**Supplemental Methods**

**Combined genetic and splicing analysis of *BRCA1* c.[594-2A>C; 641A>G] highlights the relevance of naturally occurring in-frame transcripts for developing disease gene variant classification algorithms.**

de la Hoya et al.

**Splicing Analyses**

**1.1 Samples**

RNA extraction protocols from lymphoblastoid cell lines (LCLs), leukocytes (LEU) (3-6 days cultured), and biopsy samples were performed with standard methods that have been described previously ([1-4](#_ENREF_1)). In the case of peripheral blood collected in PAXgene tubes (Qiagen, Hilden, Germany), RNA extraction was performed with the PAXgene Blood RNA kit according to the manufacturer’s instructions (Qiagen, Hilden, Germany). When indicated, Puromycin (Puro) or Cycloheximide (Cyclo) was added to cell cultures (end concentration of 100µg/ml) 6-8 hours prior to cell harvest. In addition, we used commercially available total RNA (guanidium thiocyanate isolation method) from Human Mammary Gland and Human Ovary (Clontech, Mountain View, CA). More precisely, Human Mammary Gland RNA (Clontech, catalog No: 636576) was obtained from a 27 year-old Caucasian female. Human Ovary RNA (Clontech, catalog No: 636555) was pooled from 3 Caucasian females, aged 40, 55 and 51. In all cases, RNA samples were treated with RNase-Free DNase previous to Reverse Transcription (RT) reactions. For RT-PCR reactions, different contributors used different commercially available kits (SuperScriptII Reverse Transcriptase, Invitrogen, Carlsbas, CA; High Capacity RNA-to-cDNA Mater Mix, Applied Biosystems, Foster City, CA; Primescript RT reagent kit, Takara Biotechnology, Shiga, Japan; Transcriptor High Fidelity cDNA synthesis Kit, Roche Applied Science, Mannheim, Germany; One-Step RT-PCR kit, Qiagen, Hilden, Germany) in combination with oligodT and/or random hexamers. Previously, we have shown that different RNA extraction/cDNA synthesis protocols are indistinguishable for the purpose of BRCA1 RT-PCR analyses ([2](#_ENREF_2)).

**1.2 Capillary Electrophoresis**

Capillary electrophoresis analysis of alternative splicing has been extensively described previously ([2](#_ENREF_2), [3](#_ENREF_3)). In the present study, we have analyzed the *BRCA1* alternative splicing landscape at the vicinity of exon 10 with up to five different RT-PCR assays: E7-E11q.1, E8.1-E11p, E8.1-E11q.1, E8.1-E11q.2, E8.2-E11q.2, and E8.1-E12 (see Table 1 below for further details). To perform semi-quantitative analyses, we kept a low number of PCR cycles (depending on the contributing laboratory, 28, 33, or 36 cycles). While 28 or 33 cycle PCRs are optimal for semi-quantification, 36 cycle PCRs were needed to detect the rather minor alternative splicing event ▼10p. Depending on the contributing laboratory, capillary electrophoresis analyses were performed with 3100, 3130, 3130XL, 3500XL, or 3730XL ABI PRISM Genetic Analyzers (Applied Biosystems, Foster City, CA) using POP-7 filled capillary arrays of 36 or 50 cm. Size-calling and peak areas were analyzed with GeneMapper v4.0 or GenScan v3.7 software (both from Applied Biosystems, Foster City, CA), or GeneMarker 2.4 (Softgenetics LLC, State College, PA). As internal size-standard, we used LIZ-500 or LIZ-1000 (Applied Biosystems, Foster City, CA). Splicing Fraction (SF) of each individual alternative splicing event were measured in all cases as the ratio between the peak area of the individual events and the Σ of all peak areas (all transcripts) detected by the corresponding assay. For instance, Δ9,10SF is determined in a E8-E11q assay as the ratio between Δ9,10 peak areaand the Σ of Δ9+Δ10+▼10p+Δ9,10+full-length (FL) peak areas (depending on samples an PCR cycles, not all five peaks are necessarily observed in all determinations). Electropherograms with saturated peaks were not considered for SF analyses.

