Supplementary Figures for
Clinical, splicing and functional analysis to classify *BRCA2* exon 3 variants: application of a points-based ACMG/AMP approach

Figure S1: Mouse embryonic stem cell model used for the functional analysis of *BRCA2* variants.

Figure S2: In silico splicing predictions for the *BRCA2* exon 3 consensus acceptor site

Figure S3: Results from minigene assays. See Legends of sub-figures for more details.

Figure S4: Results from mRNA assays of patient material. See Legends of sub-figures for more details.

Figure S5: Quantitative dPCR results of the two known isoforms other than ΔE3 in patient samples.

Figure S6: RNASeq data of patient samples

Figure S7: Results from dPCR analysis of mESC

Figure S8: Capillary electrophoresis of mESC RT-PCR products for *BRCA2* c.68-2A>G

Figure S9: Cell viability of BRCA2 exon 3 variants in mESC

Figure S10: Western blot analysis of mESC

Figure S1: Mouse embryonic stem cell model

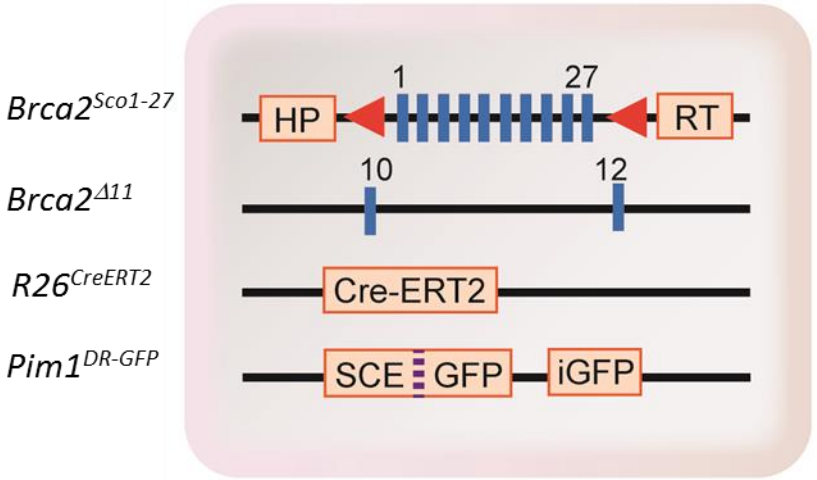
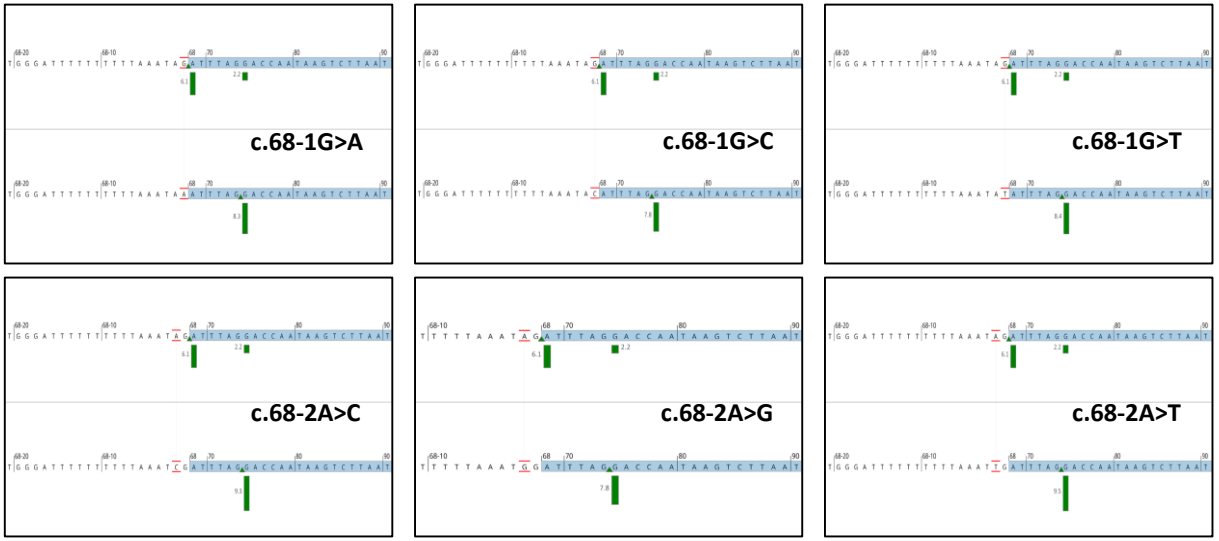


Figure S1: Mouse embryonic stem cell model used for the functional analysis of *BRCA2* variants. The Brca2 -/loxP mES cell line contains a conditional Brca2 allele and a disrupted Brca2 allele. The DR-GFP construct, containing two differently mutated inactive GFP genes, SceGFP and iGFP, was integrated at the Pim-1 locus. The Cre-ERT2 construct, integrated at the Rosa26 locus encodes for a Cre-recombinase fused to a mutated ligand-binding domain (LBD) of the estrogen receptor. Binding of 4-OHT to the LBD will results in translocation of the Cre-recombinase into the nucleus where it mediates removal of the conditional mBrca2 allele.

Figure S2: *In silico* splicing predictions for the *BRCA2* exon 3 consensus acceptor site



BRCA2 variants	SpliceAI scores (pre-mRNA position)				SpliceAI read-outs	
	AL	DL	AG	DG		
c.68-2A>T	0.97(+2)	0.00	0.98(+8)	0.02(+250)	**Δ(E3p6)**	Δ(E3)↓
c.68-2A>G	0.97(+2)	0.03(+250)	0.95(+8)	0.00	**Δ(E3p6)**	Δ(E3)↑
c.68-2A>C	0.97(+2)	0.00	0.97(+8)	0.01(+250)	**Δ(E3p6)**	Δ(E3)↓
c.68-1G>A	0.97(+2)	0.00	0.97(+8)	0.01(+249)	**Δ(E3p6)**	Δ(E3)↓
c.68-1G>T	0.97(+2)	0.00	0.97(+8)	0.02(+249)	**Δ(E3p6)**	Δ(E3)↓
c.68-1G>C	0.97(+2)	0.00	0.97(+8)	0.01(+249)	**Δ(E3p6)**	Δ(E3)↓

Figure S2. *In silico* splicing predictions for all six possible variants targeting the *BRCA2* exon 3 consensus acceptor site support ΔE3p6 (r.68_73del) as the major outcome.

MES splicing predictions (top) were performed with the analytical/visualization software AlamutVisualPlusv1.4 (Sophia Genetics) with default settings. Predictions are very similar for all six variants; the native acceptor site (score 6.1) is lost, and a cryptic acceptor site (score 2.2) is activated (scores ranging from 7.8 to 9.5), predicting ΔE3p6.

SpliceAI predictions (bottom) were performed online using the SpliceAI API (https://spliceailookup.broadinstitute.org) with the following parameters; genome version hg38, score type raw, and max distance 10000. AL, acceptor loss; DL, donor loss; AG, acceptor gain; DG, donor gain. Δ score (and pre-mRNA position) are shown. Scores >0.8 (high precision) appear in red. Scores <0.2 (below threshold for high recall) appear in grey. Predictions are very similar for all six variants; the acceptor site is lost (high precision), an alternative acceptor site is gained (high precision), and the native exon 3 donor site is neither damaged nor improved. Combined, these scores support ΔE3p6 (but not ΔE3) as the splicing outcome caused by these six variants. Δ scores below threshold are not significant, but might be compatible with a slight upregulation of the alternative splicing isoform ΔE3 for c.68-2A>G (exon 3 donor site is slightly weakened), and a similar downregulation of ΔE3 for the remaining variant (exon 3 donor site is slightly improved). Note that BRCA2 exon 3 is 249nt long.

Figure S3 : Results from minigene-assays

See following pages for Sub-figures S3a, S3b, S3c

Figure S3a

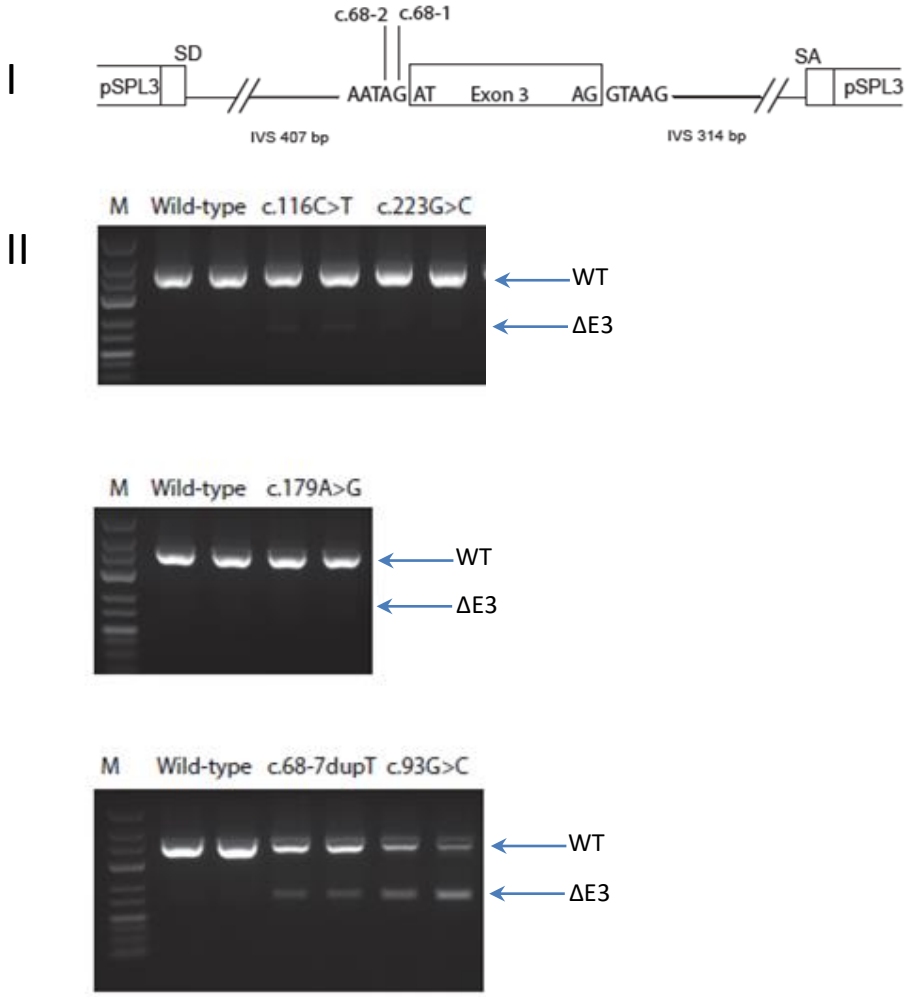


Figure S3a. Minigene assay of selected variants. I: construct used for minigene assay. II Minigene results: agarose gel of minigene products. Three different runs are shown in the three panels including wild-type controls and examined variants in duplicate. Upper band (WT) represents the wild-type full length transcript. The lower band represents ΔE3. M: Molecular marker

