherited susceptibility of HBOC. Furthermore, the pSAD plasmid and their derived minigenes are useful tools for molecular diagnostics and genetic counselling of hereditary breast/ovarian cancer as well as the basic research of the splicing process. Summarizing, splicing reporter minigenes such as MGBR2_ex19-27 have the following advantages: i) no need of patient RNA; ii) straightforward laboratory protocol; iii) systematic analysis and functional classification of DNA variants from several exons with a single minigene; iv)

analysis and quantification of the splicing outcome of a single mutant allele without the interference of the wt one; v) high reproducibility of results; vi) assay of any human disease gene in pSAD and hence it is a valuable way to functionally evaluate whichever candidate DNA variants detected, for example, in next-generation sequencing projects. Finally, a clear understanding of disease-causing events will allow the development of individualized treatments.

ACKNOWLEDGMENTS

EAV was supported by grants PI10/2910 and PI13/1749 of the Instituto de Salud Carlos III (Spanish Ministry of Economy and Competitiveness) and BIO/VA08/13 (Consejería de Sanidad, Junta de Castilla y León). AA was supported by the European Social Fund and Consejeria de Educación de la Junta de Castilla y León under the P.O. Castilla y León 2007-2013. We are grateful to the breast/ovarian cancer patients and clinicians who collaborated in this study, particularly Dr. Germán Marcos (Hospital Río Hortega, Valladolid) and Dr. Enrique Lastra (Complejo Hospitalario de Burgos). We thank Dr. Mar Infante, Dr. Mercedes Durán and Dr. Cristina Miner (Cancer Genetics Laboratory, IBGM) for their scientific support and helpful discussions, and Eugenia Fraile-Bethencourt for the critical reading of the manuscript. We sincerely thank Sandra Villafruela (Viana de Cega, Valladolid) for her enormous effort and constant and enthusiastic support. The authors declare no conflict of interest.

AUTHOR'S CONTRIBUTIONS

AA, BD-G and EAV designed and developed the pSAD plasmid. AA contributed to the experimental design, bioinformatics analysis, minigene construction and performed most of the splicing functional assays. BD-G participated in minigene construction, splicing assays and sequencing.CH-M and AC-G actively participated in the functional assays and bioinformatics analyses. EAV conceived the study, supervised all the experiments and wrote the manuscript. All authors contributed to data interpretation, revisions of the manuscript and approved the final version of the manuscript.

REFERENCES

Acedo A, Sanz DJ, Duran M, Infante M, Perez-Cabornero L, Miner C, Velasco EA. 2012. Comprehensive splicing functional analysis of DNA variants of the BRCA2 gene by hybrid minigenes. Breast Cancer Res 14: R87-

Baralle D, Lucassen A, Buratti E. 2009. Missed threads. The impact of pre-mRNA splicing defects on clinical practice. EMBO Rep 10: 810-816.

Baralle M, Skoko N, Knezevich A, De CL, Motti D, Bhuvanagiri M, Baralle D, Buratti E, Baralle FE. 2006. NF1 mRNA biogenesis: effect of the genomic milieu in splicing regulation of the NF1 exon 37 region. FEBS Lett 580: 4449-4456.

Bonatti F, Pepe C, Tancredi M, Lombardi G, Aretini P, Sensi E, Falaschi E, Cipollini G, Bevilacqua G, Caligo MA. 2006. RNA-based analysis of BRCA1 and BRCA2 gene alterations. Cancer Genet Cytogenet 170: 93-101.

Bonnet C, Krieger S, Vezain M, Rousselin A, Tournier I, Martins A, Berthet P, Chevrier A, Dugast C, Layet V, Rossi A, Lidereau R et al. 2008. Screening BRCA1 and BRCA2 unclassified variants for splicing mutations using reverse transcription PCR on patient RNA and an ex vivo assay based on a splicing reporter minigene. J Med Genet 45: 438-446.

Bryksin AV and Matsumura I. 2010. Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. Biotechniques 48: 463-465.

Buratti E, Baralle M, Baralle FE. 2006. Defective splicing, disease and therapy: searching for master checkpoints in exon definition. Nucleic Acids Res 34: 3494-3510.

Burn TC, Connors TD, Klinger KW, Landes GM. 1995. Increased exon-trapping efficiency through modifications to the pSPL3 splicing vector. Gene 161: 183-187.

Caputo S, Benboudjema L, Sinilnikova O, Rouleau E, Beroud C, Lidereau R. 2012. Description and analysis of genetic variants in French hereditary breast and ovarian cancer families recorded in the UMD-BRCA1/BRCA2 databases. Nucleic Acids Res 40: D992-1002.

Chauveau C, Rowell J, Ferreiro A. 2014. A rising titan: TTN review and mutation update. Hum Mutat 35: 1046-1059.

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 Cancer 37: 314-20.

de Garibay GR, Acedo A, Garcia-Casado Z, Gutierrez-Enriquez S, Tosar A, Romero A, Garre P, Llort G, Thomassen M, Diez O, Perez-Segura P, az-Rubio E et al. 2014. Capillary Electrophoresis Analysis of Conventional Splicing Assays: IARC Analytical and Clinical Classification of 31 BRCA2 Genetic Variants. Hum Mutat 35: 53-57.

Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. 2009. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res 37: e67-

Douglas AG and Wood MJ. 2011. RNA splicing: disease and therapy. Brief Funct Genomics 10: 151-164.

Fairbrother WG, Holste D, Burge CB, Sharp PA. 2004. Single nucleotide polymorphism-based validation of exonic splicing enhancers. Plos Biology 2: 1388-1395.

Frank TS, Deffenbaugh AM, Reid JE, Hulick M, Ward BE, Lingenfelter B, Gumpper KL, Scholl T, Tavtigian SV, Pruss DR, Critchfield GC. 2002. Clinical characteristics of individuals with germline mutations in BRCA1 and BRCA2: analysis of 10,000 individuals. J Clin Oncol 20: 1480-90.

Gaildrat P, Krieger S, Thery JC, Killian A, Rousselin A, Berthet P, Frebourg T, Hardouin A, Martins A, Tosi M. 2010. The BRCA1 c.5434C->G (p.Pro1812Ala) variant induces a deleterious exon 23 skipping by affecting exonic splicing regulatory elements. J Med Genet 47: 398-403.

Goina E, Skoko N, Pagani F. 2008. Binding of DAZAP1 and hnRNPA1/A2 to an exonic splicing silencer in a natural BRCA1 exon 18 mutant. Mol Cell Biol 28: 3850-3860.

Gomez Garcia EB, Oosterwijk JC, Timmermans M, van Asperen CJ, Hogervorst FB, Hoogerbrugge N, Oldenburg R, Verhoef S, Dommering CJ, Ausems MG, van Os TA, van der Hout AH et al. 2009. A method to assess the clinical significance of unclassified variants in the BRCA1 and BRCA2 genes based on cancer family history. Breast Cancer Res 11: R8-

Hansen TV, Bisgaard ML, Jonson L, Albrechtsen A, Filtenborg-Barnkob B, Eiberg H, Ejlertsen B, Nielsen FC. 2008. Novel de novo BRCA2 mutation in a patient with a family history of breast cancer. BMC Med Genet 9: 58-

Houdayer C, Caux-Moncoutier V, Krieger S, Barrois M, Bonnet F, Bourdon V, Bronner M, Buisson M, Coulet F, Gaildrat P, Lefol C, Leone M et al. 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. Hum Mutat 33: 1228-1238.

Houdayer C, Dehainault C, Mattler C, Michaux D, Caux-Moncoutier V, Pages-Berhouet S, d'Enghien CD, Lauge A, Castera L, Gauthier-Villars M, Stoppa-Lyonnet D. 2008. Evaluation of in silico splice tools for decision-making in molecular diagnosis. Hum Mutat 29: 975-982.

Krawczak M, Reiss J, Cooper DN. 1992. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. Hum Genet 90: 41-54.

Lopez-Bigas N, Audit B, Ouzounis C, Parra G, Guigo R. 2005. Are splicing mutations the most frequent cause of hereditary disease? FEBS Lett 579: 1900-1903.

Machackova E, Foretova L, Lukesova M, Vasickova P, Navratilova M, Coene I, Pavlu H, Kosinova V, Kuklova J, Claes K. 2008. Spectrum and characterisation of BRCA1 and BRCA2 deleterious mutations in high-risk Czech patients with breast and/or ovarian cancer. BMC Cancer 8: 140-

Michailidou K, Hall P, Gonzalez-Neira A, Ghoussaini M, Dennis J, Milne RL, Schmidt MK, Chang-Claude J, Bojesen SE, Bolla MK, Wang Q, Dicks E et al. 2013. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. Nat Genet 45: 353-361.

Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W, Bell R, Rosenthal J et al. 1994. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 266: 66-71.

Millevoi S, Bernat S, Telly D, Fouque F, Gladieff L, Favre G, Vagner S, Toulas C. 2010. The c.5242C>A BRCA1 missense variant induces exon skipping by increasing splicing repressors binding. Breast Cancer Res Treat 120: 391-399.

Mucaki EJ, Ainsworth P, Rogan PK. 2011. Comprehensive prediction of mRNA splicing effects of BRCA1 and BRCA2 variants. Hum Mutat 32: 735-742.

Narod SA and Foulkes WD. 2004. BRCA1 and BRCA2: 1994 and beyond. Nat Rev Cancer 4: 665-676.

Rajagopalan S, Andreeva A, Rutherford TJ, Fersht AR. 2010. Mapping the physical and functional interactions between the tumor suppressors p53 and BRCA2. Proc Natl Acad Sci U S A 107: 8587-8592.

Reese MG, Eeckman FH, Kulp D, Haussler D. 1997. Improved splice site detection in Genie. J Comput Biol 4: 311-323.

Sanz DJ, Acedo A, Infante M, Duran M, Perez-Cabornero L, Esteban-Cardenosa E, Lastra E, Pagani F, Miner C, Velasco EA. 2010. A high proportion of DNA variants of BRCA1 and BRCA2 is associated with aberrant splicing in breast/ovarian cancer patients. Clin Cancer Res 16: 1957-1967.