.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **RT-PCR assay** | **Forward primer** | | **Reverse primer (FAM-labeled)** | | **Capillary Electrophoresis Analysis of RT-PCR products**  (expected peak sizes expressed in bp) | | | | |
| mapping | seq | mapping | seq | Δ9,10 | Δ10 | Δ9 | FL | ▼10p |
| E7-E11q.1 | c.372\_c.393  (Exon 7) | CATCCAAAGTATGGGCTACAGA | c.799\_c.819  (Exon 11q) | TGGCTCCACATGCAAGTTTG | 324/327a | 368/371a | 401/404a | 445/448a | 466/469a |
| E8.1-E11p | c.459\_c.478  (Exon 8) | TGTCCAACTCTCTAACCTTG | c.759\_c.778  (Exon 11p) | TTTCTGGATGCCTCTCAGCT | 199 | 243 | 276 | 320 | 341 |
| E8.1-E11q.1 | c.459\_c.478  (Exon 8) | TGTCCAACTCTCTAACCTTG | c.799\_c.819  (Exon 11q) | TGGCTCCACATGCAAGTTTG | 240 | 284 | 317 | 361 | 382 |
| E8.2-E11q.2 | c.462\_c.971  (Exon 11) | CCAACTCTCTAACCTTGGAACTGTG | c.949\_c.971  (Exon 11q) | CTTCCAGCCCATCTGTTATGTTG | 389 | 433 | 466 | 510 | 531 |
| E8.1-E12 | c.459\_c.478  (Exon 8) | TGTCCAACTCTCTAACCTTG | c.4145\_c.4164  (Exon 12) | CTGAGAGGATAGCCCTGAGC | 276b | 320b | 353b | 397b | 418b |

**Table 1. Description of RT-PCR assays used in the present study.**

For each RT-PCR assay, we show the sequence of forward and reverse primers, mapping to the Ensembl sequence ENST00000357654 (NCBI NM\_007294.3), and the expected size of the peaks corresponding to different *BRCA1* alternative splicing events. The actual size calling may vary ±2bp with respect to the expected size due to factors such as local density of size-standard peaks, capillary array length, and/or Taq polymerase addition of a 3'-adenine overhang. a Capillary electrophoresis analysis of RT-PCR products generated with the E7-E11q.1 assay produces ±3bp doublet peaks due to alternative splicing at BRCA1 exon 8 NAGNAG splicing acceptor site ([3](#_ENREF_3), [5](#_ENREF_5)). b With protocols used in the present study, the E8.1-E12 RT-PCR assay do not generate products spanning the long BRCA1 exon 11 (3426nt). For that reason, the expected size peaks displayed in the table correspond to *BRCA1* Δ11q transcripts ([3](#_ENREF_3), [5](#_ENREF_5)): Δ9,10+Δ11q, Δ10+Δ11q , Δ9+Δ11q and ▼10p+Δ11q transcripts.