Figure S3b

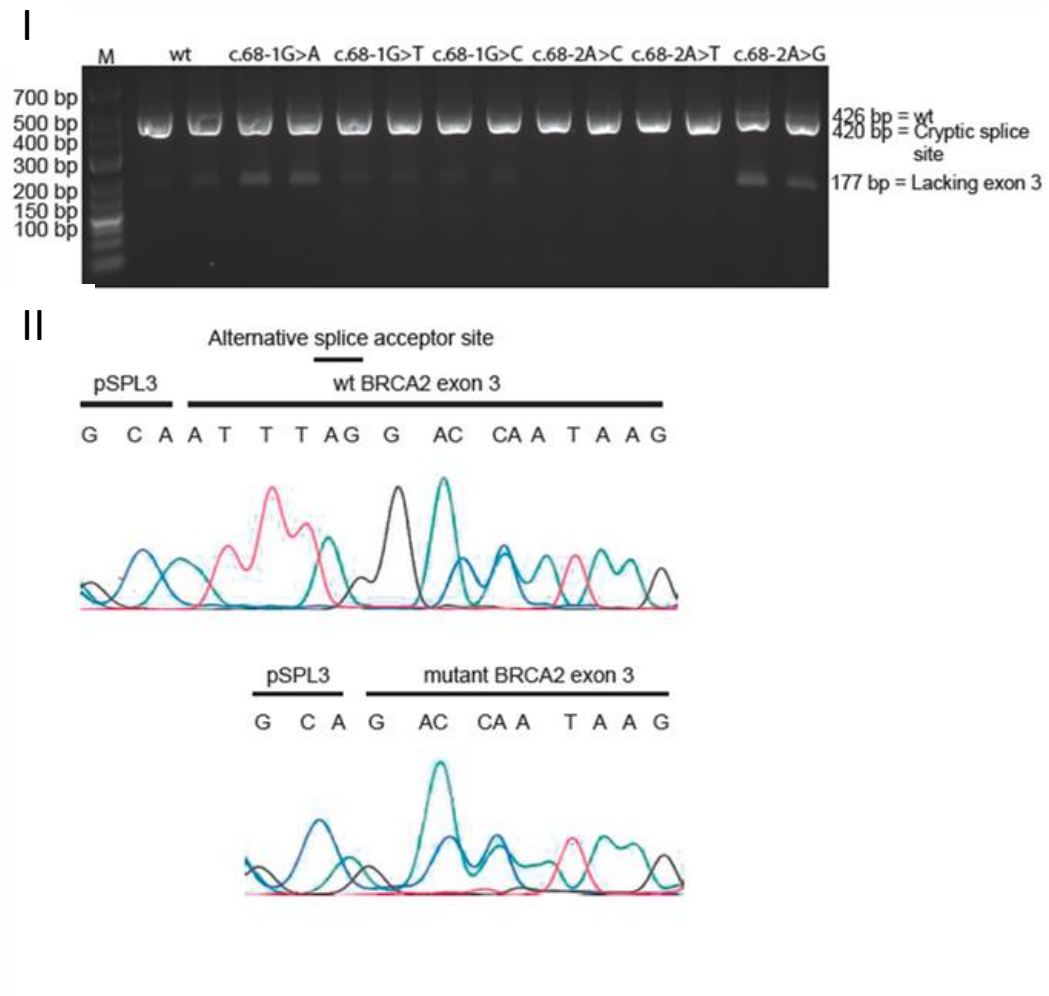


Figure S3b. Minigene assay of c.68-1 and c.68-2 variants. I Minigene results: agarose gel of minigene products. II: Sanger sequencing was performed on the large fragments excised from gel. This showed use of c.74 for all 6 variants. A representative example is shown. The wt acceptor is expected to be completely inactivated. Since we did not observe a major size difference with the wildtype fragment we assumed that this fragment was derived from the use of cryptic splice site that would generate a 6 bp smaller fragment, an invisible difference on the agarose gel. Indeed sequencing in II confirms that an alternative donor in c.74 is used, as also supported by high MaxEnt prediction (Supplementary Table 1). The band at 177bp is corresponding to ΔE3. Mw: Molecular marker with indicated band sizes.

Figure S3c

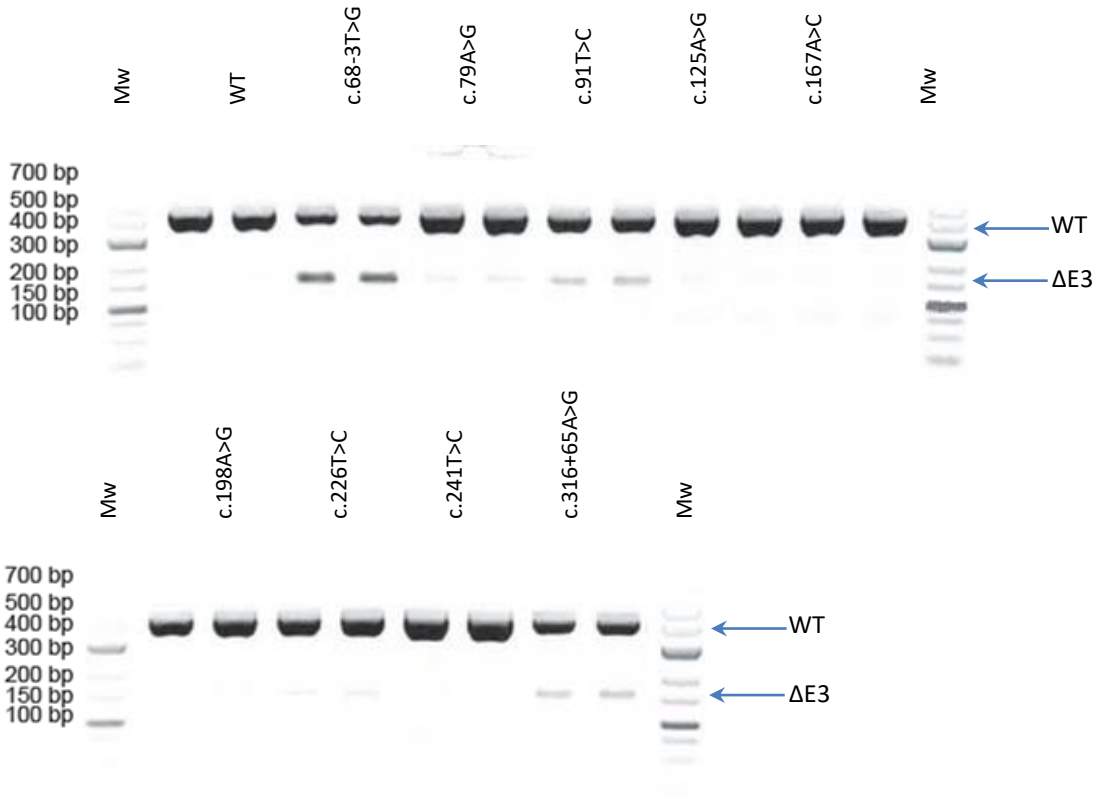
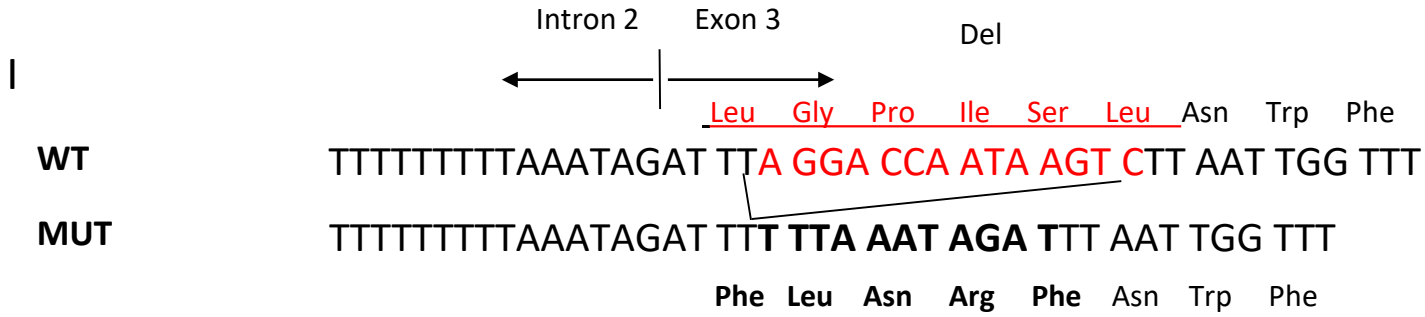


Figure S3c. Minigene assay of selected variants. Minigene assay was performed using pSPL3 vector with exon 3 as illustrated in S3a. Wild type control (WT) and variants are tested in duplicate. The position of ΔE3 and wildtype (WT) bands is shown. Mw: Molecular marker with indicated band sizes.

Figure S4: Results from mRNA assays of patient material.

See following pages for Sub-figures S4a to S4j

Figure S4a



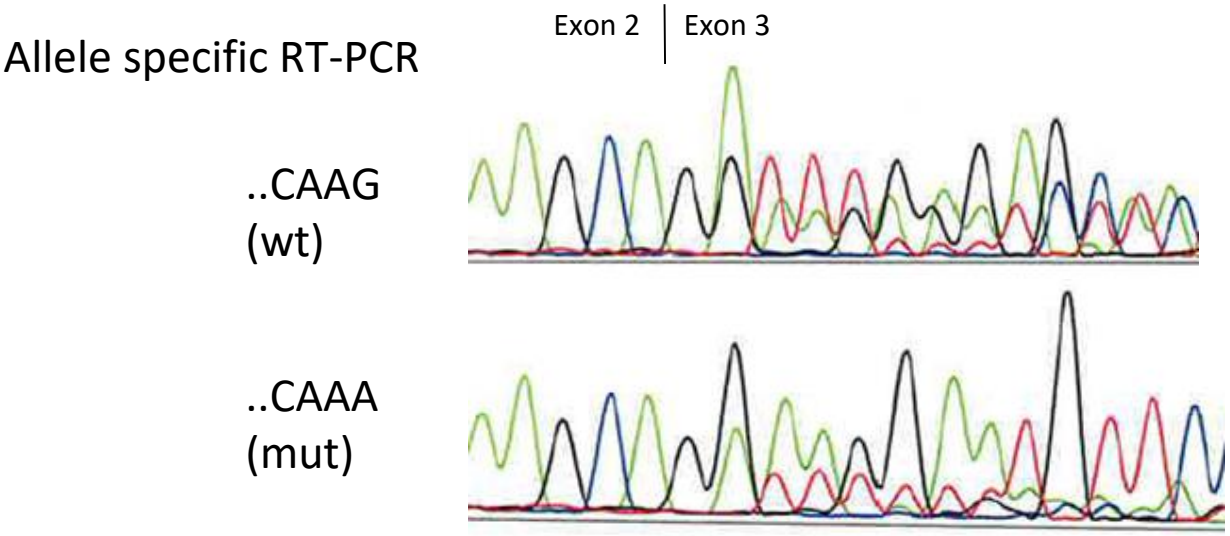


Figure S4a. Splice assay of c.72_85delinsTTTAAATAGAT variant. I: schematic figure showing wild type allele and mutant allele. The variant includes a deletion of 14 bases (red letters) and insertion of 11 bases (in bold). The affected amino acids are shown with the same color coding. II: Sanger sequencing of RT-PCR products. RT-PCR was performed using allele-specific forward primers recognizing the G and A alleles of c.-26G>A located in UTR in exon 1. This common variant is used here solely for allelle discrimination by allelle specific priming. The upper panel shows the result from the G allele. Wild type transcript and natural ΔE3 is identified from this allele. For the A-allele the c.72_85delinsTTTAAATAGAT transcript and increased ΔE3 is observed.