Spurdle AB, Couch FJ, Hogervorst FB, Radice P, Sinilnikova OM. 2008. Prediction and assessment of splicing alterations: implications for clinical testing. Hum Mutat 29: 1304-1313.

Steiner B, Truninger K, Sanz J, Schaller A, Gallati S. 2004. The role of common single-nucleotide polymorphisms on exon 9 and exon 12 skipping in nonmutated CFTR alleles. Hum Mutat 24: 120-129.

Sterne-Weiler T and Sanford JR. 2014. Exon identity crisis: disease-causing mutations that disrupt the splicing code. Genome Biol 15: 201-

Stratton MR and Rahman N. 2008. The emerging landscape of breast cancer susceptibility. Nat Genet 40: 17-22.

Thomassen M, Blanco A, Montagna M, Hansen TV, Pedersen IS, Gutierrez-Enriquez S, Menendez M, Fachal L, Santamarina M, Steffensen AY, Jonson L, Agata S et al. 2012. Characterization of BRCA1 and BRCA2 splicing variants: a collaborative report by ENIGMA consortium members. Breast Cancer Res Treat 132: 1009-1023.

Vreeswijk MP, Kraan JN, van der Klift HM, Vink GR, Cornelisse CJ, Wijnen JT, Bakker E, van Asperen CJ, Devilee P. 2009. Intronic variants in BRCA1 and BRCA2 that affect RNA splicing can be reliably selected by splice-site prediction programs. Hum Mutat 30: 107-114.

Wang GS and Cooper TA. 2007. Splicing in disease: disruption of the splicing code and the decoding machinery. Nat Rev Genet 8: 749-761.

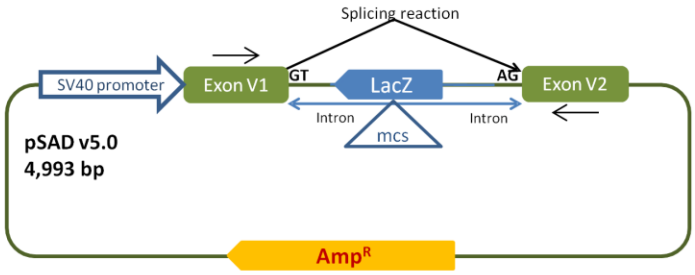
Whiley PJ, de la HM, Thomassen M, Becker A, Brandao R, Pedersen IS, Montagna M, Menendez M, Quiles F, Gutierrez-Enriquez S, De LK, Tenes A et al. 2014. Comparison of mRNA Splicing Assay Protocols across Multiple Laboratories: Recommendations for Best Practice in Standardized Clinical Testing. Clin Chem 60: 341-352.

Whiley PJ, Guidugli L, Walker LC, Healey S, Thompson BA, Lakhani SR, Da Silva LM, Tavtigian SV, Goldgar DE, Brown MA, Couch FJ, Spurdle AB. 2011. Splicing and multifactorial analysis of intronic BRCA1 and BRCA2 sequence variants identifies clinically significant splicing aberrations up to 12 nucleotides from the intron/exon boundary. Hum Mutat 32: 678-687.

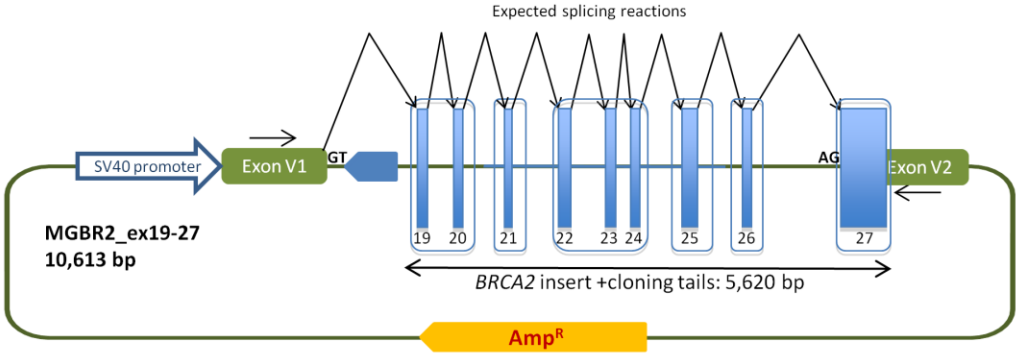
Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, Collins N, Gregory S, Gumbs C, Micklem G, Barfoot R, Hamoudi R et al. 1995. Identification of the Breast-Cancer Susceptibility Gene Brca2. Nature 378: 789-792.