**1. 3 Digital PCR (dPCR)**

All dPCR experiments were performed in a QuantStudio 3D Digital PCR 20K platform according to the manufacturer’s instructions (Applied Biosystems, Foster City, CA). As indicated, we performed different assays combining FAM-labeled and VIC-labeled TaqMan assays. FAM-labeled assays included an Applied Biosystem pre-designed assay (Hs01556198) specific for the E8/E11 junction (Δ9,10 assay), and a custom designed TaqMan assay specific for the *BRCA1* E9/E10 junction (FL assay). The FL assay was designed with the Applied Biosystem proprietary on-line pipeline. VIC-labeled assays included an Applied Biosystem pre-designed assay (Hs01556193) specific for *BRCA1* E23/24 junction (a proxy for overall *BRCA1* expression), and a pre-designed assay (Hs00609073) specific for *BRCA2* E26/27 junction (a proxy for overall *BRCA2* expression). dPCR chips were analyzed in the cloud-based QuantStudio 3D Analysis Suit v2.0 (Applied Biosystem, Foster City, CA) to review quality (only green and yellow flag chips were considered for further analyses), and calculate copies/µl of FAM and VIC molecules (and the precision of these measures). Default settings were used in all cases. Subsequently, data was exported to an excel file to calculate the FAM/VIC ratio. To measure Δ9,10SF, we performed experiments combining Δ9,10 and Hs01556193 assays. To measure FLSF we performed experiments combining FL and Hs01556193 assays. To measure Δ9,10 and FL relative expression levels, we performed experiments combining Δ9,10 (or FL) assays with Hs00609073. In relative expression level experiments, the Δ9,10 (or FL) relative expression level of each sample was normalized to the average Δ9,10 (or FL) relative expression level as measured in control samples.

**1.4 Absolute quantification of alternative splicing events by real-time PCR (qPCR)**

In order to specifically amplify *BRCA1* full-length transcripts (transcripts containing exons 9,10, and 11q), transcripts containing six naturally occurring alternative splicing events (∆9, Δ9,10, Δ9\_11, ∆10, ∆11, and ∆11q), or transcripts combining two splicing events (Δ9,10+∆11q), we designed various primers in specific exonic regions, or overlapping specific exon/exon junctions. Specificity of primer combinations (we tested 3 primer sets peralternative splicing event ) were determined by temperature gradient PCR. Pooled cDNA from healthy controls was used as template to amplify all specific alternative splicing event fragments, with the single exception of ∆10 fragments, which were obtained using variant carrier cDNA as template. Cycling conditions were 95°C for 2mins followed by 35 cycles of 95°C for 30secs, gradient annealing temperature (annealing temperature of 55°C - 65°C at intervals of 2°C) for 30secs, 72°C for 60secs and a final extension time of 7mins. Gel electrophoresis was used to visualize the fragments. PCR products amplified using a primer combination and lowest annealing temperature for which only one PCR product could be observed were considered optimal (see table at the end of this section) and selected for cloning. PCR products were cloned using the pGEM-T Vector System (Promega, Auburn, Victoria, Australia). Recombinant clones were selected from a single colony and sequence confirmed. Plasmid preparations containing the PCR products for each of the eight fragments were quantified using a Qubit 3.0 Fluorometer (Aplied Biosystems, Foster City, CA) and a serial dilution was made. Dilutions were used as template for deriving a standard curve for quantitative PCR (qPCR).

A standard curve was determined using pGEM-T clones carrying the eight isoforms and the specific primers/annealing temperature for each transcript. Real-time PCR reactions were carried out in a Lightcycler 480 (Roche, Castle Hill, NSW, Australia) using Platinum SYBR Green qPCR SuperMix-UDG (Invitrogen). Cycling conditions were: 50°C for 2mins, 45 cycles of 95°C for 2mins, 95°C for 20secs, optimized annealing temp (see table at the end of this section)15secs, 72°C for 20secs. Normalized expression values (using GAPDH as an internal reference) were obtained using the Lightcycler 480 Gene Scanning software for cDNA isolated from the variant carrier and a set of 11 non-variant carrying controls. Crossing point (CP) values were plotted against the standard curve. The number of molecules of the 3.2Kb plasmid was estimated at 2.9x108/ng. Given that there are approximately 290 million molecules in 1ng plasmid, the starting number of molecules in each reaction could be estimated enabling a comparison within a sample and across the controls.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Forward** | | **Reverse** | | **product (bp)** | **Annealing Temp (Cº)** |
| **Full-length** | ex10F | CTCAAGGAACCAGGGATGAA | ex11pR | ACTGGGTTGATGATGTTCAGT | 101 | 63 |
| **∆11** | ex10/12F | TGGATTCTGCAAAAAAGGGTGAA | ex12R | CTGAGTGGTTAAAATGTCACTCTGA | 107 | 61 |
| **∆11q** | ex11p/12F | ATCCAGAAAAGTATCAGGGTGAA | ex12R | CTGAGTGGTTAAAATGTCACTCTGA | 107 | 59 |
| **∆9\_11** | ex8/12F | TGTCTACATTGAATTGGGTGAAGCA | ex12R | CTGAGTGGTTAAAATGTCACTCTGA | 106 | 63 |
| **∆9,10** | ex8/11F | TGTCTACATTGAATTGGCTGCTTGT | ex11pR | GCACGCTTCTCAGTGGTGT | 101 | 65 |
| **∆10** | ex9/11F | AATAAGGCAACTTATTGCAGCTGCTTGT | ex11pR | GCACGCTTCTCAGTGGTGT | 104 | 61 |
| **∆9** | ex8/10F | GTCTGTCTACATTGAATTGGTGTGG | ex10/11R | CTCAGAAAATTCACAAGCAGCCTTT | 113 | 61 |
| **∆9,10+∆11q** | ex8/11pF | TCTGTCTACATTGAATTGGCTGCTT | ex11p/12R | AGATGCTGCTTCACCCTGAT | 146 | 61 |