Figure S4b

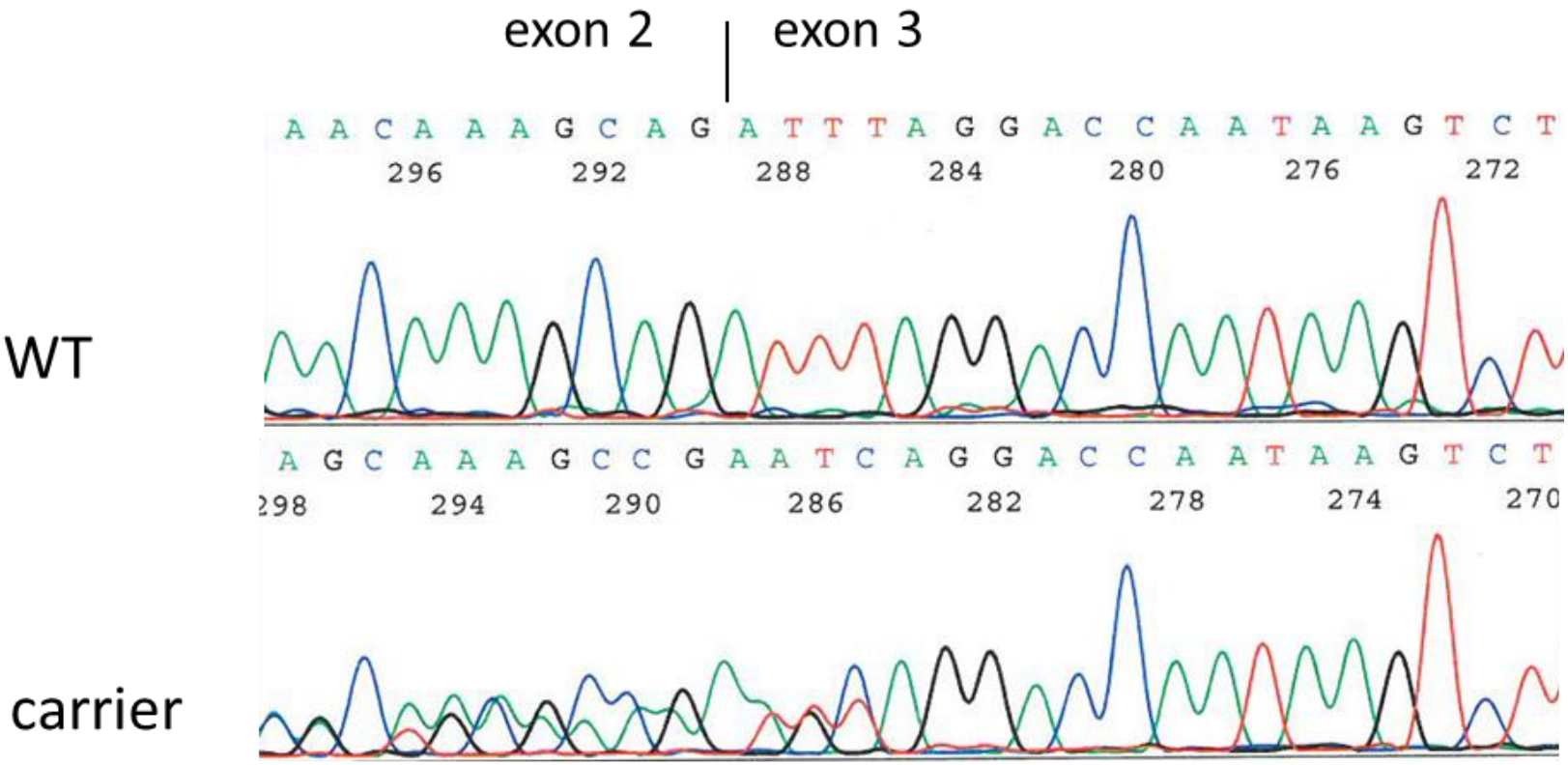


Figure S4b. Splice assay of c.68-2A>G variant. Sanger sequencing of RT-PCR products using PCR primers in exon 2 and 4 using cDNA from Paxgene RNA as template. Sequencing is performed with reverse primer. ΔE3 is not identified in this part of the sequence because the product is shorter. The upper panel shows result from non-carrier (WT). The lower panel shows result from a c.68-2A>G carrier; in addition to the WT transcript, an isoform lacking 6 bases corresponding to 2 amino acids is observed.

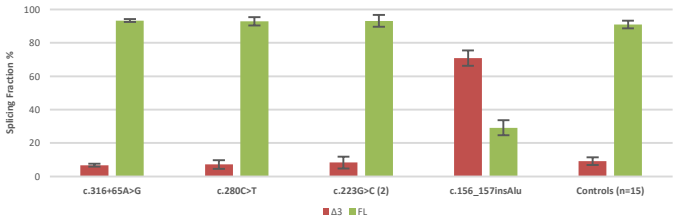
Figure S4c

Figure S4c. Capillary electrophoresis analysis of RT-PCR products from LCL. RNA was harvested from peripheral blood of patients that harbor the c.316+65A>G, c.280C>T, c.156_157insAlu and two samples with c.223G>C. RNA from 15 controls individuals were also analysed. Data represent the mean of at least two independent PCRs, and provide a semi-quantitative analysis of transcript proportions. The relative proportion of the wild type and ΔE3 transcripts, as determined from the area under the curve for the peaks, is expressed as a percentage. I: summary of fragment analysis. II: pseudo gel image generated from capillary electrophoresis (ABI 3730, Life Technology). III: Electropherograms for variant carriers and controls (non-carriers). The position of ΔE3 and wildtype (WT) bands/peaks are shown

I

Summary fragment analysis:

	delta3		FL	
Variant	x̄	σ	x̄	σ
c.316+65A>G	6,7431881	0,8883811	93,256812	0,8883811
c.280C>T	7,1901945	2,5436402	92,809805	2,5436402
c.223G>C (N=2)	8,3481218	3,5218101	93,098636	3,5218101
c.156_157insAlu	70,835856	4,5329644	29,164144	4,5329644
Controls (N=15)	9,1273063	2,3110675	90,872694	2,3110675



II

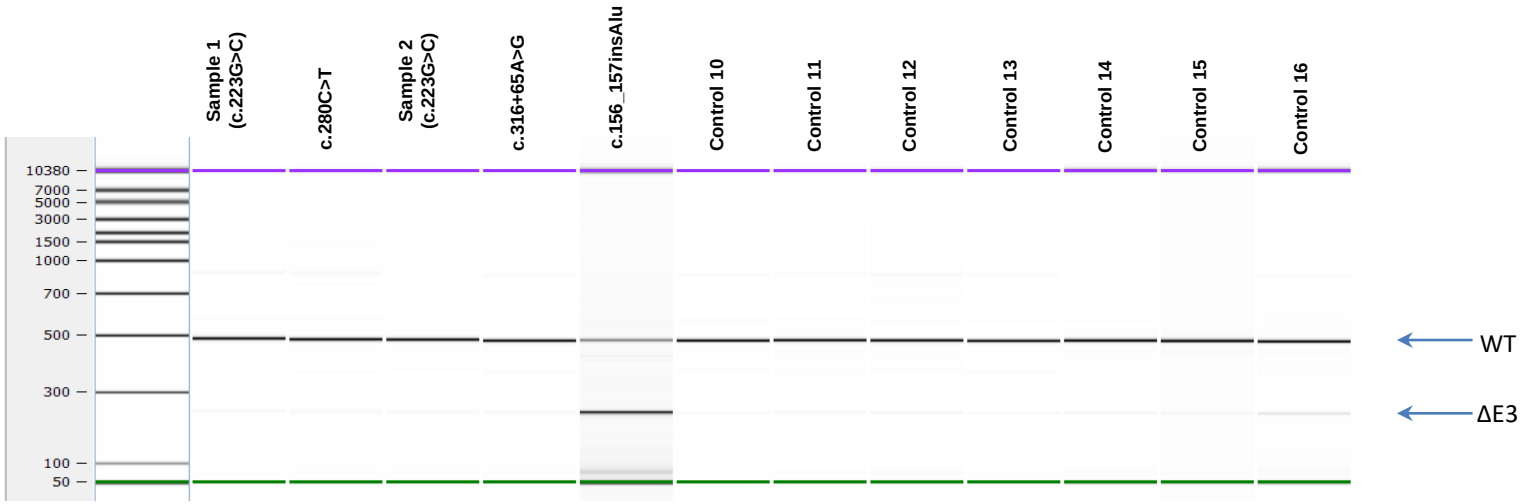
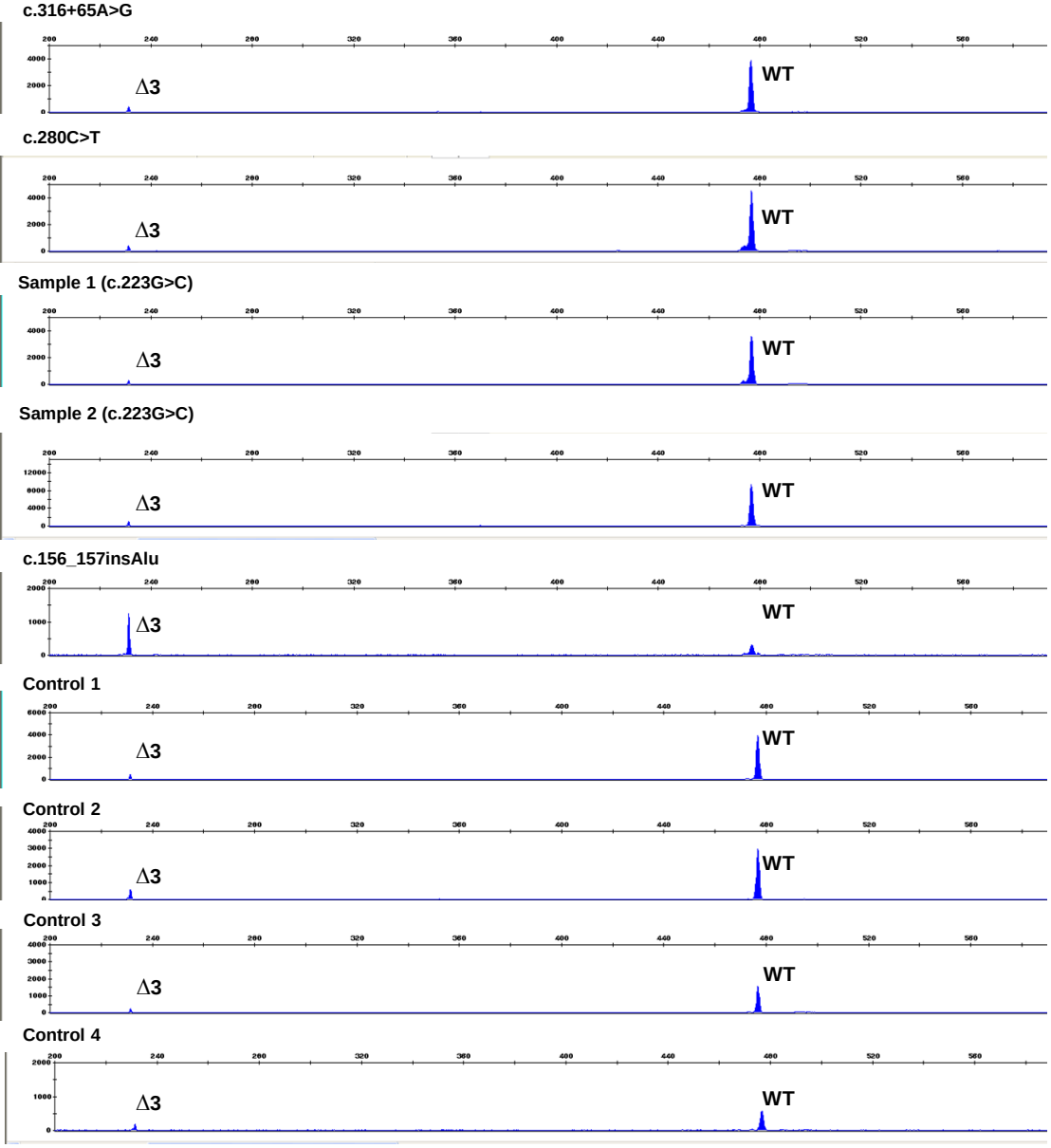