Figure 1. Structures and functional analysis of the splicing vector pSAD and the minigene MGBR2_ex19-27. A) The pSAD vector contains a SV40 transcription promoter, two constituve exons (V1 and V2), and two selection markers: ampicillin resistance and β-galactosidase (LacZ) with a multiple cloning site (mcs). Specific primers to amplify minigene transcripts are indicated by arrows in V1 and V2 exons; B) Structure of the minigene MGBR2_ex19-27: [IVS18 (247 bp) - EX19 (156 bp) - IVS19 (398 bp) - EX20 (145 bp) - IVS20 (207 bp) // IVS20 (90 bp) - EX21 (122 bp) - IVS21 (262 bp) // IVS21 (262 bp) - EX22 (199 bp) - IVS22 (234 bp) - EX23 (164 bp) - IVS23 (93 bp) - EX24 (139 bp) - IVS24 (147 bp)// IVS24 (271 bp) - EX25 (245 bp) - IVS25 (431 bp) // IVS25 (341 bp) - EX26 (147 bp) - IVS26 (344 bp) // IVS26 (221 bp) - EX27 (723 bp)]. The expected splicing reactions in eukaryotic cells are indicated by arrows. The acceptor site and part of exon V2 was replaced by that with 221 bp of intron 26 and 723 bp of exon 27 with its corresponding acceptor site but maintaining the V2 sequence to anneal specific reverse RT-PCR primer. C) Splicing Functional Assays of the pSAD® v5.0 vector, the intermediate minigenes MGBR2i_EX19-24, MG BR2i_EX19-25, MG BR2i_EX19-26 and the final construct MGBR2_EX19-27. RNA was retrotranscribed and amplified with vector specific primers. Sizes of each transcript are indicated below each band.

A)



B)



C)

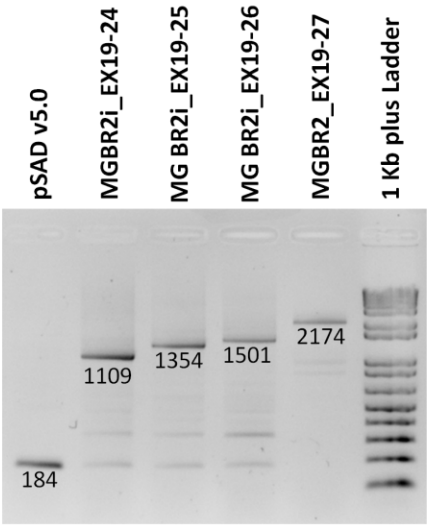


Figure 1

Figure 2. Capillary electrophoresis of fluorescent RT-PCR products from wild type MGBR2_ex19-27 and derived mutant minigenes. RT-PCR products were amplified with one FAM-primer forward (priming on exon V1) or reverse (V2) of the pSAD vector and one exonic BRCA2 primer. All samples were run on an ABI3130 sequencer with Genescan LIZ1200 ("orange/faint"peaks) as size standard. Screenshots of electropherograms visualized with the Peak Scanner software are shown. Fragment sizes and relative fluorescent units are indicated on the x- and y-axes, respectively. WT designates the expected canonical transcript with each primer pair. A) Analysis of exons 19 and 20 (primers RTpSAD-FW and RTBR2_ex22-RV); B) exons 21 and 22 (RTpSAD-FW and RTBR2_ex23-RV); C) exons 23 and 24 (RTpSAD-FW and RTBR2_ex25-RV); D) exon 25 (RTBR2_ex24-FW and RTpSAD-RV); E) exons 26 and 27 (RTBR2_ex25-FW and RTpSAD-RV). The exon composition and the splicing reaction of the different isoforms are shown on the right. Transcripts that were not fully characterized (ins23 or 436-nt) are not represented.