**Table 2. Optimal primer sequences and annealing temperatures for absolute quantification of alternative splicing eventsby qPCR**

**1.5 RNAseq experiments.**

The TruSeq targeted RNA expression kit (Illumina, San Diego, CA) was used to target exon regions across *BRCA1*. Each probe pair targets a specific splice junction or coding SNP. Coding SNP probes are situated within the same exon, either side of the variant of interest, whereas splice junction probes are positioned near the 3’ and 5’ end of two adjacent exons, respectively. This design allows for the detection of certain alternative splicing events, and small splice junction aberrations, as the probes do not need to bind specifically with the other in their original pair. Any exon skipping event has the potential to be detected providing there is an upstream and a downstream probe flanking the deletion. Briefly, TruSeq Targeted RNA Expression chemistry involves pre-designed oligonucleotide probes that hybridise to the target *BRCA1* cDNA region followed by an extension-ligation reaction then takes place to connect the probes, and an amplification step to create the template strand. This is PCR amplified to add indices prior to sequencing. Sequencing was performed on Illuminas MiSeq platform. All *BRCA1* exons had predesigned probes situated on either end to allow detection of non-aberrant mRNA splice junctions, with the single exception of *BRCA1* 16-17. In addition, one probe in the *BRCA1* 3-5 pair spans the splice site inhibiting detection of splicing events involving this junction. TruSeq targeted RNA expression kit allowed us to also quantify each detected splice junction. After normalizing the read counts, the expression of each splicing event was compared across samples to determine expression differences. All samples were sequenced with and without treatment of Cycloheximide.

Splice junction *BRCA1* 2-3 was used as the full length reference transcript. We assumed that this junction is present in all alternative splicing events that don't overlap this junction. We also assumed that the alternative events do not co-occur. Under these assumptions we subtracted all alternative splicing reads, that didn't overlap the exon 2-3 junction, from the total 2-3 junction reads. Some junctions were exempt from this as they are common NAGNAG events (8p, 13p, 14p), which are likely to be present in the full length transcript. Δ9,10 was also excluded as it returned questionable read depths.

The resulting 2-3 read depth, together with the sum of all alternative events(excluding those mentioned above), gave the total expression, from which the proportions of each alternative event were determined for each sample.The resulting proportions were back transformed (95% CI) for the log data. This provided the standard deviation and mean for the expression of the control samples for each junction, onto which the variant sample relative expression was overlayed. Delta 10 and FL data was extracted for the figures