Figure S4c – III: electropherograms for variant carriers and controls (non-carriers)



Fragment analysis for individual samples:

Samples	%delta3	%WT
c.316+65A>G	6,74	93,26
c.280C>T	7,19	92,81
c.223G>C	6,90	93,10
c.223G>C	9,79	90,21
c.156_157insAlu	70,84	29,16
control1	9,75	90,25
control2	11,35	88,65
control3	12,13	87,87
control4	7,41	92,59
control5	9,00	91,00
control7	8,03	91,97
control9	11,32	88,68
control 10	7,14	92,86
control 11	7,76	92,24
control 12	6,05	93,95
control 13	5,99	94,01
control 14	9,98	90,02
control 15	9,93	90,07
control 16	13,77	86,23
control 17	7,30	92,70

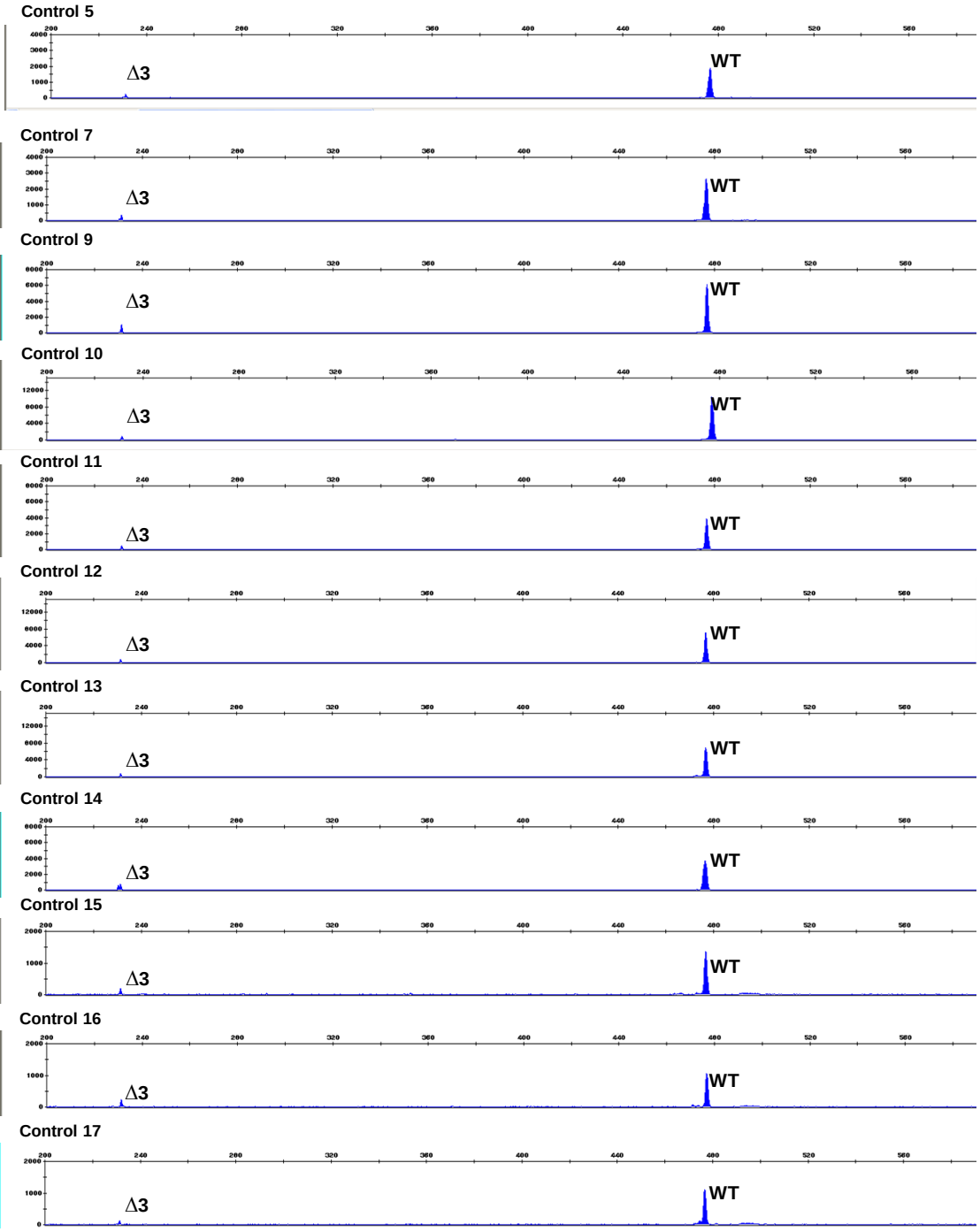
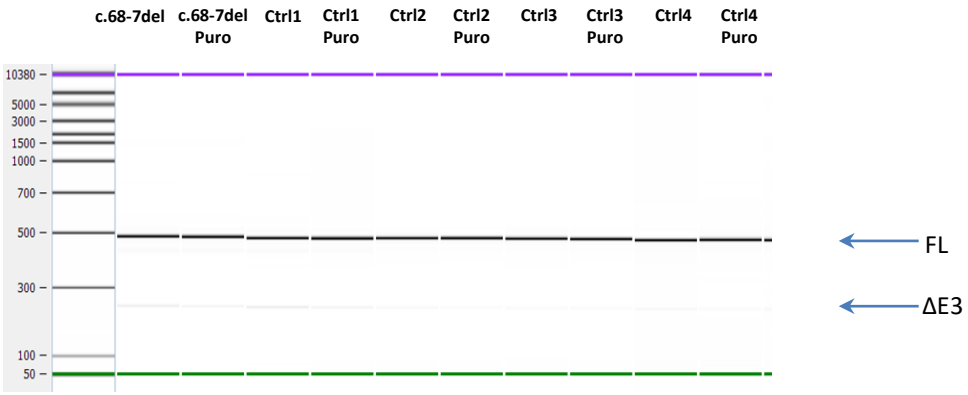


Figure S4c – III: continued

Figure S4d

Figure S4d. Capillary electrophoresis analysis of RT-PCR products from LCL. RNA was harvested from peripheral blood of two patients that harbor the c.68-7del variant. RNA from 10 control individuals were also analysed. I: pseudo gel image generated from capillary electrophoresis data. Full length and ΔE3 bands are indicated. II: Sanger sequencing of bands excised from agarose gel. III: Electropherograms for variants and controls (non-carriers). IV: Summary of fragment analysis. The proportion of the wild type and ΔE3 and ΔE3-4 transcripts was expressed as a percentage of the total area of them. Data represent the mean of at least two independent PCR reactions. Puro/P=puromycin-treated sample. ND: not determined. x̄: mean of two independent PCR, σ: standard deviation of two independent PCR reactions.

I



II

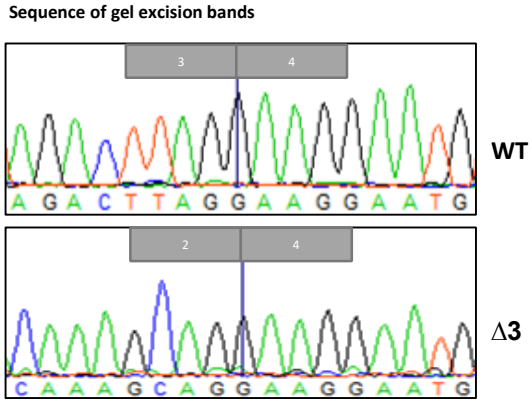
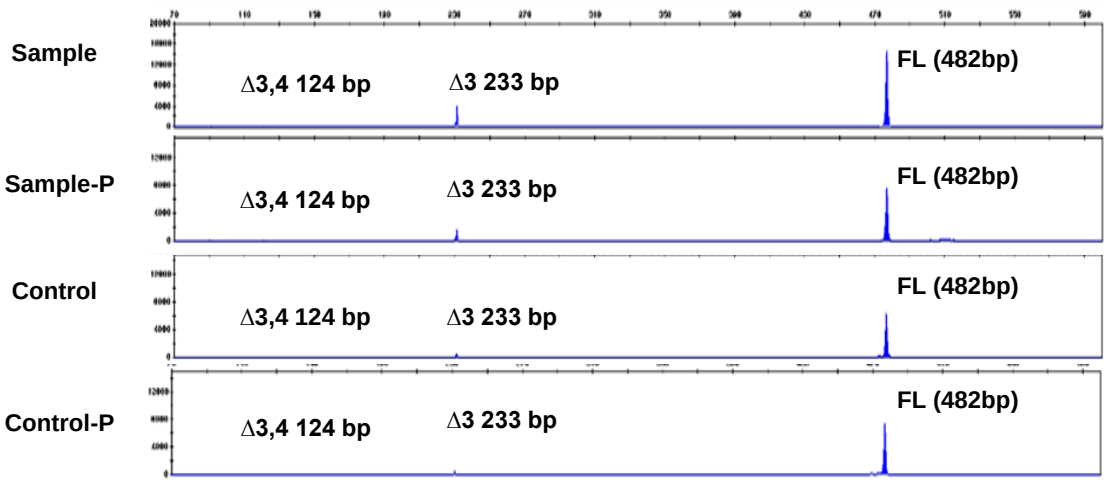


Figure S4d: continued

III



IV

	Δ3,4	Δ3	FL
Sample	0,86	15,07	84,08
Sample-P	0,75	12,62	86,63
Controls mean	0,87	5,71	94,03
Controls-P mean	0,69	3,91	95,95

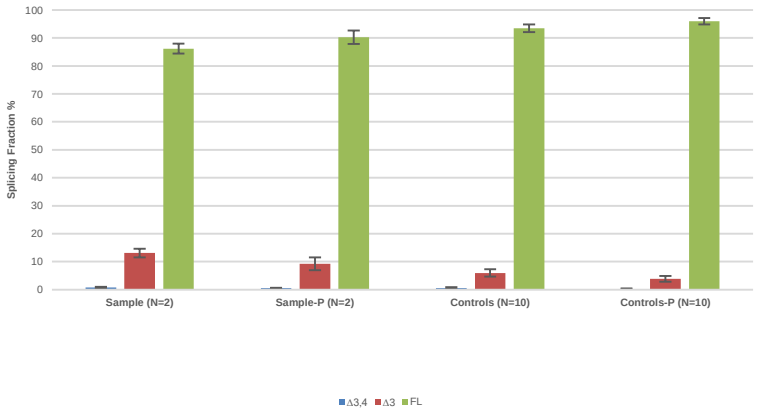


Figure S4e

ID	Variant	Protein	Splicing
BR659	c.198A>G	Gln66Gln	Δ4
BR1008			Δ4, Δ3-4
BR1292			Δ3
BR1001	c.223G>C	Ala75Pro	Δ3, Δ4
BR1045			Δ3
BR1155	c.125A>G	Tyr42Cys	Δ3

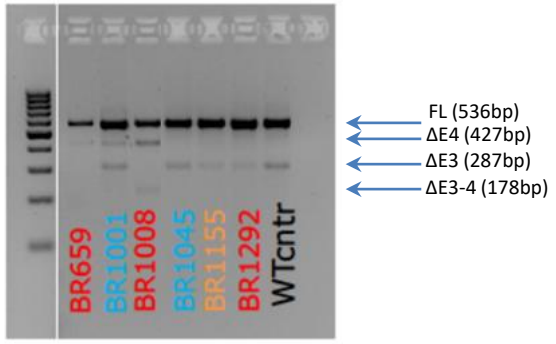


Figure S4e. Splice assay of c.198A>G, c.223G>C and c.125A>G variants. RT-PCR products generated with forward primers spanning exon 1-2 junction and reverse primers spanning exon 5-6 junction were analyzed by agarose gel electrophoresis. ΔE3, ΔE4, ΔE3-4 and full length (FL) bands are marked by arrow. WTcntr = WT control. The isoforms ΔE3, ΔE4, ΔE3-4 all represent naturally occurring splicing events previously reported in control samples.