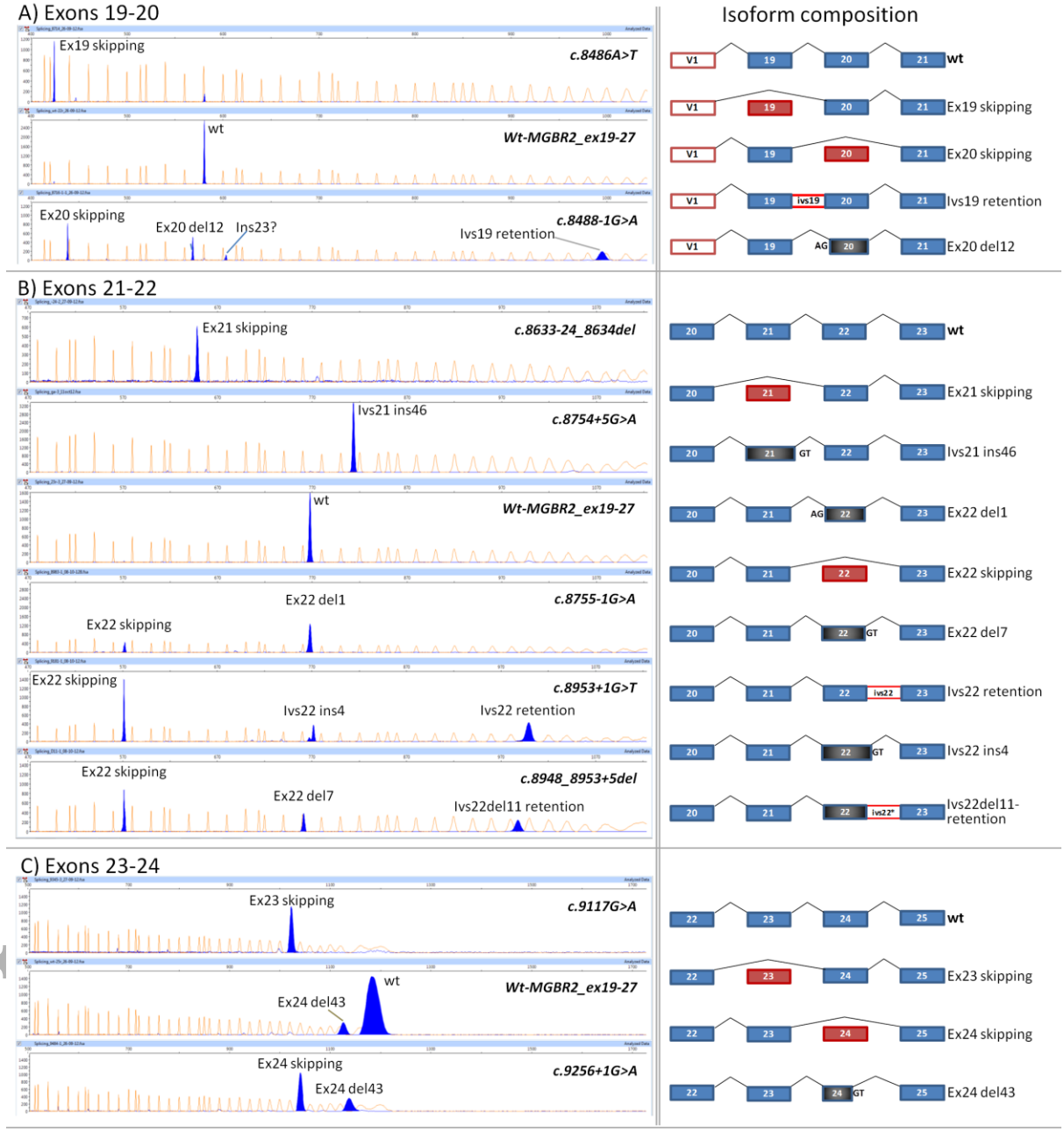


Figure 2

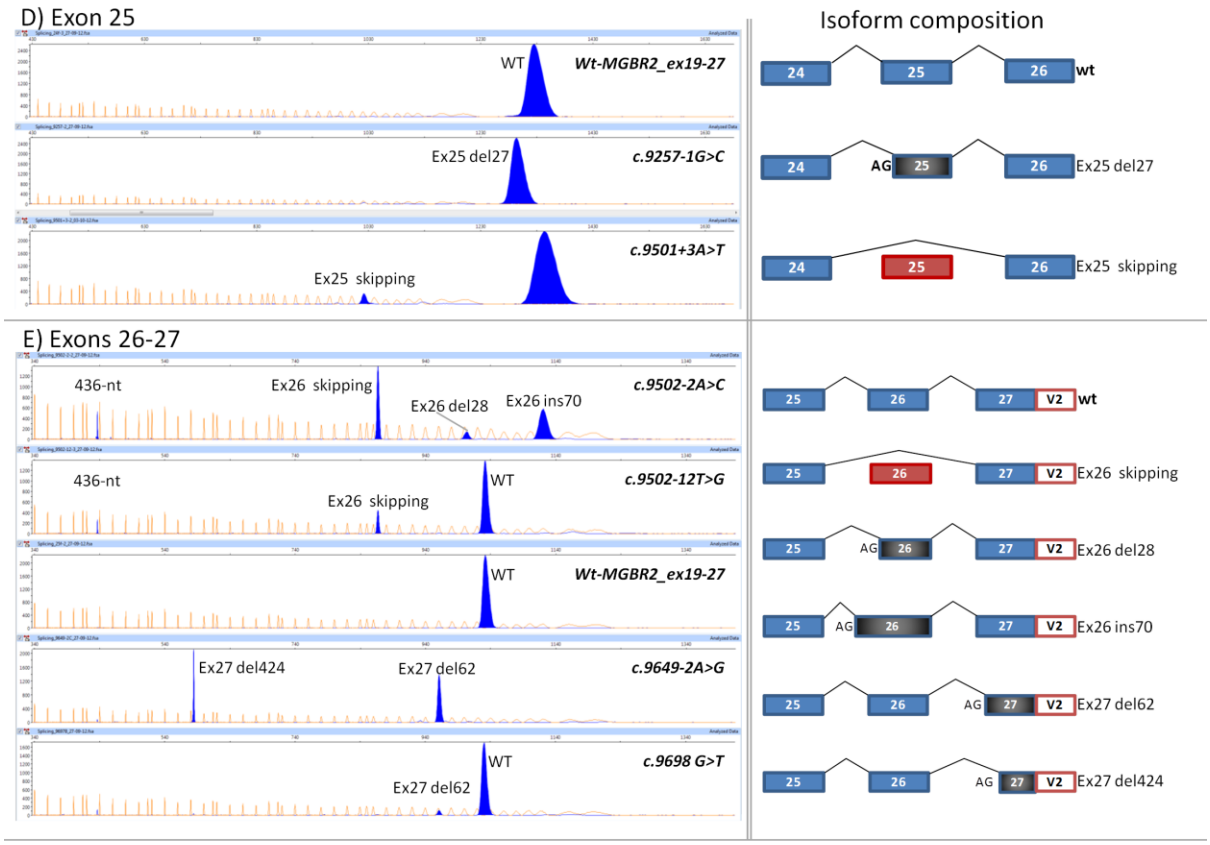


Figure 2

Figure 3. Effect of Nonsense-mediated Decay inhibitors on splicing outcome of DNA variant c.8948_8953+5del. Wild type and mutant minigenes were transfected into HeLa cells and treated (+CHX) or not (-CHX) with cycloheximide. RNA was retrotranscribed and cDNA was amplified with primers RTPSAD-FW and RTBR2_ex23-RV. A) An agarose gel electrophoresis is shown on the left where the expected canonical transcript is indicated by an arrow and asterisks indicate novel transcripts or significant changes in transcript amounts after NMD