**1.6 RNA interference experiments.**

All small interference RNAs (siRNAs) used represent validated sequences in several previous publications from our and other labs ([6](#_ENREF_6)). 25 nM final concentration of A1 (hnRNP A1) siRNA (5’-CAGCUGAGGAAGCUCUUCAdTdT-3’), Tra2β siRNA (5’-GCAUGAAGACUUUCUGAAAdTdT-3’), SF2 (SRSF1) siRNA (5’-CCAAGGACAUUGAGGACGUdTdT-3’), or SC35 (SRSF2) siRNA (5’-AAUCCAGGUCGCGAUCGAAdTdT-3’) were transfected into MDAMD231 cells with INTERFERin (Polyplus transfections, Illkirch, France) following the manufacturer’s instructions. As a control, we transfected a scrambled sequence (the luciferase siRNA 5’-CGUACGCGGAAUACUUCGAdTdT-3). Forty-eight hours later, the RNA was extracted using RNAeasy plus kit (Qiagen, Hilden, Germany). The overall strategy has been described previously ([6](#_ENREF_6)).

**1.7 In silico predictions of variant-induced alterations in Exonic Splicing Regulatory sequences (ESRseq)**

We analyzed the potential impact of *BRCA1* c.641A>G on RNA splicing by using an *in silico* approach based on the calculation of total ESRseq score changes (ΔtESRseq) as described previously ([7](#_ENREF_7)).

**1.8 Splicing Reporter Minigene Experiments**

*pCAS2-BRCA1-exon 10 minigene assay*

The pCAS2 splicing vector has been previously described ([8](#_ENREF_8)). The pCAS2-*BRCA1*-exon-10 minigene constructs (Figue 3A) were generated as follows. First, the wild-type genomic segment *BRCA1* c.594-147\_c.670+173 was amplified by PCR from genomic DNA using forward primer BR1-10-BamHI-F (GACCGGATCCCTAAAGGAGAGAG) and reverse primer BR1-10-MluI-R (GACCACGCGTTTTAAATCTATCAG), carrying 5’ tails with *BamHI* and *MluI* restriction sites, respectively (underlined). PCR-amplified genomic segments encompassed *BRCA1* exon 10 (77 bp) and part of the 5’ and 3’ flanking intronic sequences (147 bp and 173 bp, respectively). After digestion with *BamHI* and *MluI*, the PCR products were inserted into the *BamHI* and *MluI* cloning sites of pCAS2, a two-exon splicing reporter vector, in order to produce the three-exon pCAS2-BRCA1-exon10-WT minigene. Then, variants of interest were introduced into this construct by site-directed mutagenesis using the two-stage overlap extension PCR method ([9](#_ENREF_9)). The inserts of the mutant minigenes were sequenced to ensure that no unwanted mutations were introduced during amplification or cloning.

Wild-type and variant minigene constructs were transiently transfected in parallel into HeLa cells using the FuGENE 6 transfection reagent, according to manufacturer’s instructions (Roche Applied Science, Mannheim, Germany). Transfections were performed in 12-well plates by using 6x104 cells/well (at ~60% confluence) and 400 ng of each minigene construct. Cells were then collected 24 h post-transfection. Total RNA was extracted using the NucleoSpin RNA II kit (Macherey Nagel, Düren, Germany), according to the manufacturer’s instructions, including a DNase treatment. The RT-PCR reactions were performed in a 25 µl reaction volume by using the OneStep RT-PCR kit (Qiagen, Hilden, Germany), and 200 ng RNA as template. Reactions were performed using the vector-specific forward primer pCAS-KO1F (5’-TGACGTCGCCGCCCATCAC-3’) and the reverse primer pCAS2R (5’-ATTGGTTGTTGAGTTGGTTGTC-3’), with 30 cycles of amplification. RT-PCR products were separated by electrophoresis on 2% agarose gels containing ethidium bromide and visualised by exposure to ultraviolet light under conditions of non-saturating exposure. RT-PCR products were gel-purified and fully sequenced to determine their identity.