Figure S4f

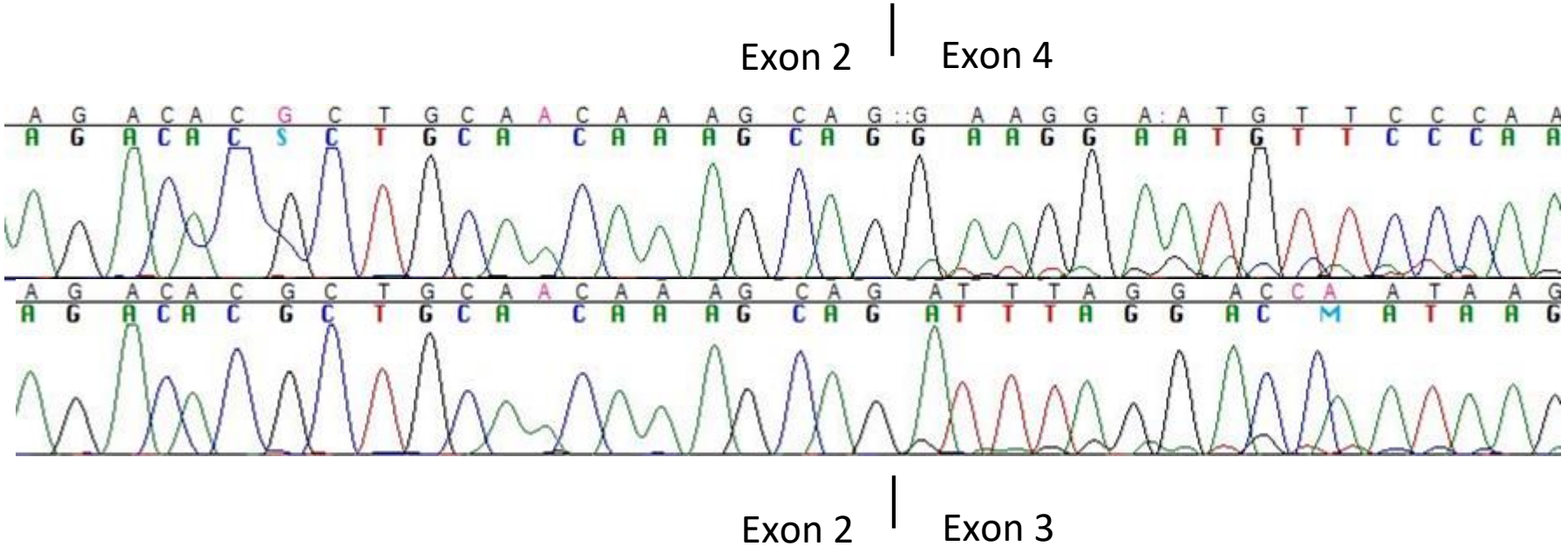


Figure S4f. Splice assay of c.277_317-726delinsCCAT variant. RNA was purified from blood using Trizol. Sanger sequencing of RT-PCR product using forward primer. Upper panel show analysis from a c.277_317-726delinsCCAT carrier. Lower panel represents a non-carrier. The analysis shows high level of ΔE3 and low level of WT transcript from the carrier sample and no alternative transcripts.

Figure S4g

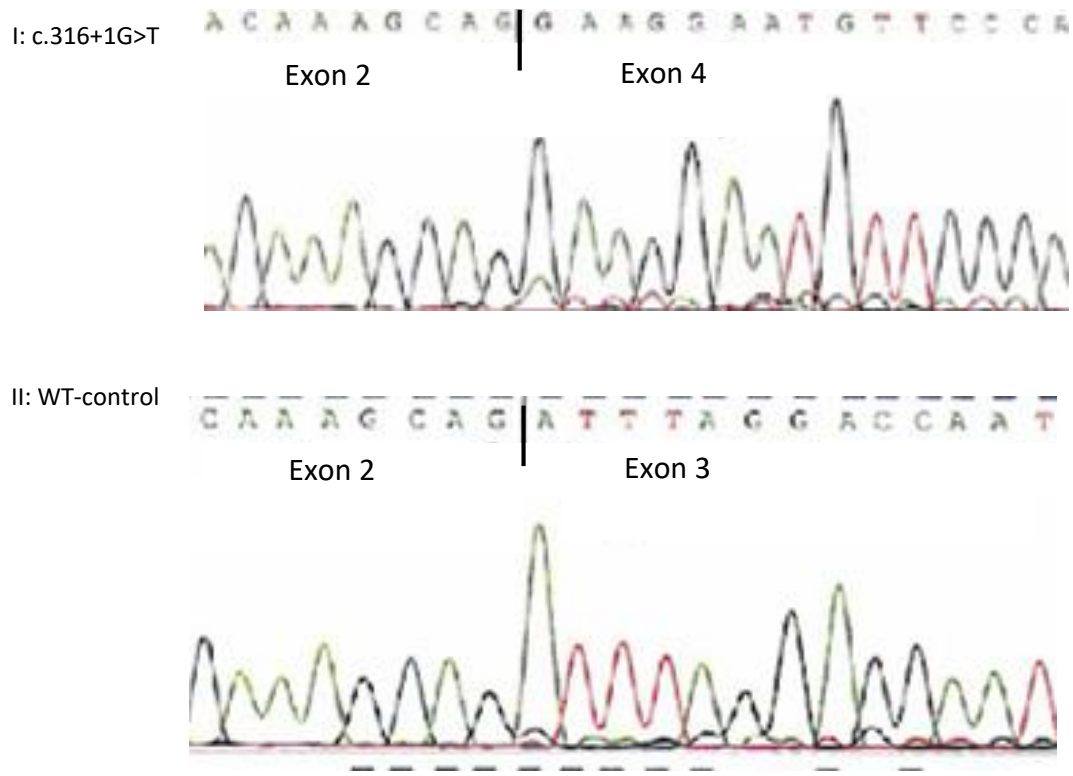


Figure S4g. Splice assay of c.316+1G>T variant. RNA was purified from Paxgene tube. Sanger sequencing of RT-PCR product using forward primer. Upper panel show analysis from a c.316+1G>T carrier. Lower panel represents a non-carrier sample. The analysis shows high level of ΔE3 and low level of WT transcript from the carrier sample and no alternative transcripts.

Figure S4h

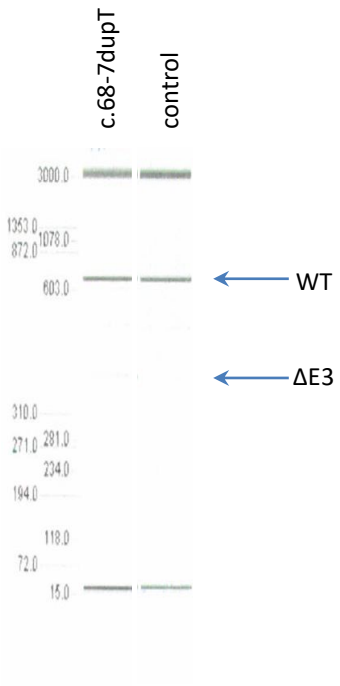


Figure S4h. Splice assay of c.68-7dupT. RT-PCR products were analyzed by QiaXcel capillary electrophoresis. Position of ΔE3 bands (not detectable) is marked by arrow.

Figure S4i

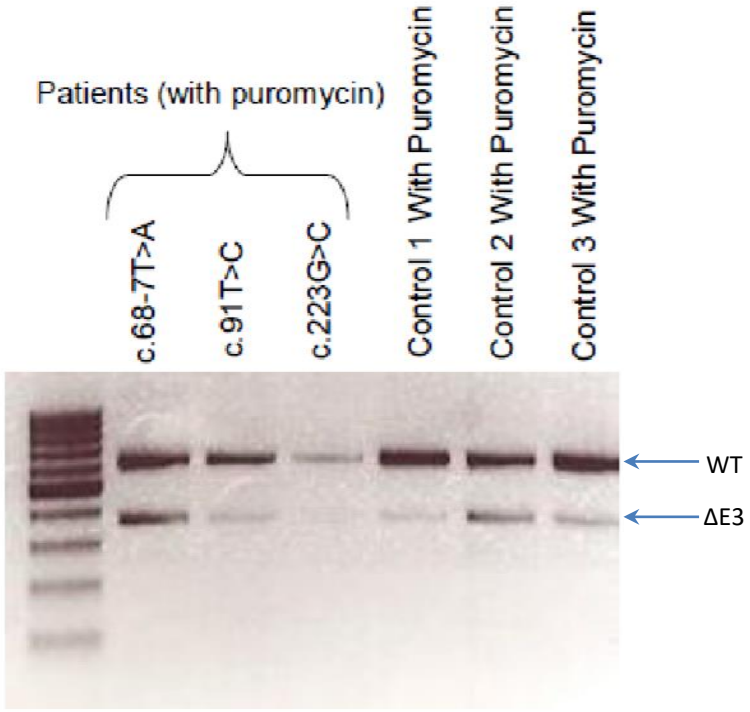


Figure S4i. Splice assay of c.91T>C and c.223G>C. RT-PCR products were analyzed by agarose gel electrophoresis. ΔE3 and wild type bands are marked by arrow.