inhibition. Electropherograms of fluorescent RT-PCR products of variant c.8948_8953+5del are shown on the right. B) Effect of NMD on the relative proportions of each transcript with (grey bars) and without cycloheximide (white bars). Values are averages of three independent experiments. The size of each transcript is indicated between parentheses.

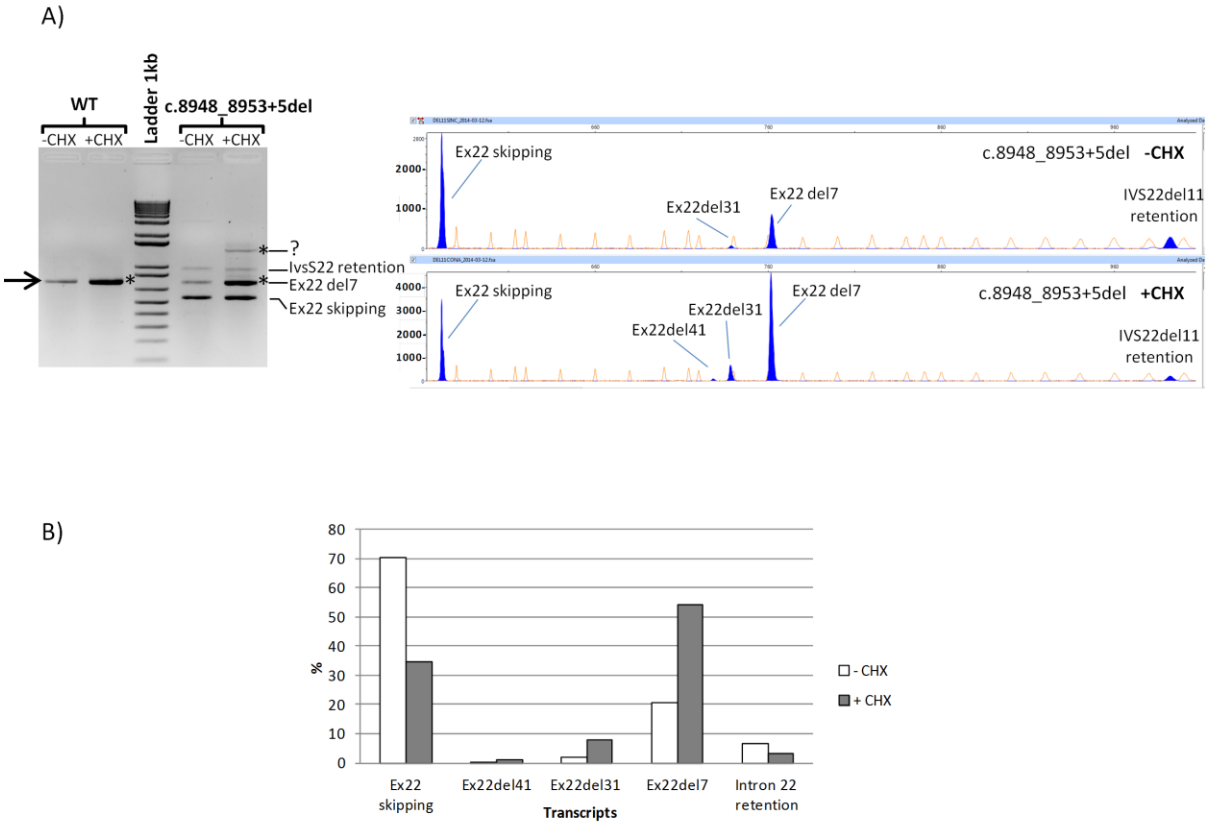


Figure 3

Figure 4. Mapping of splicing regulatory sequences by functional assays of exonic microdeletions of exons 19 (A) and 20 (B). Deletions are designated according to the HGVS nomenclature. Arrows indicate possible abnormal transcripts.