*pB1 minigene assay*

The pB1 minigene splicing reporter vector has been previously described ([10](#_ENREF_10)). This construct has a pcDNA3(+) backbone and contains the exon 1 of α-globin followed by *BRCA1* exons 8 to 12 (and part of their flanking intronic regions) in under the control of the CMV promoter. Mutant pB1 minigenes were prepared by site-directed mutagenesis through a two-step PCR overlap extension method ([9](#_ENREF_9)). After digestion with XhoI and HindIII, the *BRCA1* fragments (c.594-174\_c.670+345) containing the variants of interest were inserted into pB1 *in lieu* of the corresponding WT sequence. Then, the inserts of the mutant minigenes were sequenced to ensure that no unwanted mutations were introduced during amplification or cloning. Wild-type and variant minigene constructs were transiently transfected into HeLa, MCF7, HBL100, and IgrOV1cells Transfection, RNA extraction and RT-PCR analysis were performed as described above for pCAS2-BRCA1-exon10, with the exception that the RT-PCR primers used here were: pB1-V.8(2)-F (,5’-GAGGCCCTGGAGAGGACA-3‘,a vector-specific forward primer) and RT.1.8-11-R (, 5’-ACGCTTCTCAGTGGTGTTCA-3’, a reverse primer on *BRCA1* exon 11).

**References**

1 Vreeswijk, M.P. and van der Klift, H.M. (2012) Analysis and interpretation of RNA splicing alterations in genes involved in genetic disorders. *Methods Mol Biol*, **867**, 49-63.

2 Whiley, P.J., de la Hoya, M., Thomassen, M., Becker, A., Brandao, R., Pedersen, I.S., Montagna, M., Menendez, M., Quiles, F., Gutierrez-Enriquez, S. *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.

3 Colombo, M., Blok, M.J., Whiley, P., Santamarina, M., Gutierrez-Enriquez, S., Romero, A., Garre, P., Becker, A., Smith, L.D., De Vecchi, G. *et al.* (2014) Comprehensive annotation of splice junctions supports pervasive alternative splicing at the BRCA1 locus: a report from the ENIGMA consortium. *Hum Mol Genet*, **23**, 3666-3680.

4 Romero, A., Martin, M., Cheang, M.C., Lopez Garcia-Asenjo, J.A., Oliva, B., He, X., de la Hoya, M., Garcia Saenz, J.A., Arroyo Fernandez, M., Diaz Rubio, E. *et al.* (2011) Assessment of Topoisomerase II alpha status in breast cancer by quantitative PCR, gene expression microarrays, immunohistochemistry, and fluorescence in situ hybridization. *Am J Pathol*, **178**, 1453-1460.

5 Dosil, V., Tosar, A., Canadas, C., Perez-Segura, P., Diaz-Rubio, E., Caldes, T. and de la Hoya, M. (2010) Alternative splicing and molecular characterization of splice site variants: BRCA1 c.591C>T as a case study. *Clin Chem*, **56**, 53-61.

6 Raponi, M., Smith, L.D., Silipo, M., Stuani, C., Buratti, E. and Baralle, D. (2014) BRCA1 exon 11 a model of long exon splicing regulation. *RNA Biol*, **11**, 351-359.

7 Di Giacomo, D., Gaildrat, P., Abuli, A., Abdat, J., Frebourg, T., Tosi, M. and 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. *Hum Mutat*, **34**, 1547-1557.

8 Gaildrat, P., Krieger, S., Di Giacomo, D., Abdat, J., Revillion, F., Caputo, S., Vaur, D., Jamard, E., Bohers, E., Ledemeney, D. *et al.* (2012) Multiple sequence variants of BRCA2 exon 7 alter splicing regulation. *J Med Genet*, **49**, 609-617.

9 Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. and Pease, L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, **77**, 51-59.

10 Raponi, M., Douglas, A.G., Tammaro, C., Wilson, D.I. and Baralle, D. (2012) Evolutionary constraint helps unmask a splicing regulatory region in BRCA1 exon 11. *PLoS One*, **7**, e37255.