Figure S4j

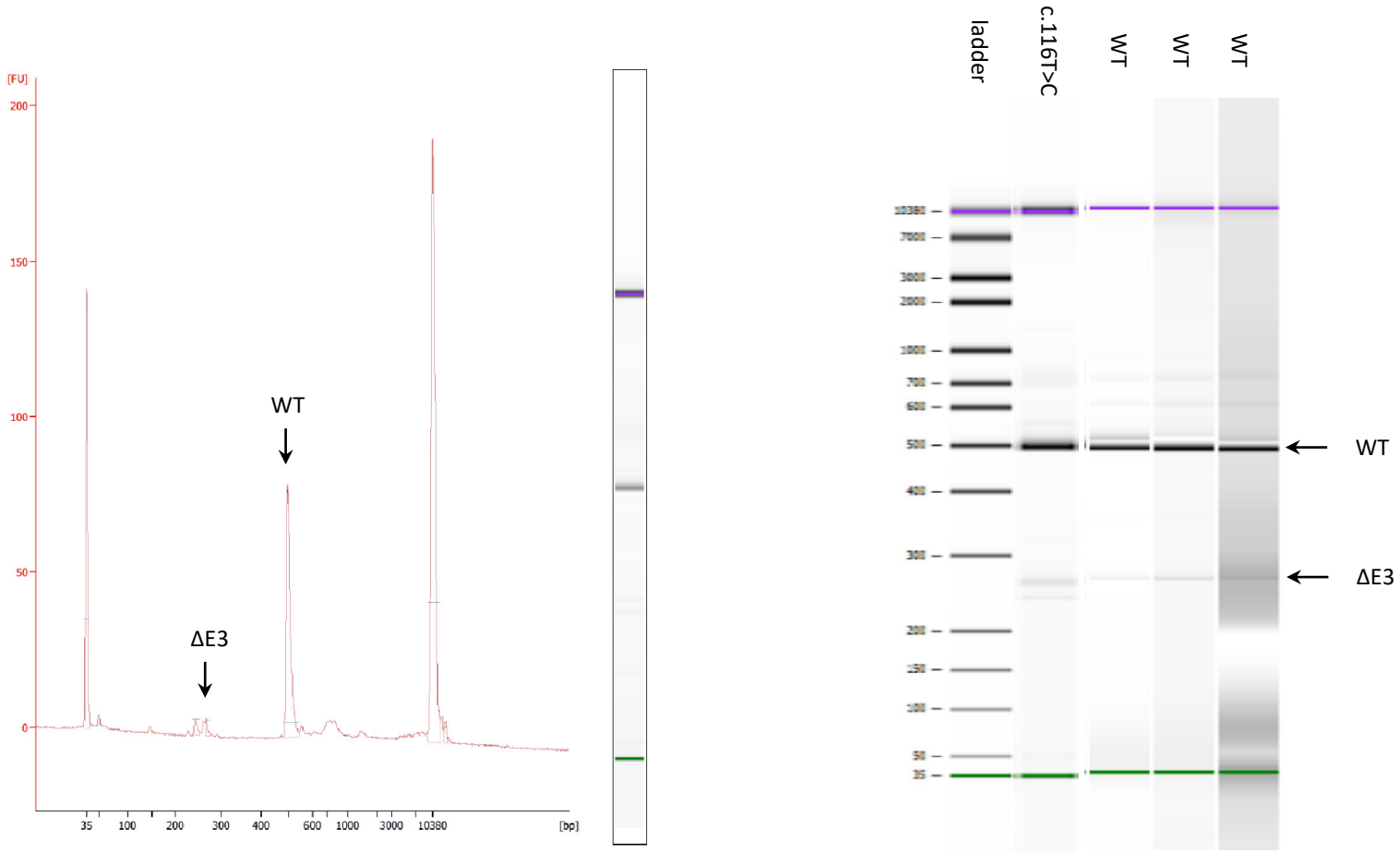


Figure S4j. Splice assay of c.116T>C. RT-PCR products were analyzed by Bioanalyzer. ΔE3 and wild type bands are marked by arrow. Left panel shows electropherogram and right panel shows pseudo-gel image including 3 non-carrier control samples (WT).

Figure S5: Quantitative dPCR of ΔE3-4 and ΔE3-7 isoforms

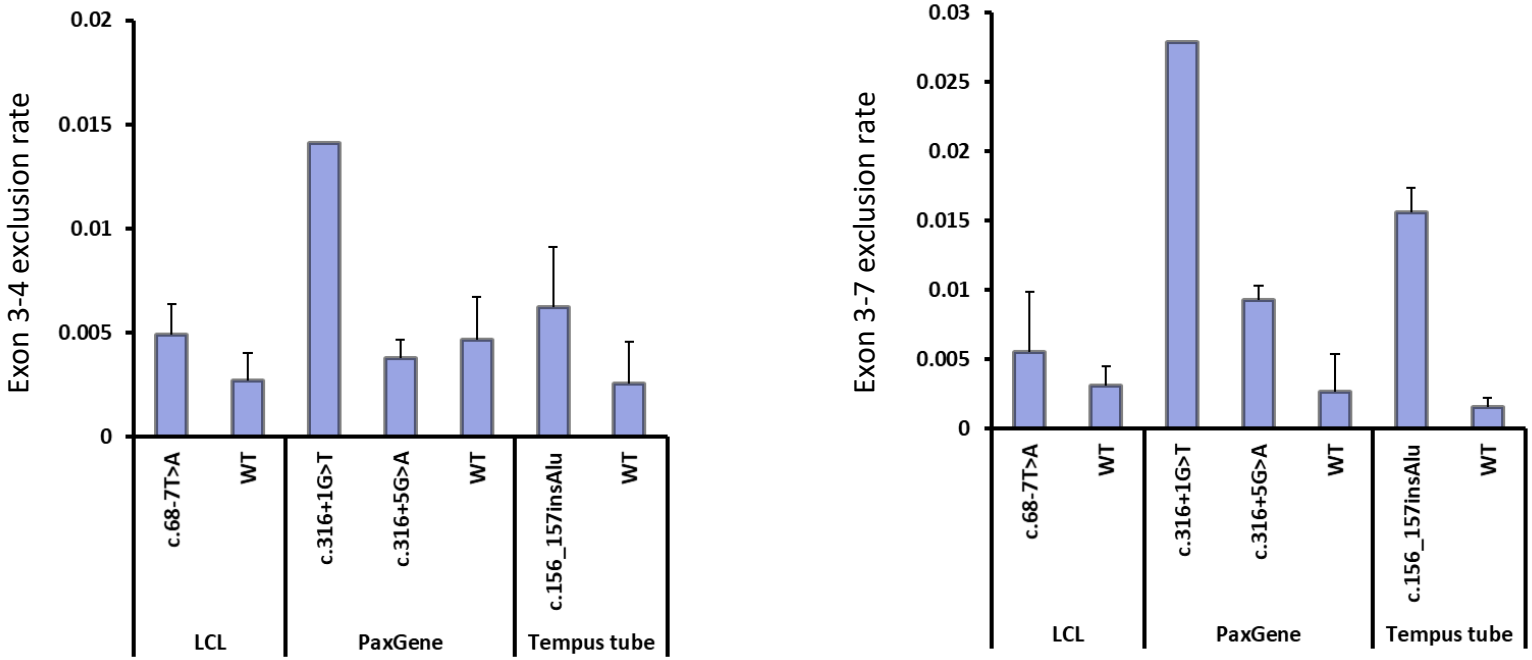
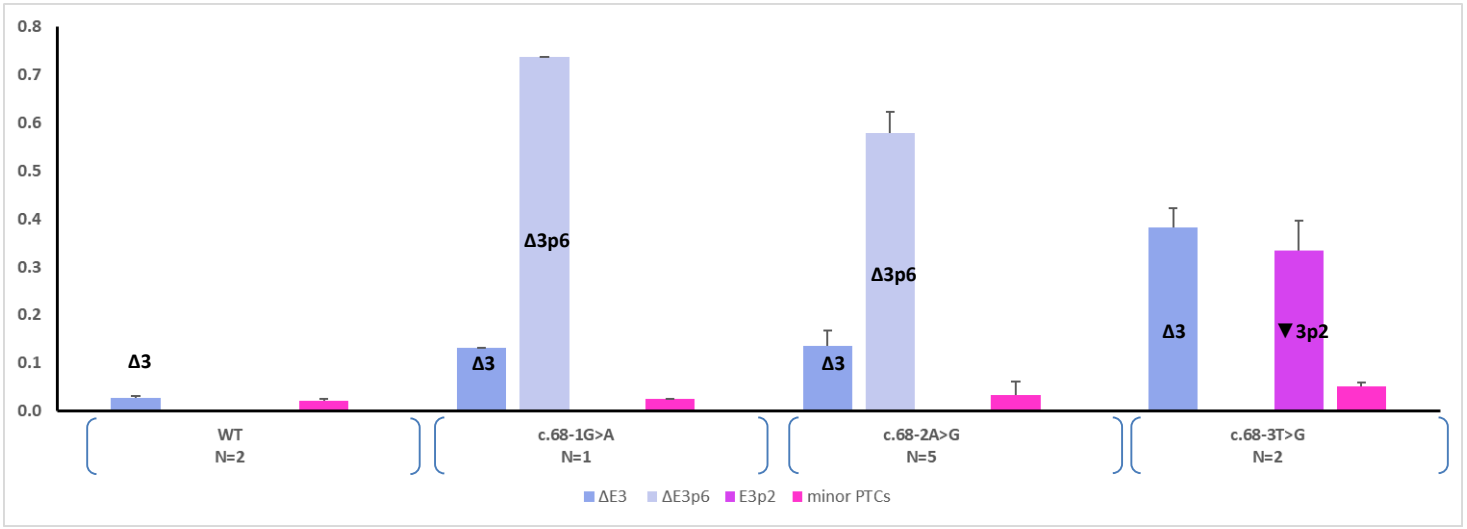


Figure S5. Quantitative dPCR of the two known other isoforms than ΔE3 in patient samples: ΔE3-4 (left panel) and ΔE3-7 (right panel) isoforms. The exclusion rate is shown for full exon 3 skipping control variant c.156_157insAlu, known benign partial skipping control variant c.68-7T>A, and variant c.316+1G>T previously reported to lead to full exon skipping in a minigene assay (PMID:29707112, Caputo et al (2018) Oncotarget 9: 17334). Data is shown for 3 different sample (collection) types: Lymphoblastoid cell line cultures (LCL), Paxgene blood, and Tempus blood. The error bars show standard deviation of 3 or more measurements. See methods for details on exclusion rate calculation. A single WT non-carrier control was assayed for each sample type.

Figure S6: RNASeq data of patient samples

See following pages for Sub-figures S6a to S6d

Figure S6a RNASeq data



All PTC transcripts	WT N=2	c.68-1G>A N=1	c.68-2A>G N=5	c.68-3T>G N=2
r.68_425del, p.D23Vfs*10	0.49%	1.06%	1.74%	4.33%
r.68_73del, p.D23_L24del	0.02%	73.64%	57.85%	0.04%
r.64_67del, p.A22Ifs*2	0.28%	0.47%	0.00%	0.00%
r.-39_67del, p.M1?	0.05%	0.00%	0.00%	0.22%
r.317_425del, p.G106Vfs*10	0.03%	1.07%	0.38%	0.12%
r.317_320del, p.R107Mfs*13	1.12%	0.00%	0.62%	0.43%

Figure S6a Quantitative RNA sequencing of patient RNA samples. Upper panel: major detected transcripts. Minor protein transcript expected to cause protein truncation, NMD or not to be expressed are summarized as minor PTC's. The bars show inferred per allele expression relative to total transcript amount and error bars are standard deviations. Lower panel: percentage expression of all PTC-NMD-non-coding transcripts. Data per variant are presented in Figure S6b-d.