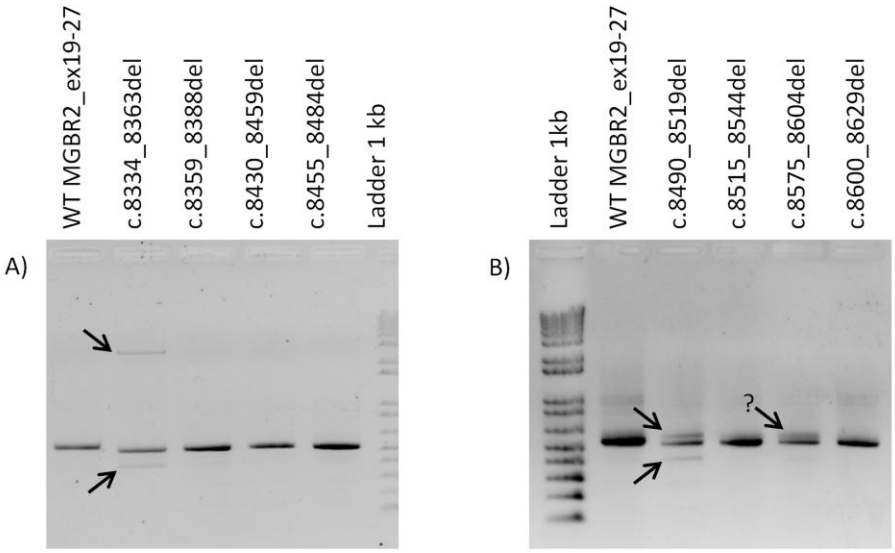


Figure 4

Table1. Primers for amplification and cloning of fragments containing exons 19 to 27.

Primers	Exons	Sequence 5′→3′
Mgbr2_ex19-fw[1]	19-24	5′ CACACACTCGAGCACACATCCGGAATAGCATTAAGAACTTGTAGCA 3′
Mgbr2_ex24-rv[1]		5′ CACACAGGATCCCACACAGGTACCAAATCAGAGGTTCAAAGAGGCT 3′
Mgbr2_ex25-fw[2]	25	5′ gtaagcctctttgaacctctgatttAATCTGTACTCCTGTTAGCAAT 3′
Mgbr2_ex25-rv[2]		5′ caaaagggaacaaaagctgggtaccTCAGTGTCTCTTATCTGGGATT 3′
Mgbr2_ex26-fw[1]	26	5′ CACACAGGTACCAGTAAATAGAGCTAGGACTTGA 3′
Mgbr2_ex26-rv[1]		5′ CACACAGGTACCTTGATTTGCATCTACTGTGATT 3′
Mgbr2_ex27-fw[2]	27	5′ ccagcttttgttccctttgtgagggtCTTCCTTCCTTTTCATGTCATTT 3′
Mgbr2_ex27-rv[2]		5′ atctgtctctgtctctctctccaccttAATTTCTTTTCTCATTGTGCAAC 3′

[1] Cloning sites are underlined.

[2] Insertion tails for overlap extension PCR cloning are in lower case.

Table 2. Bioinformatics analysis and splicing outcome of DNA variants from exons 19 to 27 with impact on splicing.