Figure S6b

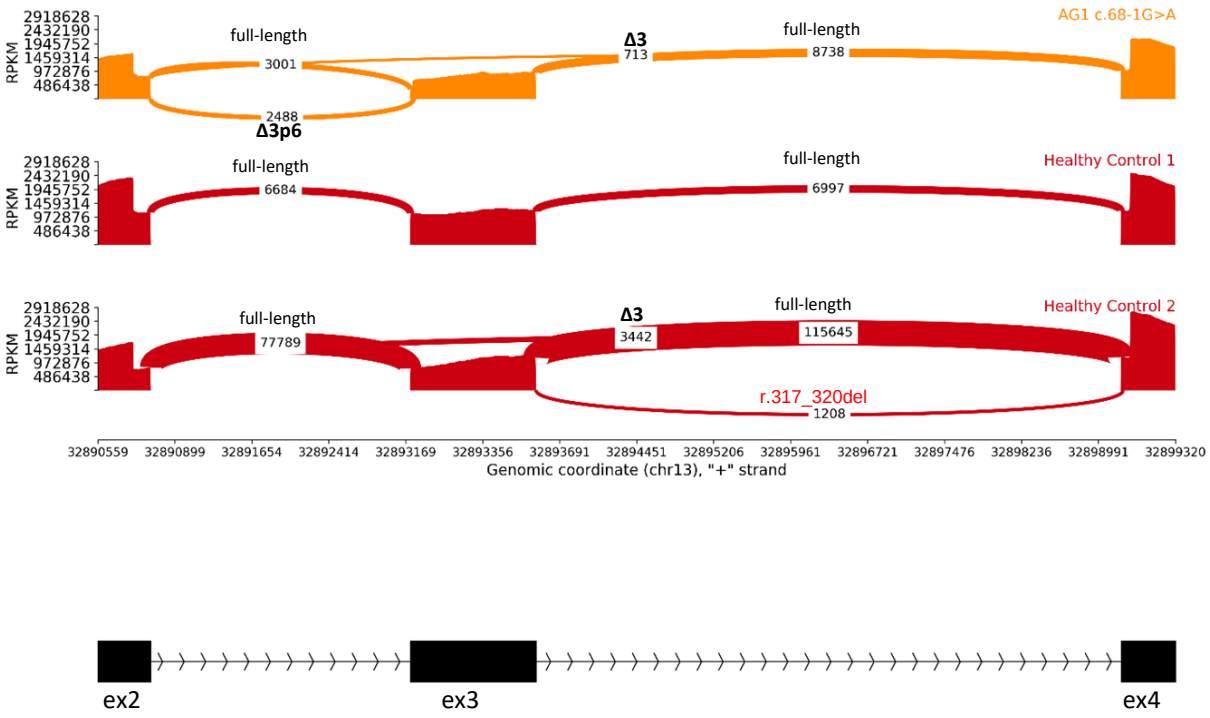


Figure S6b Quantitative RNA sequencing of patient RNA sample with c.68-1G>A. Upper panel: Individual read connections and number of reads supporting the same transcript are shown. Transcript annotation is provided. X-axis: chromosomal position, Y-axis: reads per kilobase per million of total reads (RPKM). Patient sample from carrier of c.68-1G>A (orange) and 2 WT samples (red) are shown. Lower panel: gene structure for exon 2-4 region.

Figure S6c

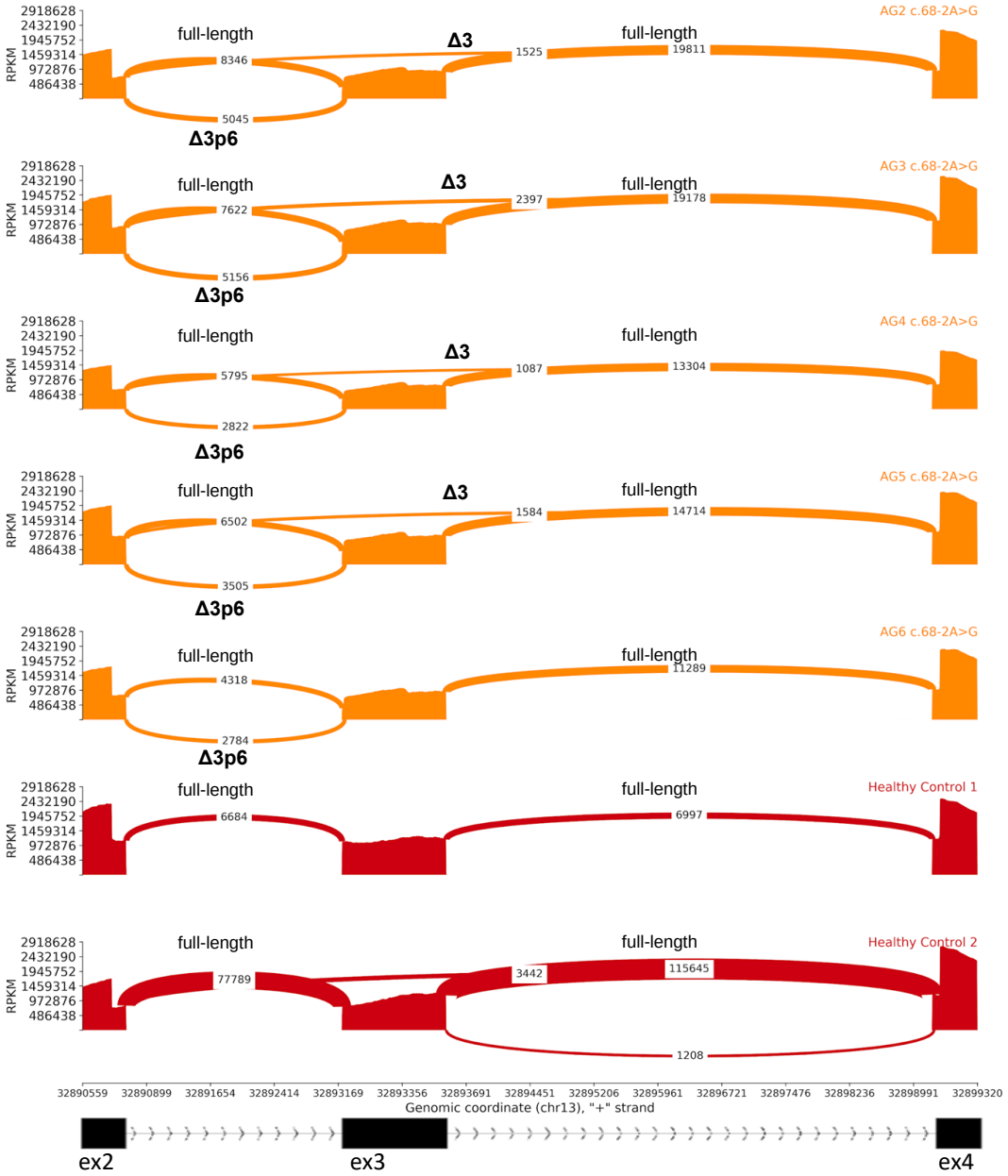


Figure S6c Quantitative RNA sequencing of patient RNA samples with c.68-2A>G. Upper panel: Individual read connections and number of reads supporting the same transcript are shown. Transcript annotation is provided. X-axis: chromosomal position, Y-axis: reads per kilobase per million of total reads (RPKM). Results are shown for patient samples from 5 carriers of c.68-2A>G (orange) and 2 WT samples (red). Lower panel: gene structure for exon 2-4 region.

Figure S6d

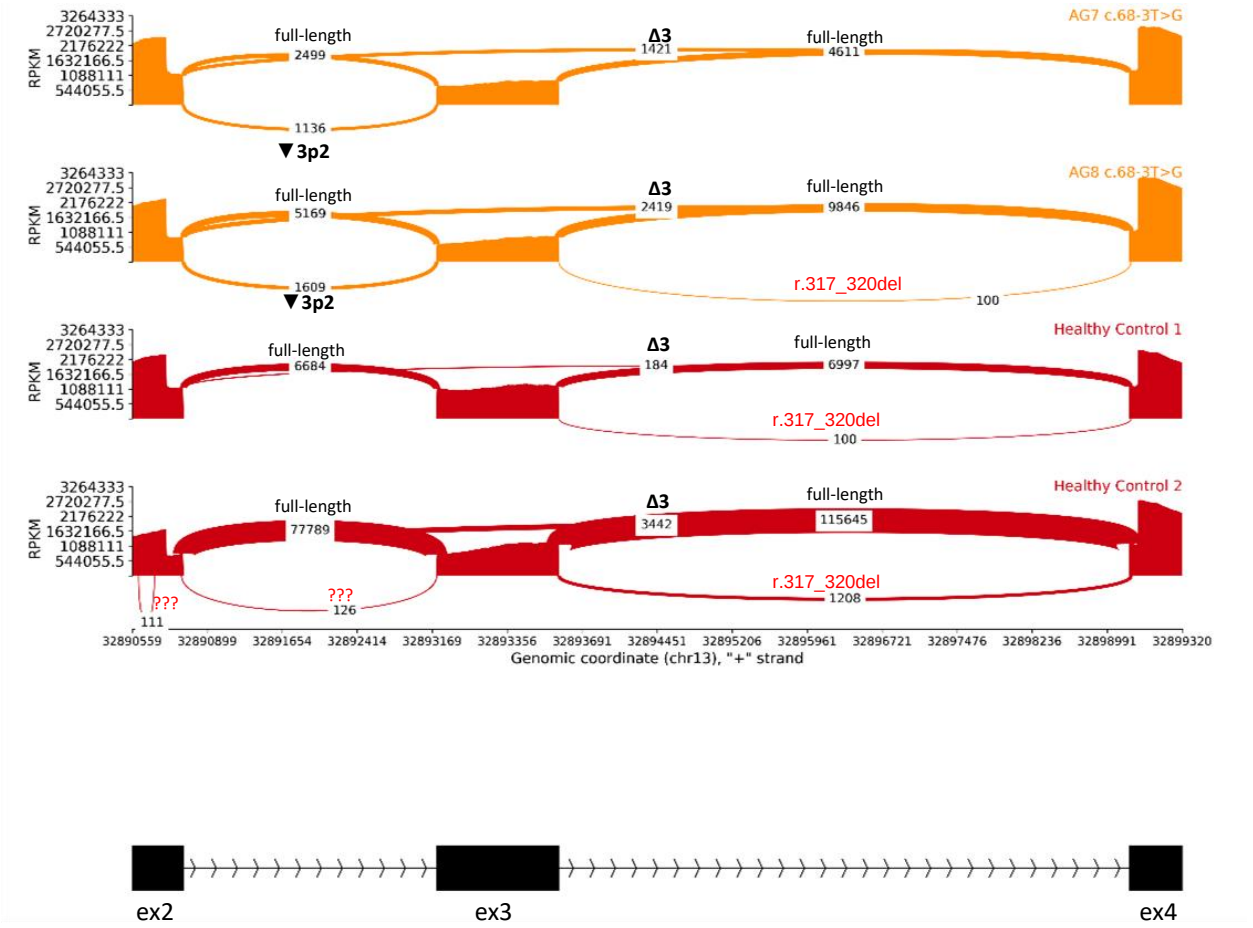


Figure S6d Quantitative RNA sequencing of patient RNA samples with c.68-3T>G. Upper panel: Individual read connections and number of reads supporting the same transcript are shown. Transcript annotation is provided. X-axis: chromosomal position, Y-axis: reads per kilobase per million of total reads (RPKM). Patient samples from 2 carriers of c.68-3T>G (orange) and 2 WT samples (red). Lower panel: gene structure for exon 2-4 region.