DNA variant[a]	Bioinformatics (wt→mutant scores)[b]	Splicing outcome	RNA effect[c]	Protein Effect[c]
c.8486A>T	Alteration of the canonical donor site/ESE Disruption NNS: 0.95→0.87/ HSF: 88.86→84.11 / MaxEnt: 9.46 →6.69 [-] SF2/ASF (77.23)	Exon 19 skipping (80%) Canonical transcript (20%)	r.[=;8332_8487del]	p.[Gln2829Leu; p.Ile2778_Gln2829del]
c.8488-1G>A	Alteration of the canonical acceptor site HSF: 82.32→53.37/ MaxEnt: 3.3 → -5.44	Intron 19 retention (43%) Exon 20 skipping (28%) Exon 20 del12 (20%) Ins 23 nt ? (7%) Canonical transcript (<2%)	r.[=; 8487_8488ins8487+1_8488-1; 8488_8632del;8488_8499del,?]	p.[=;Gln2829fs*2;Trp2830Lysfs*13;Trp2830_Lys2833del;?]
c.8633-24_8634 del	Alteration of the canonical acceptor site NNS: 0.99→<0.4 / HSF: 82.1→38.03/ MaxEnt: 6.97→-17.62	Exon 21 skipping (100%)	r.8633_8754del	p.Glu2878Glyfs*5
c.8754G>A	Alteration of the canonical donor site NNS: 0.98→<0.4 / HSF: 94.75→84.17/ MaxEnt: 7.66→3.22	Ivs21- ins46 (100%)	r.8754_8754+46ins	p.Gly2919Valfs*4
c.8754+4A>G	Alteration of the canonical donor site : NNS: 0.98→<0.4 / HSF: 94.75→86.41/ MaxEnt: 7.66→4.3	Ivs21- ins46 (100%)	r.8754_8754+46ins	p.Gly2919Valfs*4
c.8754+5G>A	Alteration of the canonical donor site: NNS: 0.98 →0.43/HSF: 94.75→82.58/MaxEnt: 7.66→2.91	Ivs21- ins46 (100%)	r.8754_8754+46ins	p.Gly2919Valfs*4
c.8754+5G>T	Alteration of the canonical donor site NNS: 0.98 →0.42/ HSF: 94.75→82.43/MaxEnt: 7.66→3.13	Ivs21- ins46 (100%)	r.8754_8754+46ins	p.Gly2919Valfs*4
c.8755-1G>A	Alteration of the canonical acceptor site HSF: 81.68→52.73/MaxEnt: 3.78→-4.76	Exon 22 del1 (83%) Exon 22 skipping (17%)	r.[8755del;8755_8953del]	p.[Gly2919Valfs*8;Gly2919Leufs*3]
c.8948_8953 +5del	Alteration of the canonical donor site. New donor site NNS: 1.0→0.94 (new)/HSF: 100→74.99 (new)/MaxEnt: 10.86→-0.48	Exon 22 skipping (52%) Exon 22-del7 (23%) Intron 22 retention (25%)	r.[8755_8953del;8947_8953del; 8947_8953+5delins8953+6_8954-1]	p.[Gly2919Leufs*3;Asp2983Leufs*3;Asp2983Valfs*29]
c.8953+1G>T	Alteration of the canonical donor site. New donor site NNS: 1.0→0.79 (new)/HSF: 100→73.16 (new)/MaxEnt: 10.86→2.35	Exon 22 skipping (44%) Intron 22 retention (39%) Ivs22 – ins4 (17%)	r.[8755_8953del;8953+1_8954-1_ ins; 8953+1_8953+4ins]	p.[Gly2919Leufs*3;Val2985Glyfs*15; Val2985Glyfs*34]
c.9117G>A	Alteration of the canonical donor site. NNS: 0.57 →< 0.4/HSF: 71.95→ 61.37/MaxEnt: 4.28→-4.93	Exon 23 skipping (100%)	r.[8954_9117del]	p.[Val2985Glyfs*3]
c.9256+1G>A	Alteration of the canonical donor site. NNS: 0.95 →< 0.4/HSF: 85.75 →58.91/MaxEnt: 9.43→1.25	Exon 24 skipping (74%) Exon 24 del43 (26%)	r.[9118_9256del;9214_9256del]	p.[Val3040Aspfs*18;Val3072Aspfs*18]
c.9257-1G>C	Alteration of the canonical acceptor site. NNS: 0.98→< 0.4/HSF: 87.14 →58.19/MaxEnt: 13.32→5.26	Exon 25del27 (100%)	r.[9257_9283del]	p.[Gly3086_Ser3094del]
c.9501+3A>T	Alteration of the canonical donor site NNS: 0.99 →0.41/HSF: 96.31→91.28 /MaxEnt: 10.28→4.36	Canonical transcript (87%) Exon 25 skipping (13%)	r.[=;9257_9501del]	p.[=;Gly3086Glufs*3]
c.9502-12T>G	Alteration of the canonical acceptor site (Polypyr.tract) NNS: 0.93→0.74/HSF: 81.99→78.82 /MaxEnt: 8.35→6.72	Canonical transcript (92%) Exon 26 skipping (8%)	r.[=;9502_9648del]	p.[=;Asn3168_Leu3216del]
c.9502-2A>C	Alteration of the canonical acceptor site NNS: 0.93→<0.4/HSF: 81.99→53.04 /MaxEnt: 8.35→0.31	Ivs25 – ins70nt (57%) Exon 26 skipping (33%) Exon 26 del28 (8%) 436 nt isoform ?	r.[9501_9502ins9502-70_9502-1; 9502_9648del; 9502_9529del]	p.[Asn3168Glufs*7;Asn3168_Leu3216del;Asn3168Lysfs*40]
c.9649-2A>G	Alteration of the canonical acceptor site NNS: 0.96→<0.4/HSF: 90.97→62.02 /MaxEnt: 8.54→0.59	Exon 27 del62 (72%)	r.[9649_10072del;9649_9710del]	p.[Met3217Glufs*25;Met3217Glufs*17]

		Exon 27 del424 (28%)		
c.9698G>T	Creation of putative silencers [+] 1 Sironi Motif 2 (62.22)/1 Fas-ESS / 2 PESS (34.8/44.7)	Canonical transcript (95%) Exon 27 del424 (4%) Exon 27 del62 (<1%)	r.[=; 9649_10072del;9649_9710del]	p.[=; Met3217Glufs*25;Met3217Glufs*17]

[a] Description according to HGVS guidelines (http://www.hgvs.org/mutnomen). Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the GenBank reference sequence NM_000059.1.

[b] [+] and [-] symbols indicate creation or disruption, respectively, of splicing regulatory sequences. Cut-offs of NNSPLICE and Human Splicing Finder are indicated in Supp. Table S2.

[c] RNA and protein effects follow the nomenclature guidelines of the Human Genome Variation Society.