Figure S7: Results from dPCR analysis of mESC

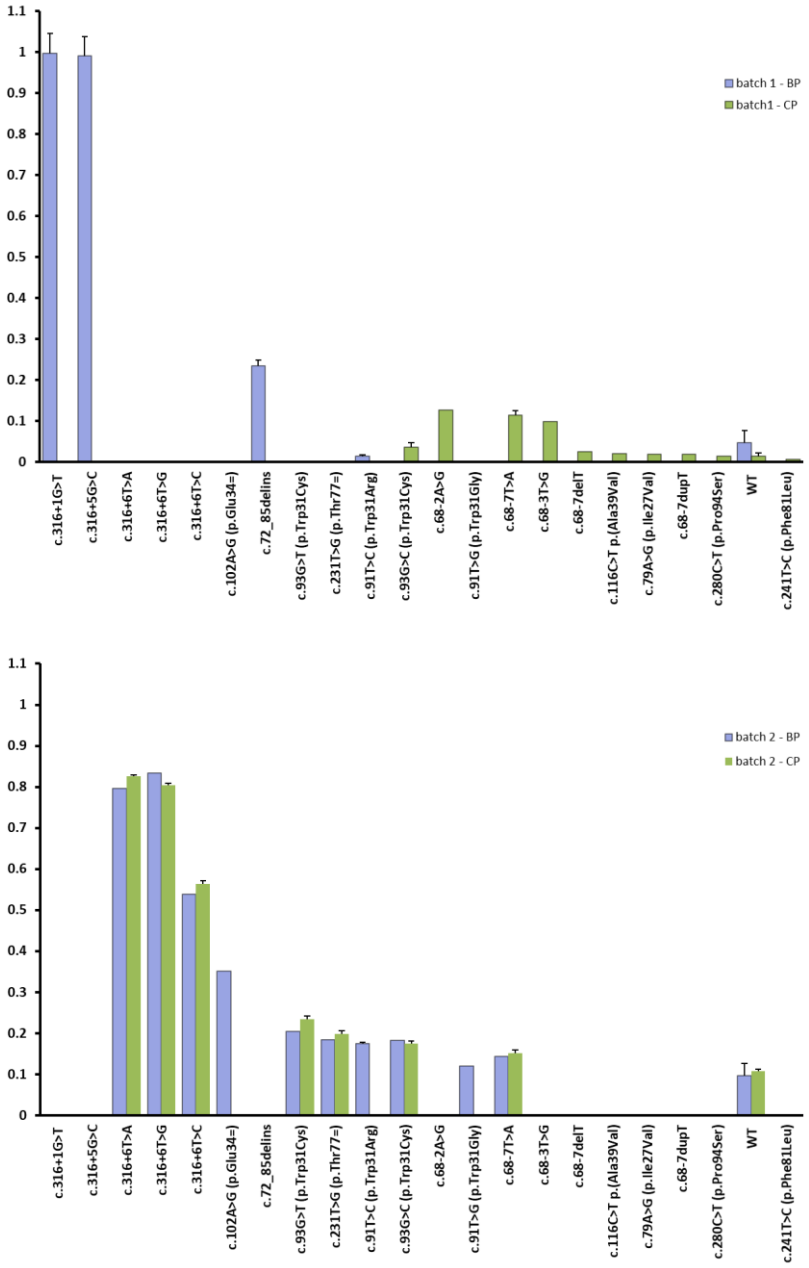


Figure S7. Quantitative dPCR of exon 3 skipping in mouse embryonic stem cells (mESC). The analyses were performed at two different times, batch 1 (upper panel) and batch 2 (lower panel). Exclusion rate is shown for a number of variants including a complete exon 3 skipping control (c.316+5G>C) and wildtype control (WT). Data from two experimental conditions are included: BP where mouse BRCA2 is still expressed and CP where mouse *BRCA2* is depleted. The error bars are confidence intervals from the Poisson distribution of dPCR data.

Figure S8: Capillary electrophoresis of mESC RT-PCR products for c.68-2A>G and c.68-3T>G

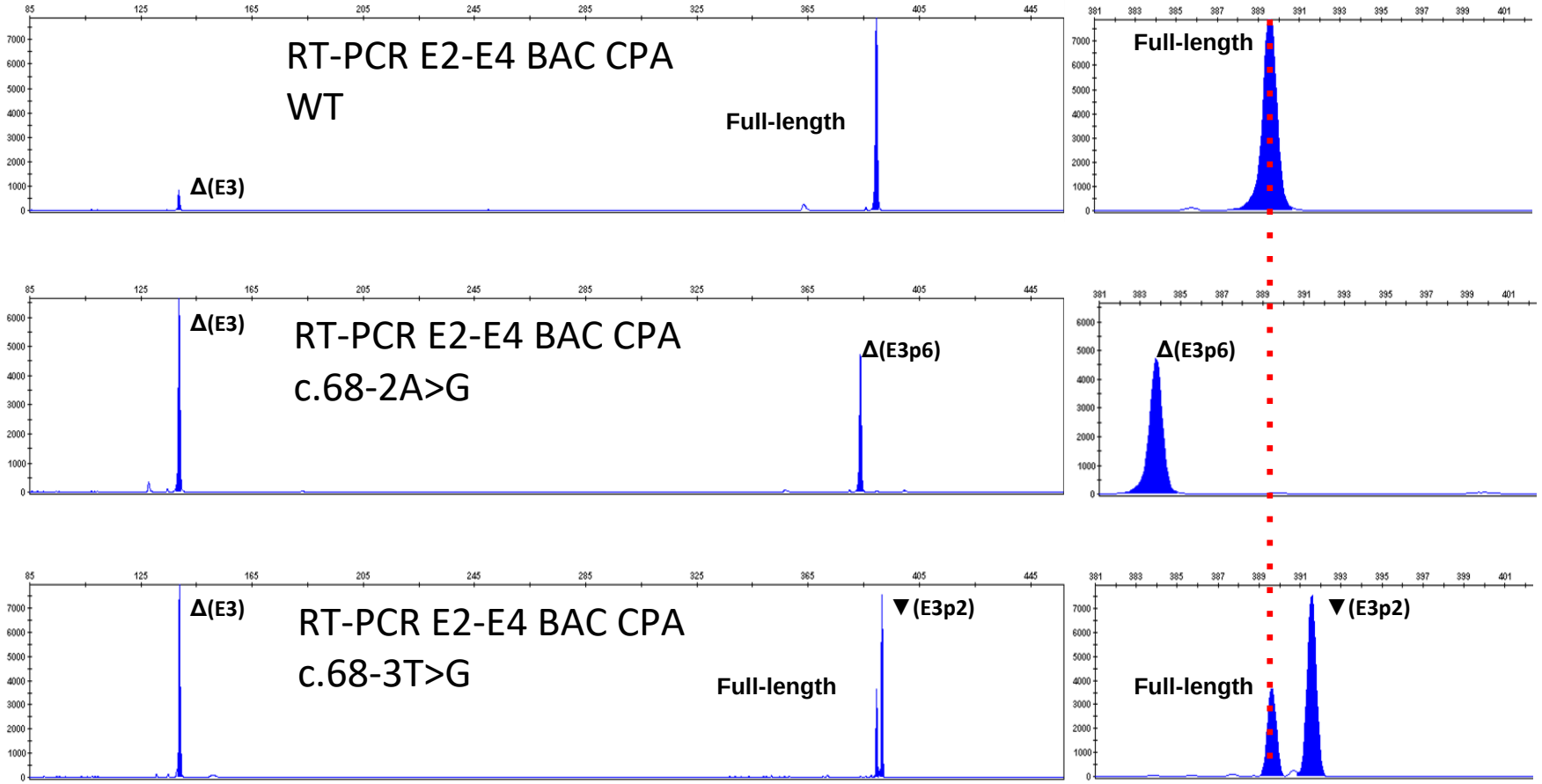


Figure S8: Capillary electrophoresis of mESC RT-PCR products for *BRCA2* c.68-2A>G. Upper panel shows result for cDNA from *BRCA2* WT BAC control expressed in mESC system. Middle panel, result for *BRCA2* c.68-2A>G; the variant shows solely expression of ΔE3 and the transcript lacking 6 bases (ΔE3p6). Lower panel, result for *BRCA2* c.68-3T>G showing Full length, ΔE3 and out of frame transcript including two additional bases (E3p2).

Figure S9: Cell viability of BRCA2 exon 3 variants in mESC.

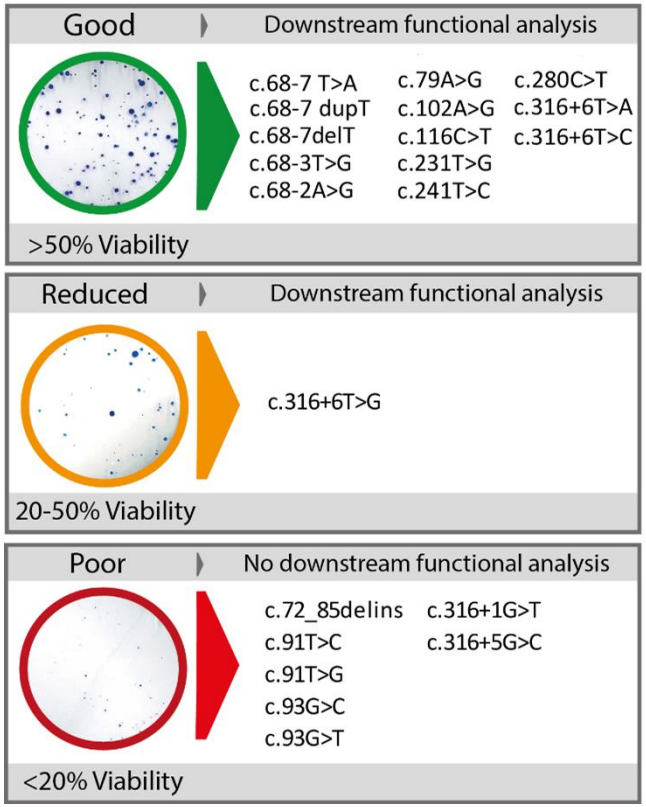


Figure S9: Cell viability of *BRCA2* exon 3 variants in mESC. HAT-resistant clones formed post Cre-mediated removal of the conditional mBrca2 allele were visualized by methylene blue staining. Representative complementation phenotypes observed in the cell viability assay are shown for the three categories. For each variant, the number of clones was compared to the complementation of WT *BRCA2* expressing cells, and based on that categorized in one of the three categories: poor, no complementation or <20% viability; reduced, 20-50% viability; good, >50% viability. Variants of known pathogenicity were c.68-7T>A (benign) and c.316+5G>C (pathogenic). Downstream functional characterization was performed for all variants in the reduced and good complementation category).

Figure S10: Western blot analysis of mESC

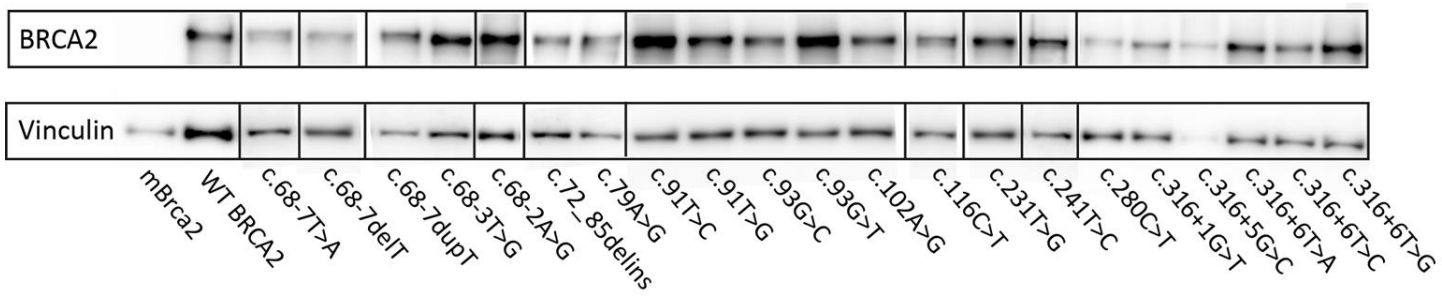


Figure S10. Western blot analysis of *BRCA2* exon 3 variants in mESC. Vinculin was used as loading control. It is important to note that the ΔE3 protein isoform (3335 aa) cannot be distinguished from the full-length *BRCA2* protein (3418 aa) by western blot analysis due to the small difference in size.